



Digital pathology assessment of kidney glomerular filtration barrier ultrastructure in an animal model of podocytopathy

Aksel Laudon^{1,2,†}, Zhaoze Wang^{1,†}, Anqi Zou^{3,†}, Richa Sharma², Jiayi Ji², Winston Tan², Connor Kim¹, Yingzhe Qian¹, Qin Ye¹, Hui Chen⁴, Joel M Henderson⁴, Chao Zhang^{3,*}, Vijaya B Kolachalama^{3,5,*} , and Weining Lu^{2,4,*} 

¹Department of Biomedical Engineering, Boston University, Boston, MA 02215, United States

²Nephrology Section, Department of Medicine, Boston University Chobanian & Avedisian School of Medicine, Boston Medical Center, Boston, MA 02118, United States

³Computational Biomedicine Section, Department of Medicine, Boston University Chobanian & Avedisian School of Medicine, Boston Medical Center, Boston, MA 02118, United States

⁴Department of Pathology and Laboratory Medicine, Boston University Chobanian & Avedisian School of Medicine, Boston Medical Center, Boston, MA 02118, United States

⁵Department of Computer Science and Faculty of Computing & Data Sciences, Boston University, Boston, MA 02215, United States

*Corresponding authors. Weining Lu, Pathology and Laboratory Medicine Nephrology Section, Department of Medicine, Boston University Chobanian & Avedisian School of Medicine Boston Medical Center, 650 Albany Street, Boston, MA 02118, United States. Tel: 617-414-1770. E-mail: wlu@bu.edu; Vijaya B Kolachalama, Computational Biomedicine Section, Department of Medicine, Boston University Chobanian & Avedisian School of Medicine, Boston Medical Center, 650 Albany Street, Boston, MA 02118, United States. Tel: 617-358-7253; E-mail: vkola@bu.edu; Chao Zhang, Computational Biomedicine Section, Department of Medicine, Boston University Chobanian & Avedisian School of Medicine, Boston Medical Center, 650 Albany Street, Boston, MA 02118, United States. Tel: 617-358-8535; E-mail: chz2009@bu.edu

[†]These authors contributed equally to this work.

Abstract

Transmission electron microscopy (TEM) images can visualize kidney glomerular filtration barrier ultrastructure, including the glomerular basement membrane (GBM) and podocyte foot processes (PFP). Podocytopathy is associated with glomerular filtration barrier morphological changes observed experimentally and clinically by measuring GBM or PFP width. However, these measurements are currently performed manually. This limits research on podocytopathy disease mechanisms and therapeutics due to labor intensiveness and inter-operator variability. We developed a deep learning-based digital pathology computational method to measure GBM and PFP width in TEM images from the kidneys of Integrin-Linked Kinase (ILK) podocyte-specific conditional knockout (cKO) mouse, an animal model of podocytopathy, compared to wild-type (WT) control mouse. We obtained TEM images from WT and ILK cKO littermate mice at 4 weeks old. Our automated method was composed of two stages: a U-Net model for GBM segmentation, followed by an image processing algorithm for GBM and PFP width measurement. We evaluated its performance with a 4-fold cross-validation study on WT and ILK cKO mouse kidney pairs. Mean [95% confidence interval (CI)] GBM segmentation accuracy, calculated as Jaccard index, was 0.73 (0.70–0.76) for WT and 0.85 (0.83–0.87) for ILK cKO TEM images. Automated and manual GBM width measurements were similar for both WT ($P = .49$) and ILK cKO ($P = .06$) specimens. While automated and manual PFP width measurements were similar for WT ($P = .89$), they differed for ILK cKO ($P < .05$) specimens. WT and ILK cKO specimens were morphologically distinguishable by manual GBM ($P < .05$) and PFP ($P < .05$) width measurements. This phenotypic difference was reflected in the automated GBM ($P < .05$) more than PFP ($P = .06$) widths. Our deep learning-based digital pathology tool automated measurements in a mouse model of podocytopathy. This proposed method provides high-throughput, objective morphological analysis and could facilitate podocytopathy research.

Keywords: glomerular basement membrane width; podocyte foot process width; podocytopathy; digital pathology; artificial intelligence; machine learning; image segmentation

Introduction

Chronic kidney disease impacts over 13% of the global population [1]. Proteinuric kidney diseases comprise a large portion of this massive health burden and often result in kidney failure [2]. Podocytopathy is a proteinuric kidney disease caused by injury to the podocytes in the glomeruli [3]. The podocyte is a critical part of the kidney's glomerular filtration barrier that serves as a selectively permeable membrane to filter blood plasma while retaining blood cells and large serum proteins such as albumin [1]. The

glomerular filtration barrier also consists of the glomerular basement membrane (GBM) and fenestrated endothelium, which together perform kidney glomerular ultrafiltration function and prevent leaking of essential proteins into the urine (i.e. proteinuria) [2–4]. Podocytes are postmitotic cells with interdigitating, finger-like protrusions called foot processes that adhere to and compress the underlying GBM [4]. Between podocyte foot processes (PFPs) are special cell-cell junctions with narrow gaps of about 40 nm called slit diaphragms (SD) [4].

Received: 4 November 2024; **Revised:** 16 March 2025; **Editorial decision:** 17 March 2025; **Accepted:** 26 March 2025

© The Author(s) 2025. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

In both research and clinical pathology settings, transmission electron microscopy (TEM) images of kidney tissues are commonly used to observe the glomerular filtration barrier ultrastructure and detect morphological changes. TEM images provide a grayscale, cross-sectional visualization of the GBM, PFPs, and SDs, which are difficult to resolve under light microscopy [5]. The glomerular filtration barrier morphology can be quantified using TEM images by measuring the widths of the GBM and PFPs. GBM width is measured as the orthogonal distance between the basal cell membrane of the endothelium and the basal surface of the PFPs [6]. On the other hand, PFP width is measured as the distance between subsequent SDs parallel to the GBM [7].

These measurements taken from TEM images of patient kidney biopsies facilitate renal pathologists' diagnosis and classification of glomerular diseases. An increase in PFP width, also called PFP effacement, has been associated with podocytopathies such as focal segmental glomerulosclerosis, minimal change disease, and membranous nephropathy [8, 9]. Recently, restored PFP width was identified as an independent indicator of treatment outcome and remission in patients with proteinuric kidney disease such as lupus nephritis [10]. Increased GBM width is often observed in podocytopathies such as membranous nephropathy [11] and other proteinuric kidney diseases such as diabetic kidney disease [12, 13] and lupus nephritis [5, 14], while decreased GBM width may suggest thin basement membrane nephropathy with type IV collagen mutations [5, 14]. These metrics are also used experimentally as image biomarkers for podocytopathy in pre-clinical animal models or clinical trials as they enable quantification of glomerular filtration barrier ultrastructural changes due to podocyte injury, genetic mutation, or therapeutic intervention [7, 15–17].

Despite the utility of GBM and PFP widths, efforts to fully automate their measurements in TEM images have been limited. Researchers and clinical renal pathologists continue to rely on manual measurement of TEM images, which is labor-intensive and limits the efficiency and accuracy of glomerular disease research and clinical diagnosis [7, 12, 14, 16, 18–22]. In addition, manual measurement introduces significant inter-operator variation and bias [6, 19, 22, 23], which affects reproducibility and can be addressed by automated measurement of kidney glomerular ultrastructure. Computation-aided tools have measured either GBM or PFP width, though not both [20, 24]. Lastly, recent methods for identification of glomerular filtration barrier ultrastructure were developed on human biopsies and may have limited generalizability and utility for research with animal models [20, 24, 25].

Any potential method for automated GBM and PFP measurements in TEM images requires a key step for the identification of these structures. Deep learning refers to machine learning with multiple layers of neural networks, designed to leverage large computing power to extract information and optimize predictions. The recent advent of deep learning revolutionized computer vision and enabled semantic segmentation [25, 26], which refers to an algorithm that places every pixel of an image into discrete categories, such as GBM or non-GBM. We opted for a U-Net deep learning framework for efficiency and accuracy on a small dataset without reliance on pretrained models. U-Net was developed for segmentation in biological applications, using skip-connections to minimize data loss and enable higher resolution segmentation, along with data augmentation to improve performance on small image sets [26, 27]. Segmentation accuracy of biological images is commonly evaluated in terms of Dice

coefficient or Jaccard index [28], both calculated from counts of true or false and positive or negative pixel predictions.

In this study, we leveraged a U-Net based deep learning framework to segment the GBM and measure GBM and PFP widths in kidney TEM images from Integrin-Linked Kinase (ILK) podocyte-specific conditional knockout (cKO) mice, an animal model of podocytopathy and severe proteinuria [29, 30], and their wild-type (WT) littermate mice.

Methods

Animal models

ILK podocyte-specific cKO mice and their WT littermates were generated using the ILK^{flox} and NPHS2-Cre mice as previously described [29]. The tail DNAs were isolated from 3-week-old mice produced from the breeding of $ILK^{flox/+}$; NPHS2-Cre⁺ heterozygous mice. The $ILK^{flox/flox}$; NPHS2-Cre⁺ homozygous mice (ILK cKO) and their WT littermates ($ILK^{+/+}$; NPHS2-Cre⁺) were subsequently identified by polymerase chain reaction genotyping methods, as previously described [29]. At 4 weeks of age, WT and ILK cKO mice were euthanized, and their kidneys were dissected and fixed in Karnovsky's fixative. TEM images were obtained using a Joel JEM-1011 TEM instrument (JEOL USA Inc., USA) as described previously [7, 31].

Manual measurement of the GBM and PFP width on TEM images

A total of 6–12 TEM images at 12 000× magnification (3484-by-2672 pixels) were captured for each of the WT ($n=4$) and ILK cKO ($n=4$) mice, for a total of $n=40$ TEM images from mice of each genotype. An expert operator manually measured GBM width on $n=30$ images from WT and $n=23$ from ILK cKO mice and PFP width on $n=23$ images from WT and $n=28$ images from ILK cKO mice using ImageJ software (version 1.51g; National Institutes of Health, USA). GBM width was measured between the basal surfaces of the endothelium and PFPs, orthogonal to the GBM contour [5], while PFP width was measured between subsequent SDs parallel to the GBM contour (Fig. 1C) [7]. The results were then pasted on an Excel sheet, and the harmonic mean for each image was calculated, which has been proven in previous publications to better estimate GBM width [32].

U-Net model and GBM segmentation

The first step of our automated method was GBM semantic segmentation with a U-Net deep learning model [27]. Ground truth GBM labels were produced from manually annotated TEM images with open source Qu-Path software [33]. The GBM was identified as described above, with the work divided among three operators [6, 21]. The resulting binary GBM labels classified each pixel as either GBM or background. Each TEM image and its corresponding GBM label were shrunk row- and column-wise by a factor of two before a 100-pixel sliding window was applied to produce 8586 overlapping 512-by-512-pixel label/image tile pairs from each whole image. By utilizing overlapping crops and horizontal flips with a 50% probability, we augmented the size of our relatively small dataset. We trained and tested a U-Net model for GBM segmentation on this preprocessed dataset of TEM label/image tile pairs [27]. This U-Net model is designed for image segmentation tasks, featuring four encoding blocks and four decoding blocks. It includes batch normalization after each convolution, ReLU activations, and the “same” convolution mode to maintain spatial dimensions. During training, 25% of label/image tile pairs in the training set were randomly selected as the validation set to

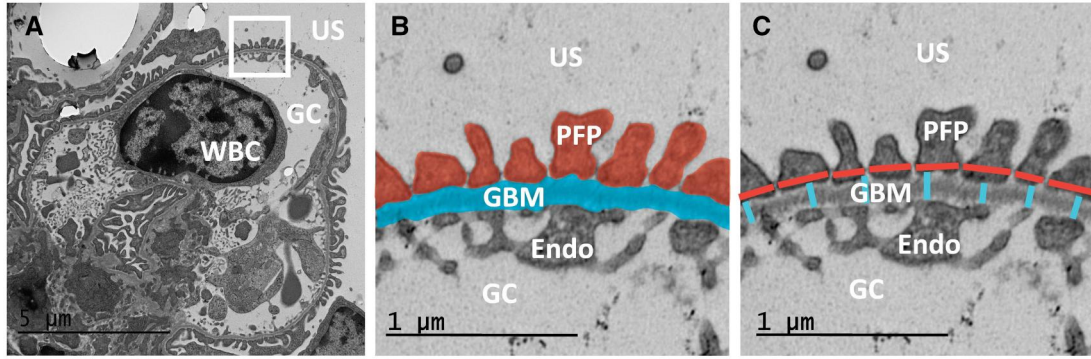


Figure 1 Illustration of kidney TEM image ultrastructure and measurement. (A) Kidney TEM image of a WT mouse with a region of the glomerular filtration barrier in a white box. GC, glomerular capillary; US, urinary space; WBC, white blood cell. (B) The enlarged white box area in A shows the same glomerular filtration barrier region with highlighted PFP (red) and GBM (blue). Endo, endothelial cell. (C) Demonstration of manual measurement of PFP width (horizontal red lines) and GBM width (vertical blue lines)

assess prediction accuracy after each training epoch and guide the update of model parameters. With each training epoch, our model seeks to minimize a dice loss function applied to the validation set predictions.

$$\text{Dice loss}(\text{prediction}, \text{label}) = 1 - \frac{2|\text{prediction} \cap \text{label}|}{|\text{prediction}| + |\text{label}|} \quad (1)$$

$$= 1 - \frac{2 \times \text{true positive}}{2 \times \text{true positive} + \text{false positive} + \text{false negative}}$$

We selected forty-five epochs as training length by plotting validation and training learning curves and observing the approximate epoch at which the losses were no longer reduced by additional training. A constant learning rate of 0.01 was selected with an open source learning rate finder library by plotting validation loss as a function of the learning rate and choosing the approximate rate at which it was minimized. Additionally, we used a batch size of two and optimized the model using Stochastic Gradient Descent. U-Net models were trained with these optimized hyper-parameters and saved after the final epoch. We then applied the U-Net to the test set TEM image tiles to perform GBM segmentation, which was used to inform image processing algorithms for measuring GBM and PFP width.

GBM width measurement algorithm

The second stage was an image processing algorithm with the input of the predicted GBM segmentations. To measure mean GBM width, we first refined the segmentation results. We smoothed the GBM segmentation tiles with a narrow Gaussian kernel (standard deviation one pixel), applied Otsu-thresholding to generate a binary image, and removed small artifacts with areas less than 1000 pixels, as these generally corresponded to false positive GBM. We then applied a skeletonizing algorithm to extract the center line of each GBM segment, as described in a previous publication [6]. Finally, the mean GBM width for each tile was calculated as its total GBM area divided by its total GBM central length:

$$\text{GBM width} = \frac{\text{GBM segmented area}}{\text{GBM central length}} \quad (2)$$

Our GBM width measurement algorithm borrows the approximation of GBM width as the quotient of the segmented area and central length [34], though we did not map the central length to a straight line, so we obtained a larger approximation. The mean GBM width for a given specimen was then calculated as the

mean of its tiles' GBM widths weighted by GBM area, then converted to length by a factor of 8.33 nm/pixel based on the scale bar. We neglected to apply a correction factor for uncertain non-orthogonality of the sectioning plane to the GBM, given that the primary aim of this article was to assess manual versus automated measurement, performed on the same images, and WT versus ILK cKO morphology, as opposed to reporting absolute values for mouse GBM width [21]. This same principle applied to the PFP and SD. As such, the reported GBM and PFP widths were likely wider than reality.

PFP width measurement algorithm

The predicted GBM segmentation tiles also guided a novel image processing algorithm for measuring PFP width. The segmentation tiles were first smoothed and widened by applying a Gaussian kernel (standard deviation 9 pixels) and then Otsu-thresholding. GBM tiles that had multiple distinct GBM segments with two edges were split into tiles with just one segment each. Tiles with discontinuous segments, with segments intersecting the tile borders at points other than the segment terminus, or without any segments, were discarded. This produced a refined dataset of GBM segmentation tiles cleanly traversed by the GBM segment. When these segmentation tiles with a widened, smoothed GBM were superimposed on the original TEM image tiles, the urinary side of GBM segments transected PFPs and SDs. The TEM image pixel values were extracted from along both edges of this widened GBM segmentation. A Gaussian Mixture Model classified the pixels along both edges into two classes by grayscale values corresponding to PFPs and SDs [35]. The Gaussian Mixture Model assumed that each class's values were normally distributed, which we confirmed was a rough approximation from histograms of several representative edges. The Gaussian Mixture Model predicted each class's mean grayscale values and variances and applied these hyperparameters to classify the edge pixels into lighter versus darker classes. The urinary side was predicted as the edge with a higher proportion of pixels belonging to the darker class due to the dark staining of PFPs relative to the background. Distinguishing between the urinary and capillary edges of the GBM was a key step that required manual intervention in the most recent semi-automated PFP measurement effort [23].

The pixel values from the urinary edge were plotted to identify the SDs. First, a smoothing Savitzky-Golay filter was applied to the one-dimensional data to reduce the impact of signal noise [36]. The SDs were then identified from the smoothed edge values using a local maxima selection algorithm. The minimum

distance between SDs was set as 15 pixels, which corresponded to 125 nm and represented a conservative minimum PFP value from our expert manual measurement data. The normalized local maxima prominence threshold was set as the quotient of the standard deviation of the predicted Gaussian Mixture Model lighter/SD class and the edge mean pixel value. The mean PFP width for a given GBM segment tile was calculated as the quotient of the total urinary edge length and the number of identified SDs. To calculate the mean PFP width for a given specimen, the mean PFP width of its tiles was weighted according to the tile's total urinary edge length and then converted to length by a factor of 8.33 nm/pixel.

Validation study

We assessed our proposed automated measurement method using a 4-fold cross-validation study in which we trained four U-Net GBM segmentation models on four nonoverlapping partitions of our preprocessed dataset of TEM image tiles and corresponding GBM labels. We first sorted our dataset specimens into four random pairs of one WT and one ILK cKO mouse. Four-fold cross-validation was conducted by (i) selecting one mouse pair as the test set and the remaining three pairs as the training set, (ii) preprocessing the training and test set data, (iii) training a U-Net GBM segmentation model, (iv) predicting GBM segmentation of the test set, (v) using the predictions to measure GBM and PFP widths, (vi) iterating the process three more times until each mouse pair served as the test set. In this manner, we approximated 75/25 training/test partitions and avoided overlap between training and test sets while maximizing the test data from our relatively small dataset. As a result, we obtained GBM segmentation for each TEM image and automated GBM and PFP width measurements for each mouse.

GBM segmentation accuracy for each TEM image was reported as the Jaccard index, or intersection over union, often used to evaluate segmentation in biomedical images [5, 28]. We also used the Dice coefficient to evaluate the accuracy. Both were calculated based on true positive, false positive, and false negative by comparing the predicted GBM segmentation tiles to their corresponding GBM label tiles.

$$\begin{aligned} \text{Jaccard}(\text{prediction}, \text{label}) &= \frac{|\text{prediction} \cap \text{label}|}{|\text{prediction} \cup \text{label}|} \\ &= \frac{\text{true positive}}{\text{true positive} + \text{false positive} + \text{false negative}} \end{aligned} \quad (3)$$

$$\begin{aligned} \text{Dice coefficient}(\text{prediction}, \text{label}) &= \frac{2|\text{prediction} \cap \text{label}|}{|\text{prediction}| + |\text{label}|} \\ &= \frac{2 * \text{true positive}}{2 * \text{true positive} + \text{false positive} + \text{false negative}} \\ &= \frac{2 * \text{Jaccard}(\text{prediction}, \text{label})}{1 + \text{Jaccard}(\text{prediction}, \text{label})} \end{aligned} \quad (4)$$

The Jaccard indices for each TEM image were averaged from its corresponding tiles. The resultant automated mean GBM or PFP width measurements for mouse specimens were compared with corresponding expert manual measurements for both WT and ILK cKO groups. Lastly, the manual and automated mean GBM and PFP width measurements for each mouse were compared between WT and ILK cKO groups.

Statistical analysis

The mean Jaccard index values for both WT and ILK cKO TEM images were roughly normally distributed with a relatively large

sample size ($n=40$). Hence, we selected the parametric two-sample t-test to compare GBM segmentation accuracy between the two genotypes with an $\alpha = 0.05$ level of significance and reported the mean (95% CI) image Jaccard index and Dice coefficient for both. The remaining statistical comparisons occurred between small ($n=4$) groups with unknown distributions, so we selected the nonparametric Wilcoxon rank sum test with an $\alpha = 0.05$ level of significance. Data analyses were performed using RStudio (RStudio Inc., Boston, MA).

Results

The normal kidney glomerular filtration barrier ultrastructure image and PFP and GBM measurement under TEM are illustrated in Fig. 1.

We generated ground truth GBM labels from the manual annotation of the corresponding mouse kidney TEM images (Fig. 2A). These TEM image and GBM label pairs were tiled with a sliding window (Fig. 2B) and used to train and test our U-Net segmentation model (Fig. 2C). Segmented tiles subsequently informed image processing algorithms that measured GBM and PFP width, as in Fig. 2D and E, respectively. The image processing algorithm is demonstrated for a representative TEM image tile (Fig. 3A). After segmentation with the trained U-Net model, we skeletonized the GBM (Fig. 3B); GBM width was calculated as the GBM skeleton length divided by the GBM area. The urinary edge of the glomerular filtration barrier (Fig. 3C) and SD along it (Fig. 3D) were predicted to measure PFP width.

We evaluated our automated measurement of GBM and PFP width on our dataset of TEM images from four WT and four ILK cKO littermate mice with a 4-fold cross-validation study (Fig. 4). GBM segmentation accuracy relative to the annotated labels for each TEM image is presented for WT and ILK cKO groups in terms of the Jaccard Index in Fig. 5A. The resulting mean (95% CI) Jaccard indices for TEM images from WT and ILK cKO specimens were 0.73 (0.70–0.76) for WT and 0.85 (0.83–0.87) for ILK cKO specimen TEM images. This corresponds with Dice coefficients of 0.81 (0.79–0.84) and 0.88 (0.87–0.89), respectively. GBM segmentation accuracy was significantly greater for ILK cKO than WT images (95% CI, 0.07–0.14; $P < .01$, Fig. 5A). TEM image tiles from WT and ILK cKO mouse kidneys are visualized in Fig. 5B and D, respectively, alongside their corresponding predicted GBM segmentation in Fig. 5C and E. On visual inspection, TEM image tiles with lower contrast and more irregular GBM morphology were generally less accurately segmented (Fig. 5D and E). Representative examples of GBM and SD identification are provided in Supplementary Fig. 1. SD were often over-identified, and fidelity of SD identification was qualitatively better in regions of higher SD density.

The automated measurements from the 4-fold cross-validation were compared with corresponding manual measurements of GBM (Fig. 6A) and PFP width (Fig. 6B) for both WT and ILK cKO groups. Manual and automated GBM width measurements were similar for WT ($P = .49$) and ILK cKO ($P = .06$) mice (Fig. 6A). However, PFP width manual and automated measurements differed significantly for ILK cKO ($P = .03$) but not WT ($P = .89$) mice (Fig. 6B).

We also compared the GBM and PFP width between WT and ILK cKO groups with both automated (Fig. 7A) and manual measurements (Fig. 7B). Manual measurements differed significantly between genotype groups for both GBM ($P = .03$) and PFP width ($P = .03$) (Fig. 7B). Similarly, automated GBM width measurements for WT mice differed significantly from ILK cKO mice ($P = .03$)

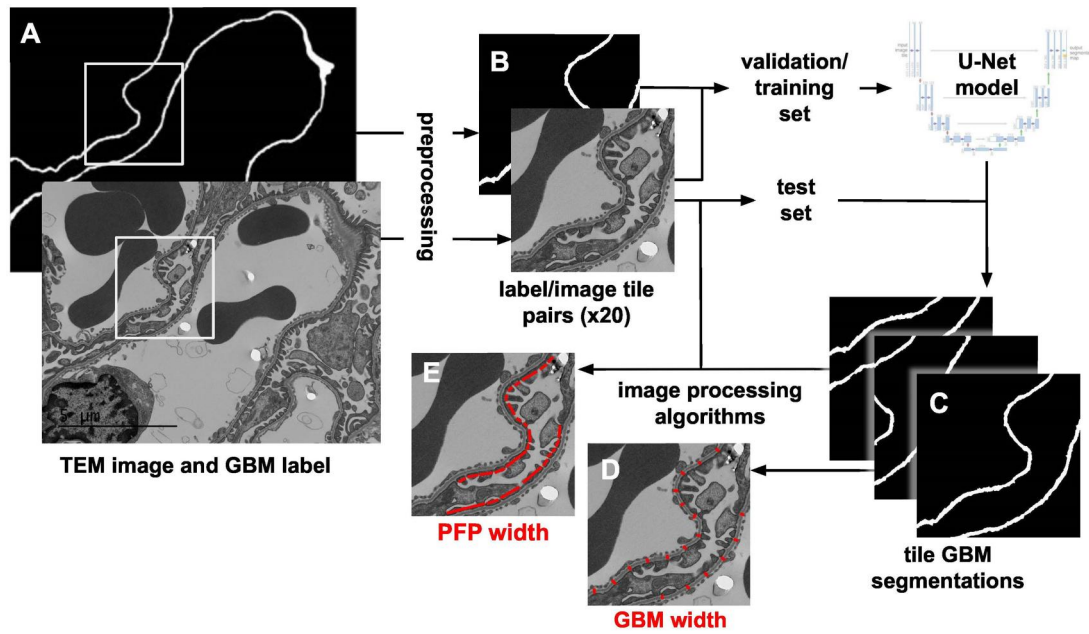


Figure 2 Block diagram of automated measurement method for GBM and PFP width. (A) TEM image and corresponding annotated GBM label. (B) Preprocessing results in 20 label/image tile pairs per label/image, which are sorted into validation/training and test sets, and the former are used to train a U-Net model for GBM segmentation. (C) The U-Net model is applied to test set TEM image tiles to generate GBM segmentation tiles. Resulting GBM segmentation tiles inform image processing algorithms for (D) GBM width, and along with image tiles, and for (E) PFP width

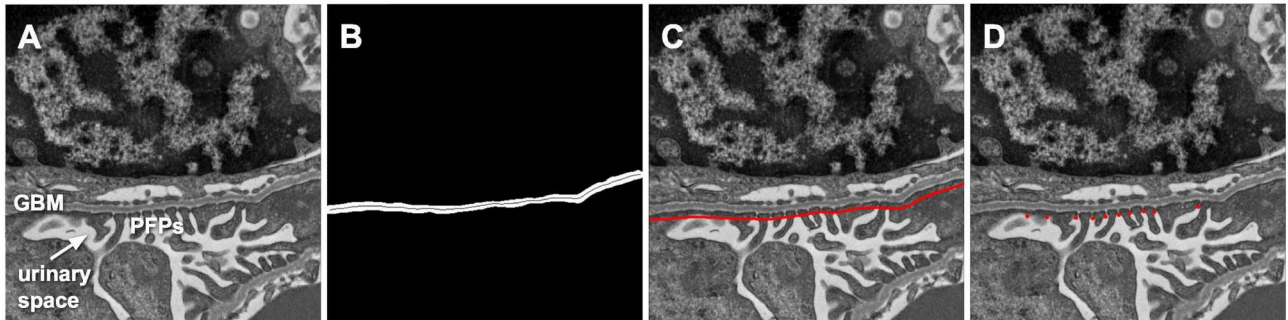


Figure 3 GBM and PFP width measurement image processing algorithms. (A) Original TEM image tile with labeled GBM, PFPs, and urinary space. (B) Refined GBM segmentation results with one segment separated and fitted with a central line for GBM width calculation. (C) Urinary edge of smoothed GBM segmentation result. (D) Example identified SDs for PFP width calculation

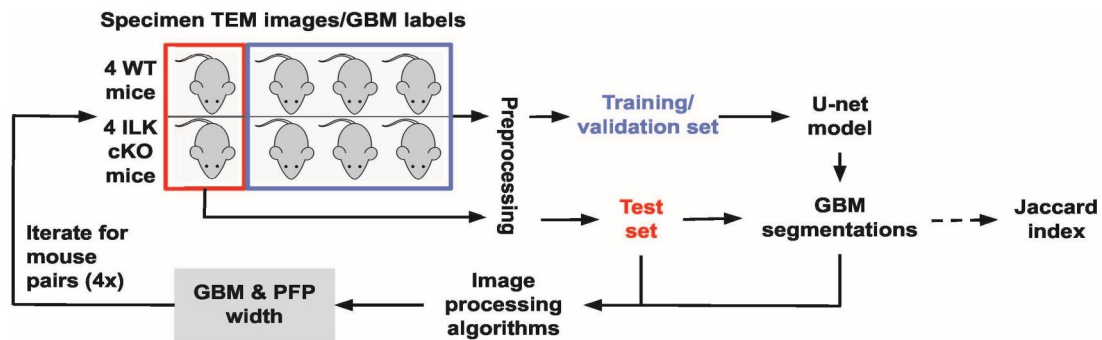


Figure 4 Four-fold cross-validation study of the proposed automated measurement method

(Fig. 7A), although automated PFP width measurements differed less substantially ($P = .06$) (Fig. 7A), which is consistent with the previously reported pathology finding in 4-week young ILK cKO mice [30].

We applied another recent, publicly available, U-Net-based, automated method for PFP width measurement developed on human kidney biopsies to our TEM images [24]. We compared the resulting proportion of image tiles with predicted PFP width and

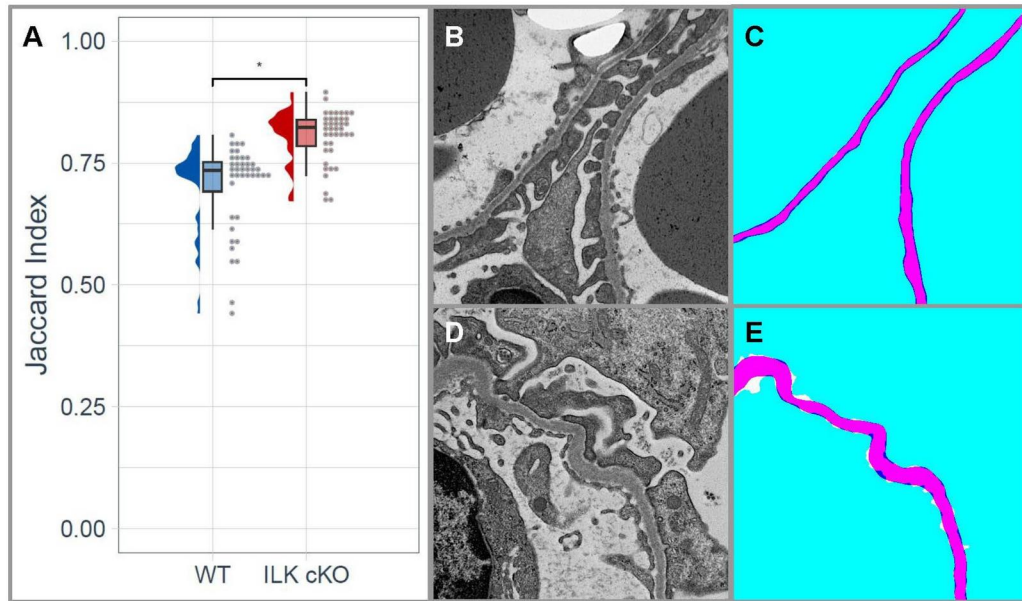


Figure 5 Accuracy of TEM image GBM segmentation with Jaccard index. (A) TEM image GBM segmentation accuracy for WT and ILK cKO groups in terms of Jaccard index, where significance is denoted by $P < .05$ (*) in a two-sample t-test. (B) Representative TEM image tile from a WT specimen and (C) corresponding GBM segmentation results with true positive (pink), true negative (light blue), false negative (white), and false positive (dark blue). (D) Representative TEM image tile from an ILK cKO specimen and (E) corresponding GBM segmentation results with true positive (pink), true negative (light blue), false negative (white), and false positive (dark blue)

specimen PFP width measurements to our method in [Supplementary Table 1](#) and [Supplementary Fig. 2](#), respectively.

Note that the values of both manual and automated GBM and PFP widths were likely overestimated given the lack of correction for segmentation angle.

Discussion

In this work, we developed a method of automating GBM and PFP width measurement in mouse kidney glomerular TEM images using a U-Net framework for GBM identification, followed by image processing measurement algorithms. We evaluated the segmentation and measurement performance through a 4-fold cross-validation study. Our automated GBM and PFP width measurements were similar to corresponding manual measurements for WT but not ILK cKO mouse kidney specimens. However, our automated method distinguished early pathological changes in the GBM width and PFP width in young 4-week-old ILK cKO mouse kidneys. This result was consistent with findings in a previously published report of GBM derangements in ILK cKO mice, including increased GBM thickness, without a significant change in PFP width at 4 weeks [30]. Our approach thus represents an advancement in quantifying glomerular TEM images in mouse models of podocytopathy, enhancing the speed and accuracy of analysis compared to current manual and semiautomated methods.

The variable morphology and poor image contrast of the GBM in TEM images make its segmentation difficult, although prior results support its selection as a superior target for deep learning segmentation relative to other components of the glomerular filtration barrier [20, 25]. In addition, with Jaccard indices of 0.73 (0.70–0.76) for WT and 0.85 (0.83–0.87) for ILK cKO mice, our segmentation accuracy exceeded the 0.64 of the random forest stacks-based machine learning by Cao et al. [5]. Although GBM morphology appeared more variable and indistinguishable in ILK cKO kidney specimens, segmentation accuracy was greater for this group. This finding may be explained by the mathematical

tendency of the Jaccard index to favor images with higher counts of true positive pixels, given the wider mean GBM width of ILK cKO kidneys.

Our approach offers several novel aspects. Its automation greatly improved efficiency relative to semiautomated methods that have been piloted for GBM [6, 12, 34] and PFP measurement [18, 23] and require operator intervention for each image. Three-dimensional structured illumination microscopy and confocal and stimulated emission depletion super-resolution microscopy were recently used for GBM and PFP width, SD length, and PFP circularity measurement [19, 37–40]. However, unlike TEM utilized by our method, these forms of microscopy are not commonly used in clinical pathology diagnosis. To our knowledge, we are the first to fully automate both PFP width and GBM width measurement simultaneously in kidney TEM images.

Recently, automated methods were proposed using variants of U-Net to measure GBM or PFP width in TEM images of human kidneys by Wang et al. and Smerkous et al., respectively [20, 24]. Wang et al. segmented both the GBM and neighboring electron dense deposits, and their RADS-Net model provided a marginal 2%–4% improvement in GBM segmentation accuracy over more traditional U-Net models [20]. The advantages of our GBM measurement approach include robustness to local variations and segmentation noise, given integration over the entire GBM. Our segmentation requires less data and training time relative to such vision transformer-based models while achieving sufficient accuracy to support subsequent GBM width calculations. Smerkous et al. segmented both the podocyte-GBM interface and SDs to measure PFP width [24]. Each method required segmentation of additional morphological features on substantially larger human image datasets and thus greater manual annotation labor [20, 24]. Additionally, at present our automated GBM and PFP width measurements appear unique in their performance on mouse specimens, including a model of podocytopathy, which remains vital to ongoing therapeutic research. As an example, when we applied the method of Smerkous et al. to our data,

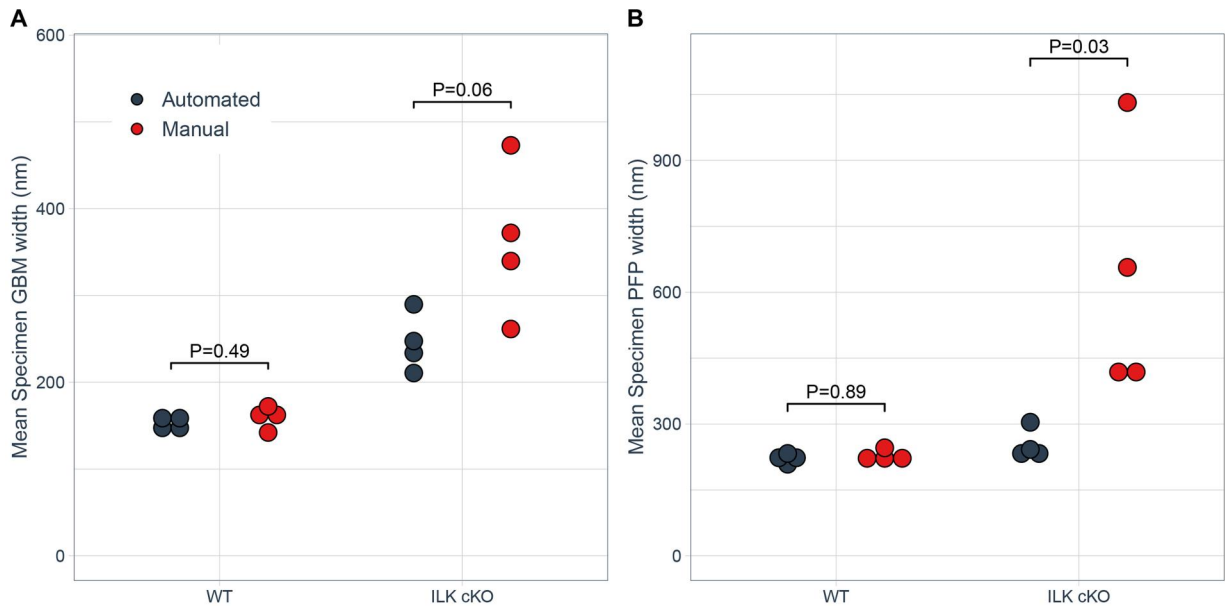


Figure 6 Comparison of automated versus corresponding manual measurements of GBM and PFP width in the same WT and ILK cKO mouse kidneys. **(A)** Comparison of GBM width measured by automated versus manual method for WT and ILK cKO groups. **(B)** Comparison of PFP width measured by automated versus manual method for WT and ILK cKO. The exact P-values are shown, which are calculated from the Wilcoxon rank sum test with an $\alpha = 0.05$ level of significance.

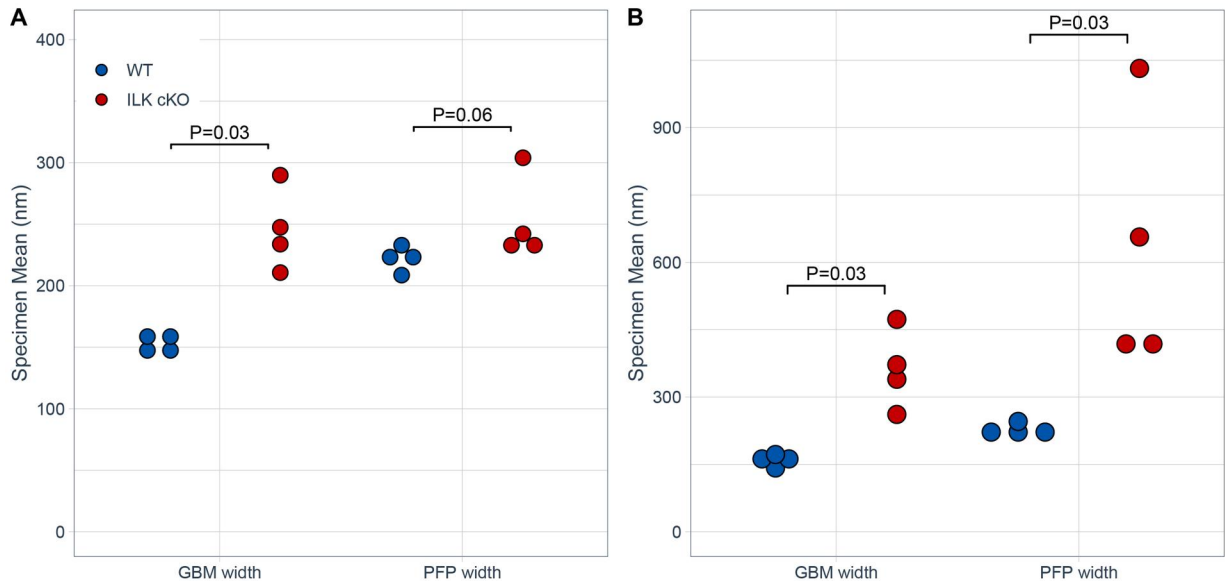


Figure 7 Distinction of subtle differences in GBM width and PFP width in 4-week-old WT versus ILK cKO mouse kidneys measured by automated and manual methods. **(A)** Automated measurements of GBM width and PFP width in 4-week-old WT versus ILK cKO littermate kidneys. **(B)** Manual measurements of GBM width and PFP width in 4-week-old WT versus ILK cKO littermate kidneys. The exact P-values are shown, which are calculated from the Wilcoxon rank sum test with an $\alpha = 0.05$ level of significance.

which was excellently developed for human biopsies, it struggled to obtain PFP width measurements and those obtained differed more substantially from manual measurements [24] (Supplementary Table 1 and Supplementary Fig. 2). However, we also fully acknowledge that U-Nets do not generalize well, and we do not intend to compare our general approaches.

Though our method performed well on a relatively small image dataset, the burden of GBM manual annotation for training could potentially benefit from self-supervised GBM segmentation. Lin *et al.* developed a self-supervised model for glomerular filtration barrier segmentation with a similar U-shaped

convolutional network [25]. However, this model nonetheless required fine-tuning on 311 manually annotated images, highlighting the importance of a well-annotated dataset.

SD identification remains a key area of potential improvement, as visual inspection of algorithm-identified SDs suggested frequent misidentification, especially overidentification, that may have driven the discrepancy between corresponding automated and manual PFP measurements among ILK cKO mice. This could be partially explained by our measurement algorithm's minimum width, which would have been less influential for the effaced PFPs in ILK cKO mice. Misidentification was more

evident in regions of the glomerular filtration barrier with wider or effaced PFPs, and, conversely, less so in areas of higher SD density. The exact explanation for discrepancies between automated and manual measurements cannot be precisely identified as, unlike manual measurements, the previous manual measurement process cannot be tracked. Unfortunately, this issue remains a meaningful yet not fully elucidated source of inaccuracy for our PFP measurement performance.

We recognize a couple other key limitations of our approach and its potential application. Our method was developed on high-quality glomerular TEM images and may not perform well in lower-contrast images, as often result from clinical biopsies. Our ability to draw firm conclusions about the accuracy of our automated measurements and their effectiveness in distinguishing pathological tissue was also limited due to only four specimens per genotype and a single disease model.

Future studies of our general approach may incorporate larger image datasets with more diverse pathologies, other animal models such as rats, and human kidney biopsies, and explore semi-supervised or reinforcement learning approaches to reduce the burden of labeling larger image datasets. Our long-term goal is to develop an open source, efficient, accurate tool that researchers and pathologists can use to input TEM images and obtain automated GBM and PFP width measurements, thereby facilitating clinical diagnosis and research.

Acknowledgments

We thank Dr. Diane Joseph-McCarthy and Dr. Darren Roblyer for their support on the Boston University Biomedical Engineering Senior Design Course.

Author contributions

Aksel Laudon (Conceptualization [equal], Data curation [equal], Formal analysis [equal], Investigation [equal], Methodology [equal], Writing—original draft [equal], Writing—review & editing [equal]), Zhaoze Wang (Data curation [equal], Formal analysis [equal], Investigation [equal], Methodology [equal], Software [equal], Writing—review & editing [equal]), Anqi Zou (Formal analysis [equal], Investigation [equal], Methodology [equal], Software [equal], Writing—review & editing [equal]), Richa Sharma (Data curation [equal], Formal analysis [equal], Writing—review & editing [equal]), Jiayi Ji (Formal analysis [equal], Investigation [equal], Methodology [equal], Writing—review & editing [equal]), Winston Tan (Formal analysis [supporting]), Connor Kim (Formal analysis [equal], Methodology [equal], Writing—review & editing [equal]), Yingzhe Qian (Formal analysis [equal], Methodology [equal], Writing—review & editing [equal]), Qin Ye (Formal analysis [equal], Methodology [equal], Writing—review & editing [equal]), Hui Chen (Data curation [equal], Methodology [equal], Writing—review & editing [equal]), Joel Henderson (Data curation [equal], Supervision [equal], Writing—review & editing [equal]), Chao Zhang (Formal analysis [equal], Methodology [equal], Supervision [equal], Writing—review & editing [equal]), and Vijaya Kolachalama (Funding acquisition [equal], Methodology [equal], Supervision [equal], Writing—original draft [equal], Writing—review & editing [equal]), and Weining Lu (Conceptualization [equal], Data curation [equal], Formal analysis [equal], Funding acquisition [equal], Investigation [equal], Methodology [equal], Project administration [equal], Resources [equal], Supervision [equal], Writing—original draft [equal], Writing—review & editing [equal])

Supplementary data

Supplementary data is available at *Biology Methods and Protocols* online.

Conflicts of interest statement. The authors have declared that no conflict of interest exists for this work.

Funding

This work was supported in part by the National Institutes of Health grant R01-DK133940 (WL), R43-DK134273, R01-HL159620, R21-CA253498, and RF1-AG062109 (VBK), the DOD grant E01HT9425-23-1-1058 (WL), grants from Boston University Biomedical Innovation Technologies Affinity Research Collaboratives (BIT-ARC) and the B4D-ARC (Novel Biomarkers for Diagnosis and Drug Development—A Data Science Approach), a Boston University Clinical & Translational Science Institute Integrated Pilot Grant (CZ, WL), and funding support from Boston University Undergraduate Research Opportunities Program (AL). This publication is also supported in part by the National Center for Advancing Translational Sciences, National Institutes of Health, through Boston University Clinical & Translational Science Institute Grant Number 1UL1TR001430. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

Data availability statement

The segmentation model and code are publicly available in GitHub at https://github.com/CZCBLab/GBM_UNET. The data set used and/or analyzed during the current study is available upon request.

References

1. Assady S, Benzing T, Kretzler M et al. Glomerular podocytes in kidney health and disease. *Lancet* 2019;**393**:856–58. [https://doi.org/10.1016/S0140-6736\(18\)33000-9](https://doi.org/10.1016/S0140-6736(18)33000-9)
2. Tian X, Ishibe S. Targeting the podocyte cytoskeleton: from pathogenesis to therapy in proteinuric kidney disease. *Nephrol Dial Transplant* 2016;**31**:1577–83. <https://doi.org/10.1093/ndt/gfw021>
3. Kopp JB, Anders H-J, Susztak K et al. Podocytopathies. *Nat Rev Dis Primers* 2020;**6**:68. <https://doi.org/10.1038/s41572-020-0196-7>
4. Benzing T, Salant D. Insights into glomerular filtration and albuminuria. *N Engl J Med* 2021;**384**:1437–46. <https://doi.org/10.1056/NEJMr1808786>
5. Cao L, Lu Y, Li C et al. Automatic segmentation of pathological glomerular basement membrane in transmission electron microscopy images with random forest stacks. *Comput Math Methods Med* 2019;**2019**:1684218. <https://doi.org/10.1155/2019/1684218>
6. Rangayyan RM, Kamenetsky I, Benediktsson H. Segmentation and analysis of the glomerular basement membrane in renal biopsy samples using active contours: a pilot study. *J Digit Imaging* 2010;**23**:323–31. <https://doi.org/10.1007/s10278-009-9188-6>
7. Pisarek-Horowitz A, Fan X, Kumar S et al. Loss of roundabout guidance receptor 2 (Robo2) in podocytes protects adult mice from glomerular injury by maintaining podocyte foot process structure. *Am J Pathol* 2020;**190**:799–816. <https://doi.org/10.1016/j.ajpath.2019.12.009>
8. Deegens JKJ, Dijkman HBPM, Borm GF et al. Podocyte foot process effacement as a diagnostic tool in focal segmental

- glomerulosclerosis. *Kidney Int* 2008;**74**:1568–76. <https://doi.org/10.1038/ki.2008.413>
9. da Silva CA, Monteiro MLGDR, Araújo LS et al. In situ evaluation of podocytes in patients with focal segmental glomerulosclerosis and minimal change disease. *PLoS One* 2020;**15**:e0241745. <https://doi.org/10.1371/journal.pone.0241745>
 10. Ichinose K, Kitamura M, Sato S et al. Podocyte foot process width is a prediction marker for complete renal response at 6 and 12 months after induction therapy in lupus nephritis. *Clin Immunol* 2018;**197**:161–68. <https://doi.org/10.1016/j.clim.2018.10.002>
 11. Duan S, Sun L, Zhang C et al. The thickness of glomerular basement membrane predicts complete remission in primary membranous nephropathy. *Ren Fail* 2023;**45**:2179335. <https://doi.org/10.1080/0886022X.2023.2179335>
 12. Kamenetsky I, Rangayyan RM, Benediktsson H. Analysis of the glomerular basement membrane in images of renal biopsies using the split-and-merge method: a pilot study. *J Digit Imaging* 2010;**23**:463–74. <https://doi.org/10.1007/s10278-009-9233-5>
 13. Marshall CB. Rethinking glomerular basement membrane thickening in diabetic nephropathy: adaptive or pathogenic? *Am J Physiol Renal Physiol* 2016;**311**:F831–43. <https://doi.org/10.1152/ajprenal.00313.2016>
 14. Kyriacou K, Nearchou M, Zouvani I et al. The many faces of thin basement membrane nephropathy; a population based study. In: *Topics in Renal Biopsy and Pathology*. London: IntechOpen, 2012. <https://www.intechopen.com/books/788>
 15. Beck LH, Berasi SP, Copley JB, Jr et al. PODO: trial Design: phase 2 Study of PF-06730512 in Focal Segmental Glomerulosclerosis. *Kidney Int Rep* 2021;**6**:1629–33. <https://doi.org/10.1016/j.ekir.2021.03.892>
 16. Veron D, Reidy KJ, Bertuccio C et al. Overexpression of VEGF-A in podocytes of adult mice causes glomerular disease. *Kidney Int* 2010;**77**:989–99. <https://doi.org/10.1038/ki.2010.64>
 17. New LA, Martin CE, Scott RP et al. Nephritin tyrosine phosphorylation is required to stabilize and restore podocyte foot process architecture. *J Am Soc Nephrol* 2016;**27**:2422–35. <https://doi.org/10.1681/ASN.2015091048>
 18. Liu J. *Implementation of a Semi-Automatic Tool for Analysis of TEM Images of Kidney Samples*. Uppsala: Uppsala University, 2012.
 19. Unnersjö-Jess D et al. Deep learning-based segmentation and quantification of podocyte foot process morphology suggests differential patterns of foot process effacement across kidney pathologies. *Kidney Int* 2023;**103**:1120–30. <https://doi.org/10.1016/j.kint.2023.03.013>
 20. Wang Y, Liu Y, Fu Y et al. Segmentation and thickness calculation of glomerular basement membrane using RADS-Net in glomerular microscopic images. *Biomed Signal Process Control* 2024;**88**:105557. <https://doi.org/10.1016/j.bspc.2023.105557>
 21. Viana DL, Alladagbin DJ, Dos-Santos WLC et al. A comparative study of human glomerular basement membrane thickness using direct measurement and orthogonal intercept methods. *BMC Nephrol* 2022;**23**:23. <https://doi.org/10.1186/s12882-021-02634-1>
 22. Geraghty M, Thompson MP, Gorman D et al. Sources of variability in podocyte foot process width measurements and approaches to mitigation: PO2004. *J Am Soc Nephrol* 2020;**31**:617.
 23. Vargas SR et al. *Deep Learning for Quantitative Image Analysis of Morphological Changes in Podocyte Foot Processes in Electron Micrographs of Rat Model of Passive Heymann Nephritis*. New York: Pfizer, 2021.
 24. Smerkous D, Mauer M, Tøndel C et al. Development of an automated estimation of foot process width using deep learning in kidney biopsies from patients with Fabry, minimal change, and diabetic kidney diseases. *Kidney Int* 2024;**105**:165–76. <https://doi.org/10.1016/j.kint.2023.09.011>
 25. Lin G, Zhang Z, Long K et al. GCLR: a self-supervised representation learning pretext task for glomerular filtration barrier segmentation in TEM images. *Artif Intell Med* 2023;**146**:102720. <https://doi.org/10.1016/j.artmed.2023.102720>
 26. Yang J, Hu X, Pan H et al. Multi-scale attention network for segmentation of electron dense deposits in glomerular microscopic images. *Microsc Res Tech* 2022;**85**:3256–64. <https://doi.org/10.1002/jemt.24182>
 27. Ronneberger O, Fischer P, Brox T. Medical image computing and computer-assisted intervention. In: *MICCAI 2015: 18th International Conference*, Munich, Germany, October 5–9, 2015, proceedings, part III 18. 234–41. Boston, MA: Springer.
 28. Eelbode T, Bertels J, Berman M et al. Optimization for medical image segmentation: theory and practice when evaluating with Dice score or Jaccard index. *IEEE Trans Med Imaging* 2020;**39**:3679–90.
 29. Dai C, Stolz DB, Bastacky SI et al. Essential role of integrin-linked kinase in podocyte biology: bridging the integrin and slit diaphragm signaling. *J Am Soc Nephrol* 2006;**17**:2164–75. <https://doi.org/10.1681/ASN.2006010033>
 30. El-Aouni C, Herbach N, Blattner SM et al. Podocyte-specific deletion of integrin-linked kinase results in severe glomerular basement membrane alterations and progressive glomerulosclerosis. *J Am Soc Nephrol* 2006;**17**:1334–44. <https://doi.org/10.1681/ASN.2005090921>
 31. Graham L, Orenstein JM. Processing tissue and cells for transmission electron microscopy in diagnostic pathology and research. *Nat Protoc* 2007;**2**:2439–50. <https://doi.org/10.1038/nprot.2007.304>
 32. Ramage IJ, Howatson AG, McColl JH et al. Glomerular basement membrane thickness in children: a stereologic assessment. *Kidney Int* 2002;**62**:895–900. <https://doi.org/10.1046/j.1523-1755.2002.00527.x>
 33. Bankhead P, Loughrey MB, Fernández JA et al. QuPath: open source software for digital pathology image analysis. *Sci Rep* 2017;**7**:16878. <https://doi.org/10.1038/s41598-017-17204-5>
 34. Wu HS, Dikman S, Gil J. A semi-automatic algorithm for measurement of basement membrane thickness in kidneys in electron microscopy images. *Comput Methods Programs Biomed* 2010;**97**:223–31. <https://doi.org/10.1016/j.cmpb.2009.07.002>
 35. Reynolds DA. Gaussian mixture models. In: Li SZ, Jain AK (eds), *Encyclopedia of Biometrics*. Cham, Switzerland: Springer, 2009, 741.
 36. Steinier J, Termonia Y, Deltour J. Smoothing and differentiation of data by simplified least square procedure. *Anal Chem* 1972;**44**:1906–09. <https://doi.org/10.1021/ac60319a045>
 37. Siegerist F, Ribback S, Dombrowski F et al. Structured illumination microscopy and automatized image processing as a rapid diagnostic tool for podocyte effacement. *Sci Rep* 2017;**7**:11473. <https://doi.org/10.1038/s41598-017-11553-x>
 38. Butt L, Unnersjö-Jess D, Hohné M et al. A mathematical estimation of the physical forces driving podocyte detachment. *Kidney Int* 2021;**100**:1054–62. <https://doi.org/10.1016/j.kint.2021.06.040>
 39. Unnersjö-Jess D, Butt L, Hohné M et al. A fast and simple clearing and swelling protocol for 3D in-situ imaging of the kidney across scales. *Kidney Int* 2021;**99**:1010–20. <https://doi.org/10.1016/j.kint.2020.10.039>
 40. Unnersjö-Jess D, Ramdedovic A, Hohné M et al. Three-dimensional super-resolved imaging of paraffin-embedded kidney samples. *Kidney360* 2022;**3**:446–54. <https://doi.org/10.34067/KID.0005882021>

© The Author(s) 2025. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Biology Methods and Protocols, 2025, 10, 1–9

<https://doi.org/10.1093/biomethods/bpaf024>

Methods Article