Performance of an AI-based cancer diagnosis system in France’s largest network of pathology institutes

Daphna Laienfeld1, Judith Sandbank1,2, Chaim Linhart1, Lilach Bien1, Cobi Reouven1, Roei Harduf1, Mahul Amin3, Olivier Levrer1, Stéphane Rossat4 & Delphine Raoux4

1 Ibex Medical Analytics, Israel 2 Institute of Pathology, Maccabi Healthcare Services, Israel 3 University of Tennessee Health Science Center, USA 4 MédiPath, France

BACKGROUND

- Prostate cancer is the second leading cause of cancer-related deaths, compounded by complexity of diagnosis, insufficient diagnostic reproducibility, and a growing shortage in pathologists.
- Ibex Medical Analytics, focusing on AI-based cancer diagnostics, has developed an algorithm that identifies various cell types and tissue structures within whole slide images of prostate core needle biopsies (PCNBs), such as cancerous glands (of Gleason patterns 3, 4 and 5), high-grade PIN, inflammation and atrophic glands.
- MédiPath is the largest network of pathology institutes in France, servicing 370,000 patients including 3,000 prostate biopsies annually.

Objective

This study was aimed to assess the performance of Ibex’s algorithm on benign cases from 4 labs within the network, and to detect and characterize missed cancer cases.

METHODS

Data

- Calibration set included 150 anonymized cases (100 benign, 50 cancer) including 1,140 H&E slides with various clinical features (e.g., Gleason grade).
- 40 slides from the Calibration set were annotated and used as part of the algorithm’s training data.
- Study set included 100 benign cases (801 H&E slides).

Algorithmic Analysis

- Calibration set was used to determine a threshold score optimizing for specificity and sensitivity.
- Ibex prostate algorithm was run on the Study set.

RESULTS

DISTRIBUTION OF CANCER SCORES

- Calibration Set (100 benign & 50 cancer cases)
- Study Set (100 benign cases)

EXAMPLES OF CANCER

- Low-Grade Cancer: Diagnosed G3+3 (both reviewers).
- High-Grade Cancer: Diagnosed G4+3 (both reviewers).

EXAMPLES OF OTHER FEATURES

- High-Grade PIN: Low Probability
- Inflammation: Low Probability
- Atrophy: Low Probability

RESULTS OF EXPERTS REVIEW

Diagnoses Distribution

<table>
<thead>
<tr>
<th>Reviewers’ Diagnosis</th>
<th># slides</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASAP/ ASAP</td>
<td>7</td>
</tr>
<tr>
<td>ASAP/ Benign</td>
<td>1</td>
</tr>
<tr>
<td>ASAP/ ASAP</td>
<td>1</td>
</tr>
<tr>
<td>Cancer/ ASAP</td>
<td>4</td>
</tr>
<tr>
<td>Cancer/ Cancer</td>
<td>50</td>
</tr>
</tbody>
</table>

Characterization of Missed Cancers

<table>
<thead>
<tr>
<th>Case</th>
<th>Original Diagnosis</th>
<th>Review</th>
<th>Cancer Grade</th>
<th>Tumor Size (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASAP</td>
<td>Benign</td>
<td>Benign</td>
<td>3.3</td>
<td>3</td>
</tr>
<tr>
<td>ASAP</td>
<td>ASAP</td>
<td>ASAP</td>
<td>4.6</td>
<td>4</td>
</tr>
<tr>
<td>ASAP</td>
<td>Cancer</td>
<td>Cancer</td>
<td>5.2</td>
<td>5</td>
</tr>
<tr>
<td>ASAP</td>
<td>ASAP</td>
<td>ASAP</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>ASAP</td>
<td>ASAP</td>
<td>ASAP</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>ASAP</td>
<td>ASAP</td>
<td>ASAP</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>ASAP</td>
<td>ASAP</td>
<td>ASAP</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>ASAP</td>
<td>ASAP</td>
<td>ASAP</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Threshold: was set to 0.86255 to optimize for sensitivity & specificity.

Specificity = 99.8% (11 benign slides out of 917 pass the threshold).
Sensitivity = 98.8% (118 cancer slides out of 225 pass the threshold).

Calibration set.
25 of the Study set slides, reported as benign, passed the threshold and were sent to review.
After correcting diagnoses based on the review, the specificity of the Study set was similar to that of the Calibration set.

Total 25 Cancer was diagnosed in 12 unique cases by at least one reviewer.

CONCLUSIONS

1. The Ibex algorithm for detecting prostate cancer is extremely accurate, and it can identify additional features that are clinically important, such as high-grade PIN, atrophy and inflammation.
2. Prostate cancer, typically low-grade tumors, might be missed by pathologists. High-grade tumors are also misdiagnosed, especially when they are small and appear in a single slide.
3. Pathology institutes can utilize AI-based systems to increase the accuracy, efficiency and turnaround time of the diagnostic process.