Progression of chronic kidney disease at the tissue level is the result of accumulating extracellular matrix. This progression may be assessed by changes in GFR, but estimated GFR (eGFR) from serum creatinine is relatively inaccurate. Accurate assessment of parenchymal fibrosis is one key variable for evaluating chronic injury and response to intervention and perhaps should be considered the gold standard to assess progression. However, definitions and methods to assess fibrosis accurately are lacking and have been surprisingly difficult to achieve. Early sea-faring explorers faced similar challenges because the lack of accurate measurements of longitude greatly hampered navigation. Only when relative measurements of time at sea were made possible in the 1700s by development of a reliable chronometer by the self-taught watchmaker John Harrison could this obstacle be overcome. Similar challenges face us in navigating the obstacles in assessing tissue fibrosis.

In this issue of *JASN*, the studies by Farris et al. tackle some of the important obstacles to accurate assessment of fibrosis. First, they compare various methods of determining fibrosis in the same tissue sample within one pathology group using morphometric methods, visual scoring, and various stains. These studies show excellent interassay and interobserver reproducibility by visual and morphometric techniques, although morphometric techniques showed slightly less variation between observers. Fibrosis, whether assessed by morphometry of collagen type III immunohistochemistry or unpolarized Sirius red–stained sections or by visual scoring of trichrome-stained slides, correlated with eGFR. Compared with previous studies, the agreements indeed are impressive. When Furness et al. examined interobserver agreement of fibrosis in renal biopsies among a large group of pathologists, $\kappa$ values were poor (0.306). In that study, there were no didactic training sessions or specific detailed definitions, but graphical performance feedback did not improve the reproducibility; it worsened actually. Other studies have found more robust interobserver correlations for fibrosis assessment when precise definitions were offered, followed by head-to-head microscope sessions to refine definitions. This latter approach was used by the IgA Nephropathy Oxford classification pathologists and resulted in reasonable reproducibility of broad assessments of fibrosis. However, correlation with outcome was limited to estimations of whether fibrosis involved more or less than 25 or 50% of renal parenchyma. Finer precision assessing fibrosis linked to outcome could not be demonstrated. In the International Fabry Working Group assessment of fibrosis, there was also reasonably good interobserver reproducibility, and the degree of fibrosis assessed visually also correlated highly with morphometry.

None of these previous efforts addressed or discussed key issues of the distribution and quality of fibrosis and how to differentiate between various types of fibrotic processes. Thus, the percentage of cortex that is abnormal or the percentage of fibrotic tissue within the biopsy represent quite different parameters. Similarly, early fine diffuse fibrosis may activate different pathophysiologic mechanisms than localized areas of dense mature fibrous tissue. The constituents of fibrosis may also vary in different settings. Edema interlaced with fine

---

**Navigating the Challenges of Fibrosis Assessment: Land in Sight?**

Agnes B. Fogo* and Charles E. Alpers†

*Department of Pathology, Vanderbilt University School of Medicine, Nashville, Tennessee; and †Department of Pathology, University of Washington School of Medicine, Seattle, Washington

strands of collagen and inflammatory cells differs fundamentally from mature dense fibrous tissue with minimal inflammation, suggesting a need for differing therapeutic interventions.

The methods for assessment also have inherent limitations and challenges. The studies by Farris et al.\textsuperscript{5} are a particularly noteworthy advance in the field by addressing several of these issues to a degree not previously accomplished in the literature. Exclusion of normal connective tissue is a technical challenge for assessing fibrosis. The novel morphometric method used by Farris et al. may be useful in distinguishing fibrosis from normal connective tissue, combining trichrome and periodic acid Schiff staining and deriving a differential staining area morphometrically.

Interestingly, the studies by Farris et al.\textsuperscript{5} did not demonstrate greater precision of the polarized Sirius red stain (often presumed to be the most accurate measure of fibrosis available) or better correlation with eGFR, perhaps related in part to shorter staining times and more variability. Most surprising, in a head-to-head test of visual assessment of fibrosis by pathologists armed with good definitions and a reliable trichrome stain versus computerized techniques, the “eyes” proved equally advantageous in terms of reproducibility and correlation with eGFR. Furthermore, visual assessment is relatively cheap with only modest cost beyond the time of the pathologists doing the assessment and is certainly faster than morphometric analysis.

Thus, the studies by Farris et al.\textsuperscript{5} define many issues, suggest some answers, but clearly require further substantial validation involving much larger sample sizes. This is particularly true for the surprising result that visual inspection of scanned images of trichrome-stained tissues can provide such a reliable measure of fibrosis, when this reliability has been difficult to achieve by others using trichrome stains, albeit with less precise definitions of fibrosis and less controlled conditions.

Some remaining issues include staining variability, tissue sampling, and limitations of eGFR as a clinical standard. All stains in these study were done in the same laboratory, so the issue of variability between laboratories could not be addressed. The trichrome stain, which stains all forms of collagen, may be insensitive for detection of very early forms of fibrosis. The length of tissue fixation may also dramatically influence the staining: this variable is particularly challenging for clinical biopsies that may come to the pathology laboratory from multiple sites and distances. The picro Sirius red stain, regardless of whether polarized, lends itself easily to computerized image analysis and thus theoretically should be useful for more precise quantification, although the staining process is more time-consuming than for the trichrome stain. Of note, shorter staining time leads to increased variability, which may be overcome with overnight (or 8 hours) staining (Paul Grimm, personal communication). Polarization must also be done under rigorously controlled conditions for optimal accuracy.

The size of the tissue sample also greatly influences sampling assessment of fibrosis. Even the best technical measures will fall short if applied to a sample that is not representative of the entire kidney parenchyma. Future studies should therefore also assess variability of measures between cores in a biopsy and establish minimum thresholds for optimal reproducible assessments, an approach used by the Cooperative Clinical Trials in Transplantation group for establishing optimal sample size for diagnosis and threshold of acute rejection.\textsuperscript{9} Future studies will also need to address such issues as whether morphologic techniques can reliably distinguish types/stages of fibrosing injury, whether measures of integrity of interstitial vasculature are also essential to understanding the severity of any given fibrotic injury, and whether the specific composition of the fibrotic parenchyma is a critical prognostic feature because the studies Farris et al.\textsuperscript{5} rest largely on an untested presumption that accumulations of collagen and fibrosis are equivalent processes.

Additional studies, such as those under way by the Banff Working Group of Fibrosis,\textsuperscript{10} will be necessary for further accurate and dynamic assessment of both function and structural injury. Proteomic, transcriptomic, or mass spectrometry methods may also offer novel accurate measures of fibrosis and related mechanisms. The studies by Farris et al.\textsuperscript{5} are an important step in the navigation toward accurate assessment of fibrosis.

DISCLOSURES
None.

REFERENCES


