The Leeds Teaching Hospitals NHS
NHS Trust

UNIVERSITY OF LEEDS

Digital Pathology in the NHS : The Evidence Base, Validation and Deployment

Dr Bethany Williams MBBS BSc

Digital Pathology Fellow

Leeds Teaching Hospitals NHS Trust, University of Leeds

bethany.williams2@nhs.net

www.virtualpathology.leeds.ac.uk

@LeedsPathology

Acknowledgement

Leeds Teaching Hospitals NHS Trust and Leica Biosystems have formed a collaborative partnership for the clinical deployment of digital pathology

The clinical use claims described in the information provided have not been cleared or approved by the U.S. FDA nor are the products available in the United States.

Overview

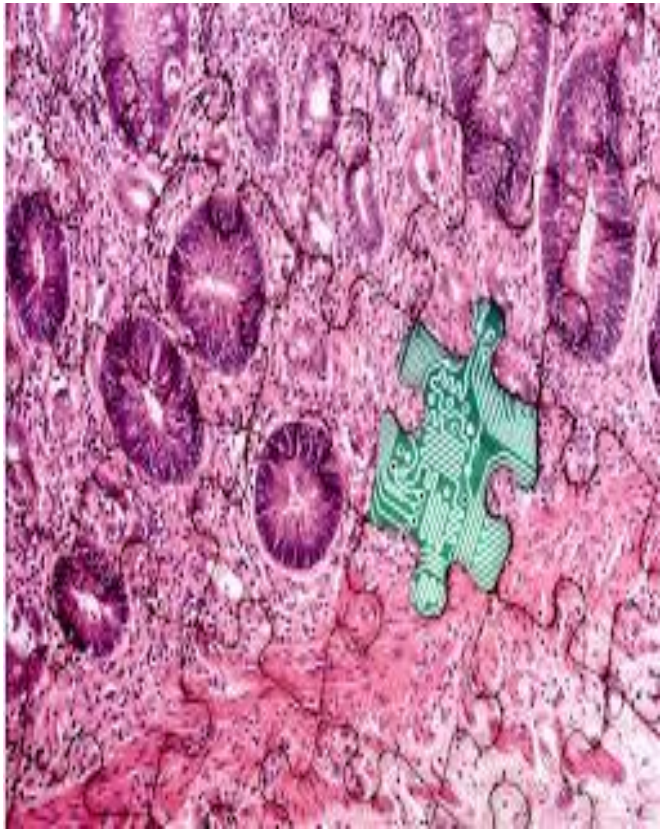
An introduction to Leeds

Concordance

Discordance

Validation and training protocol

Results from the breast validation pilot



Overview

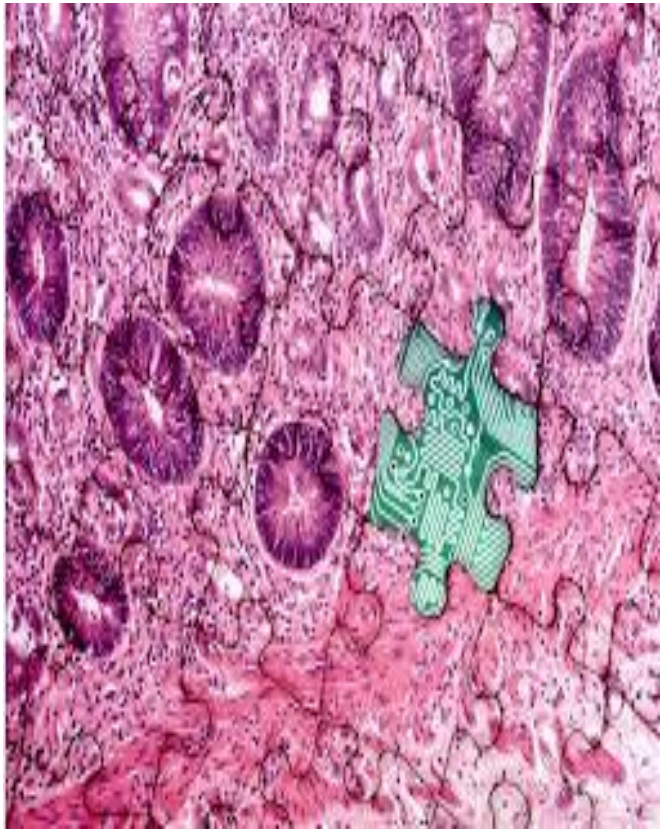
An introduction to Leeds

Concordance

Discordance

Validation and training protocol

Results from the breast validation pilot



Welcome to Leeds….

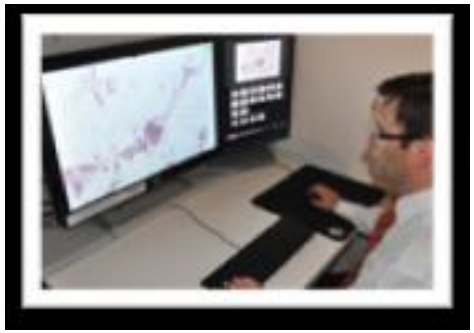
- Leeds Teaching Hospitals NHS Trust
  - Single site laboratory
  - Fully sub-specialised
  - 45 consultant pathologists
  - 30 trainee pathologists
  - 250,000 H&E slides/ year
  - University department on same site
- Scanning since 2003
  - 6 Aperio scanners
  - >332,000 slides
  - 160TB of image data

Expertise in Digital Pathology

***Website**: 10,000 + virtual slides, slide library, e-learning, QA materials, papers, videos and more*

www.virtualpathology.leeds.ac.uk

***Powerwall**: 12 megapixels*



Digital Workstation In Trainee Area

The clinical use claims described in the information provided have not been cleared or approved by the U.S. FDA nor are the products available in the United States.

Clinical usage

- Currently - 20% of laboratory slides scanned for primary diagnosis
- 3 pathologists validated for primary diagnosis, all pathologists validated for immunohistochemistry assessment
- 2x Leica Aperio AT2 - 400x standard slides
- 2x Leica Aperio CS2 - large slides
- 6MP Barco monitors / 8MP Eizo
- Medical grade
- Aperio eSlide Manager
- Leeds Virtual Microscope

The clinical use claims described in the information provided have not been cleared or approved by the U.S. FDA nor are the products available in the United States.

Overview

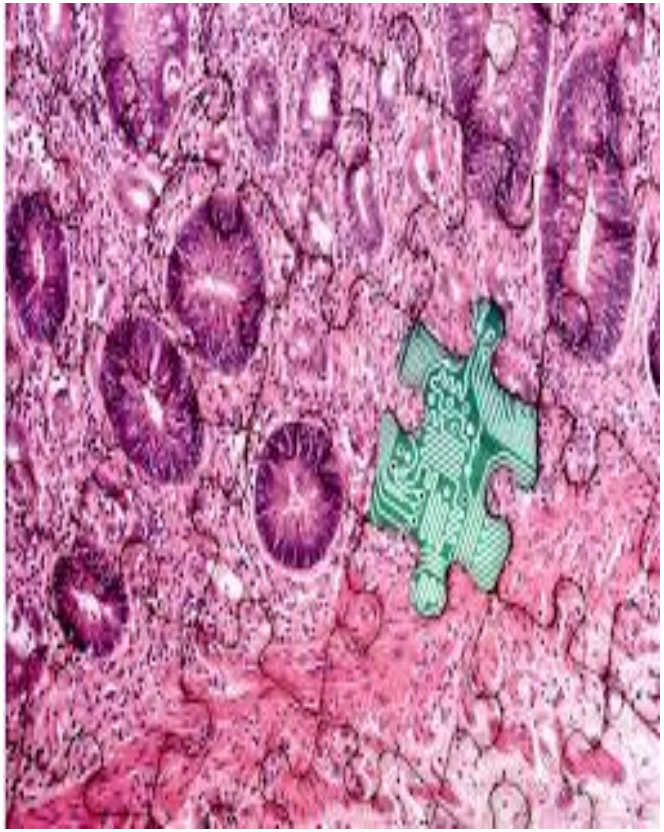
An introduction to Leeds

Concordance

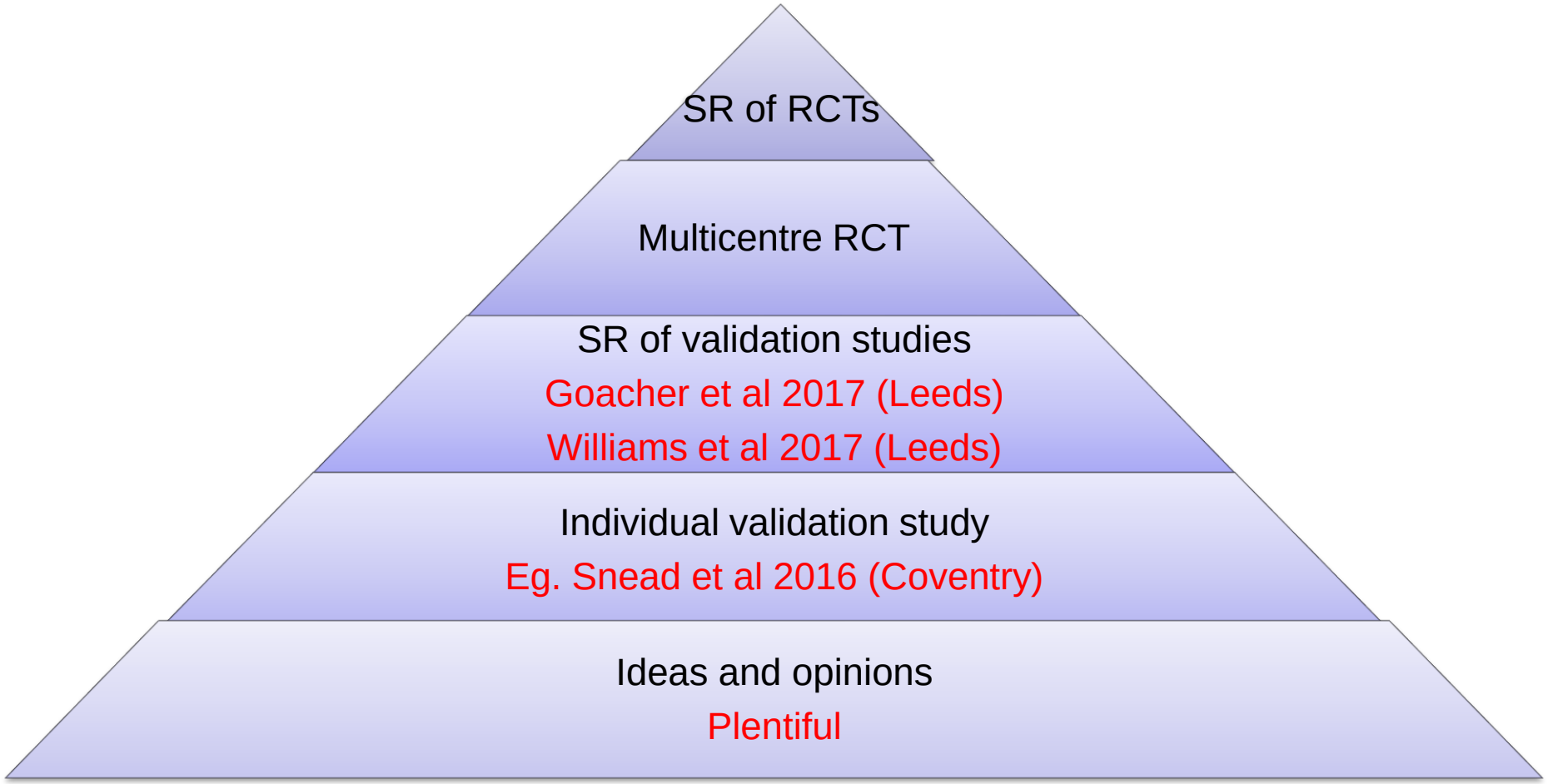
Discordance

Validation and training protocol

Results from the breast validation pilot



Evidence based medicine and digital pathology ....



Review Article

The Diagnostic Concordance of Whole Slide Imaging and Light Microscopy

A Systematic Review

Edward Goacher, BSc; Rebecca Randell, PhD; Bethany Williams, MBBS; Darren Treanor, MB, BSc, PhD, FRCPath

• *Context.*—Light microscopy (LM) is considered the reference standard for diagnosis in pathology. Whole slide imaging (WSI) generates digital images of cellular and tissue samples and offers multiple advantages compared with LM. Currently, WSI is not widely used for primary diagnosis. The lack of evidence regarding concordance

and the Cochrane Library (Wiley, London, England), between 1999 and March 2015.

Conclusions.—Thirty-eight studies were included in the review. The mean diagnostic concordance of WSI and LM, weighted by the number of cases per study, was 92.4%. The weighted mean κ coefficient between WSI and LM was

Methods: Inclusion/Exclusion Criteria

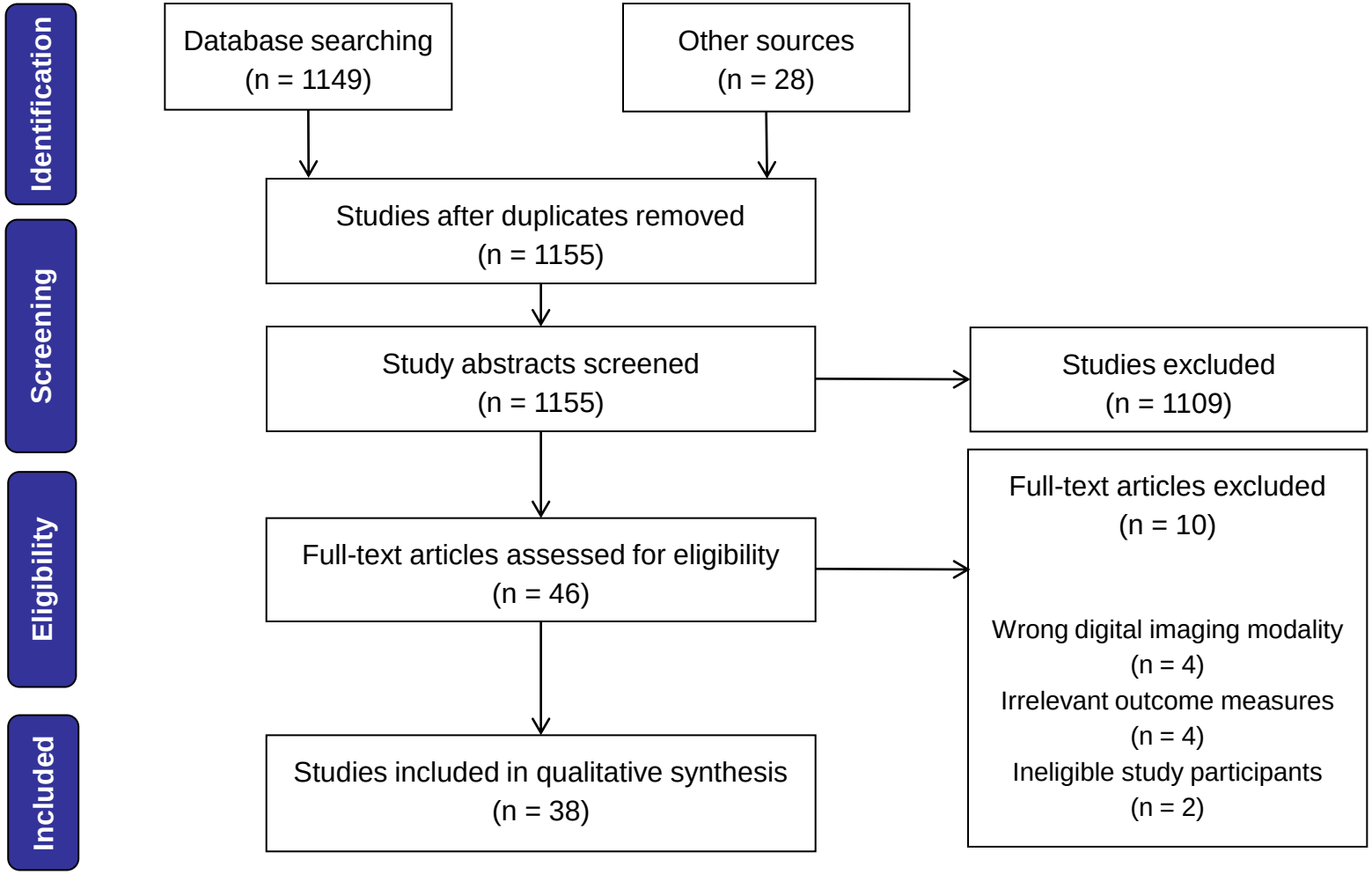
Include	Exclude
Assesses the diagnostic concordance of whole slide imaging and light microscopy	Descriptive technology studies
Majority of participants must be fully qualified pathologists	Cytology and frozen section studies
Includes primarily H&E stained slides	Educational and research studies
Quantified outcome measures	Digital image analysis software studies

Methods: Quality Assessment of Studies

- The quality of included studies was assessed using the QUADAS-2 tool

Item	Yes	No	Unclear
1. Patient selection			
1.1. Risk of bias			
1.1.1. Could the selection of patients have introduced bias?	( )	( )	( )
1.1.1.1. Was a consecutive or random sample of patients involved?	( )	( )	( )
1.1.1.2. Did the study avoid inappropriate exclusions?	( )	( )	( )
1.2. Applicability			
1.2.1. Are there concerns that the included patients and setting do not match the review question?	( )	( )	( )
2. Index test			
2.1. Risk of bias			
2.1.1. Could the conduct or interpretation of the index test have introduced bias?	( )	( )	( )
2.1.1.1. Were the index test results interpreted without the knowledge of the results of the reference standard?	( )	( )	( )
2.1.1.2. Were the corresponding clinical details provided for each case?	( )	( )	( )
2.1.1.3. Are participants trained in using the index test?	( )	( )	( )
2.2. Applicability			
2.2.1. Are there concerns that the index test, its conduct, or its interpretation differs from the review question?	( )	( )	( )
3. Reference standard			
3.1. Risk of bias			
3.1.1. Could the reference standard, its conduct, or its interpretation, have introduced bias?	( )	( )	( )
3.1.1.1. Is the reference standard likely to correctly classify the target condition?	( )	( )	( )
3.1.1.2. Were the reference standard results interpreted without knowledge of the results of the index test?	( )	( )	( )
3.1.1.3. Were the corresponding clinical details provided for each case?	( )	( )	( )
3.2. Applicability			
3.2.1. Are there concerns that the target condition as defined by the reference standard does not match the question?	( )	( )	( )
4. Flow and timing			
4.1. Risk of bias			
4.1.1. Could the patient flow have introduced bias?	( )	( )	( )
4.1.1.1. Was there an appropriate interval between the index test and the reference standard?	( )	( )	( )
4.1.1.2. Did all patients receive the same reference standard?	( )	( )	( )
4.1.1.3. Were all patients included in the analysis?	( )	( )	( )

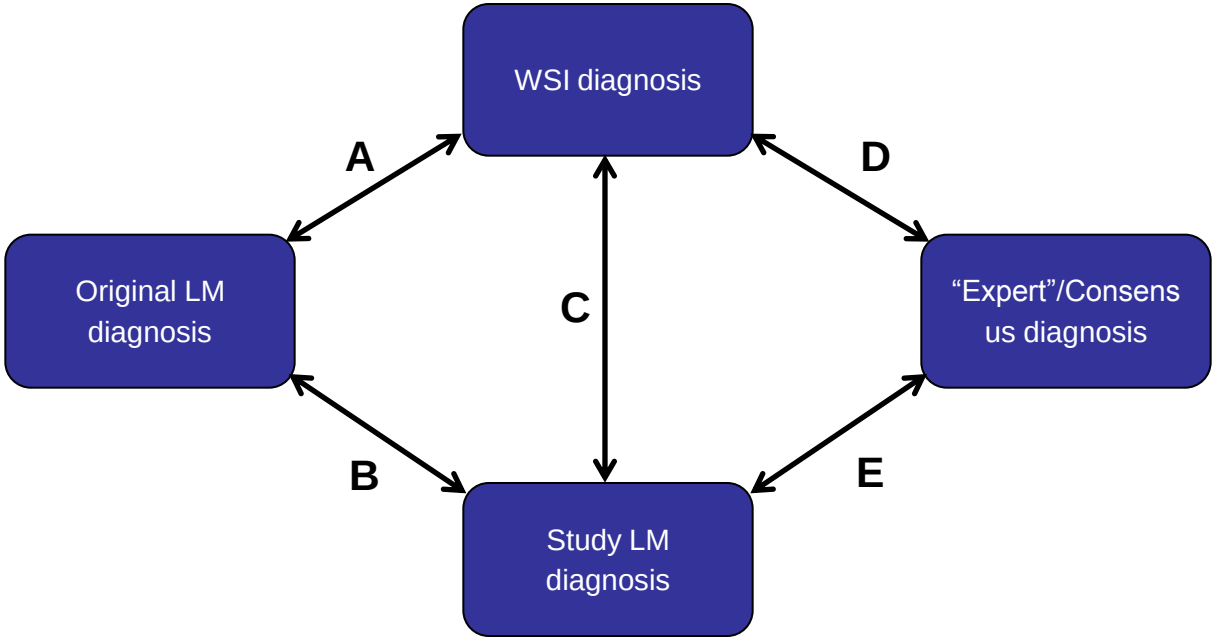
Results: PRISMA Flow Diagram



Results

- 38 studies identified
- 2006-2015
- No. study participants: 1 – 26
- No. cases: 20 - 524

Results: Study Design



LM: Light microscopy
WSI: whole slide imaging

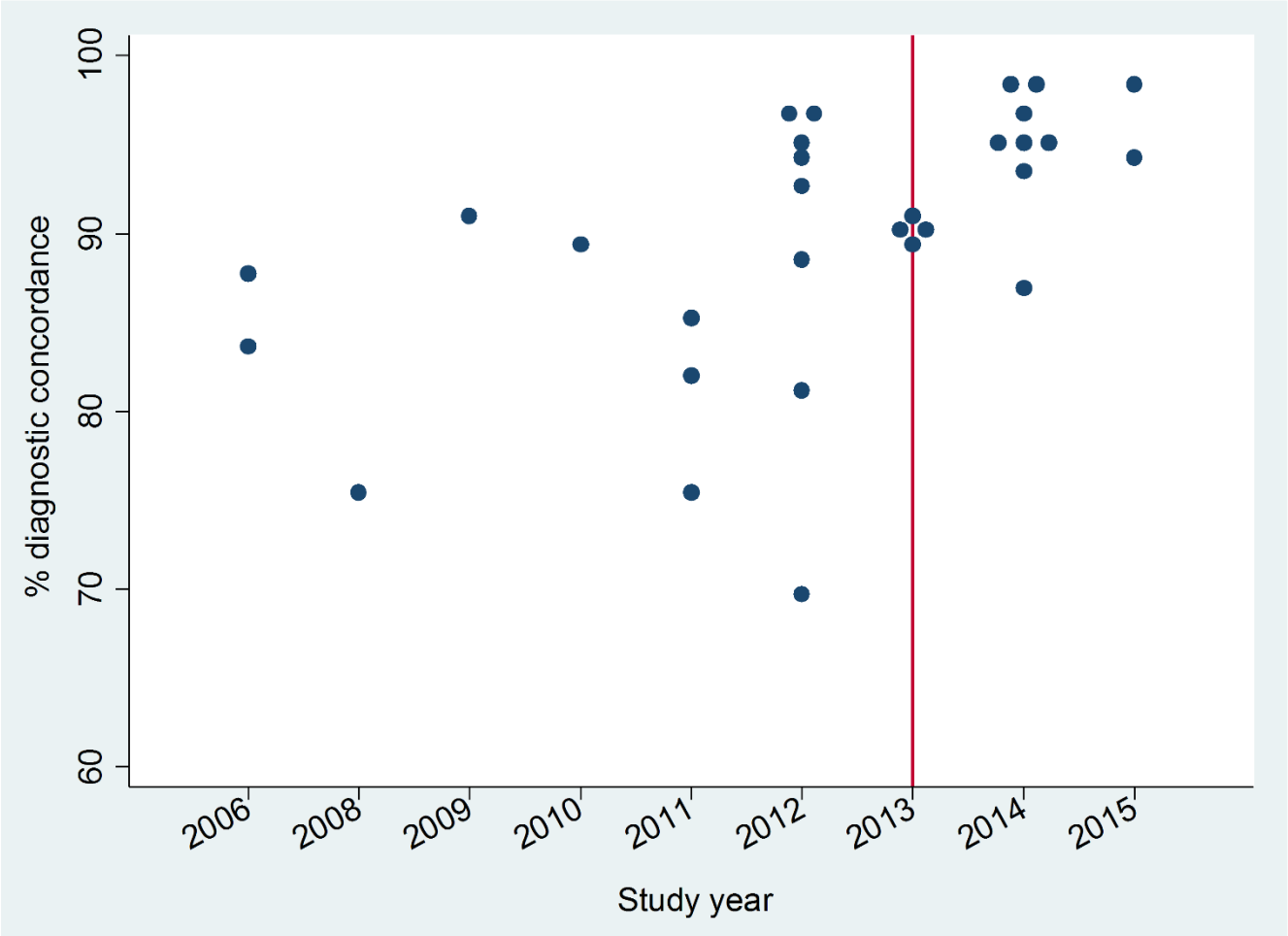
The diagnostic concordance of whole slide imaging and light microscopy : a systematic review

- Mean diagnostic concordance of whole slide imaging and light microscopy **92.4%**

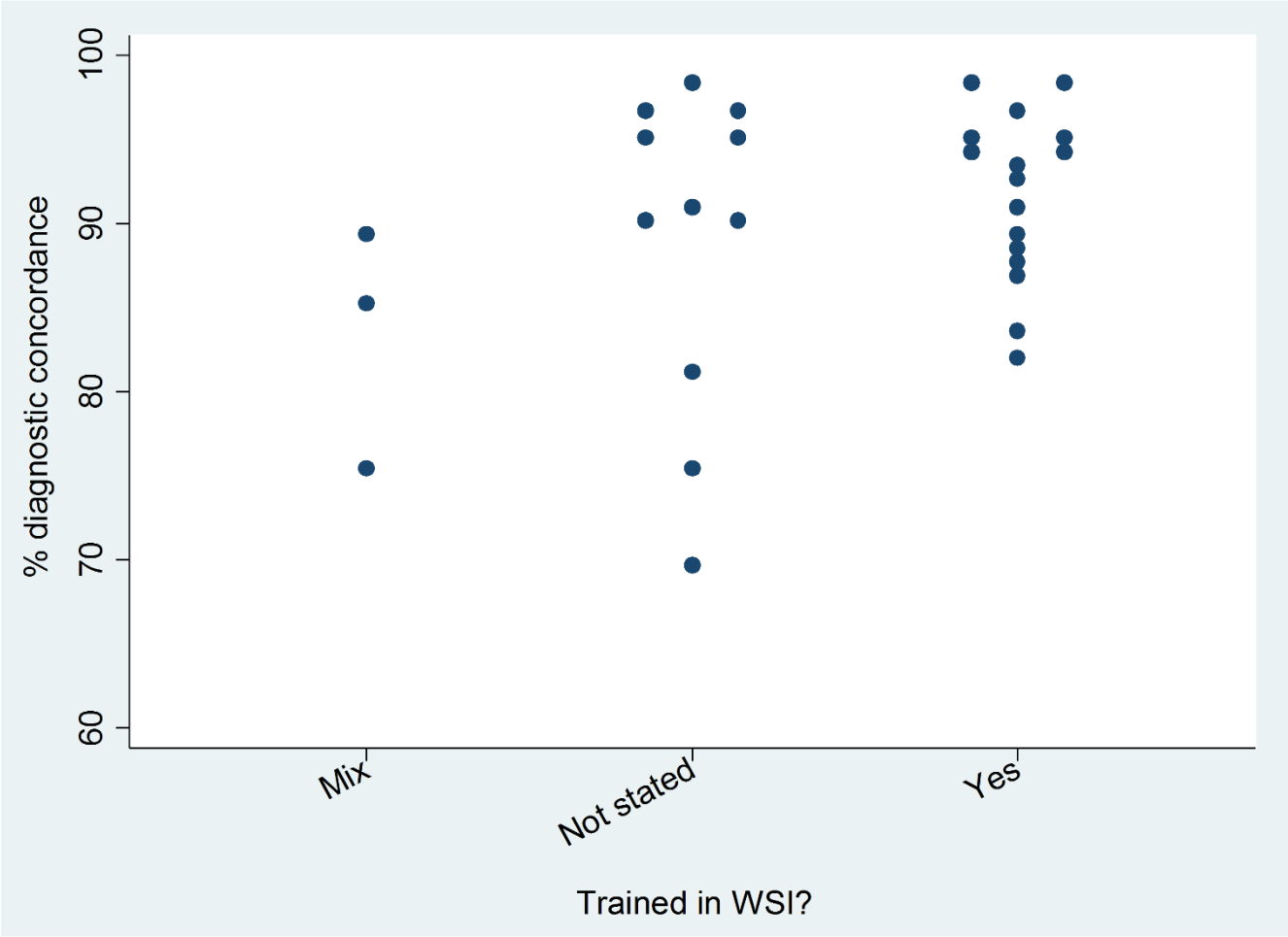
- 60% of included studies showed a concordance **≥90%,** of which 33% showed a concordance **≥95%**

- **Supports a high level of diagnostic concordance**

Results: Concordance by Year of Study



Results: Concordance by Training



Mean diagnostic concordance

92.4%

But are we looking at this the right way?

The clinical use claims described in the information provided have not been cleared or approved by the U.S. FDA nor are the products available in the United States.

Overview

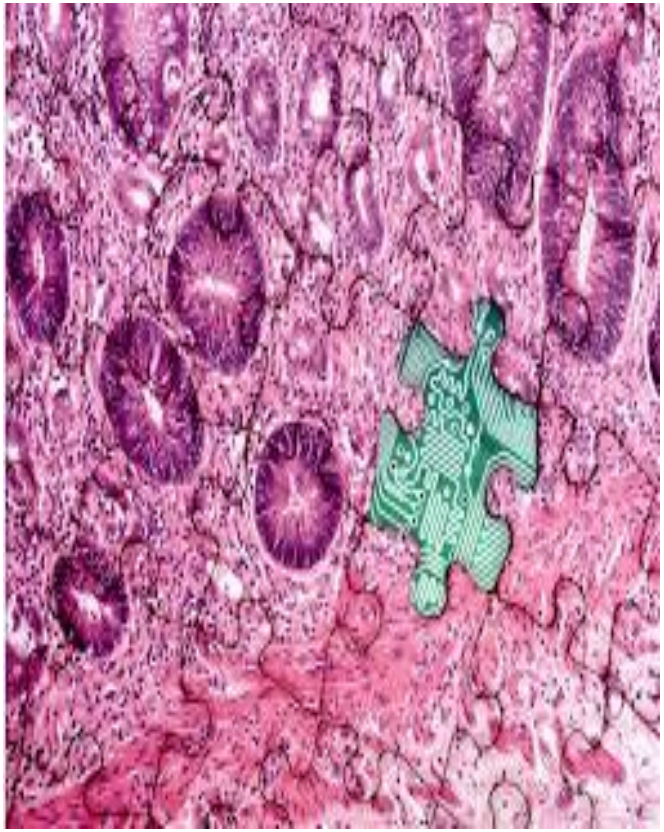
An introduction to Leeds

Concordance

Discordance

Validation and training protocol

Results from the breast validation pilot



Diagnostic discordance ...

- Mean diagnostic discordance of whole slide imaging and light microscopy

7.6%

- Does this matter?
  - To patients
  - To clinicians
  - To the pathologist

Archives of Pathology and Laboratory Medicine 2017

Original Article

A Systematic Analysis of Discordant Diagnoses in Digital Pathology Compared With Light Microscopy

Bethany J. Williams, MB, BS, BSc; Philip DaCosta, MBBS, MRCS, LRCP, FRCPath; Edward Goacher, BSc; Darren Treanor, MB, BSc, PhD, FRCPath

• *Context.*—Relatively little is known about the significance and potential impact of glass-digital discordances, and this is likely to be of importance when considering digital pathology adoption.

Objective.—To apply evidence-based medicine to col-

glass was the preferred diagnostic medium in 286 (85%), and digital in 44 (13%), with no consensus in 5 (2%). Twenty-eight discordances had the potential to cause moderate/severe patient harm. Of these, glass was the preferred diagnostic medium for 26 (93%). Of the 335

Royal College of Pathologists System of Categorisation for Discrepancies (2014)

Category	Description
A	Inadequate dissection, sampling or macroscopic description
B	Discrepancy in microscopy 1. A diagnosis which one is surprised to see from any pathologist 2. A diagnosis which is clearly incorrect, but which one is not surprised to see a small percentage of pathologists suggesting 3. A diagnosis where inter-observer variation is known to be large (eg. Difficult diagnosis, difference between 2 tumour grades)
C	Discrepancy in clinical correlation
D	Failure to seek a second opinion in an obviously difficult case
E	Discrepancy in report (includes misidentification)

Royal College of Pathologists – Categorisation of Discrepancies for Duty of Care Reviews

The Leeds Teaching Hospitals NHS
NHS Trust

UNIVERSITY OF LEEDS

Category	Description
1	**No impact on care**
2	**Minimal harm, no morbidity** (Delay in diagnosis or therapy only, < 3/12, unnecessary non-invasive further diagnostic efforts, unnecessary therapy without morbidity)
3	**Minor harm, minor morbidity** (Delay in diagnosis or therapy only, >3/12,unnecessary invasive further diagnostic efforts, delay in therapy with minor morbidity)
4	**Moderate harm, moderate morbidity** (Due to delay in diagnosis, due to otherwise unnecessary diagnostic efforts, due to otherwise unnecessary therapeutic efforts)
5	**Major harm, major morbidity** (Loss of limb, organ or function of organ system due to unnecessary diagnostic efforts, severe morbidity due to delayed or unnecessary therapeutic efforts, death)

Example of discordance analysis

The Leeds Teaching Hospitals NHS
NHS Trust
UNIVERSITY OF LEEDS

Colon biopsy	
True diagnosis	Focal active colitis
Discordant diagnosis	Normal colon
Preferred diagnostic modality	Glass
RCPath expression of concern	B2 (clearly incorrect, but would expect small proportion of pathologists to make same error)
RCPath potential for harm	3 (minor harm, minor morbidity)
Diagnostic categories	GI, Benign, Inflammation
Discordance	Missed diagnosis
Diagnostic tasks/objects/features	Finding small object, Neutrophils
Detail from paper	Granulocytes difficult to discern on digital. Improved at 40x.

Results

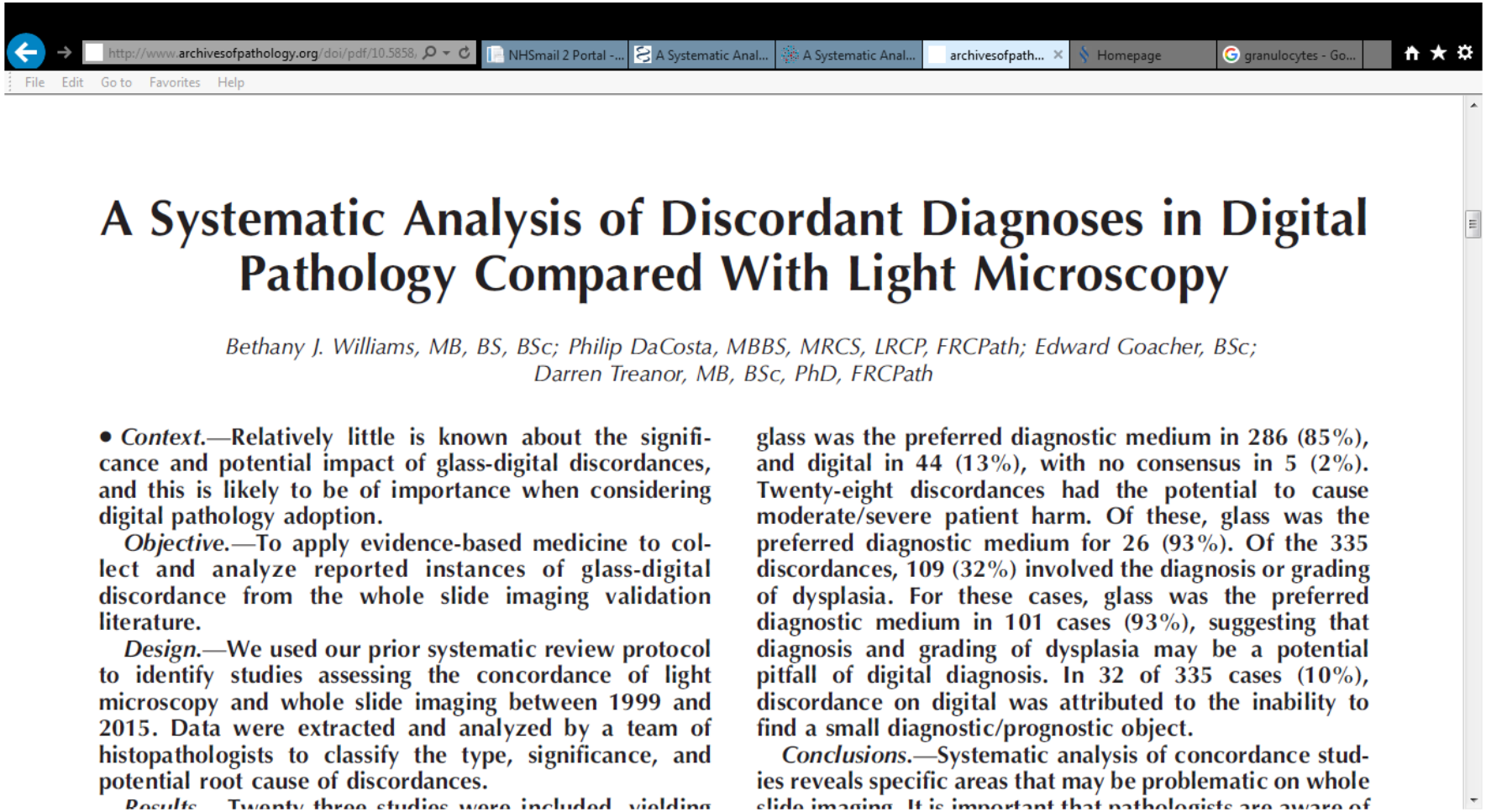
- 1300 abstracts reviewed
- 39 papers extracted
- 23 with detailed, extractable discordant diagnostic pairs
- Published 2006-2016, 82% post 2010

- 8069 glass: digital read pairs
- 335 instances of discordance recorded
- 4% of glass: digital comparisons

In brief:

- The digital diagnosis was preferred to the glass diagnosis in 15% of discordances
- An inferior diagnosis made on digital was no more likely to result in moderate/major patient harm than an inferior diagnosis made on glass
- An inferior diagnosis made on digital was no more likely to be a B1 “surprising error” than an inferior diagnosis made on glass

For more detail.....



http://www.archivesofpathology.org/doi/pdf/10.5858/... NHSmail 2 Portal -... A Systematic Anal... A Systematic Anal... archivesofpath... Homepage granulocytes - Go...

File Edit Go to Favorites Help

A Systematic Analysis of Discordant Diagnoses in Digital Pathology Compared With Light Microscopy

Bethany J. Williams, MB, BS, BSc; Philip DaCosta, MBBS, MRCS, LRCP, FRCPath; Edward Goacher, BSc; Darren Treanor, MB, BSc, PhD, FRCPath

• *Context.*—Relatively little is known about the significance and potential impact of glass-digital discordances, and this is likely to be of importance when considering digital pathology adoption.

Objective.—To apply evidence-based medicine to collect and analyze reported instances of glass-digital discordance from the whole slide imaging validation literature.

Design.—We used our prior systematic review protocol to identify studies assessing the concordance of light microscopy and whole slide imaging between 1999 and 2015. Data were extracted and analyzed by a team of histopathologists to classify the type, significance, and potential root cause of discordances.

Results.—Twenty-three studies were included, yielding glass was the preferred diagnostic medium in 286 (85%), and digital in 44 (13%), with no consensus in 5 (2%). Twenty-eight discordances had the potential to cause moderate/severe patient harm. Of these, glass was the preferred diagnostic medium for 26 (93%). Of the 335 discordances, 109 (32%) involved the diagnosis or grading of dysplasia. For these cases, glass was the preferred diagnostic medium in 101 cases (93%), suggesting that diagnosis and grading of dysplasia may be a potential pitfall of digital diagnosis. In 32 of 335 cases (10%), discordance on digital was attributed to the inability to find a small diagnostic/prognostic object.

Conclusions.—Systematic analysis of concordance studies reveals specific areas that may be problematic on whole slide imaging. It is important that pathologists are aware of

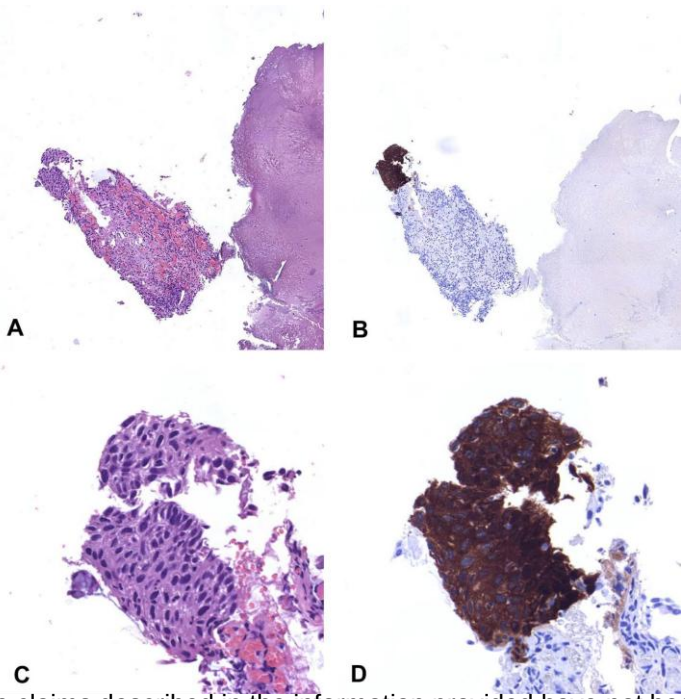
The clinical use claims described in the information provided have not been cleared or approved by the U.S. FDA nor are the products available in the United States.

Discussion: Possible pitfalls

- Diagnosing and grading dysplasia
- Finding small diagnostic/prognostic objects
- Specific items/textures

Dysplasia

- 108 discordances concerning diagnosis of dysplasia
- Glass diagnosis preferred in **94%** of all dysplasia related discordances
- Tendency to undercall dysplastic lesions



Ordi et al
Glass read – CIN3
Digital read - reactive

Diagnosing and grading dysplasia

- Low power – regions of interest
- High power – nuclear detail

- Potential for risk mitigation on digital:
  - *Education of pathologists*
  - *Individual learning curve/validation procedure*
  - *Workflow modification – eg. mandatory glass check*
  - *Routine scanning at 40x for diagnostic bx*

Finding small objects

Category	Object	Glass preferred	Digital preferred
Neoplasia	Small primary tumour	1	2
	Lymph node metastasis	3	
	Tumour microsatellite	1	
	Focal tumour invasion	3	3
Inflammation	Foci of inflammatory activity	8	
	Granulomas	3	
Micro-organisms	H Pylori	2	
	Candida	3	
Sparse cells	RS cells	2	
Focal benign features		5	2
Focal immunopositivity		1	

Finding small objects

- Navigation within slide
- Navigation between slides
- Display resolution and quality
- Scanning magnification

- Potential for risk mitigation on digital:
  - *Pathologist training – navigation, efficiency*
  - *Use of thumbnails*
  - *Indication of slides viewed*
  - *Incorporate checks into workflow*
  - *40x – micro-organisms*

Specific entities causing difficulty in included studies

- Neutrophils (19)
- Eosinophils (7)
- Mast cells (1)

- Nucleated red blood cells (8)
- Eosinophilic granular bodies (1)

- Blue mucin (1)
- Amyloid (2)

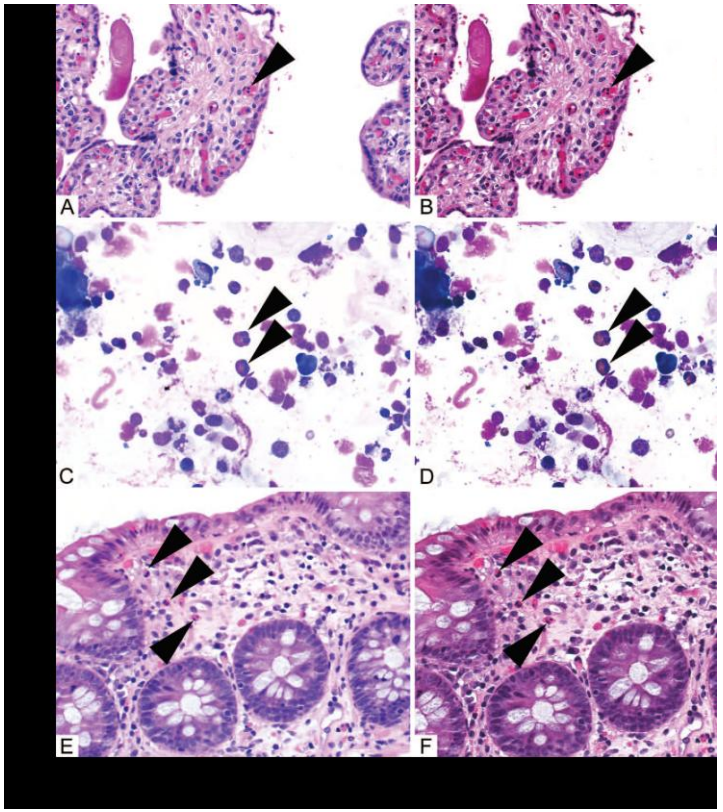


Image from Arnold et al

Mitigating "problem" cases

- Awareness of specialty specific pitfalls
- Familiarisation with specialty specific digital images
- Default 40x scanning / request for 40x rescan for specific cases

Where do we go from here?

- Collection and dissemination of glass:digital discordance data

- Importance of education for pathologists – navigation training, specialty specific pitfalls
- Incorporation of workflow modifications for “difficult” specimens

- The bottom line : Important that individual pathologists are competent and confident in digital diagnosis

Overview

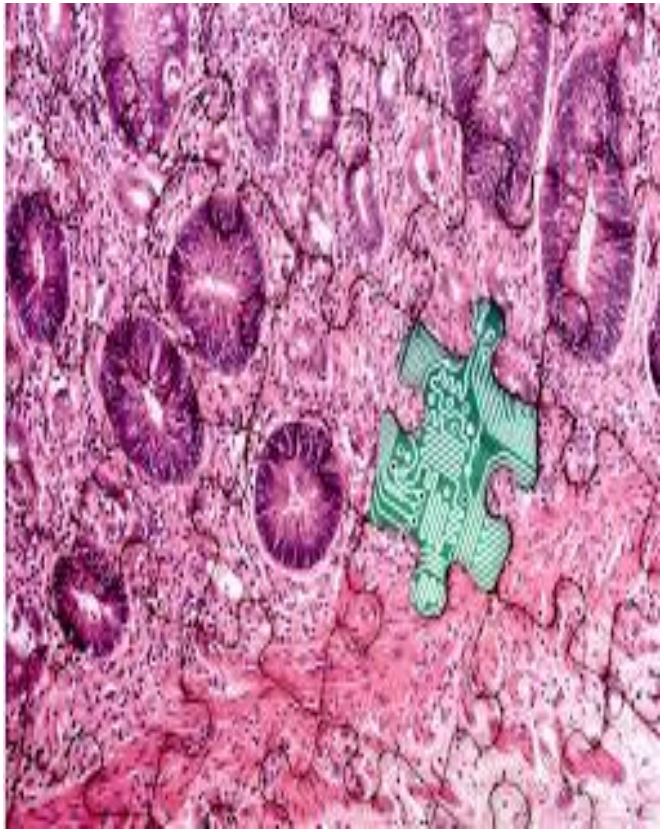
An introduction to Leeds

Concordance

Discordance

Validation and training protocol

Results from the breast validation pilot



The Leeds approach to clinical validation....

The Leeds Teaching Hospitals NHS
NHS Trust
UNIVERSITY OF LEEDS

Pragmatic

Patient safety focused

Professional engagement and education

The clinical use claims described in the information provided have not been cleared or approved by the U.S. FDA nor are the products available in the United States.

How are we validating digital pathology?

- Novel validation protocol
- Pathologist-centric
- Utilises evidence base

- Enables self-identification of digital diagnosis pitfalls, allows pathologist to gain confidence in risk mitigated environment, with early exposure to live digital reporting

- Allows development of workflow modifications to aid confident and competent diagnosis

Royal College Guidelines

The Royal College of Pathologists

Pathology: the science behind the cure

The clinical use claims described in the information provided have not been cleared or approved by the U.S. FDA nor are the products available in the United States.

Validation summary

Training

- 1:1 formalized training in digital microscope use
- Observed practice with feedback

Validation 1
Training set

- Test set of 20 challenging and informative specialty specific cases
- View on digital, make notes, compare with glass immediately

Validation 2
Live cases

- Entire workload scanned (2 months)
- Diagnosis made on digital immediate glass check before sign out

The clinical use claims described in the information provided have not been cleared or approved by the U.S. FDA nor are the products available in the United States.

Sample breast validation training set scope:

The Leeds Teaching Hospitals NHS
NHS Trust

UNIVERSITY OF LEEDS

1.1 Specimen type:
Core biopsies
Mammotome biopsies
Diagnostic excision specimens
Therapeutic excision specimens

1.2 Tissue type:
Nipple
Breast
Lymph node

1.3 Stains:
H and E
Ck/myoepithelial markers
ER/PR immuno
Her2 immuno

1.4: Diagnoses
Benign breast disease/inflammatory conditions
Calcification (need to check glass for weddelite)
Epithelial atypia
In situ carcinoma
Invasive carcinoma incl. identification of special types
Benign fibroepithelial lesions
Malignant fibroepithelial lesions

1.5: Tasks
Diagnosis
Grading inc. mitotic count and nuclear assessment
Identification of microinvasion/LVSI
Identification of calcium in screening specimens
Interpretation of immunostains incl. Her2, PR
Interpretation of myoepithelial/ck immunos
Identification of micrometastases

2. Potential pitfalls
Recognising epithelial atypia
Identifying microinvasion/LVSI
Identification of micrometastases
Identification of calcium (weddelite)
Granulomatous inflammation
Lobular carcinoma

Sample training set case:

Breast Test Set
Case 1 (3 slides)
Specimen type - Left breast lump

Clinical details - Diagnostic excision biopsy. Left breast 12 o'clock. Elective excision U2 C2 fibroadenoma.

Macro - Lump of fibrofatty breast parenchyma which weighs 10.1g and measures 35x24x20mm. 3 representative pieces.

Sample training set case:

Digital diagnosis

How would you rate your confidence in your diagnosis?

1	2	3	4	5	6	7
Not confident At all						Very Confident

Sample test set case:

Glass diagnosis - has your impression of the case changed?

How would you rate your confidence in your glass diagnosis?

1	2	3	4	5	6	7
Not confident At all						Very Confident

Comments

Stage 2 Validation – live cases

The Leeds Teaching Hospitals NHS
NHS Trust
UNIVERSITY OF LEEDS

- All cases scanned prospectively
- Diagnosis made on digital, with immediate glass reconciliation prior to sign out
- At least 2 months WTE workload
- Approximately 200-250 cases (appropriate case mix of biopsies:resections)
- Regular meetings with specialty group and trainer (roughly every 20% of workload)
- Discuss discordances, difficulties, feedback
- Plan workflows
- Difficult/discordant cases collected and assembled in digital library and shared within department

Validation Documentation

- Training performed
- Meetings attended
- Summary of glass-digital correlation throughout validation
- List and comments on discordances noted throughout

VALIDATION SUMMARY AND OUTCOME

Pathologist

Specialty

Trainer

The pathologist has completed a course of digital pathology training and validation procedure. As a result of discussion between the pathologist and trainer, the mutually agreed outcome of this validation procedure is:

1. Fully validated for digital primary diagnosis in the specified diagnostic area.
2. Validated for digital primary diagnosis in the specified diagnostic area, with some exceptions/workflow modifications (see below)
3. Not validated for digital primary diagnosis in the specified diagnostic area at this time.

For outcome 2, please describe exceptions/modifications:

Pathologist _____

Trainer _____

Date _____

The clinical use claims described in the information provided have not been cleared or approved by the U.S. FDA nor are the products available in the United States.

Overview

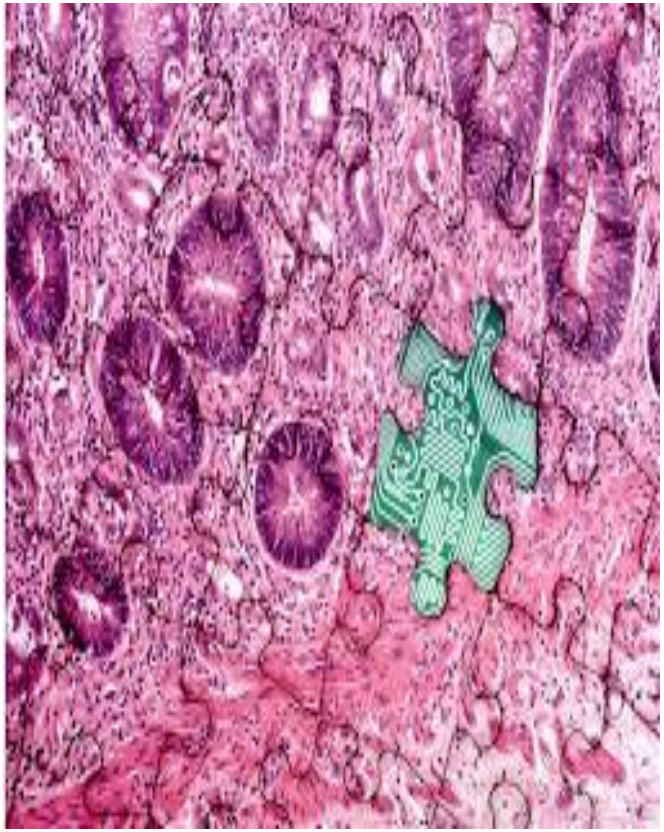
An introduction to Leeds

Concordance

Discordance

Validation and training protocol

Results from the breast validation pilot



Clinical deployment – breast pathology

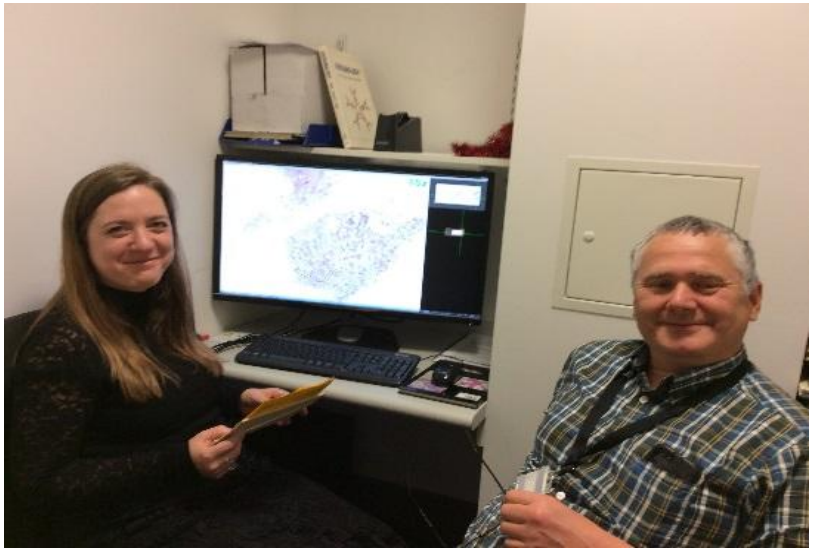
The Leeds Teaching Hospitals NHS
NHS Trust
UNIVERSITY OF LEEDS

- Pilot study to inform wider digital deployment

- 4 consultant breast pathologists at SJUH
- 3 have completed validation, 1 in progress

- Enthusiasts and skeptics!

- 40x scans
- 6MP medical grade displays



The clinical use claims described in the information provided have not been cleared or approved by the U.S. FDA nor are the products available in the United States.

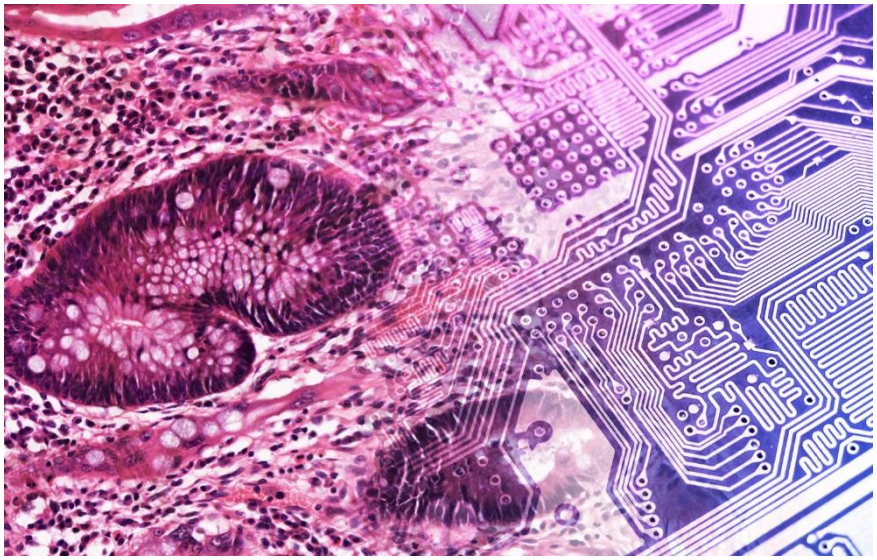
Training set areas of difficulty:

- Mitotic figure counting
- Micrometastases
- Weddelite
- Nuclear atypia

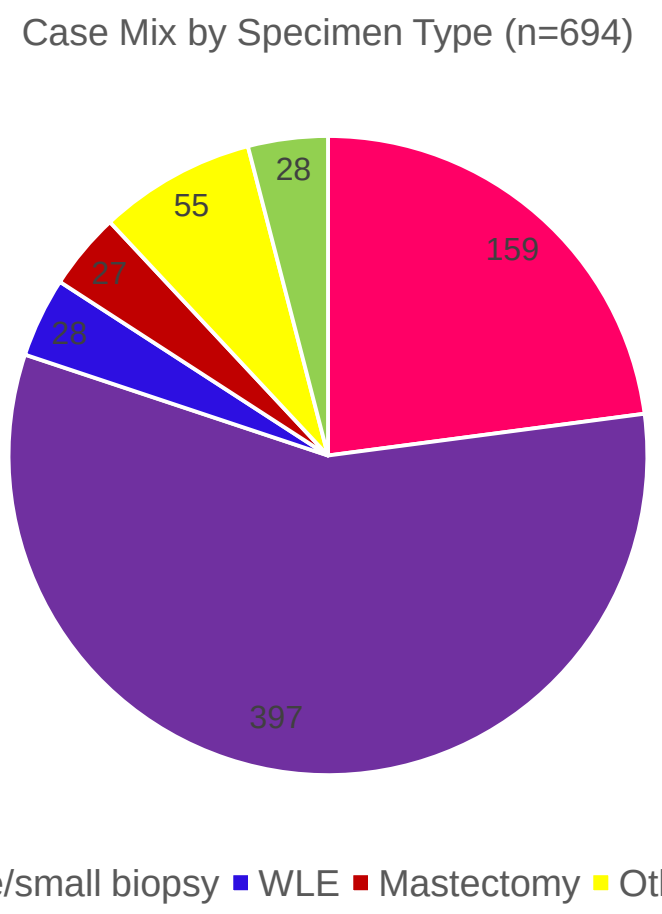
Live reporting

- 694 cases
- 3,500 slides

- Standard, large
- H&E, immuno

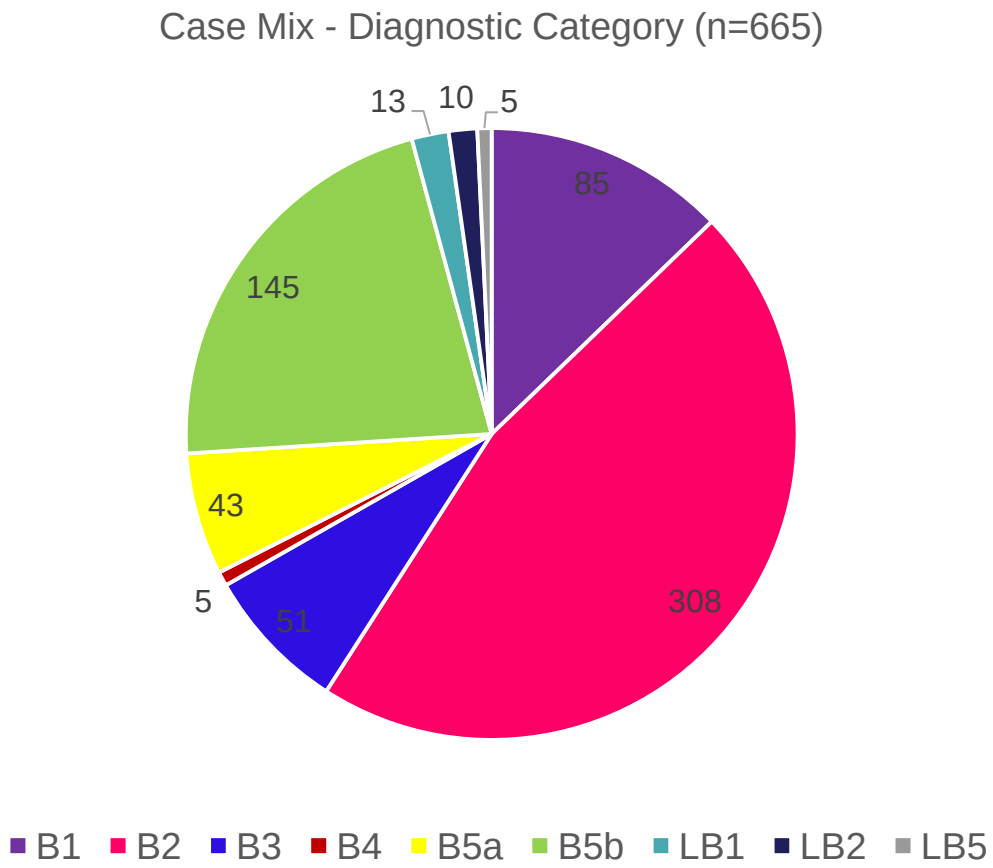


Live Cases Specimen type



The clinical use claims described in the information provided have not been cleared or approved by the U.S. FDA nor are the products available in the United States.

Live Cases : Diagnostic category



Concordance statistics

The Leeds Teaching Hospitals NHS
NHS Trust

UNIVERSITY OF LEEDS

	Pathologist 1	Pathologist 2	Pathologist 3	All pathologists
Technical failure rate	0.7%	1.4%	1.0%	1.0%
Complete concordance	95.0%	96.2%	97.4%	96.2%
Any observable difference	5.0%	3.8%	2.6%	3.8%
Complete clinical concordance	99.3%	99.1%	98.5%	98.8%
Clinically significant observable difference	0.7%	0.9%	1.5%	1.2%

The clinical use claims described in the information provided have not been cleared or approved by the U.S. FDA nor are the products available in the United States.

True discordances

Specimen	Digital diagnosis	Glass diagnosis
Core biopsy	G2 invasive ductal carcinoma	G3 invasive ductal carcinoma
Vacuum biopsy	Benign phyllodes tumour (B3)	Fibroadenoma with inflammation (B2)
Vacuum biopsy	Benign (B2)	plus AIDP (B3)
Vacuum biopsy	Sclerosing adenosis (B3)	Sclerosing adenosis, small focus DCIS (B5a)
Vacuum biopsy	Microcysts	Microcysts and weddelite
Sentinel node	Benign	Isolated tumour cells
Vacuum biopsy	Columnar cell change, calcification (B2)	Columnar cell change, calcification, single focus atypical cells (B3)
Vacuum biopsy	Small focus DCIS	DCIS missed on initial glass review, pathologist had to refer back to digital slide to locate lesion on glass and confirm its presence

Diagnostic confidence

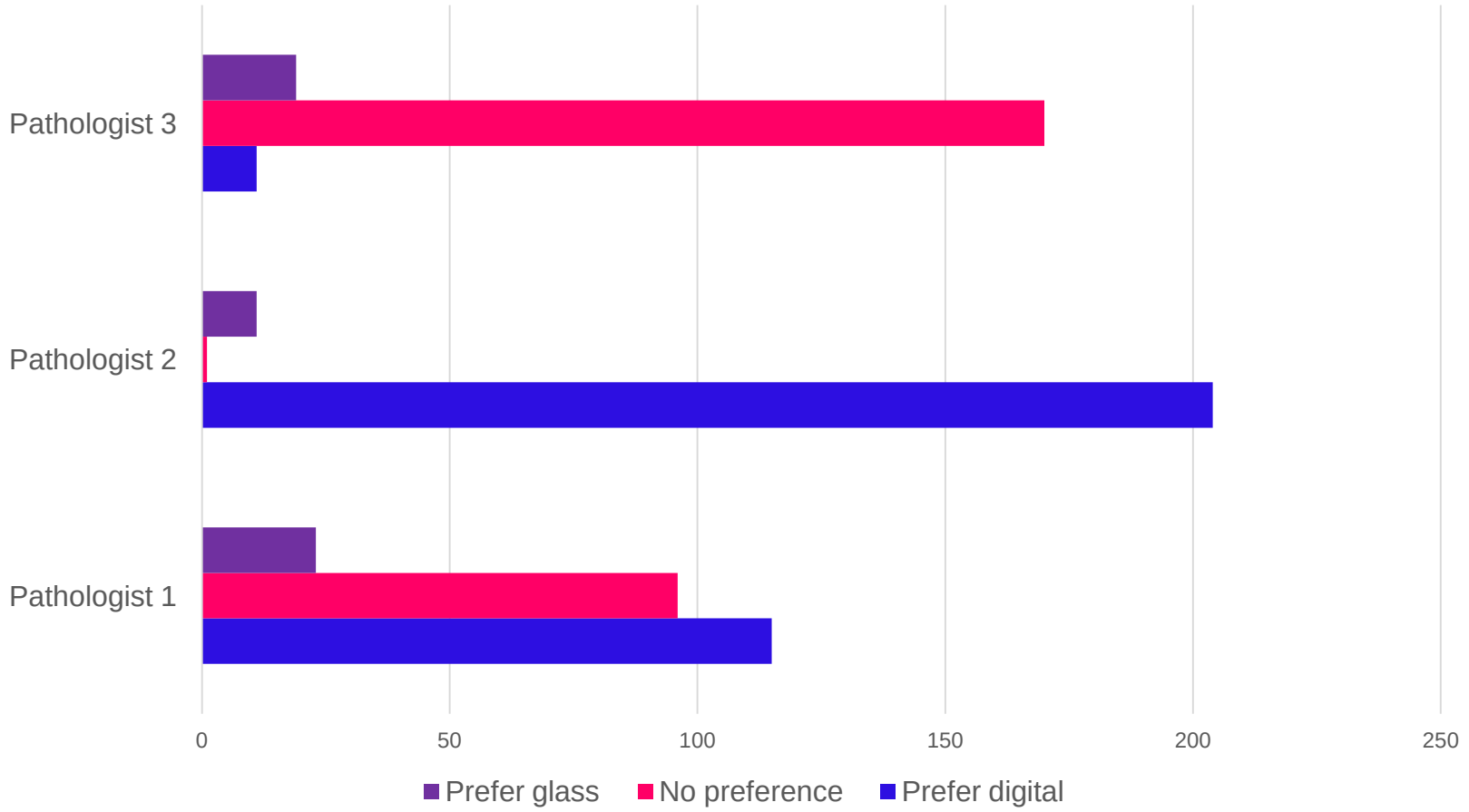
How would you rate your confidence in your diagnosis?

1 2 3 4 5 6 7

Not confident At all — Very Confident

	Digital slides		Glass slides	
	Mean confidence	**Range**	**Mean confidence**	**Range**
Pathologist 1	6.70	4-7	6.80	4-7
Pathologist 2	6.90	4-7	6.90	4-7
Pathologist 3	6.79	0-7	6.99	6-7

Reporting preferences



The clinical use claims described in the information provided have not been cleared or approved by the U.S. FDA nor are the products available in the United States.

Validation Completion

- 1. Fully validated for digital primary diagnosis in the specified diagnostic area.
- 2. Validated for digital primary diagnosis in the specified diagnostic area, with some exceptions/workflow modifications (see below)
- 3. Not validated for digital primary diagnosis in the specified diagnostic area at this time.

Exceptions/modifications

- 1. Invasive cancers where mitotic count would affect overall grade
- 2. Core biopsies containing fibroepithelial lesions with apparently cellular stroma
- 3.Vacuum biopsies for calcification where no calcification is seen on digital slides, if there is radiological evidence that calcium is present in the specimen
- 4.Any rare/challenging case, including difficult referral cases

What did they think?

- “the low power view makes things much easier”
- “for large cases and multiple levels, the time saving is huge”
- “great for teaching – I can do live diagnosis with trainees sat by my side”
- “Validation for each new tissue type category and each new user is essential so the user can understand the limitations, adjust to how things look and tailor a suitable workflow”

What did they think?

The clinical use claims described in the information provided have not been cleared or approved by the U.S. FDA nor are the products available in the United States.

On the horizon in Leeds....

The Leeds Teaching Hospitals NHS
NHS Trust

UNIVERSITY OF LEEDS

2017

- Immunohistochemistry validation and deployment
- Neuropathology primary diagnosis validation and deployment
- MDT

2018

- 100% digitisation – all slides scanned
- Specialty by specialty validation

Final thought….

"My interest is in the future because I'm going to spend the rest of my life there…."

The clinical use claims described in the information provided have not been cleared or approved by the U.S. FDA nor are the products available in the United States.

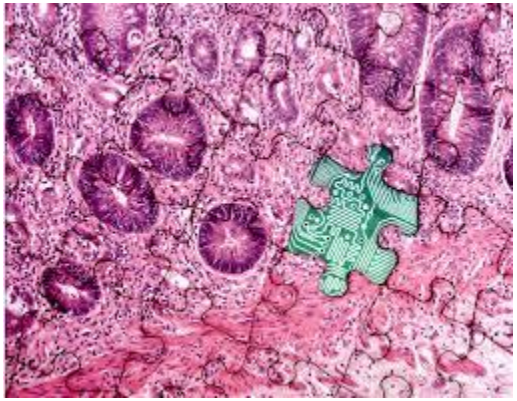
Acknowledgements

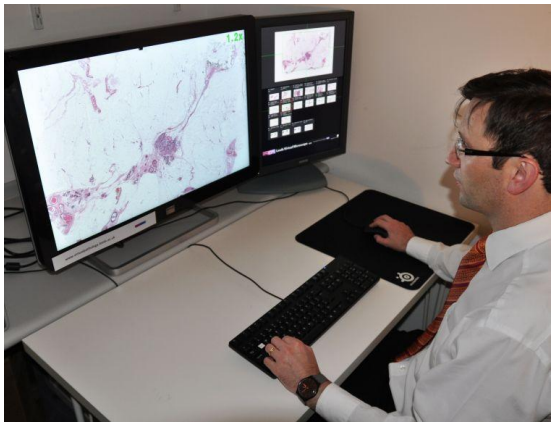
- Thanks to our participants, Prof. Andy Hanby, Dr Rebecca Millican-Slater, Dr Anju Nijhawan, Dr Eldo Verghese

- Thanks to the digital pathology team at Leeds – Dr Darren Treanor, Ms Chloe Lockwood, Mr Basharat Hussain

- This work was supported by funds from Leica Biosystems, as part of a strategic partnership

Thank you!





bethany.williams2@nhs.net

17565 REV.A IMC-8119-REV-A 10/2017