PodoCount: A Robust, Fully Automated, Whole-Slide Podocyte Quantification Tool

Check for updates

Briana A. Santo[1], Darshana Govind[1], Parnaz Daneshpajouhnejad[2], Xiaoping Yang[2], Xiaoxin X. Wang[3], Komuraiah Myakala[3], Bryce A. Jones[4], Moshe Levi[3], Jeffrey B. Kopp[5], Teruhiko Yoshida[5], Laura J. Niedernhofer[6], David Manthey[7], Kyung Chul Moon[8], Seung Seok Han[9], Jarcy Zee[10], Avi Z. Rosenberg[2] and Pinaki Sarder[1]

[1]Department of Pathology and Anatomical Sciences, University at Buffalo, Buffalo, New York, USA; [2]Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA; [3]Department of Biochemistry, Molecular & Cellular Biology, Georgetown University, Washington, District of Columbia, USA; [4]Department of Pharmacology and Physiology, Georgetown University, Washington, District of Columbia, USA; [5]Kidney Disease Section, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland, USA; [6]Institute on the Biology of Aging and Metabolism, Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota, Minneapolis, Minnesota, USA; [7]Kitware Inc., Clifton Park, New York, USA; [8]Department of Pathology, Seoul National University College of Medicine, Seoul, Republic of Korea; [9]Department of Internal Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea; and [10]Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania, Perelman School of Medicine, Philadelphia, Pennsylvania, USA

Introduction: Podocyte depletion is a histomorphologic indicator of glomerular injury and predicts clinical outcomes. Podocyte estimation methods or podometrics are semiquantitative, technically involved, and laborious. Implementation of high-throughput podometrics in experimental and clinical workflows necessitates an automated podometrics pipeline. Recognizing that computational image analysis offers a robust approach to study cell and tissue structure, we developed and validated PodoCount (a computational tool for automated podocyte quantification in immunohistochemically labeled tissues) using a diverse data set.

Methods: Whole-slide images (WSIs) of tissues immunostained with a podocyte nuclear marker and periodic acid–Schiff counterstain were acquired. The data set consisted of murine whole kidney sections ($n = 135$) from 6 disease models and human kidney biopsy specimens from patients with diabetic nephropathy (DN) ($n = 45$). Within segmented glomeruli, podocytes were extracted and image analysis was applied to compute measures of podocyte depletion and nuclear morphometry. Computational performance evaluation and statistical testing were performed to validate podometric and associated image features. PodoCount was disbursed as an open-source, cloud-based computational tool.

Results: PodoCount produced highly accurate podocyte quantification when benchmarked against existing methods. Podocyte nuclear profiles were identified with 0.98 accuracy and segmented with 0.85 sensitivity and 0.99 specificity. Errors in podocyte count were bounded by 1 podocyte per glomerulus. Podocyte-specific image features were found to be significant predictors of disease state, proteinuria, and clinical outcome.

Conclusion: PodoCount offers high-performance podocyte quantitation in diverse murine disease models and in human kidney biopsy specimens. Resultant features offer significant correlation with associated metadata and outcome. Our cloud-based tool will provide end users with a standardized approach for automated podometrics from gigapixel-sized WSIs.

Correspondence: Avi Z. Rosenberg, Department of Pathology, Johns Hopkins University School of Medicine, 720 Rutland Avenue, Ross Building, Room 632D, Johns Hopkins Medical Institutions, Baltimore, Maryland 21205, USA. E-mail: arosen34@jh.edu; or Pinaki Sarder, Department of Pathology and Anatomical Sciences, University at Buffalo, 955 Main Street, Room 4204, Buffalo, New York 14203, USA. E-mail: pinakisa@buffalo.edu

Chronic kidney disease is a state of prolonged and progressive reduction in kidney function that may evolve to end-stage kidney disease (ESKD). Driven by increasingly prevalent conditions with high incidence (e.g., diabetes, hypertension), chronic kidney disease accounts for unprecedented mortality and socioeconomic burden.[1,2] To mitigate this, biomedical initiatives aim to identify disease-related biomarkers with

improved precision for early detection and intervention. For glomerular diseases, some biomarker studies have focused on podocytes, highly specialized epithelial cells that maintain the kidney filtration barrier. The podocyte depletion hypothesis proposes[3–8] that podocyte loss in the setting of podocytopathic injury from hyperfiltration, hyperglycemia, or hypertension is an early determinant of proteinuria and glomerulosclerosis.[9–11] Thus, podocyte enumeration offers a measurable indicator of irreversible glomerular injury and therapeutic success in these states.

Unfortunately, existing podocyte estimation methods or podometrics are semiquantitative and not scalable or offer morphologic assessment.[3,4,12–16] Furthermore, podocyte identification on routine and special stains viewed under brightfield microscopy remains difficult.[12] In a recent study, Venkatareddy *et al.*[16] present a novel methodology for podometric estimation, including podocyte count and density from a single histologic section. Recognizing that the number of nuclear profiles per glomerulus cross-section is not a true estimate of podocyte count, this single-section method applied podocyte nuclear labeling, manual annotation, image analysis software, and stereological equations to arrive at a correction factor (CF) which modulates podocyte count and volume density estimates based on section thickness. Designed for single glomerulus quantitation from image patches, this estimation method lacks scalability to the WSI context. To achieve big-data podocyte studies that facilitate early detection and intervention, accurate and automated methods for brightfield whole-slide podocyte quantification must be established. Therefore, we developed PodoCount (an automated podometric tool for single-section estimation from WSIs) and used it to evaluate WSIs of kidney sections immunostained with a podocyte marker (see the workflow in Supplementary Figure S1).

METHODS

Data Sources
Human data collection followed protocols approved by the Institutional Review Board at the Seoul National University College of Medicine (H-1812-159-998), Seoul, Republic of Korea. All experiments were performed according to federal guidelines and regulations. Animal studies were performed in accordance with protocols approved by the Institutional Animal Care and Use Committee at the Georgetown University, National Institutes of Health, University of Minnesota, and Johns Hopkins University, are consistent with federal guidelines and regulations, and are in accordance

with recommendations of the American Veterinary Medical Association guidelines on euthanasia.

Description of Murine and Human Data
We analyzed 135 whole kidney sections from 135 mice across 6 kidney disease models (Figure 1a,c) and 45 kidney biopsy specimens from patients with DN (Figure 1b,d), as detailed in the subsequent sections. These multi-institutional, murine and human, male and female data feature highly variable sample preparation, staining, imaging, and pathology, comprising a diverse data set to assess robustness and reproducibility (Supplementary Figure S2).

Models 1 and 2 (Diabetes)
There were 2 models of type 2 diabetes mellitus used (db/db and KKAy). Model 1 consisted of db/db (leptin receptor mutation) mice on BKS background ($n = 4$) that developed spontaneous/congenital disease and were compared with wild type (WT) BKS mice ($n = 6$).[17] Model 2 consisted of the KKAy mice ($n = 10$) (described previously[18]) that developed spontaneous diabetes of polygenic origin and were compared with WT mice ($n = 8$).

Model 3 (Focal Segmental Glomerular Sclerosis)
A postadaptive form of focal segmental glomerular sclerosis (FSGS) was induced in WT FVB/N mice ($n = 8$) by a combination of 4 interventions (SAND) (0.9% saline drinking water, angiotensin II infusion [osmotic pump], unilateral nephrectomy, and deoxycorticosterone [subcutaneous pellet][19,20]) and were compared with untreated mice ($n = 11$).

Model 4 (HIV-Associated Nephropathy)
Tg26 mice (gag-pol-deleted HIV-1 genome25) ($n = 13$) on an FVB/N × 129S F1 background with a collapsing glomerulopathy were compared with littermate controls ($n = 11$).

Model 5 (Aging)
In this aging model, 21-month-old C57BL/6 male mice ($n = 5$) were compared with 4-month-old controls ($n = 6$), both obtained from the National Institute on Aging rodent colony.[21]

Model 6 (Progeria)
An $Ercc1^{-/\Delta}$ progeroid mouse model ($n = 30$) and WT littermate controls ($n = 20$) (15–18 weeks old) on a C57BL/6J:FVB/N f1 background were used for this study.[22,23] Mice were bred and genotyped as previously described.[24]

Human Cohort
Human tissues consisted of needle biopsy samples from human patients with type 2 diabetes mellitus ($n = 45$) (biopsied 2011–2017 at the Seoul National University Hospital, which were collected from the
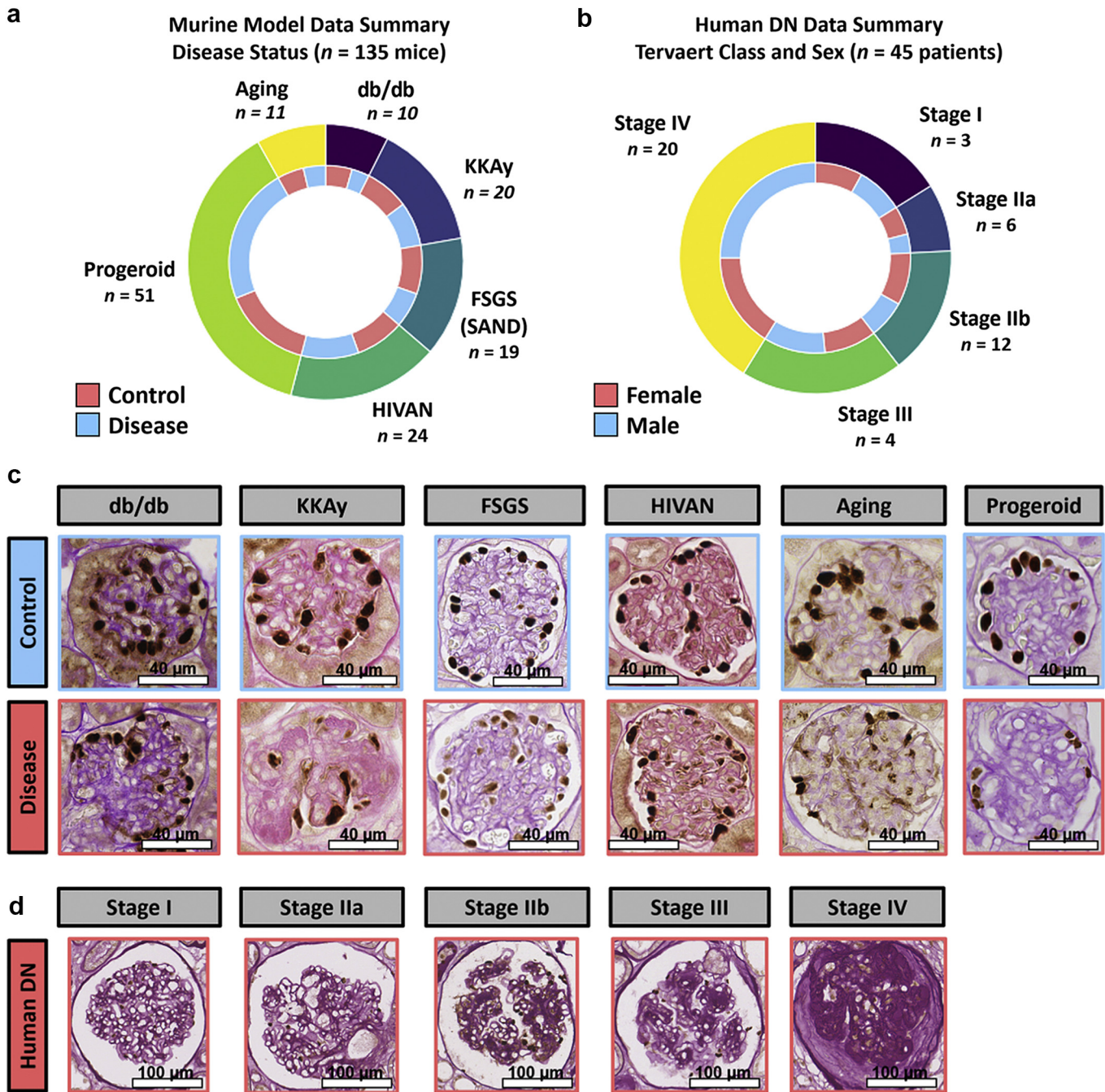
Figure 1. Summary of data sets. The image data set contains light microscopic images of kidney tissues from 6 mouse models of glomerular disease and 5 stages of human DN. (a) The murine cohort was composed of tissues from 135 mice with control and diseased specimens for each model. Two distinct models of type II diabetes mellitus were studied (db/db and KKAy). The SAND intervention (saline, angiotensin II, uninephrectomy, and deoxycortisone) models postadaptive FSGS (FSGS [SAND]). Samples from SAND, HIVAN, and Progeroid syndrome models included male and female mice; those from the db/db, KKAy, and Aging mouse models consisted only of males. (b) The human DN study consisted of 45 patients ($n = 35$ male and $n = 10$ female subjects). Representative glomerular p57-PAS image from (c) each mouse model and (d) each Tervaert stage of the human DN cohort. DN, diabetic nephropathy; FSGS (SAND), a postadaptive model of FSGS, focal segmental glomerular sclerosis; HIVAN, HIV-associated nephropathy.

Seoul National University Human Biobank). Biopsies were graded by a renal pathologist based on the Tervaert classification scheme.[25] Clinical metadata including serum creatinine and estimated glomerular filtration rate (eGFR)[26] were measured at the time of biopsy and at 1 year and 2 years postbiopsy. For the purposes of our study, progression to ESKD within 2 years after biopsy was the primary end point. Patients with type 1 diabetes mellitus and additional glomerulonephritis and patients without need for antidiabetic agents were excluded from the cohort.

Sample Preparation and Imaging

All samples were formalin-fixed, paraffin-embedded tissues cut at 2 μm in thickness. Podocyte nuclei were immunohistochemically labeled for p57^{kip2}, a marker of podocyte terminal differentiation[27] (catalog number ab75974, Abcam, Cambridge, United Kingdom), and detected with horseradish peroxidase and diaminobenzidine chromogen substrate (catalog number RU-HRP1000, Diagnostic BioSystems, Pleasanton, CA, and catalog number BSB0018A, Bio SB, Santa Barbara, CA, respectively). A periodic acid–Schiff poststain was applied without hematoxylin counterstain.

Brightfield WSIs were captured using an Aperio AT2 microscope (Leica Microsystems, Buffalo Grove, IL) or a NanoZoomer S360 slide scanner equipped with a 40× objective (Hamamatsu Photonics, Bridgewater, NJ). The full image data set consisted of WSIs of 135 whole murine kidney sections and 45 DN biopsy specimens, all from discrete mice and human participants (Figure 1).

Whole-Slide Segmentation of Renal Parenchyma

Images of podocytes and renal tissue compartments were extracted from WSIs through image segmentation by selecting image regions of interest based on differences in color, texture, and shape.[28] Segmented structures included whole tissue sections, glomerular boundaries (glomerular capillaries, Bowman's space, and Bowman's capsule), and podocyte nuclei. Select image processing techniques that require a more extensive explanation are italicized and defined in the glossary of terms in Supplementary Table S1. Structural segmentation involved several sequential steps. First, a global mean-based threshold was applied to segment the tissue section from the WSI background. Glomerular boundaries were then detected using our published Human-AI-Loop tool,[29] a convolutional neural network developed for WSI segmentation. For WSIs with >1 tissue section, tissue boundary segmentations were used to partition the whole-slide glomerular population into groups linked to their respective tissue section. Immunohistochemically positive podocyte nuclei were then segmented from detected glomerular units using stain deconvolution[30] and local mean-based thresholding. Morphologic image processing techniques,[28] including hole filling, size exclusion, and marker-controlled watershed,[31–33] were applied to refine segmentations and separate overlapping nuclei.

Computational performance of WSI segmentation was evaluated. Tissue section and glomerulus boundaries were manually annotated in 12 randomly selected WSIs, equally sampled from each murine data set. A stratified randomized sampling was performed, wherein for each cohort and disease state, samples (or images) were given a number and then selected via the random number generator in MATLAB ("rand" command). Per WSI, automated segmentations were compared pixel wise against manual ground truth. To quantify tissue and glomerulus boundary segmentation performance, the pixel-wise sensitivity, specificity, precision, and accuracy were calculated across WSIs, defined as,

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

where TP, true positive; TN, true negative; FP, false positive; and FN, false negative.

Glomerulus images were randomly and equally sampled from murine ($n = 40$ per cohort, 20 each from control and disease) and human ($n = 40$ per DN stage) data, and podocyte nuclei were manually annotated in each image. Pipeline segmentations were compared pixel wise against manual ground truth. To quantify podocyte nuclear segmentation performance, sensitivity, specificity, precision, and accuracy were computed per glomerulus image (as defined previously).

Performance of podocyte detection was assessed with Hit-Miss analysis to determine the frequency at which podocyte nuclei were positively identified. Median performance was computed per cohort (median of $n = 40$ murine glomerulus images or $n = 200$ human glomerulus images) and then across all data sets.

Computational Podocyte Count and Density Estimation From WSIs

Segmented podocyte nuclei were automatically enumerated in each glomerulus image as a raw count. Image analysis techniques were then applied in sequence to emulate the single-section method[16] (Supplementary Figure S3). First, the bounding box of each nuclear profile was derived, and the box length and width were averaged to find the apparent caliper diameter per profile ($d_{bounding\ box}$). Then, $d_{bounding\ box}$ values were averaged to d, and established equations[16] were used to estimate D, CF, the corrected podocyte count, and podocyte density (number per 10^6 μm^3).

Table 1. Histologic image feature definitions for podometrics

Features	Definition
PC	Corrected podocyte count after application of the single-section method's CF. Computed as number of podocyte nuclear profiles times the CF.
GA	Cross-sectional area of the glomerulus unit (μm^2).
GPD	Podocyte volume density is computed as the ratio of the corrected podocyte count to the glomerulus volume and approximates the spatial density of podocytes (number per 10^6 μm^3).
TPA	Total podocyte nuclear area is computed as the cumulative area of podocyte nuclear profiles for a given glomerular unit (μm^2).
GPC	Glomerular podocyte nuclear coverage is computed as the ratio of total podocyte nuclear area to glomerulus unit cross-sectional area.

2D, two-dimensional; CF, correction factor; GA, glomerulus area; GPC, glomerular podocyte nuclear coverage; GPD, glomerular podocyte density; PC, corrected podocyte count; TPA, total podocyte nuclear area.
Podocyte morphometrics are invaluable tools for prognostication. Podometric methodologies compute podocyte nuclear count, size, and spatial density, relative to glomerulus area, to provide quantitative modeling of progressive glomerular disease. These features are incorporated into PodoCount as PC, GA, GPD, TPA, and GPC to quantify podocyte depletion through image features engineered from digitized renal histopathology. All reported podometric feature values are based on 2D quantification from glomerulus profiles in whole kidney sections.

The performance of podocyte count and density estimates, which were output as continuous values, was assessed by calculation of error. To do this, computational estimates of podocyte count and density were compared against manual ground truth generated using MATLAB as described in Venkatareddy et al.[16] (Estimation of D Using Image-Pro Software and the Quadratic Equation). For a detailed explanation of ground truth generation, see the Supplementary Methods.

Whole-Slide Podocyte and Glomerulus Feature Extraction

Built-in morphologic operations were applied to derive whole-slide coordinate locations and geometric features from podocyte nuclei and glomerulus profiles.[34] Geometric features included image object area (μm^2), bounding box area (μm^2), convex area (μm^2), eccentricity, equivalent diameter, extent, major and minor axis lengths (μm), orientations, perimeters (μm), and solidities. A brief description of each feature is provided in Supplementary Table S1. Feature statistics were computed per podocyte, per glomerulus podocyte population, and per WSI glomerulus population.

Biologically Inspired Podocyte Feature Engineering

Additional morphologic and spatial features were engineered from podocyte nuclear profiles and glomerulus units. Features validated in this work included total podocyte nuclear area (μm^2) and podocyte nuclear coverage (Table 1). Validation of hand-crafted features (e.g., podocyte nuclear distance to the glomerulus unit edge, Supplementary Table S2) was the scope of future work, and thus further discussion was omitted.

Determination of the Number of Glomerulus Profiles Required to Reliably Estimate Podocyte Density

The number of glomerular profiles required to arrive at an estimate of podocyte density within 10% of the whole-slide mean value with 90% and 95% confidence was assessed per mouse model (Supplementary Figure S4). For each mouse (WSI), the true whole-slide podocyte density was calculated. Podocyte density estimates for $n = 2$ to 50 randomly sampled glomeruli were then computed, for 1000 trials apiece, and compared against the true value. The probability of arriving at an estimate within 10% of the true value (n successes of 1000 trials) was recorded per sampling, and the minimum n glomeruli required to achieve 90% and 95% confidence was determined. The study was repeated for estimates within 20% of the true value.

Statistical Analysis

Data were analyzed with Minitab Statistical Software version 19 (Minitab 17 Statistical Software 2010, Minitab, State College, PA) and MATLAB's Statistics and Machine Learning Toolbox. Feature means and SDs were reported to allow for interpretation of biological and/or clinical significance. The normality of each feature distribution was assessed with the Anderson-Darling test statistic.[35] Nonparametric tests were used for non-normal feature distributions. Differences between groups (e.g., WT vs. transgenic mice) were assessed with unpaired, 2-sample t tests. Nonparametric tests and Welch's correction for unequal variances were performed when appropriate.

Correlation analyses between murine histologic image features and terminal urinary albumin-to-creatinine ratio (UACR) were completed using the Spearman rank-order correlation measure (Spearman's ρ).[36] Spearman correlation was selected over Pearson correlation because of the non-normality of image feature data.

Using the Statistics and Machine Learning Toolbox in MATLAB, linear mixed-effects models were implemented to assess whether image features differed between disease statuses while accounting for multiple glomeruli from the same mouse kidney section.[37] In this study, WT and transgenic disease statuses consisted of the pooled glomerulus populations across WT and transgenic mice, respectively. In each linear mixed-effects model, the image feature was used as the outcome and the model included a fixed effect for binary disease status and a random intercept term for each mouse. Parameter estimation was conducted by maximizing the restricted log-likelihood of the model. Linear mixed-effects models were selected over unpaired, 2-sample t tests (i) to account for potential clustering of glomerulus subpopulations derived from

the same mouse and (ii) because linear mixed-effects models can handle missingness at random and therefore different numbers of glomeruli per mouse.

For analysis of human data, differences among 3 or more groups (e.g., DN stages) were assessed with Kruskal–Wallis nonparametric tests followed by *post hoc* Dunn's tests, as the data were not normally distributed and violated analysis of variance criteria.[38,39] Logistic regression was used to study ESKD outcome. First, a model was fit using only eGFR because eGFR is a benchmark outcome indicator.[40] To assess whether image features improved the prediction of ESKD beyond eGFR alone, subsequent models were fit using a single-image feature and adjusted for eGFR.

To correct for multiple hypothesis testing, we implemented the Benjamini-Hochberg procedure, with a false discovery rate of 0.05[41] for all analyses. A corrected *P* value (*q*) <0.05 was considered statistically significant. The Benjamini-Hochberg procedure was selected over the Bonferroni correction because of its robustness to high volume and nonindependence of tests.

Residual and absolute error were used to compare computational podocyte count and density estimates, respectively, against manual ground truth. Departures in computational estimates compared with ground truths were visualized with Bland-Altman plots. Residual error was calculated as the difference between ground truth and computational estimates ($y - \widehat{y}$), and absolute error was calculated as the absolute value of the difference between ground truth and computational estimates ($|y - \widehat{y}|$). Correlation between computational and ground truth estimates was assessed using Pearson's correlation analysis (Pearson's *R*).[36]

Deployment of Whole-Slide Podocyte Analysis With Cloud Computation

HistomicsUI,[42] a distributed system with RESTful application programming interface, was developed by Kitware (Clifton Park, NY) and was used to deploy our algorithm as a plugin, thereby creating an online platform that would enable multiple users to detect and quantify podocytes via a web interface. The algorithm was packaged in the form of a Docker image using *Docker* software (Palo Alto, CA),[43–45] a framework that enables users to build and run applications in containers. The generated container conforms to the Slicer CLI workflow interface, which allows *HistomicsUI* to display a user interface to adjust algorithm parameters.

Hardware

Computational processing was performed on a Linux distribution operating system (Ubuntu 16.04) with 2 Intel Xeon Silver 4114 processors, each with 10 cores, running at 2.20 GHz and equipped with 64 GB of random-access memory. Neural network training and predictions for glomerulus boundary detection were performed using a NVIDIA Quadro RTX 5000 GPU (16 GB of memory). *HistomicsUI* plugin is made available for end users in a research computer with Intel i5 6-core processor, running at 3.1 to 4.5 GHz and equipped with 16 GB of random-access memory.

Data Availability

To support reproducibility, we released fully annotated pipeline codes along with sample image data, glomerulus annotation files, and all pipeline output (i.e., podocyte nuclear annotations files, feature files). We also launched our cloud-based *PodoCount* plugin for the end user community and created an instructional video for first-time users (Supplementary Movie S1) accessible via http://hermes.med.buffalo.edu:8080. All codes and documentation, Docker image of the web cloud interface, and data are available via http://bit.ly/3rdGPEd and our Github repository,[46] https://github.com/SarderLab/PodoCount.

RESULTS

Qualitative and Quantitative Performance Analysis

Visual inspection of pipeline-derived podocyte nuclear, glomerular, and tissue boundaries confirmed successful region detection and segmentation (Supplementary Figure S5). Computational performance evaluation was first completed to assess the quality of image segmentation. Segmented image regions included tissue sections, boundaries of glomerular units, and podocyte nuclei. Across all randomly sampled images/ROIs, we observed high performance for all segmentation tasks (Table 2 and Supplementary Figure S6). The median sensitivity, specificity, precision, and accuracy in tissue boundary, glomerulus boundary, and podocyte nuclear segmentation tasks were 0.99/0.99/0.99/0.99 (tissue boundary), 0.97/0.99/0.92/0.99 (glomerulus boundary), and 0.85/0.99/0.93/0.99 (podocyte nuclear segmentation), respectively (see Table 2 for additional results, including averages and ranges of each performance metric). Furthermore, the Hit-Miss analysis determined that the frequency at which podocyte nuclei were positively identified was 0.98.

The accuracy of podocyte count and density in PodoCount was then assessed. For each cohort, automated, continuous counts and density estimates were compared against manual ground truth. Podocyte count error was bounded by 1 podocyte per glomerulus (Table 3 and Figure 2a). Automated and manual counts were strongly and significantly correlated across all cohorts. Departure of automated counts from

Table 2. Computational performance of podocyte detection and image segmentation

Segmented region(s)	Assessment of image segmentation and podocyte detection				
	Sensitivity med. (avg./[range])	Specificity med. (avg./[range])	Precision med. (avg./[range])	Accuracy med. (avg./[range])	Hit-miss percent
Tissue boundary	0.996 (0.994/0.985–0.997)	0.998 (0.998/0.996–1.000)	0.990 (0.990/0.977–0.998)	0.997 (0.998/0.996–0.999)	—
Glom boundary	0.966 (0.965/0.954–0.977)	0.999 (0.999/0.999–1.000)	0.916 (0.917/0.875–0.943)	0.999 (0.999/0.999–1.000)	—
Podocyte nuclei	0.846 (0.834/0.425–0.993)	0.997 (0.994/0.980–1.000)	0.931 (0.933/0.763–1.000)	0.997 (0.992/0.943–1.000)	0.980

Avg., average; DN, diabetic nephropathy; Glom, glomerulus; Med., median.
Sensitivity, specificity, precision, and accuracy were computed for image segmentation tasks, and percentage accuracy in podocyte nuclear detection was assessed by Hit-Miss (i.e., frequency at which podocyte nuclei were positively identified). Performance analysis for segmentation of tissues and glomerulus unit boundaries was completed for $n = 12$ randomly selected whole-slide images, equally distributed across data sets and disease states. Assessment of podocyte detection and image segmentation tasks was completed using $n = 240$ glomerulus images randomly and equally sampled from murine ($n = 40$ per cohort, 20 each from control and disease) and human ($n = 40$ per DN stage) data. High performance was observed for segmentation of tissue and glomerulus unit boundaries. Although podocyte nuclear segmentation was less sensitive and precise, podocytes were positively detected 98% of the time. Lesser performance in podocyte nuclear segmentation was attributed to the challenge of manually delineating a consistent boundary about immunohistochemistry-labeled nuclei in brightfield images.

ground truth was greatest for the DN cohort, where histologic manifestation of disease was most pronounced (Figure 1d). Absolute errors in PodoCount density estimates were near zero (Figure 2b), and across all cohorts, automated and ground truth density estimates were strongly and significantly correlated (Table 4).

Podocyte and Glomerulus Feature Significance Across Murine Models

Quantified image features were compared based on their ability to differentiate diseased tissue from normal tissue. Statistical analysis focused on the following image features: glomerulus area, podocyte density, total podocyte nuclear area, and podocyte nuclear coverage. For each model, statistical tests compared the image features of (i) diseased mice against WT mice and (ii) the glomerulus populations of diseased mice against those of WT mice. For glomerulus-level comparisons, see Supplementary Figure S7 and Supplementary Tables S3 to S8, part B. Strength and association of image features and UACR were assessed for those models where complete UACR data were available.

1. db/db model. Although no podocyte feature proved significant when comparing db/db and WT mice (Supplementary Table S3), the distributions

were different across all features with greater variance observed in disease (Figure 3a). In db/db mice, glomerular area and total podocyte nuclear area were greater, whereas podocyte density was less. Moderate strength of correlation ($|\rho| > 0.4$) was observed pairwise between these image features and UACR; with increasing UACR, glomerular and total podocyte nuclear area increased whereas podocyte density decreased (Supplementary Table S9). Nuclear coverages were similar in diabetes and had little correlation with UACR.

2. KKAy model. In the KKAy model of diabetes, podocyte density was significantly less, whereas glomerular and total podocyte nuclear area were significantly greater (Supplementary Table S4, part A and Figure 3b). Strong, significant correlations ($|\rho| > 0.75$) were also observed between these image features and UACR (Supplementary Table S10). UACR increased with increasing glomerular and total podocyte nuclear area, as well as decreasing podocyte density.

3. FSGS model. SAND-treated mice showed significantly lesser podocyte densities and nuclear coverages with accompanying increases in the glomerular area. Total podocyte nuclear area was similar

Table 3. Comparison of podocyte counts by PodoCount and the single-section method

Cohort	Comparison of PodoCount automated counts vs. the single-section method				
	Median error in estimation		Pearson correlation analysis		
	Residual	Absolute	R^2	95% CI for R^2	P value
db/db	−0.01	0.26	0.95	(0.91–0.98)	<0.001
KKAy	−0.36	0.47	0.82	(0.68–0.90)	<0.001
FSGS	−0.17	0.25	0.89	(0.80–0.94)	<0.001
HIVAN	0.06	0.25	0.97	(0.94–0.98)	<0.001
Aging	−0.28	0.37	0.96	(0.92–0.98)	<0.001
Progeroid	0.10	0.16	0.94	(0.87–0.97)	<0.001
DN	0.10	0.17	0.57	(0.35–0.57)	<0.001

DN, diabetic nephropathy; FSGS, focal segmental glomerular sclerosis; HIVAN, HIV-associated nephropathy.
PodoCount estimates of corrected podocyte counts were compared against manual ground truth generated using the single-section method. Error in automated counts was bounded by 1 podocyte. Tendency toward over- or underestimation was cohort dependent. Automated counts were strongly and significantly correlated with ground truth counts across all cohorts. Correlation results between automated and ground truth podocyte density estimates were evaluated with parametric Pearson analysis (R^2 value and 95% CIs reported, in addition to P values).
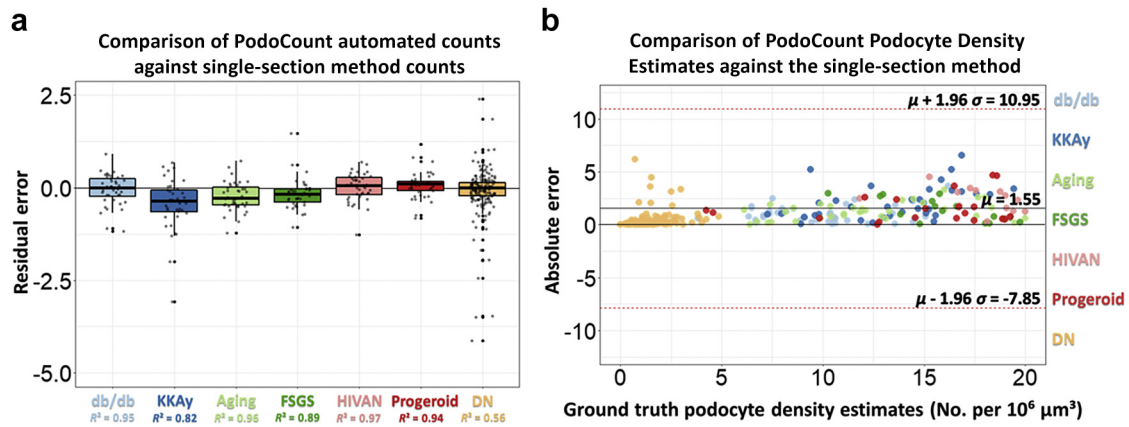
Figure 2. Comparison of podometric estimates by PodoCount and the single-section method. PodoCount estimates of corrected podocyte count and podocyte density were compared against those from manual ground truth measurements using the single-section method. (a) Error in automated counts was bounded by 1 podocyte. Tendency toward over- or under- estimation was cohort dependent. (b) The modified Bland-Altman plot highlights the departure in PodoCount podocyte density estimates from ground truth. No., number; FSGS (SAND), focal segmental glomerular sclerosis; HIVAN, HIV-associated nephropathy.

between WT and FSGS-affected mice (Supplementary Table S5, part A). The podocyte density distributions of control and FSGS mice were markedly different, with control mice featuring a broad range of average podocyte densities and FSGS mice featuring a narrow range of reduced podocyte densities (Figure 3c). Moderate-to-strong correlations described the relationships between UACR and image features (Supplementary Table S11). With increasing UACR, significant reduction in podocyte density was observed.

4. HIV-associated nephropathy (HIVAN) model. Image features did not significantly differentiate Tg26 mice from WT mice (Supplementary Table S6, part A). Before correction for multiple testing, the podocyte counts of Tg26 and WT mice were significantly different as demonstrated by the marked difference in their respective distributions (Figure 3d). All other feature values and distributions were comparable. In a transgenic model of HIVAN, variable

penetrance among mice and within a single animal contributes to subtler histomorphologic phenotypes.

5. Aging model. In contrast with young mice, old mice featured significantly greater glomerular area, as well as significantly reduced podocyte density and podocyte nuclear coverage (Figure 3e, Supplementary Table S7, part A). Total podocyte nuclear area did not significantly differ with increased age.

6. Progeria ($Errc1^{-/\Delta}$) model. Significantly lower glomerular areas and podocyte densities were observed in $Errc1^{-/\Delta}$ mice (Supplementary Table 8, part A and Figure 3f). The podocytes' nuclear coverages of WT and $Errc1^{-/\Delta}$ mice were comparable. Before correction for multiple testing, significantly lower total podocyte nuclear areas were noted in $Errc1^{-/\Delta}$ mice.

For each model, a single estimate of the apparent mean caliper diameter (d), true mean caliper diameter (D), and CF was determined (Supplementary Table S12). The CFs per model were found to be 0.19 in db/db, 0.20

Table 4. Comparison of podocyte density by PodoCount and the single-section method

| | Comparison of PodoCount automated counts vs. the single-section method | | | | |
| | Median error in estimation | | Pearson correlation analysis | | |
Cohort	Residual	Absolute	R^2	95% CI for R^2	P value
db/db	−0.01	0.75	0.96	(0.93–0.98)	<0.001
KKAy	−1.88	2.18	0.92	(0.85–0.96)	<0.001
FSGS	−1.59	2.29	0.97	(0.93–0.98)	<0.001
HIVAN	0.51	1.80	0.98	(0.96–0.99)	<0.001
Aging	−0.93	1.38	0.97	(0.95–0.99)	<0.001
Progeroid	0.59	1.59	0.94	(0.88–0.97)	<0.001
DN	0.11	0.15	0.39	(0.24–0.48)	<0.001

DN, diabetic nephropathy; FSGS, focal segmental glomerular sclerosis; HIVAN, HIV-associated nephropathy.
PodoCount estimates of podocyte density were compared against manual ground truth generated using the single-section method. Absolute error in automated estimates was on average 1.55 (number per 10^6 μm^3). Tendency toward over- or underestimation was cohort dependent. Automated estimates were strongly and significantly correlated with ground truth estimates across all cohorts. Correlation results between automated and ground truth podocyte density estimates were evaluated with parametric Pearson analysis (R^2 value and 95% CIs reported, in addition to P values).
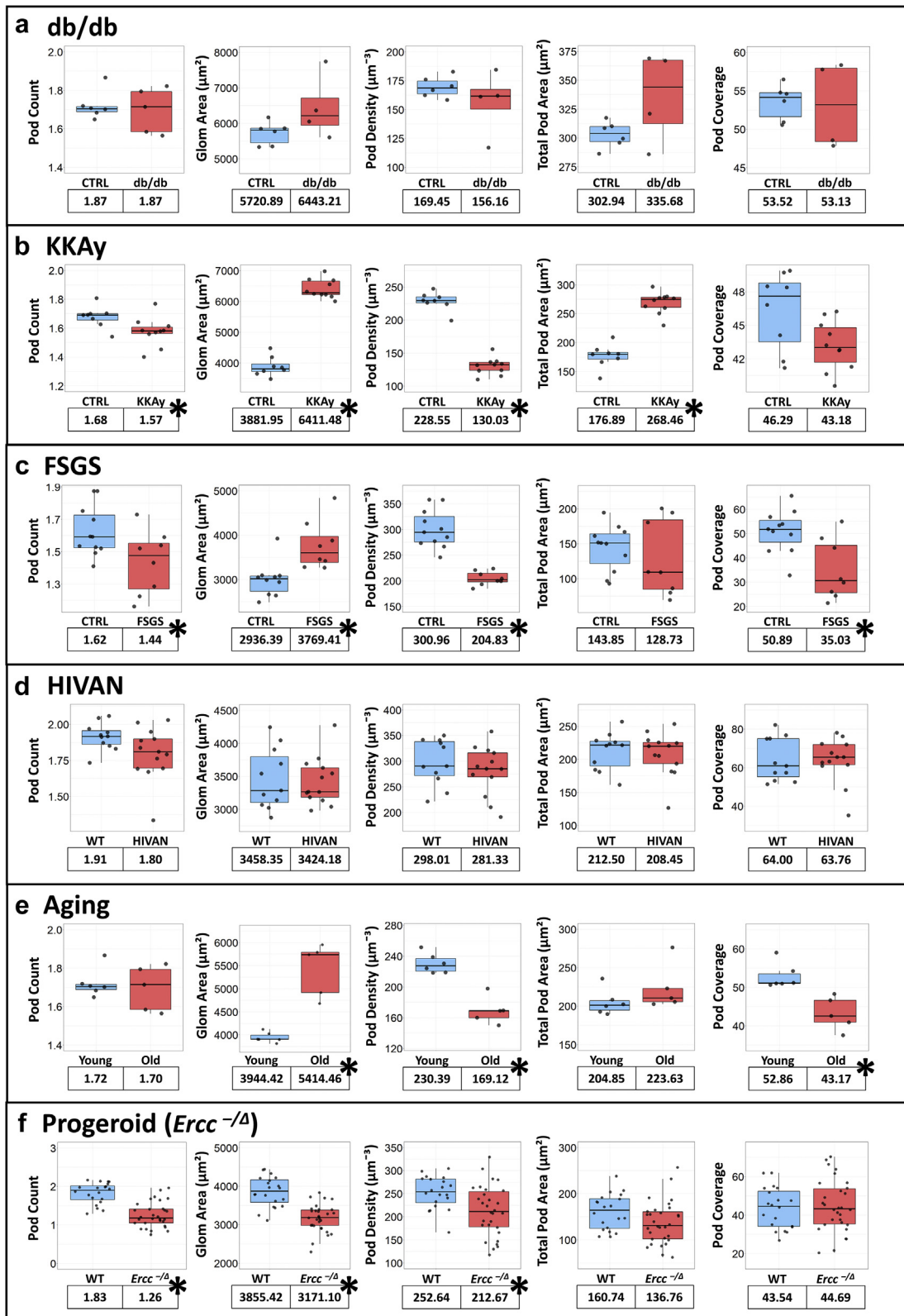
Figure 3. Podocyte and glomerular morphometrics of control and disease mice across murine models. Distribution of podocyte feature values across disease states with each black dot corresponding to a single mouse in the (a) db/db model of type II diabetes mellitus, (b) KKAy model of type II diabetes mellitus, (c) FSGS model, (d) HIVAN model, (e) Aging model, and (f) Progeroid ($Ercc^{-/\Delta}$) model. All podometric values are based on 2D quantification from glomerulus profiles in whole kidney sections. *$q < 0.05$. 2D, 2-dimensional; CTRL, control; FSGS, focal segmental glomerular sclerosis; Glom =, glomerulus; HIVAN, HIV-associated nephropathy; Pod, podocyte; WT, wild type.

Table 5. Podocyte morphometrics significantly differentiated diabetic nephropathy stages IIb and III

	Diabetic nephropathy cohort				
	(A) Summary of feature values across patients' DN stages according to the Tervaert classification scheme				
	Mean ± SD				
Image feature	Stage I	Stage IIa	Stage IIb	Stage III	Stage IV
n	3	6	12	4	20
PC	1.09 ± 0.24	1.15 ± 0.72	1.40 ± 0.56	0.52 ± 0.31	0.85 ± 0.51
GA	29660.26 ± 4644.71	10085.43 ± 16639.20	31116.08 ± 8381.64	22629.21 ± 4758.69	23256.96 ± 9121.63
GPD	16.82 ± 3.85	19.96 ± 9.79	19.89 ± 6.31	12.55 ± 8.10	17.85 ± 7.91
TPA	95.76 ± 30.14	99.69 ± 70.29	155.42 ± 98.43	28.34 ± 27.70	84.94 ± 86.35
GPC	3.09 ± 0.65	2.82 ± 1.44	4.14 ± 2.09	0.98 ± 0.80	2.73 ± 1.66

	(B) Kruskal–Wallis and *post hoc* Dunn's tests comparing DN stages at the patient level	
Image feature	Significantly different DN stages	*q* value
PC	IIb from III	0.007[a]
GA	IIb from IV	0.067
GPD	No distinct groups	0.646
TPA	IIb from III	0.012[a]
GPC	IIb from III	0.017[a]

DN, diabetic nephropathy; GA, glomerular area; GPC, glomerular podocyte nuclear coverage; GPD, glomerular podocyte volume density; *n*, number of unique patient thin-needle biopsies per DN stage; PC, corrected podocyte count; TPA, total podocyte nuclear area.

[a]$q < 0.05$.

Pipeline-computed features were ranked based on their ability to differentiate between DN stages defined by the Tervaert classification scheme. (A) The table summarizes the mean and SD of each feature across Tervaert stages. (B) Corrected podocyte count was the most significant indicator of disease in DN, followed by total podocyte nuclear area and podocyte nuclear coverage. Pairwise tests revealed that differences in podocyte counts and nuclear morphometrics are consistently observed between DN stages IIb and III. Podometric values are based on 2D quantification from glomerulus profiles in whole kidney sections. Feature units: GA, μm^2; GPD, number per $10^6 \mu m^3$; TPA, μm^2. Statistical conclusions are based on Kruskal–Wallis and *post hoc* Dunn's tests comparing population medians at a significance level of 0.05. H_0: At least one population mean is different.

in KKAy, 0.22 in FSGS, 0.21 in HIVAN, 0.20 in Aging, and 0.22 in Progeroid. These CF values align well with those previously reported.[16]

Glomerular Sampling for Accurate Podocyte Density Estimates

The number of glomerular profiles required to arrive at an estimate of podocyte density within 10% of the true (whole-slide) value with 90% and 95% confidence was studied for each model based on murine phenotype. Across all cohorts, accurate estimation in the diseased state required more glomerular profiles (Supplementary Table S13). The number of profiles required to achieve an estimate within 10% of the true value with 90% confidence for each cohort's control and disease groups, respectively, was found to be 32 and 38 in db/db, 27 and 29 in KKAy, 30 and 37 in FSGS, 27 and 28 in HIVAN, 28 and 33 in Aging, and 35 and 44 in Progeroid, respectively. As expected, the number of required profiles increased for estimates with 95% confidence. Achievement of a 90% confident estimate with 10 or fewer profiles required relaxation of the constraint to estimation within 20% of the true value. These data demonstrate the value of a robust pipeline that assesses all the available glomeruli to eliminate model-specific variability and potential sampling bias.

Podometrics in Clinical Human DN Biopsies

Biopsy-level features were compared among DN subjects based on their Tervaert classification[25] (Table 5) and outcome (Table 6 and Figure 4a and b). Corrected podocyte count was the lead indicator of DN stage at

the patient level, with a marked reduction in glomerular podocyte number (from 1.4 to 0.52) defining the transition from DN stage IIb to III (part B in Table 5). Total podocyte nuclear area and podocyte nuclear coverage were also significantly different across DN stages and were characteristic of the transition from stage IIb to III. Meanwhile, significantly greater glomerular area differentiated stage IIb from IV. These observations are consistent with established histopathologic classification criteria[25,47] and suggest a relationship with progressive podocyte injury, glomerular hypertrophy, and transition between DN stages. Single estimates of d, D, and CF were also determined for the human DN cohort (Supplementary Table S12). The CF value representative of the entire data set was 0.22. This human CF value is comparable to what has been previously reported.[16]

Podometrics Predict ESKD in DN

A logistic regression model fit using the DN cohort's eGFRs at time of biopsy significantly predicted ESKD with an odds ratio of 0.75 (Table 7). Each additional unit increase in eGFR was associated with a 25% decrease in the odds of patient progression to ESKD, a finding supported by clinical standards that uphold eGFR as a key indicator of outcome.[40] A series of logistic regression models each fit using an engineered image feature with adjustment for eGFR demonstrated that glomerular area and total podocyte nuclear area were also significant predictors of ESKD. For each 1000 μm^2 increase in mean biopsy glomerular area, the odds of ESKD was increased by 43%. Similarly, each 10 μm^2

Table 6. Biopsy nuclear podocyte pathology is predictive of progression to end-stage kidney disease

	Diabetic nephropathy cohort				
	Two sample *t* test comparing nonprogressors (*n* = 31) against progressors to ESKD (*n* = 14)				
	Feature summary (mean ± SD)		Difference of means (ESKD – no ESKD)		
Image feature	No ESKD	ESKD	Difference	95% CI	*q* value
PC	3.44 ± 1.93	1.97 ± 1.05	−1.47	(−2.37 to −0.57)	0.002[a]
GA	28432.41 ± 9450.26	21426.01 ± 5384.96	−7006.40	(−11502.52 to −2510.29)	0.003[a]
GPD	19.22 ± 7.52	15.73 ± 7.38	−3.49	(−8.40 to 1.43)	0.156
TPA	124.26 ± 95.79	50.74 ± 29.88	−73.52	(−111.85 to −35.18)	<0.001[a]
GPC	3.01 ± 2.10	2.21 ± 1.31	−0.80	(−2.00 to 1.31)	0.002[a]

ESKD, end-stage kidney disease; GA, glomerular area; GPC, glomerular podocyte nuclear coverage; GPD, glomerular podocyte density; PC, corrected podocyte count; TPA, total podocyte nuclear area.

[a]$q < 0.05$.

The table summarizes the mean and SD of each feature across patient outcomes. Significant reduction in corrected podocyte count, glomerular area, total podocyte nuclear area, and glomerular podocyte nuclear coverage was characteristic of progressor biopsies. Podometric values are based on 2D quantification from glomerulus profiles in whole kidney sections. Feature units: GA, μm^2; GPD, number per 10^6 μm^3; TPA, μm^2. Statistical conclusions are based on 2-sample *t* tests comparing murine population means at a significance level of 0.05.

increase in biopsy total podocyte nuclear area was associated with a 25% decrease in the odds of ESKD. Odds ratio-associated *P* values also suggested that, together, total podocyte nuclear area and eGFR were more significantly related to ESKD incidence than eGFR alone. When adjusted for eGFR, podocyte count, podocyte density, and glomerular podocyte nuclear coverage were not found to be significant predictors of ESKD (odds ratios 0.91, 0.80, and 0.89, respectively). These studies highlight the potential for encoded features in histomorphology to increase the predictive power of clinical metrics.

PodoCount in the Cloud

PodoCount was deployed as a cloud-based plugin on the Sarder Laboratory's Digital Slide Archive[48] (Supplementary Figure S8). This integration was facilitated by HistomicsTK, a web-based tool that allows for installation of user-defined algorithms as plugins in a virtual user interface, HistomicsUI, and is supported by the OpenSlide[49] library for handling proprietary digital pathology WSI formats. PodoCount end users need only upload a WSI and glomerulus annotation file and select the option for PodoCount analysis. Predicted podocyte nuclear annotations are displayed in the cloud and available for download (.xml. for standard desktop pathology viewers) along with image features files (.csv). An instructional video for end users was shared along with open-source documentation and codes to maximize plugin accessibility (Supplementary Movie S1, http://bit.ly/3rdGPEd).

DISCUSSION

In this work, we introduced PodoCount, a novel tool for automated, whole-slide assessment of podocyte depletion and nuclear morphometry. Compared with
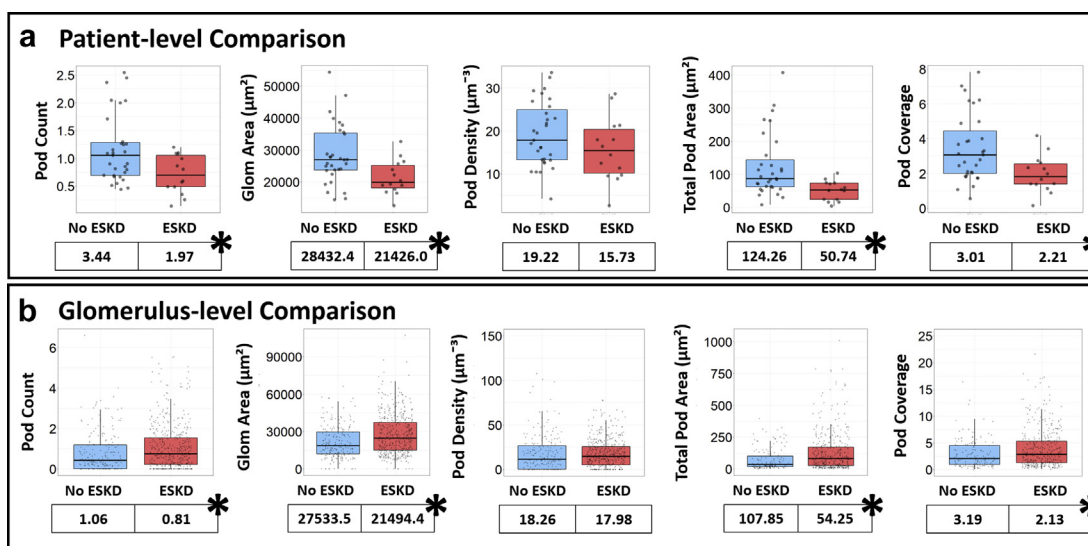

Figure 4. Podocyte and glomerular morphometrics in diabetic nephropathy kidney biopsy specimens predict outcome. Distribution of podocyte feature values comparing those with diabetes with progression to ESKD to those without with each black dot corresponding to a single patient (a) or glomerulus (b). All podometric values are based on 2D quantification from glomerulus profiles in whole kidney sections. *$q < 0.05$. 2D, 2-dimensional; ESKD, end-stage kidney disease; Pod, podocyte; Glom, glomerulus.

Table 7. Nuclear indicators of podocyte pathology may improve patient prognostication from time of biopsy

	Diabetic nephropathy cohort					
	Logistic regression for prediction of patient progression to ESKD ($n = 14$ ESKD, 31 no ESKD)					
	eGFR			Feature		
Image feature (adjusted for eGFR)	OR	95% CI	P value	OR	95% CI	q value
eGFR	0.75	0.58–0.96	0.020[a]	—	—	—
eGFR + PC	0.54	0.23–1.29	0.167	0.91	(0.83–0.99)	0.285
eGFR + GA	0.74	0.57–0.96	0.021[a]	1.43	(0.34–6.06)	0.023[b]
eGFR + GPD	0.69	1.01–1.78	0.047[a]	0.80	(0.54–1.20)	0.389
eGFR + TPA	0.74	0.58–0.96	0.022[a]	0.75	(0.60–0.95)	0.018[b]
eGFR + GPC	0.74	0.58–0.95	0.022[a]	0.89	(0.31–2.61)	0.836

eGFR, estimated glomerular filtration rate at time of biopsy; ESKD, end-stage kidney disease; GA, glomerulus area; GPC, glomerular podocyte nuclear coverage; GPD, glomerular podocyte density; OR, odds ratio; TPA, total podocyte nuclear area.
[a]Logistic regression models were evaluated for feature-based prediction of patient outcome in diabetic nephropathy. Each pipeline-computed image feature was evaluated in combination with eGFR as predictors of ESKD incidence (response variable). Response frequency was 14 of 45 patients. The values and 95% CIs for odds ratios were reported, in addition to their associated P values. Select image features were rescaled to provide interpretable unit changes in OR, including GA by 1000, GPD by 10, and TPA by 10. Feature units: GA, μm^2; GPD, number per 10^6 μm^3; TPA, μm^2.
[b]$q < 0.05$.

existing automated methods,[50,51] PodoCount takes a deterministic approach to podocyte quantification, using classical image analysis in place of deep learning. Benefits of this approach include computational simplicity, strong performance without big-data requirements, and thus no need for expensive GPUs or training time. PodoCount's computational image analysis approach was informed by the benchmark method for estimation of podocyte density from a single histologic section developed by Venkatareddy et al.[16] We validated this tool using WSIs from 6 distinct mouse models ($n = 135$) and human DN ($n = 45$) biopsy specimens from male and female mice and participants. These data were curated from multiple institutions and feature lab-to-lab technical variability, aspects that are known to hinder generalizable computational frameworks,[52–56] including highly variable sectioning and staining methods, as well as image and tissue quality (Supplementary Figure S2). PodoCount navigated this challenging data set well, achieving precise segmentation and accurate detection of podocyte nuclei (Table 2), as well as highly accurate podocyte density estimates (Tables 3 and 4 and Figure 2). When paired with strategic, literature-informed, feature engineering, this computational performance facilitated robust quantification of podocyte depletion across varied renal pathologies using a podocyte nuclear marker and nuclear-based quantification from 2-dimensional cross-sections.

Our nuclear-based approach was informed by prior studies which demonstrated that podocyte nuclear quantification is sufficient to estimate glomerular podocyte density and predict podocyte pathology.[15,16] Emphasis on nuclear quantification guided our choice of podocyte label. p57[kip2] was a podocyte-exclusive nuclear label within the glomerular microenvironment enabling reproducible detection amidst proximate cell populations. Two-dimensional analysis was opted for over 3-dimensional podocyte visualization to maximize feasibility and efficiency while maintaining accuracy in podocyte estimation. Prior studies validated that podocyte estimation from single sections is comparable to 3-dimensional, multisection methods, irrespective of disease state.[3,12,14–16] Murine data CFs and podocyte density estimates computed using PodoCount aligned well with the literature,[16,50,57–59] as did human CF values.[16] Podocyte density estimates for human DN data were slightly lower than those values reported for prior studies in early diabetes and clinical nephropathy.[60]

According to the podocyte depletion hypothesis, podocyte depletion may manifest (i) absolutely, as a reduction in glomerular podocyte count, or (ii) relatively, when pathologic increase in glomerular area reduces podocyte spatial density.[3] In this work, we found that computational podocyte features reproducibly predict disease-specific histopathology. Although the db/db model emulates early changes in human DN,[61] the KKAy model, renowned for rapid glomerular basement membrane thickening,[62] recapitulates the later, morphologically advanced stages of DN. Quantified feature trends in the db/db and KKAy models reflected these tendencies toward mild and advanced diabetic changes, respectively. Similarly, computed image features from the aging and progeroid models aligned well with our current understanding of glomerular senescence. Studies demonstrated that with age relative podocyte depletion from compensatory podocyte hypertrophy peaks at a 2.5-fold increase, giving way to absolute podocyte loss and progressive glomerulosclerosis. Observed feature trends in the aging and progeroid models aligned well with the literature, with marked increase in total podocyte nuclear area, in the presence of reduced podocyte volume

density, reflecting peak podocyte hypertrophic capacity, followed by overt podocyte loss.

The postadaptive FSGS phenotype is characterized by glomerular hypertrophy and podocyte loss[63,64] and is best described as a combination of absolute and relative podocyte depletion. Computed, significant decrease in podocyte density, as well as increase in glomerular area, for SAND-treated mice in the FSGS model was consistent with FSGS pathology. Similarly, quantified podocyte loss in Tg26 mice was consistent with HIVAN pathology.[65,66]

From the human DN cohort, we learned that podocyte nuclear image features have the potential to be valuable diagnostic and prognostic tools. Podocyte nuclear metrics differentiated patient biopsy specimens ($n = 45$) according to Tervaert class, highlighting the transition from stage IIb to III as a key turning point in DN pathology (Table 5) and outcome (Table 6 and Figure 4). Intriguingly, logistic regression analysis demonstrated that podocyte morphometrics derived from patient biopsy specimens have the potential to improve prediction of ESKD beyond eGFR alone (Table 7). Total podocyte nuclear area was more significantly associated with ESKD than eGFR alone, and together, these 2 measures significantly predicted progression to ESKD. Glomerular sampling studies for reliable estimation of podocyte density were not completed for human data wherein the number of glomerulus profiles per patient biopsy was limited (Supplementary Table S14). We recognize that the number of human samples is a limitation of the study and emphasize that our findings warrant future studies with greater statistical power.

As the irreversible end point of chronic kidney disease, ESKD is characterized by complete loss of kidney function and patient dependence on dialysis or transplant for survival. The presence of significant feature-response relationships underscores the potential for biopsy podocyte features to increase the precision of clinical metrics in chronic kidney disease prognostication, and thus improve patient outcomes. With potential to augment both experimental and clinical workflows, PodoCount was launched as an open-source cloud-based tool to maximize accessibility and promote standardization of podocyte morphometrics.

DISCLOSURE

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

BAS conceptualized and performed the quantitative analyses, designed and conducted the computational methods, completed all statistical analyses, interpreted the results, and wrote the manuscript. DG and DM assisted with the adaptation of *PodoCount* codes to the cloud format. PD assisted with completion of ground truth podocyte counts. XY assisted in manuscript preparation. db/db, KKAy, and Aging cohorts were generated by the Levi Lab (XXW, KM, BAJ, ML). The FSGS (SAND) and HIV mouse cohorts were generated by the Kopp Lab (JBK, TY). JBK also assisted in manuscript editing. The *Ercc1*$^{-/\Delta}$ Progeroid cohort and littermate controls were generated by the Niedernhofer Lab (LJN). SSH and KCM provided human biopsy and clinical data for patients with diabetes. JZ assisted with statistical design and analytical considerations. AZR conceived and optimized the staining technique that enables this computational method. AZR also assisted in manuscript preparation. PS conceived the idea of automated computational quantification and enumeration of podocytes from digital whole-slide images in the

domain of renal pathology and making this tool available for nephrology community via an open-source cloud resource. PS also assisted in manuscript preparation, coordinated with the study team, assisted in study design, supervised the computational implementation, and critically analyzed the results.

SUPPLEMENTARY MATERIAL

Supplementary File (PDF)
Supplementary Methods.
Figure S1. Flowchart of the PodoCount pipeline.
Figure S2. PodoCount manifests scientific rigor and reproducibility.
Figure S3. Schematic overview of PodoCount's computational image analysis approach to engineer podocyte count and density from glomerulus images.
Figure S4. Schematic overview of confidence study designed for determination of the number of glomerulus profiles required to arrive at a reliable estimation of podocyte density.
Figure S5. Nuclear enumeration and quantitation from segmented whole-slide images.
Figure S6. Computational performance distribution across podocyte detection and image segmentation tasks.
Figure S7. Podocyte and glomerular morphometrics of control and disease phenotypes across murine models.
Figure S8. PodoCount in the Cloud.
Table S1. Digital image processing terms and definitions.
Table S2. Feature definitions provided for ease of biological interpretation of computational findings.
Table S3. Trends consistent with early diabetic nephropathy histopathology were observed in the db/db cohort.
Table S4. Podocyte depletion and podocyte nuclear hypertrophy were observed in the KKAy cohort.
Table S5. Podocyte depletion was observed in a model of postadaptive focal segmental glomerulosclerosis.
Table S6. Absolute podocyte depletion indicated disease in a model of HIV-associated nephropathy.
Table S7. Glomerular hypertrophy drove reduction in podocyte density and coverage with increased age.
Table S8. Reduced podocyte density was driven by absolute podocyte depletion in a model of Progeria.
Table S9. Significant correlations were not observed between histologic morphometrics and terminal UACR in the db/db model.
Table S10. Significant correlations were observed between histologic morphometrics and terminal UACR in the KKAy model.
Table S11. Significant correlations were observed between histologic morphometrics and terminal UACR in the FSGS mouse model.
Table S12. Reference values for estimated true mean caliper diameter (D) across study data sets.

Table S13. Reliable estimation of podocyte density in disease requires sampling more glomeruli per mouse.
Table S14. Distribution of glomerulus profiles across the thin needle biopsy data set biopsy ($n = 45$ patients).
Supplementary Video (MOV)
Movie S1. PodoCount in the Cloud video.

REFERENCES

1. Carney EF. The impact of chronic kidney disease on global health. *Nat Rev Nephrol*. 2020;16:251–251. https://doi.org/10.1038/s41581-020-0268-7

2. Alicic RZ, Rooney MT, Tuttle KR. Diabetic kidney disease: challenges, progress, and possibilities. *Clin J Am Soc Nephrol*. 2017;12:2032–2045. https://doi.org/10.2215/CJN.11491116

3. Puelles VG, Bertram JF, Moeller MJ. Quantifying podocyte depletion: theoretical and practical considerations. *Cell Tissue Res*. 2017;369:229–236. https://doi.org/10.1007/s00441-017-2630-z

4. Puelles VG, Bertram JF. Counting glomeruli and podocytes: rationale and methodologies. *Curr Opin Nephrol Hypertens*. 2015;24:224. https://doi.org/10.1097/MNH.0000000000000121

5. Kikuchi M, Wickman L, Hodgin JB, Wiggins RC. Podometrics as a potential clinical tool for glomerular disease management. *Semin Nephrol*. 2015;35:245–255. https://doi.org/10.1016/j.semnephrol.2015.04.004

6. Dai H, Liu Q, Liu B. Research progress on mechanism of podocyte depletion in diabetic nephropathy. *J Diabetes Res*. 2017;2017:2615286. https://doi.org/10.1155/2017/2615286

7. Armelloni S, Mattinzoli D, Ikehata M, et al. Urinary mRNA expression of glomerular podocyte markers in glomerular disease and renal transplant. *Diagnostics*. 2021;11:1499. https://doi.org/10.3390/diagnostics11081499

8. Naik AS, Aqeel J, Wang SQ, Chowdhury M, He K, Wiggins RC. Urine marker analysis identifies evidence for persistent glomerular podocyte injury across allograft lifespan. *Clin Transplant*. 2021;35:e14457. https://doi.org/10.1111/ctr.14457

9. Minakawa A, Fukuda A, Sato Y, et al. Podocyte hypertrophic stress and detachment precedes hyperglycemia or albuminuria in a rat model of obesity and type2 diabetes-associated nephropathy. *Sci Rep*. 2019;9:18485. https://doi.org/10.1038/s41598-019-54692-z

10. Shankland S. The podocyte's response to injury: role in proteinuria and glomerulosclerosis. *Kidney Int*. 2006;69:2131–2147. https://doi.org/10.1038/sj.ki.5000410

11. Yu D, Petermann A, Kunter U, Rong S, Shankland SJ, Floege J. Urinary podocyte loss is a more specific marker of ongoing glomerular damage than proteinuria. *J Am Soc Nephrol*. 2005;16:1733–1741. https://doi.org/10.1681/ASN.2005020159

12. Lemley KV, Bertram JF, Nicholas SB, White K. Estimation of glomerular podocyte number: a selection of valid methods. *J Am Soc Nephrol*. 2013;24:1193–1202. https://doi.org/10.1681/ASN.2012111078

13. Nicholas SB, Basgen JM, Sinha S. Using stereologic techniques for podocyte counting in the mouse: shifting the paradigm. *Am J Nephrol*. 2011;33:Suppl 1(suppl 1):1–7. https://doi.org/10.1159/000327564

14. Puelles VG, Douglas-Denton RN, Cullen-McEwen L, et al. Design-based stereological methods for estimating numbers of glomerular podocytes. *Ann Anat Anat Anz*. 2014;196:48–56. https://doi.org/10.1016/j.aanat.2013.04.007

15. Puelles VG, van der Wolde JW, Schulze KE, et al. Validation of a three-dimensional method for counting and sizing podocytes in whole glomeruli. *J Am Soc Nephrol*. 2016;27:3093–3104. https://doi.org/10.1681/ASN.2015121340

16. Venkatareddy M, Wang S, Yang Y, et al. Estimating podocyte number and density using a single histologic section. *J Am Soc Nephrol*. 2014;25:1118–1129. https://doi.org/10.1681/ASN.2013080859

17. Wang B, Charukeshi Chandrasekera P, Pippin J. Leptin-and leptin receptor-deficient rodent models: relevance for human type 2 diabetes. *Curr Diabetes Rev*. 2014;10:131–145. https://doi.org/10.2174/1573399810666140508121012

18. Tomino Y. Lessons from the KK-Ay mouse, a spontaneous animal model for the treatment of human type 2 diabetic nephropathy. *Nephrourol Mon*. 2012;4:524–529. https://doi.org/10.5812/numonthly.1954

19. Basting T, Lazartigues E. DOCA-salt hypertension: an update. *Curr Hypertens Rep*. 2017;19:32. https://doi.org/10.1007/s11906-017-0731-4

20. Rosenberg AZ, Roshanravan H, Delsante M, et al. An APOL1-induced FSGS mouse model that mimics human FSGS nephropathy [abstract]. *J Am Soc Nephrol*. 2018;29:48.

21. Palliyaguru DL, Teixeira CVL, Duregon E, et al. Study of longitudinal aging in mice: presentation of experimental techniques. *J Gerontol A Biol Sci Med Sci*. 2021;76:552–560. https://doi.org/10.1093/gerona/glaa285

22. Weeda G, Donker I, de Wit J, et al. Disruption of mouse ERCC1 results in a novel repair syndrome with growth failure, nuclear abnormalities and senescence. *Curr Biol*. 1997;7:427–439. https://doi.org/10.1016/s0960-9822(06)00190-4

23. Gurkar AU, Niedernhofer LJ. Comparison of mice with accelerated aging caused by distinct mechanisms. *Exp Gerontol*. 2015;68:43–50. https://doi.org/10.1016/j.exger.2015.01.045

24. Ahmad A, Robinson AR, Duensing A, et al. ERCC1-XPF endonuclease facilitates DNA double-strand break repair. *Mol Cell Biol*. 2008;28:5082–5092. https://doi.org/10.1128/MCB.00293-08

25. Tervaert TWC, Mooyaart AL, Amann K, et al. Pathologic classification of diabetic nephropathy. *CJASN*. 2010;21:556–563. https://doi.org/10.1681/ASN.2010010010

26. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate [published correction appears in *Ann Intern Med*. 2011;155:408]. *Ann Intern Med*. 2009;150:604–612. https://doi.org/10.7326/0003-4819-150-9-200905050-00006

27. Nagata M, Shibata S, Shigeta M, Yu-Ming S, Watanabe T. Cyclin-dependent kinase inhibitors: p27Kip1 and p57kip2 expression during human podocyte differentiation. *Nephrol Dial Transplant*. 1999;14(suppl 1):48–51. https://doi.org/10.1093/ndt/14.suppl_1.48

28. Gonzales RC, Wintz P. *Digital Image Processing (Book)*. Addison-Wesley Publishing Co, Inc; 1977.

29. Lutnick B, Ginley B, Govind D, et al. An integrated iterative annotation technique for easing neural network training in medical image analysis. *Nat Mach Intell*. 2019;1:112–119. https://doi.org/10.1038/s42256-019-0018-3

30. Ruifrok AC, Johnston DA. Quantification of histochemical staining by color deconvolution. *Anal Quant Cytol Histol*. 2001;23:291–299.

31. Parvati K, Rao P, Mariya Das M. Image segmentation using gray-scale morphology and marker-controlled watershed transformation. *Discrete Dyn Nat Soc*. 2008;2008.

32. Veta M, Huisman A, Viergever MA, van Diest PJ, Pluim JP. Marker-controlled watershed segmentation of nuclei in H&E stained breast cancer biopsy images. Paper presented at: 2011 IEEE international symposium on biomedical imaging: from nano to macro 2011.

33. Ginley B, Lutnick B, Jen K-Y, et al. Computational segmentation and classification of diabetic glomerulosclerosis. *J Am Soc Nephrol*. 2019;30:1953–1967. https://doi.org/10.1681/ASN.2018121259

34. Van der Walt S, Schönberger JL, Nunez-Iglesias J, et al. scikit-image: image processing in Python. *PeerJ*. 2014;2:e453. https://doi.org/10.7717/peerj.453

35. Nelson LS. The Anderson-Darling test for normality. *J Qual Technol*. 1998;30:298.

36. Benesty J, Chen J, Huang Y, Cohen I. Pearson correlation coefficient. In: *Noise Reduction in Speech Processing*. Springer; 2009:1–4.

37. Wiley RW, Rapp B. Statistical analysis in Small-N Designs: using linear mixed-effects modeling for evaluating intervention effectiveness. *Aphasiology*. 2019;33:1–30. https://doi.org/10.1080/02687038.2018.1454884

38. Miller Jr RG. *Beyond ANOVA: Basics of Applied Statistics*. CRC Press; 1997.

39. Stoline MR. The status of multiple comparisons: simultaneous estimation of all pairwise comparisons in one-way ANOVA designs. *Am Stat*. 1981;35:134–141.

40. Hoefield RA, Kalra PA, Baker PG, et al. The use of eGFR and ACR to predict decline in renal function in people with diabetes. *Nephrol Dial Transplant*. 2011;26:887–892. https://doi.org/10.1093/ndt/gfq526

41. Thissen D, Steinberg L, Kuang D, Kuang D. Quick and easy implementation of the Benjamini–Hochberg procedure for controlling the false positive rate in multiple comparisons. *J Educ Behav Stat*. 2002;27:77–83.

42. Gutman DA, Khalilia M, Lee S, et al. The digital slide archive: a software platform for management, integration, and analysis of histology for cancer research. *Cancer Res*. 2017;77:e75–e78. https://doi.org/10.1158/0008-5472.CAN-17-0629

43. Docker AC. [software engineering]. *IEEE Softw*. 2015;32:102–c103.

44. Boettiger C. An introduction to Docker for reproducible research. *ACM SIGOPS Oper Syst Rev*. 2015;49:71–79.

45. Docker MD. lightweight Linux containers for consistent development and deployment. *Linux J*. 2014;2014(239):2.

46. The Sarder Lab GitHub Repository. Updated 2021. Accessed March 30, 2022. https://github.com/SarderLab

47. Hughson MD, Hoy WE, Douglas-Denton RN, Zimanyi MA, Bertram JF. Towards a definition of glomerulomegaly: clinical–pathological and methodological considerations.

Nephrol Dial Transplant. 2011;26:2202–2208. https://doi.org/10.1093/ndt/gfq688

48. The Sarder lab whole slide image archiver and viewer. Updated 2021. Accessed March 30, 2022. http://ptyhisto.med.buffalo.edu:8080/

49. Goode A, Gilbert B, Harkes J, Jukic D, Satyanarayanan M. OpenSlide: a vendor-neutral software foundation for digital pathology. *J Pathol Inform*. 2013;4:27. https://doi.org/10.4103/2153-3539.119005

50. Govind D, Becker J, Miecznikowski J, et al. PodoSighter: a cloud-based tool for label-free podocyte detection in kidney whole slide images. *J Am Soc Nephrol*. 2021;32:2795–2813. https://doi.org/10.1681/ASN.2021050630

51. Zimmermann M, Klaus M, Wong MN, et al. Deep learning–based molecular morphometrics for kidney biopsies. *JCI Insight*. 2021;6:e144779. https://doi.org/10.1172/jci.insight.144779

52. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology—new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol*. 2019;16:703–715. https://doi.org/10.1038/s41571-019-0252-y

53. Holzinger A, Malle B, Kieseberg P, et al. Towards the augmented pathologist: challenges of explainable-ai in digital pathology. arXiv. Published December 18, 2017. Accessed March 30, 2022. https://arxiv.org/abs/1712.06657

54. Madabhushi A. Digital pathology image analysis: opportunities and challenges. *Imaging Med*. 2009;1:7–10. https://doi.org/10.2217/IIM.09.9

55. Santo BA, Rosenberg AZ, Sarder P. Artificial intelligence driven next-generation renal histomorphometry. *Curr Opin Nephrol Hypertens*. 2020;29:265–272. https://doi.org/10.1097/MNH.0000000000000598

56. Janowczyk A, Zuo R, Gilmore H, Feldman M, Madabhushi A. HistoQC: an open-source quality control tool for digital pathology slides. *JCO Clin Cancer Inform*. 2019;3:1–7. https://doi.org/10.1200/CCI.18.00157

57. Bruggeman LA, Wu Z, Luo L, et al. APOL1-G0 protects podocytes in a mouse model of HIV-associated nephropathy.

PLoS One. 2019;14:e0224408. https://doi.org/10.1371/journal.pone.0224408

58. Kaverina NV, Eng DG, Miner JH, Pippin JW, Shankland SJ. Parietal epithelial cell differentiation to a podocyte fate in the aged mouse kidney. *Aging (Albany NY)*. 2020;12:17601–17624. https://doi.org/10.18632/aging.103788

59. Schneider RR, Eng DG, Kutz JN, Sweetwyne MT, Pippin JW, Shankland SJ. Compound effects of aging and experimental FSGS on glomerular epithelial cells. *Aging (Albany NY)*. 2017;9:524–546. https://doi.org/10.18632/aging.101176

60. Pagtalunan ME, Miller PL, Jumping-Eagle S, et al. Podocyte loss and progressive glomerular injury in type II diabetes. *J Clin Invest*. 1997;99:342–348. https://doi.org/10.1172/JCI119163

61. Alpers CE, Hudkins KL. Mouse models of diabetic nephropathy. *Curr Opin Nephrol Hypertens*. 2011;20:278. https://doi.org/10.1097/MNH.0b013e3283451901

62. Diani AR, Sawada GA, Zhang NY, et al. The KKAy mouse: a model for the rapid development of glomerular capillary basement membrane thickening. *Blood Vessels*. 1987;24:297–303. https://doi.org/10.1159/000158706

63. Nishizono R, Kikuchi M, Wang SQ, et al. FSGS as an adaptive response to growth-induced podocyte stress. *J Am Soc Nephrol*. 2017;28:2931–2945. https://doi.org/10.1681/ASN.2017020174

64. Rosenberg AZ, Kopp JB. Focal segmental glomerulosclerosis [published correction appears in *Clin J Am Soc Nephrol*. 2018;13:1889]. *Clin J Am Soc Nephrol*. 2017;12:502–517. https://doi.org/10.2215/CJN.05960616

65. Wyatt CM, Klotman PE, D'Agati VD. HIV-associated nephropathy: clinical presentation, pathology, and epidemiology in the era of antiretroviral therapy. *Semin Nephrol*. 2008;28(6):513–522. https://doi.org/10.1016/j.semnephrol.2008.08.005

66. Medapalli RK, He JC, Klotman PE. HIV-associated nephropathy: pathogenesis. *Curr Opin Nephrol Hypertens*. 2011;20:306–311. https://doi.org/10.1097/MNH.0b013e328345359a