ORIGINAL ARTICLE

Gray Matter Atrophy in Multiple Sclerosis: A Longitudinal Study

Elizabeth Fisher, Ph.D.,[1] Jar-Chi Lee, M.S.,[2] Kunio Nakamura, B.S.,[1] and Richard A. Rudick, M.D.[3]

Objective: To determine gray matter (GM) atrophy rates in multiple sclerosis (MS) patients at all stages of disease, and to identify predictors and clinical correlates of GM atrophy.

Methods: MS patients and healthy control subjects were observed over 4 years with standardized magnetic resonance imaging (MRI) and neurological examinations. Whole-brain, GM, and white matter atrophy rates were calculated. Subjects were categorized by disease status and disability progression to determine the clinical significance of atrophy. MRI predictors of atrophy were determined through multiple regression.

Results: Subjects included 17 healthy control subjects, 7 patients with clinically isolated syndromes, 36 patients with relapsing-remitting MS (RRMS), and 27 patients with secondary progressive MS (SPMS). Expressed as fold increase from control subjects, GM atrophy rate increased with disease stage, from 3.4-fold normal in clinically isolated syndromes patients converting to RRMS to 14-fold normal in SPMS. In contrast, white matter atrophy rates were constant across all MS disease stages at approximately 3-fold normal. GM atrophy correlated with disability. MRI measures of focal and diffuse tissue damage accounted for 62% of the variance in GM atrophy in RRMS, but there were no significant predictors of GM atrophy in SPMS.

Interpretation: Gray matter tissue damage dominates the pathological process as MS progresses, and underlies neurological disabillity. Imaging correlates of gray matter atrophy indicate that mechanisms differ in RRMS and SPMS. These findings demonstrate the clinical relevance of gray matter atrophy in MS, and underscore the need to understand its causes.

Ann Neurol 2008;64:255–265

Recent imaging and pathology studies have demonstrated that multiple sclerosis (MS) affects both cerebral gray matter (GM) as well as white matter (WM). Characteristic findings in GM include focal regions of demyelination, activated microglia, apoptotic neurons, and atrophy of cortical and deep GM structures.[1–4] Focal GM lesions are difficult to detect using conventional imaging because of low contrast and small lesion size. However, GM atrophy can be reliably measured from standard magnetic resonance images (MRIs). Quantitative analysis of MRIs has shown that GM tissue volumes are lower in MS patients than in healthy age-matched control subjects, and that GM atrophy begins early in the course of disease.[5–10] Prior studies of GM atrophy in MS patients have been designed to be cross-sectional, short-term longitudinal, or limited to a particular disease stage. The evolution of GM atrophy over the course of MS and how it relates to progression of disability have not been fully described. An important unanswered question is whether GM atrophy occurs as a direct result of GM pathology, or whether it is secondary to tissue damage within WM lesions.

The objective of this study is to characterize MS-related GM atrophy in a real-world setting. Patients with clinically isolated syndromes (CIS), relapsing-remitting MS (RRMS), and secondary progressive MS (SPMS), together with age- and sex-matched healthy individuals who served as concurrent control subjects, were entered into a prospective longitudinal study. This article addresses the pattern of brain tissue loss in these patients over the course of 4 years. The correlations between GM tissue loss and other MRI measures of tissue damage, and between GM tissue loss and clinical worsening were investigated.

Subjects and Methods

Subjects

Patients were recruited from the Mellen Center for Multiple Sclerosis Treatment and Research at Cleveland Clinic. Age-

From the [1]Department of Biomedical Engineering; [2]Quantitative Health Sciences Lerner Research Institute; and [3]Mellen Center for Multiple Sclerosis Treatment and Research, Neurologic Institute, Cleveland Clinic, Cleveland, OH.

Received Feb 15, 2008, and in revised form Mar 28. Accepted for publication May 12, 2008.

This article includes supplementary materials available via the Internet at http://www.interscience.wiley.com/jpages/0364-5134/suppmat

Published online July 23, 2008, in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/ana.21436

Address correspondence to Dr Fisher, Department of Biomedical Engineering ND20, Cleveland Clinic Foundation, 9500 Euclid Avenue, Cleveland, OH 44195. E-mail: fisherc@ccf.org

and sex-matched healthy control subjects were recruited by inviting the research subjects with MS to invite a spouse or friend into the study. The study was reviewed and approved by the Cleveland Clinic Institutional Review Board (IRB study numbers 2612 and 3709), and each study subject reviewed and signed an informed consent document. A financial payment of $50 per study visit was given to each patient and control subject for their participation.

Patients with MS met the International Panel criteria,[11] and each had a cranial MRI scan demonstrating lesions consistent with MS. Patients were classified as RRMS if they had two or more discrete relapses with significant neurological recovery in the prior 3 years, and as SPMS if they experienced continued deterioration for at least 6 months, with or without superimposed relapses in a patient who had a prior history of at least two relapses. Patients with CIS had an episode of neurological dysfunction typical for an initial MS presentation (eg, optic neuritis, transverse myelitis, brainstem syndromes). Conversion of CIS patients to clinically definite MS was based on a clinical relapse. Disease therapy with interferon-β, glatiramer acetate, methotrexate, or azathioprine was allowed. Age- and sex-matched healthy control subjects were required to have a normal neurological examination, a normal brain MRI, and no history of symptoms suggestive of MS. Patients and healthy control subjects were excluded if they received corticosteroid therapy within 2 months, were on bimonthly corticosteroid pulses, required therapy for hypertension, or had a history of transient ischemic attack or stroke, heart disease, pulmonary disease, diabetes, or chronic renal insufficiency.

Visit Schedule

Healthy control subjects were evaluated annually; CIS and MS patients were evaluated biannually. Demographic information (sex, age, racial background, educational status, and birthplace) and MS disease history (date and nature of first symptom, date of diagnosis, clinical pattern of disease from study onset, and clinical pattern of disease in the year before study entry) were recorded at baseline. At each visit, clinical assessments included Kurtzke Extended Disability Status Scale (EDSS), timed ambulation, 9-hole peg test, 3-second Paced Auditory Serial Addition Test, relapse history, and medications. The timed ambulation, 9-hole peg test, and Paced Auditory Serial Addition Test scores were transformed into the Multiple Sclerosis Functional Composite (MSFC) by normalization to a published MS reference group.[12] At each study visit, subjects underwent standardized MRI examinations (see later for descriptions). For subjects requiring corticosteroids for relapses, study visits were postponed until at least 6 weeks after steroid therapy to avoid confounding effects. Disease worsening among the MS subjects was defined in two ways: (1) subjects who progressed to a more severe stage of MS (eg, CIS patients who converted to RRMS; RRMS patients who converted to SPMS) were considered worse; and (2) subjects whose conditions worsened by 1.0 EDSS point (or 0.5 point for those who started the study with EDSS > 5.0) sustained for two consecutive 6-month visits were considered worse.

Magnetic Resonance Imaging Examinations

Images were acquired on a 1.5-Tesla magnetic resonance scanner and consisted of a T2-weighted, fluid-attenuated inversion recovery image (FLAIR), proton density-weighted images acquired with and without a magnetization transfer pulse for calculation of magnetization transfer ratio (MTR), and T1-weighted images acquired before and after injection of standard-dose gadolinium godopentetate dimeglumine (Gd-DTPA) (T1 and T1gad). The details of the image acquisitions are provided in Supplementary Table 1. (www.mrw.interscience.wiley.com/suppmat/0364-5134/suppmat/ana.21436.html)

MRIs were analyzed to calculate brain parenchymal fraction (BPF), gray matter fraction (GMF), white matter fraction (WMF), T2 lesion volume (T2LV), T1 hypointense lesion volume (T1LV), gadolinium-enhancing lesion volume, mean MTR of normal-appearing brain tissue (NABT MTR), and mean lesion MTR relative to normal-appearing tissue (lesion MTR ratio). All software for MRI analysis was developed at the Cleveland Clinic Department of Biomedical Engineering.

The whole brain was segmented and BPF was calculated from FLAIR images as described previously.[13,14] The FLAIR images were also used to segment T2 lesions as described later. GM voxels were segmented automatically from the T1 images using a new method that combines an intensity-based probability map and two types of regional probability maps (K. Nakamura and E. Fisher, unpublished data). First, the FLAIR image was registered to the T1 image, and the whole-brain and T2 lesion masks were applied to mask out non-brain and lesion voxels from the T1 image. The intensity-based GM probability map was calculated using unsupervised clustering (modified fuzzy c-means)[15] applied to the masked T1 image. An anatomic GM probability map was derived from a brain atlas coregistered to the patient's MRI. Lastly, an individualized GM morphological probability map was created using the brain surface and lateral ventricles as landmarks to define regions that have a high likelihood for GM. The three probability maps were then combined to create a final GM map. An example of the GM segmentation is shown in Figure 1. The lesion-masking step and the use of the regional probability maps effectively prevent the misclassification of lesions and partial volume voxels as GM, a common problem with GM segmentation algorithms.[16]

GM volume was calculated from the final GM map using a three-compartment model to account for partial-volume effects. The measured GM volume was then adjusted to correct for an artifact related to the use of fuzzy c-means. It was determined in a separate study that as T2 lesions enlarge, there is a reduction in the measured GM volume caused by slight shifts in the probabilities assigned to voxels in the intensity range between GM and WM. This results in an unwanted, but consistent and linear, dependence of GM volumes on T2LVs. The GM volumes were adjusted to account for this technical issue as follows: adjusted_GM_volume = measured_GM_volume + 0.26*T2LV. GM fraction was calculated as the final adjusted GM volume divided by the volume within the outer contour of the brain (the same volume as the denominator for BPF). WMF was calculated as BPF

Table 1. Baseline Characteristics

Characteristics	HC (n = 17)	CIS (n = 7)	CIS→MS (n = 8)	RRMS (n = 28)	RR→SPMS (n = 7)	SPMS (n = 19)
Mean age[a] (SD) [range], yr	41.6 (8.1) [32–56]	44.9 (10.1) [27–53]	36.1 (7.1) [26–49]	39.7 (8.4) [17–53]	42.2 (7.0) [33–50]	49.7 (7.4) [39–65]
Female sex, n (%)	10 (59)	6 (86)	5 (63)	23 (82)	4 (57)	14 (74)
Mean symptom duration[b] (SD), yr	NA	0.28 (0.07)	0.51 (0.70)	6.7 (5.1)	18.5 (11.1)	17.4 (5.0)
Mean EDSS score[b] (SD)	NA	0.86 (0.85)	1.19 (0.37)	2.0 (1.5)	4.79 (1.58)	5.39 (1.34)
Mean MSFC[a] (SD)	0.55 (27)	0.36 (0.37)	0.51 (0.46)	0.39 (0.62)	−0.34 (0.78)	−1.04 (1.49)
Mean T2 lesion volume[c] (SD), ml	NA	2.5 (2.3)	7.6 (8.3)	20.7 (17.8)	42.4 (24.1)	43.8 (26.0)
Mean T1 lesion volume[d] (SD), ml	NA	0.15 (0.22)	0.29 (0.48)	1.74 (2.49)	8.17 (6.89)	8.73 (9.00)
Gd+	NA	14.3%	37.5%	17.9%	28.6%	15.6%
Mean NABT MTR[a] (SD)	35.6 (0.75)	35.5 (0.98)	35.9 (0.81)	35.3 (1.01)	34.6 (0.35)	34.35 (1.19)
Mean lesion MTR ratio (SD)	NA	0.89 (0.12)	0.92 (0.05)	0.92 (0.04)	0.88 (0.03)	0.90 (0.05)
Mean BPF[e] (SD)	0.862 (0.012)	0.861 (0.008)	0.855 (0.023)	0.840 (0.027)	0.810 (0.02)	0.801 (0.04)
Mean GMF[f] (SD)	0.554 (0.015)	0.551 (0.010)	0.555 (0.016)	0.537 (0.018)	0.519 (0.017)	0.528 (0.032)
Mean WMF[g] (SD)	0.308 (0.011)	0.309 (0.009)	0.300 (0.017)	0.304 (0.016)	0.291 (0.015)	0.280 (0.016)

HC = healthy control subjects; CIS = patients who had a clinically isolated syndrome and did not meet the criteria for a diagnosis of clinically definite multiple sclerosis (MS) over the course of 4 years; CIS→MS = CIS patients who converted to clinically definite MS over the course of 4 years; RRMS = relapsing-remitting MS patients (throughout study); RR→SPMS = relapsing-remitting MS patients who converted to secondary progressive MS over the course of 4 years; SPMS = secondary progressive MS patients (throughout study); SD = standard deviation; EDSS = Expanded Disability Status Scale; MSFC = Multiple Sclerosis Functional Composite; Gd+ = patients with gadolinium-enhancing lesions; NABT = normal-appearing brain tissue; MTR = magnetization transfer ratio; BPF = brain parenchymal fraction; GMF = gray matter fraction; WMF = white matter fraction.
Significant differences between groups ($p < 0.05$): [a]CIS→MS versus SPMS; [b](CIS, CIS→MS, RRMS) versus (RR→SPMS, SPMS); [c](CIS) versus (RRMS, RR→SPMS, SPMS); [d](CIS, CIS→MS) versus (RR→SPMS, SPMS); [e](HC, CIS, CIS→MS) versus (RR→SPMS, SPMS); [f](HC, CIS, CIS→MS) versus (RR→SPMS, SPMS), RR versus SPMS; [g](HC, CIS, CIS→MS, RRMS) versus SPMS.

minus GMF. The accuracy and reproducibility of the GM segmentation method were evaluated in a separate study. The mean GM volume errors were determined to be 1.2% when assessed with BrainWeb (http://www.bic.mni.mcgill.ca/brainweb/) and 3.1% when compared with manual tracings. In a scan-rescan evaluation consisting of nine patients imaged three separate times over 2 weeks, the mean coefficient of variation for GMF was 1.1% (K. Nakamura and E. Fisher, submitted).

Baseline T2 lesions were automatically segmented in brain-masked FLAIR MRIs using a modified version of the Iterated Conditional Modes algorithm.[17] The T2 lesion mask was used to guide the automated segmentation of T1 hypointensities and gadolinium-enhancing lesions in coregistered T1 and T1gad images, as described previously.[18] Segmentation of T2 lesions in the follow-up images utilized the baseline lesion mask and a registration and subtraction method to detect both persistent and new T2 lesions. Lesion

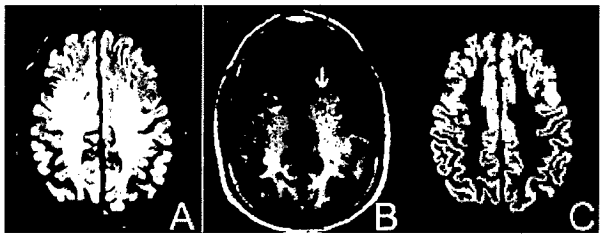

Fig 1. Example of gray matter (GM) segmentation results. (A) Fluid-attenuated inversion recovery (FLAIR) image used for segmentation of brain versus nonbrain structures and segmentation of T2 lesion versus normal-appearing brain tissue. (B) T1-weighted image used for segmentation of GM versus white matter. Arrow *indicates a T1-hypointense lesion with similar intensity to GM that would be potentially misclassified. (C) Final GM segmentation results. Note that lesion is not classified as GM.*

segmentation results were visually verified and semiautomatically corrected, using interactive software. Partial-volume effects were accounted for in the calculation of all brain and lesion volumes using either two- or three-compartment mixture models as appropriate.[19]

MTR images were calculated from the proton density image pair acquired with and without an MT pulse.[20] The mean MTR of NABT and mean lesion MTR were calculated from voxels included within the brain and T2 lesion masks, respectively. To ensure a consistent measure of the degree of abnormality of MTR within lesions, we normalized the mean lesion MTR by the mean MTR of normal-appearing WM (lesion MTR ratio).

Statistical Analysis

Baseline and on study changes were assessed and compared among disease subgroups and healthy control subjects using analysis of covariance. For categorical variables, a χ^2 test was performed. Spearman's rank correlation was used for correlations between GM atrophy and clinical disability. Pearson's correlation was used to assess correlations between GMF and WMF, and between GMF and age.

Multiple regression models were developed for whole-brain, GM, and WM atrophy using a set of baseline MRI predictors and their 4-year changes. The regression models were developed on 2 datasets: 36 RRMS patients, including both the CIS→RRMS and the stable RRMS groups, and 27 SPMS patients, including both the RRMS→SPMS and the SPMS groups. For each regression model, we followed a two-step process. First, we utilized bootstrap bagging method for predictor selection.[21,22] This method used automated forward stepwise selection to identified MRI predictors to include in 1,000 bootstrapped samples, which were 100% the size of the original dataset. All variables that were significant predictors in more than 50% of the bootstrap runs were retained. In the final regression models, only selected predictors were included. Adjusted R^2 is reported for the final models. All analyses were performed using SAS version 8.2 (SAS Institute, Cary, NC).

Results

One hundred six research subjects were enrolled in the study. Nineteen of these subjects (18%) discontinued the protocol for various reasons; this report provides information on the 87 subjects who remain in the protocol. These subjects were observed for a mean of 4.1 years (range, 3.4–4.8 years). At each study visit, the disease category for subjects with MS was assessed de novo, without reference to the disease category previously assigned. Eight of 15 subjects who initially entered with a diagnosis of CIS transitioned to RRMS; 7 of 35 initially categorized as RRMS transitioned to SPMS; and 1 of 20 initially categorized as SPMS was classified at every visit subsequent to baseline as RRMS. Of the eight CIS patients who converted to RRMS, five had new T2 lesions on the year 4 MRI, whereas only one of the seven CIS patients who did not convert to RRMS based on clinical criteria had a new T2 lesion.

Table 1 shows baseline characteristics for each subgroup, *as classified at the final visit.* As expected, compared with RRMS group, the average SPMS patient was significantly older, had greater symptom duration, greater EDSS score, lower MSFC score, greater T2LV and T1LV, lower NABT and lesion MTR score, and lower BPF, GMF, and WMF. Compared with RRMS patients who did not progress to SPMS, patients who converted had longer symptom duration, greater baseline EDSS, more EDSS change, greater baseline T2LV and T1LV, and lower baseline NABT MTR, BPF, GMF, and WMF. Table 2 shows the mean changes over 4 years for each subgroup. RRMS patients who converted to SPMS had greater changes in T1LV, BPF, GMF, and WMF than the RRMS patients who did not convert.

The changes in fractional brain volumes (BPF, WMF, and GMF) are plotted in Figure 2 for each subgroup. Whole-brain atrophy rates were similar in stable CIS patients and healthy control subjects, but steadily increased as disease severity increased. Increasing atrophy was driven entirely by increasing rates of GM atrophy. Expressed as a fold increase compared with the concurrently studied healthy control subjects, GMF change was 3.4-fold greater than normal in patients converting from CIS to RRMS, 8.1-fold greater in RRMS patients, 12.4-fold greater in patients converting from RRMS to SPMS, and 14-fold greater in SPMS patients. The GM atrophy rates in the combined set of RRMS patients (CIS→RRMS and RRMS stable) and combined set of SPMS patients (RRMS→SPMS and SPMS) were significantly greater than GM atrophy rate in healthy control subjects ($p = 0.05$ and $p = 0.005$, respectively). In contrast, WM atrophy rates were similar in all disease categories, at approximately threefold greater than in healthy control subjects.

Table 2. **On-Study Changes**

Characteristics	HC (n = 17)	CIS (n = 7)	CIS→MS (n = 8)	RRMS (n = 28)	RR→SPMS (n = 7)	SPMS (n = 19)
% Time on DMT[a]	NA	32.5 (47.4)	89.6 (24.6)	77.6 (41.1)	87.2 (21.7)	52.8 (42.0)
Mean ΔEDSS score (SD)	NA	0.50 (1.32)	0.63 (1.71)	−0.036 (0.78)	0.64 (0.69)	0.08 (0.38)
Mean ΔMSFC[b] (SD)	0.23 (0.14)	0.49 (0.39)	0.34 (0.42)	0.13 (0.28)	−1.44 (2.50)	0.12 (0.71)
Mean ΔT2 lesion volume (SD), ml	NA	−0.26 (0.23)	0.54 (1.56)	1.34 (3.21)	1.06 (2.92)	0.64 (5.28)
Mean ΔT1 lesion volume[c] (SD), ml	NA	0.007 (0.21)	−0.06 (0.15)	0.84 (1.06)	1.88 (1.54)	1.01 (1.37)
Gd+ during study[d]	NA	14.3%	50.0%	35.7%	100%	36.8%
Mean ΔNABT MTR (SD)	0.74 (1.03)	0.43 (0.59)	−0.55 (1.1)	−0.29 (0.98)	−0.22 (1.02)	−0.32 (0.75)
Mean Δlesion MTR (SD)	NA	0.032 (0.07)	−0.01 (0.04)	−0.015 (0.03)	−0.012 (0.03)	−0.016 (0.03)
Mean BPF %Δ/year[e] (SD)	−0.066 (0.22)	−0.003 (0.15)	−0.15 (0.14)	−0.23 (0.32)	−0.35 (0.18)	−0.39 (0.31)
Mean GMF %Δ/year (SD)	−0.028 (0.24)	−0.028 (0.25)	−0.096 (0.23)	−0.23 (0.34)	−0.35 (0.37)	−0.39 (0.50)
Mean WMF %Δ/year (SD)	−0.076 (0.35)	0.11 (0.25)	−0.24 (0.29)	−0.24 (0.72)	−0.33 (0.53)	−0.25 (0.49)

HC = healthy control subjects; CIS = patients who had a clinically isolated syndrome and did not meet the criteria for a diagnosis of clinically definite multiple sclerosis (MS) over the course of 4 years; CIS→MS = CIS patients who converted to clinically definite MS over the course of 4 years; RRMS = relapsing-remitting MS patients (throughout study); RR→SPMS = relapsing-remitting MS patients who converted to secondary progressive MS over the course of 4 years; SPMS = secondary progressive MS patients (throughout study); DMT = disease-modifying therapy; EDSS = Expanded Disability Status Scale; MSFC = Multiple Sclerosis Functional Composite; Δ = change; Gd+ = patients with gadolinium-enhancing lesions; NABT = normal-appearing brain tissue; MTR = magnetization transfer ratio; BPF = brain parenchymal fraction; GMF = gray matter fraction; WMF = white matter fraction. Significant differences between groups ($p < 0.05$): [a]CIS versus (CIS→MS, RR→SPMS); [b]RR→SPMS versus (CIS, CIS→MS, RR→SPMS, SPMS); [c]CIS→MS versus RR→SPMS; [d]RR→SPMS versus (CIS, RRMS, SPMS); [e]CIS versus SPMS.

Correlations between Atrophy and Age

The differences in GM atrophy rates between groups could not be explained by differences in age. GM atrophy rate was not significantly correlated with age in the patients as a whole or in any of the MS subgroups. After adjusting for age, the GM atrophy rate in RRMS patients was still significantly greater than in healthy control subjects ($p = 0.048$). The age-adjusted GM atrophy rate in the SPMS patients was also significantly greater than in healthy control subjects ($p = 0.019$).

Correlations between Atrophy and Disability

Table 3 shows clinical characteristics and atrophy measurements for RRMS and SPMS patients according to whether they had sustained EDSS worsening over the 4 years. Other than EDSS change, there were no significant differences between the subgroup of 13 subjects with sustained EDSS worsening compared with more stable patients, probably because of the small number of subjects in the comparison groups. However, there was a consistent trend toward more whole-brain and GM atrophy in patients with EDSS progression. Of interest, a similar pattern was not observed for change in WMF.

GMF was correlated with clinical disability scores, as shown in Figure 3. Correlations were greatest with MSFC (Spearman's rank correlation coefficient = 0.52) and were similar between RRMS and SPMS subgroups. GMF was also moderately correlated with

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.