Deployment of a Deep Learning Model to Assist Pathologists with Donor Kidney Biopsy Evaluation

Joseph P. Gaut, MD, PhD1, Jon Marsh, PhD1, Mike Blackford2, Victoria Swamidass2, Rick Berry2, and S. Joshua Swamidass, MD, PhD1

1Washington University School of Medicine in St. Louis, Dept. of Pathology and Immunology; 2PlatformSTL

INTRODUCTION / BACKGROUND

• Over 100,000 patients are currently on the waiting list for a kidney transplant in the United States (1).
• 13 patients will die every day awaiting this life-saving treatment. There is a chronic, global shortage of donor kidneys available to meet this high demand (1,2).
• Unfortunately, there is a high rate of donor kidney discard, most commonly related to the biopsy evaluation (3,4).
• Acceptable donor kidneys are being discarded because of variable and inaccurate donor biopsy results (5).
• The current study proposes a web-based digital pathology workflow implementing a deep learning model to assist pathologists with quantitative evaluation of donor kidney biopsy frozen section whole slide images.

MATERIALS AND METHODS

• A deep learning model was previously developed to quantify non-sclerosed and globally sclerosed glomeruli in frozen section whole slide images (5, 6).
• A cloud-based software design to 1) deploy the deep learning model 2) display the deep learning model results in an accessible manner, and 3) provide a means for the pathologist to review and modify the deep learning model output.
• 20 consecutive donor kidney wedge frozen section biopsies from the Washington University Image Database were utilized for this study.
• Ground truth was established by serial annotation by 2 renal pathologists using previously described methods.
• Three pathologists utilized the cloud-based software to evaluate 20 kidney biopsies for percent global glomerulosclerosis.
• The pathologist and model results were compared with ground truth. Intra-class correlation was used to assess agreement.

RESULTS

Figure 1. Guided Review Software

• A software program was developed with an intuitive user-interface design to upload and analyze digital whole slide images of donor kidney biopsies using the previously published deep learning algorithm for percent global glomerulosclerosis quantitation.
• The guided review feature was designed specifically with the clinical workflow in mind to optimize efficiency and performance (Figure 1).
• This feature populates the scanned whole slide image with the deep learning model output using blue and red circles, representing non-sclerotic and sclerotic glomeruli, respectively.
• A window insert is displayed on the right-hand side of the screen ensuring the pathologist reviews the entire tissue. As the pathologist views the dashed-line box area of the image, they may remove or add glomeruli with simple, intuitive features.
• The model results showed excellent correlation with ground truth assessment (r=0.97)(Figure 2).
• Pathologists used the software to quantify percent global glomerulosclerosis on 20 frozen section cases and showed excellent agreement with one another and the model (Figure 2).

Figure 2. Pathologist and Model Correlation

<table>
<thead>
<tr>
<th>#GGS</th>
<th>#NGS</th>
<th>#GGG</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.94</td>
<td>0.94</td>
<td>0.95</td>
</tr>
</tbody>
</table>

#GGS – no global glomerulosclerosis
#NGG – no non-globally sclerosed glomeruli
#GGG – percent global glomerulosclerosis

CONCLUSIONS/FUTURE DIRECTION

Conclusions
• Deployment of a deep learning model to quantify percent global glomerulosclerosis using intuitive software designed to ensure complete and efficient tissue examination is a feasible approach to leveraging deep learning in clinical practice.
• The model showed excellent agreement with ground truth as established by expert renal pathologists.
• Reviewing pathologists were able to intuitively navigate the software and evaluate percent global glomerulosclerosis.
• The reviewing pathologists showed excellent correlation with ground truth, the model, and one another when using the software.
• Further studies are planned to evaluate the variability and accuracy of donor kidney biopsy examination using this approach with the ultimate aim of decreasing inappropriate donor kidney discard.

REFERENCES


Disclosure Statement

The work was supported in part by funding from the New and Mid-America Transplant Foundation. J. Gaut, J. Marsh, and S.J. Swamidass may receive royalty income based on technology developed by J. Gaut, J. Marsh, and L. Sawardoss (platformSTL) by Washington University in St. Louis. The technology is evaluated in this research. S.J. Swamidass has a financial ownership interest in platformSTL and may financially benefit if the company is successful in marketing its product(s) that are related to this research.