Digital Pathology and Image Analysis for Robust High-Throughput Quantitative Assessment of Alzheimer Disease Neuropathologic Changes

Janna Hackett Neltner, MD, Erin Lynn Abner, MPH, Frederick A. Schmitt, PhD, Stephanie Kay Denison, BS, Sonya Anderson, BBA, Ela Patel, and Peter T. Nelson, MD, PhD

Abstract
Quantitative neuropathologic methods provide information that is important for both research and clinical applications. The technologic advancement of digital pathology and image analysis offers new solutions to enable valid quantification of pathologic severity that is reproducible between raters regardless of experience. Using an Aperio ScanScope XT and its accompanying image analysis software, we designed algorithms for quantitation of amyloid and tau pathologies on 65 β-amyloid (6F/3D antibody) and 48 phospho-tau (PHF-1)–immunostained sections of human temporal neocortex. Quantitative digital pathologic data were compared with manual pathology counts. There were excellent correlations between manually counted and digitally analyzed neuropathologic parameters ($R^2 = 0.56–0.72$). Data were highly reproducible among 3 participants with varying degrees of expertise in neuropathology (intraclass correlation coefficient values, $>0.910$). Digital quantification also provided additional parameters, including average plaque area, which shows statistically significant differences when samples are stratified according to apolipoprotein E allele status (average plaque area, 380.9 μm$^2$ in apolipoprotein E ε4 carriers vs 274.4 μm$^2$ for noncarriers; $p < 0.001$). Thus, digital pathology offers a rigorous and reproducible method for quantifying Alzheimer disease neuropathologic changes and may provide additional insights into morphologic characteristics that were previously more challenging to assess because of technical limitations.

Key Words: Alzheimer disease, Autopsy, Digital pathology, Image analysis, Neuropathology.

INTRODUCTION
The association between the severity of Alzheimer disease neurologic changes (ADNCs) and cognitive impairment has long been established (1–19). The burden of neurofibrillary tangles (NFTs) and neuritic β-amyloid (Aβ) plaques (NPs), as well as their distribution throughout the cortex, are the bases for the Braak and Consortium to Establish a Registry for Alzheimer’s Disease (CERAD) staging schemes, respectively (1, 20, 21). Previous studies from the University of Kentucky Alzheimer’s Disease Center (UK-ADC) have demonstrated the use of quantitative assessment of ADNCs. Manual quantitation of AD pathologies has been performed at the UK-ADC for more than 20 years. The quantitation of ADNCs helped demonstrate the association of the different lesion type densities with cognitive impairment across the spectrum of disease from preclinical AD (11, 22) and mild cognitive impairment (5) to end-stage disease (23). Using the quantitative manual counts, the severity of ADNCs can also be correlated with possible risk factors such as diabetes (24) and to related diseases such as tangle-only pathology lacking amyloid plaques (25); they can also highlight differences in the anatomic distribution of ADNCs (26). Furthermore, having quantitative neuropathologic assessment enabled us to develop a model of how each pathologic subtype contributes to cognitive impairment (27). In these studies, we previously found evidence of a “leveling off” of amyloid plaques in late-stage disease (18), but our confidence in these results was limited because of the technical drawbacks of the manual counting method (i.e. with caps on amyloid plaque counts, as described [23, 25, 28]). Work from other centers has also showed the benefits of quantitative assessment of AD neuropathologic changes (29–33). A common thread in these published studies is that the benefits of quantitative pathology can best be obtained if they are used in combination with detailed data about the patients’ non-AD pathologic findings, medical histories, and cognitive status.

Although insights can be gained by quantitative assessment of ADNCs, manual methods as performed at UK-ADC are painstaking efforts that suffer from certain drawbacks. This process, which involves manual inspection and quantitation of $8$ different areas of the brain processed via the modified Bielschowsky silver impregnation protocol, is extremely time consuming. In addition, despite our long-standing expertise in this area, there is still suboptimal inter-rater reliability, which limits the ability to generalize our findings to those of other institutions. A final drawback includes the practice of limiting the counted area to $5$ microscopic fields and implementing a cap value to plaques. Although this is necessary to allow manual counting to be manageable, it introduces a source of error to our quantitation results.
The advent of digital slide scanners and software analysis packages is revolutionizing the practice of pathology and the research questions that can be addressed using pathologic material. Digital slide scanners create digitized images of glass slides that can be viewed from a computer at varying magnifications and stored indefinitely. Such files can then be used for high-throughput sophisticated image analyses, which offer quantitation capabilities that far exceed those that can be achieved manually in terms of reproducibility, unbiased image recognition algorithms, and sheer volume. Previous studies have demonstrated some of the advantages of digital image analysis within AD research either through montaged still images (34) or a variety of digital slide systems (35, 36). One such software package is Aperio Genie Histologic Pattern Recognition Tool (37). The Genie software has the ability to “learn” objects of our specification, in this case, NFTs and NPs, based on a set of representative slides. Once this “teaching” is complete, the software has the ability to identify NFTs and NPs on any phospho-tau (PHF-1)–stained section in a very reproducible manner. This pattern recognition software has been successfully used in a variety of pathologic conditions, ranging from kidney failure to breast carcinoma (38, 39).

We took advantage of the strengths of digital pathology to update our quantitation protocol for ADNCs. In addition to parameters similar to those that we also obtain manually, we anticipated finding additional changes within plaque morphology that were impossible to quantify rigorously by manual methods. Additional data provided routinely by the algorithms include average plaque area and varying staining intensities. Such variables applied over many cases may reveal new insights into the pathologic expression of the disease that may never have been possible through manual methods alone.

**MATERIALS AND METHODS**

**Case Selection and Scanning**

Slides and paraffin-embedded tissue from the superior and middle temporal gyri were analyzed from which manual quantitation of diffuse Aβ plaques (DPs), NPs, and NFTs had already been performed on modified Bielschowsky impregnations of near-serial sections via methods previously described in detail (23, 25, 28). Briefly, paraffin-embedded tissue was processed, and 8-µm-thick sections were cut and processed using a modified Bielschowsky method (34). Diffuse Aβ plaques were counted through a 10 objective (field size, 2.35 mm²) in the 5 most involved fields, with an arbitrary “cap” at 250 plaques over all 5 fields. Neuritic Aβ plaques were also counted in a similar manner but without a cap. Neurofibrillary tangles were counted in the 5 most severely affected fields through a 20 objective (field size, 0.586 mm²). An arithmetic mean was then calculated for each parameter, resulting in a DP, NP, and NFT count per field.

For digital amyloid quantitation, a convenience sample of 65 cases was chosen based on the manual DP counts. Cases were selected with the goal of ensuring an adequate distribution of DP counts, ranging from 0 and 50 DP per field, that is, not “anchored” in very high or very low DP counts. For each case, Aβ immunohistochemical (IHC) staining had already been performed via previously described methods (28) with a monoclonal NCL-Aβ 6F/3D antibody (Novocastra, Newcastle, UK).

---

**FIGURE 1.** Analytic workflow for high-throughput quantitation of Alzheimer disease–type neuropathologic changes (A), with a more detailed explanation of the image analysis steps (B).
Further analyses of amyloid plaque size distributions were performed on those cases with known apolipoprotein E (APOE) status (APOE e4, n = 19; APOE non-e4, n = 41), as well as a subset of these with a final Mini-Mental State Examination (MMSE) score of 27 to 30 (APOE e4, n = 6; APOE non-e4, n = 24).

For digital quantification of tau (neurofibrillary) pathologies, a convenience sample of 48 cases were chosen based on the Braak stage to ensure a full range of tau levels (Braak 0, 4 cases; Braak I, 5 cases; Braak II, 4 cases; Braak III, 3 cases; Braak IV, 4 cases; Braak V, 14 cases; and Braak VI, 14 cases). The PHF-1 (a kind gift from Dr Peter Davies, Bronx, NY) IHC was performed on the superior and middle temporal gyri. Additional analyses to compare digital amyloid and tau pathologies were performed on a subset of the cases in which both PHF-1 and Aβ IHC stains were performed. Cases with confounding pathologies (e.g. neocortical Lewy bodies, hippocampal sclerosis, vascular disease, frontotemporal dementias, and progressive supranuclear palsy) were excluded from this analysis, leaving a total of 24 cases across the full spectrum of ADNC severity. Slides were then loaded into an Aperio ScanScope XT and scanned at 40 magnification (0.25 μm/pixel) via the semi-automated method and then stored on a dedicated server. Slides were checked for image quality using an Aperio “quality factor greater than 90” and visual inspection.

**Analysis Region Selection**

Whole slide analysis was impractical because it was overly time consuming and severely biased by differential distribution of gray/white matter in different cases. To enable a more focused analysis that could be reproducible between users, a square analysis region (4 mm²) was created. For each case, that box was placed within the gray matter at the site of highest concentration of pathologic findings. Subsequent boxes were then placed as far from the existing boxes as possible without overlapping other analysis regions. Areas of poor stain quality were avoided. To minimize analysis alterations caused by portions of folded tissue or foreign material, these were eliminated from the analysis boxes with the negative pen tool.

In the amyloid analyses, meningeal vessels were also excluded by the negative pen tool, so as to limit the burden and plaque analyses to parenchymal content. Parenchymal vessels that were involved by amyloid angiopathy were included in the burden analysis; these vessels were not picked up by the plaque density algorithm.

**Analysis Workflow**

Once the analysis regions were selected, the next step was to identify the appropriate quantitation algorithms to use. Cases in which there was no need to subtype pathologies within a given immunostain, for example, with the Aβ IHC stain, the pattern recognition software adds an unnecessary and time-consuming step. Therefore, analysis of either amyloid plaque density or overall amyloid burden could be performed by simple modification of the preexisting algorithm templates. In contrast, within a single PHF-1 immunostain, we wanted to quantitate NFTs and NPs separately. The pattern recognition software allows each of these structures to be highlighted in isolation. Once identified, they could then be quantitated by a variety of mechanisms. A representative schematic of the digital workflow is in Figure 1.

**Aβ Quantitation**

Two different parameters were calculated for amyloid pathologies: an overall amyloid burden and an amyloid plaque density. For the overall amyloid burden, the Aperio Image Analysis Toolbox Positive Pixel Count (PPC), version 9.1, was used. Modifications were made to the input parameters of the algorithm to optimize our staining protocol based on comparison of sample analysis markups with manual counting results by visual inspection in multiple areas on 3 representative slides (Table, Supplemental Digital Content 1, http://links.lww.com/NEN/A398). The modified PPC was then run on the 10 4-mm² boxes selected on each slide as previously described. The amyloid burden was calculated by adding up all of the weak (1+, yellow), positive (2+, orange), and strong (3+, red) pixels from the data and dividing by the overall analysis area (40 mm²). For the amyloid plaque density, the Aperio Image Analysis Toolbox Nuclear algorithm, version 9.1, was used with subsequent modifications, derived in a similar manner to the amyloid burden (Table, Supplemental Digital

---

**FIGURE 2.** Determination of appropriate analysis area. (A) A 4-mm² annotation was initially placed in the gray matter region with the highest density of pathology (*). Subsequent regions were then randomly selected, up to 30 regions. The plaque density was then calculated for each individual box and then sequentially averaged together in a stepwise fashion. (B) The average density approached an asymptote by 10 boxes in all test cases studied. Scale bar = 2 mm.
Content 2, http://links.lww.com/NEN/A399). This modified algorithm was then run on the same 10 boxes chosen for the amyloid burden analysis. The amyloid plaque density was calculated by adding up all the weakly (1+, yellow), moderately (2+, orange), and strongly (3+, red) positive “nuclei” and dividing by the overall analysis area (40 mm²).

**Tau Quantitation**

Three different parameters were quantified for tau pathologies: NFT density, NP burden, and overall tau burden. For the NFT and NP quantitation, a single Genie algorithm was first developed to separate NFTs and NPs from the background. Approximately 10 to 15 representative samples for each of the 3 classes (NFTs, NPs, and background) were highlighted using a digital pen tool. In selecting sample structure, care was taken to make sure that the range of possible morphologies was covered for any given pathology (e.g. making sure that multiple areas with white matter, vasculature, or meningeal tissue were included for the “background” training). Then the Genie training algorithm, consisting of 2,000 iterations, was run with the apparent 5 magnification. The algorithm was then evaluated by visual inspection in an unmarked area to assess its accuracy. This process was repeated until the algorithm or “classifier” could identify both structures to our satisfaction. The NFT/NP Genie algorithm was then used in subsequent analyses for quantification.

The NFT density was calculated via a modified nuclear algorithm, version 9.1, with our NFT/NP Genie classifier limited to NFTs. An input parameter that was important to address was the minimum nuclear cutoff size (in square micrometers). To find the nuclear size that would both maximize the number of NFTs and minimize the number of non-NFT particles counted (i.e. oligodendroglia and glial inclusions), the algorithm was run multiple times on a subset of 10 cases using a series of “minimum nuclear size” cutoffs ranging from 35 to 100 μm². The NFT density for each cutoff was calculated and then compared with both the manual counts and the final MMSE scores. Using this method, it was found that a minimum nuclear size of 40 μm² provided the most accurate result. The algorithm input parameters are listed in Table, Supplemental Digital Content 3, http://links.lww.com/NEN/A400. The modified nuclear algorithm was run on 10 4-mm² boxes selected by the method previously stated within the gray matter in the PHF-1-immunostained sections. The NFT density was calculated by dividing the total number of nuclei counted by the overall analysis area (40 mm²).

For the NP burden, the PPC, version 9.1, was used with our NFT/NP Genie classifier limited to NPs and modified for our staining protocol via visual inspection. The algorithm input parameters are listed in Table, Supplemental Digital Content 4, http://links.lww.com/NEN/A401. The modified PPC was then run on the same boxes selected for NFT analysis. The NP burden was calculated by adding up all the weak (1+, yellow), moderately (2+, orange), and strongly (3+, red) positive “nuclei” and dividing by the overall analysis area (40 mm²).

**FIGURE 3.** Amyloid quantification. (A) The Aβ immunostain was performed on the superior and middle temporal gyri (SMTG) of all cases. (B) Digital analysis of amyloid plaque density in the number of diffuse Aβ plaques (DP number) per square millimeter via a modified nuclear algorithm. In addition to quantitation of the amyloid plaques, the digital analysis also separated the plaques by varying degrees of intensity (0, blue; 1+, yellow; 2+, orange; 3+, red). Scale bar = 50 μm.
positive (2+, orange), and strong (3+, red) pixels from the data and dividing by the overall analysis area (40 mm²).

The overall tau burden was determined using a PPC algorithm, version 9.1, similar to the amyloid burden quantitation but with stain-specific modifications based on visual inspection. The algorithm input parameters are listed in Table, Supplemental Digital Content 5, http://links.lww.com/NEN/A402. The modified PPC was then run on the same 10 4-mm² boxes selected for the other tau analyses. The tau burden was calculated by dividing the “strong” (3+, red) pixels by the overall analysis area (40 mm²).

**Interobserver Variability**

To test the reproducibility of the algorithms, subsets of 30 Aβ- and 30 PHF-1–stained sections were given to 2 additional individuals with varying degrees of neuropathology expertise, that is, the brain bank coordinator involved in the manual counting of these pathologies and an undergraduate student with no neuropathology experience. Each individual

**TABLE 1.** Ordinary Least-Squares Regression $R^2$ Correlation Coefficients for Comparisons Between Manual Counts and Digital Parameters

<table>
<thead>
<tr>
<th>Digital Amyloid (n = 65)</th>
<th>Correlation With Manual DPs $R^2$ (p) Values</th>
<th>Correlation With Manual DPs + NPs $R^2$ (p) Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amyloid burden</td>
<td>0.62 (&lt;0.001)</td>
<td>0.60 (&lt;0.001)</td>
</tr>
<tr>
<td>Amyloid a</td>
<td>0.72 (0.001)</td>
<td>0.67 (0.09)</td>
</tr>
<tr>
<td>Digital tau (n = 48)</td>
<td>Manual NFTs $R^2$ values</td>
<td>Manual NPs $R^2$ values</td>
</tr>
<tr>
<td>NFT density</td>
<td>0.57 (0.04)</td>
<td>—</td>
</tr>
<tr>
<td>Tau burden</td>
<td>0.56 (NA)</td>
<td>0.27 (NA)</td>
</tr>
<tr>
<td>NP burden</td>
<td>—</td>
<td>0.31 (&lt;0.001)</td>
</tr>
</tbody>
</table>

NA, not applicable.

![Figure 4](http://jnen.oxfordjournals.org/)

**FIGURE 4.** Amyloid quantification correlations (n = 65). (A) There was good correlation between the digital quantification via digital amyloid plaque density and the manually derived DP counts. (B) Inclusion of the manually counted NPs, which were also picked up by the amyloid immunohistochemistry, also showed excellent correlation. (C) Correlation of the digital plaque density to digital amyloid burden.
was given the same set of digital slides, along with instructions for analyses as previously described, with no preexisting analysis boxes selected. This method was only performed a single time. All 5 digital analyses were then run on these selected regions and compared with the original data received from the fields selected by a neuropathologist.

**Statistical Analysis**

Ordinary least-squares regression was conducted to assess the relationship between manual and digital counts. Group comparisons involved paired $t$ tests, with graphed results of amyloid plaque morphology shown as arithmetic mean ± SEM, using MS Excel. Intraclass correlation coefficients were calculated to assess inter-rater reliability for the 3 users. One-way analysis of variance tables (with autopsy case as the factor) were constructed for each measure using PROC GLM in SAS/STAT 9.3, and estimated intraclass correlation coefficients and 95% confidence intervals (95% CIs) were calculated based on the mean squared errors for within and between subjects.

**RESULTS**

After scanning, the next step was to identify the regions for analysis. Although initial analyses were run on the entire slide, it was found to be inefficient, taking hours to identify 1 parameter when the Genie software was used. Limiting the analysis area to a preset region of known size and shape led to far faster analyses. To balance accuracy and efficiency, a series of 30 4-mm$^2$ annotation boxes were placed in the manner previously described (Fig. 2A) on 15 representative Aβ-stained slides. Amyloid quantitation via the modified nuclear algorithm was then performed in each of the square regions. By adding the plaque numbers within each box together in a sequential manner and then dividing by the total area analyzed (e.g. Box 1 plaque number/4 mm$^2$, Box 1 + 2 plaque number/8 mm$^2$, Box 1 + 2 + 3 plaque number/12 mm$^2$, etc), an additive amyloid plaque density was calculated. When these additive densities were graphed with respect to box number, it became clear that the density approached an asymptote (Fig. 2B). In all cases tested, this asymptote was approached by 8 to 10 analysis boxes. To ensure that the box selection procedure was not significantly biased, we also performed the same analyses on a subset of 5 cases, choosing the first square in a region with the least involvement of pathology. In all cases, the analyses approached the same asymptote as was reached in the original selection. A second algorithm, the overall amyloid burden, was also tested and showed similar results. Because of these findings, we used 10 4-mm$^2$ boxes for each slide to quantitate amyloid and tau pathologies. This dropped the analysis time from an average of 1 and 6.5 hours to 5 and 45 minutes for each amyloid and tau algorithm, respectively.

**Amyloid Quantitation**

Both the amyloid burden and amyloid plaque density (Fig. 3) correlated strongly with manual DP counts, with $R^2$ values of 0.62 and 0.72, respectively (Table 1, Fig. 4A). By altering our quantitation protocol from the modified Bielschowsky to amyloid IHC, the ability to distinguish DPs from NPs during amyloid quantitation was lost. Because the amyloid IHC should pick up both plaque types, the digital parameters were correlated to the sum of the manually derived DP and NP counts. This resulted in similar correlations between our methods ($R^2$ values 0.60 and 0.67 for amyloid burden and plaque density, respectively) (Table 1, Fig. 4B). An additional confounding issue involved the artificial capping of manual counts at 50 plaques per field. When a similar cap was used on the digitally derived numbers, purely for the sake of correlation, there was marked improvement of correlation between the amyloid plaque density and manual DP counts, with an $R^2$ value...
FIGURE 6. Digital quantification of tau pathologies. (A) Analyses were performed on PHF-1–immunostained sections. (B) After using the crafted Genie NFT/NP algorithm to isolate the NFTs, the NFT density (NFTs per square millimeter) was determined by a modified nuclear algorithm, with NFTs ranging in staining intensities from 0+ (blue), 1+ (yellow), 2+ (orange), and 3+ (red). (C) In a similar manner, the NP burden was calculated by first using the same crafted Genie algorithm to isolate the NPs and then running a modified positive pixel count to highlight each pixel based on a similar staining intensity as above. (D) An overall tau burden was also calculated (red, positive IHC staining; blue, negative IHC staining). Scale bar = 25 μm.
of 0.82, as well as with the sum of DP and NP manual counts ($R^2 = 0.79$). Finally, the 2 separate digital amyloid parameters were correlated to each other to confirm concordance, which was found to be the case ($R^2 = 0.76$, Fig. 4C).

**Tau Quantitation**

Correlations between tau pathologies were more variable (Table 1). Representative images of the Genie NFT/NP recognition and the subsequent tau algorithms are shown in Figures 5 and 6, respectively. Digital NFT density correlated well with the manual NFT counts ($R^2 = 0.57$, Fig. 7). Whereas there was no official capping policy of tau pathologies, the manual NFTs were never more than 50 NFTs per field, despite the Aperio density often being significantly larger. We hypothesize that a bias caused by counting fatigue likely affected the manual data. For correlation purposes, the Aperio data were capped at 50 NFTs per square millimeter, as was done on the amyloid plaque density. This resulted in a significant improvement in the correlation, with an $R^2$ value of 0.7369. Overall tau burden showed a good correlation with manual NFT counts ($R^2 = 0.56$) and the digitally derived NP burden ($R^2 = 0.73$) and NFT density ($R^2 = 0.86$) but weaker correlation with the manual NP counts ($R^2 = 0.27$). The NP burden showed the weakest correlation with manual counts, with an $R^2$ value of 0.31. Even implementing a cap (as with the NFT and amyloid plaque densities), there was only a mild improvement in correlation ($R^2 = 0.37$) (Fig. 8). By visual inspection, the NP burden algorithm seemed to appropriately highlight NPs. Further visual comparison between the original silver impregnation preparations (on which the manual counts had been obtained) and the new PHF-1 immunostain highlighted the markedly increased sensitivity of the PHF-1, which identified more NPs than were seen with the Bielschowsky method (Fig. 8B, C).

**Interobserver Variation**

The observed agreement among the 3 users was very good for all measures assessed (Table 2). Intraclass correlation coefficients ranged from 0.910 (95% CI, 0.855–0.945) for the NFT density to 0.986 (95% CI, 0.977–0.992) for the tau burden. The $R^2$ correlation coefficients were also calculated via least-squares regression and found to range from 0.93 to 0.98 between the individual users.

**Additional Analyses**

By examining subsets of the above cases, additional comparisons were made using information gathered from the above algorithms. The first subset of cases was limited to those with only ADNCs and had both PHF-1 and Aβ immunostaining performed (n = 24). Digital amyloid plaque densities (as well as the manual sum of DPs and NPs) were examined as a function of increasing tau burden. The results showed the digital amyloid plaque density began decreasing when the tau burden approached 500,000 3+ pixels/mm² (Fig. 9); this trend was not identified through the manual counts because of the capping procedure.

The second additional analysis involved the comparison of amyloid plaque area for cases with known APOE allele status. When separated into 2 cohorts based on the presence (n = 7) or absence (n = 41) of the APOE e4 allele, those cases with at least 1 APOE e4 allele had statistically larger plaque sizes (mean, 380.9 μm²) versus those lacking the APOE e4 allele (mean, 274.4 μm²; p < 0.001) (Fig. 10A). This finding maintained significance even when the analysis was limited to those cases with a final MMSE score of 27 to 30. Those with APOE e4 (n = 5) had an average plaque size of 333.20 μm², whereas those without APOE e4 alleles (n = 25) had a significantly smaller plaque size (233.48 μm²; p = 0.023) (Fig. 10B).

**DISCUSSION**

The goals of this study were to establish a set of computer algorithms to replace manual quantitation at our center and to highlight the possible benefits that digital pathology may provide over the traditional counting methods at other research centers. The results indicate that amyloid plaque parameters, NFT density, and overall tau burden correlate strongly with the manual count data that have been used for years and in dozens of published studies at the UK-ADC. We also were able to elucidate new features of plaque morphology that would be practically impossible to evaluate using most other methods.

Digital pathology offers multiple benefits that surpass both semiquantitative methods and manual counts. Digital algorithms offer superior reproducibility and higher throughput performance that enables a far more standardized approach to the assessment of ADNCs. If individual centers begin to use a standard algorithm for quantitation, results could be used across institutions, thereby exponentially increasing the statistical power available to all centers involved. The digital approach is relatively efficient when it comes to manpower. Although it does take additional time to scan the slide and set up the analysis windows (~45 minutes to prepare and scan at 40 via the semiautomated method and an additional 5 to 10 minutes to select the analysis windows per slide), the bulk of the analysis work is done by the server alone. These analyses can be set up during the day and then allowed to run overnight without interruption. In addition, neuropathologic expertise is not a requirement for this method. Anyone at any level of expertise can be taught to select analysis areas in a few minutes. Because the analysis algorithms are held constant, regardless of who sets up the windows, the data will be consistent. This could be expanded
to involve algorithm sharing between institutions and thus improve the inter-rater reliability between the different research centers to help standardize the field of quantitative ADNCs.

Previous work showed that digital pathology could help elucidate different subtypes of AD cases based on quantifiable patterns of ADNCs (29). We confirmed that digital pathology can be used to discover new and interesting trends that we were not able to identify before at our center despite decades of work in quantitative assessment of ADNCs. Although our manually quantified ADNC numbers suggested that amyloid

**FIGURE 8.** The NP quantitation correlation (n = 48). (A) Even with a virtual cap, the correlation between the digital NP burden and the manual NP counts was below that of the neurofibrillary tangles (NFTs). On review of the cases with high digital NP burden and low manual NP counts, much of this was thought to be caused by the increased sensitivity of the PHF-1 immunostain when compared with the modified Bielschowsky method. (B, C) A representative 10 field in the superior and middle temporal gyri (SMTG) of the case highlighted with an arrow in (A). Those that were likely counted as DPs on the silver stain (B) were actually NPs by PHF-1 IHC (C). Scale bar = 50 μm.
plaque burden leveled off with increasing pathology, we failed to identify that it actually decreases with increasing tau burden by our manual methods alone. Intuitively, it makes sense that individuals with APOE ε4 alleles would have more amyloid plaque pathology (40–43), and this might correspond with having larger amyloid plaques; however, using manual counts, we could not demonstrate this trend reliably. These data may enable other new insights into the pathologic changes seen in AD.

Despite the benefits of the digital pathology methods, they also entail potential drawbacks. The up-front cost of digital pathology could be problematic for some centers and hospitals; the system described here cost almost $300,000 in 2010. Whole slide analysis was the theoretical goal; however, the massive amount of analysis time this required made this impractical. Our protocol still quantitates far more area than was previously examined manually. The most problematic pathology to quantify were the NPs, as indicated by the weaker correlations between digitally and manually counted numbers. Because of the heterogeneous nature of the NPs, single plaques could not be counted individually as they could by manual methods. Thus, we had to convert to an NP burden, which gives a picture of the overall NP surface area rather than an individual plaque number. We also switched our staining protocol from the modified Bielschowsky histochemical method to the PHF-1 IHC stain. This stain has a better sensitivity for tau-related abnormalities (44–47) and thus highlighted many more NPs than were seen with the silver method. It is likely that this combination of variables led to the weaker correlation seen between our NP parameters.

Digital pathology offers a valuable resource for quantitative pathology in neurodegenerative disease. With these algorithms, more AD-type pathology can be counted faster and more reproducibly than by manual inspection alone. In addition, more parameters can be rigorously examined from staining intensity to plaque size and more. As our use of this technology advances, it will open up new understanding of the pathologies in human brain aging.

**ACKNOWLEDGMENT**

We are deeply grateful to all of the participants in our longitudinal aging study and to the patients with Alzheimer disease in our Alzheimer’s Disease Center’s research clinic.

**REFERENCES**


**TABLE 2.** ICCs With 95% CIs to Assess Inter-Rater Reliability for 3 Users

<table>
<thead>
<tr>
<th>Measure</th>
<th>Estimated ICC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amyloid burden</td>
<td>0.973 (0.956–0.984)</td>
</tr>
<tr>
<td>Amyloid plaque density</td>
<td>0.974 (0.957–0.984)</td>
</tr>
<tr>
<td>Tau burden</td>
<td>0.986 (0.977–0.992)</td>
</tr>
<tr>
<td>NFT density</td>
<td>0.910 (0.855–0.945)</td>
</tr>
<tr>
<td>NP density</td>
<td>0.970 (0.951–0.982)</td>
</tr>
</tbody>
</table>

ICC, intraclass correlation coefficient.

**FIGURE 9.** Comparison of varying amyloid plaque densities as overall tau burden increases (n = 24). With the increased counting capabilities offered by digital quantitation, it was clear that plaque number tended to decrease as tau burden increased, a trend that could not be seen using our manual data alone.

**FIGURE 10.** Examination of amyloid plaque area in patients with known APOE allele status (APOE ε4, n = 7; non-APOE ε4, n = 41). (A) Patients with APOE ε4 alleles had a significantly larger average plaque area versus those without an APOE ε4 allele. (B) This finding holds true even when the degree of cognitive impairment is controlled by limiting the final MMSE scores to 27 to 30 (APOE ε4, n = 5; non-APOE ε4, n = 25). Graphs are mean ± SE. *p < 0.001, **p < 0.023.
34. Grimmer T, Tholen S, Yousefi BH, et al. Progression of cerebral amyloid load is associated with the apolipoprotein E epsilon4 genotype in Alzheimer’s disease. Biol Psychiatry 2010;68:79–84

© 2012 American Association of Neuropathologists, Inc.