

A Review of the Automated Detection of Change in Serial Imaging Studies of the Brain

Julia Patriarche and Bradley Erickson, MD, PhD

Serial imaging is frequently performed on patients with diseases of the brain, to track and observe changes. Magnetic resonance imaging provides very detailed and rich information, and is therefore used frequently for this application. The data provided by MR can be so plentiful; however, that it obfuscates the information the radiologist seeks. A system which could reduce the large quantity of primitive data to a smaller and more informative subset of data, emphasizing change, would be useful. This article discusses motivating factors for the production of an automated process to this effect, and reviews the approaches of previous authors. The discussion is focused on brain tumors and multiple sclerosis, but many of the ideas are applicable to other disease processes, as well.

KEY WORDS: Change detection, magnetic resonance, brain tumor, multiple sclerosis

WHEN A PATIENT is found to have a brain tumor, a treatment protocol is initiated which typically includes imaging studies to localize the pathology and narrow the differential diagnosis and biopsy to determine the type of tumor. Interventions are chosen depending on the biopsy-derived diagnosis, which may include any combination of surgical resection, generalized or localized radiation, and chemotherapy. During and after intervention, patients receive periodic imaging examinations to watch for and monitor changes. In serial imaging, the radiologist is presented with an enormous amount of data. The changes may be subtle. Tracking disease course is of great clinical importance; however, with current methods there is frequently ambiguity as to whether or not a change has occurred, and it is not typically readily apparent what changes have occurred, where, and to what degree. The

exploration of methods of technological automation to improve this process is therefore one of active investigation.

There are many acquisition-related factors which make manual examination of studies for the detection of pathology-related change difficult. Perfect image alignment is not possible from one scan to the next, or even within one acquisition (due to patient motion). Therefore, one cannot assume a one-to-one correspondence between slices from one acquisition to the next to make side-by-side comparisons. Furthermore, a different scanner may be used in a followup scan which will have different signal characteristics and may even have different operating software. Often, magnetic resonance (MR) or computer tomography (CT) scanning parameters are not the same from one acquisition to the next. The gradients may have decayed over time. Radiofrequency (RF) field heterogeneities may have changed over time. An essential component in the process of change detection is therefore not simply the detection of change but the separation of acquisition-related change from disease-related change. Methods that help to separate acquisition-related change from disease-related change, methods that aid in reducing the quantity of data presented to the radiologist, and methods that produce objective, reproduc-

From the Department of Radiology, Mayo Clinic and Foundation, 200 First Street SW, 55905, Rochester, MN.

Correspondence to: Bradley Erickson, MD, PhD; e-mail: bje@mayo.edu

Copyright © 2004 by SCAR (Society for Computer Applications in Radiology)

Online publication 29 June 2004

doi: 10.1007/s10278-004-1010-x

ible, and accurate metrics of disease course are of great interest. The introduction of serial studies does add data volume, but understanding the change in appearance over time is essential to understanding disease course.

DETAILED REVIEW OF THE LITERATURE

Manual Inspection

The most common approach for the detection of change on imaging studies is visual inspection; however, this approach suffers from a myriad of problems. One of the most important of these is the quantity of data presented to the radiologist. At each acquisition, there are multiple pulse sequences, each of which consists of many slices. Change can present itself in many ways, to different degrees spread across different pulse sequences. The radiologist is required to assimilate all of this data in order to render a decision, which often is quite difficult. One group of authors showed that volumetric changes of up to 59% were not appreciated by visual inspection in their study.¹ Some authors have used subtraction to aid in the process of manual inspection. One group applied a registration and subtraction technique which reportedly achieves subvoxel registration accuracy in part through the use of sinc interpolation.^{2,3} In summary, when certain conditions are met (such as the data is band-limited), sinc interpolation is a method of interpolation which provides reduced error compared with other interpolation techniques, eg, linear interpolation. The sinc function is the in-slice point spread function for MR because the data as acquired in k -space is band-limited. In Hajnal et al.,² the authors demonstrate the ability of their technique to aid in the observation of many kinds of subtle change. They show deformations in normal brain structure between a volunteer with his head resting in the scanner on the left side versus the right side (with the deformations induced by gravity); brain growth patterns in a normal child; changes in sulci, cisterns, and sinuses when a normal subject breathes 100% oxygen versus carbogen; the growth of a brain tumor; progression of multiple sclerosis; and postoperative changes in a patient after undergoing endarterectomy (to

name a few). Though their technique represents a valuable step in change detection, it is decidedly manual and subjective: The actual process of interpreting the subtraction images is conducted entirely by inspection. Furthermore, these authors subtract only one pulse sequence at a time, which still leaves the radiologist to assimilate and mentally manage a potentially large amount of information in order to yield a clinical judgment. An important part of their report, though, is that they assert, demonstrate, and explain that with subtraction techniques spatial changes smaller than one voxel can be detected.

Some authors of articles on change detection by subtraction incorporate modifications to the basic approach. One group utilizes fast sinc resampling, a modified version of sinc interpolation as described above, to improve computational tractability.⁴ They additionally incorporate a linear stretching term in the registration, as they remark that, during the subtraction phase, small changes in voxel size (eg, due to variations in field gradients of an MR scanner) can be misinterpreted as being due to pathologic change. Another group has applied change detection to patients with multiple sclerosis.⁷ They utilize a surface registration approach, followed by subtraction. Subvoxel registration is likely to be an essential step in any change detection approach. However, due to the anisotropy of most current practical MR pulse sequences, postacquisition registration of nonisotropic pulse sequences necessarily reduces the acuity of interacquisition comparison, particularly when the prescribed registration involves a rotation through the plane of acquisition. Furthermore, Rusinek et al.⁵ present subtraction of T1-weighted images only; however, change detection of diseases typically requires the examination of many pulse sequences together. One group applies a test of statistical significance after subtraction to reduce false positives, using a $3 \times 3 \times 3$ window.⁶

Ettinger et al.⁷ combines classification and subtraction in order to detect change.⁷ They align serial images and subtract the classified volumes to observe changes in patients with multiple sclerosis. They remark that the subtraction of classified volumes helps to address a number of issues related to acquisition such as

changes induced by changing the scanner from baseline to followup and RF heterogeneity effects. This is important, because it is a step toward the separation of acquisition-related changes from pathology-related changes. Their classification, however, is a hard classification—this unnecessarily reduces the ability of the system to detect many types of changes (eg, subtotal changes in tissue character and subvoxel changes in boundary location). Serial scanning, in principle, should be able to detect subvoxel shifts in boundaries; in many cases, hard classification loses this information. As with many articles on change detection, one focus of the Ettinger et al.⁷ article is their serial exam registration algorithm, which is a surface registration strategy that uses the intracranial cavity (ICC) as an invariant structure. This approach reportedly only achieves a root mean square (RMS) error of 1.96 mm, which is insufficient to detect subtle changes. This low accuracy likely results because this approach does not take into consideration the position and shape changes which the brain undergoes relative to the ICC. The brain's shape is not invariant; it is pliable and can move within the cranial cavity as a result of its suspension in cerebrospinal fluid (CSF). The Hajnal et al.² article provides good demonstrations of changes in brain position and shape that can be observed in normal and pathological cases. For serial imaging in particular, registration is possible which yields accuracy of a fraction of a voxel, since large portions of the anatomy will be entirely unchanged. A technique which overlooks this potential loses change detection sensitivity unnecessarily. Furthermore, the authors use trilinear interpolation to effect the final transformation, which introduces errors. Interpolation methods, such as sinc interpolation, have less of a potential for introducing errors; and in the ideal case (volumetric 3D acquisition and infinite sinc window), sinc does not introduce errors at all.

Measurement Sampling

The measurement of tumor diameter is one approach that has been used to assess patient status. Two standard methods are those of the World Health Organization⁸ and the Southwest

Oncology Group.⁹ Both of these methods use largest diameter and largest corresponding perpendicular. More recent guidelines, such as RECIST, promote the use of a single maximal tumor diameter.¹⁰⁻¹⁴ There are significant problems associated with maximal diameter approaches. RECIST, for example, completely ignores lesions smaller than 1 cm and limits the overall measure to include a maximum of 5 individual lesions per organ.¹⁰ This method also sums up computed diameters over multiple lesions, which can tend to obfuscate the responses of individual lesions. Furthermore, the RECIST guideline dictates that measurements should only be taken within the acquisition plane, which is problematic as lesions do not grow strictly in-plane. Finally, RECIST assumes a single image type and does not address how to handle disparate boundaries seen on different MR pulse sequences. For serial image inspection, the method would be completely insensitive to tumors that grow in the through-plane direction. Another problem with these approaches is the intended generality of their application. Compared with other organs, the brain is particularly sensitive to tumor location. Some authors remark that the purpose of objective tumor response quantification is to develop a surrogate marker for the prediction of clinical events such as symptom control, time to death, or disease progression.¹²

Particularly with brain lesions, however, tumor location greatly influences the relationship between such a surrogate and its intended clinical correlates. To disregard location information is to potentially obfuscate this relationship beyond intelligibility. These techniques are additionally designed to apply only to solid tumors. Many features of brain tumors, such as edema, necrosis, and in many cases the tumors themselves, do not obey demarcations. It is not surprising, therefore, that RECIST measurements applied to brain tumors do not correlate well with either radiologist impressions or clinical status.

Although these approaches attempt to construct an objective and quantitative metric and make it one which the radiologist can compute quickly and easily, the multiplicity of assumptions and the use of metrics which are so sensitive to measurement uncertainty serve to

undermine the ability to detect small changes. Thiesse et al¹⁵ found major disagreements between observers for 40% of evaluations, and minor disagreements for a further 10.5% of evaluations. Much of this disagreement stems from the sensitivity of the measures to potentially small spatial measurement errors. Filipek et al¹ found volumetric changes of up to 145% associated with radius changes of only 2.7 mm, which may be difficult to appreciate, particularly in light of the irregular borders exhibited by many tumors. Clarke et al¹⁶ investigated the two-diameter approach with serial MR images of six glioma patients. They asked radiologists to compute the maximum area using the two-diameter method; they then determined that the coefficient of variability of the area measurements was 16%, which they attributed primarily to human difficulty in determining the tumor margin. They then asked the radiologists to use their area measures to rate each patient using the standardized categories: complete response, partial response, stable disease, and progressive disease. In 30% of the cases, the radiologists were unable to reach a consensus because the categories were so sensitive to the variability of the area metric. The one- and two-diameter measures are sometimes thought of as surrogate markers for tumor volume, which in some studies has been shown to be predictive of survival.¹⁷ In contrast, Chow et al¹⁷ also found that the area of enhancing tumor was not predictive of survival. This may be explained by Clarke et al¹⁶ who demonstrated that there was no significant relationship between the two-diameter-derived area measure and the tumor volume, which may be due to variability in tumor shape and lack of interobserver consistency. In light of the findings of these studies, it seems that the one- and two-diameter measurement sampling approaches are suboptimal. The use of automation, as opposed to shortcuts, should prove much more efficacious in establishing a clinically usable and quantitative technique to predict response and survival from radiologic studies.

Volumetrics

Some authors have computed global volume of each tissue of interest at each acquisition in a

series, with the intention of subtracting this metric from one scan to the next. Examples of this method are available for tumors, multiple sclerosis,¹⁸ and Alzheimer's diseases.^{5,19} A significant problem with this approach is that, from a mathematical standpoint, computing a measure (such as volume of a structure of interest) in one scan and then computing the corresponding measure in a followup scan and finally subtracting the two necessitates an increase in uncertainty. This is not the case when the change is computed in one step. Computing volumes over unnecessarily large regions in an image also affects uncertainty. The summed error over many voxels in a volume (eg, resulting from the determination of the boundary position) can be prodigious and can easily overwhelm small measures of interest, such as those related to minimally detectable change.

Jack et al¹⁹ used volumetric measurements of the hippocampus and temporal horn to compare a control group with a group of patients with Alzheimer's disease. They had an expert outline these anatomical structures, and the volumes for each patient were computed. Each patient was reimaged approximately one year later and the process of outlining and volume computation was repeated. From these values, annual percent change in volume of each anatomical structure was computed. This approach overlooks partial volume effects to some degree, and so a subtle spatial shift along a boundary might be overlooked, even if it extended for many voxels and therefore included a relatively large total volumetric change. It additionally increases the uncertainty in the change metric, in the sense that hard classification amounts to "rounding up" and "rounding down" of the fuzzy membership of the boundary voxels. Information exists in the image to delineate the boundary between structures with subvoxel precision. One person's partial volume artifact is another person's subvoxel information. This information is being discarded when boundaries are manually delineated with only voxel-size precision.

Volumetrics finds even greater obstacles when applied to processes such as tumors and multiple sclerosis because of the use of manual boundary delineation of the structure of inter-

est, and consequently the assignment of a hard classification. Brain gliomas are usually infiltrative and, therefore, within a voxel there exists a gradation in membership. In this sense, hard classification is unnatural to many tumors. White matter can exhibit greater or lesser enhancement and can be more or less edematous. These changes in tissue character are at least as significant as frank tumor growth, when change is of interest. It is also possible for a tumor to change shape without changing volume. For example, it is possible for part of a tumor to become necrotic while another part grows (yielding no net volumetric change). In each of these cases, volumetric techniques would not indicate that a change had occurred when in fact one had. Another issue arises from the fact that, from a clinical standpoint, knowing where and how changes have occurred is as important as knowing that they have occurred. This will have a profound impact upon how symptoms will manifest and upon prognosis. If a tumor is invading eloquent regions of the brain, the implication for treatment and prognosis is quite different compared with when a tumor is invading comparatively expendable regions. Conventional volumetric techniques do not provide information in this regard. Which anatomic and functional structures are involved will impact the choice of repeat intervention, or even the possibility of repeat intervention. Assuming mechanical considerations have been taken into account,²⁰ the location and degree of change could also impact the choice of biopsy site, since localized description of changes may suggest that one region of a tumor was more aggressive than another.

In a discussion of volumetrics with regard to tumor imaging, it is important to discuss the reason for computing volume change as opposed to static single volume. There is a lack of consensus in the literature as to whether static tumor volume is predictive of outcome. Some authors have found volume to be predictive of outcome,¹⁷ whereas other authors have not.^{21,22} One possible explanation for this disagreement is that abnormal signal demonstrated in imaging studies is not fully reflective of the extent of pathology from a histologic standpoint. Kelly et al^{23,24} attempted to correlate MRI and biopsy findings in a series of patients with previously

untreated intracranial glial neoplasms. Although biopsy of tissue with normal MR imaging characteristics was not performed with all patients for ethical reasons, when it was done, approximately half of the biopsies demonstrated isolated tumor cells in regions of parenchyma with normal imaging characteristics. Burger et al²⁵ examined postmortem sections of patients with glioblastoma multiforme (GBM) and found isolated tumor cells in tissue that imaged normally in MR, as distant as the hemisphere opposite to that of the principal lesion. Johnson et al²⁶ examined four formalin-fixed brain specimens from patients who had had a previous diagnosis of GBM using postmortem MRI and microscopic studies of whole-brain sections cut at the same level. They found tumor cells up to 5 cm from the nearest MR signal abnormality. Tovi et al^{27,28} compared pre- and postmortem MR examinations of five patients who had malignant glial brain tumors with histopathologic examination of whole-brain sections. In 4 of 5 cases, tumor cells were found in distant areas that had appeared normal in the MR examinations. In one of their cases, a well-demarcated 1 × 5-cm tumor was found, and in two of their cases, tumor cells were in the contralateral hemisphere. One reason these findings are particularly significant is that they strongly suggest that quantifying the volume of abnormal imaging will not necessarily provide an accurate representation of the volume and extent of histologic tumor burden. However, it is likely that change seen in imaging is reflective of change from a histologic standpoint. In a sense, therefore, change seen in serial imaging is likely to be more strongly correlated with clinical status than imaging-based static volume measurements. This point is made more concretely by Filipek et al¹ who describe that the speed of change is an indication of the aggressiveness of the tumor and has been shown statistically to be a predictor of response.

Haney et al^{29,30} emphasizes the importance of detection of subtle changes. They describe techniques for the computation of contrast-enhancing tumor volume in serial MR and growth rate secondarily, in order to predict therapeutic response. They describe two methods, one based upon nearest-neighbor classification and the other based upon surface

modeling. In essence, they identify the tumor and compute its volume at each acquisition. They then subtract the volume found for one acquisition from the volume found at the other acquisition and divide by the time interval between scans to obtain the growth rate. The authors apply these methods, in conjunction with short-interval scanning (image acquisition interval once per week), in order to assess the efficacy of temozolamide, a chemotherapeutic agent, in a patient with glioblastoma multiforme.²⁹ They demonstrate that short-interval scanning and growth rate computation allows investigators to observe the slowing and arrest of tumor growth, following the administration of their trial chemotherapeutic agent. They suggest that long-interval scanning could miss subtleties of the response, particularly if the tumor later progressed. For example, if the next followup scan had been taken after a response and then progression, investigators might be led to erroneously conclude that the tumor had never responded. The authors further assert that volume and growth rate computation using such short-interval scanning may provide a mechanism to disambiguate between radiation necrosis and recurrent tumor in ambiguous lesions; the patient in that particular study had an ambiguous lesion at the commencement of the short-interval scanning protocol. One reason this study is so germane to the current discussion is that a short time interval between acquisitions means volume changes will be small. Small errors in determining the boundary or region of the tumor, accumulated over a large tumor volume, would have the potential to obscure the presence, absence, and magnitude of any actual change. A system able to localize and characterize minimally detectable change, and only the change, has the potential to make growth rate determination more accurate and less obscured by uncertainty.

Warping

Rey et al³¹ have used nonlinear registration (warping) as a method for detecting change. There are advantages to this approach. Thompson et al³² examined patterns of growth and development in the brains of normal children using continuum mechanical tensor maps.

They used serial T1-weighted fast SPGR MR images, applied a bias field correction algorithm, rigidly registered interacquisition images and resampled using Chirp-Z (in plane) and linear interpolation (through plane), and applied histogram matching. They constructed surface models of the cortex, deep sulci, corpus callosum, caudate, and ventricular surfaces based upon manually defined points. They used an elastic image registration algorithm to match surfaces from one acquisition to the next; the resulting surface deformations were used to derive volumetric deformation fields from which local measures of three-dimensional tissue dilation and contraction were quantified. They emphasized that volumetric techniques overlook very important information: When they reduce change to a small set of numbers and attempt to use those numbers to describe change, significant localized changes may be overlooked. For example, between ages 7 and 11 years, the global measurements of these authors showed an overall 22.4% increase in midsagittal callosal cross-sectional area (where cross-sectional area is an indication of the number of white matter fibers crossing at that location); however, they found local growth to be as high as 80%. With regard to tumors, this is a critical point if growth rate is a marker of therapeutic response and prognosis. The method described above is an example of a class of warping algorithms where the transformation is derived based on the surfaces of structures which have typically sharp boundaries but are relatively homogenous internally. However, brain tumors are inherently not discrete structures: Tumor, edema, and necrosis inherently blend into structures of which they are a part. It is therefore not appropriate to use surface warping approaches in the case of tumors.

There are a number of problems with using nonlinear registration for change detection. One is that nonlinear registration is underconstrained. For a given pair of acquisitions, there are an infinite number of displacement fields that will yield a match. This problem is typically partly addressed by specifying the constraints under which the displacement field will be derived, for example, by the use of continuum mechanical models. A viscous fluid model of Freeborough et al³³, was used to register serial

MR images of patients with Alzheimer's disease. The use of a fluid model is an approximation; limited work has been done to describe how tumors grow and how infiltration and growth are balanced. These effects make it very difficult to disambiguate the underconstrained warping problem. In addition, current warping algorithms also require a one-to-one tissue correspondence, based upon intensity. Warping algorithms are based on the assumption that a particular region of tissue exists in both acquisitions (and appears identical in terms of intensity) and has simply moved. This is a requirement which tumors frequently do not meet. In tumor progression and regression, a given region of tissue can not only move, it can change character: white matter can become enhancing or edematous; enhancing tumor can become necrotic, etc. Infiltration and expansion can coexist with this change of character as well. The case of white matter becoming edematous is an example of this, because edema is known to result in a 10%–40% volumetric expansion of white matter.³⁴ This case is particularly underconstrained because the warping algorithm is left to determine which region of the image has become edematous and which region represents white matter which was previously edematous and expanded. A similar issue exists with infiltrative tumor. If new enhancement arises adjacent to a region which was previously enhancing, the algorithm must determine whether previously nonenhancing white matter has become infiltrated, or whether the tumor itself has pushed the white matter away, or a combination of these effects. To some degree it may do this by inference, through examination of mass effect in surrounding tissue.

An approach that attempts to accomplish this disambiguation of frank growth from tissue character change has been described by Thirion et al.³⁵ The central theme provided by these authors is that if new tissue is being deposited (as in the case of tumor growth), then mass effect will be present in adjacent structures. They use a warping algorithm to develop a vector displacement field from the serial imaging studies. The user then gives the coordinates of the center of a lesion to be examined, and the algorithm places concentric shells (which may

be spheres or isocontours) around this point. For each shell, the divergence of the warping field is integrated over the volume within; this yields the change in volume (ΔV) for that shell. If character change is occurring within the shell without deposition of new tissue, then as the shells are made larger and larger, the computed ΔV will return to zero, indicating that there is no net change in volume within the shell. If, on the other hand, ΔV reaches a plateau as the shells are made larger and larger around the lesion center, then real mass effect is present and new tissue is being deposited. This approach of integrating divergence over the volume of a series of concentric shells centered on a lesion center has a great deal of merit. The authors demonstrate the algorithm's ability to detect changes where no changes in enhancement are present at all, only mass effect. On the other hand, there are a few significant drawbacks. In the ideal case, if mass effect were present, the computed value of ΔV would remain constant as soon as the shells were made large enough to completely encompass the lesion. Because of noise, however, this is not the case. As the shells become larger than the size necessary to completely encompass the lesion, the computed value of ΔV oscillates around the theoretical ΔV plateau, and in fact the magnitude of this oscillation becomes larger and larger as the shell is made larger and larger. The authors address this issue by setting all values of their vector displacement field lower than a specified threshold to zero in order to control the noise. This is one critical problem since it strictly limits the algorithm's ability to detect small changes. Furthermore, while this algorithm works very well when a lesion generates mass effect without a change in enhancement, it works much less well at the other extreme: change in enhancement with no mass effect. If a region of lesion becomes more enhancing, for example, a small level of enhancement increase would not be recognized by the algorithm at all, while larger enhancement increases would not be differentiable from each other. Another problem is related to the use of continuum mechanical models for warping. The forces of the continuum mechanical model resist the forces generated by the intensity differences between the two serial images being matched.³⁶

This is intentional. The purpose of these models is to enforce smoothness and penalize locally large deformations. Without these forces there would be no penalty, for example, to prevent the warping algorithm from moving a group of collocated points from the first acquisition so that they ended up on top of each other in the second acquisition. However, particularly when tissue class change occurs, large and irregular deformations would be required if the transformations are to be treated as if they were due to deformation. The simple fact is that tissue class changes are not due to deformation, so attempting to model them as if they were, while enforcing continuum mechanical constraints, is inconsistent. The degree to which the mechanical model is enforced will impact the system's ability to detect and characterize these kinds of changes. Additionally, inferring change from measured secondary mass effect is problematic if the lesion exists in a large homogeneous region of white matter, since in a three-dimensional volume the impact of mass effect decays as the square of the distance. Small changes, in particular, are quickly overwhelmed by noise and generate no measurable mass effect in neighboring structures. Another problem is that this approach would not be particularly sensitive to many small distributed changes, for example, the development of distributed necrosis. If small regions of necrosis appeared within a previously enhancing tumor, the warping algorithm would align enhancement with enhancement; however, the displacement and the divergences might not be large, even though large changes in intensity had occurred. Significant mass effect would not be expected since the development of necrosis does not involve the deposition of large amounts of new tissue. This change could be overlooked, which is significant given that the development of necrosis is a significant prognostic indicator.²²

Temporal Analysis

Some authors have described the analysis of serial sets of MR images of patients with multiple sclerosis as 4D datasets.^{37,38} Gerig et al³⁷ compute various metrics for each voxel, such as difference between maximum and minimum value over the time series, mean and standard

deviations over the time series, number of crossings of the mean, maximum deviation from the mean, minimum deviation from the mean, maximum time gradient, etc. They then combine these metrics to obtain a metric representing the probability of a given voxel being involved in lesion-related change over the given time series. Knowledge of the disease process in question, and the expected periodicity, allows specific analysis of the fourth dimension (time). There are differences between the analysis of serial images of patients with multiple sclerosis as described and the analysis of patients with brain tumors. In the former, lesions are expected to wax and wane over time, so metrics used to detect lesions and changes in lesions will be attuned to this fact. Nevertheless, these authors raise a very interesting point: In the analysis of serial MR datasets, one may be at an advantage if one is able to leave behind notions of static numerical analysis and think of the data in terms of a 4D dataset. One may be at an even greater advantage if one is able to match the metrics to the behavior of the disease process being studied.

Steps in a Practical Change Detection System

In practical terms, a change detection system should not be thought of as an algorithm but rather as a pipeline of algorithms and steps.³⁸ Figure 1 shows a possible sequence of these steps. Some of these steps are possible with methods currently described in the literature (eg, step 4: interscan scaling correction based upon the size invariance of the intracranial cavity), and some are not possible with currently available techniques. A good implementation of some steps could ameliorate the need for others. For example, if perfect prospective registration were possible, then retrospective registration would be unnecessary. Likewise, if control could be obtained over acquisition-related variability, then the need for subsequent correction of inhomogeneities and intensity normalization would be ameliorated. This discussion is not intended to be a final one; it is intended only to address some of the issues concerning practical change detection and to point out that change detection represents a

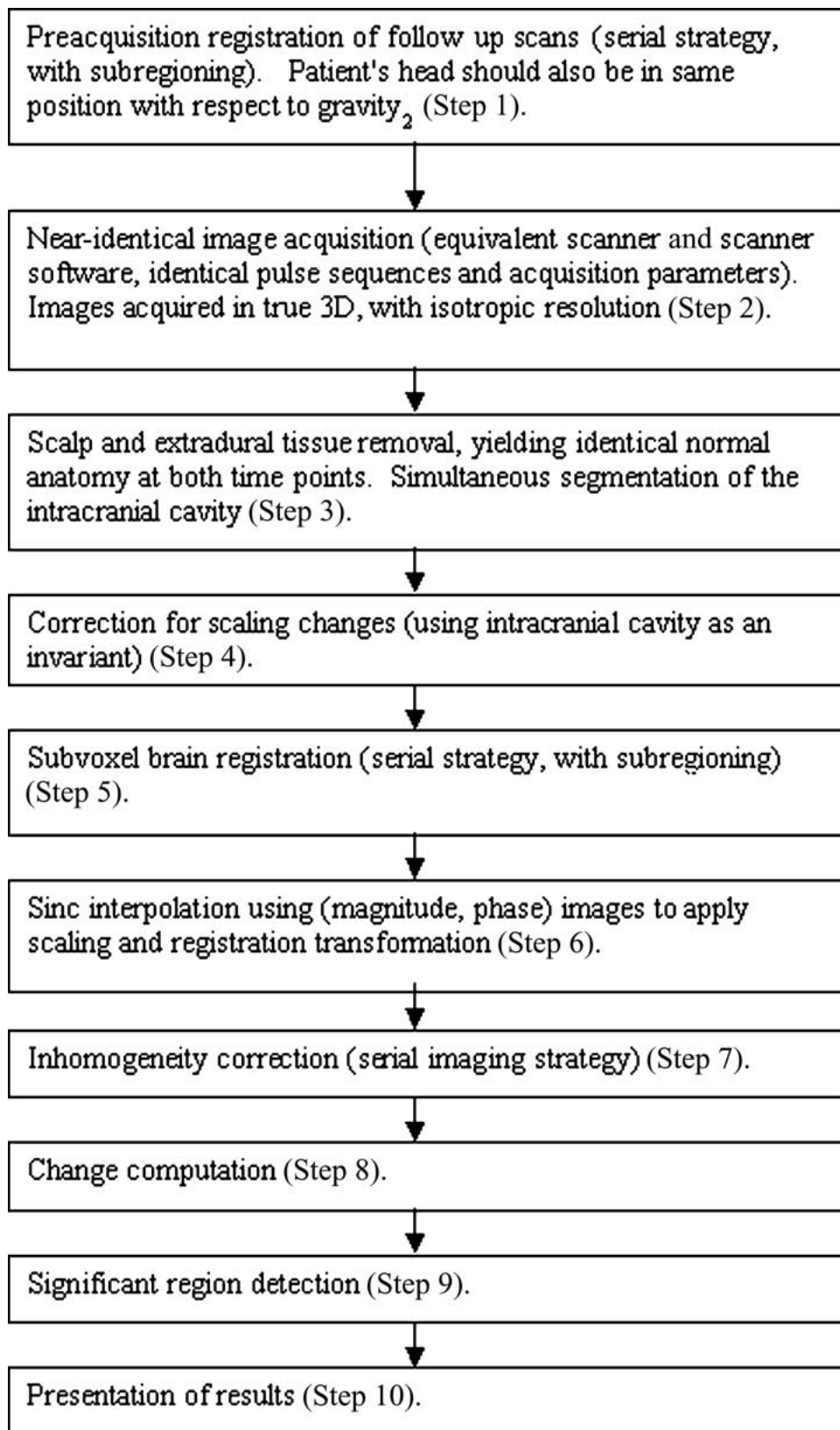


Fig 1. Possible steps in a practical change detection system.

series of issues and not just one. To undertake change detection necessitates developing approaches for separating acquisition-related changes from the pathology-related changes, which are of interest to the radiologist. This section suggests some possible technical approaches, but this list is by no means exhaustive.

Preacquisition Registration of Scans

Kikinis et al.³⁹ performed 24 serial exams of a patient with multiple sclerosis over a one-year period in order to track lesion evolution. As a part of their study, they evaluated the variations in head position. No effort was made to recreate prior physical positioning from study to study. They defined four points, “the nasion, the external auditory canals, and a point at the vertex.” They used a rigid body registration algorithm using the intracranial cavity as a basis for registration and determined that, over the year, the average displacement without registration was 0.9 cm; the minimum was 0.1 cm and the maximum was 1.9 cm. This is of concern because interpolation to effect a registration transformation and bring two studies into alignment results in artifacts, particularly when thick slices, nonisotropic voxels, slice gaps, and non-true-3D pulse sequences are used. These artifacts limit the change detection algorithm’s acuity. Another issue of positioning relates to partial volume effects. Guttmann et al.⁴⁰ studied the impact of various factors on lesion load in patients with multiple sclerosis. They found that typical patient repositioning in the scanner (completely unrelated to interpolation) had a median impact of 5.4% on measured lesion burden. Gawne-Cain et al.⁴¹ reported an even larger median difference of 9.9% (the difference may have been partly attributable to the fact that the first group of authors used 3-mm contiguous slices while the second group of authors used 5-mm contiguous slices). Both groups used hard classification as their method of identifying tumors (the first group used an automatic classification technique and the second group used manual contouring), and they further reason that changes in partial volume effects probably account for the bulk of the changes in measured volume. It would be

interesting to investigate whether a fuzzy classification-based method for volume measurement would suffer to the same degree from changes in scan orientation in serial acquisition. If nothing else, such a study would emphasize how important fuzzy methods are in computing volumes from an imaging modality for which partial volume effects are endemic. Nevertheless, this variability in volume quantification with positioning between scans is of critical importance since it places a very significant floor on the size of changes that can be detected with serial imaging, even before any image processing (even registration) has been undertaken. Although acquiring serial scans already aligned would not eliminate the errors these authors describe (partial volume effects are always present), it would control the variability they induce, and therefore measures of change would be more reliable.

Kikinis et al.³⁹ also reported that the act of interpolating to register images that were acquired out of registration had a statistically significant impact on their measured lesion load, particularly when the lesion load was small. They used trilinear interpolation which is far from optimal, but even sinc interpolation, which theoretically should not introduce errors when the image is a true 3D volumetric acquisition, is generally windowed spatially (due to computational considerations) and therefore introduces artifact. Additionally, if the acquisition is not 3D, sinc interpolation cannot be used for through-slice interpolation. Furthermore, when the slice thickness is greater than the in-plane voxel dimension, the act of rotating through the through-plane direction introduces additional blurring into the in-plane dimension. Preacquisition registration would help to ameliorate these problems.

Preacquisition registration can take two forms: mechanical and electronic. The first is possible (to a certain degree) with technology which is readily available. The community of radiation therapy providers, for example, uses masks molded to the shape of the patient’s face to hold the head in the correct position and orientation from one treatment session to the next. Pilipuf et al.⁴² describes a system that reportedly achieves submillimeter control over positioning (mean = 0.6 mm, SD = 0.1 mm,

max = 1.0 mm), with negligible patient discomfort and low cost. Noninvasive head holders that use dental molds which the patient bites on and which are rigidly affixed to the head holder (which is in turn affixed to the scanner) are also described.^{43,44} A noninvasive method of head fixation that uses two ear plugs and one rest situated on the patient's nasion is also described.⁴⁵ A head fixation device using a plaster mold of the head has also been described.⁴⁶ Many additional approaches for head fixation exist; some key design considerations have been described by Strother et al.⁴⁷ Given the significance of change detection and the cost of serial MR imaging, the relatively low incremental cost and difficulty of using head fixation devices probably justifies their use. An additional benefit is that some approaches help reduce motion artifacts. There is resistance to their adoption, however, due to practical considerations, such as claustrophobia and nausea and vomiting.

Electronic registration would consist of acquiring the followup volumes already in registration with the baseline scan by reorienting the acquisition planes. An approach to this end has been described by Oshio et al.⁴⁸, in which the authors use a noninvasive head frame with two ear pieces in the patient's auditory canals and a nose piece on the patient's nasion. The head frame provides standard fiducial reference points in a scouting image. Prior to acquisition of the final followup scan, the slices are reoriented so that they are in registration with the corresponding slices of the baseline scan. The authors report that error using this system is < 1 mm, which they attribute to uncertainty in the placement of the device, error in locating the reference points, and patient motion.⁴⁸ A possible alternative approach would be to use the patient's own brain as a basis for volumetric registration and reorientation of the slices prior to final image acquisition. At the time of the followup scan, a test scan could be acquired and the brain would be registered to the baseline scan. The acquisition of the final scan would be adjusted accordingly. It is likely that this would be a preferable technique to the head frame approaches described above, since it would require no additional hardware, would not introduce patient discomfort, and would hope-

fully result in a more accurate registration. In a comparison of registration strategies, it was demonstrated that noninvasive stereotactic head frames with fiducial markers performed poorly compared with intensity-based approaches.⁴⁹ Although Strother et al.⁴⁹ noted that the fiducial landmarks themselves could be aligned with high accuracy, they remarked that the low performance observed was attributable to the fact that the head holder is not completely stationary with regard to the skull, and the skull is not completely stationary with regard to the brain. Logically speaking, therefore, if the brain is what is of interest, an optimal approach would be one which acquires slices using the brain itself as the basis for alignment.

Acquisition Considerations

For the purposes of change detection, reducing slice thickness, eliminating interslice gaps, having isotropic voxels, and preregistration are all desirable. These are interrelated to some degree. For example, if one could perform true 3D volumetric acquisitions with very small voxels in all pulse sequences, and one was using sinc interpolation with an infinitely large window, this would, to a certain extent, obviate the need for preacquisition registration. In contrast, if one had perfect registration prior to acquisition so that voxels at each time represented the same portion of anatomy, there would be less need for isotropic voxels. Overall, the following factors should be maintained during serial scanning: equivalence of the scanner used, the pulse sequences used, the acquisition parameters, and the scanner software. Slice gaps should never be present because they represent areas in which the change detection algorithm will be completely unable to detect change, and more importantly they make it impossible to establish a one-to-one anatomical correspondence from one acquisition to the next. They also make it impossible to apply arbitrary linear or nonlinear spatial transformations to the images. Other acquisition-related changes, such as decay of the gradient coils and inhomogeneities, should be ameliorated to the greatest degree possible as these represent changes which confound the pathology-related changes of interest. If they are present in the images, then the algorithm

will be forced to attempt to remove them, and this will not always be possible. Identical administration of contrast (type, quantity, method, time before acquisition, etc.) should also be ensured.

Segmentation of Skull and Brain (Yielding Identical Anatomy in Each Scan)

The subsection "Subvoxel Registration" below discusses registration of the serial studies. Several structures (the ears, the lower mandible, the eyelids, the tongue, etc.) completely unrelated to the brain are acquired in a typical head MRI. They are not invariant and therefore have the potential to mislead the registration algorithm. Since several of the algorithms below require that each voxel represent exactly homologous portions of anatomy at both acquisitions (neglecting pathology-related changes), it is essential that these structures be prevented from misleading the registration algorithm. Additionally, in the subsection "Scaling Correction" below, scaling is performed based upon the inner table of the skull and therefore this must also be segmented. The actual act of segmentation will not be addressed in detail here; a variety of algorithms and approaches exist to perform this task with varying degrees of automation.⁵⁰

Scaling Correction

It is commonly recognized that gradient amplifiers of magnetic resonance imagers change over time, resulting in a corresponding scaling distortion of acquired images if proper quality control does not correct this. In order to detect changes associated with disease processes, it is essential that these changes in the gradient amplifier be reversed to prevent their effects from being misinterpreted as changes due to disease. A useful invariant for this purpose is the inner table of the skull.⁴ A standard linear registration algorithm could be used to compute the scaling factors.

Obviously, however, the preferred method would be for quality control personnel to prevent these changes from entering into the imaging studies by making adjustments to the scanner prior to acquisition, because applying a

scaling factor would require use of interpolation which, as discussed previously, almost invariably introduces errors. It should be noted that it is inadvisable to incorporate this scaling into the registration algorithm proper—scaling correction should be derived based upon a different part of the anatomy (inner table of the skull) than the rigid body registration (the brain itself). Derivation of a scaling correction transformation presumes that the structure upon which the correction is based is invariant over time. If this scaling correction were derived as a part of a registration algorithm based upon the brain itself, this assumption would be violated since the brain may change under a pathological process. If the scaling transformation were derived based upon the brain, the transformation could unwittingly undo the process the algorithms were created to measure. Likewise, construction of a rigid body registration algorithm should be based on the brain rather than on the inner table of the skull, because the position of the brain is not fixed with respect to the skull because of its suspension in CSF. On the other hand, it would be undesirable to apply interpolation twice since, as discussed previously, interpolation almost invariably introduces errors. Therefore, a preferable approach would be to begin by computing the scaling transformation. Then at each iteration of the registration algorithm (in the subvoxel registration step), the trial transformation matrix (containing rotation and translation components) would be multiplied by the previously computed scaling matrix. The unified transformation would be applied as one and the cost function computed.

Subvoxel Registration (Subvoxel Approach with Subregioning)

Examples of linear registration algorithms intended for serial studies which reportedly achieve subvoxel accuracy are described in the literature.²⁻⁴ Bosc et al.⁶ report using a nonlinear registration step following an initial linear registration step to correct for acquisition-related distortion of the brain. It is important to note that these algorithms assume identical underlying anatomy, which is exactly contrary

to the assumption that patients have changes occurring in their brain. Since the patient is the same in each case, large portions of anatomy are expected to be invariant; this is a great boon to the registration algorithm. However, since the subject under discussion is the detection of change, some regions of anatomy may not be invariant. A registration algorithm attempting to achieve subvoxel accuracy should therefore be able to base its transformation on only the portions of the anatomy that are invariant from the baseline to the followup scans. This can be done quite simply. When the volumes are out of alignment, the value of the cost function over all voxels will be fairly uniformly poor. As the volumes become aligned, the values of the cost function for more and more voxels will be more and more optimal. When the brains are perfectly aligned (assuming invariant anatomy), the value of the cost function at all voxels will be expected to occupy a distribution about some optimal level. Image noise will obviously prevent it from ever becoming a singleton. In the case of disease process-induced anatomy changes, a population of voxels will approach a distribution about the optimum level when the two volumes are registered, while a population of voxels will remain as outliers. In this case, these outlier voxels may be excluded from the overall cost function computation at each trial. The volume-wide cost function at each iteration could then be computed as a combination of the number of voxels included in the cost function (which should be maximized) and the traditional metric such as the sum of the square of the differences (which should be minimized).

Sinc Interpolation

Once the scaling and linear registration transformations have been computed, they should be combined, as discussed above, in order that at most one interpolation procedure be applied to the data. The transformation should be effected using an algorithm that introduces as little artifact as possible. Sinc interpolation^{2,3} could be used in three dimensions if the volumes were acquired in true 3D, or sinc interpolation could be applied in-plane with another algorithm out of plane (an interpolation algorithm which does not possess the

band-limited constraint of sinc interpolation, such as trilinear interpolation). If computational considerations were a factor, fast sinc could be used instead,⁴ since, with the use of a head fixation device and preacquisition registration, the required transformation should be small. If a volumetric intensity-based preacquisition registration approach is used to acquire the scans in registration (as described in the subsection "Preacquisition Registration of Scans" above), the step of interpolation may not be necessary at all.

Inhomogeneity Correction (Serial Imaging Strategy)

As mentioned previously, the availability of serial imaging studies affords a great opportunity, since major portions of the anatomy will remain invariant from one scan to the next. Therefore, applying knowledge of the ways in which inhomogeneities are likely to manifest, the baseline scan may be used as a standard in order to detect inhomogeneities in the followup scan and viceversa.⁶ If inhomogeneity correction parameters were being optimized at the same time as linear registration parameters (as described above), the invariant structures could be detected for both purposes and the simultaneous correction of inhomogeneities would aid the registration algorithm. These two processes would be symbiotic.

Change Computation

The step of change computation relates to conversion of corrected image volumes into measures of change. This process is somewhat related to classification in the sense that it corresponds to a process of assigning each voxel in an image to various classes, depending upon the voxel's location in feature space. There are important differences, however. In classification, what is of interest is static membership in particular tissues/clusters. In change detection, two things are of interest: what tissue(s) a particular voxel contains (similar to classification) and, just as important, how that voxel moves through feature space, from one scan to the next.

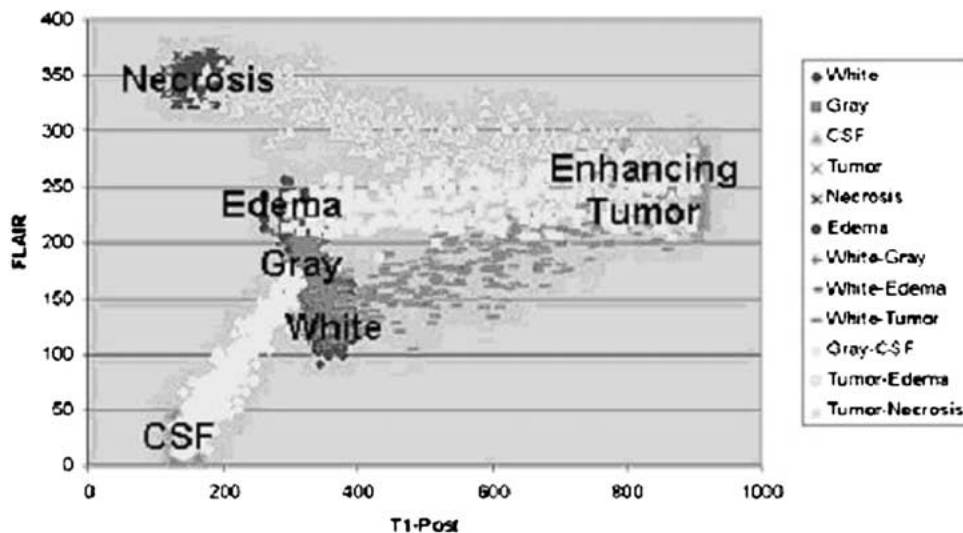


Fig 2. The gray matter cluster underlies the white edema line of transition. Using a static method, such as classify-subtract, voxels moving along this transition would appear to change dramatically in gray matter membership. By focusing on direction of movement and using knowledge of the way transitions occur, rather than static feature space location at each time point, the problem of spurious gray matter change is obviated.

An obvious approach to the process of change computation could consist of the comparison of two static classified membership volumes, essentially by subtraction. This is a valid way to approach this problem; however, great advantages result from considering both images together as a 4D dataset, specifically with regard to the management of noise. Considering the data as a multiple acquisition/pulse sequence volume also leads to a series of realizations, mostly relating to the availability of important knowledge regarding the problem. In classification, knowledge is available that voxel locations in feature space tend to center around clusters, each of which corresponds to a tissue. In change detection, it is analogously important to recognize that change tends to occur along lines in feature space, between particular pairs of cluster centroids. This knowledge is important because it allows both the reduction of noise and the simultaneous prevention of problems before they enter the data. In most voxels it can be assumed that at most two tissues are present. With this assumption in hand, voxels can be projected onto the line connecting the two centroids involved, under the assumption that any deviation from this line results solely from noise. Knowledge of what classes of transitions are physically possible results in a

Table 1. Current Methods of Change Detection and Quantification

Method	Subtypes
Manual inspection	Without subtraction
	With subtraction
Measurement sampling	Maximum diameter
	Maximum area
Volumetrics	Manual outlining
	Classification
	Surface modeling
Warping	
Temporal analysis	

reduction of the number of possible “lines” in feature space to a manageable number, and the availability of two acquisitions helps to disambiguate which line is relevant for a particular voxel, in the face of noise. As mentioned, there are some very specific problems which this approach helps to prevent from ever entering into the data. An important example is given by edema/white matter partial-volumed voxels, which are located in feature space quite close to the gray matter cluster in brain MRI. With the knowledge that static voxel location in feature space is not relevant, only the direction of motion is, this problem is overcome by avoidance (Fig 2).

A simultaneous thread of logic involves the quantification of change. Since partial-volume (and mixing) effects manifest as linear combinations of intensities over each pulse sequence, in order to reverse the linear combination and determine the relative proportions of the constituent components, a Euclidean classifier is necessary. In particular, when more than two tissues are considered, Euclidean classifiers are very susceptible to noise (a voxel will likely end up having nonnegligible membership in all tissues). However, the domain-specific knowledge just described results in the ability to narrow the consideration to at most two tissues per voxel. The problem of change computation therefore consists of two interrelated problems.

The first is the determination of which of a short list of transition types each voxel contains, which is an inherently crisp problem. The second is the determination of distance traveled through feature space of the voxel, in the direction of that line, which is inherently a fuzzy problem.

Significant Region Detection

At this point, data has been transformed from imaging data to change data. Inevitably, noise will result in false positives—voxels that appear to have changed but actually have not. A method is needed, therefore, to reduce the number of false positives. This method may be quite simple. For example, thresholding would discard all changes smaller than a certain size. A more appealing approach could take into account the fact that changes rarely occur in isolation; when real changes occur, they usually manifest as a group of spatially contiguous voxels undergoing the same type of change. An implementation of this type could require not only that a voxel be above a certain magnitude but that a minimum number of its neighbors also be above that magnitude.⁵¹ Another approach used in the field of fMRI for allowing groups of spatially colocated voxels to reinforce each other involves using intelligent filters to “smear” large magnitude changes into their neighbors⁵² (however, this results in errors, both false positive and false negative, as one might expect). Another approach is to use the likelihood ratio to test whether a group of

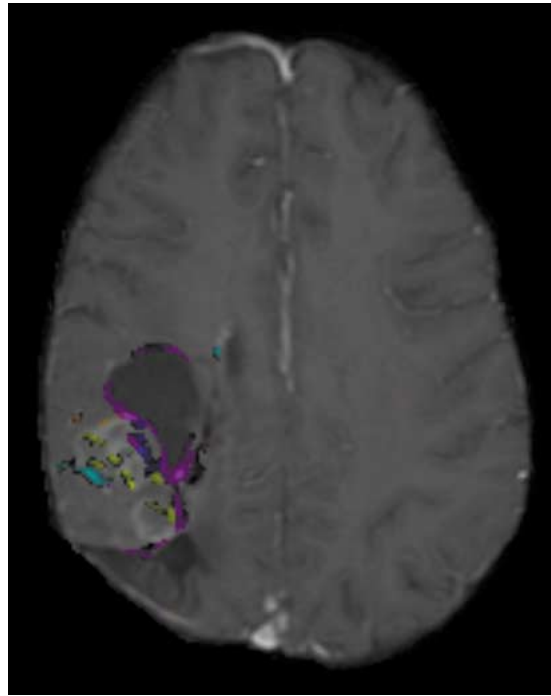


Fig 3. A sample change detection image. Type of change is encoded by color; magnitude of change is encoded by the intensity.

voxels is changing,⁵³ which allows smaller clusters to be detected as change as long as their magnitude is sufficiently high, as well as larger clusters to be detected as change with a smaller change requirement. A threshold based upon cluster size is not only able to separate changes of large magnitude from noise, but also separate changes of much smaller magnitude consisting of spatially contiguous groups of voxels undergoing the same type of change.

Presentation of Results

The change membership volume generated in the preceding step could be masked by the significant region detection generated in this step. The changes remaining—those computed to be significant—could be displayed in a volume, where each change type could be color-coded. The magnitude of the change in each voxel could be represented by the intensity of the color. The colors could be superimposed upon an anatomical image to help the radiologist orient him or herself. In addition, quantitative summations

could be made for each change type by adding the change membership of each type over the entire volume and multiplying by the volume of each voxel. This would yield a volumetric representation of how much white matter had become enhancing, how much enhancing tumor had become necrotic, and soon, based upon fuzzy metrics. From this, net change in each tissue could also be computed (Fig 3).

DISCUSSION AND SUMMARY

The field of change detection has been described in which computational methods assimilate multiple image types from two acquisitions in a series and determine which regions have changed, in what way, and to what degree. It would seem to be an optimum use of human and computer resources if the task of acquisition to-acquisition comparison and quantitative analysis should be assigned to the computer, to which these tasks come naturally. It is the process of interpretation and judgment of those quantitative results which should be given to the radiologist. Computational change detection methods promise, to some degree, to relieve the radiologist of "slice/information overload", while simultaneously improving his or her ability to interpret the data, in terms of both acuity and speed.

REFERENCES

- Filipek, PA, Kennedy, DN, Caviness, VS: Volumetric analyses of central nervous system neoplasm based on MRI. *Pediatr Neurol* 7:347-351, 1991
- Hajnal, JV, Saeed, N, Oatridge, A, et al: Detection of subtle brain changes using subvoxel registration and subtraction of serial MR images. *J Comput Assist Tomogr* 19:677-691, 1995
- Hajnal, JV, Saeed, N, Soar, EJ, et al: A registration and interpolation procedure for subvoxel matching of serially acquired MR images. *J Comput Assist Tomogr* 19:289-296, 1995
- Freeborough, PA, Woods, RP, Fox, NC: Accurate registration of serial 3D MR brain images and its application to visualizing change in neurodegenerative disorders. *J Comput Assist Tomogr* 20:1012-1022, 1996
- Rusinek, H, Leon, MJ de , George, AE, et al: Alzheimer disease: measuring loss of cerebral gray matter with MR imaging. *Radiology* 178:109-114, 1991
- Bosc, M, Heitz, F, Armspach, JP, et al: Automatic change detection in multimodal serial MRI: application to multiple sclerosis lesion evolution. *Neuroimage* 20:643-656, 2003
- Ettinger GJ, Grimson WE, Lozano-Perez T, et al.: Automatic registration for multiple sclerosis change detection. IEEE Workshop on Biomedical Image Analysis, Seattle, WA, 1994
- Miller, AB, Hoogstraten, B, Staquet, M, et al: Reporting results of cancer treatment. *Cancer* 47:207-214, 1981
- Green, S, Weiss, G: Southwest oncology group standard response criteria, endpoint definitions and toxicity criteria. *Invest New Drugs* 10:239-253, 1992
- Padhani, AR, Ollivier, L: The RECIST (Response Evaluation Criteria in Solid Tumors) criteria: implications for diagnostic radiologists. *Br J Radiol* 74:983-986, 2001
- Gehan, EA, Tefft, MC: Will there be resistance to the RECIST (Response Evaluation Criteria in Solid Tumors)? *J Natl Cancer Inst* 92:179-181, 2000
- Therasse, P, Arbuck, SG, Eisenhauer, EA, et al: New guidelines to evaluate the response to treatment in solid tumors. *J Natl Cancer Inst* 92:205-216, 2000
- Tsuchida, Y, Therasse, P: Response evaluation criteria in solid tumors (RECIST): new guidelines. *Med Pediatr Oncol* 37:1-3, 2001
- Padhani, AR, Husband, JE: Are current tumour response criteria relevant for the 21st century?. *Br J Radiol* 73:1031-1033, 2000
- Thiesse, P, Ollivier, L, Di Stefano-Louineau, D, et al: Response rate accuracy in oncology trials: reasons for interobserver variability. *J Clin Oncol* 15:3507-3514, 1997
- Clarke, LP, Velthuizen, RP, Clark, M, et al: MRI measurement of brain tumor response: comparison of visual metric and automatic segmentation. *Magn Reson Imaging* 16:271-279, 1998
- Chow, KL, Gobin, YP, Cloughesy, T, et al: Prognostic factors in recurrent glioblastoma multiforme and anaplastic astrocytoma treated with selective intra-arterial chemotherapy. *Am J Neuroradiol* 21:471-478, 2000
- Weiner, HL, Guttman, CR, Khoury, SJ, et al: Serial magnetic resonance imaging in multiple sclerosis: correlation with attacks, disability, and disease stage. *J Neuroimmunol* 104:164-173, 2000
- Jack, CR, Petersen, RC, Xu, Y, et al: Rate of medial temporal lobe atrophy in typical aging and Alzheimer's disease. *Neurology* 51:993-999, 1998
- Eaves, G: The invasive growth of malignant tumours as a purely mechanical process. *J Pathol* 109:233-237, 1973
- Kowalczyk, A, Macdonald, RL, Amidei, C, et al: Quantitative imaging study of extent of surgical resection and prognosis of malignant astrocytomas. *Neurosurgery* 41:1028-1036, 1997
- Hammoud, MA, Sawaya, R, Shi, W, et al: Prognostic significance of preoperative MRI scans in glioblastoma multiforme. *J Neurooncol* 27:65-73, 1996
- Kelly, PJ, Daumas-Duport, C, Kispert, DB, et al: Imaging-based stereotaxic serial biopsies in untreated intracranial glial neoplasms. *J Neurosurg* 66:865-874, 1987
- Kelly, PJ, Daumas-Duport, C, Scheithauer, BW, et al: Stereotactic histologic correlations of computed tomography- and magnetic resonance imaging-defined abnormalities in patients with glial neoplasms. *Mayo Clin Proc* 62:450-459, 1987

25. Burger, PC, Heinz, ER, Shibata, T, et al: Topographic anatomy and CT correlations in the untreated glioblastoma multiforme. *J Neurosurg* 68:698-704, 1988
26. Johnson, PC, Drayer, BP, Rigamonti, D, et al: Defining the extent of infiltrating glioblastoma multiforme (GBM): A comparison of postmortem magnetic resonance imaging (MRI) with histopathology. *J Neuropathol Exp Neurol* 46:389, 1987
27. Tovi, M, Hartman, M, Lilja, A, et al: MR imaging in cerebral gliomas. Tissue component analysis in correlation with histopathology of whole-brain specimens. *Acta Radiol* 35:495-505, 1994
28. Tovi, M: MR imaging in cerebral gliomas analysis of tumour tissue components. *Acta Radiol Suppl* 384:1-24, 1993
29. Haney, SM, Thompson, PM, Cloughesy, TF, et al: Mapping therapeutic response in a patient with malignant glioma. *J Comput Assist Tomogr* 25:529-536, 2001
30. Haney, SM, Thompson, PM, Cloughesy, TF, et al: Tracking tumor growth rates in patients with malignant gliomas: a test of two algorithms. *Am J Neuroradiol* 22:73-82, 2001
31. Rey, D, Subsol, G, Delingette, H, et al: Automatic detection and segmentation of evolving processes in 3D medical images: Application to multiple sclerosis. *Med Image Anal* 6:163-179, 2002
32. Thompson, PM, Giedd, JN, Woods, RP, et al: Growth patterns in the developing brain detected by using continuum mechanical tensor maps. *Nature* 404:190-193, 2000
33. Freeborough, PA, Fox, NC: Modeling brain deformations in Alzheimer disease by fluid registration of serial 3D MR images. *J Comput Assist Tomogr* 22:838-843, 1998
34. Kyriacou, SK, Davatzikos, C, Zinreich, SJ, et al: Nonlinear elastic registration of brain images with tumor pathology using a biomechanical model. *IEEE Trans Med Imaging* 18:580-592, 1999
35. Thirion, JP, Calmon, G: Deformation analysis to detect and quantify active lesions in three-dimensional medical image sequences. *IEEE Trans Med Imaging* 18:429-441, 1999
36. Thirion, J: Elastic matching: continuum mechanical and probabilistic analysis In: Toga, A ed *Brain Warping Academic Press, San Diego, CA, 1999*, pp 183-197
37. Gerig, G, Welti, D, Guttmann, CRG, et al: Exploring the Discrimination Power of the Time Domain for Segmentation and Characterization of Lesions in Serial MR Data Boston: MICCAI, 1998
38. Meier, DS, Guttmann, CR: Time-series analysis of MRI intensity patterns in multiple sclerosis. *Neuroimage* 20:1193-1209, 2003
39. Kikinis, R, Guttmann, CR, Metcalf, D, et al: Quantitative follow-up of patients with multiple sclerosis using MRI: technical aspects. *J Magn Reson Imaging* 9:519-530, 1999
40. Guttmann, CR, Kikinis, R, Anderson, MC, et al: Quantitative follow-up of patients with multiple sclerosis using MRI: reproducibility. *J Magn Reson Imaging* 9:509-518, 1999
41. Gawne-Cain, ML, Webb, S, Tofts, P, et al: Lesion volume measurement in multiple sclerosis: how important is accurate repositioning?. *J Magn Reson Imaging* 6:705-713, 1996
42. Pilipuf, MN, Goble, JC, Kassell, NF: A noninvasive thermoplastic head immobilization system. Technical note. *J Neurosurg* 82:1082-1085, 1995
43. Bettinardi, V, Scardaoni, R, Gilardi, MC, et al: Head holder for PET, CT, and MR studies. *J Comput Assist Tomogr* 15:886-892, 1991
44. Carini, S, Calcagno, E, Tortori-Donati, P, et al: A new model for non-invasive, reproducible fixation of a stereotaxic frame using an orthodontic resin plate. *Acta Neurochir* 118:159-161, 1992
45. Laitinen, L: Noninvasive multipurpose stereoadapter. *Neur Res* 9:137-141, 1987
46. Edward, V, Windischberger, C, Cunningham, R, et al: Quantification of fMRI artifact reduction by a novel plaster cast head holder. *Hum Brain Mapp* 11:207-213, 2000
47. Strother, S, Perlmutter, J: Headholders for functional brain imaging. *J Cereb Blood Flow Metab* 7:S16-S18, 1987
48. Oshio, K, Panych, LP, Guttmann, CR: A simple noninvasive stereotactic device for routine MR head examinations. *J Comput Assist Tomogr* 20:588-591, 1996
49. Strother, SC, Anderson, JR, Xu, XL, et al: Quantitative comparisons of image registration techniques based on high-resolution MRI of the brain. *J Comput Assist Tomogr* 18:954-962, 1994
50. Kikinis, R, Shenton, ME, Gerig, G, et al: Routine quantitative analysis of brain and cerebrospinal fluid spaces with MR imaging. *J Magn Reson Imaging* 2:619-629, 1992
51. Lemieux, L, Wieshmann, UC, Moran, NF, et al: The detection and significance of subtle changes in mixed-signal brain lesions by serial MRI scan matching and spatial normalization. *Med Image Anal* 2:227-242, 1998
52. Friman, O, Borga, M, Lundberg, P, et al: Detection of neural activity in fMRI using maximum correlation modeling. *Neuroimage* 15:386-395, 2002
53. Hsu, YZ, Nagel, HH, Rekers, G: New likelihood test methods for change detection in image sequences. *Comput Vision Graphics Image Process* 26:73-106, 1984