Validation and Clinical Deployment of an AI-Based Solution for Detection of Gastric Adenocarcinoma and Helicobacter pylori in Gastric Biopsies

Judith Sandbank1, Geraldine Sebag2, Ayala Arad3, Rachel Mikulinsky4, Tali Feinberg5, Issar Yazbin6, Inbal Gross7, Ronen Heled8, Maya Grinwald9, Manuela Vescler10, Chaim Linhart11

1 Institute of Pathology, Maccabi Healthcare Services, Rehovot, Israel; 2 Ibex Medical Analytics, Tel Aviv, Israel

Presented at USCAP 2022

BACKGROUND

- Computer-assisted diagnostic solutions to evaluate gastric biopsies hold promise to increase accuracy and efficiency in diagnosis
- AI-based solutions for clinical use in the diagnosis of gastric biopsies should accurately detect the following:
  - Carcinoma / HG Dysplasia
  - Helicobacter pylori
  - Neuroendocrine neoplasms
  - Adenoma, Low-Grade Dysplasia and more
- Maccabi Healthcare Services (MHS) is a large healthcare provider with 200,000 histology accessions annually, of which ~ 18,000 are gastric biopsies

Objective

- To clinically validate the performance of an AI-based algorithm in the detection of gastric adenocarcinoma (AdC), high-grade (HG) dysplasia, and Helicobacter pylori (H. pylori), and to implement it in routine clinical workflow

METHODS

Algorithm Development

- Based on fully-supervised, multi-layered convolutional neural networks (CNNs) and specifically designed to classify and analyze whole slide images (WSI)
- Trained on >750,000 image samples from multiple labs and geographies
- Multiple classes of benign and pathologic features manually annotated by a team of senior pathologists

Algorithm Validation

- Dataset of 1,845 gastric biopsies (AdC-enriched) from MHS, independent from the training set
- Slides were scanned using a Philips IntelliSite Scanner at 40x magnification (0.25 µm/pixel)
- Run in a blinded manner and results were compared to ground truth diagnoses (based on the original pathology reports)

Clinical Deployment

- The algorithm was implemented as a Second Read system at MHS, running in parallel to the pathologist’s routine workflow and raising alerts when encountering discrepancies between the automated analysis and the pathologist’s diagnosis, prompting a second pathologist’s review
- Galen Gastric raises two types of alerts:
  1. Slides from benign cases with high probability for Adenocarcinoma or High-grade Dysplasia
  2. Slides from benign cases with high probability for Helicobacter pylori and related gastritis.

RESULTS

Galen Gastric Performance Study

The Galen Gastric algorithm demonstrated very high performance for the detection of Adenocarcinoma (AdC), High-grade Dysplasia and Helicobacter pylori.

<table>
<thead>
<tr>
<th>Study</th>
<th>Analysis</th>
<th>Number of cases (N)</th>
<th>AdC</th>
<th>Specificity</th>
<th>Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benign cases</td>
<td>AUC (95% CI)</td>
<td>180 (160-200)</td>
<td>0.914</td>
<td>97.28%</td>
<td>96.96%</td>
</tr>
<tr>
<td>AdC</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>HG</td>
<td>0.982</td>
<td>87.28%</td>
<td>97.40%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>True positives</td>
<td>0.946</td>
<td>91.7%</td>
<td>91.61%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Algorithm performance in the retrospective and prospective studies

*AUC for AdC / HG Dysplasia was based only on seven positive cases, out of which six reached the highest scores in the dataset

Galen Gastric Clinical Deployment

<table>
<thead>
<tr>
<th>Study</th>
<th>Total</th>
<th>Adenocarcinoma (%)</th>
<th>High-grade dysplasia (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrospective</td>
<td>692</td>
<td>0.994</td>
<td>97.28%</td>
</tr>
<tr>
<td>Prospective</td>
<td>639</td>
<td>0.932</td>
<td>87.86%</td>
</tr>
</tbody>
</table>

Table 2: Pathologist’s diagnoses for the cases analyzed by Galen Gastric

- Galen Gastric can support optimization of H. pylori workflow:
  - Automatic screening for H. pylori diagnosis, streamlining H. pylori-IHC stain ordering, and reducing pathologists review and turnaround time (TAT)
  - Implementation of a comprehensive and efficient quality control process for increased diagnostic accuracy

DISCUSSION & CONCLUSIONS

1. Although gastric cancer is rare, it has poor prognosis, and therefore timely diagnostic prevention of diagnostic mistakes is of high importance
2. Galen Gastric (AI) demonstrated accurate detection of a broad range of pathological features in clinical use, including adenocarcinoma, HG dysplasia, H. pylori, being an effective and user-friendly diagnostic support tool for pathologists
3. The AI-powered Galen Gastric proposes a much more cost-effective diagnosis workflow, enabling efficient detection of Helicobacter pylori together with reduction in turnaround time and minimizing ordering of additional stains
4. AI allows accurate detection of multiple pathological features beyond cancer detection, such as adenoma, LG dysplasia, neuroendocrine neoplasms and more