CT Based Semi-Automated Method for Pneumonia Severity in Mice

Ansh Johri 1,*, Lewis Hu 2

1Department of Biomedical Engineering, Columbia University, New York, NY 10027; *National Institutes of Health, National Heart, Lung, Blood Institute, Washington, DC 20010.

Abstract

Misdagnosis of community-acquired pneumonia is an important clinical problem, leading to a high rate of mortality. Diagnoses are typically conducted using two-dimensional chest x-rays, which have shown to be time-consuming and inaccurate. In an effort to improve the current diagnostic method, we utilized Micro-Computed Tomography (MicroCT) and image analysis software to develop a diagnostic algorithm that can quantitatively assess the severity of pneumonia in mice. We believed this method would provide more immediate, precise, and accurate diagnoses as opposed to the qualitative assessments done by radiologists at present, because MicroCT provides opportunities for non-invasive radiographic endpoints for pneumonia studies. A quantitative scoring of previously obtained Computed Tomography (CT) scans of pneumonia infected and control mice was developed. At the endpoint of 168 hours, each of the mice was categorized as either a) a Saline (control)-injected mouse (total=13), a Pneumonia-injected Survivor (total=11), or a Pneumonia-non-injected Non-survivor (total=11). Comparison tests were then completed, including Saline vs. All Pneumonia Injected Mice, Pneumonia Survivors vs. All Survivors (both Saline & Pneumonia) vs. Pneumonia Non-survivors. In all three comparisons, the semi-automated algorithm was better able to distinguish between the different groups than radiologists using two-dimensional chest x-rays of the mice’s lungs, with p-values of 0.001, 0.039, and 0.001 for the semi-automated algorithm, and 0.004, 0.581, 0.058 for the radiologists, respectively.

Key Words: Community-acquired pneumonia, Computed Tomography

Introduction

In the United States, pneumonia is the sixth leading cause of death and the number one infectious disease killer (M. S. Niederman, 1998). The incidence is an inflammation of one or both lungs caused by an infection from bacteria, viruses or fungi. The infection causes the alveoli of the lungs to become inflamed and filled with fluid, which leads to symptoms such as cough, fever and respiratory breathing difficulties (Jelic, 2005). Often times, pneumonia occurs as a secondary infection when the immune system of a person is already weakened due to prior infection, such as an upper respiratory tract infection. This primary infection causes inflammation in the inner lining of airways that leaves the patient susceptible to the secondary infections such as pneumonia (Boone, 2004).

Pneumonia can be classified according to the population affected. Hospital-acquired pneumonia is acquired when a patient breathes germs during a hospital stay for another illness. People are most prone to hospital-acquired pneumonia while on a mechanical ventilator, since potentially pneumonia-causing bacteria may be blown directly into the lungs. The most common type of pneumonia is community-acquired. Community-acquired pneumonia is acquired outside of hospitals and other health care settings, with about 5.6 million people getting infected every year in the USA and 1.1 million requiring hospitalization (M. S. Niederman, 1998). Community-acquired pneumonia is an important clinical problem, with high rates of misdiagnosis and mortality. Current methods to diagnose pneumonia rely on two dimensional (2D) chest x-rays, which are known to have low sensitivity early in the course of pneumonia (Mohd). Radiologists typically score six lung zones (upper, lower, anterior, posterior, and lower, on the right and left sides) for each patient on a scale of 0 to 4, such that zero is normal, and the maximum possible abnormal score is 24 for the combined zones: 0 represents 0% pneumonia involvement, 1 represents 25% involvement, 2 represents up to 50%, 3 represents up to 75%, and 4 represents up to 100% (Armbrust, 2005). These chest x-rays may take days to diagnose the severity of pneumonia, in which time immunocompromised patients, such as patients with HIV/AIDS, cancer, diabetes, or sickle cell anemia, may reach a severity beyond curing (Smergal, 2008; Stuart, 2008). For example, immuno-compromized patients with pneumonia have a mortality rate of 12% (Mohd). Furthermore, radiologists are often inconsistent with their diagnoses; two radiologists may judge the severity of pneumonia in patients very differently, leading to possible misdiagnosis (I. Hsu, 2007). Thus, imaging techniques to evaluate pneumonia earlier and with more accuracy would be important diagnostic tools for clinicians. Imaging information could also be used to guide decisions on the clinical care needed, such as whether to hospitalize or to treat the patient at home, thus improving pneumonia diagnosis.

In order to address these limitations of inaccuracy, inconsistency, and delayed diagnosis, a different diagnostic method is required. Computed Tomography (CT) scans use x-rays that pass through the specimen and are received by sensors on the other end. Denser portions of the specimen result in a reduced amount of radiation received by the other end, since the specimen hinders the radiation. This disparity in densities, or attenuation, can be reconstructed to produce a 3D image with different grayscale values (Figure 1). Hounsfeld Units (HU) are grayscale values that correspond to the density of each voxel. In the Hounsfield scale, -1000 represents air, 0 represents water, and 1000 represents bone density. Notably, fluid or pus in the alveolar sacs would be approximately 0 HU, normal lung alveoli have a mixture of air and tissue reading near -500 HU, and voxels in lung with a mixture of air and fluid would be between -500 and 0 HU. CT scans, which can visualize the entire lung as opposed to the 2D projection scans in a chest radiograph, might have the sensitivity to assess the severity of pneumonia as early as 24 hours after onset. This earlier timeframe for treatment would allow immunocompromised patients to receive immediate treatment, thus decreasing their mortality rate. Since CT scans provide a more detailed depiction of the lung, they are potentially more accurate than the current chest x-ray method. Finally, by developing a semi-automated method that uses CT scans to diagnose the severity of pneumonia, more precise diagnoses can be conducted, since the procedure is more automated and less prone to human error (Muller, 2006).

The purpose of this particular research was to develop an algorithm to measure the severity of pneumonia in mice through Micro-Computer Tomography (MicroCT) Scan Analysis and test its effectiveness through comparison with radiologists’ diagnosis. MicroCT works in the same way as a regular CT scanner, but is typically used to image smaller specimens, such as rodents, as opposed to human beings. There were three goals for the image analysis algorithm. The first was to achieve high reproducibility in repeated analysis of the same MicroCT scan. Current methods typically involve having two radiologists independently score the chest x-rays; the final score is then the average of the independent scores. The second goal was to achieve higher accuracy using image segmentation algorithm to quantify the amount of pneumonia in the lungs. This is different from current methods which require radiologists to qualitatively assess multiple images of pneumonia.

Materials and Methods

The project used MicroCT scans to evaluate a murine model of bacterial pneumonia through image analysis by semi-automated segmentation and comparison of results to
radiologists’ interpretation. The following steps were used in the research to prepare mice with pneumonia, and then determine and validate the severity of their pneumonia:

1. Inoculation of mice with different severities of pneumonia bacteria and acquisition of CT scans (L. Hsu, 2007).
2. Radiologists’ diagnosis of pneumonia (L. Hsu, 2007).
3. Pneumonia diagnosis using semi-automated diagnostic algorithm

The first two steps were done prior to this project, where as the third step was conducted specifically in this research, and performed by one individual, with three trials for each mouse. Mercury Computer Systems’ Amira 4.1 software was used to perform image analysis of the CT scans. Amira can perform image segmentation, 3D visualization, and other image processing (Schimel, 2007). Since the severity of pneumonia can be assessed by the volume distribution of materials within the lung due to inflammation, it was expected that more lung volume distribution in the range of -510 to -350, -300 to -250, -250 to -200, -200 to -150, -150 to -100, -100 to -50, and -50 to 0, respectively. These materials were entered into the “Label Voxel” tool, generating these 8 distinct density ranges within the entire CT scan. Figure 5 illustrates this procedure: The materials “Well-Aerated Lung,” “B,” and “C,” encapsulate all voxels inside the entire scan within the density ranges of -510 to -350, -300 to -300 and -300 to -250, respectively (Figure 5). The tool can only analyze three materials at a time, but this had a negligible effect on the quantification analysis because the analysis is almost instantaneous.

We were only interested in the volume of each material within the lung, not the entire scan, which includes bone, muscle, fat, fur, and other tissues. It should be noted that the units of the volume measurements are irrelevant, since we eventually determined the percentage distribution of each material within the lung. Using Amira, we were able to measure how much of each of the eight materials was present in the lung volume. In Table 1, the volume of each of the first three materials can be seen: outside the lung (Exterior), inside the lung (Lung), and in the whole scan (Total) (Table 1).

Essentially, the semi-automated method was broken down into the three following steps. The first was the isolation of the lung from the rest of the CT scan. The second was the removal of extraneous anatomical features, such as the stomach bubble, trachea, and mediastinum, which cannot get affected by pneumonia and would therefore skew our calculations. The final step involved finding the percentage of each of the lung materials (Well-Aerated, B, C, …). The voxel distribution was analyzed for each of the survival groups.

In the development of a semi-automated method, three trials were conducted for each of the 35 mice, and the coefficient of variation (CV) was calculated to evaluate reproducibility. The semi-automated segmentation was reproducible, with the trials for each MicroCT scan resulting in the same segmentation volumes within a coefficient of variation of 2%. Although all three trials were done by the same researcher, it is doubtful that there was much bias. The procedure to do the quantification is very structured and standard. If another researcher performed the experiment, assuming that the researcher followed the exact same procedure, there should be a negligible amount of variation, since all the quantifications are automatically done by the computer, and are therefore human-independent. Grouping the MicroCT segmentation results showed the expected findings for the three experimental groups (Table 2). In the lower ranges of attenuation, Saline-Inoculated mice had the greatest percentage of lung...
Comparisons (1) and (3) were significantly different than the radiologists’ scores, thus indicating that the semi-automated method was indeed more reproducible than the radiologists’ method (Table 4). All the materials showed more significance than the radiologists’ interpretations; however, only H had a p-value under 0.05 for all three group comparisons.

A qualitative assessment of the scans was also made. Figures 6-8 show frontal pictures of three mice, each from the saline, pneumonia survivors, and pneumonia non-survivors groups, respectively (Figures 6, 7, 8). The red color represents voxels in the highest third density range of the lung, the orange color represents voxels in the middle third density range, and the blue color represents voxels in the lowest third density range. As it can be seen, the saline mouse had the greatest amount of orange and least amount of red, the pneumonia survivor group had the least orange and more red, and finally, the pneumonia non-survivor group had the least orange and most red.

Discussion

Preliminary statistical analysis of the semi-automated segmentation of MicroCT detected differences between the groups of pneumonia survivors versus pneumonia deaths. There was good reproducibility of the semi-automated segmentation, with less than 2% variability with repeated application of the methods. Radiologists’ average scores provided statistically significant differences between the mice inoculated with pneumonia vs. the mice inoculated with saline, but not between the groups of pneumonia survivors vs. pneumonia non-survivors, nor between all surviving mice vs. pneumonia deaths. This comparison suggests that the semi-automated segmentation may provide a better method for quantitative scoring of pneumonia severity by CT scans, compared to scoring by radiologists.

Only a few other studies have applied quantitative image analysis to CT scans of pneumonia in animal models (Amigoni, 2008). For example, one examined another type of bacterial pneumonia and used different methods to score the pneumonia. Another examined acid-aspiration pneumonia to score the lung injury. Their studies did not examine severity in terms of survival, and did not compare the image analysis to scoring by radiologists.

Previous applications of quantitative image analysis of lung CT scans have focused on different clinical problems, rarely pneumonia: (1) emphysema, in which lungs have abnormal pockets of low attenuation, (2) lung tumors, which have much higher attenuation than normal lungs, and (3) normal and abnormal physiologic distribution of aeration and blood flow, often using tracer materials to detect flow (Ritman, 2008). Tracers to detect blood flow include xenon, intravenously injected contrast, and microspheres, but since such tracers were not used in this project, the problem of distinguishing normal lung vs. high-attenuation pneumonia involvement was more difficult.

Clinical use of CT imaging is widespread throughout the world because of the wealth of information that CT scans provide about abnormal fluid, tumors, infection, etc. Clinical High Resolution CT (HRCT) is now starting to be used to diagnose pneumonia in patients, but is rarely used for community-acquired pneumonia, which affects more people than any other type of pneumonia (Jelic, 2005). Although the spatial resolution of MicroCT is better than clinical HRCT on an absolute scale (60 microns vs. 500-1000 microns), the spatial detail is better with HRCT when considering the anatomic size of mouse lungs vs. human lungs (apex to diaphragm -5cm vs. -50cm, alveoli 80 microns vs. 210 microns, respectively). Thus, it can be concluded that transferring the technique that was successful in distinguishing normal lung from pneumonia in animal models (Amigoni, 2008) to human disease is now starting to be used in patients.
Comparison Groups | P-value from t-test of percent lung by Semi-Automated Segmentation in material H | P-value from t-test of Radiologists scores
--- | --- | ---
Saline vs. Bacteria | 0.001 | 0.004
Pneumonia Survivors vs. Pneumonia Deaths | 0.039 | 0.581
All Survivor vs. Bacterial Deaths | 0.001 | 0.058

Table 4 Comparison of Micro-CT quantification by Manual H to qualification by radiologists - A comparison of the ability of the semi-automated method to tell groups apart versus radiologist ability. Scoring by the semi-automated method is far more precise than the scoring by the radiologists.

from mouse to humans should make the algorithm more accurate due to increased relative resolution. The current HRCT technique, though an improvement to the more common 2D chest X-ray method, may still provide insufficient detail for quantitative analysis of a 3-dimensional reconstruction because it generally captures only a few 2D slices (Uchiyama, 2003). Newer techniques of human spiral CT will capture enough information for the 3D reconstruction as used in our algorithm.

The principal limitation of this method was the accuracy of manually isolating the lung "volume of interest" from the rest of the mouse. It was sometimes difficult to separate the lung from nearby structures, such as the mediastinum. In the future, indinated or other contrast materials may help more finely define these anatomical structures, until the algorithm is applied to HRCT, which provides finer resolution and, therefore, clearer anatomical structure contrast. This way, it would be easier to isolate the lung during the steps using Amira. However, the extraneous materials have a relatively small percentage in the original lung constructed, so there is still a little room for error in terms of having a slightly poor exclusion of these anatomical features. It was found that on average, even if quantifications were done having all three principal anatomical features still in the volume of interest, there would be a coefficient of variation of less than 3%, which is still well under the statistically significant threshold of 5%. The removal of the anatomical features is for lung infection consistency and improved precision.

Another limitation of this animal study was the small sample size; we intend to extend this technique to another group of mice with pneumonia. Finally, these studies examined only one strain of bacteria and mice, with the inclusion of antibiotics and supportive care. Extending these techniques to other mice will require additional validation, but may help to provide a non-invasive endpoint for studies with experimental pneumonia in transgenic animals. This method of quantitative assessment of pneumonia severity by CT has potential for application in clinical trials in community-acquired pneumonia, as well as other lung diseases.

References


L Hsu, ES, R Summers, D Koziol (2007). Computed Tomography Scans To Assess Severity of Bacterial Pneumonia in a Mouse Model.


