Visual assessment of perfusion and metabolism in neurodegenerative dementia

DAVID FÄLLMAR
Abstract

A worldwide demographic shift is currently occurring, with rapidly increasing numbers of elderly individuals. Since the incidence of neurodegenerative disease generally increases with age, this entails an increase in dementia prevalence. There are several strong incentives for establishing robust and widely available imaging methods for the early diagnosis of these diseases. Atrophy patterns are evident only late in the disease process, and the distinction from healthy ageing can often be elusive. For early diagnosis, physiologic parameters such as perfusion or metabolism must be assessed. The available modalities all have restricted clinical usefulness. The main aim of this thesis was to advance the clinical usefulness of perfusion and metabolism imaging in patients with neurodegenerative dementia, with a focus on visual assessment.

A cohort of patients with neurodegenerative dementia was included, along with an age-matched control group. All subjects underwent MRI, including a pseudocontinuous ASL sequence and FDG-PET. In papers II and III, a subgroup containing both patients and controls underwent a second FDG-PET with reduced dose. In paper IV, the material was combined with a similar cohort from Amsterdam.

Paper I showed that spatial smoothing increased the correlation between visually assessed perfusion and metabolism levels as displayed with FDG-PET. However, the distinction between patients and healthy controls was less satisfactory due to false positives.

Paper II showed that differences in regional standard uptake value ratios between normal- and low-dose FDG-PET were small and without clinically significant bias.

Paper III showed that the diagnostic performance of Z-score maps showing regions of significant deficits in metabolism was highly similar in normal- and low-dose FDG-PET images.

Paper IV showed that ASL perfusion-based Z-score maps can be used for diagnostic purposes with high specificity, but inferior sensitivity, compared to FDG-PET.

In conclusion, the included studies address aspects of the visual assessment of perfusion and metabolism neuroimaging, with a focus on clinical usefulness in diagnosing neurodegenerative dementia.

Keywords: Dementia, Neurodegenerative, Brain imaging, Diagnostics, Neuroradiology, Neuroimaging

David Fällmar, Department of Surgical Sciences, Radiology, Akademiska sjukhuset, Uppsala University, SE-75185 Uppsala, Sweden.

© David Fällmar 2016

ISSN 1651-6206
urn:nbn:se:uu:diva-304731 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-304731)
“It is always fun to look at brains”
-Danielle van Westen

To my family.
Design and layout: Håkan Pettersson
Cover artwork: Original drawing by Albert Engström, montage by David Fällmar and Håkan Pettersson
List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


Reprints were made with permission from the respective publishers.
## Contents

Introduction ........................................................................................................ 11
Dementia ........................................................................................................... 11
  Diagnostics .................................................................................................. 11
  Causes of dementia ....................................................................................... 12
Neuroimaging .................................................................................................... 14
  Registration and spatial normalization ......................................................... 14
  Smoothing ................................................................................................... 15
  Morphology ................................................................................................ 15
  Perfusion .................................................................................................... 16
  Metabolism ................................................................................................. 16
  Two sides of the same coin? ....................................................................... 17
Quantification ................................................................................................... 17
  Morphology ................................................................................................ 17
  Perfusion .................................................................................................... 18
  Metabolism ................................................................................................. 19
Visual assessment .............................................................................................. 19
  Morphology ................................................................................................ 19
  Perfusion .................................................................................................... 20
  Metabolism ................................................................................................. 20
Aims ................................................................................................................... 23
  General aim ................................................................................................ 23
  Specific aims ............................................................................................... 23
    Paper I ...................................................................................................... 23
    Paper II .................................................................................................... 23
    Paper III ................................................................................................ 24
    Paper IV ................................................................................................. 24
Methods .............................................................................................................. 25
  Subjects ....................................................................................................... 25
  MRI protocol ............................................................................................... 26
    MRI post-processing .................................................................................. 26
  PET protocol ............................................................................................... 28
    PET post-processing .................................................................................. 28
    Visual assessments ................................................................................... 31
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>18F-FDG</td>
<td>Fluorodeoxyglucose marked with 18F</td>
</tr>
<tr>
<td>AAL</td>
<td>Automated anatomic labeling</td>
</tr>
<tr>
<td>AD</td>
<td>Alzheimer’s disease</td>
</tr>
<tr>
<td>ADD</td>
<td>Alzheimer’s disease dementia</td>
</tr>
<tr>
<td>ASL</td>
<td>Arterial spin labeling</td>
</tr>
<tr>
<td>bvFTD</td>
<td>Behavioural variant (of) frontotemporal dementia</td>
</tr>
<tr>
<td>CBD</td>
<td>Corticobasal degeneration</td>
</tr>
<tr>
<td>CBF</td>
<td>Cerebral blood flow</td>
</tr>
<tr>
<td>CT</td>
<td>Computed tomography</td>
</tr>
<tr>
<td>DLB</td>
<td>see LBD</td>
</tr>
<tr>
<td>FDG</td>
<td>Fluorodeoxyglucose</td>
</tr>
<tr>
<td>FDG PET</td>
<td>Fluorodeoxyglucose positron emission tomography</td>
</tr>
<tr>
<td>FLAIR</td>
<td>Fluid attenuation inversion recovery</td>
</tr>
<tr>
<td>FTD</td>
<td>Frontotemporal dementia</td>
</tr>
<tr>
<td>FTLD</td>
<td>Frontotemporal lobe degeneration</td>
</tr>
<tr>
<td>LBD</td>
<td>Lewy body dementia</td>
</tr>
<tr>
<td>MCI</td>
<td>Mild cognitive impairment</td>
</tr>
<tr>
<td>MMSE</td>
<td>Mini-mental state examination</td>
</tr>
<tr>
<td>MPR</td>
<td>Multi-planar reconstruction</td>
</tr>
<tr>
<td>MRglu</td>
<td>Metabolic rate of glucose consumption</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic resonance imaging</td>
</tr>
<tr>
<td>MTA</td>
<td>Medial temporal lobe atrophy</td>
</tr>
<tr>
<td>pCASL</td>
<td>Pseudocontinuous arterial spin labeling</td>
</tr>
<tr>
<td>PET</td>
<td>Positron emission tomography</td>
</tr>
<tr>
<td>ROI</td>
<td>Region of interest</td>
</tr>
<tr>
<td>SD</td>
<td>Semantic dementia</td>
</tr>
<tr>
<td>SPECT</td>
<td>Single-photon emission computed tomography</td>
</tr>
<tr>
<td>SSP</td>
<td>Stereotactical surface projection</td>
</tr>
<tr>
<td>SUV</td>
<td>Standardized uptake value</td>
</tr>
<tr>
<td>VaD</td>
<td>Vascular dementia</td>
</tr>
<tr>
<td>VOI</td>
<td>Volume of interest</td>
</tr>
</tbody>
</table>
Introduction

Dementia

Dementia is derived from the Latin de- for “away”, and mens, meaning, “mind”. Dementia is usually defined as a chronic, acquired syndrome involving multiple cognitive impairments that are sufficient to interfere with activities of daily living [1]. Memory impairment is a well-recognized feature of dementia, but other cognitive domains such as executive function, language, motor planning, and visuospatial perception are often involved as well, which can affect the functions of daily living to a large extent. Involvement of the frontal lobes can infer changes in behaviour and personality, a great burden to those near and dear to the afflicted.

Dementia is not a disease but a condition that can be secondary to a number of different pathogenic factors, mainly vascular or neurodegenerative diseases, or a combination of causes.

A worldwide demographic shift is currently occurring, with rapidly increasing numbers of elderly individuals, sometimes dubiously referred to as the gloomy “grey ing of the world” [2], or the clairobscure “silver tsunami”. Since the incidence of neurodegenerative disease generally increases with age, this entails an increase in dementia prevalence. In the year 2015, there were an estimated 46.8 million people with dementia worldwide – which is projected to rise to 131.5 million in 2050 [3].

The direct and indirect costs of caring for these individuals are enormous [2]. If dementia care was a country, it would be the 18th largest in terms of gross domestic product; the annual cost is estimated to be 7000 billion Swedish Crowns, exceeding the market values of companies such as Apple or Google [3].

Diagnostics

With the current diagnostic definitions, the definitive diagnosis is based on histopathology if a brain biopsy is done or an autopsy is carried out (which is seldom). In practice, the patient is given a clinical probability diagnosis in vivo, often by a general practitioner, geriatrician, or neurologist. The clinical probability diagnosis is based on a combination of a structured interview, cognitive screening tests such as the Mini-Mental State Examination
(MMSE) [4], a neurological examination, and morphological neuroimaging. In some cases, especially in younger patients, the work-up is expanded through analysis of cerebrospinal fluid biomarkers, functional neuroimaging, and neuropsychological tests. In the early stages of a neurodegenerative disease, repeated assessments are required since the early symptoms are often nonspecific. Early symptoms that do not interfere with the ability to function at work or during activities of daily life are not sufficient for fulfilling dementia criteria [1, 5], and in those cases, the heterogeneous term mild cognitive impairment (MCI) is used. Patients with MCI can later convert to dementia, remain stable, or improve.

Currently, there is no curative pharmaceutical agent for the treatment of any of the neurodegenerative dementia disorders, and relatively limited options for symptom alleviation. Still, diagnoses are important for several reasons. For the patient and his/her relatives, an understanding of the symptoms is helpful in order to deal with the progressive deterioration and behavioural changes. A reasonable prognosis enables planning of housing requirements and what level of care will be needed. In some cases, the question of heredity needs to be discussed.

For a long time now, the advent of disease-modifying drugs has been close, at least for Alzheimer’s disease (AD). Many substances have been tried on a limited scale with various results, but so far, none have made it all the way to clinical use [1, 6]. This research field is very active, and it is likely that some means of slowing down or stopping the progress of the disease will eventually be available. The damage already done, however, is less likely to be cured, which gives us a strong incentive for establishing robust and widely available methods for early diagnosis.

Causes of dementia

Alzheimer’s disease (AD) is the most common cause of dementia in all age groups, and the incidence rises significantly with age – afflicting approximately 2.5% of 65-year olds, 20% of 85-year olds, and 40% of 90-year olds (in Western Europe and the USA) [3]. There are familial forms that have received much scientific attention, but the vast majority of cases are sporadic [7]. The disease was originally described based on the histopathological findings of cortical extracellular “senile plaques” and intracellular “neurofibrillary tangles” [8], which much later were found to be composed of beta-amyloid protein and hyperphosphorylated tau protein, respectively. The pathological accumulation of these proteins precedes symptoms by many years. Due to progressive cognitive deterioration, a patient with AD will eventually fulfil dementia criteria, and the term “Alzheimer’s disease de-
mentia” (ADD) can be used to specify this. An alternative term would be “dementia due to Alzheimer’s disease”.

The most prominent symptoms in each patient vary, especially in the early stages, but the classic features are impairment of short-term episodic memory, abstract reasoning, language, and visuospatial functions. This corresponds to a typical degeneration pattern involving the temporal and parietal lobes.

The presence of low levels of beta amyloid and high levels of tau protein in cerebrospinal fluid is a strong supportive feature of AD [1]. A new diagnostic framework (IWG-criteria) has been proposed to allow for diagnosis of AD in vivo, based on a typical clinical phenotype and the presence of positive biomarkers [9]. Specific diagnosis criteria that allow for an early diagnosis enables earlier intervention and facilitates research studies of AD in preclinical states [10].

Vascular dementia (VaD) is chronic cognitive dysfunction secondary to cerebrovascular disease. This entity is not within the scope of this thesis, but important in clinical diagnostics. The prevalence of cerebrovascular disease increases significantly with age, and in the elderly, VaD is the second most common cause of dementia after AD. Many patients have a combination of vascular and neurodegenerative dementia, so called mixed dementia, which sometimes clouds the pattern for the diagnosing physicians.

Frontotemporal dementia (FTD), sometimes described as frontotemporal lobar degeneration (FTLD), is a group of clinical neurodegenerative syndromes with different underlying pathophysiological properties [7]. Three subgroups of FTD have been described: a behavioural variant (bvFTD), progressive non-fluent aphasia, and semantic dementia (SD), each with distinct symptoms and patterns of neurodegeneration (overlap is common) [11]. The older term Pick’s disease was previously used synonymously with FTLD, but is now reserved for a subtype with the specific histopathological feature of Pick bodies [12]. It is possible that future classification will consist of distinct subgroups based on pathogenesis.

The main symptoms in FTD are a combination of frontal symptoms such as apathy, loss of empathy, or disinhibition, with altered speech output such as decrease in the amount of speech or perseverations [12]. SD can have prominent features of loss of general knowledge and word-finding difficulties, sometimes with distinct involvement of the most anterior part of the (left) temporal lobe.

Another classification of the three conditions mentioned above is bvFTD versus primary progressive aphasia, where the latter is divided into a non-fluent variant and a semantic variant [12].
Corticobasal degeneration (CBD) is a rare disease that is also associated with the FTD/FTLD syndrome. CBD is associated with asymmetrical involvement of the frontoparietal cortex and substantia nigra, causing asymmetrical apraxia and extrapyramidal motor symptoms.

Lewy body dementia (LBD, or DLB for “dementia with Lewy bodies”) has clinical features that can be similar to AD, but pathophysiological features in common with Parkinson’s disease. If dementia precedes motor symptoms or if they develop during the same year, the term LBD is used, but if motor symptoms precede dementia by more than one year, the term Parkinson’s disease with dementia is used. Parkinson’s disease with dementia and LBD can be indistinguishable in neuroimaging in late stages. LBD can infer deficits in memory and visuospatial ability similar to AD, but is more often fluctuant. Other characteristics that can distinguish LBD from AD are recurrent visual hallucinations, extrapyramidal symptoms, and sleep-behaviour disorders [13].

**Neuroimaging**

Neuroimaging, often in the form of computed tomography (CT), has traditionally been a part of the diagnostic work-up of suspected dementia, with the main purpose of ruling out surgically treatable differential diagnoses, such as tumours and subdural haematomas. With improving image quality and an increase of available imaging modalities, the focus has changed to the intention of demonstrating positive disease markers [7]. MRI and functional neuroimaging are superior to CT in this endeavour.

**Registration and spatial normalization**

The term registration refers to the process of bringing one (source) imaging volume into anatomical correspondence with another (target) [7]. This is commonly performed when handling multimodal or repeated imaging clinically and during scientific comparisons. Registration can be performed with various degrees of geometrical freedom, which means various degrees of manipulation of the imaging data, with the intention of maximal correlation. A registration without manipulation of geometry is called rigid.

A registration using shifting and rotation in three dimensions and zooming and shearing in three dimensions has 12 degrees of freedom and is referred to as an “affine registration”. When more degrees of freedom are allowed, using polynomial models, the process is referred to as “deformation based analysis” [7].
When the target volume of a registration is a standardized anatomical space or coordinate system, the process is referred to as “spatial normalization”. This enables a large array of tools such as automated anatomical labeling and comparisons at group level.

**Smoothing**

Spatial smoothing is a post-processing tool that averages each pixel with adjacent pixels. Traditional methods have been 5-point smoothing for 2D images and 9-point smoothing 3D, in which all neighbouring pixels are given equal weight. Gaussian smoothing is a more sophisticated method in which surrounding pixels are weighted according to distance from the pixel being smoothed [14]. The weighting function is described by its “full width at half maximum” (FWHM), where a larger FWHM means more extensive smoothing.

Spatial smoothing increases signal-to-noise ratio by averaging out noise at the expense of spatial resolution. A smoothed image is often more appealing to the eye, with round contours and smudged noise, but also results in blurring and potential loss of image detail [14].

Smoothing of images has been a necessary and obvious post-processing step in nuclear medicine because of noisy images. Radiology has traditionally depended on spatial resolution and image detail, and smoothing has not been beneficial. With an increasing amount of functional imaging, especially in MRI, radiologists will need to consider how to use smoothing in the clinical setting.

**Morphology**

Assessing patterns of brain atrophy on morphological imaging is a necessary diagnostic step in the work-up of patients with suspected neurodegenerative disease. Global atrophy can be an unspecific finding, but focal patterns of cortical atrophy can suggest AD (medial temporal lobes, parietal lobes), FTD (anterior temporal lobes, prefrontal cortex), CBD (asymmetrical frontoparietal), progressive supranuclear palsy (mesencephalon, superior cerebellar peduncles), and other less common conditions [15].

Adding a functional dimension to imaging adds more information and facilitates early detection [16, 17].

White matter changes and signs of cerebrovascular lesions are important for assessing the presence of vascular causes. The sensitivity of this assessment is greater on MRI than CT [18], although modern CT scanners have been
shown to provide reliable information that is comparable with that obtained with MRI in a memory clinic setting [19].

Perfusion
Cortical perfusion correlates with neuronal and synaptic activity. Patterns of focally reduced perfusion can reveal areas afflicted by neurodegenerative disease, antedating detectable atrophy [16, 20].

Historically, regional brain perfusion in human subjects was first assessed with isotope injections in the carotid artery by pioneers such as David Ingvar and Niels Lassen, and later by using inhalation of $^{133}$Xenon [21]. During recent decades, single-photon emission computed tomography (SPECT) using $^{99m}$Tc-HMPAO has been an established method of assessing regional cerebral blood flow.

Arterial spin labeling (ASL) is an MRI sequence that non-invasively measures cortical perfusion. A radio frequency pulse tags ascending blood in the carotid and vertebral arteries, and the arrival of tagged blood in cortical areas (i.e. perfusion) can be measured. After a number of repeated imaging volumes of the entire brain with and without blood tagging, perfusion maps can be calculated [22].

Since the introduction of ASL in the 1990s [23], its potential use in the diagnostic work-up of patients with cognitive dysfunction has begun to be systematically explored [20, 24, 25]. Several types of ASL labeling approaches are available, as reviewed in [26], but pseudocontinuous ASL (pCASL) is recommended for clinical purposes by a major consensus paper [22].

The gold standard assessment for brain perfusion is often considered to be $^{15}$O-water positron emission tomography (PET). Due to practical issues such as short half-life and need for complex analysis, this examination is mainly used for research purposes at expert centres with cyclotron access.

Metabolism
Cortical glucose metabolism correlates with neuronal and synaptic activity. Patterns of focally reduced glucose consumption can reveal areas afflicted by neurodegenerative disease, antedating detectable atrophy [17, 27].

The positive finding of impaired glucose metabolism has become an important diagnostic supportive feature of a dementia work-up [1], and is often a valuable support for the clinician in the differential diagnosis of patients with early and unspecific symptoms.
\(^{18}\)F-fluorodeoxyglucose (FDG) is a glucose analogue with the positron-emitting fluorine-18 substituted for one of the hydroxyl groups. It is distributed similar to glucose in the body and is taken up by cells with high glucose consumption. When intracellular, the molecule will be phosphorylated but not metabolized, and the radioactive isotope will remain in the cell. It will eventually emit the positron, with a half-life of 110 minutes, leading to an annihilation that produces detectable photons. Thus, a PET scanner can trace the presence and location of tissues with high levels of glucose consumption.

FDG-PET is a robust and commonly used method, and it has been shown to have superior sensitivity and specificity in diagnosing dementia compared to SPECT [28]. It has to a large extent replaced SPECT clinically but to some extent is still restricted by availability. Hybrid PET/CT scanners with sequential CT and FDG-PET scanning combine morphological and functional information. PET/MRI scanners with simultaneous acquisition of both modalities are emerging for research and also offer promising prospects for clinical use in the near future.

PET scanners can also be used for the visualization of amyloid deposits in the brain, which is an essential part of the pathophysiology in Alzheimer’s disease. This is an interesting and potentially useful part of neuroimaging in suspected dementia but is not covered in this thesis.

Two sides of the same coin?

There is a good correspondence between the pattern of hypoperfusion detected with SPECT and the pattern of hypometabolism detected with FDG-PET [29]. Several studies describe qualitative and quantitative comparisons of ASL perfusion and FDG-PET in healthy controls and patients with neurodegenerative disease, and a close correlation has been described [30, 31]. Thus, there is a strong covariance between the presence of hypoperfusion and the presence of hypometabolism in neurodegenerative disease [32]. This implies that similar diagnostic information can be retrieved from these two physiological parameters.

Quantification

Morphology

The diameters and volumes of anatomical regions can be measured manually. If the spatial resolution is sufficient, images can be spatially normalized, and automated anatomical regions can be defined, resulting in individual volumetrical measurements [33].
Quantification of morphology can be used to show volume decrease within a subject if repeated measurements are available. The rate of yearly volumetric reduction has been suggested as a way to distinguish healthy ageing from pathology [15]. Volumetry can also be used to compare an individual to a cohort of normal controls, but this has so far been restricted clinically because of cumbersome post-processing and because the interpretation of the volumes in an individual is limited by large intersubject variations. However, due to intensive software development, clinically useful implementations can be imminent [34].

Perfusion

Much effort has been invested in the quantification of ASL data to achieve absolute cerebral blood flow (CBF) values, and considerable progress has been made. Different ASL protocols have been validated against $^{15}$O-water PET, with various and interesting results (e.g. steady state ASL in [35], pulsed ASL in [36], pCASL in [37]). Several studies have shown that pCASL can provide whole-brain CBF results similar to those obtained with $^{15}$O-water PET, e.g. [37, 38]. The test-retest reliability of ASL has been examined in several protocols [37, 39], most of which have shown acceptable or good reliability.

Even so, several questions are currently unanswered. ASL-derived CBF data seem to be hampered by inter-subject differences [40], probably due to a subject-method interaction [36]. Estimation of labeling efficiency on a subject-specific basis would improve the accuracy [41] but is impractical. Artefacts due to variations in transit time are common [42, 43], as is the focal decrease in ASL signal in watershed areas [44]. The relationship between the timing of the subjects’ arterial transit time and the postlabel delay time of the protocol is an important factor. Too short a time delay results in apparent incomplete perfusion of the cortex and spots of high signal intensity representing remaining intravascular signal. Too long a time delay results in loss of signal strength. A few protocols have been tried out using multiple time delays [45], and this might have benefits but is more complex and is not yet suitable for routine purposes.

Various stimuli, such as caffeine intake, demanding cognitive tasks, or opening of the eyes, affect CBF and can therefore be regarded as confounders when the results are interpreted [46, 47].

Absolute quantification also requires measurements or assumptions about tissue magnetization interactions, longitudinal relaxation time, et cetera [25]. These and other methodological problems must be taken in consideration when performing and assessing ASL images, especially when handling quantified data. On-going research with novel quantitative ASL protocols is not covered in this thesis.
From a clinical perspective, region-by-region quantification of CBF values is currently quite complex, and there is no consensus regarding perfusion cut-off values that distinguish healthy ageing from dementia [44, 48]. The use of intensity-normalized perfusion maps showing intrasubject regional deficits has been suggested [49].

**Metabolism**

When an arterial input function (or an approximation) in the form of repeated blood samples during FDG-PET scanning is used, it is possible to calculate the metabolic rate of glucose consumption (MRglu) in each voxel. This is often used in research, but seldom in clinical practice due to its impracticality. A unit called standard uptake value (SUV) is commonly used in PET imaging, but this is less relevant for brain studies. Instead, a SUV ratio is calculated by dividing the SUV in a specific region of interest with a reference region from the same image, such as the pons, whole brain average, or cerebellar cortex.

**Visual assessment**

**Morphology**

An essential challenge in assessing morphological images is to discriminate pathological atrophy from what is expected or “normal for age group”. Healthy ageing includes some degree of parenchymal atrophy, but the variation in the normal population is large [33, 50].

For proper visual assessment of global and focal atrophy, a structured evaluation is strongly recommended. On modern CT and morphological MR images, the use of 3D protocols allows for multiplanar reconstruction. Assessment of medial temporal lobe atrophy is performed on coronal images, usually after angulating images perpendicular to the anterior-posterior commissure line [51].

Many visual rating scales have been described for assessing atrophy, and several are commonly used clinically, such as the four-grade rating of global cortical atrophy [52], the Schelten’s score for medial temporal lobe atrophy [53], and the Koedam score for posterior atrophy [54]. These assessments are the workhorses of clinical dementia neuroimaging, but fall outside the scope of this thesis. Cut-off limits for clinical use of these rating scales have been suggested [55].
Perfusion

Currently, only a few publications describe visual assessment of ASL [48, 56]. The use of ASL in clinical routine is presently under establishment [20]. Sample images of ASL perfusion are shown in Figure 1.

![Sample ASL images from a healthy control. The same slice is depicted as a raw subtraction image to the left, as a motion corrected perfusion map in the middle, and with additional smoothing and dichromatic colour scheme to the right.](Image)

Metabolism

Visual assessment of FDG-PET images for detecting neurodegenerative disease has been an established method for many years, and it is an important supportive feature for diagnostics [1]. Numerous studies have described the capability of FDG-PET to distinguish different neurodegenerative disorders based on the patterns of hypometabolism. Pleasant and useful overviews are available [27, 57].

Accurate visual assessment of FDG-PET images requires a high level of expertise and may be challenging, especially in patients with early stages of neurodegenerative disease. Since its introduction in 1995 [58], the use of automated 3D-stereotactic surface projection maps (3D-SSP maps) has been an increasingly important part of clinical evaluation of these patients. The 3D-SSP method defines a large number of surface points on a standardized brain surface and determines the maximum cortical uptake for each point by sampling the registered image to a pre-defined depth. The maximum value from each sampling point is then mapped to a three-dimensional representation image. The resulting 3D-SSP maps can be thresholded to show cortical areas with statistically significant deficits, when the individual is compared to a database of healthy controls. This highlights areas with metabolism levels at the far left end of a standardized normal distribution curve – in effect values
with a Z-score below -1.96. For this reason, such maps are commonly referred to as Z-score maps. Z-score maps reveal patterns of pathology, which facilitate differential diagnosis. They have been shown to improve the accuracy of diagnosing neurodegenerative disease [17, 59], especially for novice readers [60].

There are several commercially available software packages with an SSP function, e.g. HybridViewer BRASS (Hermes medical solutions), CortexID (GE Healthcare), Neurostat 3D-SSP (University of Michigan, USA), and SPM (Wellcome Dept. London, UK). 3D-SSP is used clinically, both for FDG and amyloid imaging, and method optimization is ongoing [61]. Currently, none of the commercially available software has been adapted for perfusion images.

Sample Z-score images of three of the included patients are provided in Figure 2 a–c.
Figure 2a–c. Sample images of FDG-PET Z-score maps. a) Patient with SD, with significant deficits in the anterior temporal lobes and anterior cingulate cortex. b) Patient with AD, with typical deficits in the temporoparietal, precuneal, and posterior cingulate cortices, but with sparing of the occipital lobes. c) Patient with LBD, with widespread temporoparietal deficits and involvement of the occipital lobes. Images were created using HybridViewer BRASS 2.6A.10, Hermes Medical Solutions.
Aims

General aim
The general aim of the thesis is to advance the clinical usefulness of perfusion and metabolism imaging in patients with suspected neurodegenerative disease. Currently, there is a gap between modern functional neuroimaging research and the clinical use of these modalities. This gap is partly due to the complex and time-consuming post-processing commonly employed in research, and to the difficulties of interpreting quantified data in the work-up of an individual patient. A general aim of this thesis is to explore the role of visual assessment in the clinical use of these powerful modalities.

Specific aims
Paper I
Spatial smoothing is commonly used on ASL data before statistical group analyses, but its application in a visual assessment setting is insufficiently described in the literature. The aims of paper I were to compare visual assessment of ASL based perfusion maps with and without additional smoothing to FDG-PET data and to evaluate the effects of beta estimate filtering.

Paper II
FDG-PET of the brain has been studied extensively during the late twentieth century, mostly on scanners with inferior resolution and sensitivity compared to those currently used. Dosage of FDG remains largely unchanged despite significant improvement in scanner quality and reconstruction algorithms, which theoretically should allow for dose reduction. Evidence that clinically acceptable diagnostic images can be obtained with low FDG dose is needed before dose reductions can be implemented. This is especially important in order to increase the availability of this potentially valuable imaging tool for use in younger patients and research protocols based on serial scans of healthy volunteers. The purpose of paper II was to test whether the dose of FDG could be considerably decreased while maintaining SUV...
ratio characteristics in healthy controls and in patients with neurodegenerative disorders.

**Paper III**

Visual assessment of FDG-PET images requires a high level of experience and can be challenging, especially for novice readers. The clinical usefulness is enhanced by the use of Z-score maps that show regions with significant deficits. The aim of paper III was to evaluate whether Z-score maps based on low-dose FDG-PET images were sufficient for diagnostic purposes. Together with paper II, the aim was to validate the clinical and scientific use of low-dose FDG-PET.

**Paper IV**

The clinical role of ASL-based perfusion maps is not yet established, and their implementation is limited by the absence of cut-off values separating normal from pathological perfusion. While the use of Z-score maps has been established for the analysis of FDG-PET, they remain unexplored in ASL. The aim of this study was therefore to compare ASL-based Z-score maps with FDG-PET-based Z-score maps for visual assessment of neurodegenerative dementia.
Methods

Subjects
A total of 21 patients who had undergone a thorough diagnostic procedure and received a diagnosis of neurodegenerative disease in the stage of mild cognitive impairment or mild dementia were enrolled from the memory clinic at the geriatric department of the Uppsala University Hospital. Most of the included patients had previously been examined with FDG-PET and/or neuropsychological tests as part of their diagnostic work-up.

Eight patients were diagnosed as having a behavioural variant FTD (bvFTD) and two as SD according to the criteria of Neary et al. [62]. Two patients had LBD [13], seven others had AD [5, 63], and one patient had CBD according to clinical features and areas of hypometabolism detected with FDG-PET. The mini-mental state examination (MMSE) [4] was performed upon inclusion, and in most cases, a thorough neuropsychological test battery was also administered. The diagnoses imply that all patients had neurodegenerative diseases with assumed regional hypoperfusion and hypometabolism, potentially identifiable on both FDG-PET and ASL.

An age-matched group of 40 cognitively normal controls was recruited through advertising. A structured interview was performed to exclude subjects with cognitive dysfunction or history of cerebral disease, and a physical examination was done to eliminate subjects with signs of neurological deficits or extrapyramidal symptoms. They did not show clinical signs of depression and were not on drugs with psychoactive or cognitive side effects. Inclusion criteria included MMSE ≥ 26 in combination with normal results on the Trail Making Test A and the “7 minute screen™” [64].

All included subjects underwent MRI and FDG-PET examinations as described below. A subset of 16 controls and 8 patients also performed a second PET examination with low dosage, making them eligible for inclusion in paper II and paper III.

In paper IV, the ASL scans and normal-dose PET scans from controls, AD patients, and FTD patients were combined with a similar cohort recruited in Amsterdam, described in detail by Verfaillie et al. [32].
MRI protocol

The subjects were scanned in a 3T MRI scanner (Achieva; Philips Medical Systems, the Netherlands) using an eight-channel head coil, with a protocol that included a 3D T1-weighted gradient echo sequence (3D turbo field echo), FLAIR, and pseudo-continuous ASL (pCASL). The pCASL scanning parameters were repetition time (TR) = 4100 ms, echo time (TE) = 14 ms, multi-transmit, FOV = 230 x 230 mm, pixel size = 2.75 x 2.75 mm, 30 dynamic scans, 20 slices of 5 mm. The flip angle was 90 degrees, and SPIR fat suppression was used. The labeling duration was 1650 ms, and post-labeling delay was set to 1600 ms. The multi-slice 2D imaging readout leads to an effective delay of 1600 ms for the first slice, which increases to 2300 ms for the last acquired slice. No vascular crushing was used. Background suppression was included using inversion pulses at 1710 and 2860 ms after a presaturation pulse preceding labeling. The total pCASL scan time was 4 min 14 s.

MRI post-processing

Paper I

Two different post-processing methods were used in paper I: simple control-label subtraction, and spatial smoothing in combination with General Linear Modelling (GLM)-based beta estimate functions. For the smoothing/beta estimate post-processing, the images were initially smoothed using a three-dimensional Gaussian filter with a FWHM of 3.5 mm. In an attempt to reduce low frequency noise in the ASL data, a number of low frequency cosine basis functions were added to the GLM design matrix, acting as a high pass filter. Sample images are shown in Figure 3.
Paper II
To allow for automatic quantification of PET data, the T1-weighted MR images from each subject were spatially normalized to a common reference space. The MNI standard space [65] was used in combination with a T1 template from the International Consortium for Brain Mapping (ICBM) [66].

Paper III
In clinical routine, many patients examined with FDG-PET only have a prior CT scan. In order to validate the typical clinical situation, the MRI scans were not used in paper III, and the registration was PET-driven.

Paper IV
Perfusion maps were created from the raw ASL data at each site (Amsterdam, Uppsala). The perfusion maps were normalized to template space using an MRI-driven registration from each subject’s T1-weighted image. The healthy controls from each site were compiled to constitute separate control groups. The perfusion map for each subject was then compared pixel-wise to the respective control group of the corresponding site, and Z-scores were
calculated for pixels with a value below the average of the control group. These Z-score maps were shown as colourized overlay images depicting significant deficits. The same procedure was also carried out using the FDG-PET data.

**PET protocol**

All subjects underwent a PET examination starting 25 min after injection of 3 MBq/kg $^{18}$F-FDG (mean administered activity 235 MBq, effective dose 4.47 mSv), either on a Discovery ST (GE Healthcare, Waukesha, WI, USA) PET/CT scanner or on an ECAT Exact HR+ PET scanner (Siemens, Knoxville, TN, USA).

Images from the PET/CT were reconstructed using ordered subsets expectation maximization (OSEM; 2 iterations, 21 subsets), applying all appropriate corrections including attenuation correction based on a low-dose CT scan acquired prior to PET scanning.

Images from the ECAT scanner were reconstructed using normalization and attenuation-weighted OSEM (6 iterations, 8 subsets), with attenuation corrections based on a 10-min transmission scan with rotating $^{68}$Ge rod sources.

Spatial resolution of the resulting images was similar for both scanners using the chosen reconstruction settings, with reconstructed voxel sizes of 2x2x3.7 and 2x2x2.4 mm for PET/CT and PET images, respectively, and a spatial resolution of approximately 6 mm.

Sixteen of the healthy controls and eight of the patients were enrolled for inclusion in paper II and paper III. These subjects underwent a second PET examination on a separate occasion. The second scan was either a low dose (LD) scan using 25% of the normal dose (mean administered activity 57 MBq, effective dose 1.1 mSv), or an ultra low dose (ULD) scan using 10% of the normal dose (mean 24 MBq, effective dose 0.46 mSv).

The LD dosage was 0.75 MBq/kg and the ULD dosage was 0.3 MBq/kg. The scanner and examination routine were the same for both scans for each subject.

**PET post-processing**

**Paper I**

Standardized uptake value ratios were used as the gold standard for evaluating ASL perfusion assessments. All PET images were registered to Montreal Neurological Institute (MNI) standard space [65, 67], and a volume of inter-
est (VOI) template derived from the automated anatomical labeling (AAL) atlas [68] was applied. SUV ratios normalized to cerebellum grey matter were computed. From the AAL, smaller regions were combined to larger regions correlating to the assessment protocol (frontal, temporal, parietal, occipital, and a combined region for precuneus and posterior cingulum), using weighted averages.

**Paper II**

Twenty-minute summations (25–45 min) were used. Normal-dose and low-dose FDG-PET images were first co-registered using rigid registration to the subject’s T1-weighted MR image. The T1 images were spatially normalized as described earlier. The calculated transforms for MR images were then applied to the co-registered PET images. Once PET images were transformed to template space, they were intensity-normalized using a global reference value derived from thresholded mean values of all voxels within a brain mask. See sample images in Figure 4.

Regional SUV ratios were calculated by dividing each region of interest with the global mean value of the thresholded brain mask, using an atlas derived from the automated anatomic labeling (AAL) atlas [68]. Volumes of interest were the anterior cingulate, frontal, lateral temporal, parietal, and a combined (precuneus + posterior cingulate) cortex from each side.
Figure 4. Three sample subjects from paper II. The top two rows show a healthy control and a sample patient, comparing normal- and low-dose images. The bottom row shows a healthy control comparing normal and ultra-low-dose. Images are shown in Sokoloff colour scheme.
Paper III
Ten-minute summations (35–45 min) from the normal- and low-dose images described in paper II were used. The images were spatially normalized using a PET-driven registration. Z-score values using cerebellum as reference region were calculated with volumes of interest generated from an automatic atlas and used for quantitative analysis. As a second step, 3D-SSP images showing Z-scores of < -1.96 were created, and presented randomized and blinded in two separate qualitative reading sessions as described below.

Paper IV
Intensity-normalized FDG-PET images were used to create Z-score maps based on the same post-processing strategy as described in the corresponding section in “MRI-processing”, described above. Thus, two Z-score maps were created from each subject, one from each modality.

Visual assessments

Paper I
Three experienced diagnostic physicians visually assessed the ASL images. The readers had access to only the ASL images and not to other MRI or PET images or clinical information. The subjects were presented blinded, with patients and healthy controls mixed in the same reading session. The assessment was done separately in nine brain regions for each subject: frontal, temporal, parietal and occipital lobes in each hemisphere, and a separate region for precuneus + posterior cingulum bilaterally. Each of the nine brain regions was qualitatively graded according to: 0 = normal perfusion, 1 = suspected hypoperfusion, 2 = hypoperfusion, or X = not able to evaluate. Each perfusion rating from ASL (0, 1, or 2) was compared with the FDG-PET SUV ratio in the corresponding anatomical region, and correlation coefficients were calculated. The coefficient thus indicates whether the reported hypoperfusion corresponds to measured hypometabolism in the same region or not. Thus, the FDG-PET was used as the gold standard for regional neurodegeneration.

Furthermore, the assessment results from beta smoothing images were compared to those from ASL subtraction in order to evaluate potentially improved characterization.

Each time a reader assessed a subject, a forced decision was also made as to whether the subject was a patient with any neurodegenerative disease or a healthy control, and discriminative performance was calculated.
Paper II
No visual assessments were performed in this paper.

Paper III
In addition to quantitative analysis, two experienced nuclear medicine specialists visually assessed the images on two separate sessions. In the first session, the readers viewed randomized Z-score maps from each subject individually, without knowing whether the images were from a patient or a control, normal-dose, or low-dose. The reader classified each image as normal or pathological, giving a confidence score of 0 to 5. If the image was classified as pathological, the reader also noted a specific diagnosis (AD/FTD/LBD/CBD).

The second session was carried out on a subsequent occasion. The readers viewed two sets of Z-score images at a time: the normal- and low-dose images from each subject. In half of the cases (randomly decided), the normal-dose images were placed above the low-dose images and vice versa. The clinical diagnosis of the subject (Control/AD/FTD/LBD/CBD) was provided next to the images. For each subject, the readers decided which set of images was most consistent with the clinical diagnosis, together with a subjective quantification of how large the differences between the images were, on a scale from 0 to 5. Sample images are shown in Figure 5. Note that a difference of zero, meaning that the images were equally consistent with the diagnosis, was a valid option for the readers as some pairs of images were identical or nearly identical.
Figure 5. (A-C) Comparisons of Z-score maps obtained in subjects given a normal dose or a low dose. The upper rows in each section show Z-score maps derived from subjects given a normal dose, the lower rows from low-dose. A) Healthy control with an unspecific artefact along the midline. B) Sample patient (SD) with similar images between doses. C) The patient (CBD) with the highest difference score from the assessments in session 2.

**Paper IV**

All Z-score maps were presented using an in-house developed interface software (Open Mira), providing all relevant features of a medical imaging viewer, including sagittal, transversal, and coronal views; scrolling through the slices of all views; and zoom/pan functionality. All maps were randomly intermixed and blinded with regard to their origin. The readers did not have access to any other images such as the original PET or T1, nor to any clinical information. Thresholded Z-scores were shown as overlays in a dichromatic colour scale, over an averaged and heavily smoothed greyscale brain image without specific anatomical information, to avoid assessment bias based on atrophy patterns. The default Z-score thresholds shown in the over-
lay were 2.0 (lower) and 6.0 (upper). The thresholds were manually adjustable by the readers. Six sample patients are shown in Figure 6, with the ASL-based and FDG-based Z-score maps placed pairwise for comparison. Clinically relevant sections have been chosen, and the Z-scores are manually thresholded for similarity. The selected cases show varying degrees of concordance between ASL-based and FDG-based Z-score maps.

Figure 6. (A–F): Sample Z-score maps from six patients. Z-score maps from six subjects are shown, with the maps from each subject placed pairwise for comparison. The left column (A–C) shows AD patients and the right column shows FTD. The selected subjects have varying degrees of concordance between ASL-based maps (upper row in each box) and FDG-based (lower rows). Clinically relevant sections were chosen, and the Z-scores were manually thresholded for similarity.

Four experienced physicians with a specific interest in dementia imaging visually assessed and rated all images, two neuroradiologists and two nuclear medicine specialists. Each reader assessed each Z-score map and classified it as either a healthy control or patient, together with a confidence level
(range 0–4) representing how convincing the control/patient discrimination was. For subjects rated as patients, the readers also made a differential diagnostic decision between AD and FTD.

Statistical analysis

**Paper I**
Non-parametric Pearson correlation was used for comparison of the ASL assessment median scores with FDG-PET SUV ratios from the corresponding brain regions, and correlation coefficients (c) and p values were calculated. The specific test was chosen due to the comparison of a parametric to a non-parametric (ordinal) variable. A significance level of < 0.01 was chosen for coefficient tables due to multiple comparisons.

Kappa values were calculated using Fleiss kappa. Separate interreader kappa values were determined for the readers pairwise per region and postprocessing method. Paired Wilcoxon tests of median values were used to compare differences in the grades for subjective image quality, for each post-processing type.

**Paper II**
SUV ratios extracted from the normal- and low-dose scans were compared pairwise. The differences were tested for normal distribution using the Shapiro–Wilk test and visualized with normal probability plots for detection of outliers. Bland–Altman plots and boxplots were calculated for visual evaluation of variance. Single sample t-tests were used to determine whether there was constant bias between the normal and low-dose SUV ratios, and simple regression was used to quantify proportional bias, expressed as an \( r^2 \) and a p value.

**Paper III**
In the quantitative part of the study, the Z-score differences were tested for normal distribution using the Shapiro–Wilk test and visualized with normal probability plots for detection of outliers. Single sample t-tests were used to determine whether there was constant bias between the normal and low-dose Z-scores, and simple regression was used to quantify proportional bias, expressed as an \( r^2 \) and a p value. The procedure was repeated after exclusion of the ULD group and all non-neocortical regions.

In the qualitative session one, Cohen’s kappa test was used for inter-reader agreement, and to describe the contingency tables in conjunction with the Pearson \( \chi^2 \) test. The Wilcoxon matched pairs test was used to compare the confidence scores. A proportion test was used for the rate of correct diagno-
ses. In session two, Wilcoxon matched pairs and mean values with confidence intervals were used to determine whether there was any significant preference for normal or low-dose-derived images.

**Paper IV**

Contingency tables were constructed with Cohen’s kappa test and Pearson chi$^2$ tests. Proportion tests, calculated for each modality by adding the assessments from all readers, were used to compare performance. Interreader agreement per modality was evaluated pairwise for all readers using Cohen’s kappa and collectively for all four readers using Fleiss’ kappa. Wilcoxon’s matched pairs test was used to compare the confidence levels.

**Ethical considerations**

The basis of research ethics on human beings is the informed consent of the volunteers. With children and patients with cognitive impairment, the term “informed” is elusive. The patients enrolled in these studies were first considered by a senior geriatrician with long experience in dementia research (LK), and they were deemed to be in a condition to decide whether to participate or not. Nevertheless, almost all of them (not the mildest cases) had a family member present during inclusion who co-signed the documents. The local ethics committee approved the study plan.

PET and PET/CT examinations infer a dose of ionizing radiation to the subject. The subjects who underwent two PET examinations (paper II and paper III), received radiation doses of 5 to 6 mSv, which approximately corresponds to two times the annual radiation dose to the average person living in Sweden. Since the stochastically increased risk of harm from radiation decreases significantly with age, this was considered to be acceptable. A specific approval from the local committee for radiation protection was retrieved prior to scanning.

If the validation of low-dose FDG-PET results in an acceptance of decreased FDG tracer dosage, much has been won for future patients and research controls.
Results

Paper I

Correlation coefficients between median scores from visual ASL assessments and FDG-PET SUV ratios were calculated for all regions. For subtraction images, the correlation was -0.18, with a $p$ value of 0.436. For the beta-smoothed images, the correlation was -0.54, with a $p$ value of 0.012. A negative correlation value indicates that high scores in visual assessment (i.e. apparent hypoperfusion) correlated to low SUV ratios (i.e. hypometabolism). The results from frontal and parietal regions are shown in Figure 7.

![Figure 7](image.png)

*Figure 7.** Comparison of ASL and FDG-PET for frontal and parietal lobes. On the x-axis in each of the eight columns, a median assessment score ranging from 0 to 2 is shown for each region and post-processing method separately. The y-axis shows the FDG SUV ratio. A downward slope on the line shows a correlation between high score (visually assessed hypoperfusion) and low SUV ratio (hypometabolism). At the bottom of each column, the correlation coefficient (c) and $p$ value are shown. Regions with a correlation coefficient of $p < 0.01$ are marked with an asterisk.

Regarding the discrimination of patients from healthy elderly, spatial smoothing of ASL images increased false positive results, as shown in Table 1. The subjective image quality was graded higher for beta-smoothed images than for subtraction.
<table>
<thead>
<tr>
<th>Image rated as:</th>
<th>Subtraction</th>
<th>Beta smoothing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Controls</td>
<td>Patients</td>
</tr>
<tr>
<td>Normal</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>Pathological</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Neg pred</td>
<td>73%</td>
<td>61%</td>
</tr>
<tr>
<td>Precision</td>
<td>100%</td>
<td>63%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>60%</td>
<td>50%</td>
</tr>
<tr>
<td>Specificity</td>
<td>100%</td>
<td>73%</td>
</tr>
<tr>
<td>Accuracy</td>
<td>81%</td>
<td>62%</td>
</tr>
</tbody>
</table>

**Table 1.** Discrimination of patients from healthy controls.

According to the forced decision, “Normal” was assigned to subjects with no suspicion of hypoperfusion, and “Pathological” to subjects with suspected neurodegenerative disease, visually assessed on ASL. K and p represents kappa- and chi²-derived p values for the contingency tables, respectively. The highest kappa value is marked in bold, and p value of < 0.01 is marked with an asterisk.

**Paper II**

Regional SUV ratios comparing normal-dose and low-dose/ultra-low-dose were examined using Bland-Altman analysis, revealing no clinically significant bias. Absolute values of the difference in each region (ND-LD) were calculated and expressed in percent. The low-dose protocol produced a mean absolute difference in regional SUV ratios of 0.015 (1.32%) in controls, and 0.019 (1.67%) in patients with dementia disorders, as shown in Table 2. The ultra-low-dose protocol produced a slightly higher mean difference of 0.023 (2.10%), with the largest differences noted in the anterior cingulate cortices.
<table>
<thead>
<tr>
<th>Region</th>
<th>LD controls</th>
<th>LD patients</th>
<th>ULD controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left ant. cingulate</td>
<td>1.02% (0.05-2.65)</td>
<td>2.14% (0.44-5.65)</td>
<td><strong>4.71%</strong> (2.96-7.32)</td>
</tr>
<tr>
<td>Right ant. cingulate</td>
<td>1.39% (0.48-3.64)</td>
<td>2.08% (0.02-5.66)</td>
<td><strong>4.20%</strong> (0.21-8.39)</td>
</tr>
<tr>
<td>Left frontal</td>
<td>0.79% (0.18-1.51)</td>
<td>1.10% (0.02-3.33)</td>
<td>1.05% (0.32-1.57)</td>
</tr>
<tr>
<td>Right frontal</td>
<td>0.77% (0.23-2.72)</td>
<td>0.83% (0.18-2.11)</td>
<td>1.54% (0.53-3.16)</td>
</tr>
<tr>
<td>Left lateral temporal</td>
<td>1.31% (0.74-2.60)</td>
<td>1.58% (0.70-2.24)</td>
<td>2.05% (0.42-3.82)</td>
</tr>
<tr>
<td>Right lat temporal</td>
<td>0.55% (0.01-2.04)</td>
<td>1.64% (0.13-3.93)</td>
<td>1.42% (0.44-2.93)</td>
</tr>
<tr>
<td>Left occipital</td>
<td>1.85% (0.21-4.37)</td>
<td>1.76% (0.54-3.26)</td>
<td>0.94% (0.33-1.55)</td>
</tr>
<tr>
<td>Right occipital</td>
<td>1.95% (0.37-4.42)</td>
<td>1.35% (0.05-2.87)</td>
<td>1.58% (0.84-2.28)</td>
</tr>
<tr>
<td>Left parietal</td>
<td>1.63% (0.17-3.48)</td>
<td>1.68% (0.38-4.38)</td>
<td><strong>3.61%</strong> (0.47-6.21)</td>
</tr>
<tr>
<td>Right parietal</td>
<td>1.70% (0.12-5.04)</td>
<td>1.76% (0.23-2.87)</td>
<td>1.01% (0.10-2.08)</td>
</tr>
<tr>
<td>Left precuneus + post cingulate</td>
<td>1.25% (0.52-2.90)</td>
<td>2.04% (0.44-3.76)</td>
<td>1.68% (0.07-5.43)</td>
</tr>
<tr>
<td>Right precuneus + post cingulate</td>
<td>1.62% (0.50-3.89)</td>
<td>2.12% (0.49-5.12)</td>
<td>1.42% (0.17-3.50)</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td><strong>1.32%</strong> (0.01-5.04)</td>
<td><strong>1.67%</strong> (0.02-5.66)</td>
<td><strong>2.10%</strong> (0.07-8.39)</td>
</tr>
</tbody>
</table>

Table 2. Average absolute differences per region.

The columns represent the low-dose healthy controls, the low-dose patients, and the ultra-low-dose controls, respectively. Each row represents the mean absolute difference in SUV ratios from the normal dose in a cortical region, expressed as a percentage, with the range in parentheses. All mean values higher than 3% are marked in bold for easy reference. Ant = anterior. Post = posterior.

**Paper III**

Regarding the quantitative section, the differences in Z-scores between normal- and low-dose were normally distributed. Considering -1.96 as a general cut-off limit to separate normal from pathological, 6.33% of the data points would be misclassified (false negative, false positive) by the low-dose protocol compared to normal-dose.

Bland–Altman comparisons between the methods were performed with and without the results from the ULD group and non-neocortical regions. Using the LD group and neocortical regions only, there was a constant bias of 0.206 (p < 0.001), and a negative proportional bias (r² = 0.014, p = 0.002). The combination of a constant positive bias and a negative proportional bias signifies that negative Z-scores were reported as less negative by the low-
dose protocol compared to normal-dose (e.g. LD showing -2.34, while ND showed -2.54), whereas no such bias was found for positive Z-scores.

Regarding the discrimination between patients and healthy controls in the qualitative section, the sum of assessments in each category was identical, resulting in equal discriminative performance between doses, as shown in Table 3. There was a slight difference in the rate of correct differential diagnosis (67% vs. 56%), not reaching statistical significance (p = 0.49). Pair-wise matching of confidence scores showed no significant difference. The rate of uncertain cases was 6% in both dose groups.

<table>
<thead>
<tr>
<th>Normal-dose</th>
<th>Low-dose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image rated as:</td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td>Patients</td>
</tr>
<tr>
<td>Normal</td>
<td>15</td>
</tr>
<tr>
<td>Pathological</td>
<td>3</td>
</tr>
<tr>
<td>Of which uncertain</td>
<td>1</td>
</tr>
</tbody>
</table>

| Sensitivity | 72% | 72% |
| Specificity | 83% | 83% |
| Neg pred v | 75% | 75% |
| Precision | 81% | 81% |
| Accuracy | 78% | 78% |
| Correct ddx | 67% | 56% |

| n:36 | n: 36 |
| K: 0.556 | K: 0.556 |
| p: 0.0008 | p: 0.0008 |

Table 3 shows the results from the diagnostic assessment in paper III. Assessments from both readers are combined (n = 18 x 2). “Neg pred v” represents negative predictive value. The assessments were not identical between normal- and low-dose-derived images, but the numbers of assessments in each category were, thus resulting in identical discriminative performance between the methods. “Uncertain” signifies assessments with a discriminative confidence score < 3. “Correct ddx” refers to the rate of correct diagnoses among the true positive cases. K and p represents kappa- and chi²-derived p values for the contingency tables, respectively.

In the second session of the qualitative section, the readers compared which set of images was most consistent with the clinical diagnosis. The normal-dose images were preferred in 8% of the assessments, the low-dose images in 13%, and they were deemed to be equal in 79%. The difference was not significant. The data was also translated to a semi-continuous scale of 11 steps, where “-5” represented strongest difference in favour of LD, “0” represented no difference, and “5” strongest difference in favour of ND. The mean of the assessments on this scale was -0.05, and the 95% confidence interval included zero.
In a first step, we analyzed the readers’ performance in distinguishing between patients and controls during assessments. Contingency tables of distinction for all assessments are shown by modality in Table 4. Performance results and other comparisons are listed in Table 5.

<table>
<thead>
<tr>
<th>Measure</th>
<th>ASL</th>
<th>FDG-PET</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>53% (53%)</td>
<td>96% (99%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Specificity</td>
<td>84% (85%)</td>
<td>54% (54%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>60% (61%)</td>
<td>91% (97%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>80% (81%)</td>
<td>71% (76%)</td>
<td>0.007</td>
</tr>
<tr>
<td>Accuracy</td>
<td>68% (68%)</td>
<td>77% (80%)</td>
<td>0.008</td>
</tr>
<tr>
<td>Uncertain cases, controls</td>
<td>7%</td>
<td>25%</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Uncertain cases, patients</td>
<td>8%</td>
<td>7%</td>
<td>0.741</td>
</tr>
<tr>
<td>Inter-reader agreement</td>
<td>0.58</td>
<td>0.38</td>
<td>N/A</td>
</tr>
<tr>
<td>Unanimous distinction</td>
<td>66%</td>
<td>57%</td>
<td>0.235</td>
</tr>
<tr>
<td>Unanimous correct dist.</td>
<td>49%</td>
<td>52%</td>
<td>0.607</td>
</tr>
<tr>
<td>Distinction of AD cases</td>
<td>57%</td>
<td>95%</td>
<td>0.002</td>
</tr>
<tr>
<td>Distinction of FTD cases</td>
<td>49%</td>
<td>96%</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Higher confidence score</td>
<td>125</td>
<td>87</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 5. Patient/control performance results by modality.

Values that were significantly higher than the corresponding value from the other modality have been marked in bold. The performance results in parentheses show the results after exclusion of uncertain cases.

Comments: 1) Rate of cases with confidence level < 2. 2) Fleiss’ kappa describing inter-reader agreement for all readers collectively. 3) “Unanimous distinction” means that all readers agreed on the patient/control distinction. 4) All readers agreed on the correct distinction. 5) The rate of patients from each diagnosis category that were correctly identified as patients. AD = Alzheimer’s disease, FTD = frontotemporal dementia. 6) The number of instances with higher confidence score than the corresponding image from the other modality, examined with Wilcoxon matched pairs.
For ASL-based maps, kappa values for pairwise interreader agreement were 0.71, 0.70, 0.67, 0.54, 0.54, and 0.40, and the total kappa value for all readers collectively was 0.58. For FDG-PET-based maps, kappa values for pairwise agreement were 0.51, 0.51, 0.39, 0.37, 0.36, and 0.33, and the total kappa value for all readers collectively was 0.38. No specific differences were identified between the readers with a background in nuclear medicine, compared to the neuroradiologists.

As shown in Table 5, ASL-based images yielded higher specificity, positive predictive value, inter-reader agreement, and confidence scores. FDG-based images yielded higher sensitivity, negative predictive value, and accuracy. However, the FDG-based images of controls contained a large number of uncertain cases. Exclusion of uncertain cases resulted in modest improvements in both modalities, shown in parentheses in the table.

In a second step, we analyzed the readers’ performance in assessing the correct differential diagnosis (AD vs FTD). Results are given in Table 6.

<table>
<thead>
<tr>
<th>Measure</th>
<th>ASL</th>
<th>FDG-PET</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct differential diagn.</td>
<td>76%</td>
<td>(80%)</td>
<td>0.168</td>
</tr>
<tr>
<td>Unanimous diff. diagn.</td>
<td>24%</td>
<td>60%</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Correct diagn. rate in AD</td>
<td>89%</td>
<td>82%</td>
<td>0.482</td>
</tr>
<tr>
<td>Correct diagn. rate in FTD</td>
<td>56%</td>
<td>87%</td>
<td>0.030</td>
</tr>
</tbody>
</table>

Table 6. Differential diagnosis performance results by modality. Values that were significantly higher than the corresponding value from the other modality have been marked in bold. The performance results in parentheses show the results after exclusion of uncertain cases. “Diff.” = Differential. “Diagn.” = Diagnosis.

Comments: 1) Rate of correct differential diagnosis among true positive cases. 2) All readers gave correct differential diagnosis. 3) Rate of correct diagnosis among true positive cases sorted by diagnosis (AD = Alzheimer’s disease, FTD = fronto-temporal dementia).

The patient groups were examined separately. ASL yielded a higher rate of correct differential diagnosis in AD cases compared to FTD cases (89% vs. 56%, p < 0.001). FDG-PET yielded similar results between patient groups (data shown in Table 6).

The cohorts were also examined separately. The rate of false positive findings was higher in ASL images from the Uppsala cohort compared to the Amsterdam cohort (16% vs. 3%, p = 0.015). In other respects, the cohorts were similar (data not shown).
A sample healthy control with false positive findings is shown in Figure 8. Both the ASL-based and FDG-PET-based Z-score maps from this subject show artefactual Z-score patterns representative for the respective modality in this study.

![Figure 8. Healthy control with false positive findings. Both the ASL-based and FDG-PET-based Z-score maps show artefactual Z-score patterns representative for the respective modality in this study. The ASL-based map from this particular healthy control was read as Alzheimer’s disease (AD) by one reader. The FDG-based images were read as AD by two readers, and frontotemporal dementia by one reader. The clusters at the junction between grey and white matter in the lower row were present in several of the FDG-PET images of healthy controls, contributing to the number of uncertain cases and false positive assessments.](image)

A representative discordant patient is shown in Figure 9. The ASL-based map of this AD patient was read as false negative by three readers, while all four readers correctly diagnosed the FDG-PET-based map. Several of the false negative ASL cases had disease-typical deficit patterns that were less evident, or not at all evident, using the default settings, but could be revealed by individually adjusting the Z-score threshold.
Figure 9. Representative discordant case. The ASL-based image of this AD case (Age: 67, MMSE: 25) was assessed as false negative by three of four readers, while the corresponding FDG-PET image was unanimously correctly diagnosed. The top row shows the ASL-based image with the default Z-score thresholds (lower 2.0, upper 6.0). The middle row shows the same image after threshold adjustment (0.77, 1.87). The bottom row shows the FDG-PET-based images of the same subject, with default threshold settings (2.0, 6.0). This is a representative case of a disease-typical deficit pattern evident on ASL-based images only after individual adjustment of the threshold.
Discussion

Visual assessment

In research articles, quantitative subjective grading is often referred to as the “Likert scale”. The original Likert scale was designed to measure attitudes and is widely used. The scale yields an ordinal dataset with arbitrary numbers that can be used to establish a rank order [69]. Since the data is ordinal, non-parametric statistics are appropriate. Subjective grading or semi-quantification is a suboptimal tool due to the imperfect test-retest reliability of human readers, although sometimes unavoidable in imaging research that includes visual assessments. Pairwise comparisons have been shown to be more reliable than subjective grading alone [69], which motivated the second reading session in paper III. In research projects with limited numbers of patients, this issue needs to be considered during study design. Clear definitions of the steps in the ordinal scales are also important.

Agreement is a common end-point in radiology studies, often expressed with a kappa value that describes the level of similarity of the ratings by two or more readers. The kappa values included in this thesis are generally low, compared to many other publications. An important reason for this is the exceptionally high degrees of freedom associated with the performed assessments. If the assessments were based on a less complex task, e.g. grading opacifications on a chest X-ray, the readers would base their assessments on the exact same view and greyscale information. In papers I and IV, the readers could scroll through stacks of images and manually adjust the window level/contrast. This was deliberately chosen to imitate a clinical setting and to draw upon the clinical diagnostic skills and experience of the readers. However, each increase of freedom for the readers reduces the level of similarity of the conditions in which the assessments are made. As a consequence, the inter-reader agreement is challenged by complexity. In both of these papers, tutorial training sessions were performed prior to the reading sessions, with the ambition to reduce this complexity and to promote a common assessment strategy. In order to completely avoid this problem, joint reading sessions with consensus-based ratings might be considered.
Paper I

Post-processing with Gaussian smoothing strengthened the correlation between visual ASL rating and FDG-PET SUV ratios for frontal and parietal regions. However, it decreased inter- and intra-observer agreement, and discrimination of patients from healthy controls was less satisfying due to false positives. Signal changes due to small movements, pulsations due to the cardiac cycle, breathing, and potentially other physiological activity, might add up and constitute a source of low frequency noise in standard subtraction images. We hypothesized that using a general linear modelling filter could remove such low frequency noise, but were unable to show significant improvements in this visual assessment setting. Limitations of this study include small group size and that the readers may have been insufficiently familiar with the types of images included in the study.

Paper II

The main finding in this study is that 0.75 MBq/kg (56 MBq for a 75 kg subject) is a sufficient FDG dose for conducting brain PET scans with retained SUV ratio characteristics in adults with or without neurodegenerative disease, without increasing scan time. The discrepancies in SUV ratios between normal- and low-dose subjects were minor, both for patients and for healthy controls, and no clinically significant bias or trend was found. Considering the technological advantages of more modern equipment, the results from this study should be applicable to any PET facility.

The ultra-low-dose protocol should be considered for research applications, especially in young individuals and for paradigms requiring serial scanning over time.

Limitations include the lack of analysis of absolute values (e.g. MRglu).

Paper III

Visual assessment of 3D-SSP Z-score maps from low-dose FDG-PET scans provided reliable diagnostic information, although the quantitative values had some degree of variance compared to Z-scores from normal-dose examinations. The findings include a slight constant bias of negative Z-score values from the low-dose protocol, but this does not seem to be a clinically relevant issue. Regarding differential diagnosis, there was a minor trend towards an advantage for the ND images, but the difference was not statistically significant. The results from this study in conjunction with paper II validates a substantial FDG-dose reduction in clinical routine.
The main limitation of this study is the limited number of included patients. Adding a cohort of virtual low-dose images created from short summations of normal-dose images would be an interesting way of consolidating the results. Another way of strengthening this validation would be to repeat the design using another 3D-SSP software.

Paper IV

The main finding in this study is that ASL-based Z-score maps can be visually assessed with high specificity and positive predictive value in patients with AD and FTD. Sensitivity and accuracy were lower than in the FDG-based images. ASL can easily be added to the morphological MRI protocol of a routine dementia work-up, and creating Z-maps constitutes a relatively minor additional effort. Due to the high specificity and positive predictive value of ASL, an abnormal finding on ASL can indicate a specific diagnosis, with little additional effort. If ASL is negative or inconclusive but the clinical suspicion of neurodegenerative dementia persists, the addition of FDG-PET could be considered due to its excellent sensitivity and superior accuracy.

The lower sensitivity of ASL in this study may imply that regional perfusion decrease is evident at a later stage than the regional decrease of glucose metabolism seen with FDG-PET. Thus, deficits on ASL may be a later sign with lower sensitivity compared to FDG-PET. Since the ASL scans were acquired prior to the FDG-PET scans, there is also a theoretical possibility that disease progression between scans could contribute to the difference in sensitivity, but this was considered unlikely. The low sensitivity of ASL in this study is also associated with the Z-score threshold, as discussed separately.

ASL was more efficient in correctly diagnosing AD cases compared to FTD cases, whereas there was no difference between patient groups with regard to FDG-PET. This could possibly represent a more disease-specific change of perfusion in AD than in FTD. This issue should be addressed in a larger study with another approach. Differential diagnosing was less robust in ASL than in FDG-PET.

When the cohorts were compared, a higher rate of false positive findings was found among the ASL images in the Uppsala cohort. This could be caused by a methodological difference between the ASL protocols at the two sites [22]. Another possible cause is the higher mean age in the Uppsala control group, which can be associated with increased ASL variability due to
slow circulation (the expected arterial transit time increases with age, partly due to concurrent cardiovascular pathology [22]).

The rate of uncertain cases and false positive results were higher in the FDG-based images of healthy controls, associated with method-dependent artefacts. As a consequence, the FDG-based Z-scores in this study showed poor specificity and inter-reader agreement, which is not in agreement with studies based on the 3D-SSP method [17, 70]. This presumably depends on a crucial methodological difference: 3D-SSP maps are based on specific post-processing operations not used in this study. Instead, the current study used a voxel-wise comparison of all voxels after spatial normalization. Thus, the method used in this study is less sophisticated, and may be more vulnerable to registration mismatch and partial volume effects.

The 3D-SSP tools were not used in the current investigation because none of the existing variations of the SSP technique known to the authors is adapted to perfusion data and a method-dependent bias could not be excluded. We therefore used a standard voxelwise comparison in the current investigation, with the aim to directly compare ASL versus FDG-PET Z-maps. As a next step, assessing the utility of SSP analysis of ASL data would be an interesting pursuit.

Evaluation of ASL perfusion has several potential pitfalls such as artefacts and variability. The ASL quantification models currently available rely partly on approximations and assumptions, and development is still on-going [38, 71]. There is considerable variance in cerebral perfusion levels in healthy elderly, which makes cut-off values between normal and pathological perfusion difficult to establish [36, 72, 73]. The inter-individual variance in ASL perfusion is large compared to the variance of SUV ratios in FDG-PET, with some subjects having twice the perfusion of others [36, 72, 73]. This is an issue when a single subject is compared with a control group of limited size. It is plausible that a larger normal database, preferably with subsampling of age and gender, would facilitate the distinction between normal and pathological. This is probably even more important for ASL than for FDG-PET. As shown in Figure 9, the default threshold was not always suitable for the ASL-based images. In certain cases, adjusting the threshold revealed a distinct disease-typical pattern. The issue with thresholding is associated with the inter-individual variation in perfusion, and with the size and quality of the control group, as discussed above. A user-independent solution to this issue would be beneficial. Applying individual intensity normalization has been successfully used in other studies to display diseasespecific perfusion patterns in the presence of large variations [49, 74]. One possibility could be to use an intracranial reference region and display perfusion ratios instead of absolute values. This would be analogous to the proce-
dure in FDG-PET, where SUV ratios are obtained using a reference region such as the pons, cerebellum, or whole-brain (used in this study). However, ASL is not yet fully established and is still undergoing intensive development. It is plausible in the near future that improved ASL sequences can provide more robust data with less variance, which would alleviate the issue discussed above.

The differences in scanning protocols between the two investigative sites can be considered a limitation due to possible bias. Another limitation of this investigation is the sizes of the control groups. The ideal control group would not only be larger, but also perfectly matched to the patients with regard to gender and age or large enough to sub-sample accordingly.
Future perspectives

ASL perfusion

In addition to its applicability in cerebrovascular and neuro-oncological disease, the ASL technique has great potential in the field of neurodegeneration. The possibility of achieving diagnostic information comparable to that obtained by FDG-PET using a four-minute non-invasive MRI sequence is inspiring. The adaptation of ASL from an experimental research method to a clinical routine method is currently taking place, and its modus operandi and place in the clinic are far from settled. When established, ASL will provide a more widespread availability compared to FDG-PET. Newer ASL sequences often use 3D readout, with partially different results from the 2D method used in this thesis.

The considerable inter-individual variance of brain perfusion in healthy controls is obviously an issue when a single subject is compared to a control group of limited size. Many confounding factors can affect the perfusion, e.g. coffee, smoking, exercise, medications, and time of the day. It is plausible that a reference database of normal perfusion needs to be large for such comparisons to be robust. Even then, an adapted thresholding strategy may be necessary for robust implementation. Another possibility would be to use an intracranial reference region and display ratios instead of absolute values, as is done in dynamic susceptibility contrast (DSC) perfusion MRI or in FDG-PET, where pons, cerebellum, or whole-brain can be used as reference.

No 3D-SSP software known to the authors can currently be easily customized enough to use perfusion data. Implementation of ASL in a 3D-SSP tool would probably improve the performance compared to the voxel-wise comparison used in paper IV of this thesis. This would be an interesting pursuit for future development.

Low-dose FDG-PET

Two standard deviations below normal is a classical diagnostic cut-off, leaving 97.5% of the normal data above the cut. Due to the current and expected interest in early diagnosis, a more nuanced assessment would be appealing.
In this pursuit, Z-score values between 0 and -1.96 could be clinically relevant if they comprise a disease-typical pattern in a patient with mild cognitive impairment.

In the search for methods to quantify disease progression rate, many studies have explored the rate of atrophy and the rate of ventricular dilatation, as summarized by Frisoni et al. [15]. Annual reduction of hippocampal volume has been found to be larger in AD patients than in age-matched controls. Repeated low-dose FDG-PET examinations may provide a similar quantitative “rate-of-progression” measurement for on-going degeneration in cortical metabolism. This could be useful for identifying pathology at early stages of disease, and for monitoring disease progression rate [75]. This may also be a useful tool for validating the effects of disease-modifying pharmaceutical agents. A high level of reproducibility is required for this approach, as well as identical conditions between scans. Using 3D-SSP maps to calculate and visualize the rate of progression is probably a useful approach.

The potential future significance of the dose reduction described in this thesis reaches beyond the field of neurodegenerative dementia. Ethical issues and concerns regarding radiation to volunteers would be considerably less relevant in all kinds of research involving FDG-PET in healthy controls. This is especially true for young subjects, even children, and this opens up for countless research projects previously considered difficult due to radiation dose.

**Visual assessment**

Medical imaging has a long tradition of visual assessment, but its form and conditions need to be further explored and improved. Meanwhile, automated analyses such as multi-class classifiers, machine learning algorithms, and other computerized multimodal decision supportive tools are undergoing an explosive evolution [34]. It is very likely that many exciting new features will increase the usefulness of perfusion and metabolism imaging shortly. Quantification and machine learning algorithms can add information and offer decision support, but in the foreseeable future, visual assessment will play an important role for quality control and in confirming and refining computer-based analysis results.
Conclusions

Paper I
Spatial smoothing of ASL images increased false positive results in the discrimination of patients with neurodegenerative dementia from healthy elderly. However, regional characterization and subjective perception of image quality was improved.

Paper II
Low-dose FDG, 0.75 MBq/kg, is sufficient for evaluating regional standard uptake value ratios in brain PET scans in adults with or without neurodegenerative disease, and its use results in a reduction of total effective dose from 4.5 to 1.2 mSv. This opens for a reduced radiation burden to patients, research volunteers, and nuclear medicine staff, and enables more liberal inclusion of healthy volunteers in research studies.

Paper III
Reliable 3D-SSP Z-score maps with sufficient diagnostic information can be created from low-dose FDG-PET scans of the brain, after administering 0.75 MBq/kg FDG.

Paper IV
ASL perfusion-based Z-score maps can be used as a diagnostic tool in patients with suspected neurodegenerative disease. In the current study, ASL provided high specificity and positive predictive value, but inferior sensitivity compared to FDG-PET. FDG-PET might be reserved for those cases where ASL is negative and high clinical suspicion persists.
Sammanfattning på svenska

Titel
Visuell granskning av perfusion och metabolism vid neurodegenerativ demens.

Bakgrund


Demenstillstånd som är orsakade av kärlsjukdom och infarkter i hjärnan kallas vaskulär demens och omfattas inte av denna avhandling. I klinisk praxis ses inte sällan en blandbild av vaskulär och neurodegenerativ demens.

Neuroradiologiska hjärnavbildningsmetoder, oftast med datortomografi, har traditionellt använts främst för att utesluta andra, åtgärdbara orsaker till symptomen, som till exempel hjärntumör. I takt med att bildkvaliteten och kunskapen om sjukdomarna har ökat, så har fokus glidit över till att även påvisa fynd som kan stärka den kliniska diagnoserstanken, eftersom en del direkt och indirekta tecken kan ses på bilderna. Det tydligaste är att man kan se olika månader av atrofi, det vill säga volymförlust, som kan tala för eller mot vissa diagnoser. I tidigt sjukdomsskede är detta inte särskilt uttalat, och ibland är det svårt att skilja sjukdomstecken från normalt åldrande. Därför krävs det andra avbildningsmetoder med större känslighet. Vissa funktioner i hjärnan såsom blodgenomströmning (perfusion) och ämnesomsättning (metabolism) kan mätas och avbildas med avancerade, funktionella metoder. Sådana metoder har större chans att hitta sjukdomarna i tidigt skede. Perfusion har i många år avbildats med en metod som kallas SPECT, och de sen-
Aste åren har en ny, magnetkamerabaserad metod som kallas ASL blivit aktuell. Metabolism i hjärnbarken kan avbildas genom att spåra omsättningen av injicerat radioaktivt socker, FDG-PET. Ofta omvandlar man PET-bilderna till så kallade Z-score-kartor, som tydligt visar vilka områden i hjärnan som har signifikant nedsatt upptag av socker, jämfört med en grupp normala individers hjärnor.

Alla tre metoder som nämns ovan har begränsad klinisk användbarhet av olika orsaker. SPECT är inte tillräckligt känsligt och bilderna har inte tillräckligt bra upplösning. ASL är en ny metod som än så länge bara använts i forskning, och metoden håller för närvarande på att etableras som en rutinmässig metod, främst vid enstaka expertcentra. FDG-PET är robust och mycket användbar kliniskt, men är inte tillgänglig överallt, och har nackdelar i form av pris, tidsåtgång, samt att patienterna utsätts för en viss mängd joniserande strålning.

**Målsättning**

Det huvudsakliga målet med den här avhandlingen är att förbättra den kliniska användbarheten av att avbilda perfusion och metabolism hos patienter med misstänkt neurodegenerativ demenssjukdom, med fokus på visuell granskning.


**Metoder**

En grupp patienter med neurodegenerativa demenssjukdomar inkluderades från minnesmottagningen på Akademiska sjukhuset i Uppsala. Genom afsöjering rekryterades en åldersmatchad grupp med friska kontroller. Alla individer som inkluderades i studien genomgick både en MR-undersökning som inkluderade ASL, samt en FDG-PET undersökning. En delmängd av
deltagarna gjorde en ytterligare FDG-PET-undersökning, denna gång med låg dos FDG. Dessa undersökningar användes i Artiklar 2 och 3. Till artikel 4 kombinerades materialet med ett liknade material från Amsterdam, för att uppnå ett högre antal inkluderade patienter.

**Resultat**

Artikel 1. Effekten av den bildutjämnande funktionen förstärkte sambandet mellan visuellt bedömd perfusion med uppmätt metabolism, vilket användes som referensmetod. Dock så var bedömarna bäst på att skilja patienter från friska kontroller när de granskade ej utjämnade bilder.

Artikel 2. Vid jämförelse av regionala SUV-kvoter var variansen mellan normaldos och lågdosundersökningarna låg, och inga systematiska felvärden noterades.


**Diskussion**

När nya diagnostiska metoder övergår från experimentell forskning till klinisk rutin behöver man etablera pragmatiska och användbara metoder för att granska och utvärdera resultaten. ASL är ännu inte en etablerad och välbeprövad klinisk metod, och vissa frågetecken kvarstår kring hur metoden ska användas. ASL kan användas för att få fram kvantitativa mätvärden på hjärnans perfusion, men detta är komplex och den kliniska användbarheten av metoden är ännu begränsad. Artikel 1 och 4 belyser detta genom att utforska två olika metoder som kan användas vid visuell granskning i kliniskt rutinarbete.

Resultaten från artikel 1 är tvetydiga, men belyser att bildbehandlingsmetoder såsom smoothing har inverkan på hur bilderna bedöms, vilket är viktigt att ha med sig vid kliniskt arbete.

Resultaten från artikel 4 visar att Z-score-kartor kan bli ett användbart sätt att implementera ASL i kliniskt bruk, och kan i vissa fall vara ett alternativ

Acknowledgements

The journey that starts with a general idea that a thesis on dementia “would be fun” and ends in expressing gratitude for its fulfilment is long. Over the grassy plains of brain-storming study design, through the dense forest of ethical permission, on to the long road of searching for healthy controls, interviewing and examining patients, leaping the loose rocks of finding scanner time, then deep down the treacherous marshlands of post-processing images. Up again, to climb the steep mountains of writing and submitting, before the end is within reach. After six and a half years of tireless wandering, I now get to throw down my ring and turn to all of you who helped me along this rich adventure.

My supervisors:

Elna-Marie Larsson – my Gandalf and locomotive. Your astonishing ability to move forward and find solutions is unmatched. I will always be grateful for the meeting at Anja Smits’ office where we instantly found each other, and my road to radiology and dementia research began. With your executive skills, your respect and loyalty to your PhD students, and your sensitive leadership, you set the example that we all want to follow.

Lena Kilander – my hero in the clinical field. You combine the academic know-how of ethical permissions and solid writing with admirable patient skills, and you are brightly inspirational for all clinically active colleagues with your artful combination of knowledge and respect towards the patients and their families. If my path had not led me to radiology, I would have liked to follow you to the world of cognitive medicine.

Mark Lubberink – the rising star of nuclear medicine. Your fantastic ability to explain incomprehensible information at an adequate level has helped me to enter the field with curiosity and confidence. Your strong drive for research stimulates us all, and I hope we will collaborate on new exciting projects in the future.

This thesis would not have been possible without:

Jens Sörensen – my “fourth supervisor”. The Richard Gere of nuclear medicine. You combine a cultivated wisdom while radiating an awesome atmos-
phere of amazement and curiosity. You have the ability to answer any question with philosophically rich reasoning, and your generous enthusiasm is invaluable. You initiated the low-dose projects, and I am very proud and grateful to be able to include them in my thesis.

All the volunteers, both patients and healthy controls. You have boldly provided your brains, veins, and your time, and it is thanks to you and your generous participation that this research could be performed.

Torsten Danfors – my guide to the world of brain PET and neurodegenerative diagnostics. I enjoy and respect your teachings more than you know.

Johan Lilja – my squire and support, always willing to run that extra mile to help out, despite drowning in your own work. Thank you for providing excessive data crunching and necessary insights into the world of spatial normalization.

Sven Haller – the Deus ex Machina. Your involvement in this project has been of true anastrophical significance, and I hope we can continue our collaboration, which I appreciate enormously.

Käthe Ström – the solid rock of inclusion logistics. In the early phases of this project, I depended entirely on your rigorous and reliable handling, performed with humbleness and highly professional service-mindedness.

Mimmi Lidholm, Britt-Mari Bolinder and Eva Broström – for struggling with the ungrateful puzzle of finding scanner time for all the subjects. Thank you for your patience and stubbornness. Without the effort from the likes of you, no clinical research could be performed, and I hope you take pride in your involvement.

Dag-Lotta – the third parent to our children. Trusting the daily care of Harry and Björn to your attentive expertise allowed me to concentrate on my work to a higher extent.

I would also like to thank:

Ulf, Maria, Johannes, Niclas, Ruta, Vilma; everyone in my fellowship - for the laughs and the sense of coherence you brought.
Christina Lundberg for inspiring me in so many ways.
Johan Virhammar and Anna Falk, for setting a respectable standard regarding the thesis.
Johan Wikström – my clinical supervisor and new professor, for watching my back and supporting my research.
LG and Allina – for arranging so I could follow this through.
Håkan Pettersson – for all your help with computers, layout and figures. I hope you take pride in your contributions.
SAFT – for providing breaks from the common routine, as well as insights to our present and future world, set in constant change.
Kerstin Lundström – for awakening my interest in dementia, a long time ago. Without your gentle push, I might have ended up doing something far less interesting and meaningful.
Eva Kumlien, Anja Smits and the other colleagues at the Neurology department for introducing me to both clinical work and clinical research.
Robert Elster a.k.a. Vibrasphere. Much of the writing has been done with your fantastic music in my ears. You are a genius and I look forward to new releases.
The UNS group – for adding flavour to my time as a PhD student.
Sara Shams, Mana Shams, Peter Nyrén – I look forward to exploring the future of neuroradiology with you.
Christl Richter-From, the nestor of bureaucracy.
Markus Fahlström and Mats Rynnes – for showing up and fixing things. We need that!
Minnesmottagningen – for your kind reception and your help with the inclusion of patients
Arvid Morell – for your kind help when I was green, and for inspiring me to organize my files efficiently.
Andre Ahlgren – for your generous help with handling ASL.
Atle Björnerud – for helping me with Nice and kindly providing opportunities (to be continued).
Edwin Stanton Spencer, for providing excellent and quick proofreading.
Anja and Håkan – for your time and support.

My parents – for the perfect combination of love, freedom and constructive support you have given through the years.

Last but not least: Helena Fällmar, my precious wife and lifelong companion. Your endless love and support made this possible and enjoyable.

Thank you.
References


A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine”.)