Published in final edited form as:

Acad Radiol. 2010 April ; 17(4): 479–488. doi:10.1016/j.acra.2009.10.026.

Computer-Assisted Quantitative Evaluation of Therapeutic Responses for Lymphoma Using Serial PET/CT Imaging

Xin Gao, Ph.D.¹, Zhong Xue, Ph.D.^{1,*}, Jiong Xing, MD¹, Daniel Y. Lee, MD, Ph.D.², Stephen M. Gottschalk, MD³, Helen H. Heslop, MD³, Catherine M. Bollard, MD³, and Stephen T.C. Wong, Ph.D.¹

¹ Center for Bioengineering and Informatics, The Methodist Hospital and The Methodist Hospital Research Institute, Weill Medical College of Cornell University, Houston TX

² Department of Radiology, The Methodist Hospital and The Methodist Hospital Research Institute, Weill Medical College of Cornell University, Houston TX

³ Department of Pediatrics, Section of Hematology/Oncology, Texas Children's Hospital, Baylor College of Medicine, Houston TX

Abstract

Rationale and Objectives—Molecular imaging modalities such as PET/CT have emerged as an essential diagnostic tool for monitoring treatment response in lymphoma patients. However, quantitative assessment of treatment outcomes from serial scans is often difficult, laborious, and time consuming. Automatic quantization of longitudinal PET/CT scans provides more efficient and comprehensive quantitative evaluation of cancer therapeutic responses. This study develops and validates a Longitudinal Image Navigation and Analysis (LINA) system for this quantitative imaging application.

Materials and Methods—LINA is designed to automatically construct longitudinal correspondence along serial images of individual patients for changes in tumor volume and metabolic activity via regions of interest (ROI) segmented from a given time-point image and propagated into the space of all follow-up PET/CT images. We applied LINA retrospectively to nine lymphoma patients enrolled in an immunotherapy clinical trial conducted at the Center for Cell and Gene Therapy, Baylor College of Medicine. This methodology was compared to the readout by a diagnostic radiologist, who manually measured the ROI metabolic activity as defined by the maximal Standardized Uptake Value (SUVmax).

Results—Quantitative results showed that the measured SUVs obtained from automatic mapping are as accurate as semi-automatic segmentation and consistent with clinical examination finding. The average of relative squared differences of SUVmax between automatic and semi-automatic segmentation was found to be 0.02.

Conclusion—These data support a role for LINA in facilitating quantitative analysis of serial PET/CT images to efficiently assess cancer treatment responses in a comprehensive and intuitive software platform.

Zhong Xue, PhD, 6565 Fannin ST, B5-018, Houston TX, Tel: 713-4412577, Fax: 713-4418696, zxue@tmhs.org.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

Lymphoma; quantitative evaluation of treatment outcomes; PET/CT; longitudinal registration of serial images

Lymphoma is a hematologic malignancy of lymphocyte origin, accounting for approximately 5% of all cancers in the United States (1). This diverse group of diseases is broadly classified as Hodgkin's disease (HD) or non-Hodgkin's lymphoma (NHL) with several sub-classification schemes to describe various cellular, genetic, and clinical subtypes (2). Treatment for lymphoma is dependent on its type and stage, as well as the age and general clinical status of the patient. For most early-stage lymphomas, the standard first-line treatment for lymphoma includes chemotherapy with or without radiation therapy (3). Immunotherapy, radioimmunotherapy, and hematopoietic stem-cell transplantation have added to the therapeutic armamentarium especially for patients with aggressive, recurrent, or advanced disease. Accurate monitoring of patients undergoing any form of treatment is critical for evaluating response and disease recurrence. However, the diagnostic gold standard, tissue biopsy, is both invasive and logistically difficult to perform on all patients at various time intervals.

Currently, treatment responses are assessed by a canonical approach that integrates clinical examination, laboratory findings, and imaging data. PET/CT combining metabolic information using the positron-emitting sugar analog, [F-18]-fluorodeoxyglucose (FDG), with morphologic changes in tumor size (4–12), has been well studied in the evaluation of treatment response and predictive value in patients with lymphoma (11–20). Combined modality PET/CT has been shown to improve the diagnostic accuracy over either modality alone, particularly for lymphoma (21–23). Relative FDG uptake in lesions of active high grade lymphomas are typically high thus allowing for quantitative assessment of disease status (24). Despite the growing clinical imaging knowledgebase, interpretation of PET/CT studies is largely dependent on imprecise criteria for measuring mass lesions and relative tumor metabolism, both of which are manually acquired. Moreover, multiple PET/CT scans obtained over time generate massive amounts of data, most of which are analyzed manually adding significant inefficiency and risk for errors in the interpretation process. Clinical investigators are thus hampered by the absence of validated quantitative methods for evaluating therapy response in an efficient and robust format.

Although computer-aided methods for lymphoma segmentation (25,26) and longitudinal CT analysis exist (27), these methods apply global affine transformation or pairwise free-form deformable (FFD) registration for processing longitudinal image data and there is a lack of longitudinal stability in the quantitative analysis. Traditional pairwise (28,29) and groupwise (30) image registration algorithms have been used for image alignment, however the pairwise algorithms warp each image separately and often cause relatively unstable measures of the serial images because no temporal information of the serial images has been used in the registration procedure. Groupwise image registration methods simultaneously process multiple images but consider the images as a group, not a time series. Thus the temporal information has not been used efficiently. In fact, for longitudinal images, the relationship between temporally neighboring images is much more important than that of the images with larger time intervals.

In this work the recently developed joint serial image registration and segmentation algorithm for longitudinal CT data (31) is used in the computer-assisted quantitative analysis for lymphoma treatment monitoring. This Longitudinal Image Navigation and Analysis (LINA) software tool facilitates the quantitative evaluation of treatment outcomes for lymphoma

patients using a computer-assisted serial image analysis approach and automatically constructs the longitudinal correspondences along serial images for each individual patient. In this way, it is possible to automatically determine and track ROIs in the serial images after defining them at one time-point using a semi-automated segmentation algorithm.

In experiments, we applied the proposed methodology to the datasets of patients with lymphoma in the clinical trial investigating LMP-specific CTLs in patients with relapsed EBV-positive lymphoma (ClinicalTrials.gov Identifier: NCT00671164). Two sets of experiments were performed to validate the method. In the first experiment we tested the accuracy of the registration method using simulated serial lung CT images with known ground truth about the deformation. The results show that the average error using the longitudinal deformable registration on 10 randomly simulated images was 3.3 mm. The second experiment calculated the errors of all the maximum of Standardized Uptake Value (SUVmax) within each ROI determined by LINA and the semi-automatic/manual segmentation method, respectively. For all the ROIs of the nine patients studied in this paper, the average of relative squared differences of SUVmax between the automatic and the semi-automatic segmentation results was 0.02. Based on these measures, the longitudinal quantitative index curves showing treatment responses using the proposed method are also consistent with the semi-automatic/manual method. The proposed serial image quantification approach and the LINA toolkit can therefore supply more accurate, complete and intuitive assessments of the treatment responses in patients with lymphoma after CTL therapy.

Methods

Data

Longitudinal PET/CT images of nine patients in the clinical trial, administering LMP-specific cytotoxic T-lymphocytes to patients with relapsed EBV-positive lymphoma, were acquired in an integrated PET/CT system (GE Discovery ST) with conventional dosing of [F-18] FDG (0.21 mCi/kg). Serial images of patients consisted of 2 to 6 complete PET/CT datasets obtained at multiple time points. The voxel size of CT images was $0.98\text{mm} \times 0.98\text{mm} \times 3.75\text{mm}$, and that of PET was $5.47\text{mm} \times 5.47\text{mm} \times 3.75\text{mm}$.

Patient

The clinical protocol was approved by the Food and Drug Administration (FDA), the Recombinant DNA Advisory Committee and the hospitals' Institutional Review Boards. Nine patients had active disease at the time of CTL therapy. Details of the selection criteria and patient population are described by Bollard *et al.* (32).

Computer-assisted quantitative analysis

A flowchart describing the methodology for the computer-assisted quantitative analysis is shown in Fig. 1. We developed a LINA software tool to semi-automatically segment lesions, automatically calculate the longitudinal deformations for each image series, and map the ROIs from one time-point onto all other time-points. Both the volumetric shape changes and longitudinal quantitative PET measurements of each ROI can be visualized using LINA. The detailed steps are described in detail as follows.

Co-registration of PET and CT images—Precise identification of the lymphoma region also requires the aid of CT image, which supply information regarding the precise anatomical structure and the lymphoma boundary. To facilitate their interpretation, it is necessary to co-register PET and CT images, thereby relating the metabolic activity (uptake pattern) from the PET image to the morphologic information provided by CT. PET/CT images were co-registered based on maximization of Normalized Mutual Information (NMI) using a global

optimization method developed by M. Jenkinson *et al.* (33). Rigid transformation and tri-linear interpolation were used in the multi-resolution implementation of the algorithm by combining a fast local optimization Powell's method (34). Denoting the transformation between the CT I_t and PET P_t images at time-point t as A_t , a voxel v in the CT image will correspond to voxel $u = A_t(v)$ in the PET image. These transformation matrices were stored for constructing the longitudinal correspondences among the PET images by combining with the longitudinal deformations.

Segmentation of ROI—To quantitatively analyze the treatment response in lymphoma patients, who received adoptive CTL therapy, ROIs of mass lesions needed to be visually identified in PET/CT images. These ROIs included lymph nodes and other lymphoid organs with increased PET activity. In our application, we segment two types of ROIs: (1) shapes or boundaries of ROIs that are clearly detectable, e.g., discrete lymph nodes; and (2) ROIs without discernable boundaries. We therefore used a level set-based semi-automatic ROI segmentation for this study. For the ROIs with clear boundaries on CT, we used the image boundary-based level set algorithm (35) to extract the ROI shapes with a proper initialization; for ROIs without discernable boundaries manual segmentation will be used if the level set method fails.

Determining longitudinal deformation of serial CT images—Even with the help of the semi-automatic segmentation tool, the extraction of lymphoma regions is laborious, especially for patients requiring extensive serial imaging with a relatively large number of mass lesions (ROIs) in many sets of PET/CT images. Moreover, in order to calculate the longitudinal PET activity changes of each ROI, it is also necessary to construct their temporal correspondences. Generally, the images taken at the first time-point are chosen as the baseline datasets, and the ROIs (possible lymphoma mass regions and the liver region) are semi-automatically segmented or manually marked from the fused PET/CT images at the baseline. Since the longitudinal deformations among each image series are known, these segmented ROIs can be deformed onto the images at other time-points so that all corresponding ROIs in the follow-up images can be automatically extracted. Similarly, any new ROI that is only observable in a follow-up image can also be marked or segmented, which can then be automatically mapped onto the baseline and other time-point images.

We developed a serial CT images registration algorithm (31) to automatically align all the CT images of the same patient at different time-points to automatically calculate the longitudinal correspondences. For serial image registration, the relationship between temporally neighboring images is much more important than that of the images with larger time intervals, since both anatomical structure and tissue properties of neighboring images tend to be more similar for neighboring images than others; moreover these temporal changes might be characterized using specific physical processes models. Therefore, we formulated the serial image registration so that the registration of the current time-point image is related to not only the previous but also the following images (if available). Given a series of CT images I_t , $t = 0, 1, \dots, T$ (I_0 is called the baseline) all the subsequent images were first globally aligned onto the space of the baseline by applying the rigid registration (33). Thus matrix $R_{0 \rightarrow t}$ or R_t will reflect the transformation from time-point 0 to time-point t . In addition to the global transformation we needed to estimate the deformations from the baseline onto each image, i.e., $\mathbf{f}_{0 \rightarrow t}$ or simply denoted as $\mathbf{f}_t$. For the current CT image I_t , if the deformation of its previous image I_{t-1} , i.e., $\mathbf{f}_{t-1}$, and that of the next image I_{t+1} , $\mathbf{f}_{t+1}$, are known, the deformation $\mathbf{f}_t$ can be calculated by jointly considering both the previous and the next images and by minimizing,

$$\begin{aligned}
E_t = & \sum_{v \in \Omega} \{ |e[I_t(\mathbf{f}_t(R_t(v)))] \\
& - e[I_0(v)]|^2 \\
& + \sum_{i=-1,1} |e[I_t(\mathbf{f}_t(R_t(v)))] - e[I_{t+i}(\mathbf{f}_{t+i}(R_{t+i}(v)))]|^2 \},
\end{aligned} \tag{1}$$

where $e[]$ is the operator for calculating the features around each voxel of the image, and Ω is the image domain. In this work, the feature vector for each voxel consisted of the image intensity, gradient magnitude, and the fuzzy membership functions obtained by performing a 4-D fuzzy c-mean (FCM) algorithm on the images (assuming three tissue types: bone, low-intensity and high-intensity tissues), i.e., $e[v] = [I(v), |\nabla I(v)|, \mu_{v,1}, \mu_{v,2}, \mu_{v,3}]$. Since the cubic B-Spline was used to model the deformation field, the continuity and smoothness was guaranteed, and the smoothness regularization term of the deformation field was omitted. Further, we used a topological regularization step to ensure that the Jacobian determinants of the deformations fields were positive. Thus the topology of the deformation field did not change from one image onto the subsequent images. The serial image registration algorithm then iteratively refines the deformation field $\mathbf{f}_t$ of each time-point image by minimizing the energy function in Eq.(1) and performs 4-D clustering of the image series until convergence. Notice that in the first iteration, since the registration results for neighboring images were not available, only the first term of Eq.(1) was used, which is essentially a pairwise FFD (36).

Automatic ROI delineation and quantitative indexing—After global co-registration and longitudinal deformable registration, correspondences among the CT image series were constructed, and all the ROIs segmented at one time-point image were automatically mapped onto other time-points. Suppose an ROI r_0 has been obtained from the baseline CT image, the corresponding region in the baseline PET image will be $A_0(r_0)$. The corresponding region in the PET image at time-point t will be $r_t = A_t[\mathbf{f}_t(R_t(r_0))]$, where the global transformation from time-point 0 to time-point t , R_t , is first applied, followed by a deformable transformation $\mathbf{f}_t$, and finally the deformed ROI is transformed onto the PET space at time point t using A_t . If the ROI is not segmented at the baseline, it can also be automatically mapped onto other time-points in a similar combined global and deformable transformation manner. The reason that we combine the global transformations between CT and PET images, the global transformations between the baseline CT image and other time-point images, as well as the longitudinal deformation fields for the globally aligned CT images for mapping the ROI is that no interpolation on the PET images need to be performed in order to obtain accurate SUV measures from PET images.

Our next step is to quantify the signals in PET image, given each region r_t at time t . In PET imaging, a radioactive material (e.g. FDG) is systemically administered by intravenous injection and allowed a period of cellular uptake of the compound (up to 40 minutes). Following this period, the patient is scanned in an instrument that detects positron annihilation 511 keV photons. FDG accumulates in cell with through a mechanism involving glucose transport and phosphorylation by hexokinase, rendering it trapped inside cells. Thus FDG concentration is proportional to the first biochemical steps in glucose metabolism. One widely applied method for measuring FDG activity is the maximum of Standardized Uptake Value (SUVmax) (37–42), which relates the maximal value of activity concentration found in a certain tissue to the injected activity per the patient body weight. The SUV of a voxel v in region r_t is calculated using attenuation-corrected images, injected dose of FDG and FMT, patient's body weight, and the cross calibration factor between PET and dose calibration (43) and is defined as follows:

$$SUV(v) = \frac{\text{Radioactive concentration at } v \text{ (MBq/g)}}{\text{Injected dose (MBq)/Patient body weight (g)}}. \quad (2)$$

As expected, higher SUV corresponds to greater metabolic activity. There is some evidence (41,44) suggesting that an SUV greater than 2.5 indicates metastatic disease identified as suspected tumor sites. However due to the inter-individual differences and diffusion ability of radioactivity at different time points, a relative SUV value in an ROI should be compared with those in normal brain, liver or heart organ region. Therefore, it is suggested that quantitative evaluation to determine the response to therapy in patients with lymphoma needs to refer to the SUV in normal and tumor tissue since different image reconstruction, processing techniques, patient conditions, and dose of radioactive drug may affect the specific numerical value of the SUV.

In our study, the ratio of maximum of SUV in an ROI vs. the mean SUV of a liver region is employed as the Quantitative Index (QI) to quantify the therapeutic effect, which can be expressed by the following equation,

$$QI(r_i) = \frac{SUV_{\max}}{SUV_{\text{mean}}} = \frac{\max_{v \in r_i} \{SUV(v)\}}{\text{mean}(SUV(\text{liver}_i))}. \quad (3)$$

The IRB is approved from our institute to conduct this quantitative study.

Results

The LINA toolkit has been applied to lymphoma patient follow-up studies in the experiments. Patients included 6 men and 3 women. The median age was 29 (range 20–63). The total number of lesions identified on imaging was 34 and the mean number of lesions per patient was 3.7 (range, 1–7 lesions). In the first experiment, we evaluated the accuracy of the longitudinal registration algorithm using chest lung CT images, and in the second experiment, the quantitative follow-up results of the proposed method are compared with the manual results.

Quantitative evaluation of the algorithm

The proposed longitudinal deformable registration algorithm has been evaluated by using randomly simulated lung CT images. First, for each series of the sample images, the demon's algorithm (45) was used to register the subsequent images onto the first time-point image. Then, one sample image was selected as the template image, and all the other first time-point images of other subjects were deformed onto the template space using the same deformable registration algorithm. Finally the statistical model for the spatially warped/normalized temporal deformations were calculated, which was then used to randomly generate the longitudinal deformation of the template image. In this paper, the statistical model of deformation (SMD) presented in (46) was used to train such a model. In this experiment we used the chest CT images obtained from The Methodist Hospital as our datasets and the resolution is the same as our patient data. One image was selected as the template and other 19 images were selected as the training samples, and ten CT images were randomly simulated. In a similar way, the longitudinal changes of these CT images can then be simulated and applied to the ten CT images (four serial images each). The registration errors with the ground truth were calculated reflecting the accuracy of the registration:

$$\text{Registration Error} = \frac{1}{N|\Omega|} \sum_{i=1, \dots, N} \sum_{v \in \Omega} |\mathbf{f}_{i,v} - \mathbf{f}_{i,v}^*|, \quad (4)$$

where $N=10$ is the number of testing images, and $|\Omega|$ represents the number of voxels in the template image domain. $\mathbf{f}_{i,v}$ is the resultant deformation at voxel v for testing image i , and $\mathbf{f}_{i,v}^*$ is the corresponding ground truth from the simulated deformation fields.

In order to evaluate the registration accuracy, we compared the results using the LINA algorithm and the FFD-based registration (used in (25)). The distribution of the registration errors for the LINA algorithm and the FFD-based registration algorithm are plotted in Fig. 2. The mean registration error for LINA was 3.3mm , and the mean registration error for FFD-based registration was 4.6mm . The results showed that the registration errors were found in less than the largest voxel spacing of the simulated images. Compared to the voxel size of the PET image, $5.47\text{mm} \times 5.47\text{mm} \times 3.75\text{mm}$. Therefore we can achieve sub-voxel average registration accuracy, which is accurate enough to calculate quantitative measures from PET images. As for the segmentation error, the average overlapping between the automatically segmented ROIs and the manually marked ROIs are $\sim 85\%$, indicating relatively high accuracy in terms of ROI segmentation and automatic mapping.

In the second experiment we compared the proposed automatic ROI mapping based on longitudinal registration with the semi-automatic/manual segmentation methods for ROI. First, the FLIRT program (47) was used to co-register the PET/CT images at same time-point. For each CT image, a radiologist delineated all the ROIs using our semi-automatic segmentation tools. In the automatic ROI mapping method, the segmented ROIs from the baseline were automatically mapped onto all the CT images in the follow-up time-points. Finally, the SUVs of each ROI were calculated in the original PET image space by transforming the corresponding regions to PET, so that no interpolation of the PET images is performed.

For a dataset with PET (image size $128 \times 128 \times 300$) and CT (image size $512 \times 512 \times 300$), the average time for co-registration was about 8 minutes on HP xw4400 workstation by using the FLIRT program. The time for deformable registration of two CT images was dependent on the sub-volume to be registered. For sub-volumes cut from the CT with size $256 \times 256 \times 70$, the calculation time is less than 200 seconds.

We compared the results of automatic ROI mapping with the original semi-automatic/manual segmentation results, using SUV values of each region. The mean of the relative squared differences between two SUVmax values for all the regions of all the patients are calculated as follows,

$$\text{error} = \sum_{i \in N} \frac{(SUV_{\max}(r_i) - SUV_{\max}(\hat{r}_i))^2}{(SUV_{\max}(r_i))^2} / N, \quad (5)$$

where r_i and $\hat{r}_i$ denote the segmentation results using the semi-automatic/manual segmentation and the proposed automatic mapping for the i th, respectively. N is the total number of ROIs under evaluation. For all the ROIs of the nine patients we studied, this average normalized squared difference is 0.02. Therefore, the computer-assisted quantitative analysis method obtains similar results with the “gold standard.”

Segmentation and visualization of longitudinal ROI mapping

First, the semi-automatic and manual ROI segmentation are illustrated in Fig. 3 and Fig. 4, respectively. From Fig. 3(a) we can visually identify the region of interest, and then manual points are initialized inside the lymph nodes (Fig. 3(b)), and the algorithm automatically extracts the boundaries of the lymph nodes as shown in Fig. 3(c). When the boundaries of an ROI are blurry, e.g., lymph nodes are connected with muscles, the level set algorithm normally leaks into a larger region, and a manual segmentation of the ROI is then applied. From Fig. 3 (a) we identified a region of interest but from the initial point shown in Fig. 3(b), the level set segmentation leaks into a larger region in the CT image. Therefore manual segmentation of the ROI is necessary. Fig. 4(c) shows the manually marked region.

It is worth noting that the PET signal can also be used to facilitate the segmentation process. For example the level set evolution can stop to grow when the PET signal is weak. However, since the PET signal is the one that we need to measure, it will introduce biased quantitative calculation if the selection of ROI is dependent on the PET signal itself. Thus in this study we used the strategy that the ROIs are only selected from the CT images.

Based on the proposed algorithm, a longitudinal imaging navigation and analysis tool (LINA) is developed. The key functions include a sophisticated data management system to facilitate quantitative longitudinal studies and statistical analysis; an exceptional 4-D Viewer of longitudinal images with underlying longitudinal deformations and 4-D ROIs; and the quantitative results for the follow-up study of each patient. In this section, we introduce the visualization of some results. Compared to the radiological reports, the new tool will provide a platform for comprehensive dataset management.

Fig. 5 shows an example of the results with two time-points, and the lesion is in the lung. Fig. 5(a) and (b) show the fused PET/CT images at the baseline and the second time-point, respectively. Fig. 5(c) and (d) show the semi-automatic segmentation of the lesion, respectively. After deformable registration, the baseline image can be warped onto the image at second time-point, and Fig. 4(e) shows the registered result, in which the green curve illustrates the automatically transformed ROI from the baseline CT. Comparing Fig. 5(d) and Fig. 5(e), we can see that after warping the baseline image can be very similar with the second time-point image, and the shape and volume of the transformed ROI is also very similar to the semi-automatically determined ROI at the second time-point. Regarding SUVmax values at the second time-point, the semi-automatic results for the lesion and liver are 2.04 and 1.81, separately, and the SUVs of the automatic mapping results at second time-point are the same. This is because the manual and automatically mapped ROIs are highly overlapping. The SUVmax consistency of automatic and semi-automatic segmentation results indicates that the automatic segmentation method is effective for SUV calculation of ROI. In addition, QI of the ROI at the two time-points are 2.03 and 1.13 respectively, which indicate the treatment outcome of lymphoma. This decrease demonstrates that treatment had a good effect on lymphoma.

Fig. 6 shows another example with four time-points, where lymph nodes can be clearly segmented. The first row shows the fused PET/CT images, where lymphoma region has obvious high intensity. The CT images are shown in the second row, and the yellow curves illustrate the semi-automatic lymph node segmentation results. For the follow-up CT images, we applied the serial image registration algorithm and warped the baseline image onto each subsequent CT image shown in row 3. Correspondingly, the ROI at baseline can be transformed onto these CT images, shown as the green contours in row 3. At each follow-up time-point, the warped baseline image appears similar to the CT image captured at that time-point. By using the automatically mapped ROIs, SUVmax and QI measurements can be automatically generated for each case. These data can be used to quantitatively assess metabolic activity in specific lesions over multiple imaging datasets which is critical for evaluating therapy

response. From the QI values given, we can see the positive response of the treatment quantitatively in this case.

After calculation of QI (the ratio of maximal SUV of lesion to the mean SUV of the liver), they can be plotted longitudinally and supplies as quantitative response of treatment. Fig. 7 shows some results in our datasets. Each figure shows the plots of QIs of all ROIs of one patient at different time-points. For the same longitudinally corresponding ROI, we used lines with the same color to link them, and solid lines show the semi-automatic/manual results and dashed lines indicate the results of the automatic mapping results. The big solid circle markers indicate that semi-automatic/manual segmentations at that time-point are automatically mapped onto other time-points for automatic SUV calculation.

Fig. 7(a) and (b) show two examples where semi-automatic segmentation of ROI is performed from the baseline. By comparing the solid lines with the dashed lines, it can be seen that the longitudinal QI values obtained by using both methods are in reasonable agreement. The longitudinal change of QI of each ROI indicates that these ROIs are obvious response to treatment within the first five and half months, while the situation might become worse between 5.5 to 8.5 months. In Fig. 7(b) the computing results of another patient with one ROI at three time-points are shown. Similarly, we see very close measure between the two methods, and the longitudinal plots of the QI values suggest a positive response of the treatment.

Fig. 7(c) and (d) show two examples where it was difficult to identify ROIs in some images. Fig. 7(c) shows the results of one patient with four ROIs at five time-points, in which it is difficult to identify the ROIs at the fifth time-point except for one. All the lesions are determined from the baseline image and the computer-assisted method can determine the corresponding lesions for all the follow-up studies, while it is difficult to even manually mark the lesion for some scans. The results show that our method can delineate other three ROIs at the fifth time-point by using serial CT registration, which assists oncologist to quantitatively evaluate the treatment response. In Fig. 7(d), two ROIs can be detected at the first time-point (blue and red point markers), and the other two ROIs indicated by green and magenta markers can be identified at the second time-point. By using the proposed computer-assisted quantitative approach, we can automatically calculate the SUV values of along the image series and hence prove a comprehensive quantitative evaluation of the treatment outcomes of lymphoma. Fig. 7(c) and (d) indicate that our proposed method enable to provide pro- and post-perspective analysis for treatment outcomes of lymphoma.

From the experimental results, one can see clearly that this method has the following advantages. Since lymphomas are often gathered together in some patients, it is difficult to map one-to-one correspondence among the lymphoma regions segmented by hand at different time-points. Our method is able to automatically segmented ROIs from follow-up CT images by using ROI mapping based on the longitudinal correspondences, and it is more efficient to accomplish the goal of quantitative analysis; the novel analysis tool enables us to show the longitudinal image data with ROI in 4-D. The volume and QI of ROIs can be visualized in time sequence. The plots of QI supplies a quantitative evaluation means for treatment response of lymphoma. All these functions simplify the analysis and improve the efficiency of analysis.

It is impractical to only show the longitudinal curves like Fig. 7 to draw conclusions of the therapeutic responses. Our rationale is that by using the LINA tool, users/experts not only visualize the lesions from the aligned CT images and overlap PET images onto the CT, but also visualize the quantitative values (such as SUVmax, QI) for selected lesions. Thus it would be more efficient and informative to conduct quantitative imaging for follow-up study of lymphoma patients by combining both of ROIs in longitudinal images and corresponding longitudinal PET values.

It is worth noting that computer-assisted analysis method might occasionally cause computation errors. In our experiments, the only one error occurred when the lesion size is small ($<12\text{ mm}$). There are several factors affecting this error. First, the precision of PET/CT co-registration will undoubtedly affect the accuracy of SUV calculation of lymphoma region especially when the lesion is small and does not have clear boundary from CT images. This was the case why the error occurred. Another factor is that whether the automatically mapped ROIs are accurate enough to cover the PET peak region. Based on our validation of the algorithms, the matching accuracy for longitudinal CT series is about 3.3 mm for resolution $0.98\text{ mm} \times 0.98\text{ mm} \times 3.75\text{ mm}$, and it is less than one voxel size. Therefore, in the case where lesion size is small ($<12\text{ mm}$), we applied a relatively large ROI to cover the whole PET peak spot, i.e., grow the accurate ROI by 3.75 mm in all directions. In this way, all the 34 lesions can be successfully analyzed using our computer-assisted quantitative analysis tool.

Conclusion

We proposed a new computer-assisted method for quantitative evaluation of lymphoma response to therapy from serial PET/CT images. In this method, with the initial segmented ROI using semi-automatic or manual segmentation method at the baseline time-point, the corresponding ROI can be automatically identified and delineated by the local deformable registration of serial CT images. Thus longitudinal correspondences of ROI (lymphoma region) along serial images of the same patient are constructed. The proposed algorithm was tested and evaluated using nine patient [F-18] FDG PET/CT datasets with thirty-four lesions under a clinical trial of lymphoma treatment, administration of TGF-resistant LMP-specific cytotoxic T-lymphocytes to patients with relapsed EBV-positive lymphoma. Validation results showed accurate registration and SUV measures. The LINA software tool is the first set of quantitative tool using longitudinal image registration and segmentation for PET/CT lymphoma or other similar lesion studies.

References

1. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2008. *CA Cancer J Clin* 2008;58(2):71–96. [PubMed: 18287387]
2. Morton LM, Turner JJ, Cerhan JR, et al. Proposed classification of lymphoid neoplasms for epidemiologic research from the Pathology Working Group of the International Lymphoma Epidemiology Consortium (InterLymph). *Blood* 2007;110(2):695–708. [PubMed: 17389762]
3. Hoppe, R.; Mauch, PM.; Armitage, JO.; Diehl, V.; Weiss, LM. *Hodgkin Lymphoma*. Lippincott Williams & Wilkins; 2007.
4. Antoch G, Freudenberg LS, Debatin JF, et al. A radiologist's perspective on dual-modality PET/CT: Optimized CT scanning protocols and their effect on PET quality. *Journal of Nuclear Medicine* 2002;43(5):307.
5. Bar-Shalom R, Yefremov N, Guralnik L, et al. Clinical performance of PET/CT in evaluation of cancer: Additional value for diagnostic Imaging and patient management. *Journal of Nuclear Medicine* 2003;44(8):1200–9. [PubMed: 12902408]
6. Coleman RE, Delbeke D, Guiberteau MJ, et al. Concurrent PET/CT with an integrated imaging system: Intersociety dialogue from the joint working group of the American College of Radiology, the Society of Nuclear Medicine, and the Society of Computed Body Tomography and Magnetic Resonance. *Journal of Nuclear Medicine* 2005;46(7):1225–39. [PubMed: 16000293]
7. Ell PJ. The contribution of PET/CT to improved patient management. *British Journal of Radiology* 2006;79(937):32–6. [PubMed: 16421402]
8. Farma JM, Santillan AA, Melis M, et al. PET/CT fusion scan enhances CT staging in patients with pancreatic neoplasms. *Annals of Surgical Oncology* 2008;15(9):2465–71. [PubMed: 18551347]
9. Sironi S, Messa C, Mangili G, et al. Integrated FDG PET/CT in patients with persistent ovarian cancer: Correlation with histologic findings. *Radiology* 2004;233(2):433–40. [PubMed: 15516617]

10. Schiepers C, Filmont JE, Czernin J. PET for staging of Hodgkin's disease and non-Hodgkin's lymphoma. *Eur J Nucl Med Mol Imaging* 2003;30 (Suppl 1):S82–8. [PubMed: 12719922]
11. Hutchings M, Barrington SF. PET/CT for therapy response assessment in lymphoma. *J Nucl Med* 2009;50 (Suppl 1):21S–30S. [PubMed: 19380407]
12. Weber WA. Assessing tumor response to therapy. *J Nucl Med* 2009;50 (Suppl 1):1S–10S. [PubMed: 19380403]
13. Kostakoglu L, Goldsmith SJ. F-18-FDG PET evaluation of the response to therapy for lymphoma and for breast, lung, and colorectal carcinoma. *Journal of Nuclear Medicine* 2003;44(2):224–39. [PubMed: 12571214]
14. Kuo PH, Carlson KR, Christensen I, Girardi M, Heald PW. FDG-PET/CT for the evaluation of response to therapy of cutaneous T-cell lymphoma to vorinostat (suberoylanilide hydroxamic acid, SAHA) in a phase II trial. *Molecular Imaging and Biology* 2008;10(6):306–14. [PubMed: 18665425]
15. De Barse C, Depas G, Dresse MF, et al. Whole-body FDG PET imaging as a method for staging and early assessment of treatment response in pediatric patients with lymphoma. *Journal of Nuclear Medicine* 2003;44(5):346.
16. Eich HT, Kobe C, Dietlein M, et al. FDG-PET for treatment response assessment in advanced stage Hodgkin Lymphoma - report on the 2nd interim analysis of GHSG trial HD15. *Strahlentherapie Und Onkologie* 2008;184:11.
17. Moulin-Romsee G, Spaepen K, Stroobants S, Mortelmans L. Cost-effectiveness of early treatment response assessment by PET in non-Hodgkin lymphoma. *European Journal of Nuclear Medicine and Molecular Imaging* 2007;34:S242–S.
18. Pregno P, Chiappella A, Bello M, et al. Response assessment in aggressive non Hodgkin lymphoma disease: Predictive value of mid-treatment evaluation by 18-FDG-positron emission tomography/computed tomography (PET). *Annals of Oncology* 2008;19:249–50.
19. Pregno P, Chiappella A, Penna D, et al. Predictive value of response assessment with mid-treatment evaluation of 18-FDG-positron emission tomography/computed tomography (PET) in aggressive non Hodgkin lymphoma (NHL). *Haematologica-the Hematology Journal* 2007;92:59–60.
20. Stroobants S, Brepoels L, Thomas J, Verhoef G. PET-CT for treatment response assessment in lymphoma. *Haematologica-the Hematology Journal* 2007;92:20.
21. Fueger BJ, Yeom K, Czernin J, Sayre JW, Phelps ME, Allen-Auerbach MS. Comparison of CT, PET, and PET/CT for staging of patients with indolent non-Hodgkin's lymphoma. *Mol Imaging Biol* 2009;11(4):269–74. [PubMed: 19326177]
22. Hernandez-Maraver D, Hernandez-Navarro F, Gomez-Leon N, et al. Positron emission tomography/computed tomography: diagnostic accuracy in lymphoma. *Br J Haematol* 2006;135(3):293–302. [PubMed: 17032175]
23. Macapinlac HA. The utility of 2-deoxy-2-[18F]fluoro-D-glucose-positron emission tomography and combined positron emission tomography and computed tomography in lymphoma and melanoma. *Mol Imaging Biol* 2004;6(4):200–7. [PubMed: 15262235]
24. Ngeow JY, Quek RH, Ng DC, et al. High SUV uptake on FDG-PET/CT predicts for an aggressive B-cell lymphoma in a prospective study of primary FDG-PET/CT staging in lymphoma. *Ann Oncol* 2009;20(9):1543–7. [PubMed: 19474116]
25. Yan JY, Zhao BS, Curran S, Zelenetz A, Schwartz LH. Automated matching and segmentation of lymphoma on serial CT examinations. *Medical Physics* 2007;34(1):55–62. [PubMed: 17278490]
26. Pekar, V.; Kaus, MR.; McNutt, T. Office USPat. Book 3D image segmentation. City: Koninklijke Philips Electronics N.V; May. 2009 3D image segmentation.
27. Sofka M, Stewart CV. Location registration and recognition (LRR) for longitudinal evaluation of corresponding regions in CT volumes. *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv* 2008;11(Pt 2):989–97.
28. Duncan J, Ayache N. Medical image analysis: Progress over two decades and the challenges ahead. *IEEE Trans Pattern Anal Mach Intell* 2000;22:85–106.
29. Rueckert D, Sonoda LI, Hayes C, Hill DL, Leach MO, Hawkes DJ. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE transactions on medical imaging* 1999;18(8):712–21. [PubMed: 10534053]

30. Marsland S, Twining CJ, Taylor C. A minimum description length objective function for group-wise non-rigid image registration. *Image and Vision Computing* 2008;26(3):333–46.
31. Xue, Z.; Wong, K.; Wong, STC. Joint Registration and Segmentation of Serial Lung CT Images for Image-Guided Lung Cancer Diagnosis and Therapy Computerized Medical Imaging and Graphics. 2009. in press
32. Bollard CM, Gottschalk S, Leen AM, et al. Complete responses of relapsed lymphoma following genetic modification of tumor-antigen presenting cells and T-lymphocyte transfer. *Blood* 2007;110(8):2838–45. [PubMed: 17609424]
33. Jenkinson M, Smith S. A global optimisation method for robust affine registration of brain images. *Medical Image Analysis* 2001;5(2):143–56. [PubMed: 11516708]
34. Press, WH.; Flannery, BP.; Teukolsky, SA.; Vetterling, WT. *Numerical Recipes in C*. Cambridge, U.K: Cambridge Univ. Press; 1992.
35. Paragios N, Mellina-Gottardo O, Ramesh V. Gradient vector flow fast geometric active contours. *IEEE transactions on pattern analysis and machine intelligence* 2004;26(3):402–7. [PubMed: 15376886]
36. Rueckert D, Frangi AF, Schnabel JA. Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration. *IEEE Trans Med Imaging* 2003;22(8):1014–25. [PubMed: 12906255]
37. Ceresoli G, Zucali P, Van Hemert R, et al. Positron emission tomography with F18-fluorodeoxyglucose (FDG-PET) in malignant pleural mesothelioma (MPM): Prediction of response to chemotherapy by quantitative assessment of standard uptake value (SUV). *Lung Cancer* 2005;49:S219–S.
38. Keyes JW. Suv - Standard Uptake or Silly Useless Value. *Journal of Nuclear Medicine* 1995;36(10):1836–9. [PubMed: 7562051]
39. Luo J. Estimate standard uptake value (SUV) in F18FDG PET tumor imaging. *Medical Physics* 2006;33(6):2014–5.
40. Mahan S, Ramsey C. Automatic generation of standard uptake value (SUV) isolines for treatment planning. *Medical Physics* 2003;30(6):1429.
41. Panizo CM, Panizo E, Bosch J, et al. The maximum standard uptake value (SUVmax) of PET/CT correlates with lymphoma aggressivity. *Annals of Oncology* 2008;19:250.
42. Suzuki O, Morimoto M, Hashiguchi H, et al. Standardization of PET standard uptake value for delineating GTV in integrated PET-CT of head and neck cancer. *International Journal of Radiation Oncology Biology Physics* 2008;72(1):S409–S.
43. Thie JA. Understanding the standardized uptake value, its methods, and implications for usage. *Journal of Nuclear Medicine* 2004;45(9):1431–4. [PubMed: 15347707]
44. Rosenberg SA. National-Cancer-Institute Sponsored Study of Classifications of Non-Hodgkins Lymphomas - Summary and Description of a Working Formulation for Clinical Usage. *Cancer* 1982;49(10):2112–35. [PubMed: 6896167]
45. Vercauteren T, Pennec X, Perchant A, Ayache N. Non-parametric diffeomorphic image registration with the demons algorithm. *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv* 2007;10(Pt 2):319–26.
46. Xue Z, Shen D, Davatzikos C. Statistical representation of high-dimensional deformation fields with application to statistically constrained 3D warping. *Med Image Anal* 2006;10(5):740–51. [PubMed: 16887376]
47. Jenkinson, M. Book FLIRT. 5. Vol. 5. City: University of Oxford; 2008. FLIRT. <http://www.fmrib.ox.ac.uk/fsl/flirt/index.html>

Fig. 1.

Flowchart for the computer-assisted quantitative evaluation of therapeutic effects of lymphoma from serial PET/CT images. After global co-registration of the PET and CT images at each time-point, ROIs and liver were segmented using a level set-based semi-automatic method. Longitudinal deformations of the serial data were automatically calculated. Then any semi-automatically segmented ROI one time-point can be automatically mapped onto other time-point images to facilitate quantitative analysis. PETb and CTb stand for baseline PET/CT images, and PETt and CTt stand for PET/CT images at time t.

Fig. 2.
Comparison of the registration errors between LINA and FFD-based registration using simulated images.

Fig. 3.

Illustration of semi-automatic lymph node segmentation using level set method. (a) The region of interest is first identified by overlapping the PET with the CT; (b) an initial point is then manually marked for each lymph node from CT image; and (c) the level set method is applied to automatically extract the lymph nodes.

Fig. 4.

An illustration of manual segmentation of ROI where boundaries are not clearly discernable, in which the semi-automatic method has failed. (a) The ROI is first identified by overlapping the PET with the CT; (b) the level set segmentation leaks into a larger region in the CT image; and (c) the region of interest is then manually marked on the CT image by referring to (a).

Fig. 5.

Visual comparison of the semi-automatic segmentation of lesion and the automatically mapped lesion. (a) and (b) fused PET/CT images at the baseline and the second time-point, respectively; (c) and (d) semi-automatic segmentation of lesion; (e) warped baseline CT and ROI onto the second time-point image. The SUVs of ROIs and the ratios of SUV of lesion to that of the liver are also given.

Fig. 6.

Comparison of the semi-automatic segmentation of lymph node and the automatically mapped shapes. Row 1: fused PET/CT images at different time-points; Row 2: semi-automatically segmented lymph nodes; Row 3: the warped baseline CT image and the corresponding ROIs at subsequent time-points. It can be seen that the QI of ROI decreases along with time/treatment.

Fig. 7.

Plots of longitudinal QI values of different patients for both of semi-automatic segmentation method (dash-dot lines) and the proposed automatic mapping method (solid lines). (a) Seven lesions, four time-points; (b) one lesion, three time-points; (c) four lesions, five time-points; and (d) four lesions, two time-points. In (a) and (b) all the lesions are determined from the baseline image and no other new lesion is identified in the follow-up studies; in (c), all the lesions are determined from the baseline image and the computer-assisted method can

determine the corresponding lesions for all the follow-up studies, while it is difficult to even manually mark the lesion for some scans; in (d), new lesions are identified at follow-up studies and they can be mapped to the baseline using the proposed approach.