

Toward Robust Characterization of Lung Diseases: A Sensitivity Analysis of Lung Computed Tomography Biomarkers to Registration Error

Rahul Hingorani, Nicole L. Brown, Christopher M. Cervantes, Robert H. Brown, and Andrew S. Gearhart

ABSTRACT

Computed tomography (CT) scans, because of their ability to differentiate tissue densities, have been widely used to evaluate lung health. Recent studies such as COPDGene have collected inhalation and exhalation CT scans from thousands of subjects, promising insight into the mechanical properties of lung tissue. These paired scans must often be spatially aligned (i.e., registered) to extract biomarkers describing the movement of lung tissue that may correlate with disease. Unfortunately, the relationship between registration and biomarker error is poorly characterized, a challenge that must be addressed before registration-based biomarkers can be used in clinical practice. In our analysis, we consider three registration-based biomarkers (Jacobian determinant, anisotropic deformation index, and slab-rod index) and demonstrate their sensitivity to modeled registration error. We provide a range of errors for a given biomarker, highlighting how both the magnitude of registration error and correlations between vectors in the registration error field can influence biomarker error. We then describe a method to measure the error field for a particular registration algorithm and compare it with modeled registration error. These estimates enable selection of an appropriate registration error model, which improves understanding of biomarker uncertainty. Quantifying the relationship between registration and biomarker error is crucial because it may inform the selection of a registration algorithm to reduce error in new research studies, and in turn, result in robust imaging biomarkers for disease characterization.

INTRODUCTION

Computed tomography (CT) is a medical imaging technique that uses x-rays to determine tissue density—allowing physicians to obtain 3-D representations of patients' soft tissues and bones. Chest CT scans are

typically composed of millions of 3-D pixels, or voxels, that represent cuboids of tissue or air that are ~1 mm on a side. Because of CT's ability to differentiate tissue density, it has been used to evaluate lung health for many

Note: This work draws heavily, in many cases verbatim, from R. Hingorani, N. L. Brown, C. M. Cervantes, R. H. Brown, and A. S. Gearhart, "On characterizing the sensitivity of lung computed tomography biomarkers to registration error," in Proc. SPIE 12032, Medical Imaging 2022: Image Processing, 120321C, Apr. 4, 2022, <https://doi.org/10.1117/12.2611087>.¹

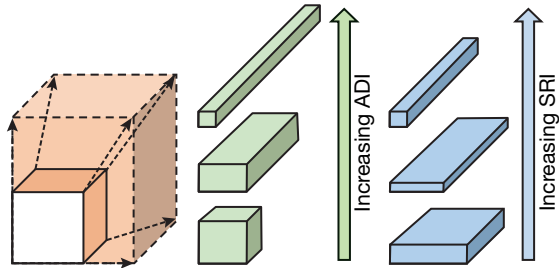


Figure 1. Intuition of JD (orange), ADI (green) and SRI (blue) values. Increasing JD is shown by the volume expansion of a voxel, increasing ADI is shown by increased stretching in one direction, and increasing SRI is shown by the transition from thin/flat “slabs” to long “rods.”

disorders—for example, chronic obstructive pulmonary disease (COPD),² asthma,³ and coronavirus disease 2019 (COVID-19).⁴

Recent studies such as COPDGene move beyond analyzing individual CT scans to collecting paired inhalation and exhalation scans from thousands of subjects,⁵ capturing the state of the lung tissue at the transition points of a respiratory cycle. As the lungs are inherently dynamic organs, these paired scans have the potential to indicate changes in the mechanical properties of diseased lung tissue, which may inform patient treatment. In particular, two classes of biomarkers have been derived from inhalation/exhalation scan pairs: those that use raw tissue density values at voxels (e.g., parametric response mapping⁶) and those that use the displacement between voxels within a scan pair (e.g., Jacobian determinant [JD], anisotropic deformation index [ADI],⁷ and slab-rod index [SRI]⁸). Our work focuses on this latter class of biomarkers. Specifically, JD is a measure of voxel-wise volume change in the lungs, while ADI and SRI measure the magnitude and orientation of directional preference in volume change, respectively.⁸ Figure 1 illustrates these concepts by showing JD as the volume change in a voxel (orange region, left), increasing ADI as increased stretching in one direction (green), and SRI as a transition from thin/flat “slabs” to long “rods” (blue).

Both classes of imaging biomarkers rely on registration, or the process of aligning voxels between scan pairs, to estimate changes between inhalation and exhalation (i.e., estimate the motion of the lungs during respiration). Specifically, registration produces a set of per-voxel vectors (a deformation field) that approximates the transformation from one paired scan to the other (Figure 2). In general, algorithmic errors can occur when registering these scans because of the complex nature of lung motion. Unfortunately, the relationship between registration error and biomarker variation is not well understood—making it difficult to control the quality of imaging data sets. There are at least two components to this problem: (1) measuring (or modeling) registration error and (2) characterizing its effect on downstream biomarkers. Recent approaches to understanding registration error use machine learning models to identify regions of poor registration quality^{9–12} and to adaptively refine registrations.¹³ Existing work that explores the link between registration error and biomarker variation focuses on JD.^{14–16} Our work discusses an approach to both modeling and measuring registration error and analyzes the impact of registration error on JD, ADI, and SRI.

This article quantifies the relationship between errors in registration and imaging biomarkers. In the next section, we analyze the sensitivity of three biomarkers (JD, ADI, and SRI) to modeled registration error. We highlight that the magnitude of this error and the correlations between the vectors in the registration error field can influence the resulting biomarkers. Then we describe a strategy to measure registration error and compare it with modeled error for a pair of algorithms. These measurements enable the selection of an appropriate registration error model, which can improve downstream biomarker uncertainty characterization. This, in turn, may guide selection and tuning of a registration algorithm to reduce biomarker error in new imaging studies. Finally, we discuss future directions. This work highlights the need for estimating error in registration-based biomarkers, which is crucial when using these biomarkers to characterize disease.



Figure 2. Intuition behind registration. The voxels in the inhalation scan (a) are spatially aligned with those in the exhalation scan (b; red arrow). The output is a deformation field representing the per-voxel vectors that approximate the transformation from the inhalation to exhalation scan. (c) The inverse deformation field (from exhalation to inhalation).

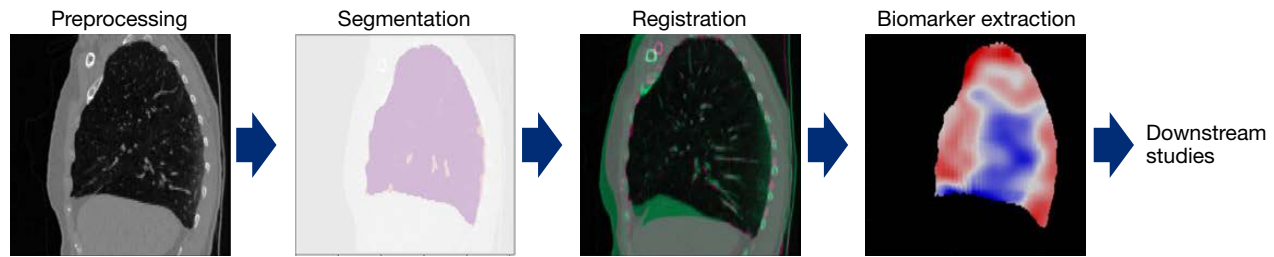


Figure 3. Processing pipeline to compute registration-based lung biomarkers. Our analysis focuses only on two of these stages: registration and biomarker extraction.

SENSITIVITY ANALYSIS

Methodology

In this section, we analyze the sensitivity of JD, ADI, and SRI in the presence of simulated registration error. To extract these biomarkers from paired scans, we first preprocess the scans and identify the voxels in the lung regions. This latter step, segmentation, is conducted using the Chest Imaging Platform (CIP).¹⁷ These segmented lung regions are then registered to determine spatial correspondence between the voxels in the inhalation/exhalation scan pair. The registration is performed using the Advanced Normalization Tools¹⁸ (ANTs) library and consists of three stages: rigid (orients the scans through rotating and translating), affine (matches the scans in size through shearing and scaling), and deformable (determines the local voxel level displacements by nonlinearly warping voxels).¹⁹ For extracting biomarkers from the displacement field, only the deformable field—computed using the Symmetric Normalization (SyN) algorithm²⁰—is used because it captures local deformations. After extraction, these biomarkers are ready for use in downstream imaging studies to characterize lung diseases (Figure 3).

While both segmentation and registration may result in downstream biomarker errors, our analysis focuses on registration error and leaves segmentation error analysis to future work. Since computing the true registration is infeasible because of the complexity of knowing the precise movement of each lung voxel, we treat the deformation field (Φ) from SyN as the ground truth. To simulate registration error, a synthetic perturbation field, \mathbf{H} , is added to this ground-truth registration to create a perturbed deformation field, $\tilde{\Phi}$, represented by

$$\tilde{\Phi} = \Phi + \epsilon \mathbf{H},$$

where ϵ is a set scalar that controls the magnitude of the synthetic error. Following the work of Shao et al.,¹⁶ the synthetic perturbation fields, \mathbf{H} , are generated using a cubic B-spline model. This approach ensures a smooth perturbation field that preserves the invertibility of the perturbed deformation and avoids situations where portions of the transformed scans appear to fold over

themselves. (Having a single smooth deformation field for all lung tissue assumes that neighboring voxels influence each other locally. Anatomically, this assumption breaks down when registering voxels in a region with sliding motion—for example, along lung lobes boundaries during the breathing cycle.²¹ Future work may explore this behavior further.) The cubic B-spline model can be interpreted as interpolating between randomly distributed weight vectors located at specific coordinates in the perturbation field, known as control points (Figure 4). In their analysis, Shao et al. assumed that control points were spaced every four voxels.

The selection of control point spacing (CPS), however, is an important decision that we believe directly relates to biomarker error (Figure 5). Registration-based biomarkers at a given voxel are computed using the vectors in the deformation field immediately neighboring that voxel. The similarity in the direction of these vectors directly impacts the biomarker value. We refer to this similarity as local correlation, or the average cosine similarity of vectors within a volume surrounding a particular deformation field coordinate. Our sensitivity analysis varies both the magnitude (controlled by ϵ) and local correlation (controlled by the CPS) of the synthetic perturbation field to simulate different levels of registration error and measure the resulting percent change in each biomarker.

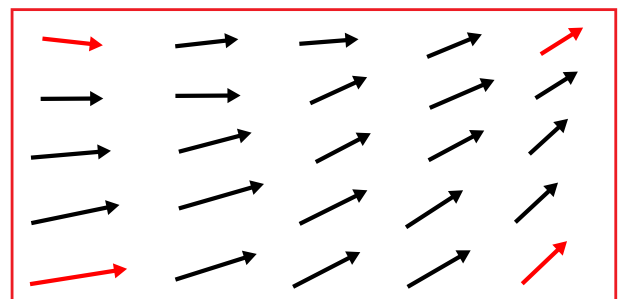


Figure 4. Visualization of cubic B-spline perturbation field generation. Interpolation (black arrows) using a cubic B-spline model occurs between randomly distributed weight vectors at control points (red arrows). Note that this is strictly an illustrative depiction and is not generated from actual splines.

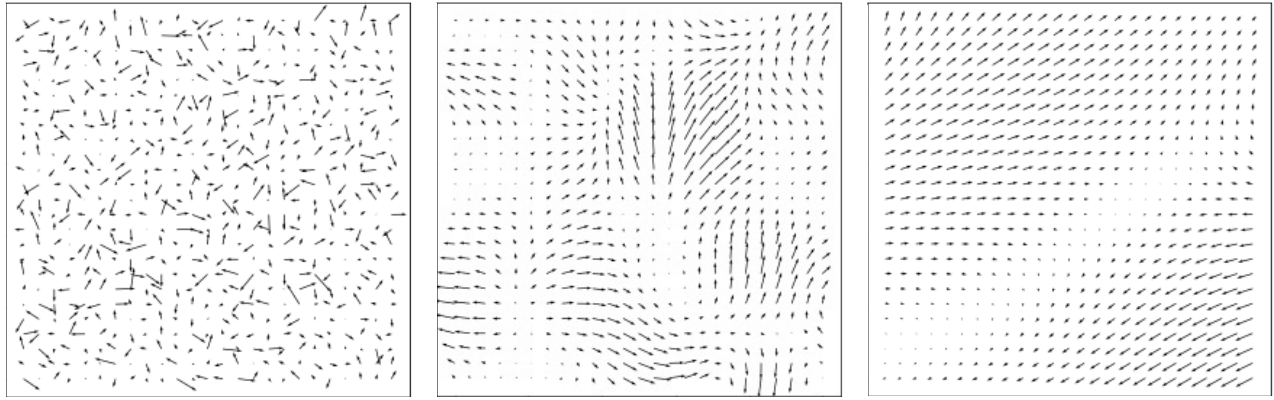


Figure 5. Examples of synthetic perturbation fields. Two-dimensional slices of a normally distributed random perturbation field (left) and cubic B-spline perturbation fields with CPS of 4 (center) and CPS of 16 (right). Neighboring vectors in the B-spline fields are more correlated with increased CPS, highlighting the impact of this parameter.

To calculate the biomarker error, we compute the mean percent change of the original biomarker values (obtained from the SyN-generated “ground truth” deformation field, Φ) and the perturbed biomarker values (obtained from $\tilde{\Phi}$). This can be expressed as

$$\% \text{ error} = \frac{1}{N} \sum_{\forall(x,y,z) \in \mathcal{S}} \frac{|B_{\tilde{\Phi}}(x,y,z) - B_{\Phi}(x,y,z)|}{B_{\Phi}(x,y,z)} * 100, \quad (1)$$

where $B_{\Phi}(x,y,z)$ and $B_{\tilde{\Phi}}(x,y,z)$ are the original and perturbed biomarker values at voxel coordinate (x,y,z) , \mathcal{S} represents the set of voxels inside the lung region (obtained via segmentation), and N is the total number of lung voxels.

The steps of the sensitivity analysis can be summarized as follows. For each scan pair:

1. Register the segmented pair to compute the ground-truth deformation field, Φ .
2. Add a cubic B-spline perturbation field to the ground-truth deformation, generating a perturbed deformation field, $\tilde{\Phi}$.
3. Compute the biomarker values for the original and perturbed deformation fields.
4. Compute the mean percent error between the original and perturbed biomarker values.
5. Repeat steps 2–4 while varying the magnitude and CPS of the B-spline field to simulate different levels of registration error.

Results

The sensitivity analysis was conducted using 10 COP-DGene inhalation/exhalation scan pairs from the DIR-Lab data set,²² each with voxel side lengths of 2.5 mm, 0.625 mm, 0.625 mm. We generated 100 synthetic

perturbation fields for each CPS (4, 8, 12, and 16 voxels) and ground-truth deformation field (10 scans), resulting in a total of $4 * 10 * 100 = 4,000$ perturbation fields. The perturbation fields were scaled by constant factor ϵ such that the average magnitude of registration error ranged from 0.19 to 0.94 mm. Each perturbation field was added to a ground-truth deformation field, and the resulting mean biomarker error for JD, ADI, and SRI was computed (Eq. 1).

Our results suggest that for all three biomarkers, small magnitudes of registration error (<1 mm) can significantly impact error in downstream biomarkers (Figure 6). For JD, the percent error can reach up to 43.8% at the largest magnitude of error considered (0.94 mm) and the lowest CPS (4 voxels). For ADI and SRI, these errors are even larger, reaching 468.1% and 123.1%, respectively. Additionally, percent error is dependent on both the registration error magnitude and CPS. Specifically, biomarker error varies directly with registration error magnitude and inversely with CPS, supporting our original hypothesis that CPS impacts the biomarker error. Smaller CPS also changes the distribution of biomarkers—an observation that may bias downstream statistical tests. This trend holds for JD, ADI, and SRI, suggesting that a robust understanding of the local correlation of actual registration error is required to characterize its impact on biomarkers.

To visualize the impact that the CPS has on biomarker values, Figure 7 shows an example of a unit volume cube (representing a voxel) deformed by two cubic B-spline perturbation fields of equal magnitude, one with a CPS of 4 voxels and the other with a CPS of 16 voxels. The cube deformed by the perturbation field with CPS of 4 voxels has a volume change that is greater in certain directions than others (i.e., anisotropic change), highlighting the fact that the local correlation of the registration error vector field impacts the nature of volume change of a given voxel, and thus, the biomarker values themselves.

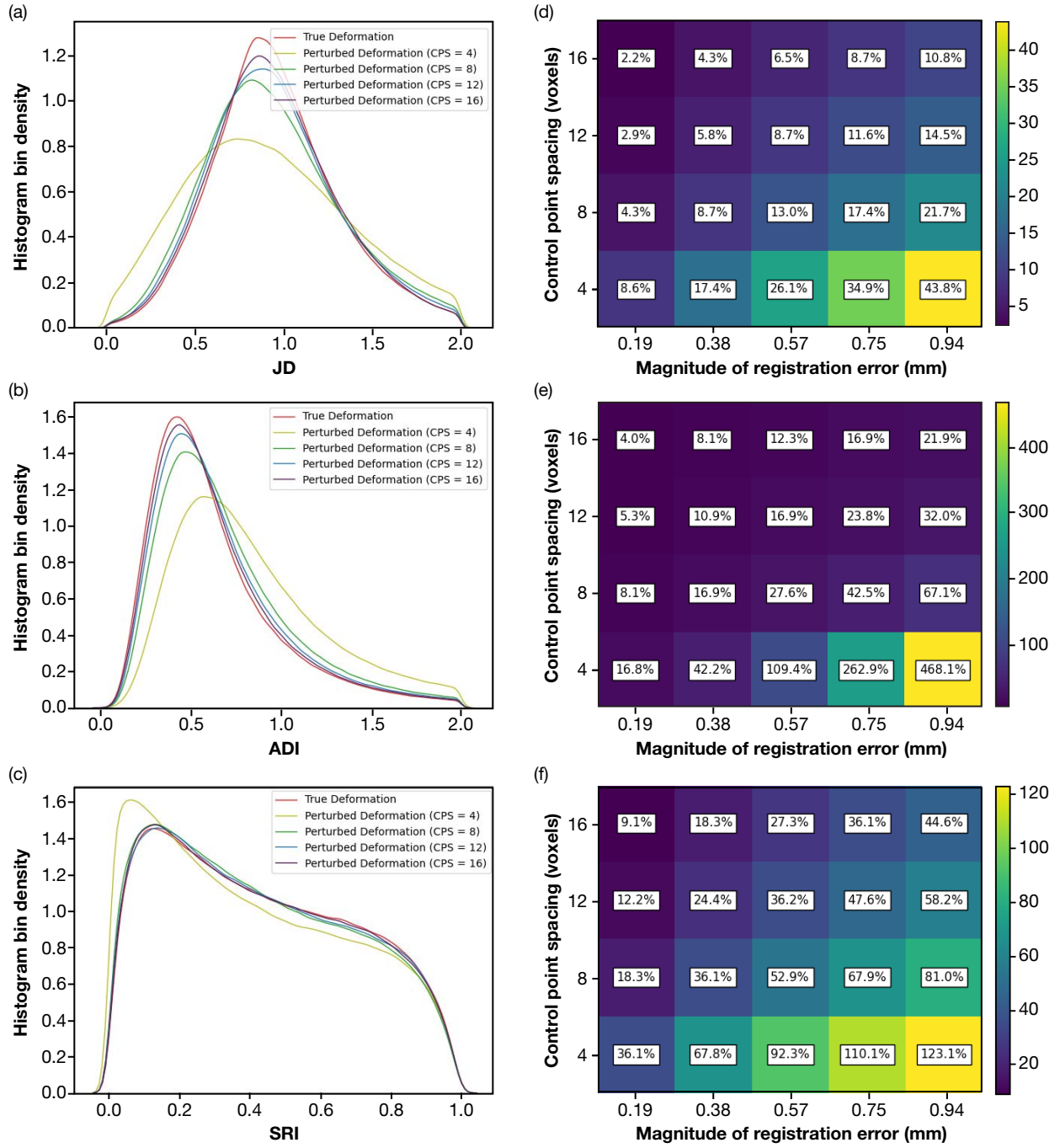


Figure 6. Example biomarker distributions (computed with kernel density estimation) and mean percent errors (computed using 100 perturbations on a single scan pair) as a function of average registration error and CPS. (a) Empirical JD distribution; (b) empirical ADI distribution; (c) empirical SRI distribution; (d) JD percent error; (e) ADI percent error; (f) SRI percent error. For all three biomarkers, the biomarker error decreases as the magnitude of registration error decreases or as the CPS increases.

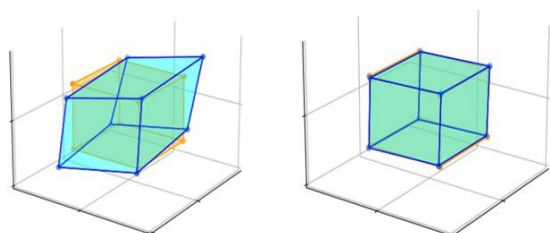


Figure 7. Unit volume voxel (orange) deformed by a $2 \times 2 \times 2$ slice of a B-spline perturbation field with a CPS of 4 (left) and a CPS of 16 (right). All voxels have been centered on the origin to show relative volume change rather than spatial shifts. The deformed voxel (blue) corresponding to the CPS of 4 has a volume change that is greater in certain directions than others, highlighting the fact that the local correlation of the registration error vector field impacts the nature of volume change.

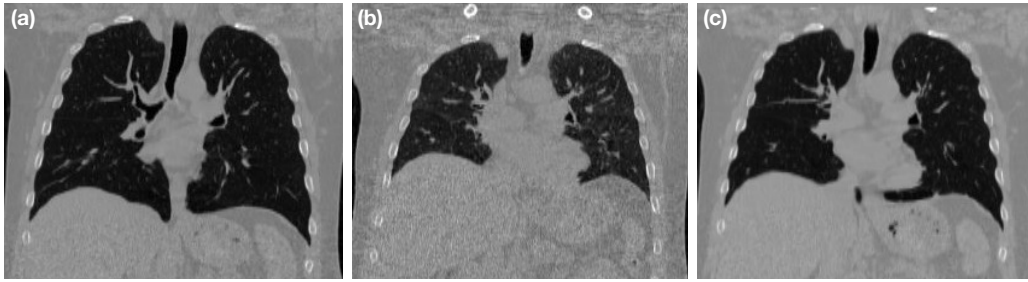


Figure 8. Visualization of synthetic CT scans. The individual scans show two-dimensional slices of original inhalation (a) and exhalation (b) scans, with a synthetic slice (c) representing a state halfway between inhalation and exhalation (DIR-Lab COPDGene patient 10).

LOCAL CORRELATION ANALYSIS

Methodology

The results in the preceding section show that biomarker error is sensitive to both the magnitude and local correlation of modeled registration error. Although these results provide an understanding of the relationship between errors in registration and biomarkers, they alone are not useful for estimating the biomarker uncertainty for a particular clinical application because of the large spread in biomarker error that is dependent on the specific parameters used to model registration error. Therefore, to truly estimate the uncertainty in biomarker values for a particular registration algorithm and data set, it is necessary to estimate both the magnitude and local correlation of *actual* registration error to understand the model parameters that should be used to simulate realistic registration error (i.e., determine the most appropriate row and column in the heatmaps in Figure 6). To estimate the registration error magnitude, one common approach is to compute the average displacement between a set of annotated landmark points in the lungs and landmark points that have been transformed via the registration field. Our work focuses on local correlation estimation since the methodology for estimating this is less defined.

To estimate the local correlation of actual registration error, we generated synthetic CT scans in which the ground-truth deformation field was known. To generate these synthetic scans, we first registered the inhalation and exhalation scans for each of the 10 DIR-Lab COPDGene scan pairs and treated the deformation field output as the ground-truth registration field. Each scan pair was registered with two different

registration algorithms, SyN and Time Varying Velocity (TVV) (both available through the ANTs library), to understand the generalizability of results to differences in the ground-truth deformation field. We then scaled the deformation field by a factor of $\frac{1}{2}$ and used this deformation field to transform the inhalation scan, creating a synthetic scan representing a state halfway between inhalation and exhalation (Figure 8). This resulted in a total of 20 synthetic scan pairs (10 pairs using SyN for the ground truth and 10 pairs using TVV for the ground truth). We then registered the inhalation scans to their corresponding synthetic exhalation scans, repeating the process for both the SyN and TVV registration algorithms (Figure 9). As the ground-truth deformation field is known for these registrations, we used the difference between these observed deformation fields and the ground-truth deformation fields to calculate the field of registration errors. Lastly, to compute the local correlation of the resulting registration error vector field, we calculated the local correlation (i.e., average pairwise cosine similarity between vectors around a particular point) averaged across 300 windows of a given size within the lungs (as indicated by a segmentation mask).

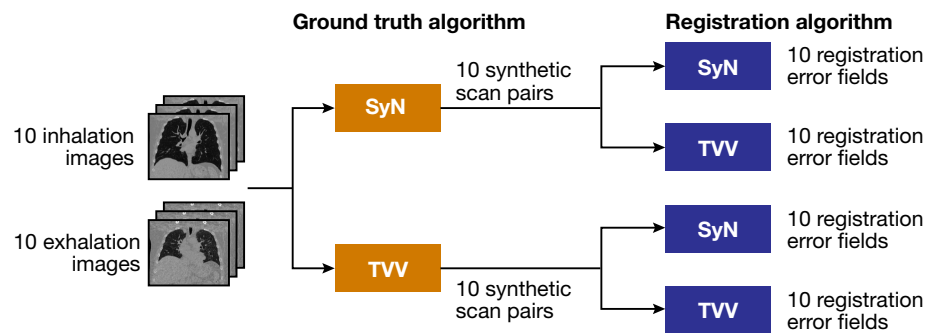


Figure 9. Experimental protocol for generating synthetic scans and registration error fields. Each of the 10 scan pairs is registered with both SyN and TVV, resulting in 10 synthetic scan pairs for each algorithm (20 total). Each synthetic scan pair is then registered with both SyN and TVV, resulting in 10 registration error fields for each unique combination of ground-truth algorithm and registration algorithm (40 total registration error fields).

The steps for the local correlation analysis can be summarized as follows for each scan pair:

1. Register the inhalation image to the exhalation image with algorithm A (either SyN or TVV) to compute the ground-truth registration field.
2. Scale the ground-truth registration field by a factor of $1/2$ and use this vector field to transform the inhalation image, generating a synthetic image representing a state halfway between inhalation and exhalation.
3. Register the inhalation image to the synthetic exhalation image with algorithm B.
4. Compute the registration error field by subtracting the deformation field from step 3 from the ground-truth deformation field from step 1.
5. Calculate the local correlation of the resulting registration error vector field averaged across 300 windows for a range of window sizes (e.g., $4 \times 4 \times 4$ to $20 \times 20 \times 20$ voxels).
6. Repeat steps 1–5 for all combinations of SyN and TVV used for algorithms A and B.

Results

For each of the 40 registration error vector fields (20 synthetic scan pairs, 2 registration algorithms), we computed the local correlation for window sizes ranging from $4 \times 4 \times 4$ voxels to $20 \times 20 \times 20$ voxels, generating a local correlation curve. These curves were averaged across the 10 registration error vector fields from each unique combination of ground-truth algorithm and registration algorithm (Figure 10). The figure also shows local correlation curves for several B-spline perturbations with different CPS. Optimally, the actual local correlation curve should match the curve for a particular B-spline perturbation to determine the most appropriate CPS to model registration error for a given algorithm, enabling a more accurate characterization of biomarker error.

In our results, the choice of registration algorithm had a significant effect on local correlation. The SyN algorithm produced output deformation fields that have similar local correlations to B-spline models with a CPS of 4. TVV, however, produced deformation fields that are typically between the curves for B-spline perturbations with spacings of 8 and 12. Choice of algorithm used for producing a ground-truth registration field did not

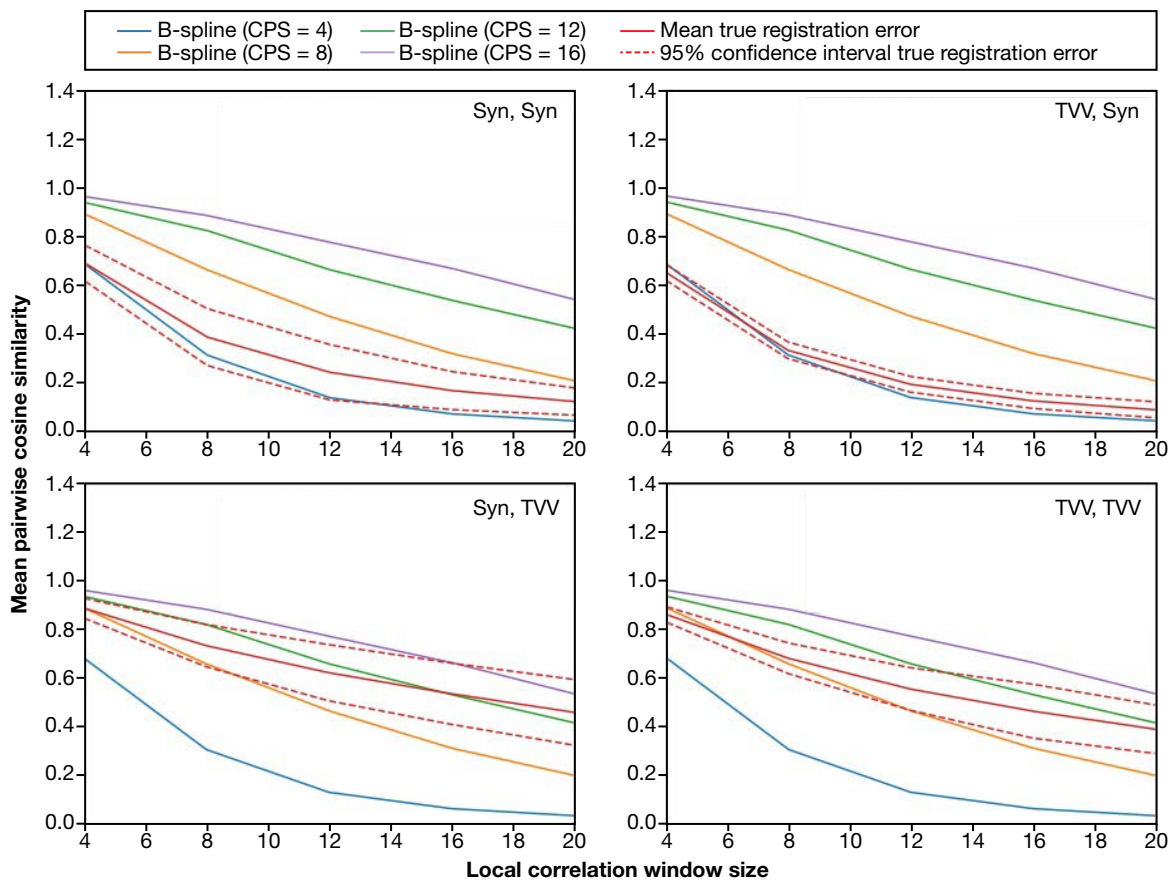


Figure 10. Local correlations (ground-truth algorithm, registration algorithm) for observed registration error and B-spline perturbation fields. The SyN algorithm produced deformation fields that have similar local correlations to B-spline models with a CPS of 4. TVV produced deformation fields that are typically between the curves for B-spline perturbations with spacings of 8 and 12.

appear to have a major effect on the local correlation of the registration error. This suggests that the local correlation is primarily a function of the registration algorithm itself, but we note that additional experimental evidence is needed to validate this conjecture. Comparing these local correlation curves with those of B-spline perturbations with different CPS allows us to combine these results with those in the sensitivity analysis section to characterize the biomarker uncertainty for a particular registration algorithm. This consequently may inform the choice of registration algorithm since it enables one to determine which algorithms have the lowest biomarker error.

CONCLUSION

In this work, we analyzed the effect of registration error magnitude and local correlation on registration-based biomarkers such as JD, ADI, and SRI. Specifically, we first measured the sensitivity of each biomarker to registration error by modeling registration error with a cubic B-spline perturbation field, varying both the magnitude and local correlation of the synthetic error to simulate different levels of registration error. We showed that biomarker error decreases as either the magnitude of registration error decreases or the local correlation of registration error increases, highlighting that even small magnitudes of registration error (<1 mm) can significantly impact biomarker error. Since biomarker error is directly dependent on the local correlation of registration error, we investigated the local correlation of *actual* registration error, introducing a method to estimate this for a particular registration algorithm. We did this by generating synthetic scans in which the ground-truth deformation field was known and then registering the synthetic scan pairs, resulting in a registration error vector field. The local correlation of the registration error vector field was compared with that of B-spline perturbation fields with different CPS to determine the

most accurate model for registration error for a given algorithm. Determining an accurate model for registration error was critical to ensure a meaningful estimate of biomarker uncertainty (i.e., determine the most relevant row in Figure 6).

To extend this analysis to future imaging protocols, one can do the following (Figure 11):

1. Conduct the sensitivity analysis using inhalation/exhalation scan pairs in the data set of interest. Apply different levels of modeled registration error (i.e., B-spline perturbation fields with varying magnitudes and CPS) to the ground-truth deformation fields and measure the mean biomarker error (generate heatmaps in Figure 6).
2. Estimate the magnitude of registration error by computing the average displacement between a set of annotated landmark points in the lungs and landmark points that have been transformed via the registration field.
3. Estimate the local correlation of registration error by generating synthetic scans in which the ground-truth deformation field is known and computing the local correlation of the registration error vector field. Compare this with the local correlation of different B-spline perturbations to determine the most similar B-spline model for registration error.
4. Using the estimated magnitude and local correlation of registration error, determine the most reasonable estimate for biomarker uncertainty for a given algorithm/data set (i.e., select a particular row/column in the heatmaps in Figure 6).
5. Using the uncertainty estimates, select the algorithm with the lowest biomarker error.

There are limitations to this pipeline. First, the impact of registration error magnitude is dependent on

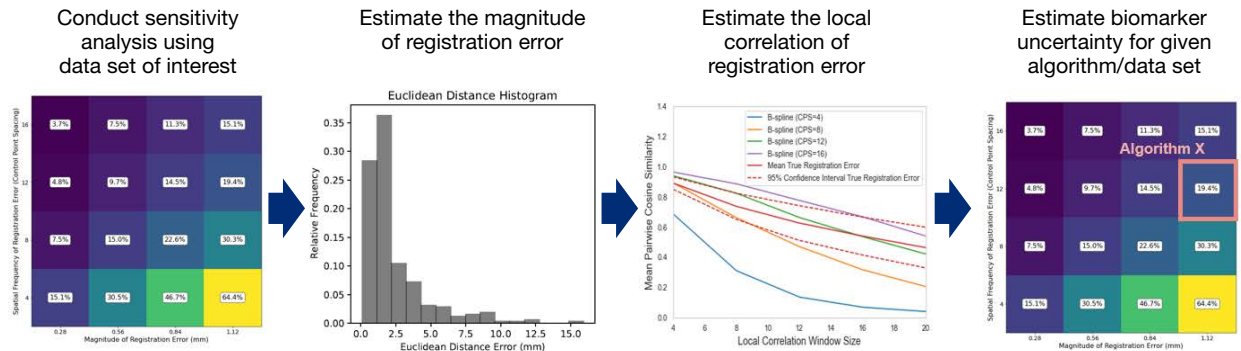


Figure 11. Proposed pipeline for estimating biomarker uncertainty for new imaging protocols. First, the sensitivity analysis is conducted to determine the biomarker error to modeled registration error. Then both the magnitude and local correlation of registration error are estimated using synthetic CT scans. Lastly, using both these estimates, the most reasonable estimate for biomarker uncertainty is determined for a given algorithm/data set.

the voxel size in the CT scans—for example, a 0.25-mm error has more of an impact on a 0.5 mm, 0.5 mm, 0.5 mm voxel than it does on a 1 mm, 1 mm, 1 mm voxel. To generalize this to data sets with different voxel sizes, the registration error magnitudes should be normalized by voxel size. Second, the method to estimate registration error magnitude is prone to human errors. The proposed method requires human annotation of key landmark points that may have labeling errors and may not be representative of the error at other regions in the scans. Lastly, the method to estimate the local correlation error may be biased to a particular algorithm. Since in some cases the same algorithm is used to first generate the synthetic scans and then register the synthetic scan pairs, the resulting registration error field may not be representative of the error in the original scan pairs. We attempted to mitigate this bias by considering two different registration algorithms (SyN and TVV). We compared the results between the case in which the same algorithm was used for generating the synthetic scans and registering the synthetic scan pairs and the case in which different algorithms were used. Our initial findings suggest that the results are generalizable since the local correlation did not vary between the two ground-truth registration algorithms (Figure 10). However, a more comprehensive analysis should be conducted in which the local correlation is compared with existing methods for estimating registration error.^{9–12}

The ability to estimate the biomarker uncertainty for a particular registration algorithm may inform the choice of algorithm in certain clinical studies since it allows one to compare algorithms in terms of biomarker error. Furthermore, our results suggest that registration-based lung biomarkers such as JD, ADI, and SRI have a complex relationship with registration error that is dependent on both the magnitude and local correlation of the error field. This relationship needs to be well understood before such biomarkers can be used in the clinical space for disease characterization. This is a rich space for further exploration—especially with the increasing availability of large imaging cohorts and the development of data-driven registration-characterization^{9–12} models (e.g., deep neural networks) that promise to greatly increase processing throughput.

REFERENCES

- R. Hingorani, N. L. Brown, C. M. Cervantes, R. H. Brown, and A. S. Gearhart, “On characterizing the sensitivity of lung computed tomography biomarkers to registration error,” in *Proc. SPIE 12032, Medical Imaging 2022: Image Processing*, 120321C, Apr. 4, 2022, <https://doi.org/10.1117/12.2611087>
- D. A. Lynch and M. L. Al-Qaisi, “Quantitative CT in COPD,” *J. Thoracic Imag.*, vol. 28, no. 5, pp. 284–290, 2013, <https://doi.org/10.1097/RTI.0b013e318298733c>.
- S. Gupta, S. Siddiqui, P. Halder, J. V. Raj, J. J. Entwisle, et al., “Qualitative analysis of high-resolution CT scans in severe asthma,” *Chest*, vol. 136, no. 6, pp. 1521–1528, 2009, <https://doi.org/10.1378/chest.09-0174>.
- T. C. Kwee and R. M. Kwee, “Chest CT in COVID-19: What the radiologist needs to know,” *Radio Graph.*, vol. 40, no. 7, pp. 1848–1865, 2020, <https://doi.org/10.1148/rg.2020200159>.
- COPDgene Study, “CT imaging in COPDGENE: Introduction for technologists,” <http://www.copdgene.org/sites/default/files/COPD-Gene%20CT%20Imaging%20Introduction.pdf>.
- C. J. Galbán, M. K. Han, J. L. Boes, K. A. Chughtai, C. R. Meyer, et al., “Computed tomography-based biomarker provides unique signature for diagnosis of COPD phenotypes and disease progression,” *Nature Med.*, vol. 18, no. 11, pp. 1711–1715, 2012, <https://doi.org/10.1038/nm.2971>.
- S. Bodduluri, S. P. Bhatt, and J. M. Reinhardt, “Computed tomography image matching in chronic obstructive pulmonary disease,” *Crit. Rev. Biomed. Eng.*, vol. 44, no. 6, pp. 411–425, 2016, <https://doi.org/10.1615/CritRevBiomedEng.2017021299>.
- R. Amelon, K. Cao, K. Ding, G. E. Christensen, J. M. Reinhardt, and M. L. Raghavan, “Three-dimensional characterization of regional lung deformation,” *J. Biomechanics*, vol. 44, no. 13, pp. 2489–2495, 2011, <https://doi.org/10.1016/j.jbiomech.2011.06.009>.
- S. M. Galib, H. K. Lee, C. L. Guy, M. J. Riblett, and G. D. Hugo, “A fast and scalable method for quality assurance of formable image registration on lung CT scans using convolutional neural networks,” *Med. Phys.*, vol. 47, no. 1, pp. 99–109, 2020, <https://doi.org/10.1002/mp.13890>.
- B. D. de Senneville, J. V. Manjón, and P. Coupé, “RegQCNNet: Deep quality control for image-to-template brain MRI affine registration,” *Phys. Medicine Biol.*, vol. 65, no. 22, art. 225022, 2020, <https://doi.org/10.1088/1361-6560/abb6be>.
- H. Sokooti, S. Yousefi, M. S. Elmahdy, B. P. Lelieveldt, and M. Starling, “Hierarchical prediction of registration misalignment using a convolutional LSTM: Application to chest CT scans,” *IEEE Access*, vol. 9, pp. 62008–62020, 2021, <https://doi.org/10.1109/ACCESS.2021.3074124>.
- K. A. Eppenhof and J. P. Pluim, “Error estimation of deformable image registration of pulmonary CT scans using convolutional neural networks,” *J. Med. Imag.*, vol. 5, no. 2, art. 024003, 2018, <https://doi.org/10.1117/1.JMI.5.2.024003>.
- G. Saygili, “Predicting medical image registration error with block-matching using three orthogonal planes approach,” *Signal Image Video Process.*, vol. 14, no. 6, pp. 1099–1106, 2020, <https://doi.org/10.1007/s11760-020-01650-2>.
- E. Castillo, R. Castillo, Y. Vinogradskiy, and T. Guerrero, “The numerical stability of transformation-based CT ventilation,” *Int. J. Comput. Assisted Radiol. Surgery*, vol. 12, no. 4, pp. 569–580, 2017, <https://doi.org/10.1007/s11548-016-1509-x>.
- E. Castillo, R. Castillo, Y. Vinogradskiy, M. Dougherty, D. Solis, et al., “Robust CT ventilation from the integral formulation of the Jacobian,” *Med. Phys.*, vol. 46, no. 5, pp. 2115–2125, 2019, <https://doi.org/10.1002/mp.13453>.
- W. Shao, S. E. Gerard, Y. Pan, T. J. Patton, J. M. Reinhardt, et al., “Sensitivity analysis of Jacobian determinant used in treatment planning for lung cancer,” in *Proc. SPIE 10574, Med. Imaging 2018: Image Process.*, vol. 10574, art. 1057418, 2018, <https://doi.org/10.1117/12.2293920>.
- “Chest Imaging Platform (CIP),” Harvard University, <https://chestimagingplatform.org/> (accessed Dec. 20, 2021).
- “ANTs, Advanced Normalizations Tools,” GitHub, <http://stnava.github.io/ANTs/> (accessed Dec. 20, 2021).
- P. Cook, “Anatomy of an antsRegistration call,” GitHub, <https://github.com/ANTsX/ANTs/wiki/Anatomy-of-an-antsRegistration-call> (accessed Dec. 20, 2021).
- B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, “Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain,” *Med. Image Anal.*, vol. 12, no. 1, pp. 26–41, 2008, <https://doi.org/10.1016/j.media.2007.06.004>.
- M. Modat, J. McClelland, and S. Ourselin, “Lung registration using the NiftyReg package,” in *Proc. MICCAI2010 Workshop: Med. Image Anal. for the Clinic—A Grand Challenge*, pp. 33–42, 2010.
- R. Castillo, E. Castillo, D. Fuentes, M. Ahmad, A. M. Wood, M. S. Ludwig, and T. Guerrero, “A reference dataset for deformable image registration spatial accuracy evaluation using the COPDgene study archive,” *Phys. Medicine Biol.*, vol. 58, no. 9, pp. 2861–2877, 2013, <https://doi.org/10.1088/0031-9155/58/9/2861>.



Rahul Hingorani, Research and Exploratory Development Department, Johns Hopkins University Applied Physics Laboratory, Laurel, MD

Rahul Hingorani is a data scientist in APL's Research and Exploratory Development Department. He received his BS and MS in electrical engineering from the University of Michigan with a focus in signal processing and machine learning. His current research interests are in implementing signal processing and machine learning algorithms for biomedical and healthcare applications. His email address is rahul.hingorani@jhuapl.edu.



Nicole L. Brown, Research and Exploratory Development Department, Johns Hopkins University Applied Physics Laboratory, Laurel, MD

Nicole L. Brown is a data scientist in APL's Research and Exploratory Development Department. She has a BS in molecular and cellular biology from Johns Hopkins University. Her recent work focuses on medical image processing and machine learning for analysis of time series data and connectomes. Her email address is nicole.brown@jhuapl.edu.



Christopher M. Cervantes, Asymmetric Operations Sector, Johns Hopkins University Applied Physics Laboratory, Laurel, MD

Christopher M. Cervantes is a data scientist in APL's Asymmetric Operations Sector. He has a BS in computer science from North Central College and an MS in computer science from the University of Illinois at Urbana-Champaign. Christopher's primary area of expertise is natural language processing (NLP). His research focuses on scene understanding—which is at the intersection of NLP and computer vision—and knowledge representation, leveraging machine learning techniques to extend and reason over

knowledge graphs. In his current role, Christopher supports the NLP, machine learning, and data analysis needs for projects in the biomedical and intelligence domains. His email address is christopher.cervantes@jhuapl.edu.



Robert H. Brown, Johns Hopkins School of Medicine and Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

Robert H. Brown is a professor of anesthesiology and critical care medicine at the Johns Hopkins School of Medicine and Johns Hopkins Bloomberg School of Public Health. He received a BA in economics from the University of Rochester and an MD from the University of Pennsylvania. His residency in anesthesiology and fellowship training in chronic pain management were at the Johns Hopkins Hospital. He subsequently received an MPH from the Johns Hopkins School of Public Health. His research interests include physiologic imaging of airway reactivity in asthma and COPD, the structure function relationship of airway-parenchymal interaction, and HIV on the acceleration of emphysema. He has additional research interests in occupational latex allergies in health care workers. More recently, he has been examining the effects of changes the levels of environmental toxins in the body after bariatric surgery.



Andrew S. Gearhart, Research and Exploratory Development Department, Johns Hopkins University Applied Physics Laboratory, Laurel, MD

Andrew S. Gearhart is a data scientist in APL's Research and Exploratory Development Department. He received a BS in computer science and mathematics and a BA in psychology from the University of Delaware, as well as a PhD in computer science from the University of California, Berkeley. His research interests include the study of software diversity, scientific machine learning, and the application of data analytics to further precision medicine. His e-mail address is andrew.gearhart@jhuapl.edu.