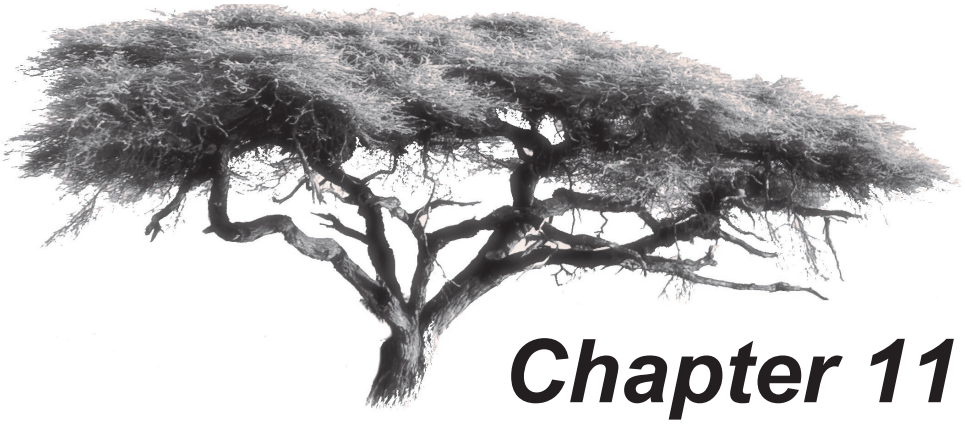

Chapter 11

Summary and general discussion

INTRODUCTION

The introduction of spiral, multidetector-row computed tomography (CT) has expanded the applications of CT by increasing the amount of obtained data within one scanning session and limiting the scanning time for a specific area of interest. For example, scanning of the entire chest within one breath hold of about 10 s can be combined with the reconstruction of thin 1.0 mm slices. However, this development was accompanied by an exponential increase in applied medical radiation dose and thereby possibly a larger hypothetical risk for the patient of developing radiation-induced cancer [1;2]. With present standard techniques, an effective dose between 4 and 8 mSv is delivered during one scanning session (UNSCEAR rapport 2001), which leads to a calculated risk of dying from radiation -induced cancer of 2-4 in 10.000 (30-year old) individuals.
Although the correlation between radiation exposure and radiation-induced cancer is still unclear for relatively low amounts of radiation, one should aim to keep the radiation dose applied during a medical examination to a minimum. For this reason the International Commission on Radiological Protection has introduced the As Low As Reasonably Achievable (ALARA) principle [3]. In the past decade radiologists started to pay attention to the possibilities of dose reduction and individually adjusted protocols have become available. Reduction of radiation dose has been shown to be feasible while maintaining enough image quality to gain the required diagnostic information.

In this thesis, we showed that low dose CT is feasible for several indications, in which the increased image noise has more or less impact on the diagnostic quality and the validity of automated quantifications of the size of abnormal areas, such as pulmonary nodules and emphysema. For quantification of structures which show high contrast with the surrounding tissue, such as a lung nodule within normal lung tissue, the diagnostic information is not limited by the amount of image noise. However, in case of a low intrinsic contrast between the area of interest and the surrounding tissue, for example in case of low extents or diffuse located emphysema, noise has a massive impact on the diagnostic process.

PART 1

In areas with a high intrinsic contrast such as the chest, radiation dose can further be reduced for specific indications. In clinical practice, many benign diseases are being followed-up by imaging. This is mainly performed with plain X-rays. Although CT scans provide more information than plain radiography, plain radiography is often the first choice due to the high radiation dose and the high costs of CT scans. However, in about 10% of the patients, who require a diagnosis for new symptoms, a CT scan is indicated after plain chest X-ray performed, because CXR did not provide the required diagnostic information. This approach

leads to more radiation dose, more costs and a larger burden for the patient. We showed in a small sample of patients from the outpatient department of pulmonology that ultralow-dose CT is feasible and can provide more information than CXR in two directions, while radiation dose is similar. More patients have to be studied to support these results and to detect particular subgroups of patients for which the combination CXR and standard chest CT could be replaced by an ultralow-dose chest CT scan.
The main disadvantage of radiation dose reduction is the accompanying increase in image noise. The detection of abnormalities with a high contrast to the surrounding normal tissue is not limited by the amount of image noise on the CT scans performed with the current minimum radiation dose [4]. However, when the structure of interest shows a low contrast to the surrounding normal tissue, image noise can hamper the detection and especially the automated size measurement of the abnormality. We demonstrated the effect of massive dose reduction on the results of the automated quantification of volume with an attenuation below a fixed threshold. This effect is mainly seen for low extents of lung destruction with a low number of voxels below the chosen attenuation threshold and a steep histogram at the level of the chosen threshold. We showed that an increase in image noise results in overestimation of the emphysema score compared to standard dose CT, but also that the application of a dedicated noise reduction filter to the reconstructed data before the automated quantification of emphysema can prevent this overestimations. We conclude that automated emphysema quantification on low dose CTs is feasible even for early stage emphysema, but only when a dedicated noise reduction filter is applied. This chapter forms the reasoning for the standard application of a noise reduction filter in chapters 7-10. Our standard dose chest CTs also show a certain amount of image noise. Since we did not have a gold standard available like histopathological specimens or emphysema scores at maximal dose, we were not able to judge the accuracy of emphysema scoring on our standard dose scans.

PART 2

The development of low dose protocols has made screening for cancer with chest CT feasible. In a screened population of current and former heavy smokers, the risk of the applied radiation dose is largely exceeded by the risk of dying from lung cancer. Detection of lung cancer in an early, curable stage could reduce lung cancer-related mortality in this population, but the benefit of screening for lung cancer has not yet been proven. The Dutch-Belgian Lung Cancer Screening Trial (NELSON) will be the first randomized controlled trial to show the benefits and limitations of screening for lung cancer in a high-risk population of current and former heavy smokers, using a highly sensitive imaging modality, i.e. low-dose , and no intervention in the control arm.
Lung cancer is today the most frequent cause of cancer death in the Western

world. On a global basis it is estimated that 1.2 million people are diagnosed with this disease every year (12.3% of the total number of cancer diagnosed), and about 1.1 million people are dying of this disease yearly (17.8% of the total cancer death) [5]. More than two-third of these people are diagnosed with locally advanced or metastatic disease, and their poor prognosis is due to a late diagnosis and lack of effective treatment of metastatic disease. Less than 15% of the patients are surviving after 5 years, and in several European countries the 5-year survival is far less.

Since the first publications of the use of the spiral CT in the early diagnosis of lung cancer, and especially the report from the Early Lung Cancer Action Program (ELCAP) in 1999, the interest in lung cancer screening by low-dose spiral CT has increased tremendously[6]. Henschke *et al* recently showed that annual screening with low dose CT can detect lung cancer in a curable stage [7].

Low dose CT has been shown to be extremely sensitive to detect early lung cancer[8;9], but the technique is not very specific[6]. About 70% of the screened subjects showed at least one non-calcified nodule at baseline-screening, which has the potential to be malignant. The vast majority of these nodules will turn out to be benign as proven by other trials [6;10;11]. We reported the way these nodules are managed within the NELSON-trial.

Growth is the most used feature to distinguish the few malignant nodules from the benign ones, but to detect growth and to calculate reliable volume doubling times, precise size measurements are of utmost importance. Revel et al showed that volumetric measurements are more accurate than two-dimensional size measurements [12], but in agreement with Wormanns et al [13], we showed that for Lungcare©, a commercially available and widely used semi-automated volume measurement program, the interobserver variability was low, but the interscan variability could be substantial. The reported limits of agreement have to be taken into account when interpreting an increase in measured nodule size on a repeat scan. We showed that the performance of the algorithm to segment the nodule completely was the most important factor contributing to the variability, using an algorithm that started from a spherical shape of a nodule. Since many nodules detected in a lung cancer screening setting are not spherical, the segmentation step often failed to include the whole nodule. We reported that an increase in volume on a repeat scan of more than 25% for non-spherical nodules and 15% for spherical nodules can with 95% likelihood be applied to a real increase in volume.

More sophisticated nodule segmentation algorithms have already been developed and they are now being released. Our study should be repeated using these new algorithms to set a new reference for measurement error. Although we did not compare the volume measurements on low dose CTs to the results on standard dose CTs, Karabulut et al already showed similar nodule size for both protocols [14].

PART 3

Since chronic obstructive pulmonary disease (COPD) and lung cancer share smoking as main risk factor, lung cancer screening trials provide a good opportunity to study the early stages and natural progression of COPD. Ezatti *et al* reported in 2003 that in the Western countries the number of people that each year dies from COPD is larger than the number of people that dies from lung cancer [15]. COPD is predicted to become the third cause of death from in developed countries in 2020 [16]. COPD is defined as an airflow limitation that is not fully reversible. However, airflow limitation is a relative late symptom of a cascade of smoking related changes caused by the exposure to noxious agents. This cascade comprises lung inflammation, lung tissue destruction, impairment of defense mechanisms that serve to limit the destruction and disruption of the repair mechanisms that may be able to restore the tissue structure in the face of some injuries (www.goldcopd.com). The results of tissue destruction are mucus hypersecretion, airway narrowing and fibrosis, destruction of the lung parenchyma (emphysema) and vascular changes.

Emphysema can easily be detected and automated be quantified on CT by highlighting voxels with an abnormally low X-ray attenuation. Several groups have shown that the results of this method demonstrate a good correlation with the gold standard, the extent of lung destruction in histopathological samples [17-20]. The method is based on the relative increase of air within a voxel due to tissue destruction. Since air has a lower attenuation than tissue, the destruction of lung tissue results in a lowered attenuation within the voxels. The number of voxels with such a lowered X-ray attenuation can be quantified and expressed as percentage of the total number of voxels within the lungs (0-100%). We have called the resulting parameter the emphysema score, but this parameter was first described for 2D datasets and known as the pixel index. Nowadays, many groups use 3D datasets and present the results as the voxel index. However, this method also suffers from some variables influencing the results other than the extent of lung destruction. The impact of slice thickness has been described by Kemerink *et al* [21]. The smaller the size of a voxel, the more different structure will end up in different voxels, resulting in more voxels at the end of the spectrum. Thinner slices result in smaller voxels and in more voxels at the lower end of the frequency histogram of CT-numbers, resulting in higher emphysema scores at a fixed attenuation threshold compared to the conventional CT reconstructed at 10mm slices.

Another variable influencing the emphysema score is the level of inspiration. Lamers *et al* [22] and Kalender *et al* [23] demonstrated the relation between the level of inspiration and the resulting emphysema score. The amount of air within the lung will increase at a higher level of inspiration and this effect results in a lower mean CT-number. We investigated the reproducibility of the quantification of emphysema scores and showed that the level of inspiration was one of the factors

influencing the reproducibility. However, the correlation between the difference in emphysema score and the difference in inspiration level was weak.
To assess the progression of the extent of moderate and severe emphysema, again knowledge of the extent of measurement error between scans is required to distinguish between progression of emphysema and measurement error. We provided the limits of agreement for three common attenuation thresholds for emphysema scores obtained in a lung cancer screening setting. The reported coefficients of variations were high, especially when compared to the coefficients of variation reported by Revel *et al* [24] for volume measurements of non-calcified nodules detected on low dose chest CT. We hypothesize that the main explanation for this difference in coefficient in variation is the difference in contrast of the abnormality of interest to the surrounding normal tissue as mentioned above. Although we used a dedicated noise reduction filter, the automated quantification technique remains sensitive to noise or small changes in CT-numbers. We scanned a phantom before and after a day of scanning NELSON participants to be able to detect and correct for changes in scanner calibration. We detected small variations in the CT-numbers in a homogeneous region of interest measured on the two scans performed on one day, but also during the NELSON-trial and when performing 5 consecutive scans. Since these variations were within the limits provided by the vendor (Philips) and independent of the time of scanning, we were not able to correct the scans performed as part of the NELSON-trial for these small fluctuations. However, these fluctuations can contribute to the coefficients of variations and the broad limits of agreement. Sophisticated methods to detect small shifts in the frequency histogram of CT-numbers may be able to correct for the level of inspiration, but also for small fluctuations in CT-numbers due to variation caused by the CT scanner.
Finally, the anatomical information about the extent and location of emphysema has been compared to the functional information provided by pulmonary function tests. Currently, pulmonologists treat only patients with pulmonary function impairment, as recommended by the Global initiative for chronic Obstructive Lung Disease (GOLD). Staging of COPD is mainly based on the percentage of predicted forced expiratory volume in one second (FEV_1), a marker of airflow limitation and to a lesser extent to the forced vital capacity (FVC). The GOLD-guidelines recommend management to:

- prevent disease progression
- relieve symptoms
- improve exercise tolerance
- improve health status
- prevent and treat complications
- prevent and treat exacerbations
- reduce mortality

The staging according to the GOLD-guidelines also describes a population at risk to develop COPD, this is the population suffering from chronic cough and sputum production, but with a normal lung function. However, not all subjects at risk will develop COPD. In this stage of disease lung function is normal, but other features of smoking related disease may already be present. One of these features is lung destruction. Detection of early lung destruction before COPD becomes symptomatic may prevent the patients from progression to the symptomatic stage. The population enrolled in the NELSON-trial is the population at risk to develop COPD, but without severe COPD, since they have to be fit enough to undergo surgery.

Although COPD is predicted to be the third cause of morbidity and mortality in Western countries in 2020 [16], little is known about the early stages of disease. A lot of research is being performed on development of airway inflammation, but little is known about the progression in disease from "at risk" to airflow limitation. Moreover, little is known about the subjects who will develop COPD and the subjects who will not. Emphysema scoring for each performed chest CT could detect emphysema in an early stage. We showed that many subjects have destructed parenchyma, but a preserved lung function. However, the gas exchange is often already impaired in these subjects, but gas exchange is not one of the hallmarks of COPD (yet). The diffusion of D_{NO} is shown to be a marker of vascular changes, which can precede macroscopic lung tissue destruction and can be impaired before CT can detect emphysema. We showed that D_{NO} is a good sensitive marker to detect emphysema.

The main limitation of studying early emphysema on low dose scans, is the lack of information about the accuracy of the emphysema scores to represent real lung destruction. In a screening study, scanning healthy participants, typically no histological tissue is available. For this reason, we could not judge the results of our emphysema scoring program. Müller *et al* [19] and Gevenois *et al* [17;18] have shown that the extent of emphysema quantified on CT shows a good correlation with the amount of lung destruction detected by macroscopic and microscopic techniques, but we were not able to support these data for low dose scans. A study comparing the emphysema scores obtained from low dose scans performed in lung cancer screening trials to histopathological samples, for example from patients undergoing lobectomy for a malignant nodule, can provide information about the accuracy of emphysema scores to represent real lung destruction.

In the final chapter, we showed that the distribution pattern of emphysema has an impact on the extent of airflow impairment. Participants of the NELSON-project with an apical predominance of emphysema showed more severe airflow limitation than subjects with an equal emphysema score, but a basal predominance of lung destruction. Longitudinal studies investigating the progression of emphysema according to the distribution of emphysema can answer the question if subjects with an apical distribution show more progression

of disease or if subjects showing progression of emphysema will demonstrate a shift to an apical distribution. When subjects with an apical distribution show more progression of disease, these subjects can comprise a subgroup for which a more aggressive risk-modifying approach is required. Only with this information one can predict the clinical relevance and course of low amounts of emphysema in asymptomatic subjects.

Overall Conclusion and Future Perspectives

In this thesis, we have shown that low dose chest CT is feasible and cannot only be applied for lung cancer screening, but also in daily routine and for early detection of lung destruction. We demonstrated that the increased image noise is a limiting factor for the detection of early emphysema, showing low contrast to the surrounding normal tissue, but also that this limitation can be overcome by the use of a dedicated noise reduction filter.

The detection of growth is the main feature to distinguish malignant non-calcified nodules from potentially benign ones, but inter-scan and to a lesser extent inter-observer variability can limit the detection of growth. However, when a nodule shows more than 25% increase in volume, a criterion which will be fulfilled by most malignant nodules on a short-term follow-up CT after three months, it has a likelihood of 95% to be really grown. New segmentation algorithms considering potentially malignant pulmonary nodules as non-spherical may reduce the inter-scan variability, decreasing the upper threshold of agreement. With these algorithms the growth rate can more precisely be calculated and real growth can more easily be distinguished from measurement variation.

This thesis will be the start of further research dealing with the detection and quantification of low-attenuation areas on low dose chest CTs performed in asymptomatic, but high-risk subjects within the scope of the NELSON-trial. We described the baseline emphysema scores in a screening population, but the results of the follow-up study will be more interesting. This data will provide insight in the natural course of emphysema and the impact of the extent of mild and severe emphysema at baseline, the distribution of emphysema at baseline and the impact of quitting smoking before the start of the trial, quitting smoking during the trial and continuous smoking on the progression of emphysema.

REFERENCE LIST

(1) de Gonzalez ABDS. Risk of cancer from diagnostic X-rays: estimates for the UK and 14 other countries. Lancet 2004; 363(9406):345-351.
(2) Brenner DJ, Doll R, Goodhead DT et al. Cancer risks attributable to low doses of ionizing radiation: assessing what we really know. Proc Natl Acad Sci U S A 2003; 100 (24):13761-13766.
(3) Prasad KN, Cole WC, Haase GM. Radiation protection in humans: extending the concept of as low as reasonably achievable (ALARA) from dose to biological damage. Br J Radiol 2004; 77(914):97-99.
(4) Itoh S, Ikeda M, Arahata S et al. Lung cancer screening: minimum tube current required for helical CT. Radiology 2000; 215(1):175-183.
(5) Landis SH, Murray T, Bolden S et al. Cancer statistics, 1999. CA Cancer J Clin 1999; 49(1):8-31.
(6) Henschke CI, McCauley DI, Yankelevitz DF et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. Lancet 1999; 354(9173):99-105.
(7) The International Early Lung Cancer Action Program. Survival of Patients with Stage I Lung Cancer Detected on CT Screening. N Engl J Med 2006; 355(17):1763-1771.
(8) Kaneko M, Eguchi K, Ohmatsu H et al. Peripheral lung cancer: screening and detection with low-dose spiral CT versus radiography. Radiology 1996; 201(3):798-802.
(9) Henschke CI, Naidich DP, Yankelevitz DF et al. Early lung cancer action project: initial findings on repeat screenings. Cancer 2001; 92(1):153-159.
(10) Sone S, Takashima S, Li F et al. Mass screening for lung cancer with mobile spiral computed tomography scanner. Lancet 1998; 351(9111):1242-1245.
(11) Swensen SJ, Jett JR, Hartman TE et al. Lung cancer screening with CT: Mayo Clinic experience. Radiology 2003; 226(3):756-761.
(12) Revel MP, Bissery A, Bienvenu M et al. Are Two-dimensional CT Measurements of Small Noncalcified Pulmonary Nodules Reliable? Radiology 2004; 231(2):453-458.
(13) Wormanns D, Kohl G, Klotz E et al. Volumetric measurements of pulmonary nodules at multi-row detector CT: in vivo reproducibility. Eur Radiol 2004; 14(1):86-92.
(14) Karabulut N, Toru M, Gelebek V et al. Comparison of low-dose and standard-dose helical CT in the evaluation of pulmonary nodules. Eur Radiol 2002; 12(11):2764-2769.
(15) Ezzati M, Lopez AD. Estimates of global mortality attributable to smoking in 2000. The Lancet 2003; 362(9387):847-852.
(16) Feenstra TL, van Genugten ML, Hoogenveen RT et al. The impact of aging and smoking on the future burden of chronic obstructive pulmonary disease: a model analysis in the Netherlands. Am J Respir Crit Care Med 2001; 164(4):590-596.
(17) Gevenois PA, De Vuyst P, De M, V et al. Comparison of computed density and microscopic morphometry in pulmonary emphysema. Am J Respir Crit Care Med 1996; 154(1):187-192.
(18) Gevenois PA, De M, V, De VP et al. Comparison of computed density and macroscopic morphometry in pulmonary emphysema. Am J Respir Crit Care Med 1995; 152(2):653-657.
(19) Muller NL, Staples CA, Miller RR et al. "Density mask". An objective method to quantitate emphysema using computed tomography. Chest 1988; 94(4):782-787.
(20) Madani A, Zanen J, de Maertelaer V et al. Pulmonary Emphysema: Objective Quantification at Multi-Detector Row CT--Comparison with Macroscopic and Microscopic Morphometry. Radiology 2006;2382042196.
(21) Kemerink GJ, Kruize HH, Lamers RJ et al. CT lung densitometry: dependence of CT number histograms on sample volume and consequences for scan protocol comparability. J Comput Assist Tomogr 1997; 21(6):948-954.
(22) Lamers RJ, Thelissen GR, Kessels AG et al. Chronic obstructive pulmonary disease: evaluation with spirometrically controlled CT lung densitometry. Radiology 1994; 193 (1):109-113.
(23) Kalender WA, Rienmuller R, Seissler W et al. Measurement of pulmonary parenchymal attenuation: use of spirometric gating with quantitative CT. Radiology 1990; 175(1):265-268.

(24) Revel MP, Lefort C, Bissery A et al. Pulmonary nodules: preliminary experience with three-dimensional evaluation. Radiology 2004; 231(2):459-466.