

# Nodule volume change estimation in thoracic CT using sphere fitting, morphological segmentation and image registration

Thomas Duindam<sup>1</sup>, Bartjan de Hoop<sup>2</sup>, and Bram van Ginneken<sup>1</sup>

<sup>1</sup> Image Sciences Institute, University Medical Center Utrecht, the Netherlands

<sup>2</sup> Department of Radiology, University Medical Center Utrecht, the Netherlands

**Abstract.** Three methods are presented to assess the relative volume change of a pulmonary nodule between two chest CT scans when the nodule location in both scans is provided. The first method fits a sphere around both nodules and computes the volume of dense tissue in that sphere. The second method segments both nodules using thresholding, component labeling and morphological processing. The third method applies non-rigid registration to transform the first to the second scan and applies that transformation to a segmentation of the nodule in the first scan to obtain a segmentation of the nodule in the second scan. All methods are applied to 50 nodule pairs from the VOLCANO'09 challenge. These cases are divided by a radiologist in stable and growing pairs. All methods produce lower mean volume change for the stable cases compared to the growing nodules, but the distributions overlap considerably. Moreover, the correlation between the volume change estimates produced by the three methods is modest. This shows that nodule volume change assessment is a complicated problem.

## 1 Introduction

Pulmonary nodules occur frequently in thoracic CT scans. In studies where subjects at high risk for developing lung cancer were scanned with low-dose CT, 8% to 51% of all subjects had at least one nodule [1]. The most important question that follows detection of a nodule is whether the lesion may be malignant. Comparison with available prior imaging of the same nodule is a critical step in answering this question. Malignant lesions tend to have a volume doubling time between 20 and 400 days. Nodules with more rapid growth likely represent inflammation and nodules that are stable for a long period of time are likely benign. Nodule growth rate is one of the most important characteristics in the determination of its probability of malignancy. Therefore, a substantial amount of research has been devoted to the design of methods to accurately estimate the growth rate of nodules. In this paper, growth  $\Delta V$  is defined as the proportional change in size of the lesion between the two scans relative to the size of the lesion in the first scan. Let  $V_1$  and  $V_2$  be the volume of the lesion in scan 1 and 2, respectively, then we define the growth as

$$\Delta V = \frac{V_2 - V_1}{V_1}. \quad (1)$$

In this definition  $\Delta V$  can take values from -1 (the nodule vanishes completely) to 0 (nodule is stable) to 1 (nodule doubles in volume) to infinity. Because the interval  $[-1,0]$  corresponds to  $[0,\infty]$ , values for  $\Delta V$  are often not normally distributed.

The higher the accuracy of the growth estimation method, the shorter the time window can be between a baseline scan where a nodule is identified and a follow-up scan used to determine growth rate. For example, if a nodule has a doubling time of 100 days, its  $\Delta V$  will be around 0.22 after 30 days. If the error in  $\Delta V$  is also around 0.2, one cannot conclude decisively after only one month whether a measured growth is real or a measurement error. Waiting for three months would be enough in this example scenario, but if the nodule was a fast growing cancer with a doubling time of 20 days, a three month delay in detection may be fatal. Most ongoing trials for lung cancer screening use two to six months follow-up scans in case an intermediately sized nodule (with a volume around 50 to 500 mm<sup>3</sup>) has been detected. If more accurate methods for growth assessment were available, this interval could be shortened.

The data used in this paper is from the Volcano '09 Challenge (<http://www.via.cornell.edu/challenge/>). The Volcano website states the goal of the challenge as follows:

The goal of this study is to compare the outcomes of various algorithms measuring the change in volume of pulmonary nodules from CT scans using a common dataset and performance evaluation method.

The dataset used contains two CT-scans of each case, taken at different, unknown points in time. The location of the approximate center of the nodule is provided in an accompanying file. Image resolution, nodule size and scan parameters vary from case to case.

An obvious approach for calculating the change of volume in a nodule over time is to segment the nodule in both scans (or in every scan if more than two scans are available) and compare the computed volumes. Nodule segmentation is not a trivial task however. Several studies have shown that when the same nodule is segmented in two scans taken only minutes apart,  $\Delta V$  can vary substantially. De Hoop et al. [2] have shown in an evaluation of six commercially available lung nodule segmentation toolkits that a difference of 20% is required to be able to say confidently that a nodule has truly changed (the 20% corresponded to twice the standard deviation of the  $\Delta V$  measurements).

One important reason why it is difficult to automatically segment nodules is that other dense structures are often attached to the nodule. These structures, which can be vessels, the pleural surface, fissures or dense abnormalities, could be erroneously included in the nodule segmentation. If this would happen in both segmentations the estimated volume change might still be correct. But it is possible that a nodule in a baseline scan is not in direct contact with a

vessel, but in a follow-up scan due to nodule growth or differences in inspiration level, the nodule is in contact with a vessel and that vessel is included in the segmentation. Such a segmentation inconsistency will cause the extent of nodule growth to be overestimated. This has been pointed out by Reeves et al. [3] and they have proposed a method to repair such inconsistencies. In this work we compare a nodule segmentation method (SEGM) which independently segments the nodule in both scans to compute growth with two other approaches which may be less sensitive to segmentation inconsistencies.

Another approach is to define an equally sized volume of interest in both scans and apply a threshold to the voxels inside the volume. If we assume that the volume of the vasculature surrounding the nodule remains stable over time, an estimation of nodule volume change can be made by counting the voxels remaining in the volume of interest after applying the threshold. In this work, we use a sphere for the volume of interest. We shall refer to this method as SPHERE.

One other approach is to segment the nodule only in the first scan and use the deformation field obtained by non rigidly registering both scans to transform the first segmentation. In this way, a segmentation for the second nodule is derived. This method has proven to be successful for artificial nodules [4]. The strength of this approach is that slight errors the segmentation will not lead to large errors in the measured relative volume change, as long as the nodules can be registered successfully. For example, if a part of an attached vessel is erroneously included in the first segmentation, the transformed segmentation will also include that vessel part and it will probably not be changed much. This method will be referred to as REG throughout this paper.

We present an evaluation of the three aforementioned approaches to nodule volumetry. The output of all three methods with various parameter settings has been evaluated, visually assessed, and compared with visual evaluation of nodule change by a human expert. Section 2 provides some details about the data used in the VOLCANO challenge. The methods are detailed in Section 3. Experiments and results are presented in Section 4. We discuss the results and draw conclusions in Section 5.

## 2 Data

The evaluation dataset consists of 50 pairs of nodules. Many pairs have been taken during the initial stages of a lung biopsy and should therefore not exhibit any growth. For others it is known, through follow-up, that they have grown. Data originated from Weill Cornell Medical College, New York, USA. Data is divided into three categories. The first category consists of 27 nodules visible on two scans of 1.25 mm slice thickness with little observed size change, and a range in diameter from approximately 4 to 24 mm. The second category of nodules included 13 nodules imaged on either two 2.5 mm scans or one 1.25 mm scan and one 2.5 or 5.0 mm scan to examine the effect of slice thickness on the performance. The nodules ranged in size from approximately 8 to 30 mm. The

third category consists of an additional 9 nodules on two 1.25 mm scans, but with a large size change; these nodules ranged in size from approximately 5 to 14 mm.

Only a limited number of slices were made available, and a list of locations of the nodules in both scans. It was not given to which of the three categories above the nodule pair belonged. Data was anonymized, and the scan order (which scan was made first in time) was not given. Our results were made available to the VOLCANO organizers for further analysis.

To be able to analyze the results of our methods, we asked a radiologist to visually inspect the cases. He used a side-by-side viewer that was developed in our group, which also allowed to watch the scan pairs after rigid registration. He assigned each pair to one of the following five classes: definitely shrinkage, possibly shrinkage, stable, possibly growth and definitely growth. For our subsequent analysis we grouped the cases of possible and definite growth/shrinkage, and swapped all pairs with shrinkage so that two groups remained: stable nodules (28 cases) and growing nodules (22 cases).

### 3 Method

All methods employ the result of a lung segmentation. Typically a 3D lung segmentation algorithm would be used for this purpose. However, the scans in the VOLCANO challenge only cover a limited axial field of view. Therefore a 2D lung segmentation was applied slice by slice. The segmentation method consists of a sequence of standard image processing steps (thresholding, component labeling, hole filling, morphological closing) and is described in [5].

#### 3.1 Morphological Segmentation (SEGM)

If a segmentation of the nodule is in both scans is present, computing the change in volume is a trivial task. The segmentation algorithm used for this study is a similar and somewhat simplified version of the algorithm described in [6].

The first step of the segmentation algorithm is very similar to the SPHERE approach: first a  $50 \times 50 \times 50$  mm volume of interest is super sampled to isotropic voxels of 0.5 mm, followed by a thresholding operation with threshold  $t$  using the lung segmentation as a mask. The next steps of the algorithm aim to remove vasculature and noise from the nodule segmentation. First a connected component analysis is executed to isolate the largest connected component. This removes noise and disconnected vessels from the segmentation. After this step only the nodule and vessels with connections to the nodule remain. To remove these vessels mathematical morphology is used. With an opening with a spherical kernel, vessels connected to the nodule are removed. The diameter  $d$  of this kernel is the second parameter for the segmentation algorithm. The opening will also remove voxels on the edge of the nodule, and smooth its surface. To reduce this effect a conditional dilation is used with an decreasing kernel size [6]. This will regrow the nodule back to its original size and reconstruct most features of the nodule

surface. Not all features can be reconstructed, but since relative volume change is more relevant than the absolute volume of a nodule this is deemed acceptable. It is possible that the opening will completely remove all candidate voxels from the segmentation, in this case all morphological operations are canceled, and the result of the connected component analysis is used for further computations. To ensure segmentation consistency as much as possible, we kept the segmentation parameters  $t$  and  $d$  the same for both scans in a pair.

### 3.2 Sphere fitting (SPHERE)

The underlying assumption of this technique is that apart from the nodule, the volume of anatomical structures inside the lung should not have changed significantly over time. Although it seems unlikely that this assumption will hold over the entire lung, especially in patients with gross pathology, there will likely be a small volume of interest around the nodule for which this assumption will hold. If we can find such a volume of interest for both scans, simply counting the number of dense voxels will suffice for making an estimation of nodule growth.

The first step of this algorithm is to take a  $50 \times 50 \times 50$  mm volume of interest centered at the nodule center, which is provided in the Volcano data set. This volume of interest is super sampled to isotropic voxels of 0.5 mm resolution. This reduces partial volume effects and also simplifies calculations as the voxels in both scans now have equal dimensions.

Next, a threshold  $t$  is applied over this volume of interest, excluding voxels outside the lung field, where we use the predetermined lung segmentation as a mask. This leaves us with all voxels denser than  $t$  inside the lung (we call these dense lung voxels). Effectively this means that only nodule and vessel voxels remain. If we would compute the volume change at this point we would not be able to quantify growth accurately, since the volume of the nodule might be very small compared to that of the surrounding vasculature. Therefore we refine the volume of interest as follows. We define a spherical volume of interest at the approximate nodule-center, and iteratively increase the diameter of this sphere by two voxels (1, 3, 5, and so on). We keep increasing the diameter until the next iteration would not add a sufficient number of dense lung voxels. We define this cut-off ratio  $r$  as the number of dense lung voxels added in the next iteration divided by the total number of voxels added in the next iteration. Thus this algorithm has two parameters, the threshold  $t$  and the cut-off ratio  $r$ .

This leaves us with two spherical volumes of interest, possibly with different diameters. This will introduce a bias in the measurement. The larger of the two is therefore chosen and applied to both scans. The final volume in  $\text{mm}^3$  is then computed by multiplying the number of dense lung voxels in the spheres by  $0.5^3$ .

### 3.3 Segmentation by registration (REG)

In this approach we use a segmentation of the nodule in one scan (the moving scan), and to avoid segmentation inconsistencies that could occur if the same algorithm were applied independently to the other (fixed) scan, we instead use

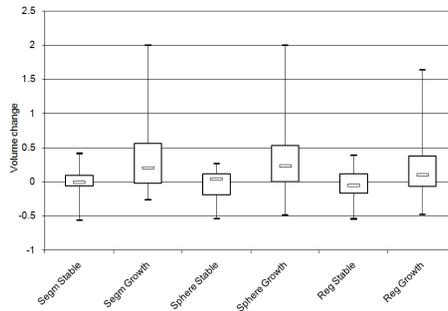
image registration to elastically deform the moving so as to resemble the fixed scan. We use the freely available package `elastix` [7], version 3.9 for registration. The registration software is able to handle anisotropic voxel sizes internally.

To force the registration software to align only nodules and not the surrounding tissue a mask is used. To make the mask, first both nodules are segmented using the morphological segmentation algorithm outlined above. This resulting segmentation is then dilated to ensure that the entire nodule is represented in the mask, even if under-segmentation has occurred. To ensure that the same structures are present in both masked areas the biggest of the two masks is used for both scans. Finally the parts of the mask that are outside the lungs are removed using the lung segmentation mentioned before.

The resulting transformation is applied to the segmentation of the nodule in the moving scan, and both volumes are compared. This method is also described in [4], and has proven to be effective for artificial nodules. It is not clear what mask, if any, was used for the experiments in [4]. The only parameters the REG method takes are the threshold  $t$  and the kernel size  $d$  for the segmentation.

Setting for the registration method were normalized mutual information as similarity measure, first a rough alignment with an affine transformation, followed by a non-rigid registration modeled by B-splines. Three resolutions were used in both affine and non-rigid stages, and always 500 iterations were performed per resolution level. The other settings were the defaults of the package.

## 4 Experiments and results



**Fig. 1.**  $\Delta V$  of all three methods for stable and growing nodules. Here parameters were determined by a user. Whiskers indicate minimum and maximum values.

Two types of experiments have been carried out. First, for all three methods, several parameter settings were evaluated. Table 1, 2 and 3 give the results.

Second, all three methods were applied using an optimal setting for the two parameters of each method, determined visually by a human observer. This

**Table 1.** Variability of  $\Delta V$  outcome for the SPHERE method using different values for threshold  $t$  and cutoff  $r$ . Stable and growth cases are treated separately. For each setting the average, standard deviation(SD), minimum, 1st quartile, median, 3rd quartile and the maximum value are reported. The row with label semi auto are the results from the semi-automatic experiments.

$t$	$r$	Avg	SD	Min	q1	Median	q3	Max
<b>Stable nodules</b>								
-600	0.5	0.118	0.389	-0.377	-0.031	0.015	0.199	1.263
-600	0.6	0.057	0.378	-0.550	-0.111	0.010	0.111	1.263
-600	0.9	0.181	0.602	-0.618	-0.087	-0.004	0.289	2.000
-500	0.5	0.015	0.261	-0.422	-0.065	0.005	0.071	0.792
-500	0.6	0.003	0.231	-0.573	-0.060	0.018	0.117	0.445
-500	0.9	0.168	0.621	-0.378	-0.116	-0.004	0.094	2.000
-400	0.5	-0.010	0.231	-0.619	-0.090	0.005	0.061	0.423
-400	0.6	-0.007	0.253	-0.668	-0.085	0.021	0.101	0.446
-400	0.9	-0.011	0.326	-0.465	-0.234	-0.029	0.082	0.928
Semi	auto	-0.024	0.217	-0.540	-0.190	0.041	0.118	0.271
<b>Growing nodules</b>								
-600	0.5	0.001	0.179	-0.433	-0.148	0.012	0.134	0.257
-600	0.6	0.000	0.205	-0.507	-0.174	0.030	0.137	0.462
-600	0.9	0.053	0.468	-0.507	-0.091	-0.002	0.039	2.000
-500	0.5	-0.015	0.201	-0.539	-0.171	0.035	0.112	0.239
-500	0.6	0.020	0.331	-0.539	-0.179	0.039	0.100	1.239
-500	0.9	0.066	0.493	-0.609	-0.107	0.016	0.035	2.000
-400	0.5	-0.023	0.224	-0.572	-0.193	0.042	0.116	0.271
-400	0.6	0.021	0.360	-0.572	-0.193	0.046	0.124	1.319
-400	0.9	0.094	0.550	-0.650	-0.163	0.014	0.043	2.000
Semi	auto	0.381	0.634	-0.485	0.003	0.232	0.532	2.000

makes our approaches semi-automatic, but the amount of interaction required is minimal. Figure 1 shows the distribution of  $\Delta V$  for growing and stable nodules for all methods. Note how the REG method works well for stable nodules (i.e.  $\Delta V$  close to zero), but not for growing nodules, where q1 is below zero. The results for the SEGM and the SPHERE method are similar in the case of growing nodules, but for stable nodules the distribution for the SEGM method has a lower variability.

Figures 2 to 4 provide examples of the results of each method, including a case where the SEGM method leads to segmentation inconsistencies. Scatterplots of the results of different methods for stable and growing cases are given in Figures 5 and 6. Correlation between SEGM and REG is 0.667, between REG and SPHERE 0.479 and the highest correlation is achieved between SEGM and SPHERE with 0.832.

**Table 2.** Variability of  $\Delta V$  outcome for the SEGM method using different values for threshold  $t$  and kernel size  $d$ . Stable and growth cases are treated separately. For each setting the average, standard deviation(SD), minimum, 1st quartile, median, 3rd quartile and the maximum value are reported. The row with label semi auto are the results from the semi-automatic experiments.

$t$	$d$	Avg	SD	Min	q1	Median	q3	Max
-600	3	-0.118	0.389	-1.263	-0.199	-0.015	0.031	0.377
-600	4	-0.057	0.378	-1.263	-0.111	-0.010	0.111	0.550
-600	6	-0.181	0.602	-2.000	-0.289	0.004	0.087	0.618
-500	3	-0.015	0.261	-0.792	-0.071	-0.005	0.065	0.422
-500	4	-0.003	0.231	-0.445	-0.117	-0.018	0.060	0.573
-500	6	-0.168	0.621	-2.000	-0.094	0.004	0.116	0.378
-400	3	-0.122	0.837	-2.000	-0.255	-0.026	0.323	2.000
-400	4	0.007	0.253	-0.446	-0.101	-0.021	0.085	0.668
-400	6	0.011	0.326	-0.928	-0.082	0.029	0.234	0.465
Semi auto		0.003	0.214	-0.560	-0.059	-0.001	0.093	0.420
<b>Growing nodules</b>								
-600	3	0.017	0.712	-1.727	-0.195	0.016	0.294	2.000
-600	4	0.033	0.729	-1.483	-0.254	0.007	0.314	2.000
-600	6	0.136	0.757	-1.392	-0.129	0.054	0.271	2.000
-500	3	-0.044	0.709	-1.681	-0.174	-0.001	0.315	2.000
-500	4	-0.064	0.724	-1.717	-0.267	-0.021	0.377	2.000
-500	6	0.044	0.750	-1.714	-0.136	0.036	0.381	1.886
-400	3	-0.122	0.837	-2.000	-0.255	-0.026	0.323	2.000
-400	4	-0.138	0.883	-2.000	-0.190	-0.012	0.300	2.000
-400	6	0.027	0.758	-2.000	-0.100	0.075	0.351	2.000
Semi auto		0.412	0.632	-0.265	-0.016	0.203	0.563	2.000

## 5 Discussion and Conclusion

It is difficult to assess the value of the three different methods that have been applied to the 50 nodule pairs of the VOLCANO'09 challenge, because we do not have access to information which nodule pairs are stable and which ones exhibit change. We therefore asked a radiologist to visually assess change in all pairs. Manually assessing small volume changes is hard for a human observer, and the radiologist repeatedly expressed his uncertainty, especially for nodules labeled as stable. It seems likely that some cases classified by the radiologist as stable are in reality slightly growing or shrinking. In some cases the automatic methods measured a change and the segmentation results of the methods were visually convincing.

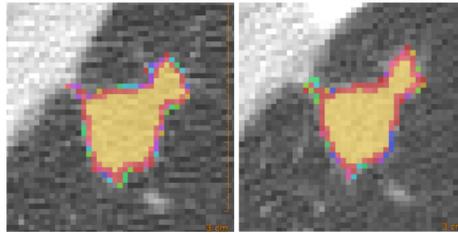
Nevertheless, even in absence of any truth, we can conclude that the methods agree only moderately well with each other. The scatter plots in Figures 5 and 6 show large disagreement between the methods in a substantial amount of cases. Figure 1 shows that in our results, SEGM and SPHERE lead to a better separation between the stable and growing nodules than the REG method.

**Table 3.** Variability of  $\Delta V$  outcome for the REG method using different values for threshold  $t$  and kernel size  $d$ . Stable and growth cases are treated separately. For each setting the average, standard deviation(SD), minimum, 1st quartile, median, 3rd quartile and the maximum value are reported. The row with label semi auto are the results from the semi-automatic experiments.

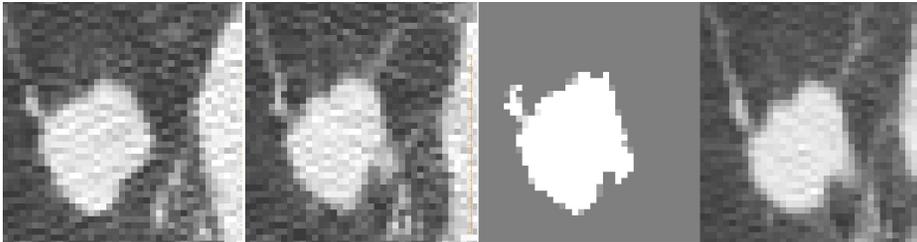
$t$	$d$	Avg	SD	Min	q1	Median	q3	Max
-400	3	0.033	0.208	-0.391	-0.116	0.037	0.148	0.415
-400	4	0.043	0.198	-0.203	-0.139	-0.022	0.137	0.446
-400	6	0.049	0.583	-0.542	-0.199	-0.050	0.181	2.000
-500	3	0.021	0.192	-0.365	-0.126	0.018	0.149	0.404
-500	4	0.030	0.208	-0.368	-0.138	0.082	0.142	0.433
-500	6	0.050	0.237	-0.332	-0.099	0.054	0.136	0.542
-600	3	0.109	0.337	-0.212	-0.112	0.054	0.188	1.308
-600	4	0.094	0.343	-0.379	-0.115	0.054	0.167	1.308
-600	6	0.104	0.326	-0.553	-0.058	0.093	0.189	1.276
Semi auto		-0.033	0.209	-0.542	-0.165	-0.050	0.120	0.390
<b>Growing nodules</b>								
-400	3	0.116	0.575	-0.851	-0.167	-0.040	0.138	2.000
-400	4	0.103	0.582	-0.918	-0.150	-0.073	0.140	2.000
-400	6	0.106	0.558	-0.897	-0.215	-0.023	0.233	1.641
-500	3	0.059	0.377	-0.517	-0.105	-0.024	0.110	1.470
-500	4	0.088	0.335	-0.722	-0.084	0.104	0.212	0.970
-500	6	0.126	0.313	-0.405	-0.044	0.102	0.242	0.936
-600	3	0.114	0.442	-0.856	-0.111	0.086	0.156	1.382
-600	4	0.143	0.392	-0.312	-0.092	0.080	0.203	1.329
-600	6	0.104	0.326	-0.553	-0.058	0.093	0.189	1.276
Semi auto		0.206	0.423	-0.478	-0.067	0.103	0.377	1.641

Visual inspection of the results from the REG method indicated that the registration results were sometimes incorrect. Finding more stable settings for the registration method is therefore an important direction for future research, especially since this method was reported to work well in the work of Kabus et al. [4]. The final metric value reported by the registration software can be used to assess the quality of the registration and thus the quality of the measured change. This can be used to construct a fully automatic method, or to report the certainty of the system that the reported volume change is correct.

SEGM is very sensitive to the parameters used for the segmentation. These parameters vary from case to case and are can be hard to find, although in most cases it is a straightforward procedure to find good settings. SPHERE is not as sensitive to over-segmentation as SEGM, but it does rely heavily on the reported center point actually being in the center of the nodule. If the reported seed point of the nodule is not in the center, SPHERE will likely fail. An obvious solution to this problem is to include the determination of the sphere center point in the fitting procedure. Another weakness of SPHERE is the assumption of a spherical nodule, and spherical nodule growth. A more complicated model for nodule shape



**Fig. 2.** Example of a successful segmentation by the SEGM method. The different colors in the image signify the amount of nodule in the voxel after sub-sampling, yellow means 100% of the voxel is nodule, red means 90%, pink means 80%, and purple means 70%. The other colors represent lower percentages of nodule in the voxel. The window center is -600 HU and the width of the window is 1600 HU (in all figures). The reported volumes by the SEGM method were 1776.9 mm<sup>3</sup> for scan 1 and 1442.5 mm<sup>3</sup> for scan 2, a  $\Delta V$  of -0.19.

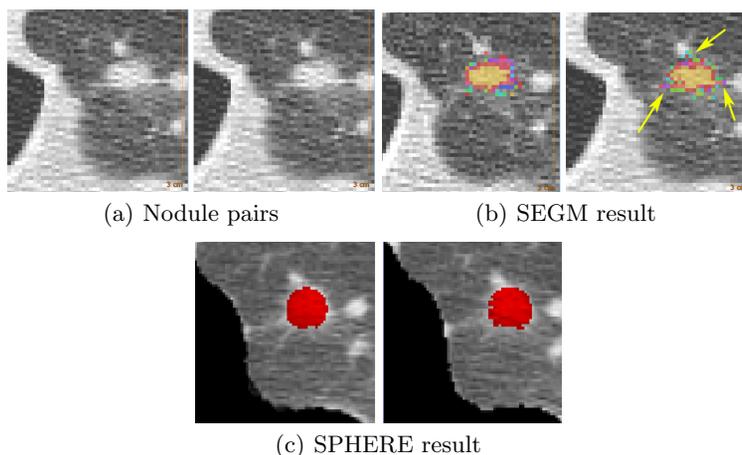


**Fig. 3.** Example of a successful segmentation by the REG method. From left to right; the nodule in the first scan, the transformed nodule after registration, the transformed segmentation and finally the nodule in the second scan. This case has been labeled shrinking by the human expert. The REG method reports a negative volume change ( $V_1 = 1511$  mm<sup>3</sup>,  $V_2 = 1336$  mm<sup>3</sup> and  $\Delta V = -0.116$ ). The parameters for the segmentation used by REG are  $t = -400$  HU and  $d = 4$  mm).

may improve the reliability of volume change assessment by SPHERE, but will also make the method much more complicated.

Finally, we note that it is not clear if measuring the relative volume change  $\Delta V$  is the most important parameter to answer the question that is clinically the most relevant: is a nodule malignant or benign? The absolute volume change, or the nodule mass (change) may be better predictors, especially when used in conjunction with other features.

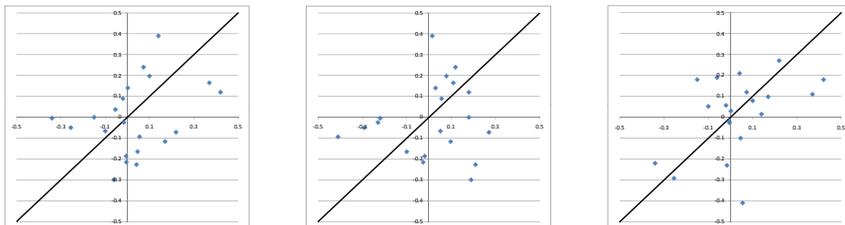
In conclusion, we have presented three simple methods for nodule volume change assessment and applied these to a public database provided by the VOLCANO'09 challenge. Although the methods produce visually convincing results in many cases, correlation between the methods is only moderately good, and the results do not show a clear separation between cases rated as stable versus growing by a radiologist.



**Fig. 4.** An example where the SPHERE method ( $t = -400\text{HU}$ ,  $r = 0.5$ ) works better than the SEGM method ( $t = -400\text{HU}$ ,  $d = 3\text{ mm}$ ). The segmentation made by SEGM grows into some vessels in the second scan (indicated with arrows), but not in the other. This leads to an overestimation of nodule volume in scan 2. The SPHERE method clearly under-segments the nodule in both scans, but it does so in a consistent manner. The radiologist labeled this case as stable. SEGM reports growth ( $V1 = 365.1\text{ mm}^3$ ,  $V2 = 416.8\text{ mm}^3$ ,  $\Delta V = 0.14$ ), SPHERE reports a nearly stable nodule ( $\Delta V = 0.015$ ). Due to super-sampling figure 4(c) appears blurred.

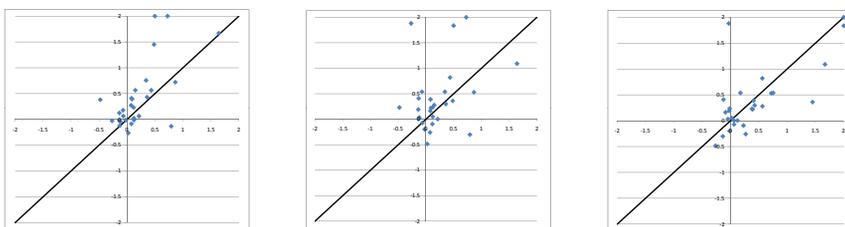
## References

1. Wahidi, M.M., Govert, J.A., Goudar, R.K., Gould, M.K., McCrory, D.C., the American College of Chest Physicians: Evidence for the treatment of patients with pulmonary nodules: when is it lung cancer?: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest* **132**(3 Suppl) (2007) 94S–107S
2. de Hoop, B., Gietema, H., van Ginneken, B., Zanen, P., Groenewegen, G., Prokop, M.: A comparison of six software packages for evaluation of solid lung nodules using semi-automated volumetry: what is the minimum increase in size to detect growth in repeated CT examinations. *European Radiology* **19**(4) (2009) 800–808
3. Reeves, A.P., Chan, A.B., Yankelevitz, D.F., Henschke, C.I., Kressler, B., Kostis, W.J.: On measuring the change in size of pulmonary nodules. *IEEE Trans Med Imaging* **25**(4) (2006) 435–450
4. Kabus, S., Müller, F., Wiemker, R., Fischer, B.: Robust lung nodule growth measurement by combining registration and segmentation. In: *The First International Workshop on Pulmonary Image Analysis*, Lulu.com 15–23
5. Sluimer, I.C., Prokop, M., Hartmann, I., van Ginneken, B.: Automated classification of hyperlucency, fibrosis, ground glass, solid and focal lesions in high resolution CT of the lung. *Medical Physics* **33**(7) (2006) 2610–2620
6. Kostis, W.J., Reeves, A.P., Yankelevitz, D.F., Henschke, C.I.: Three-dimensional segmentation and growth rate estimation of small pulmonary nodules in helical CT images. *IEEE Trans Med Imaging* **22**(10) (2003) 1259–1274
7. Klein, S., Staring, M.: Elastix. <http://www.isi.uu.nl/Elastix/> (Accessed March 30, 2009)



(a) REG versus SEGM (b) REG versus SPHERE (c) SEGM versus SPHERE

**Fig. 5.** Variability in outcome for stable nodules after semi-automatic experiments. Note that correlation is strong between segmentation and sphere, and particularly weak between registration and sphere.



(a) REG versus SEGM (b) REG versus SPHERE (c) SEGM versus SPHERE

**Fig. 6.** Variability in outcome for growing nodules after semi-automatic experiments. Note that correlation is strong between segmentation and sphere, and particularly weak between registration and sphere.