



# LUND UNIVERSITY

## Artificial Intelligence-based Assessment of Prostate Cancer Metastases in PET/CT

Lindgren Belal, Sarah

2022

*Document Version:*

Publisher's PDF, also known as Version of record

[Link to publication](#)

*Citation for published version (APA):*

Lindgren Belal, S. (2022). *Artificial Intelligence-based Assessment of Prostate Cancer Metastases in PET/CT*. [Doctoral Thesis (compilation), Department of Translational Medicine]. Lund University, Faculty of Medicine.

*Total number of authors:*

1

### General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00



# Artificial Intelligence-based Assessment of Prostate Cancer Metastases in PET/CT

SARAH LINDGREN BELAL

DEPARTMENT OF TRANSLATIONAL MEDICINE | LUND UNIVERSITY





Artificial Intelligence-based Assessment of  
Prostate Cancer Metastases in PET/CT





# Artificial Intelligence-based Assessment of Prostate Cancer Metastases in PET/CT

Sarah Lindgren Belal



**LUND**  
UNIVERSITY

DOCTORAL DISSERTATION

by due permission of the Faculty of Medicine, Lund University, Sweden.

To be defended in Room 2005/2007,

Department of Medical Imaging and Physiology,

Carl Bertil Larurells gata 9, Skåne University Hospital, Malmö,

December 9<sup>th</sup> 2022 at 1.00 PM.

*Faculty opponent*

Professor Katrine Riklund

Umeå University, Sweden

<b>Organization</b> Department of Translational Medicine LUND UNIVERSITY		<b>Document name</b> Doctoral dissertation
Author Sarah Lindgren Belal		<b>Date of issue</b>
<b>Title and subtitle</b> Artificial Intelligence-based Assessment of Prostate Cancer Metastases in PET/CT		
<b>Abstract</b>		
<p><b>Background:</b> Quantification of tumor burden from bone scan in the form of automated bone scan index (aBSI) has been validated as an imaging biomarker for patients with prostate cancer. Positron emission tomography combined with computed tomography (PET/CT) is more sensitive and accurate compared to conventional imaging such as bone scan. The evaluation of medical images including PET/CT is challenged by time-consuming and subjective analysis, and issues with intra- and inter-reader agreement. An objective and quantitative method for analysis of PET/CT, similar to aBSI for bone scan, is an unmet need. Artificial intelligence (AI) has the potential to meet this need.</p> <p><b>Aim:</b> The aim of this thesis was to develop AI-based models for automated quantification of skeletal tumor burden, and for detection of pelvic lymph node lesions, in PET/CT scans from patients with prostate cancer.</p> <p><b>Methods:</b> In paper I, standardized uptake value (SUV)-based PET indices reflecting the whole-body tumor burden in [<sup>18</sup>F]fluoride PET/CT scan were calculated in a group of patients with prostate cancer, and compared to aBSI. For proof of concept, the association of PET index with overall survival was investigated. In paper II, convolutional neural networks (CNNs) were trained to segment the axial skeleton in CT scans. In paper III, CNNs were trained to segment bone lesions in [<sup>18</sup>F]fluoride PET for automated quantification of the skeletal tumor burden. The PET indices in paper I and III were also compared to a SUV threshold for lesion segmentation. In paper IV, a CNN was trained and tested to detect pelvic lymph node lesions in [<sup>18</sup>F]PSMA-1007 PET/CT scans. The performance of all models was evaluated against readers used as reference.</p> <p><b>Results:</b> Paper I showed that the SUV-based PET index was associated with overall survival. In paper II, the AI model for bone segmentation performed equally well as an experienced reader. Paper III showed that the performance of an AI-based model for segmentation of bone lesions and quantification of skeletal tumor burden was more similar to experienced readers than a global SUV threshold. The sensitivity for the AI model in paper IV was in level with that of experienced physicians. Inter-reader disagreement among experienced physicians was seen in paper III-IV.</p> <p><b>Conclusions:</b> AI-based models for automated assessment of PET/CT were developed. Tumor burden measured from PET/CT carries prognostic information in patients with prostate cancer. The difficulty in achieving high inter-reader agreement emphasizes the need for automated and objective scan interpretation. AI-assisted quantification of tumor burden holds potential as a future prognostic imaging biomarker for patients with prostate cancer. Increasing the amount and variety of training data should further enhance the performance of the proposed models.</p>		
<b>Key words</b>		
Classification system and/or index terms (if any)		
Supplementary bibliographical information		<b>Language</b> English
ISSN and key title 1652-8220		<b>ISBN</b> 978-91-8021-325-7
Recipient's notes	<b>Number of pages</b> 79	Price
	Security classification	

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature



Date 2022-11-04

# Artificial Intelligence-based Assessment of Prostate Cancer Metastases in PET/CT

Sarah Lindgren Belal



**LUND**  
UNIVERSITY

Cover art by Sarah Lindgren Belal

Copyright pp 1-79 Sarah Lindgren Belal

Paper 1 © 2017 The Authors. Published by SpringerOpen.

Paper 2 © 2019 The Authors. Published by Elsevier B.V.

Paper 3 © The Authors (Manuscript unpublished)

Paper 4 © 2022 The Authors. Published by Springer Nature.

Faculty of Medicine  
Department of Translational Medicine

ISBN 978-91-8021-325-7

ISSN 1652-8220

Printed in Sweden by Media-Tryck, Lund University  
Lund 2022



Media-Tryck is a Nordic Swan Ecolabel  
certified provider of printed material.  
Read more about our environmental  
work at [www.mediatryck.lu.se](http://www.mediatryck.lu.se)

**MADE IN SWEDEN** 



Till Elis (X:en)

# Content

List of Papers.....	9
Abbreviations and Terms.....	10
Preface.....	12
<b>Introduction.....</b>	<b>13</b>
Prostate Cancer .....	13
Metastatic Prostate Cancer .....	17
PET/CT Imaging.....	19
Imaging Biomarkers .....	23
Artificial Intelligence .....	25
Rationale for Quantitative PET/CT Assessment.....	29
<b>Aims.....</b>	<b>31</b>
<b>Methods.....</b>	<b>33</b>
Study Populations .....	33
Image Acquisition.....	36
AI Models .....	39
Statistical Analysis .....	44
Ethical Considerations.....	46
<b>Results and Comments .....</b>	<b>49</b>
Quantification of Skeletal Tumor Burden .....	49
Detection of Pelvic Lymph Nodes.....	54
Aspects of Ground Truth.....	56
Limitations of PET/CT Scanning.....	57
<b>Conclusions and Future Scope.....</b>	<b>59</b>
<b>Populärvetenskaplig sammanfattning .....</b>	<b>61</b>
<b>Errata .....</b>	<b>63</b>
<b>Acknowledgements .....</b>	<b>65</b>
<b>About the Author.....</b>	<b>67</b>
<b>References .....</b>	<b>68</b>

# List of Papers

This thesis is based on the following four papers:

- I. **3D skeletal uptake of  $^{18}\text{F}$  sodium fluoride in PET/CT images is associated with overall survival in patients with prostate cancer.**  
Lindgren Belal S, Sadik M, Kaboteh R, Hasani N, Enqvist O, Svärm L, Kahl F, Simonsen J, Poulsen MH, Ohlsson M, Høilund-Carlsen PH, Edenbrandt L, Trägårdh E. *EJNMMI Res.* 2017;7(1)15.
- II. **Deep learning for segmentation of 49 selected bones in CT scans: First step in automated PET/CT-based 3D quantification of skeletal metastases.**  
Lindgren Belal S, Sadik M, Kaboteh R, Enqvist O, Ulén J, Poulsen MH, Simonsen J, Høilund-Carlsen PF, Edenbrandt L, Trägårdh E. *Eur J Radiol.* 2019;113:89-95.
- III. **Automated quantification of PET/CT skeletal tumor burden in prostate cancer using artificial intelligence – the PET index.**  
Lindgren Belal S, Larsson M, Holm J, Middelbo Buch-Olsen M, Sörensen J, Bjartell A, Edenbrandt L, Trägårdh E. *Manuscript.*
- IV. **Freely available artificial intelligence for pelvic lymph node metastases in PSMA PET-CT that performs on par with nuclear medicine physicians.**  
Trägårdh E, Enqvist O, Ulén J, Hvittfeldt E, Garpered S, Lindgren Belal S, Bjartell A, Edenbrandt L. *Eur J Nucl Med Mol Imaging.* 2022;49(10):3412-18.

## Abbreviations and Terms

<b>aBSI</b>	Automated bone scan index
<b>ADT</b>	Androgen deprivation therapy
<b>AI</b>	Artificial intelligence
<b>BCR</b>	Biochemical recurrence
<b>Bone scan</b>	Whole-body bone scintigraphy using <sup>99m</sup> Tc-labeled diphosphonates
<b>BSI</b>	Bone scan index
<b>CI</b>	Confidence interval
<b>CNN</b>	Convolutional neural network
<b>CoV</b>	Coefficient of variation
<b>CT</b>	Computed tomography
<b>2D</b>	Two-dimensional
<b>3D</b>	Three-dimensional
<b>EANM</b>	European Association of Nuclear Medicine
<b>EAU</b>	European Association of Urology
<b>ePLND</b>	Extended pelvic lymph node dissection
<b><sup>18</sup>F</b>	Fluorine-18
<b>[<sup>18</sup>F]DCFPyL</b>	2-(3-{1-carboxy-5-[(6-[ <sup>18</sup> F]fluoro-pyridine-3-carbonyl)-amino]-pentyl}-ureido)-pentanedioic acid
<b>[<sup>18</sup>F]FDG</b>	2-[ <sup>18</sup> F]fluoro-2-deoxy-D-glucose
<b>[<sup>18</sup>F]fluoride</b>	[ <sup>18</sup> F]sodium fluoride
<b>FN</b>	False negative
<b>FoV</b>	Field of view
<b>FP</b>	False positive
<b><sup>68</sup>Ga</b>	Gallium-68
<b>ISUP</b>	International Society of Urological Pathology

<b>L3</b>	Third lumbar vertebra
<b>mCRPC</b>	Metastatic castration-resistant prostate cancer
<b>miTNM</b>	Molecular imaging tumor, node, metastasis classification of malignant tumors
<b>mpMRI</b>	Multiparametric magnetic resonance imaging
<b><sup>99m</sup>Tc</b>	Technetium-99m
<b>OS</b>	Overall survival
<b>OSEM</b>	Ordered subset expectation maximization
<b>PCWG</b>	Prostate Cancer Working Group
<b>PERCIST</b>	Positron emission tomography response criteria in solid tumors
<b>PET/CT</b>	Positron emission tomography with computed tomography
<b>PPP</b>	Positive predictive value
<b>PSA</b>	Prostate-specific antigen
<b>PSMA</b>	Prostate-specific membrane antigen
<b>RECIST</b>	Response evaluation criteria in solid tumors
<b>RECOMIA</b>	Research consortium for medical image analysis
<b>SD</b>	Standard deviation
<b>SDI</b>	Sørensen-Dice index
<b>SUV</b>	Standardized uptake value (measured activity normalized for volume of distribution [body weight/surface area] and injected radiopharmaceutical dose)
<b>Th7</b>	Seventh thoracic vertebra
<b>TNM</b>	Tumor, node, metastasis classification of malignant tumors
<b>TP</b>	True positive
<b>TLU</b>	Total lesion uptake ( $TLV \times SUV_{\text{mean}}$ of the TLV in PET)
<b>TLV</b>	Total lesion volume (the sum of all voxels that represent tumor in PET)
<b>UICC</b>	Union for International Cancer Control



## Preface

Over the course of this PhD study, the field of prostate cancer imaging has rapidly evolved. At the time when study I was initiated, there was increasing evidence in support of the high sensitivity of [<sup>18</sup>F]fluoride PET/CT. As it was speculated to eventually replace bone scan as the first-line modality for the depiction of bone metastases in prostate cancer patients, [<sup>18</sup>F]fluoride PET/CT is the predominantly studied radiopharmaceutical in this thesis. Since then, PSMA-targeting radiopharmaceuticals have emerged as the most promising for prostate cancer imaging. At Skåne University Hospital, [<sup>18</sup>F]PSMA-1007 PET/CT has been available for clinical use since 2019, enabling the inclusion of this scan type in paper IV.

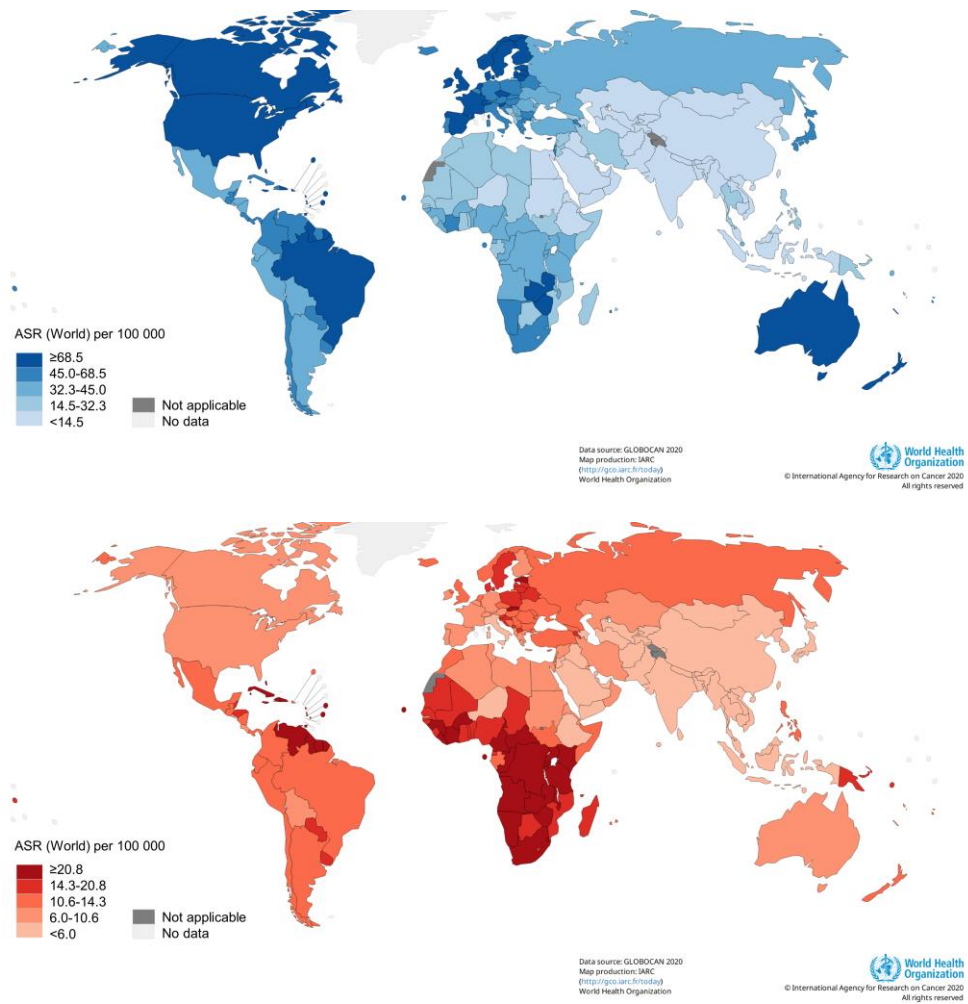
# Introduction

## Prostate Cancer

Prostate cancer is the second most common cancer and the fifth leading cause of cancer death in men worldwide, with approximately 1.4 million new cases and 375,000 deaths annually.<sup>1</sup> Australia/New Zealand, North America, and Europe have the highest incidence rates, while mortality rates are highest in Southern Africa, the Nordic countries, and parts of South America (Fig. 1). In Europe, an estimated 473,000 new cases and 108,000 deaths were diagnosed in 2020. The cumulative risk of being diagnosed with prostate cancer is 16% over a lifetime, while the risk of prostate cancer death before the age of 75 is 1% in Europe. A rapid increase in prostate cancer detection rates was seen in several high-income countries during the early to mid-1990s, with the introduction of prostate-specific antigen (PSA) testing.<sup>2</sup> After increasing for several years, prostate cancer incidence trends have been stabilizing or decreasing during the last 5-10 years. This is likely due in part to caution of PSA testing, given the possibility of unnecessary diagnosis of clinically insignificant disease.<sup>3</sup> Contrarily, mortality rates have decreased in several European countries, likely reflecting the advancements in treatment.<sup>2,3</sup>

## Histopathological Grading

Prostate cancer is commonly suspected based on an elevated serum PSA level, and/or an abnormal digital rectal examination. The final diagnosis depends on histopathological examination of prostate tissue biopsies. Grading of prostatic adenocarcinoma using the International Society of Urological Pathology (ISUP) 2014 modified Gleason system,<sup>4,5</sup> in which growth pattern and degree of differentiation of the tumor cells are assessed and assigned a score, is a prognostic factor for clinical behavior and treatment response (Table 1).<sup>6,7</sup> The ISUP grade is calculated as the sum of the most prevalent, plus the highest Gleason grade among the remaining tumor areas, in the biopsy sample. The resulting value ranges from 1-5, and with an increasing grade, the risk of prostate cancer-related mortality increases.<sup>8</sup>



**Fig. 1.** Estimated age-standardized incidence rates (top, blue) and mortality rates (bottom, red) of prostate cancer in the world in 2020. Reprinted with permission.<sup>9</sup>

**Table 1.** International Society of Urological Pathology 2014 grading of prostate cancer.<sup>4</sup>

ISUP Grade Group	Gleason Score
1	≤6
2	7 (3+4)
3	7 (4+3)
4	8
5	9-10

## Staging

Clinical and pathological staging are approaches to evaluate the prognostic characteristics of prostate cancer. Clinical tumor (cT) staging is used at initial diagnosis based on the findings during digital rectal examination and results from prostate cancer biopsies, while pathological tumor (pT) staging is evaluated in surgically removed tissues (Table 2). The 2022 European Association of Urology (EAU) guidelines on prostate cancer<sup>10</sup> recommend the use of multiparametric magnetic resonance imaging (mpMRI), which provides information on tissue density, microvascular blood flow and vessel permeability, prior to prostate biopsy to reduce overdiagnosis and allow targeted biopsies of detected lesions. The value of mpMRI for staging primary prostate cancer is still to be established. A meta-analysis showed a sensitivity of only 61% (95% CI [confidence interval] 54-67%), and a specificity of 88% (95% CI 85-91%) for T3 tumors (see below for T3 definition).<sup>11</sup> However, MRI can provide valuable information for planning local therapy.<sup>10</sup>

Prostate cancer is staged according to the eighth edition of the Union for International Cancer Control (UICC) tumor, node, metastasis classification of malignant tumors (TNM).<sup>12</sup> Regional lymph node metastases (N1) are defined as spread to the nodes of the true pelvis, located below the bifurcation of the common iliac arteries. They include pelvic, hypogastric, obturator, iliac, and sacral nodes. Distant metastases (M1) include lymph nodes outside the true pelvis including paraaortic lumbar, common iliac, superficial and deep inguinal and retroperitoneal lymph nodes, as well as bone and visceral metastases.

The European Association of Nuclear Medicine (EANM) recently proposed a molecular imaging TNM (miTNM), taking into account findings on prostate-specific membrane antigen positron emission tomography with computed tomography (PSMA PET/CT, see under 'PSMA PET/CT').<sup>13</sup> Unlike the UICC TNM classification, miTNM considers common iliac lymph nodes regional lymph nodes (miN1). As such, the N1/M1a definition is a matter of debate and is expected to be addressed in the coming ninth edition of the UICC TNM classification.<sup>14</sup>

**Table 2.** The Union for International Cancer Control tumor, node, metastasis (TNM) classification of prostate cancer.<sup>12</sup>

<b>Clinical Tumor (cT) Stage</b>	
<b>TX</b>	Primary tumor cannot be assessed
<b>T0</b>	No evidence of primary tumor
<b>T1</b>	Clinically inapparent tumor that is not palpable
<b>T1a</b>	Tumor incidental histological finding in ≤5% of tissue resected
<b>T1b</b>	Tumor incidental histological finding in >5% of tissue resected
<b>T1c</b>	Tumor identified by needle biopsy found in one or both sides, but not palpable
<b>T2</b>	Tumor is palpable and confined within the prostate
<b>T2a</b>	Tumor involves one-half of one lobe or less
<b>T2b</b>	Tumor involves more than one-half of one lobe, but not both lobes
<b>T2c</b>	Tumor involves both lobes
<b>T3</b>	Extracapsular tumor that is not fixed and does not invade adjacent structures
<b>T3a</b>	Extracapsular extension (unilateral or bilateral)
<b>T3b</b>	Tumor invades seminal vesicle(s)
<b>T4</b>	Tumor is fixed or invades adjacent structures other than seminal vesicles such as external sphincter, rectum, bladder, levator muscles, and/or pelvic wall
<b>Pathological Tumor (pT) Stage</b>	
<b>T2</b>	Organ confined
<b>T3</b>	Extraprostatic extension
<b>T3a</b>	Extraprostatic extension (unilateral or bilateral) or microscopic invasion of bladder neck
<b>T3b</b>	Tumor invades seminal vesicle(s)
<b>T4</b>	Tumor is fixed or invades adjacent structures other than seminal vesicles such as external sphincter, rectum, bladder, levator muscles, and/or pelvic wall
<b>Regional Lymph Node (N) Stage</b>	
<b>NX</b>	Regional lymph nodes not assessed
<b>N0</b>	No regional lymph node metastasis
<b>N1</b>	Regional lymph node metastasis
<b>Distant Metastasis (M) Stage</b>	
<b>M0</b>	No distant metastasis
<b>M1</b>	Distant metastasis
<b>M1a</b>	Non-regional lymph node(s)
<b>M1b</b>	Bone(s)
<b>M1c</b>	Other site(s) with or without bone metastasis

NOTE: There is no pathological T1 classification.

## Risk Stratification

Newly diagnosed patients with prostate cancer without known metastases are stratified according to their risk of biochemical recurrence (BCR) following local treatment with curative intent. According to the EAU guidelines,<sup>10</sup> patients are categorized into low, intermediate-, and high-risk groups depending on PSA levels, ISUP grade (Gleason score), and cT stage (Table 3). A higher risk group is associated with a more aggressive cancer and poorer prognosis.<sup>6, 8</sup>



**Table 3.** European Association of Urology risk groups for biochemical recurrence of localized and locally advanced prostate cancer.<sup>10</sup>

Low-risk	Intermediate-risk	High-risk	
PSA <10 ng/mL and ISUP grade 1 and cT1-2a	PSA 10-20 ng/mL or ISUP grade 2/3 or cT2b	PSA >20 ng/mL or ISUP grade 4/5 or cT2c	Any PSA  Any ISUP grade  cT3-4 or cN+
<b>Localized</b>			<b>Locally advanced</b>

ISUP: International Society of Urological Pathology; PSA: Prostate-specific antigen.

## Metastatic Prostate Cancer

### Lymph Node Metastases

The detection of regional lymph node metastases (N1) is important for prognosis and follow-up in patients with newly diagnosed prostate cancer, as they are associated with distant metastases and negatively correlated with survival.<sup>15</sup> The gold standard for diagnosing pelvic lymph node spread is through histopathological analysis of tissue obtained from an extended pelvic lymph node dissection (ePLND) and is recommended in high-risk localized and locally advanced disease.<sup>10</sup> It is an invasive procedure, with up to 20% of patients experiencing postoperative complications, with lymphocele being the most common adverse event.<sup>16</sup> Different nomograms have been developed and validated based on systemic random biopsies,<sup>17-20</sup> and more recently based on MRI-guided biopsies,<sup>21, 22</sup> that help calculate the individual risk of lymph node metastases in a patient prior to treatment. However, these nomograms do not identify the presence of lymph node metastases and cannot replace diagnostic modalities.

Non-invasive diagnostic approaches for diagnosing lymph node metastases are problematic. Conventional imaging with CT or MRI evaluates lymph nodes indirectly by morphologic characteristics such as shape and size. A lymph node short axis of >8 mm in the pelvis and >10 mm outside the pelvis are suspicious of malignancy. In prostate cancer, almost 80% of metastatic lymph nodes are too small to be detected using conventional imaging.<sup>23, 24</sup> The performance for assessment of nodal metastasis is therefore poor, with a pooled sensitivity of 42% (95% CI 26-56%) for CT, and 39% (95% CI 22-56%) for MRI.<sup>24</sup>

## Bone Metastases

In most cases, newly diagnosed prostate cancer are low-risk indolent tumors that may exist for a long time without causing symptoms or death. Some patients present with aggressive disease that grows rapidly and metastasizes to other parts of the body. Advanced prostate cancer results from any combination of hematogenous, lymphatic or contiguous local spread.

Identification of metastatic disease is essential for patient management, as distant metastases imply that the disease is beyond cure. Bone metastases are also a significant cause of morbidity due to skeletal-related events, including bone pain, pathologic fractures, and spinal cord compression.<sup>25</sup> Treatment with curative intent is only indicated in patients with localized or locally advanced disease, while patients with distant metastases are offered palliative systemic treatment with androgen deprivation therapy (ADT), with or without docetaxel or next-generation androgen receptor-targeted therapy.<sup>26-32</sup> With the increasing evidence that patients with low metastatic burden at diagnosis may benefit from a combination of systemic and local therapy, the demand for precise detection and localization of cancer spread is even higher.<sup>33-35</sup>

Bone is the most common site of distant metastases in advanced prostate cancer.<sup>36</sup> Approximately 13% of patients with newly-diagnosed disease present with bone metastases.<sup>37</sup> In metastatic castration-resistant prostate cancer (mCRPC), which is the lethal form of the disease that eventually will develop during ADT, bone metastases occur in up to 84% of patients.<sup>36, 38</sup> Spread of disease to the bone is associated with poor prognosis, with five-year survival rates dropping to 35% or less for patients with distant metastases at diagnosis, compared to >99% for localized disease.<sup>38</sup> Prostate cancer spreads to the bone via the hematogenous route, whereby cancer cells first settle in the bone marrow. Because the red marrow in adults is located in cancellous bone, the axial skeleton is the primary site of skeletal metastases.<sup>39</sup> There is a predominant upregulation of osteoblastic activity in bone metastases from prostate cancer, which causes the formation of sclerotic, mineralized woven bone seen with conventional imaging. However, metastatic processes are present in the bone marrow prior to the derangement of bone architecture.

Whole-body bone scintigraphy using technetium-99m (<sup>99m</sup>Tc)-labeled diphosphonates (bone scan) is the most widely used modality for bone imaging in prostate cancer both in primary and recurrent disease.<sup>10</sup> Suspicious metastases are visualized indirectly as increased uptake of diphosphonates due to upregulated bone turnover. However, increased bone turnover is not specific to malignancy but occurs in other conditions, such as degenerative diseases and fractures in healing. Bone scan lacks detailed anatomical information and cannot reveal lymph node or visceral metastases. On a

patient-level, the sensitivity is 79% (95% CI 73-83%), and the specificity 82% (95% CI 78-85%), for diagnosing bone metastases.<sup>40</sup> If combined with single-photon emission computed tomography and CT, the diagnostic accuracy increases.<sup>41</sup>

Of patients undergoing initial treatment with curative intent, between 20-30% will develop BCR,<sup>42, 43</sup> and of these, up to 70% will develop bone metastases within five years.<sup>44, 45</sup> Accurate diagnosis of the site(s) and extent of BCR influences further treatment planning and can be used in tailoring potential salvage treatments. Conventional imaging modalities also have limitations in this regard, especially in cases with low PSA levels.<sup>46, 47</sup>

Conventional imaging modalities, including bone scan, only carry indirect signs of existing active tumor cells. In the case of bone metastases, the depicted changes may be delayed concerning present tumor cell activity and may persist after the eradication of metastatic cells. To improve the detection of lymph node and bone metastases in both primary and recurrent prostate cancer, more modern modalities that provide quantitative functional assessment have been introduced and represent a major advance in prostate cancer imaging.<sup>48, 49</sup>

## PET/CT Imaging

PET is an imaging modality that uses radioactive isotopes that decay with positron emission. The emitted positron travels through the surrounding tissue until it reacts with an electron. The positively and negatively charged elements annihilate and produce a pair of 511 keV gamma photons traveling in opposite directions and can be detected by a PET scanner. To synthesize a diagnostic PET radiopharmaceutical, an isotope, such as fluorine-18 (<sup>18</sup>F), is bound to a biological compound. The resulting radiopharmaceutical is injected intravenously into patients and can identify functional processes within the body on a biochemical level. PET provides quantitative information in the form of standardized uptake values (SUVs), which is the measured activity normalized for volume of distribution (body weight/surface area) and injected radiopharmaceutical dose.<sup>50</sup> Combined with CT, an image of the functional information from the PET scan and the structural information from the CT scan can be obtained (PET/CT).

The advancement of PET from an instrument of research in the 1970s, the commercialization of the first PET/CT for medical use in 2001, to its present-day wide and increasing use in oncologic, cardiac, and neurological imaging has contributed instrumentation that is making a significant clinical impact. The limitations of

structural imaging in early detection and response to therapy in most diseases have become more apparent in the context of molecular imaging. As the future of healthcare lies in precision medicine and personalized care, molecular imaging with PET/CT aligns seamlessly with these concepts.

### **[<sup>18</sup>F]FDG PET/CT**

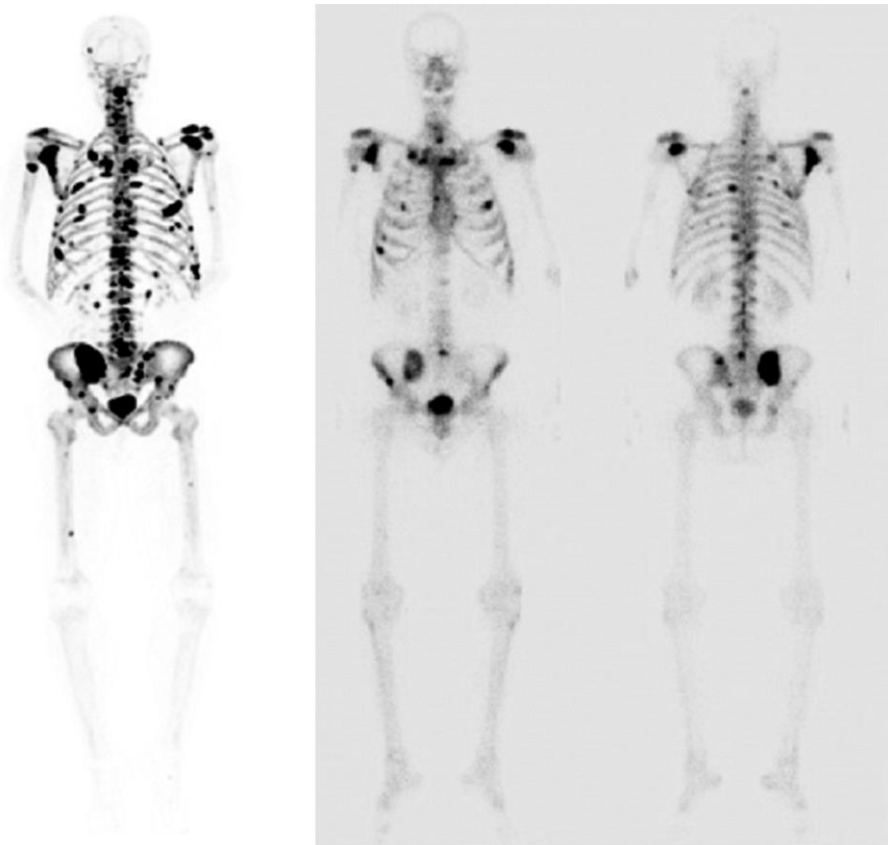
The first and still most commonly used PET tracer in oncologic imaging is 2-<sup>[18F]fluoro-2-deoxy-D-glucose</sup> (<sup>[18F]FDG</sup>).<sup>50</sup> <sup>[18F]FDG</sup> is a radioglucose analog that is trapped by phosphorylation and accumulates in tumor cells as a sign of increased metabolic activity. Despite the widespread utility in most cancers, <sup>[18F]FDG</sup> PET/CT does not play an important role in prostate cancer imaging due to its limited sensitivity in localized and early metastatic disease.<sup>51-53</sup> The sensitivity for detection of prostate cancer bone metastases is lower compared to bone scan.<sup>54</sup> <sup>[18F]FDG</sup> PET/CT may be useful in specific disease scenarios with increased glycolytic activity, such as aberrant histology (i.e., neuroendocrine) or very aggressive, poorly differentiated or undifferentiated prostate cancer.<sup>55, 56</sup> In those scenarios, <sup>[18F]FDG</sup> PET/CT may have a role in monitoring response to treatment in conjunction with bone scans. A recent randomized phase II trial comparing PSMA-targeted radionuclide therapy to cabazitaxel in patients with progressive mCRPC also used <sup>[18F]FDG</sup> PET/CT to ensure that only patients with disease concordant across <sup>[18F]FDG</sup> and PSMA PET/CT scans were recruited. The rationale was that <sup>[18F]FDG</sup>-avid areas with low PSMA expression are unlikely to benefit from therapy highly targeted towards PSMA.<sup>57</sup>

### **[<sup>18</sup>F]fluoride PET/CT**

<sup>[18F]sodium fluoride</sup> (<sup>[18F]fluoride</sup>) was introduced in the 1960s and is one of the oldest radiopharmaceuticals. It is a bone-seeking compound and depicts sites of increased bone turnover similar to <sup>99mTc</sup>-labeled diphosphonates used for bone scan. The then-available gamma cameras initially restricted the clinical use of <sup>[18F]fluoride</sup> as a bone imaging agent. In the 1970s, it was largely replaced by <sup>99mTc</sup>-labeled compounds because of their favorable physical characteristic, such as longer half-life and lower photon energy, and the widespread availability of <sup>99mTc</sup> generators.<sup>58, 59</sup> With the introduction of PET technology, it became possible to obtain high-resolution and high-contrast imaging using <sup>[18F]fluoride</sup>, and its use has been increasing during the last decades.<sup>59</sup>

Compared to <sup>99mTc</sup>-labeled diphosphonates, <sup>[18F]fluoride</sup> has higher bone uptake, faster blood clearance, and a higher target-to-background ratio. The low protein

binding and decreased background uptake allow [ $^{18}\text{F}$ ]fluoride PET/CT scanning to be performed 1 hr after intravenous radiopharmaceutical injection, compared to  $^{99\text{m}}\text{Tc}$ -labeled diphosphonate scanning, which is typically performed 3-4 hrs after administration.<sup>50</sup> In comparison to bone scan, [ $^{18}\text{F}$ ]fluoride provides superior image quality and has a higher overall sensitivity in lesion detection (Fig. 2), which is especially helpful for detecting lesions in the spine.<sup>58</sup> These properties, combined with increasing access to PET/CT, have led to the replacement of bone scan with [ $^{18}\text{F}$ ]fluoride PET/CT in several institutions. However, it remains to be seen how small findings on PET/CT that raise the suspicion of bone metastases should affect patient management. A prospective study showed no added value of [ $^{18}\text{F}$ ]fluoride PET/CT over bone scan for initial staging in patients with newly diagnosed intermediate- or high-risk prostate cancer and a negative bone scan.<sup>60</sup>



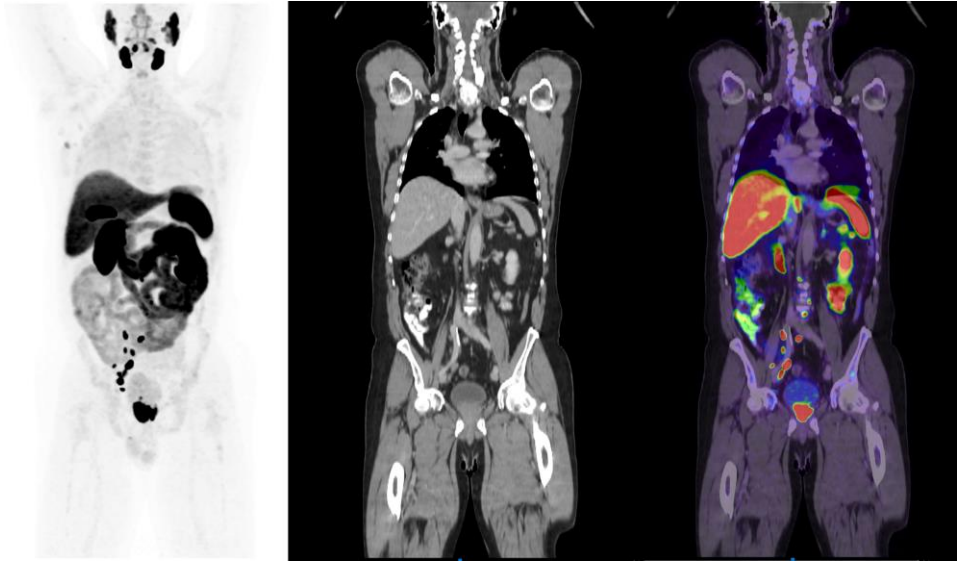
**Fig. 2.** Comparison of [ $^{18}\text{F}$ ]fluoride PET versus bone scan in a patient with metastatic prostate cancer. PET demonstrates more focal lesions than planar imaging. Left-right: PET maximum intensity projection, bone scan anterior projection, bone scan posterior projection. Reprinted with permission.<sup>39</sup>

## PSMA PET/CT

PSMA, or folate hydrolase 1, is a transmembrane glycoprotein, and its expression and localization in the prostate are associated with the cytoplasm and apical side of the epithelium surrounding the prostatic ducts.<sup>61</sup> PSMA is strongly overexpressed, approximately 100-1,000-fold, on almost all prostate cancer cells compared to normal prostatic tissues.<sup>62</sup> However, 5–10% of primary prostate cancer lesions are shown to be PSMA-negative.<sup>63</sup> PSMA expression correlates with tumor aggressiveness and increases with higher tumor grade and stage.<sup>64-66</sup> Contrarily to its name, high physiological uptake of PSMA radiopharmaceuticals is also seen in ganglia, salivary and lacrimal glands, liver, kidneys, spleen, and small intestines (Fig. 3).<sup>65-68</sup> Increased PSMA expression is also seen in other malignancies such as renal cell, urothelial, and colon carcinomas.<sup>65</sup>

PSMA ligands bind to the active site in the extracellular domain of PSMA and are internalized and endosomally recycled, leading to enhanced tumor uptake and retention and subsequent high image quality.<sup>69</sup> Both gallium-68- (<sup>68</sup>Ga) and <sup>18</sup>F-labeled PSMA radiopharmaceuticals are available for PET imaging. <sup>68</sup>Ga-labeled PSMA-targeting compounds are the most widely studied, but <sup>18</sup>F-labeled agents such as [<sup>18</sup>F]PSMA-1007 and 2-(3-{1-carboxy-5-[(6-[<sup>18</sup>F]fluoro-pyridine-3-carbonyl)-amino]-pentyl}-ureido)-pentanedioic acid ([<sup>18</sup>F]DCFPyL) have an average lower positron energy due to a shorter positron range and higher positron yield, leading to higher inherent spatial resolution. Additionally, <sup>18</sup>F is a cyclotron product with a longer half-life, which allows large-scale production and transportation.<sup>50,69</sup> Currently, there are no conclusive data about the comparison of <sup>68</sup>Ga- and <sup>18</sup>F-labeled PSMA radiopharmaceuticals.

PSMA PET/CT is currently the most sensitive imaging modality for diagnosing metastases at any stage of prostate cancer.<sup>70-72</sup> There is growing evidence underscoring the role of PSMA PET/CT imaging in BCR, especially in patients with PSA levels < 0.5 ng/mL.<sup>72,73</sup> It is the recommended modality for imaging in patients with rising PSA after initial treatment,<sup>10,74</sup> and in patients with a persistent PSA after local treatment.<sup>10</sup> PSMA PET/CT has also shown to be more accurate than CT, MRI, and/or bone scan for the staging of primary prostate cancer, particularly in high-risk disease.<sup>40, 63, 71, 72, 75</sup> It is recognized that PSMA PET/CT bears potential in evaluating treatment effect.<sup>76,77</sup> Even so, its role remains unclear as outcome data from randomized clinical trials evaluating the management and outcome based on findings on PSMA PET/CT is currently lacking but awaited.<sup>10</sup>



**Fig. 3.** [ $^{18}\text{F}$ ]PSMA-1007 PET/CT scan in a patient at initial staging showing locally advanced prostate cancer (T3b) with regional (N1) and distant (M1a) lymph node metastases. Physiological radiopharmaceutical uptake is seen in the salivary glands, kidneys, spleen, liver, gallbladder and intestines. Left-right: PET maximum intensity projection, coronal CT scan, and fused PET/CT scans.

Other PET radiopharmaceuticals are available for prostate cancer imaging, such as radiolabeled fluciclovine and choline. However, they are not within the scope of this thesis since they have not been proven superior to PSMA-targeting radiopharmaceuticals,<sup>78, 79</sup> and will therefore not be discussed further.

## Imaging Biomarkers

The localization and extent of metastatic disease is crucial for personalized and nuanced decisions to optimize clinical management strategy. Serum PSA levels do not provide accurate information about the size and number of tumor lesions, limiting its use for response assessment during treatment.<sup>80</sup> Due to the lack of serum biomarkers to independently and accurately evaluate treatment effect, monitoring metastatic spread relies on imaging. Quantitative analysis of tumor characteristics based on medical imaging can provide information to improve prognostication and prediction, and assessment and monitoring of response to therapy, through the development of imaging biomarkers.

A quantitative imaging biomarker can be defined as an objectively measured characteristic derived from an image as an indicator of normal biological processes, pathogenic processes, or response to a therapeutic intervention.<sup>81</sup> The standard imaging biomarker for assessment of tumor burden and response with conventional imaging is the Response evaluation criteria in solid tumors (RECIST) version 1.1.<sup>82</sup> In patients with prostate cancer, they are far from sufficient as osteoblastic lesions are deemed non-measurable.

Imaging biomarkers can also be semi-quantitative, for example, by using a scoring scale. The Prostate Cancer Working Group (PCWG) criteria have been proposed in an effort to make the interpretation of bone scan more standardized for clinical trials.<sup>83</sup> The most recently updated PCWG3 criteria define bone progression as  $\geq 2$  lesions on the first post-treatment scan, with  $\geq 2$  additional lesions on the next scan, and for scans after the first post-treatment scan, at least  $\geq 2$  new lesions relative to the first post-treatment scan confirmed on a subsequent scan. Thus, the PCWG3 criteria do not take into account changes in the size of lesions as indicators of an increase in total disease burden during treatment and are prone to subjective interpretation.

An objective and consistent methodology for quantifying the total skeletal tumor burden in bone scan was lacking until the development of the bone scan index (BSI).<sup>84</sup> BSI is a quantitative description of the total disease burden in the bone and represents the tumor burden expressed as a percentage of the total skeletal mass judged from a bone scan. When initially presented, the process of calculating the BSI was a manual task. With the development of a software for automated calculation of BSI, the acquisition time was reduced to seconds.<sup>85</sup> The automated BSI (aBSI) has been analytically validated as a reproducible and accurate measure of disease burden,<sup>86, 87</sup> and to be an independent prognostic imaging biomarker in patients with mCRPC in a large phase III study.<sup>88</sup> However, a major limitation for aBSI, and for all the above-mentioned methods, is the lack of sensitivity and specificity of the conventional imaging modalities to which they are applied.

### **PET/CT Imaging Biomarkers**

Similarly to RECIST 1.1 for conventional imaging, the PET response criteria in solid tumors (PERCIST) were developed for a more standardized assessment of tumor response in PET/CT.<sup>89</sup> Therapeutic response is categorized as complete response, partial response, stable disease, or progressive disease based on SUV measurement of the lesions with the highest radiopharmaceutical uptake and up to five lesions in total. This requires visual examination and manual placement of a region of interest by a reader, which is suboptimal for objective evaluation. Assessing changes in individual



lesions may also be inadequate for determining the overall response. Further, PERCIST has only been evaluated for [<sup>18</sup>F]FDG PET/CT, which has limited value for the assessment of prostate cancer.

With the growing importance of PSMA PET/CT in prostate cancer, guidelines have been developed with the aim of standardizing the reporting and harmonizing diagnostic interpretation criteria.<sup>13, 90, 91</sup> However, all encompass manual steps to various extents and therefore do not eliminate the central issue of reader subjectivity.

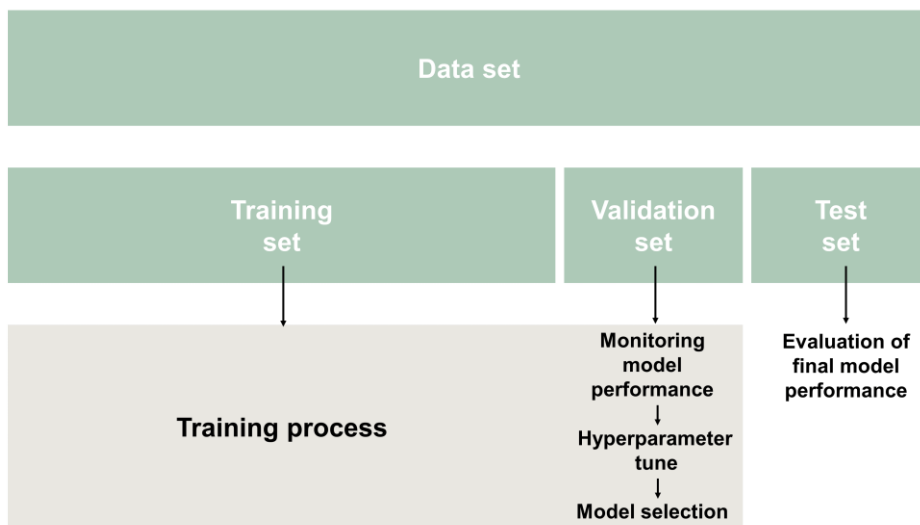
Inspired by the concept of aBSI, different methods have been used to quantify the skeletal burden and treatment response in both [<sup>18</sup>F]fluoride PET/CT<sup>92-95</sup> and PSMA PET/CT.<sup>96, 97</sup> Further, methods that also include extra-skeletal uptake suspicious of metastatic disease in PSMA PET/CT have also been proposed.<sup>98-102</sup> In small patient materials, these measures have shown a correlation to survival and promising results in predicting treatment response to systemic treatment.<sup>93, 101, 103-108</sup> Despite their computational approach, a common limitation is the built-in requirement of manual interference. This makes the assessment of changes in response to treatment subjective and time-consuming, especially in patients with widespread metastatic disease in multiple localizations, thereby restricting their potential for clinical use. Recent advancements include deep learning methods for fully automated assessment of tumor burden in PSMA PET/CT,<sup>109-111</sup> pointing to the potential of artificial intelligence (AI) to assist in this development.

## Artificial Intelligence

AI has been witnessing monumental growth in bridging the gap between machine and human capabilities. Essentially, AI is the simulation of human intelligence processes by machines, especially computer systems.<sup>112</sup> One subfield of AI is machine learning, where data-driven algorithms learn by prior examples or experience without being explicitly programmed. In machine learning, data is used to train a model on that specific information, and the trained model is subsequently used to make predictions on a new set of data. The training is an iterative process where the model is exposed to new, unfamiliar data step by step. At each step, the model makes predictions and gets feedback regarding how accurate those predictions are. The feedback is then used to correct the previous prediction errors. This predict-and-adjust process continues until the predictions of the model no longer improve.

Machine learning algorithms require data, such as medical images, to be trained on. Available data are typically split into a training, a validation (tuning), and a test set,

though variations occur (Fig. 4). A training set is used to train a network, while a validation set is used to evaluate the model under the training process. A test set is used once at the end to evaluate the performance of the final model that was fine-tuned and selected during the training process. It is important that the validation and test sets are kept separate, as some information about the validation data always leaks into the model itself during training, risking so-called overfitting. A separate, unseen test set is therefore needed to evaluate the true performance of the model and its generalizability.<sup>113</sup>

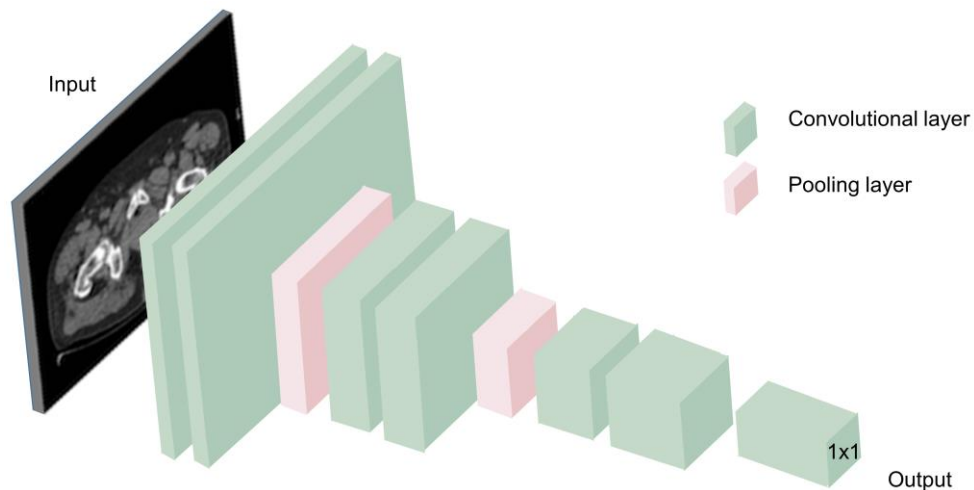


**Fig. 4.** Data is typically split into a training, a validation, and a test set. A training set is used to train a network, while a validation set is used for monitoring the model performance, fine-tuning hyperparameters, and model selection during the training process. A test set is used once at the very end to evaluate the performance of the final model that is fine-tuned and selected during the training process on the training and validation sets.<sup>113</sup>

## A Brief Introduction to Convolutional Neural Networks

Convolutional neural networks (CNNs) are deep learning models that have become dominant in various computer vision tasks. Convolution is a mathematical operation that essentially allows the network to learn features in the input image by overlapping and combining restricted portions of it. The network is thereby able to identify patterns even when they are not clearly stated in the training dataset. The most important building blocks of a CNN are convolutional layers, pooling layers, and activation functions. The term deep learning refers to the number of layers in a network, although a definition of how many layers constitute ‘deep’ is lacking. A common architecture for

a CNN model used for image classification consists of convolution and pooling layers stacked one after the other (Fig. 5).<sup>114</sup>



**Fig 5.** An overview of the architecture of a convolutional neural network (CNN) used for image classification.

Convolution is a mathematical operation in itself. One important aspect is that the same filter/kernel is applied to all parts of the input image by shifting its position. The output of the filters in the convolutional layer creates feature maps. Activation functions are added after each convolutional layer to apply a nonlinear transformation. This enables the network to learn nonlinear functions and recognize complex patterns in the data. The pooling layers are used to reduce the dimensions of the feature maps and the learnable parameters, hence reducing the amount of computing performed in the network. This is done by aggregating the features present in a region of the feature map so that further operations are not performed on precisely positioned but on aggregated features generated by the convolution layer. This makes the model more robust to variations in the position of features in the input image, which avoids overfitting. The resulting features feed into a final layer that drives the decision, such as tumor or not, or predicts the best label to describe an image, for example, vertebra or rib.<sup>114</sup>

In medical imaging, CNN models are commonly trained to accomplish three main tasks; classification, detection, and segmentation. The designation of one of these labels to a task is, however, not always clear-cut. Before the start of training, data preparation is an important step. The images are often annotated or labeled by a reader, which is a time-consuming task. Depending on the images and the task, different annotations are

required. Segmentation of an organ or an abnormal feature requires pixel-wise segmentation, which is laborious when performed manually and especially in three-dimensional (3D) images such as PET/CT. This may limit the use of a large data set. Classification of chest radiographs as normal or not only requires image labeling, which may allow for larger training material.<sup>115</sup>

### *Classification*

In classification, an image is assigned to any of predetermined classes. Different tasks can include the classification of the presence or absence of an abnormality in an image, the classification of abnormalities as benign or malignant, or the classification of malignant lesions according to their features, such as histopathological grade.<sup>116</sup>

### *Detection*

Detection refers to the task of recognizing and approximately localizing an object such as an organ, or an abnormality, such as a tumor or fracture, with, for example, a point or a bounding box. The object is given a label to help determine its exact location and orientation in relation to other structures.<sup>114, 116</sup>

### *Segmentation*

Segmentation is the process of dividing an image into different parts or regions according to its features and properties. In medical images, segmentation of organs or lesions is a common application, often as a pre-processing step for feature extraction and classification.<sup>116</sup> It can also help in the quantitative assessment of lesions or other abnormalities.<sup>114</sup> A CNN model performs segmentation by processing the whole or large parts of the image at the same time and produces an output map of labels for all pixels instead of individually labeling each pixel like previous methods. An example is the 3D U-Net, which has been very successful in solving many different segmentation tasks.<sup>117</sup> For segmentation of 3D data, 2D slices are often processed and then combined into a 3D map to reduce the amount of computational data.

## Rationale for Quantitative PET/CT Assessment

The interpretation of PET/CT scans is performed by a reader through visual examination, and findings are typically conveyed in a written clinical report. Despite criteria to help standardize assessment and reporting in research settings, intra- and inter-reader disagreement remains an issue even among experienced readers in single-center studies.<sup>118, 119</sup> There is an unmet need for an objective method to analyze PET/CT scans to increase reproducibility. The number of PET/CT examinations is expected to largely increase, especially for PSMA PET/CT, and there is a deficit of radiologists and nuclear medicine specialists to interpret these investigations. AI could help with the increasing volumes by speeding up the interpretation time and reduce intra- and inter-reader disagreement by providing standardized reporting and decision support to physicians.

The current subjective approach to scan evaluation does not allow full extraction of the information in a PET/CT scan, which inherently is a quantitative modality. Quantification of tumor burden from bone scan has been shown to correlate with overall survival (OS) in patients with advanced prostate cancer.<sup>88</sup> AI has the potential to help in the development of automated measurements reflecting the tumor burden in PET/CT which, if validated as imaging biomarkers, could be used for risk stratification and response prediction and as surrogate endpoints in clinical trials.

It has also been shown that tumor burden in PSMA PET/CT, and the corresponding therapeutic radionuclide dose received by the tumor, correlate with the treatment response of PSMA-directed radionuclide therapy.<sup>120</sup> Additionally, dose planning in a disease extent adapted fashion requires calculation of tumor volume. PSMA-directed radionuclide therapy is typically offered to patients with multiple metastases where it is impractical to manually segment all lesions. Therefore, there is a need for fast, automated, objective, and reproducible image analysis. AI has the potential to fill this need.



# Aims

The overall aim of this thesis was to develop an AI-based model for automated detection and quantification of PET/CT tumor burden in patients with prostate cancer. In a first step, an AI model for quantification of [<sup>18</sup>F]fluoride PET/CT skeletal tumor burden was developed (paper I-III). In a second step, an AI model for the automated detection of pelvic lymph node lesions was developed (paper IV). The specific aims of each paper are listed below.

## **Paper I**

To develop a PET index reflecting the skeletal tumor burden, to compare this PET index to aBSI, and to investigate the association of the PET index with OS. A secondary aim was to compare the PET index to a threshold-based model for the calculation of tumor burden.

## **Paper II**

To develop an AI model for automated segmentation of bones in CT scans and to evaluate the AI model's accuracy and reproducibility.

## **Paper III**

To develop an AI model for automated calculation of PET index and to compare the model's performance to physicians, as well as to a threshold-based model.

## **Paper IV**

To develop and evaluate an AI-based model for the detection of pelvic lymph node lesions in [<sup>18</sup>F]PSMA-1007 PET/CT scans.

**Table 4.** Study design to address the aims and objective of each study supporting this thesis (paper I-IV).

	<b>Aims</b>	<b>Design</b>	<b>Assessment</b>
<b>I</b>	<p>To calculate the skeletal tumor burden from PET/CT (PET index) reflecting the visual interpretation by a reader</p> <p>To assess the agreement of PET index with aBSI</p> <p>To explore the prognostic value of PET index</p> <p>To assess the agreement of PET index and PET index based on a SUV threshold (PET<sub>15</sub> index)</p>	<p>Calculation of tumor burden in PET:</p> <ul style="list-style-type: none"> <li>• Segmentation of bone lesions by application of different SUV thresholds chosen by a reader (PET index), and</li> <li>• based on a SUV 15 threshold (PET<sub>15</sub> index)</li> </ul> <p>Calculation of bone volume in CT:</p> <ul style="list-style-type: none"> <li>• CNN-based bone segmentation</li> </ul> <p>Calculation of skeletal tumor burden:</p> <ul style="list-style-type: none"> <li>• PET and PET<sub>15</sub> indices: Total volume of segmented lesions in PET/total bone volume in CT</li> <li>• aBSI from bone scan</li> </ul>	<p>Kaplan-Meier estimates and log-rank test to explore the presence of a survival difference between patients with high and low index values</p> <p>Univariate Cox proportional hazards regression to assess the association between indices and OS</p> <p>C-index to assess the discriminatory strengths of the different indices</p> <p>Bland-Altman plots to assess index agreement</p>
<b>II</b>	<p>To develop an AI model for bone segmentation in CT</p> <p>To evaluate the AI model's accuracy and reproducibility</p>	<p>CNN training with manual segmentations of 49 bones</p> <p>Evaluation of the AI model in two CT scans (CT<sub>1</sub> and CT<sub>2</sub>) from the same patient for reproducibility assessment</p> <p>Subgroup analysis of five bones in five patients for accuracy assessment between AI model and manual segmentations by a reader</p>	<p>SDI to assess segmentation accuracy between AI model and reader</p> <p>SD and CoV to assess the difference in bone volume calculation between AI model and reader</p> <p>Bland-Altman plots to assess bone volume reproducibility for the AI model</p>
<b>III</b>	<p>To develop an AI model for automated calculation of PET index</p> <p>To compare the AI model's performance to reader performance, and to a threshold-based model</p>	<p>Calculation of tumor burden in PET:</p> <ul style="list-style-type: none"> <li>• CNN-based segmentation of bone lesions (PET index), and</li> <li>• based on a SUV &gt;15 threshold (PET<sub>15</sub> index)</li> </ul> <p>Calculation of bone volume in CT:</p> <ul style="list-style-type: none"> <li>• CNN-based bone segmentation</li> </ul> <p>Calculation of tumor burden:</p> <ul style="list-style-type: none"> <li>• PET and PET<sub>15</sub> indices: Total volume of segmented lesions in PET/total bone volume in CT</li> </ul>	<p>Bland-Altman plots to assess index agreement</p> <p>Spearman rank correlation to assess index correlation</p> <p>TP, FP and FN number of lesions, and sensitivity and PPV, for lesion detection</p>
<b>IV</b>	<p>To develop an AI model for the detection of pelvic lymph node lesions in PET/CT</p> <p>To compare the AI model's performance to reader performance</p>	<p>Training with manual segmentations by three readers</p> <p>Evaluation against three readers</p>	<p>TP, FP and FN number of lesions, and sensitivity and PPV, for lesion detection</p>

aBSI: Automated bone scan index; C-index: Concordance index; CNN: Convolutional neural network; CoV: Coefficient of variation; FN: False negative; FP: False positive; OS: Overall survival; PPP: Positive predictive value, SD: Standard deviation; SDI: Sørensen-Dice index; TP: True positive.



# Methods

## Study Populations

Paper I-IV include patients from Odense University Hospital, Odense, Denmark, and patients from Sahlgrenska University Hospital, Gothenburg, Skåne University Hospital, Lund and Malmö, and Uppsala University Hospital, Uppsala, Sweden. The patients from the Odense, Skåne, and Uppsala study groups have been used in previous studies comparing different scan types for the detection of metastatic disease in patients with prostate cancer.<sup>121-123</sup> The PET/CT scans from the previous studies were used retrospectively to train, validate and test the AI models presented in this thesis.

### Odense Group

A study group from a previous prospective single-center study at Odense University Hospital<sup>121</sup> was used in paper I, II, and III. The objective of the original study was to compare the diagnostic accuracy of bone scan, [<sup>18</sup>F]fluoride PET/CT, and [<sup>18</sup>F]fluorocholine PET/CT using MRI as reference in the detection of spine metastases. The inclusion criteria were (1) biopsy-proven prostate cancer, (2) bone scan consistent with  $\geq 1$  malignant lesion, (3) ability to undergo MRI, and (4) ability to safely postpone the initiation of ADT until after all scans were completed. Exclusion criteria were (1) current or previous treatment with ADT and (2) pain or suspicion of medullar compression due to bone metastases. A total of 50 patients were recruited between May 2009 and March 2012. Patient characteristics are presented in Table 5.

**Table 5.** Patient characteristics for the Odense group ( $n=50$ ).

<b>Age, years</b>	
Mean (SD)	73 (8.6)
<b>PSA, ng/mL</b>	
Median (range)	84 (4-5740)
<b>Gleason Score, <i>n</i></b>	
$\leq 6$	6
7 (3+4)	8
7 (4+3)	10
8	6
9-10	19

SD: Standard deviation. NOTE: Information about Gleason score was missing in one patient.

### *Paper I*

A bone scan, a [<sup>18</sup>F]fluoride PET/CT scan, and available follow-up data at the time of data analysis were required. A total of 48 patients were included in the study group used to test the AI model.

### *Paper II*

CT scans from the [<sup>18</sup>F]fluoride PET/CT (CT<sub>1</sub>) and the [<sup>18</sup>F]fluorocholine PET/CT (CT<sub>2</sub>) scans were used. A total of 46 patients were included in the group used to test the AI model.

### *Paper III*

[<sup>18</sup>F]fluoride PET/CT scans were used. Five patients were excluded due to incorrect image labeling resulting in inaccurate SUVs or due to extremely high metastatic burden, disabling segmentation of individual foci. A total of 45 patients were included and the scans were divided into a training set ( $n=29$ ), a validation set ( $n=8$ ), and a test set ( $n=8$ ), combined with patients from other centers.

## **Sahlgrenska Group**

A study group of 100 patients from Sahlgrenska University Hospital were used to train the AI models in paper I and II in automated segmentation of the skeleton. The patients were included retrospectively and had performed [<sup>18</sup>F]FDG PET/CT examinations between 2008 and 2010. The indications for the scans were known or suspected malignant disease. The mean age was 62 (SD [standard deviation] 12) years, 40% were female, and 60% were male.

### *Paper I*

CT scans from 25 patients were randomly selected and used as a training set.

### *Paper II*

CT scans from all 100 patients were used as training set.

## **Uppsala Group**

Patients from a previous prospective study at Uppsala University Hospital<sup>123</sup> were used for training the AI model in paper III. The aim of the previous study was to evaluate the accuracy of diffusion-weighted MRI and [<sup>18</sup>F]fluoride PET/CT for the detection of bone metastases in patients with high-risk prostate cancer. The inclusion criteria were

biopsy-proven prostate cancer and Gleason score  $\geq 8$ . The exclusion criteria were (1) history of a second cancer and (2) contraindications for performing MRI. A total of 49 patients were recruited between October 2009 and March 2011. The median age was 67 (range 57-80) years.

In paper III, the absence of pathological skeletal uptake on [ $^{18}\text{F}$ ]fluoride PET/CT was required, as the material was used for training the AI model's ability to handle non-suspicious radiopharmaceutical uptake. Scans from 34 patients were included in the training set, combined with patients from other centers.

## Skåne Group

[ $^{18}\text{F}$ ]fluoride PET/CT scans from a previous study at Skåne University Hospital, Lund and Malmö,<sup>122</sup> were used retrospectively for training, validating, and testing the AI model in paper III. The objective of the previous study was to investigate whether combining information from [ $^{18}\text{F}$ ]fluorocholine and [ $^{18}\text{F}$ ]fluoride PET/CT added clinically relevant information for a group of patients with high-risk prostate cancer planned for treatment with curative intent after a normal or inconclusive bone scan. The inclusion criteria for the original study were (1) biopsy-verified high-risk prostate cancer, defined as PSA  $\geq 20$  ng/mL and/or Gleason score  $\geq 8$ , considered for curative treatment, and (2) a bone scan with normal or inconclusive findings. Exclusion criteria were (1) previous or current ADT and (2) PSA  $\geq 100$  ng/mL. A total of 90 patients were included between March 2008 and June 2010. Patient characteristics are presented in Table 6.

**Table 6.** Patient characteristics for the Skåne group (n=90).

<b>Age, years</b>	
Mean (SD)	66 (5.5)
<b>PSA, ng/mL</b>	
Mean (SD)	28 (20)
<b>Gleason Score</b>	
5-6	4
7 (3+4)	17
7 (4+3)	11
8-10	58
<b>cT Stage</b>	
T1c	14
T2	30
T3	46

SD: Standard deviation.

In paper III, one patient was excluded due to incorrect image labeling of the [<sup>18</sup>F]fluoride PET/CT scan. This resulted in the inclusion of 89 patients that were divided into a training group ( $n=53$ ), a validation group ( $n=18$ ), and a test group ( $n=18$ ), combined with patients from other centers.

## PSMA Group

Paper IV included patients with high-risk prostate cancer who were referred for initial staging with [<sup>18</sup>F]PSMA-1007 PET/CT at Skåne University Hospital, Lund and Malmö. A total of 211 patients were included and divided into a training group ( $n=125$ ), a validation group ( $n=36$ ), and a test group ( $n=50$ ).

An overview of all study groups and the combinations of patients used for training, validating and testing the AI models in Paper I-IV is presented in Table 7.

**Table 7.** Overview of the number and combinations of patients from the different study groups used for training, validating, and testing the AI models in this thesis.

Paper	Training Set ( $n$ )	Study Group	Validation Set ( $n$ )	Study Group	Test Set ( $n$ )	Study Group
I	25	Sahlgrenska	20% of training	Sahlgrenska	48	Odense
II	100	Sahlgrenska	20% of training	Sahlgrenska	46	Odense
III	116	Skåne ( $n=53$ )	26	Skåne ( $n=18$ )	26	Skåne ( $n=18$ )
		Odense ( $n=29$ )		Odense ( $n=8$ )		Odense ( $n=8$ )
		Uppsala ( $n=34$ )				
IV	125	PSMA	36	PSMA	50	PSMA

NOTE: In paper I and II, 20% of the patients in each training set were used for validation.

## Image Acquisition

### Odense Group

#### *Bone Scans*

Bone scans were acquired 3 hrs after intravenous injection of 600 MBq [<sup>99m</sup>Tc]hydroxymethylene diphosphonate. Whole-body planar scans with anterior and posterior views (scan speed 14 cm/min, 256×1024 matrix) were obtained using a dual-head gamma camera (SKYlight or PRISM XP2000, Philips Healthcare) equipped with low-energy high-resolution collimators. Energy discrimination was provided by a 20% window centered on the 140 keV of <sup>99m</sup>Tc.

### *PET/CT Scans*

PET/CT scans were obtained by a Discovery VCT 64 (GE Healthcare) 1 hr after intravenous injection of 3 MBq/kg of [<sup>18</sup>F]fluoride, with 2.5 min per bed position from the base of the skull to the midthigh. PET images were reconstructed using ordered subset expectation maximization (OSEM) (2 iterations, 28 subsets) with a 128×128 matrix, pixel size of 5.47 mm, and slice thickness of 3.27 mm. A low-dose CT scan was acquired using tube current modulation (30-150 mA, 140 kV) and was reconstructed in a field of view (FoV) of 50 cm, 512×512 matrix, slice thickness of 3.75 mm, and spacing of 3.27 mm. Within 3 weeks of the [<sup>18</sup>F]fluoride PET/CT, a diagnostic contrast-enhanced CT scan using tube current modulation (80-400 mA, 120 kV) was acquired as part of a [<sup>18</sup>F]fluorocholine PET/CT scan from the base of the skull to the midthigh. The CT images were reconstructed in a FoV of 50 cm, 512×512 matrix, slice thickness of 3.75 mm, and spacing of 3.27 mm.

### **Sahlgrenska Group**

CT scans were acquired as part of [<sup>18</sup>F]FDG PET/CT scans obtained by a Biograph 64 TruePoint (Siemens Healthineers). A low dose CT scan was acquired using tube current modulation (30-110 mA, 120 kV) from the base of the skull to the midthigh. CT images were reconstructed in a FoV of 70 cm, 512×512 matrix, slice thickness of 5 mm, and spacing of 3 mm.

### **Uppsala Group**

PET/CT scans were acquired with a Discovery ST (GE Healthcare) 1 hr after intravenous injection of 3 MBq/kg body weight of [<sup>18</sup>F]fluoride, with 2 min per bed position, from the vertex to the proximal 1/3 of the femur. PET images were reconstructed using OSEM (2 iterations, 21 subsets), a matrix size of 128×128, pixel size of 3.9 mm, and slice thickness of 3.27 mm. A low-dose CT was performed immediately before the PET using tube current modulation (10-80 mA, 140 kV) and reconstructed in a FoV of 50 cm, 512×512 matrix, slice thickness of 3.75 mm, and spacing of 2 mm.

### **Skåne Group**

PET/CT scans were acquired using a Gemini TF (Philips Medical Systems) 1-1.5 hr after intravenous injection of 4MBq/kg (max dose 400 MBq) of [<sup>18</sup>F]fluoride, with 2 min per bed position, from the vertex to the midthigh. The PET images were

reconstructed using the BLOB-OS-TOF algorithm (3 iterations, 33 subsets) to a 144×144 matrix with pixel size and slice thickness of 4 mm. A low dose CT using tube current modulation (30-160 mA, 120 kV) was obtained for attenuation correction and image fusion and reconstructed in a FoV of 60 cm, 512×512 matrix, slice thickness of 5 mm, and spacing of 5 mm. A diagnostic CT with intravenous and peroral contrast was performed 1-24 days previously as part of a [<sup>18</sup>F]fluorocholine PET/CT scan. Images were reconstructed in a FoV of 70 cm, 512×512 matrix, slice thickness of 5 mm, and spacing of 5 mm.

## **PSMA Group**

PET/CT scans were obtained using a Discovery MI (GE Healthcare) 2 hrs after intravenous administration of 4 MBq/kg (max dose 500 MBq) [<sup>18</sup>F]PSMA-1007, with 4 min per bed position, from midthigh to base of the skull. The PET images were reconstructed using a block-sequential regularization expectation maximization algorithm (Q.Clear, GE Healthcare) with a 256×256 matrix, pixel size 2.7 mm, slice thickness 2.8 mm, and a beta factor of 800.<sup>124</sup> A diagnostic CT with intravenous and oral contrast was performed for attenuation correction of the PET scans and anatomic correlation. Scans were acquired using tube current modulation (80-480 mA, 100 kV, or for patients with body mass index ≥30, 80-560 mA, 120 kV) and reconstructed using an adaptive statistical iterative reconstruction technique (ASiR, GE Healthcare) in a FoV of 70 cm, 512×512 matrix, slice thickness 5 mm, and spacing of 5 mm.

## **Manual Segmentations**

In paper I and II, the AI model was trained using manual segmentations of the following 49 bones in CT, comprising approximately 33% of the total skeletal mass;<sup>125</sup> 12 thoracic vertebrae, 5 lumbar vertebrae, sacrum, 2 hip bones, 24 ribs, 2 scapulae, 2 clavicles, and the sternum. The manual segmentations were performed by three experienced readers in paper I and by a single reader in paper III.

In paper II, manual segmentations by an experienced reader were also performed in five patients in the validation group of the following five bones; seventh thoracic vertebra (Th7), third lumbar vertebra (L3), sacrum, right seventh rib, and sternum. These bones were segmented by the same physician in both CT<sub>1</sub> and CT<sub>2</sub>, which generated a total of 25 paired segmentations of the same five bones in the five patients.

In paper III, manual segmentations of bone lesions in [<sup>18</sup>F]fluoride PET scans were performed by a reader based on the original clinical written reports. From the full set of 168 scans, 116 scans were used for training, and 26 scans were used for validation.

The remaining 26 scans were used as test set wherein three specialists in nuclear medicine (readers A, B, and C) also performed segmentations of bone lesions.

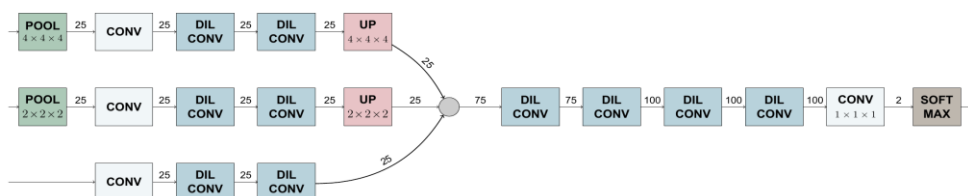
In paper IV, three physicians (readers A, B, and C) independently segmented suspected lymph node metastases below the aortic bifurcation (defined in this thesis as ‘pelvic [miN1]’) in [ $^{18}\text{F}$ ]PSMA-1007 PET/CT scans. Lymph nodes were graded according to E-PSMA guidelines.<sup>126</sup> Grade 1–2 lymph nodes were considered benign, while grade 4–5 were considered pathological. Grade 3 was considered pathological when deviating from known patterns of unspecific uptake, such as low to intermediate uptake along the external iliac vessels. From the full set of 211 scans, 50 scans with three manual segmentations each were used as a test set. The remaining 161 scans were divided into a training set ( $n=125$ ) and a validation set ( $n=36$ ).

The tools used for segmentation include the TurtleSeg software<sup>127</sup> (paper I), the cloud-based segmentation tool eScan Research (eScan Academy AB) (paper II), and the cloud-based annotation platform Research consortium for medical image analysis (RECOMIA)<sup>128</sup> (paper III and IV).

## AI Models

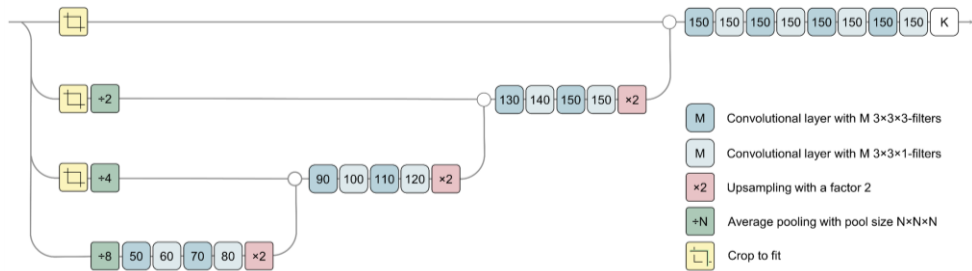
### AI Model Structure

In all studies, CNNs with architectures inspired by the 3D U-Net were used.<sup>117</sup> Fig. 6 shows the structure of the CNN in paper II used for bone segmentation.



**Fig. 6.** Structure of the fully convolutional neural network used in paper II. All convolutions are  $3 \times 3 \times 3$  except the last one. For the dilated convolutions, the first ones have a dilation rate (1,2,2), whereas all the other have (2,4,4). The difference in the first dimension is due to lower spatial resolution in the images. The numbers on the lines indicate the number of data channels after each operation. The number of channels after the last convolution is the number of output labels for each network.

In paper III and IV, the AI models consisted of two CNNs, one which segmented organs (Organ CNN) and one which segmented bone lesions (Lesion CNN, paper III) or lymph node lesions (Detection CNN, paper IV). Both networks had the same structure (Fig. 7).



**Fig. 7.** The structure of the convolutional neural network (CNN) used both for the Organ CNN and the Lesion/Detection CNN. Blue boxes are  $3 \times 3 \times 3$  convolutional layers, and the number indicates the number of filters. Pink boxes are  $2 \times$  upsampling layers, and green boxes are average pooling, where the number indicates the pool size. The effect of the pooling layers is that the network works on four different resolutions. This allows for a large receptive field with low memory cost during training. All convolutional layers use rectified linear unit activations, apart from the last one that uses a softmax activation to produce final output probabilities.

## AI Model Training

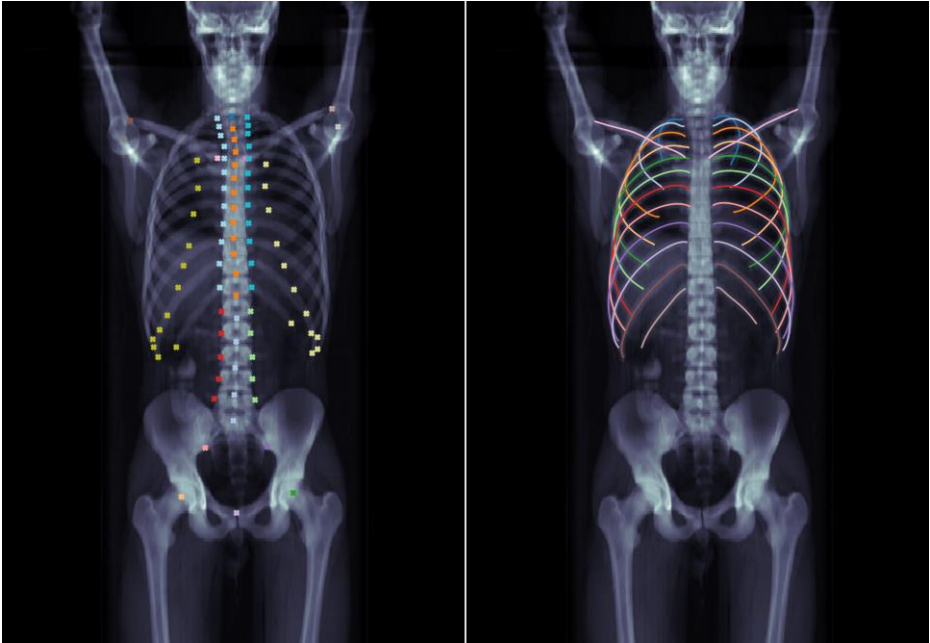
### *Paper I and II*

The training of the AI models for bone segmentation in paper I and II was based on a combination of simple geometric models and learned CNNs. For training the network in paper I, manual segmentations in 25 CT scans were used and for training in paper II, manual segmentations from 100 CT scans were used. The process can be divided into three steps:

#### 1. Landmark detection

A CNN was trained to detect a set of anatomical landmarks and a second network to detect centerlines for clavicles and ribs (paper II) (Fig. 8), as well as humeri and sternum (paper I). The CNN detects types of landmarks such as rib joints and vertebral processes, but due to the limited receptive field, it cannot detect their identity, such as which rib it belongs to. To determine this and to remove possible spurious detections, classical active shape models were used to find plausible relative positions for groups of landmarks. The second network was trained to detect centerlines for elongated bones such as ribs, and an iterative tracking scheme was used to track each bone.





**Fig. 8.** Maximum intensity projection of the CT scan together with the annotated landmarks (left) and detected center lines for ribs and clavicles (right). Landmarks with identical markers belong to the same class and are not separated by the detector.

## 2. Voxel-wise segmentation

The output landmarks from step 1 were fed as auxiliary input to a CNN that performed the final bone segmentation. In paper I, a rough segmentation based on an atlas registered using the aligned landmarks was used as input in addition to the CT scan. In paper II, a five-dimensional image mask was used as input to the final CNN, together with the original CT scan, to provide the network additional information used to correctly differentiate between similar bones, such as adjacent ribs.

## 3. Post-processing

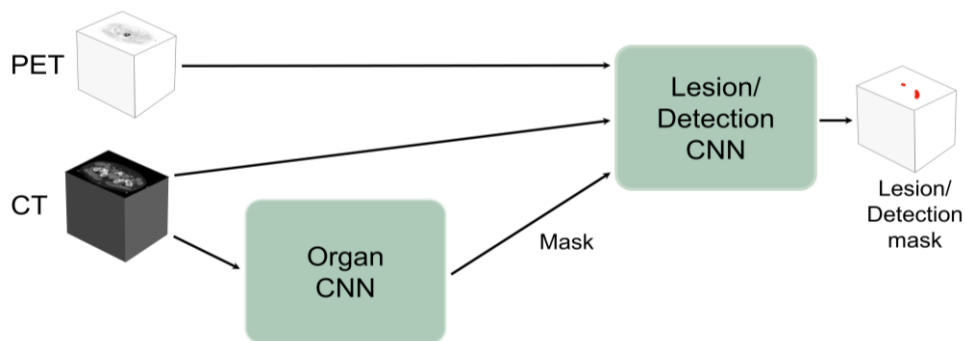
To remove spurious voxels, extraction of the largest component connected to a landmark was performed and filling holes in that component.

### *Paper III and IV*

As a first step, the Organ CNN segmented different organs using a CT scan from the test set as input (Fig. 9). The output from the Organ CNN was used to create a label mask that was added as input to the Lesion/Detection CNN to facilitate the

differentiation between malignant and benign high uptake regions. Hence, the organs marked were either organs where malignant uptake or physiological high uptake is common, or that contribute relevant information for anatomical orientation. In paper III, the marked organs in the label mask included bones, joints, kidneys, lungs, brain, skull, spleen, heart, aorta and liver. In paper IV, the marked organs in the label mask included the prostate and urinary bladder, the gastrointestinal tract, and bone.

The label mask was used as input to the Lesion/Detection CNN, together with the CT and PET data. The different input modalities were concatenated across the feature, or channel, dimension. The convolutional layer had a two output channels with softmax activation. For each voxel, the output value describes the estimated probability of that voxel belonging to a bone (paper III) or a lymph node (paper IV) lesion, or background.



**Fig. 9.** Schematic of the AI models trained for segmentation of bone lesions (Lesion CNN), and detection of pelvic lymph node lesions (Detection CNN), respectively. The Lesion/Detection CNN segments suspected metastatic uptake in the scan with a CT, PET scan and a label mask produced by the Organ CNN as input.

As post-processing in paper III, all connected components  $<0.1$  mL and areas where the high uptake regions were deemed to originate from joints were removed as a post-processing step. In paper IV, pixels above the aortic bifurcation were set to background.

## AI Model Application

### *Bone Segmentation*

In paper I, the AI model was applied for automated bone segmentation in the CT part of the  $^{18}\text{F}$ fluoride PET/CT scans in the test group. In paper II, the AI model was applied to the validation group for automated bone segmentation in the two CT scans, CT<sub>1</sub> and CT<sub>2</sub>, performed in the same patient. In paper III, the Organ CNN was applied for automated whole-body skeletal segmentation in the test group. Bone volumes were

calculated based on the segmentations. Individual volumes of segmented bones were summed to obtain the volume of the axial skeleton (paper I) and the entire imaged skeleton (paper III).

#### *Segmentation of Bone Lesions*

In paper III, the AI model was applied for fully automated segmentation of bone lesions in the test set. No manual corrections were made.

#### *Detection of Lymph Node Lesions*

In paper IV, the AI model was applied to the test set for automated detection of pelvic lymph node lesions.

## **Quantification of Skeletal Tumor Burden**

#### *Automated Bone Scan Index (aBSI)*

In paper I, EXINIbone<sup>BSI</sup> version 2 (EXINI Diagnostics AB) was used for aBSI calculation from each patient in the test and study group, respectively. The details of the automated BSI calculation and platform have been described in previous studies.<sup>85,129</sup> First, the different anatomical regions of the skeleton, except the distal parts of the extremities, are segmented. Secondly, abnormal hotspots are detected and classified as metastatic lesions. The volumes of all hotspots are then summarized and divided by the total skeletal mass, generating the aBSI.

#### *Calculation of PET Index*

The general principle for calculating PET indices was to divide the total volume of segmented suspicious uptake in the PET scan by the skeletal volume measured from the CT scan. In paper I, two methods were used to segment bone lesions in the [<sup>18</sup>F]fluoride PET scans from the patients in the test group:

### **1. Manual approach**

An individual SUV threshold was selected based on the visual interpretation by a nuclear medicine specialist. This approach aimed to reflect the clinical interpretations as closely as possible. The choice of threshold was made so that all lesions interpreted to have possible metastatic origin were segmented. After selecting a threshold, each segmented uptake was classified as caused by metastatic disease or not, based on the interpretation of the same nuclear medicine specialist. Lesions classified to originate from non-malignant conditions such as degeneration, inflammation, or fractures were excluded from the analysis. Selected thresholds

ranged between SUV 6-9. The physician was blinded to the patients' bone scans, BSI values, and survival data.

## 2. **Threshold-based approach**

A fixed threshold was applied to segment lesions with SUV >15.

The volume of each lesion located in the segmented skeletal regions in the corresponding CT scan was calculated. A PET index was then calculated by dividing the sum of all lesion volumes by the total volume of the 49 segmented bones. Two indices were calculated from each patient's scan, one based on the manual approach (PET index) and one based on the automated approach (PET<sub>15</sub> index). In order to be comparable to aBSI in the same patient, all PET indices were multiplied by 0.33 since the included bones comprised approximately 33% of the total skeletal mass.

In paper III, PET indices were obtained by dividing the total volume of all bone lesions segmented by the AI model by the total imaged skeletal volume of each patient in the test group. A threshold of SUV 15 was also applied in this study, and PET<sub>15</sub> indices were calculated by dividing the volume of all segmented lesions by the total skeletal volume.

## Statistical Analysis

Data were analyzed on a patient-level in paper I, II, and III. Lesion-based analyses were performed in paper III and IV.

### Survival Analysis

In paper I, Kaplan-Meier estimates of the survival function and the log-rank test were used to explore the presence of a significant difference in OS between high and low aBSI, and between high and low PET index groups. The group with high indices was defined as those with values above the median value, and the group with low indices as those with values below the median value. Due to the exploratory nature of the study, the choice of a median split was made as there were no known cut-offs to stratify patients based on PET indices. Univariate Cox proportional hazards regression was used to evaluate the association between the index and OS. The discriminatory strengths of the different indices were assessed using concordance index (C-index).

## Index Comparison

Bland-Altman plots were used to assess the agreement between the different indices in paper I and III. The correlations between the different indices in paper III were assessed using Spearman rank correlation.

## Assessment of AI Model Performance

In paper II, the accuracy of the AI model compared to a reader was investigated through a comparison of bone segmentation overlap and by calculation of the resulting volume differences. Segmentations by one reader performed of the same five bones in both CT<sub>1</sub> and CT<sub>2</sub>, in the subset of five patients from the test set, acted as a reference standard. The automated bone segmentations by the AI model were compared to the manual segmentations, and the accuracy between these segmentations was evaluated using Sørensen-Dice index (SDI).<sup>130</sup> The SDI evaluates how well two segmentations, A and B, of the same object agree by analyzing the overlap. A voxel is defined as belonging to the overlap of two segmentations if, in the case of paper II, it is classified as bone in both segmentations. The SDI is defined as two times the number of overlapping voxels divided by the sum of the total amount of voxels that are classified as bone in both segmentations according to the following formula. The resulting value ranges between 0-1, where 1 reflects a perfect segmentation:

$$DICE(A, B) = \frac{2|A \cap B|}{(|A| + |B|)}$$

The percentage difference between the two AI-based volume measurements of the same bone in CT<sub>1</sub> and CT<sub>2</sub> was calculated, as well as the SD and coefficient of variation.

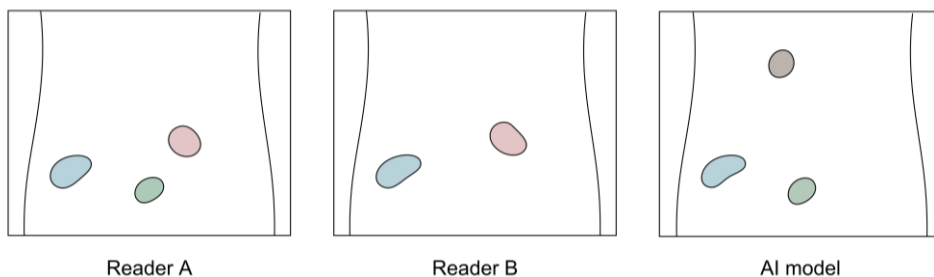
To assess the reproducibility, which in paper II was defined as the agreement in AI-based bone volume calculation in two CT scans from same patient, Bland-Altman plots were used.

## Lesion-level Analysis

For the lesion detection in paper III and IV, the sensitivity was evaluated as the percent of suspected lesions (bone or lymph node) detected by the AI model, threshold model, or a reader, respectively, out of those detected by a reader used as reference. True positive lesions for the AI model, threshold model, or a reader, were defined as either

full or partial segmentation overlap with another reader used as reference, or else they were considered false negative. Lesions detected by the AI model, threshold model or reader without segmentation overlap with another reader used as reference were considered false positive. Fig. 10 shows an example, where reader A is first used as a reference, followed by reader B, giving different values of sensitivity. The positive predictive value was calculated as the proportion of true positive lesions for the AI model, threshold model or reader, when compared to another reader used as reference, divided by true positive plus false positive lesions when compared to the same reference reader. The specificity and negative predictive value cannot be calculated on a lesion basis since it is not possible to calculate non-malignant lesions not detected by either model or reader.

For all analyses in paper III, bone segmentations by each of the four readers were alternately used as a reference and pairwise compared to the AI model, the threshold model, or another reader. The average and range of all pairwise combinations were calculated for AI model versus reader, threshold model versus reader, and reader versus reader. In paper IV, the segmentations of lymph node lesions by each of the three readers were alternately used as a reference and pairwise compared to either AI model or another reader in the same manner as in paper III.



**Fig. 10.** Example of how sensitivity was calculated using different readers as a reference. In this case, reader B detects 2/3 of the lesions marked by reader A, giving a sensitivity of 67%, whereas reader A detects both lesions marked by reader B, giving a sensitivity of 100%. Similarly, the AI model has a sensitivity of 67% with reader A as a reference and 50% with reader B as a reference.

## Ethical Considerations

Ethical approval was granted by the local research ethics committee at Lund University for the studies in paper I-IV, at Odense University Hospital for the studies in paper I-III, at Sahlgrenska University Hospital for the studies in paper I and II, and at Uppsala University for the study in paper III.

All scans were acquired as part of previous study, or because clinically indicated. No examinations were performed only for current research purposes in paper I-IV. Before transferred from a center for image analysis, all scans were pseudonymized so that they could no longer be linked to a single subject. The code key was stored separately.

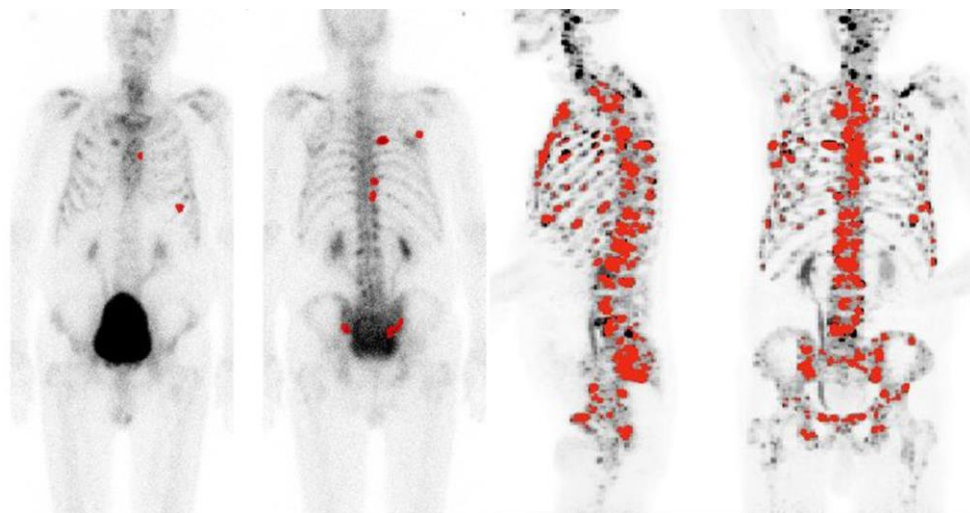




# Results and Comments

## Quantification of Skeletal Tumor Burden

A measure reflecting the skeletal tumor burden in [ $^{18}\text{F}$ ]fluoride PET/CT, coined the PET index, was introduced in paper I. A higher PET index was associated with shorter OS (hazard ratio 1.17 [95% CI 1.06-1.29]), providing proof-of-concept of its prognostic potential. The most common divergence between the indices on a patient-level was a higher PET index than aBSI, likely reflecting the superior sensitivity of [ $^{18}\text{F}$ ]fluoride PET/CT. An example of aBSI and PET index calculated from the same patient is shown in Fig. 11.



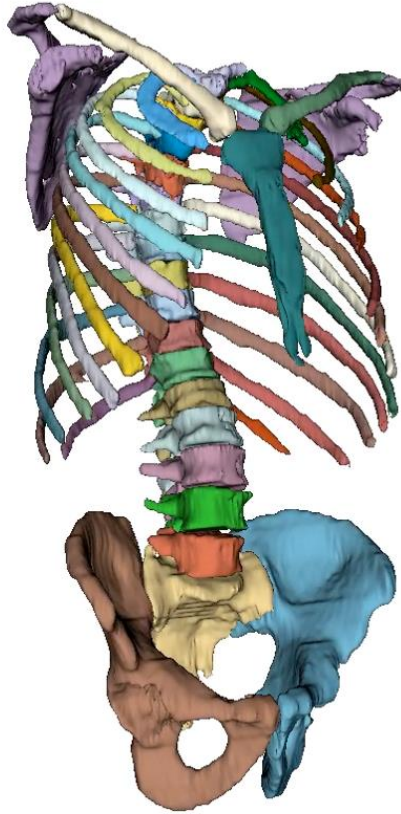
**Fig. 11.** Segmented lesions (red) in bone scan (aBSI 0.4%) and [ $^{18}\text{F}$ ]fluoride PET scan (PET index 7.9%) from a patient in the test set in paper I. Left-right: Anterior bone scan projection, posterior bone scan projection, PET maximum intensity projections.

Whole-body tumor burden measured from bone scan can stratify patients according to symptomatic progression and OS.<sup>88</sup> Skeletal tumor burden measured from PET/CT has been shown to correlate with OS in patients with advanced prostate cancer.<sup>93, 104, 108</sup> In line with these studies, the results from paper I supported the prognostic value of a

measure of skeletal tumor burden in PET/CT but warranted further development of an automated method for more efficient and accurate estimation of suspicious uptake, as judged by visual interpretation.

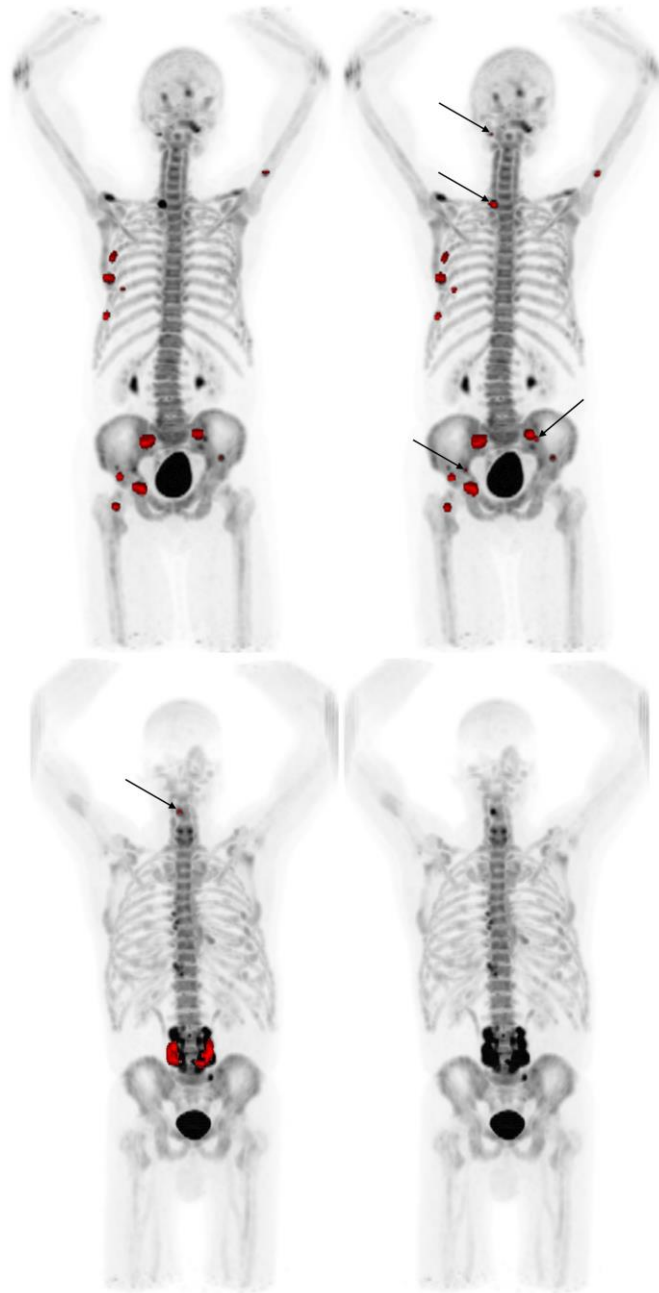
To be comparable to aBSI, the PET index presented in this thesis was based on both uptake in the PET scan and on the skeletal volume information from the CT scan. Theoretically, an indexed measure of tumor burden allows for inter-patient comparison by considering patient size. However, it requires the inclusion of the same bones, or parts of bones, in the scans, which was only sometimes the case. The development of an indexed measure of tumor burden that is comparable between patients, similar to aBSI, would benefit from standardized PET/CT scanning that consistently includes the same bones, or parts of bones, in the FoV.

In paper II, an AI model was trained for automated bone segmentation. It took the AI model approximately 2 min per CT scan to segment 49 bones (Fig. 12). The intra-reader volume difference was less with the AI-based than with the manual approach. For example, the volume difference for an AI-based segmentation of Th7 was 2% versus 14% for a physician. The segmentation of different vertebrae as Th7 in the two CT scans from the same patient explained the intra-reader disagreement for the reader. The potential to reduce intra- and inter-reader disagreement, combined with much faster segmentation time, points to the advantage of using AI for automated bone and organ segmentation. A limitation of this study was that the AI model only was validated in five bones due to the time-consuming task of performing manual segmentations. Since the publication of paper II, the model has been further developed to segment the entire skeleton,<sup>128</sup> and was used in paper III.



**Fig. 12.** A representative reconstruction of the AI-based segmentations of 49 bones from one patient in the test set in paper II.

In paper III, a fully AI-based model for automated segmentation of bone lesions in [ $^{18}\text{F}$ ]fluoride PET/CT was developed. Patient examples of segmentations by both the AI model and reader with resulting indices are shown in Fig. 13. The study showed that even among experienced specialists working at the same center, agreement on a patient-level regarding the prevalence of suspicious uptake was <75%. The AI model did not rule out the presence of suspicious uptake in any patient where the readers agreed on prevalence which was encouraging but identified  $\geq 1$  uptake in six patients in whom all readers agreed on the absence of suspicious uptake. For example, the AI model could not discriminate between suspected malignant uptake and uptake due to spinal surgery (Fig. 13). These results highlight the issue with a limited training set; an easy task for an experienced physician can be unachievable for AI due to the lack of similar training examples. More extensive and varied training material with different uptake patterns would probably improve the model's performance in this regard.



**Fig. 13.** Two patient examples from paper III of segmented bone lesions (red) by the AI model and reader in [ $^{18}\text{F}$ ]fluoride PET/CT scan. Top left: The AI model identifies several lesions in the right ribcage, pelvis, right hip, and left arm. PET index 1.34%. Top right: Reader D identifies additional lesions in the skull, spine and, pelvis in the same patient (arrows). PET index 1.42%. Bottom left: The AI model detects lesions in the lumbar spine and a small lesion in the cervical spine (arrow). PET index 0.57%. Bottom right: None of reader A-D identifies any lesions in the same patient. Of note, the patient had previously undergone lumbar spinal fusion surgery.

## **Current Results in Relation to Present Status for Tumor Volume Quantification**

Other measures of tumor burden in both [<sup>18</sup>F]fluoride and PSMA PET/CT have been presented, such as total lesion uptake (TLU) and total tumor volume (TLV). These measures have shown promising results for risk stratification, prediction of survival, and assessment of treatment response in small patient materials.<sup>95, 98, 100-103, 106, 107</sup> TLV and TLU only include the PET information and neglect the volumetric information in the CT scan. The presented methods are threshold-based and require manual correction before a final measure is determined. Therefore, their potential to alleviate the workload is limited, especially in patients with a high number of lesions, until fully automatic methods are available. In summary, different measures of quantifying tumor burden in PET/CT are currently being investigated, but there is so far no univocal evidence in favor of one. However, recent leaps forward in methodology include fully automated CNN-based models for quantification of whole-body tumor burden in PSMA PET/CT.<sup>109-111, 131</sup>

## **Current Results in Relation to Present Status for AI-based Bone Lesion Detection**

Analytical validation of an AI model is the first step toward its qualification as an imaging biomarker. It encompasses understanding the risk of detecting false positive or false negative lesions, which needs to be managed and minimized. In paper III, the average number of false positive bone lesions was 2.6 per patient, and the number of false negatives was 2.0 per patient compared to physicians. Hence, the number of false positive and false negative lesions was acceptable and relatively balanced, which should be the aim if the purpose of an AI model is to quantify the total tumor burden.

For an AI-based software for lesion detection in [<sup>18</sup>F]DCFPyL PET/CT, the number of false positive bone lesions was 8.3 per patient.<sup>110</sup> The sensitivity was higher (87%) than for the model in paper III (71%). In another study where an AI-based model for segmentation of lesions in [<sup>68</sup>Ga]Ga-PSMA-11 PET/CT was presented, the overall false positive number of lesions was 1 in every 4.3 patients, but the sensitivity was only 58% for the detection of bone lesions.<sup>109</sup> A third recent study showed an agreement of 62% in M stage assignment and 77% for identification of any distant metastases (M0 vs. M1) when compared to expert assessment.<sup>131</sup>

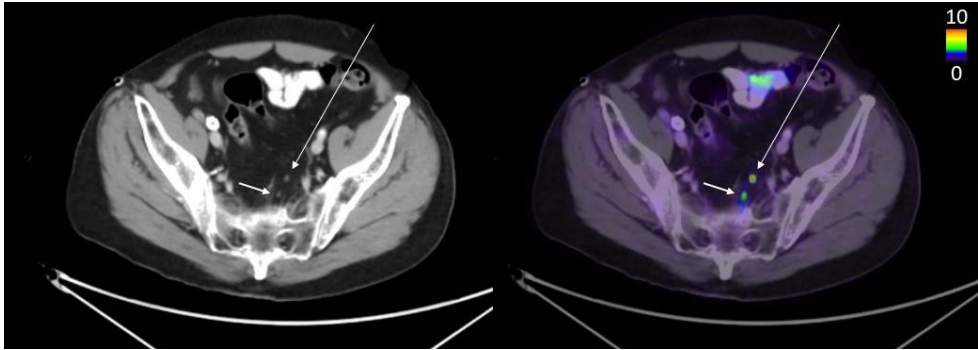
## Detection of Pelvic Lymph Nodes

As a step towards the assessment of total prostate cancer spread, an AI model for identifying pelvic lymph node lesions in [<sup>18</sup>F]PSMA-1007 PET/CT was developed in paper IV. The sensitivity of the AI model (mean 82% [range 76-90%]) was within the inter-reader range of three specialists in nuclear medicine (mean 77% [range 55-98%]). Inter-reader agreement among the specialists was not perfect despite the use of E-PSMA guidelines. Similarly, other studies have shown that even when applying criteria for standardized reporting of PSMA PET/CT, inter- and intra-reader disagreement still exists among experienced physicians.<sup>118, 119</sup> This highlights one of the implications of AI, which is the potential to help standardize scan interpretation and act as decision support to physicians.

If the aim of an AI model is to act as decision support or second opinion, it is important that it recognizes potentially metastatic uptake. At the same time, the identification of a large number of false positive lesions makes the interpretation time-consuming, thereby defeating its purpose. The effort of keeping the number of false negatives low in paper IV (0.2 per patient) came at the expense of a slightly higher number of false positives (1.8 per patient) for detecting pelvic lymph node lesions. An example of a true positive and a false negative detection is shown in Fig. 14.

### **Current Results in Relation to Present Status for AI-based Lymph Node Lesion Detection**

The number of false positives in paper IV was lower compared to another AI-based software developed in [<sup>18</sup>F]DCFPyL PET/CT, for which the average number of false positive regional and distant lymph node lesions was 19.5 and 90.8 per patient, respectively.<sup>110</sup> Another recent study showed a sensitivity of 90% in the automated detection of regional lymph node lesions, but the number of false positives was not stated and the patients had considerably more metastases than in paper IV, making the two studies difficult to compare.<sup>111</sup> In a third study, where an automated method for segmentation of lesions in [<sup>68</sup>Ga]Ga-PSMA-11 PET/CT was presented, the overall false positive number of lesions was only 1 in every 4.3 patients, but the sensitivity was only 58% for bone lesions and 68% for regional lymph node lesions.<sup>109</sup> Another CNN-based method for [<sup>68</sup>Ga]Ga-PSMA-11 PET/CT showed an agreement of 81% with expert assessment for the identification of pelvic nodal involvement (N1).<sup>131</sup>



**Figure 14.** Patient example of a  $[^{18}\text{F}]$ PSMA-1007 scan in paper IV. The long arrow shows a lymph node metastasis detected by all readers and the AI model (true positive). The short arrow shows a lymph node metastasis marked by all readers but not detected by the AI model (false negative). Left-right: axial CT scan, fused PET/CT scans.

### Role of SUV Thresholds for Lesion Segmentation

In both paper I and III, a SUV threshold was applied for automated segmentation of bone lesions to compare to reader-based segmentations. In paper I, the threshold-based  $\text{PET}_{15}$  index reflected, on average,  $1/3$  of the tumor burden as defined by the PET index based visual interpretation. In paper III, agreement with readers regarding the prevalence/absence of bone lesions was lower for the threshold model than the AI model. Application of a global SUV threshold for segmentation of suspected metastatic in  $[^{18}\text{F}]$ fluoride PET/CT is problematic, as the uptake levels in malignant disease and benign conditions are known to overlap.<sup>132</sup>

SUV thresholds have been used in recent studies for detecting lesions in PSMA PET/CT,<sup>99, 103</sup> and have found drops in sensitivity for both lymph node and bone lesions compared to readers.<sup>110</sup> In paper IV, applying a SUV threshold of 3.0 resulted in a very high number of false positive lesions that the readers did not select, on average 19.9 per patient (range 19.6–20.1 depending on which reader was used as reference).

In summary, simple SUV thresholds without manual corrections do not accurately represent tumor burden. It is also known that using different PET/CT scanners, acquisition protocols, and reconstruction algorithms result in different SUVs, especially in small lesions.<sup>133</sup> The results in paper I, III and IV support the application of AI for more accurate calculation of the skeletal tumor burden compared to a global SUV threshold.

## Aspects of Ground Truth

For the training and testing of the AI models in this thesis, methodological limitations can be discussed in view of three aspects; the segmentation of suspicious lesions, the classification of such lesions as benign or malignant, and the amount of annotated data.

Manual segmentations are commonly used when training and testing a CNN to detect or segment abnormal features such as suspected metastatic uptake in PET/CT scans.<sup>115</sup> However, manual segmentations are not a perfect reference standard as there is a known propensity for inter-reader delineation differences.<sup>134</sup> Additionally, due to partial volume effect, the voxel contours do not match the exact contours of the pathological abnormalities. In theory, only histopathology can provide ground truth for diagnosis and information on boundaries. A limitation throughout this thesis is the lack of verification of metastases by the gold standard. Given the widespread disease in many patients (Fig. 11), it would have been practically and ethically impossible to obtain a histopathological diagnosis for every bone lesion in paper I and III. In paper IV, only 27 patients in total, 6 of which in the test group, underwent eLND following PSMA PET/CT scanning, which impeded histopathological verification in most patients.

In the absence of gold standard information, study designs often use consensus readings as a reference standard for detecting and classifying lesions in PET/CT. Consensus readings are not optimal for several reasons, one being that they omit information on interpretation discrepancies by overruling the minority opinion. In this thesis, using one physician at a time as a reference and averaging the results over the choices of reference avoids giving one physician's interpretation more impact than another, and allows for a more nuanced picture of reader interpretations. However, when interpreting the results, one should bear in mind that the true information of lesion numbers and origin in the patients cannot be known with certainty.

The type and amount of annotated data required to train a CNN depend on the images and the task. Segmentation of an organ or an abnormal feature requires pixel-wise segmentation, which is laborious when performed manually, especially in 3D data such as CT and PET/CT.<sup>115</sup> Although there was a gradual expansion of the data sets with each study, the time-consuming and labor-intensive task of manually segmenting suspected metastatic lesions (paper I, III, and IV) and bones (paper I and II) was a limiting factor. As mentioned, the AI model could not distinguish between malignant uptake and uptake secondary to spondylodiosis in paper III, which was an easy task for the physicians. In paper IV, lesions in presacral lymph nodes were the most commonly missed by the AI model, as metastases in this station are rarer than in iliac lymph nodes.<sup>135</sup> The performance of the AI model in correctly recognizing less common



conditions and uptake patterns would most likely improve with larger training sets containing more varied examples.

Since publication, the AI model in paper IV has been further developed to detect primary prostate cancer tumors and distant lymph node and bone metastases.<sup>136</sup> The model was trained and validated using 540 scans, compared to 161 scans in paper IV, from [<sup>18</sup>F]PSMA-1007 PET/CT scans performed both at initial staging and because of suspected recurrent disease. A difference is that one expert reader performed the manual segmentations used for training, compared to three readers in paper IV. In this recent study, the results showed better overall performance of the AI model when using the same reader (whose segmentations had been used for training) as a reference in the evaluation, compared to the other readers used as reference. It underscores that manual segmentations by several experts are essential in achieving general applicability and avoiding bias when training a CNN, in addition to a more extensive training set.

## Limitations of PET/CT Scanning

Inherent limitations exist for both [<sup>18</sup>F]fluoride and PSMA PET/CT. [<sup>18</sup>F]fluoride PET/CT cannot detect lymph node or visceral metastases. High uptake of [<sup>18</sup>F]fluoride and PSMA-targeting radiopharmaceuticals in bone is seen in conditions other than skeletal metastases. Increased [<sup>18</sup>F]fluoride uptake due to high bone turnover occurs in degenerative joint disorders, early stages of osteoarthritis, and fractures.<sup>132, 137</sup> Non-specific bone uptake of [<sup>18</sup>F]PSMA-1007 is also a common finding, especially in the ribs and pelvis,<sup>68, 138</sup> and high uptake is also seen in malignancies other than prostate cancer.<sup>65</sup> The underlying response to therapy in prostate cancer tumor cells may not be directly or linearly reflected in PSMA uptake. Flare phenomenon following the start of ADT has been described,<sup>139, 140</sup> and PSMA expression may vary depending on ADT duration, although data are limited.<sup>141-143</sup> There are case reports of patients with metastatic neuroendocrine differentiated metastatic prostate cancer where lesions do not show uptake of PSMA-targeting radiopharmaceuticals.<sup>144, 145</sup> As with visual inspection, the diagnosis, possible ADT interferences, and timing of the scan have to be taken into account if applying an automated method for image analysis during therapy monitoring.



# Conclusions and Future Scope

In this thesis, CNNs were trained to segment bone lesions in [<sup>18</sup>F]fluoride PET/CT to measure whole-body skeletal tumor burden. This measure, coined the PET index, was shown to more closely reflect the interpretation by a reader compared to applying a SUV threshold for lesion segmentation. Another model was trained to detect pelvic lymph node lesions in PET/CT scans obtained with [<sup>18</sup>F]PSMA-1007 with a sensitivity comparable to experienced readers.

Going forward, the development of one AI model for detection and quantification of the primary tumor or local recurrence, as well as bone, lymph node, and visceral metastases in PET/CT, could provide a complete measure of the overall metastatic tumor burden. Adaption of the model's ability to estimate wide ranges of radiopharmaceutical uptake patterns at different stages of the disease would be required. A key to developing an AI model that can qualify for clinical practice and use in trials, is to enlarge the size and the variability of training data. Training a model on the collective experience embedded in a data set created by many experts could potentially elevate its performance over that of the average reader. Gathering images from different institutions obtained with different PET/CT scanners and imaging protocols, including different accumulation times and reconstruction algorithms, and with segmentations performed by multiple expert readers, are necessary for developing robust models that are broadly applicable. With these necessities come logistic and legal obstacles, which have to be overcome.

Given the limited availability of expert-annotated data, one possibility may be combining scan types for leveraging information from different radiopharmaceuticals, such as [<sup>18</sup>Ga]- and [<sup>18</sup>F]-labeled PSMA-targeting compounds, in model training. This approach may also expedite the development of AI models for future novel PET radiopharmaceuticals by not having to train them *de novo*. Adding external information to the input when training, such as results from postoperative histopathological examination after prostatectomy or eLND, may also elevate the AI model's performance by not limiting its skills to the readers used for training.

This work is part of a continuous effort to develop PET/CT imaging biomarkers as indicators of prognosis and treatment efficacy in patients with prostate cancer.

Validation in analytical and clinical studies is required to explore the true potential of AI-based assessment of PET/CT. This process has been demonstrated with the development of aBSI, and has resulted in a clinically applicable imaging biomarker.<sup>86,88</sup> Ultimately, large prospective, multicenter clinical trials are needed to evaluate whether AI-based assessment of PET/CT imparts valuable clinical information for prognosis, prediction, and evaluation of treatment response and outcome in patients with prostate cancer.

# Populärvetenskaplig sammanfattning

Bilddiagnostik är en hörnsten i handläggningen av patienter med prostatacancer. Bildtolkningen är en tidskrävande process och beroende av den enskilda tolkarens erfarenhet och kunskap. Detta begränsar effektiviteten och likvärdigheten i diagnostiken, vilket i slutändan påverkar patienters behandling och prognos. I den här avhandlingen har artificiell intelligens tränats att tolka cancers utbredning. Metoden kan leda till en standardiserad bildtolkning och därmed bättre övervakning och uppföljning av cancerbehandlingseffekten.

Prostatacancer är den vanligaste cancer hos män i västvärlden. I Sverige upptäcks cirka 10 000 nya fall varje år. Risken att dö i prostatacancer är beroende av omfattningen av sjukdomen. När cancer är begränsad till prostatan dör 4,5% av patienterna till följd av sjukdomen under de första 10 åren efter diagnos. Om cancer spritt sig utanför prostatan minskar den beräknade livslängden kraftigt till cirka 3-4 år. Vid behandling av prostatacancer är det därför viktigt både att upptäcka spridning (s. k. metastaser) tidigt och att tillförlitligt uppskatta risken för framtida metastaser, samt att kunna utvärdera förändringar i metastasbördan under behandling.

Positronemissionstomografi kombinerat med skiktröntgen (PET/DT) är idag ansedd som den mest känsliga avbildningsmetod för att undersöka spridning av prostatacancer. Bedömningen av bilderna görs av en läkare och fynden rapporteras i fritext, vilket gör att tolkning och rapportering blir subjektiv och tidsödande.

Att tolkningen av PET/DT är så beroende av den enskilda läkarens erfarenhet och kunskap är ett problem. Ett sätt att lösa detta problem är att utveckla automatiska metoder som på ett objektivt sätt beräknar de sjukliga förändringarna i en bild genom en så kallad bildbiomarkör. I sjukvården kan bildbiomarkörer få positiva konsekvenser i utvecklingen mot mer personanpassad vård.

Artificiell intelligens har potential att hjälpa till i utvecklingen av en sådan metod. Tekniken bygger på en variant av maskininlärning, så kallad "deep learning" och är samma teknik som används vid utvecklingen av till exempel självstyrande bilar eller ansiktsgenkänning. Datorn tränas med ett stort antal bilder och lär sig på sätt identifiera strukturer och eventuella avvikelser. Målet är att få datorn att tolka bilden lika bra, om inte bättre, än en människa.

I den här avhandlingen tränades artificiell intelligens att tolka PET/DT-bilder och beräkna cancerutbredningen i kroppen på enbart några sekunder. I en liten grupp med patienter visades det här måttet, som kallas PET index, vara kopplat till överlevnad. Att större tumörutbredning är kopplat till sämre överlevnad är i sig inte förvånande, utan styrkan ligger i att cancerspridningen kan presenteras på ett objektivt sätt i sifferform, snarare än i läkarens svarstext.

Den här avhandlingen visade att de är möjligt att använda artificiell intelligens för att hitta och beräkna cancerutbredning utifrån PET/DT-bilder. Artificiell intelligens kommer inte ersätta behovet av läkartolkning, men kan effektivisera arbetet och fungera som ett beslutsstöd. De objektiva och mätbara mått som artificiell intelligens kan producera utifrån medicinska bilder har potential att bidra med information om förväntad överlevnad och lämplig behandling, eller kombinationer av behandlingar. Måtten skulle också kunna användas för att utvärdera hur patienter svarar på given behandling och för att utvärdera effekt i läkemedelsstudier.

# Errata

## Paper I

The energy window for the bone scan acquisition was 140 keV  $\pm 10\%$  and not  $\pm 20\%$ .

In the training group, the CT images were reconstructed with a slice thickness of 5 mm, not 3.27 mm as stated.

The CT scan in the study group was not obtained as a diagnostic contrast-enhanced scan (80-400 mA, 120 kV), but as a low-dose scan (30-150 mA, 140 kV).

Fig. 3: The value of the lower limit of agreement is -4.79 as stated, but the line is misplaced.

Fig. 4b: The PET index was 7.9%, not 2.6%.

## Paper II

A low dose CT scan was acquired using tube current modulation 30-110 mA, not 30 mAs.

The parameters for the low-dose CT scan acquisition in the validation group was 30-150 mA, 140 kV, and not 80-400 mAs, 120 kV.

## Paper IV

The CT images were reconstructed with a slice thickness of 5 mm and not 0.625 mm.





# Acknowledgements

I would like to extend my sincere thanks to:

**Elin Trägårdh**, I am so grateful to have had you as my main supervisor. Every PhD student should be as lucky. Thank you for your encouragement and for understanding when my focus has shifted to my clinical work from time to time. I value your guidance and admire you in many ways.

**Lars Edenbrandt**, you are the mind behind this, as with so many other things. Your creativity and innovativeness is inspiring. Thank you for always taking the time, for introducing me to Elin, and for the trips to Odense.

**Anders Bjartell**, for seizing my interest in prostate cancer when I was just a medical student, for sharing your wide knowledge about the disease, and for always answering my e-mails wherever you are in the world.

**Olof Enqvist**, **Måns Larsson**, and **Johannes Ulén**, you may be the smartest people I have met, but I am too dumb to know. Thank you for the interesting and stimulating collaboration and for all the explaining. This thesis would not have been possible without your skills and your (Eigen)vision.

**David Minarik**, for the lectures on PET, the help with technical details, and for teaching me the correct Swedish translation for ‘convolutional neural network’ at my half-time review. Not a day too late.

**Fredrik Liedberg** and **Peter Leander**, for the public discussion and examination at my half-time review that helped propel the work.

**Aseem Anand** and **Mariana Reza**, who paved the way. Thank you for sharing your experience and subject expertise.

**Anna Åkesson**, for the statistical guidance.

**Tobias Axmarker**, my boss who is always supportive and has given me the required time to finish this thesis. If all bosses were like you, research residency would be a happier place.

At **Odense University Hospital, Odense, Denmark** – **Jane Simonsen** and **Poul Flemming Høilund-Carlsen**, for the visits to the Department of Nuclear Medicine and for the invitation to the Abass Alavi Meeting many years ago. A special thank you to Jane for helping with the [<sup>18</sup>F]fluoride PET scan interpretation. Poul Flemming, I especially value your input as a co-author, which has taught me a lot about the writing process.

At **NewYork-Presbyterian/Weill Cornell Medical Center, New York, USA** – **Prof. Joseph R. Osborne**, for taking me under your wings when the initial plans for my Master's fell through, and for sparking my interest in molecular imaging through your elective study.

**Pia Lindgren**, my mother, for doing your best and for the things it brought.

**Sven-Erik Lindgren**, min morfar, som har lärt mig att ska man göra något så ska man göra det ordentligt (ibland är det svårt).

**Eva Frölich Berg** and **Mauritz Berg**, for your generosity in general, and the babysitting in particular (and for your firstborn).

**Joakim Isendahl** and **Carolina Muszynska**, besides being my good friends, you are also PhD student predecessors. I have identified with, and found encouragement, in both of your journeys. Beyond that, you are also some of my favorite people to talk to about other things.

**Michaela Boström**, **Irma Mehmedagic**, **Therese Olsson**, and **Amelie Stenqvist**, a.k.a. Temporal trends in you-know-what. Thank you for allowing me to share all sorts of things with you, and for everything you have shared with me during the years. A special thank you to Amelie, for all the support during this shared final sprint.

**Mikael Larsson**, my eternally optimistic friend. Thank you for your encouragement and for all the times you generously let me borrow 91st & 3rd, where part of this work was written.

Finally and most importantly, thank you to **Johan** och **Elis**, the Berg part of the Lindgren Belal-Berg family. You are everything.

This work was supported by the Swedish government and Region Skåne under the ALF agreement (the Medical Training and Research Agreement – Avtal om Läkarutbildning och Forskning), Regional Research Support by Region Skåne, Sweden, and Knut and Alice Wallenberg Foundation, Sweden.

# About the Author



I was born in 1988 in Vallentuna, Sweden, to a Swedish mother and an Egyptian-American father. My background has made me feel at home in many countries and cultures, and I have always enjoyed traveling. In 2014, I received my medical degree at Lund University. Currently, I am working as a surgical resident at Skåne University Hospital, Sweden. My research focuses on how artificial intelligence can improve the interpretation of medical images.

My curiosity in molecular imaging began during my final year of medical school. I had taken an interest in urological surgery and especially the management of prostate cancer. With help from Prof. Anders Bjartell, I arranged to write my Master's thesis at Weill Cornell Medical Center, New York, as part of a project investigating different PET radiopharmaceuticals in the setting of neuroendocrine prostate cancer. Until then, I had primarily known diagnostic imaging to be structural. After five years of medical studies, the concept of molecular imaging was essentially new to me. I wondered why, given the ingenuity. After finishing my Master's, I prolonged my stay in New York to do an elective study in Molecular Imaging & Therapy at Memorial Sloan Kettering Cancer Center. Back in Sweden, I finished my medical degree and continued my clinical path within the field of surgery. However, my interest in nuclear medicine remained. I first met Prof. Lars Edenbrandt when he acted as an opponent for my Master's thesis, and he introduced me to Dr. Elin Trägårdh. It resulted in the project leading up to this PhD thesis, which explores how artificial intelligence can be applied for automated quantitative analysis of tumor spread in PET/CT in patients with prostate cancer.

# References

1. Sung H, Ferlay J, Siegel R, Laversanne M, Soerjomataram I, Jemal A, et al. *Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries*. CA Cancer J Clin. 2021.
2. Center MM, Jemal A, Lortet-Tieulent J, Ward E, Ferlay J, Brawley O, et al. *International variation in prostate cancer incidence and mortality rates*. Eur Urol. 2012;61(6):1079-92.
3. Culp MB, Soerjomataram I, Efstathiou JA, Bray F, Jemal A. *Recent global patterns in prostate cancer incidence and mortality rates*. Eur Urol. 2020;77(1):38-52.
4. Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA. *The 2014 International Society of Urological Pathology (ISUP) consensus conference on Gleason grading of prostatic Carcinoma: Definition of grading patterns and proposal for a new Grading system*. Am J Surg Pathol. 2016;40(2):244-52.
5. Gleason DF. *Classification of prostatic carcinomas*. Cancer Chemother Rep. 1966;50(3):125-8.
6. Epstein JI, Zelefsky MJ, Sjoberg DD, Nelson JB, Egevad L, Magi-Galluzzi C, et al. *A Contemporary prostate cancer grading system: A validated alternative to the Gleason score*. Eur Urol. 2016;69(3):428-35.
7. Berney DM, Beltran L, Fisher G, North BV, Greenberg D, Møller H, et al. *Validation of a contemporary prostate cancer grading system using prostate cancer death as outcome*. Br J Cancer. 2016;114(10):1078-83.
8. Epstein JI, W. PA, Sauvageot J, Walsh PC. *Prediction of progression following radical prostatectomy. A multivariate analysis of 721 men with long-term follow-up*. Am J Surg Pathol. 1996;20(3):286-92.
9. Ferlay J, Ervik M, Lam F, Colombet M, Mery L, Piñeros M, et al. *Global Cancer Observatory: Cancer today*. Lyon, France: International Agency for Research on Cancer. Available from: <http://gco.iarc.fr/today>, accessed 18 October 2022.
10. *EAU guidelines*. Edn. presented at the EAU Annual Congress Amsterdam 2022. ISBN 978-94-92671-16-5.
11. de Rooij M, Hamoen EH, Witjes JA, Barentsz JO, Rovers MM. *Accuracy of magnetic resonance imaging for local staging of prostate cancer: A diagnostic meta-analysis*. Eur Urol. 2016;70(2):233-45.
12. Brierley JD, Gospodarowicz MK, Wittekind C (ed.). *TNM classification of malignant tumours, 8th Edition*. John Wiley & Sons, Chichester. 2017.
13. Eiber M, Herrmann K, Calais J, Hadaschik B, Giesel FL, Hartenbach M, et al. *Prostate cancer molecular imaging standardized evaluation (PROMISE): Proposed miTNM*

- classification for the interpretation of PSMA-ligand PET/CT.* J Nucl Med. 2018;59(3):469-78.
14. Mason MD, van der Kwast TH, Mottet N, Oprea-Lager DE, Rouvière O. *Modern imaging in prostate cancer: Do we treat patients, or their scans?* Eur Urol. 2022;81(4):319-22.
  15. Daneshmand S, Quek ML, Stein JP, Lieskovsky G, Cai J, Pinski J, et al. *Prognosis of patients with lymph node positive prostate cancer following radical prostatectomy: Long-term results.* J Urol. 2004;172(6):2252-5.
  16. Fossati N, Willemse PM, Van den Broeck T, van den Bergh RCN, Yuan CY, Briers E, et al. *The benefits and harms of different extents of lymph node dissection during radical prostatectomy for prostate cancer: A systematic review.* Eur Urol. 2017;72(1):84-109.
  17. Gandaglia G, Fossati N, Zaffuto E, Bandini M, Dell'Oglio P, Bravi CA, et al. *Development and internal validation of a novel model to identify the candidates for extended pelvic lymph node dissection in prostate cancer.* Eur Urol. 2017;72(4):632-40.
  18. Eifler JB, Feng Z, Lin BM, Partin MT, Humphreys EB, Han M, et al. *An updated prostate cancer staging nomogram (Partin tables) based on cases from 2006 to 2011.* BJU Int. 2013;111(1):22-9.
  19. Godoy G, Chong KT, Cronin A, Vickers A, Laudone V, Touijer K, et al. *Extent of pelvic lymph node dissection and the impact of standard template dissection on nomogram prediction of lymph node involvement.* Eur Urol. 2011;60(2):195-201.
  20. Roach M, Marquez C, Yuo HS, Narayan P, Coleman L, Nseyo UO, et al. *Predicting the risk of lymph node involvement using the pre-treatment prostate specific antigen and Gleason score in men with clinically localized prostate cancer.* Int J Radiat Oncol Biol Phys. 1994;28(1):33-7.
  21. Gandaglia G, Martini A, Ploussard G, Fossati N, Stabile A, De Visschere P, et al. *External validation of the 2019 Briganti nomogram for the identification of prostate cancer patients who should be considered for an extended pelvic lymph node dissection.* Eur Urol. 2020;78(2):138-42.
  22. Draulans C, Everaerts W, Isebaert S, Van Bruwaene S, Gevaert T, Oyen R, et al. *Development and external validation of a multiparametric magnetic resonance imaging and International Society of Urological Pathology based add-on prediction tool to identify prostate cancer candidates for pelvic lymph node dissection.* J Urol. 2020;203(4):713-18.
  23. Heesakkers RA, Hövels AM, Jager GJ, van den Bosch HCM, Witjes JA, Raat HPJ, et al. *MRI with a lymph-node-specific contrast agent as an alternative to CT scan and lymph-node dissection in patients with prostate cancer: A prospective multicohort study.* Lancet Oncol. 2008;9(9):850-6.
  24. Hövels AM, Heesakkers RAM, Adang EM, J. JG, Strum S, Hoogeveen YL, et al. *The diagnostic accuracy of CT and MRI in the staging of pelvic lymph nodes in patients with prostate cancer: A meta-analysis.* Clin Radiol. 2008;63(4):387-95.

25. Cook GJR, Goh V. *Molecular imaging of bone metastases and their response to therapy.* *J Nucl Med.* 2020;61(6):799-806.
26. Armstrong AJ, Szmulewitz RZ, Petrylak DP, Holzbeierlein J, Villers A, Azad A, et al. *ARCHES: A randomized, phase III study of androgen deprivation therapy with enzalutamide or placebo in men with metastatic hormone-sensitive prostate cancer.* *J Clin Oncol.* 2019;37(32):2974-986.
27. Davis ID, Martin AJ, Stockler MR, Begbie S, Chi KN, Chowdhury S, et al. *Enzalutamide with standard first-line therapy in metastatic prostate cancer.* *N Engl J Med.* 2019;381(2):121-31.
28. Chi KN, Agarwal N, Bjartell A, Chung BH, Pereira de Santana Gomes AJ, Given R, et al. *Apalutamide for metastatic, castration-sensitive prostate cancer.* *N Engl J Med.* 2019;381(1):13-24.
29. Fizazi K, Tran N, Fein L, Matsubara N, Rodriguez-Antolin A, Alekseev BY, et al. *Abiraterone plus prednisone in metastatic, castration-sensitive prostate cancer.* *N Engl J Med.* 2017;377(4):352-60.
30. James ND, Sydes MR, Clarke NW, Mason MD, Dearnaley DP, Spears MR, et al. *Addition of docetaxel, zoledronic acid, or both to first-line long-term hormone therapy in prostate cancer (STAMPEDE): Survival results from an adaptive, multiarm, multistage, platform randomised controlled trial.* *Lancet.* 2016;387(10024):1163-77.
31. Sweeney CJ, Chen YH, Carducci M, Liu G, Jarrard DF, Eisenberger M, et al. *Chemohormonal therapy in metastatic hormone-sensitive prostate cancer.* *N Engl J Med.* 2015;373(8):737-46.
32. Gravis G, Fizazi K, Joly F, Oudard S, Priou F, Esterni B, et al. *Androgen-deprivation therapy alone or with docetaxel in non-castrate metastatic prostate cancer (GETUG-AFU 15): A randomised, open-label, phase 3 trial.* *Lancet Oncol.* 2013;14(2):149-58.
33. Broughman JR, Fleming CW, Mian OY, Stephans KL, Tendulkar RD. *Management of oligometastatic prostate cancer.* *Appl Radiot Oncol.* 2020;9(3):6-10.
34. Boevé LMS, Hulshof MCCM, Vis AN, Zwinderman AH, Twisk JWR, Witjes WPJ, et al. *Effect on survival of androgen deprivation therapy alone compared to androgen deprivation therapy combined with concurrent radiation therapy to the prostate in patients with primary bone metastatic prostate cancer in a prospective randomised clinical trial: Data from the HORRAD trial.* *Eur Urol.* 2019;75(3):410-18.
35. Parker CC, James ND, Brawley CD, Clarke NW, Hoyle AP, Ali A, et al. *Radiotherapy to the primary tumour for newly diagnosed, metastatic prostate cancer (STAMPEDE): A randomised controlled phase 3 trial.* *Lancet.* 2018;392(10162):2353-66.
36. Gandaglia G, Abdollah F, Schiffmann J, Trudeau V, Shariat SF, Kim SP, et al. *Distribution of metastatic sites in patients with prostate cancer: A population-based analysis.* *Prostate.* 2014;74(2):210-6.

37. Zacho HD, Barsi T, Mortensen JC, Mogensen MK, Bertelsen H, Josephsen N, et al. *Prospective multicenter study of bone scintigraphy in consecutive patients with newly diagnosed prostate cancer.* Clin Nucl Med. 2014;39(1):26-31.
38. Svensson E, Christiansen CF, Ulrichsen SP, Rørth MR, Sørensen HT. *Survival after bone metastasis by primary cancer type: A Danish population-based cohort study.* BMJ Open. 2017;7(9):e016022.
39. Li Y, Schiepers C, Lake R, Dadparvar S, Berenji GR. *Clinical utility of <sup>18</sup>F-fluoride PET/CT in benign and malignant bone diseases.* Bone. 2012;50(1):128-39.
40. Shen G, Deng H, Hu S, Jia Z. *Comparison of choline-PET/CT, MRI, SPECT, and bone scintigraphy in the diagnosis of bone metastases in patients with prostate cancer: a meta-analysis.* Skeletal Radiol. 2014;43(11):1503-13.
41. Alqahtani MM, Fulton R, Constable C, Willowson KP, Kench PL. *Diagnostic performance of whole-body SPECT/CT in bone metastasis detection using <sup>99m</sup>Tc-labelled diphosphate: A systematic review and meta-analysis.* Clin Radiol. 2020;75(12):961.e11-24.
42. Zumsteg ZS, Spratt DE, Romesser PB, Pei X, Zhang Z, Polkinghorn W, et al. *The natural history and predictors of outcome following biochemical relapse in the dose escalation era for prostate cancer patients undergoing definitive external beam radiotherapy.* Eur Urol. 2015;67(6):1009-16.
43. Caire AA, Sun L, Ode O, Stackhouse DA, Maloney K, Donatucci C, et al. *Delayed prostate-specific antigen recurrence after radical prostatectomy: How to identify and what are their clinical outcomes?* Urology. 2009;74(3):643-7.
44. Freedland SJ, Humphreys EB, Mangold LA, Eisenberger M, Dorey FJ, Walsh PC, et al. *Risk of prostate cancer-specific mortality following biochemical recurrence after radical prostatectomy.* JAMA. 2005;294(4):433-9.
45. Pound CR, Partin AW, Eisenberger MA, Chan DW, Pearson JD, Walsh PC. *Natural history of progression after PSA elevation following radical prostatectomy.* JAMA. 1999;281(17):1591-7.
46. Beresford MJ, Gillatt D, Benson RJ, Ajithkumar T. *A systematic review of the role of imaging before salvage radiotherapy for post-prostatectomy biochemical recurrence.* Clin Oncol. 2010;22(1):46-55.
47. Kane CJ, Amling CL, Johnstone PA, Pak N, Lance RS, Thrasher JB, et al. *Limited value of bone scintigraphy and computed tomography in assessing biochemical failure after radical prostatectomy.* Urology. 2003;61:607-11.
48. Beer AJ, Eiber M, Souvatzoglou M, Schwaiger M, Krause BJ. *Radionuclide and hybrid imaging of recurrent prostate cancer.* Lancet Oncol. 2010;12(2):181-91.
49. Even-Sapir E, Keidar X, Bar-Shalom R. *Hybrid imaging (SPECT/CT and PET/CT)--improving the diagnostic accuracy of functional/metabolic and anatomic imaging.* Semin Nucl Med. 2009;39(4):264-75.
50. Lin EC, Alavi A (ed.). *PET and PET/CT: A clinical guide.* Thieme Publishers, New York, Stuttgart. 2019.

51. Hofer C, Laubenbacher C, Block T, Breul J, Hartung R, Schwaiger M. *Fluorine-18-fluorodeoxyglucose positron emission tomography is useless for the detection of local recurrence after radical prostatectomy*. Eur Urol. 1999;36(1):31-5.
52. Shreve PD, Grossman Hb, Gross MD, Wahl RL. *Metastatic prostate cancer: Initial findings of PET with 2-deoxy-2-[F-18]fluoro-D-glucose*. Radiology. 1996;199(3):751-6.
53. Effert PJ, Bares R, Handt S, Wolff JM, Büll U, Jakse G. *Metabolic imaging of untreated prostate cancer by positron emission tomography with 18 fluorine-labeled deoxyglucose*. J Urol. 1996;155(3):994-8.
54. Iagaru A, Mittra E, Dick DW, Gambhir SS. *Prospective evaluation of <sup>99m</sup>Tc MDP scintigraphy, <sup>18</sup>F NaF PET/CT, and <sup>18</sup>F FDG PET/CT for detection of skeletal metastases*. Mol Imaging Biol. 2012;14(2):252-9.
55. Malaspina S, Ettala O, Tolvanen T, Rajander J, Eskola O, Boström PJ, et al. *Flare on [<sup>18</sup>F]PSMA-1007 PET/CT after short-term androgen deprivation therapy and its correlation to FDG uptake: Possible marker of tumor aggressiveness in treatment-naïve metastatic prostate cancer patients*. Eur J Nucl Med Mol Imaging. Published online ahead of print 26 September 2022.
56. Morris MJ, Akhurst T, Larson SM, Ditullio M, Chu E, Siedlecki K, et al. *Fluorodeoxyglucose positron emission tomography as an outcome measure for castrate metastatic prostate cancer treated with antimicrotubule chemotherapy*. 2005;11(9):3210-6.
57. Hofman MS, Emmett L, Sandhu S, Irvani A, Joshua AM, Goh JC, et al. *[<sup>177</sup>Lu]Lu-PSMA-617 versus cabazitaxel in patients with metastatic castration-resistant prostate cancer (TheraP): A randomised, open-label, phase 2 trial*. Lancet. 2021;397(10276):797-804.
58. Ahuja K, Sotoudeh H, Galgano SJ, Singh R, Gupta N, Gaddamanugu S, et al. *<sup>18</sup>F-sodium fluoride PET: History, technical feasibility, mechanism of action, normal biodistribution, and diagnostic performance in bone metastasis detection compared with other imaging modalities*. J Nucl Med Technol. 2020;48(1):9-16.
59. Beheshti M, Mottaghy FM, Payche F, Behrendt FFF, Van den Wyngaert T, Fogelman I, et al. *<sup>18</sup>F-NaF PET/CT: EANM procedure guidelines for bone imaging*. Eur J Nucl Med Mol Imaging. 2015;42(11):1767-77.
60. Zacho HD, Jochumsen MR, Langkilde NC, Mortensen JC, Haarmark C, Hendel HW, et al. *No added value of <sup>18</sup>F-sodium fluoride PET/CT for the detection of bone metastases in patients with newly diagnosed prostate cancer with normal bone scintigraphy*. J Nucl Med. 2019;60(12):1713-16.
61. DeMarzo AM, Nelson WG, Isaacs WB, Epstein JI. *Pathological and molecular aspects of prostate cancer*. Lancet. 2003;361(9361):955-64.
62. Sweat SD, Pacelli A, Murphy GP, Bostwick DG. *Prostate-specific membrane antigen expression is greatest in prostate adenocarcinoma and lymph node metastases*. Urology. 1998;52(4):637-40.
63. Maurer T, Gschwend JE, Rauscher I, Souvatzoglou M, Haller B, Weirich G, et al. *Diagnostic efficacy of <sup>68</sup>Gallium-PSMA positron emission tomography compared to*



- conventional imaging for lymph node staging of 130 consecutive patients with intermediate to high risk prostate cancer.* J Urol. 2016;195(5):1436-43.
64. Perner S, Hofer MD, Kim R, Shah RB, Li H, Möller P, et al. *Prostate-specific membrane antigen expression as a predictor of prostate cancer progression.* Human Pathol. 2007;38(5):696-701.
  65. Silver DA, Pellicer I, Fair WR, Heston WD, Cordon-Cardo C. *Prostate-specific membrane antigen expression in normal and malignant human tissues.* Clin Cancer Res. 1997;3(1):81-5.
  66. Israeli RS, Powell CT, Corr JG, Fair WR, Heston WD. *Expression of the prostate-specific membrane antigen.* Cancer Res. 1994;54(7):1807-11.
  67. Hvittfeldt E, Bjöersdorff M, Brolin G, Minarik D, Svegborn SL, Oddstig J, et al. *Biokinetics and dosimetry of <sup>18</sup>F-PSMA-1007 in patients with prostate cancer.* Clin Physiol Funct Imaging. 2022;42(6):443-52.
  68. Rauscher I, Krönke M, König M, Gafita A, Maurer T, Horn T, et al. *Matched-Pair Comparison of <sup>68</sup>Ga-PSMA-11 PET/CT and <sup>18</sup>F-PSMA-1007 PET/CT: Frequency of pitfalls and detection efficacy in biochemical recurrence after radical prostatectomy.* J Nucl Med. 2020;61(1):51-7.
  69. Foley RW, Redman SL, Graham RN, Loughborough WW, Little D. *Fluorine-18 labelled prostate-specific membrane antigen (PSMA)-1007 positron-emission tomography-computed tomography: Normal patterns, pearls, and pitfalls.* Clin Radiol. 2020;75(12):903-13.
  70. Fanti S, Goffin K, Hadaschik BA, Herrmann K, Maurer T, MacLennan S, et al. *Consensus statements on PSMA PET/CT response assessment criteria in prostate cancer.* Eur J Nucl Med Mol Imaging. 2021;48(2):469-76.
  71. Hofman MS, Lawrentschuk N, Francis RJ, Tang C, Vela I, Thomas P, et al. *Prostate-specific membrane antigen PET-CT in patients with high-risk prostate cancer before curative-intent surgery or radiotherapy (proPSMA): A prospective, randomised, multicentre study.* Lancet. 2020;395(10231):1208-16.
  72. Perera M, Papa N, Roberts M, Williams M, Udovicich C, Vela I, et al. *Gallium-68 prostate-specific membrane antigen positron emission tomography in advanced prostate cancer—updated diagnostic utility, sensitivity, specificity, and distribution of prostate-specific membrane antigen-avid lesions: A systematic review and meta-analysis.* Eur Urol. 2020;77(4):403-17.
  73. Pienta KJ, Gorin MA, Rowe SP, Carroll PR, Pouliot F, Probst S, et al. *A phase 2/3 prospective multicenter study of the diagnostic accuracy of prostate specific membrane antigen PET/CT with <sup>18</sup>F-DCFPyL in prostate cancer patients (OSPREY).* J Urol. 2021;206(1):52-61.
  74. Gillessen S, Armstrong A, Attard G, Beer TM, Beltran H, Bjartell A, et al. *Management of patients with advanced prostate cancer: Report from the Advanced Prostate Cancer Consensus Conference 2021.* Eur Urol. 2022;82(1):115-41.

75. Anttinen M, Ettala O, Malaspina S, Jambor I, Sandell M, Kajander S, et al. *A prospective comparison of <sup>18</sup>F-prostate-specific membrane antigen-1007 positron emission tomography computed tomography, whole-body 1.5 T magnetic resonance imaging with diffusion-weighted imaging, and single-photon emission computed tomography/computed tomography with traditional imaging in primary distant metastasis staging of prostate cancer (PROSTAGE)*. Eur Urol Oncol. 2021;4(4):635-44.
76. Gafita A, Rauscher I, Weber M, Hadaschik B, Wang H, Armstrong WR, et al. *Novel framework for treatment response evaluation using PSMA-PET/CT in patients with metastatic castration-resistant prostate cancer (RECIP 1.0): An international multicenter study*. J Nucl Med. 2022;63(11):1651-58.
77. Gafita A, Rauscher I, Fendler WP, Murthy V, Hui W, Armstrong WR, et al. *Measuring response in metastatic castration-resistant prostate cancer using PSMA PET/CT: Comparison of RECIST 1.1, aPCWG3, aPERCIST, PPP, and RECIP 1.0 criteria*. Eur J Nucl Med Mol Imaging. 2022;49(12):4271-81.
78. Afshar-Oromieh A, Zechmann CM, Malcher A, Eisenhut M, Linhart HG, Holland-Letz T, et al. *Comparison of PET imaging with a <sup>68</sup>Ga-labelled PSMA ligand and <sup>18</sup>F-choline-based PET/CT for the diagnosis of recurrent prostate cancer*. Eur J Nucl Med Mol Imaging. 2014;41(1):11-20.
79. Yadav D, Hwang H, Qiao W, Upadhyay R, Chapin BF, Tang C, et al. *<sup>18</sup>F-fluciclovine versus PSMA PET imaging in primary tumor detection during initial staging of high-risk prostate cancer: A systematic review and meta-analysis*. Radiol Imaging Cancer. 2022;4(2):e210091.
80. Sandblom G, Ladjevardi S, Garmo H, Varenhorst E. *The impact of prostate-specific antigen level at diagnosis on the relative survival of 28,531 men with localized carcinoma of the prostate*. Cancer. 2008;112(4):813-9.
81. Sullivan DC, Obuchowski NA, Kessler LG, Raunig DL, Gatsonis C, Huang EP, et al. *Metrology standards for quantitative imaging biomarkers*. Radiology. 2015;277(3):813-25.
82. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, R. F, et al. *New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1)*. Eur J Cancer. 2009;45(2):228-47.
83. Scher HI, Morris MJ, Stadler WM, Higano C, Basch E, Fizazi K, et al. *Trial design and objectives for castration-resistant prostate cancer: Updated recommendations from the Prostate Cancer Clinical Trials Working Group 3*. J Clin Oncol. 2016;34(12):1402-18.
84. Imbriaco M, Larson SM, Yeung HW, Mawlawi OR, Erdi Y, Venkatraman ES, et al. *A new parameter for measuring metastatic bone involvement by prostate cancer: The bone scan index*. Clin Cancer Res. 1998;4(7):1765-72.
85. Ulmert D, Kaboteh R, Fox J, Savage C, Evans MJ, Lilja H, et al. *A novel automated platform for quantifying the extent of skeletal tumour involvement in prostate cancer patients using the bone scan index*. Eur Urol. 2012;26(1):78-84.

86. Anand A, Morris MJ, Kaboteh R, Reza M, Trägårdh E, Matsunaga N, et al. *A preanalytic validation study of automated Bone Scan Index: Effect on accuracy and reproducibility due to the procedural variabilities in bone scan image acquisition.* J Nucl Med. 2016;57(12):1865-71.
87. Anand A, Morris MJ, Kaboteh R, Båth L, Sadik M, Gjertsson P, et al. *Analytic validation of the automated bone scan index as an imaging biomarker to standardize quantitative changes in bone scans of patients with metastatic prostate cancer.* J Nucl Med. 2016;57(1):41-5.
88. Armstrong AJ, Anand A, Edenbrandt L, Bondesson E, Bjartell A, Widmark A, et al. *Phase 3 assessment of the automated bone scan index as a prognostic imaging biomarker of OS in men with metastatic castration-resistant prostate cancer: A secondary analysis of a randomized clinical trial.* JAMA oncology. 2018;4(7):944-51.
89. Wahl RL, Jacene H, Kasamon Y, Lodge MA. *From RECIST to PERCIST: Evolving considerations for PET response criteria in solid tumors.* J Nucl Med. 2009;50:122S-50S.
90. Ceci F, Oprea-Lager DA-O, Emmett L, Adam JA, Bomanji J, Czernin J, et al. *E-PSMA: the EANM standardized reporting guidelines v1.0 for PSMA-PET.* Eur J Nucl Med Mol Imaging. 2021;48:1626-38.
91. Fanti S, Hadaschik B, Herrmann K. *Proposal for systemic-therapy response-assessment criteria at the time of PSMA PET/CT imaging: The PSMA PET progression criteria.* J Nucl Med. 2020;62(5):678-82.
92. Perk T, Bradshaw T, Chen S, Im HJ, Cho S, Perlman S, et al. *Automated classification of benign and malignant lesions in <sup>18</sup>F-NaF PET/CT images using machine learning.* Phys Med Biol. 2018;63(22):225019.
93. Harmon SA, Perk T, Lin C, Eickhoff J, Choyke PL, Dahut WL, et al. *Quantitative assessment of early [<sup>18</sup>F]sodium fluoride positron emission tomography/computed tomography response to treatment in men with metastatic prostate cancer to bone.* J Clin Oncol. 2017;35(24):2829-37.
94. Lin C, Bradshaw T, Perk T, Harmon S, Eickhoff J, Jallow N, et al. *Repeatability of quantitative <sup>18</sup>F-NaF PET: A multicenter study.* J Nucl Med. 2016;57(12):1872-79.
95. Rohren EM, Etchebehere EC, Araujo JC, Hobbs BP, Swanston NM, Everding M, et al. *Determination of skeletal tumor burden on <sup>18</sup>F-fluoride PET/CT.* J Nucl Med. 2015;56(10):1507-12.
96. Hammes J, Täger P, Drzezga A. EBONI: *A tool for automated quantification of bone metastasis load in PSMA PET/CT.* J Nucl Med. 2018;59(7):1070-75.
97. Bieth M, Krönke M, Tauber R, Dahlbender M, Retz M, Nekolla SG, et al. *Exploring new multimodal quantitative imaging indices for the assessment of osseous tumor burden in prostate cancer using <sup>68</sup>Ga-PSMA PET/CT.* J Nucl Med. 2017;58(10):1632-37.
98. Seifert R, Herrmann K, Kleesiek J, Schäfers M, Shah V, Xu Z, et al. *Semiautomatically quantified tumor volume using <sup>68</sup>Ga-PSMA-11 PET as a biomarker for survival in patients with advanced prostate cancer.* J Nucl Med. 2020;61(12):1786-92.

99. Gafita A, Bieth M, Kronke M, Tetteh G, Navarro F, Wang H, et al. *qPSMA: Semiautomatic software for whole-body tumor burden assessment in prostate cancer using <sup>68</sup>Ga-PSMA11 PET/CT*. J Nucl Med. 2019;60(9):1277-83.
100. Brito AET, Mourato FA, de Oliveira RPM, Leal ALG, Filho PJA, de Filho JLL. *Evaluation of whole-body tumor burden with <sup>68</sup>Ga-PSMA PET/CT in the biochemical recurrence of prostate cancer*. Ann Nucl Med. 2019;33:344-50.
101. Grubmüller B, Senn D, Kramer G, Baltzer P, D'Andrea D, Grubmüller KH, et al. *Response assessment using <sup>68</sup>Ga-PSMA ligand PET in patients undergoing <sup>177</sup>Lu-PSMA radioligand therapy for metastatic castration-resistant prostate cancer*. Eur J Nucl Med Mol Imaging. 2019;46(5):1063-72.
102. Schmuck S, von Klot CA, Henkenberens C, Sohns JM, Christiansen H, Wester HJ, et al. *Initial experience with volumetric <sup>68</sup>Ga-PSMA I&T PET/CT for assessment of whole-body tumor burden as a quantitative imaging biomarker in patients with prostate cancer*. J Nucl Med. 2017;58(12):1962-68.
103. Seifert R, Kessel K, Schlack K, Weber M, Herrmann K, Spanke M, et al. *PSMA PET total tumor volume predicts outcome of patients with advanced prostate cancer receiving [<sup>177</sup>Lu]Lu-PSMA-617 radioligand therapy in a bicentric analysis*. Eur J Nucl Med Mol Imaging. 2021;48(4):1200-10.
104. Lim I, Lindenberg ML, Mena E, Verdini N, Shih JH, Mayfield C, et al. *<sup>18</sup>F-Sodium fluoride PET/CT predicts overall survival in patients with advanced genitourinary malignancies treated with cabozantinib and nivolumab with or without ipilimumab*. Eur J Nucl Med Mol Imaging. 2020;47(1):178-84.
105. Grubmüller B, Rasul S, Baltzer P, Fajkovic H, D'Andrea D, Berndt F, et al. *Response assessment using [<sup>68</sup>Ga]Ga-PSMA ligand PET in patients undergoing systemic therapy for metastatic castration-resistant prostate cancer*. Prostate. 2020;80(1):74-82.
106. Schmidkonz CA-O, Cordes M, Schmidt D, Bäuerle T, Goetz TI, Beck M, et al. *<sup>68</sup>Ga-PSMA-11 PET/CT-derived metabolic parameters for determination of whole-body tumor burden and treatment response in prostate cancer*. Eur J Nucl Med Mol Imaging. 2018;45(11):1862-72.
107. Etchebehere EC, Araujo JC, Milton DR, Erwin WD, Wendt RE, Swanston NM, et al. *Skeletal tumor burden on baseline <sup>18</sup>F-fluoride PET/CT predicts bone marrow failure after <sup>223</sup>Ra therapy*. Clin Nucl Med. 2016;41(4):268-73.
108. Etchebehere EC, Araujo JC, Fox PS, Swanston NM, Macapinlac HA, Rohren EM. *Prognostic factors in patients treated with <sup>223</sup>Ra: The role of skeletal tumor burden on baseline <sup>18</sup>F-fluoride PET/CT in predicting overall survival*. J Nucl Med. 2015;56(8):1177-84.
109. Kendrick J, Francis RJ, Hassan GM, Rowshanfarzad P, Ong JSL, Ebert MA. *Fully automatic prognostic biomarker extraction from metastatic prostate lesion segmentations in whole-body [<sup>68</sup>Ga]Ga-PSMA-11 PET/CT images*. Eur J Nucl Med Mol Imaging. 2022. Published online ahead of print 17 August 2022.

110. Johnsson K, Brynolfsson J, Sahlstedt H, Nickols NG, Rettig M, Probst S, et al. *Analytical performance of aPROMISE: automated anatomic contextualization, detection, and quantification of [<sup>18</sup>F]DCFPyL (PSMA) imaging for standardized reporting.* Eur J Nucl Med Mol Imaging. 2021;49:1041-51.
111. Zhao Y, Gafita A, Vollnberg B, Tetteh G, Haupt F, Afshar-Oromieh A, et al. *Deep neural network for automatic characterization of lesions on <sup>68</sup>Ga-PSMA-11 PET/CT.* Eur J Nucl Med Mol Imaging. 2020;47(3):603-13.
112. Jarrett D, Stride E, Vallis K, Gooding MJ. *Applications and limitations of machine learning in radiation oncology.* Br J Radiol. 2019;92(1100):20190001.
113. Yamashita R, Nishio M, Do RKG, Togashi K. *Convolutional neural networks: An overview and application in radiology.* Insights Imaging. 2018;9(4):611-29.
114. Sarvamangala DR, Kulkarni RV. *Convolutional neural networks in medical image understanding: a survey.* Evol Intell. 2022;15(1):1-22.
115. Soffer S, Ben-Cohen A, Shimon O, Amitai MM, Greenspan H, Klang E. *Convolutional neural networks for radiologic images: A radiologist's guide.* Radiology. 2019;290(3):590-606.
116. Mazurowski MA, Buda M, Saha A, Bashir MR. *Deep learning in radiology: An overview of the concepts and a survey of the state of the art with focus on MRI.* J Magn Reson Imaging. 2018;49(4):939-54.
117. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. *3D U-Net: Learning dense volumetric segmentation from sparse annotation.* International Conference on Medical Image Computing and Computer Assisted Intervention: arXiv; 2016. pp. 424-32.
118. Toriihara A, Nobashi T, Baratto L, Duan H, Moradi F, Park S, et al. *Comparison of 3 interpretation criteria for <sup>68</sup>Ga-PSMA11 PET based on inter- and intrareader agreement.* J Nucl Med. 2020;61(4):533-39.
119. Derwael C, Lavergne O, Lovinfosse P, Nechifor V, Salve M, Waltregny D, et al. *Interobserver agreement of [<sup>68</sup>Ga]Ga-PSMA-11 PET/CT images interpretation in men with newly diagnosed prostate cancer.* EJNMMI Res. 2020;10(1):15.
120. Violet J, Jackson P, Ferdinandus J, Sandhu S, Akhurst T, Irvani A, et al. *Dosimetry of <sup>177</sup>Lu-PSMA-617 in metastatic castration-resistant prostate cancer: Correlations between pretherapeutic imaging and whole-body tumor dosimetry with treatment outcomes.* J Nucl Med. 2019;60(4):517-23.
121. Poulsen MH, Petersen H, Høiland-Carlsen PF, Jakobsen JS, Gerke O, Karstoft J, et al. *Spine metastases in prostate cancer: comparison of technetium-99m-MDP whole-body bone scintigraphy, [<sup>18</sup>F]choline positron emission tomography(PET)/computed tomography (CT) and [<sup>18</sup>F]NaF PET/CT.* BJU Int. 2014;114(6):818-23.
122. Kjölhede H, Ahlgren G, Almquist H, Liedberg F, Lyttkens K, Ohlsson T, et al. *Combined <sup>18</sup>F-fluorocholine and <sup>18</sup>F-fluoride positron emission tomography/computed*

- tomography imaging for staging of high-risk prostate cancer.* BJU Int. 2012;110(10):1501-6.
123. Mosavi F, S. J, Sandberg D, Turesson I, Sörensen J, Ahlström H. *Whole-body diffusion-weighted MRI compared with <sup>18</sup>F-NaF PET/CT for detection of bone metastases in patients with high-risk prostate carcinoma.* AJR Am J Roentgenol. 2012;199(5):1114-20.
  124. Trägårdh E, Minarik D, Brodin G, Bitzén U, Olsson B, Oddstig J. *Optimization of [<sup>18</sup>F]PSMA-1007 PET-CT using regularized reconstruction in patients with prostate cancer.* EJNMMI Phys. 2020;7(1):31.
  125. *Report of the task group on reference man.* ICRP Publication 23. Pergamon Press, Oxford. 1975.
  126. Ceci F, Oprea-Lager DE, Emmett L, Adam JA, Bomanji J, Czernin J, et al. *E-PSMA: the EANM standardized reporting guidelines v1.0 for PSMA-PET.* Eur J Nucl Med Mol Imaging. 2021;48(5):1626-38.
  127. Top A, Hamarneh G, Abugharbich R. *Spotlight: Automated confidence-based user guidance for increasing efficiency in interactive 3D image segmentation.* In: Menze B, Langs G, Tu Z, Criminisi A, editors. Medical Computer Vision Recognition Techniques and Applications in Medical Imaging MCV 2010 Lecture Notes in Computer Science; 2011; Berlin, Heidelberg: Springer 2011. pp. 204-13.
  128. Trägårdh E, Borrelli P, Kaboteh R, Gillberg T, Ulen J, Enqvist O, et al. *RECOMIA-a cloud-based platform for artificial intelligence research in nuclear medicine and radiology.* EJNMMI Phys. 2020;7(1):51.
  129. Sadik M, Hamadeh I, Nordblom P, Suurkula M, Höglund P, Ohlsson M, et al. *Computer-assisted interpretation of planar whole-body bone scans.* J Nucl Med. 2008;49(12):1958-65.
  130. Dice L. *Measures of the amount of ecological association between species.* Ecology. 1945;26:297-302.
  131. Capobianco N, Sibille L, Chantadisai M, Gafita A, Langbein T, Platsch G, et al. *Whole-body uptake classification and prostate cancer staging in <sup>68</sup>Ga-PSMA-11 PET/CT using dual-tracer learning.* Eur J Nucl Med Mol Imaging. 2022;49(2):517-26.
  132. Oldan JD, Hawkins AS, Chin BB. *<sup>18</sup>F sodium fluoride PET/CT in patients with prostate cancer: Quantification of normal tissues, benign degenerative lesions, and malignant lesions.* World J Nucl Med. 2016;15(2):102-8.
  133. Oddstig J, Leide Svegborn S, Almquist H, Bitzen U, Garpered S, Hedeer F, et al. *Comparison of conventional and Si-photomultiplier-based PET systems for image quality and diagnostic performance.* BMC Med Imaging. 2019;19:81.
  134. Karki K, Saraiya S, Hugo GD, Mukhopadhyay N, Jan N, Schuster J, et al. *Variabilities of magnetic resonance imaging-, computed tomography-, and positron emission tomography-computed tomography-based tumor and lymph node delineations for lung cancer radiation therapy planning.* Int J Radiat Oncol Biol Phys. 2017;99(1):80-9.

135. Briganti A, Larcher A, Abdollah F, Capitanio U, Gallina A, Suardi N, et al. *Updated nomogram predicting lymph node invasion in patients with prostate cancer undergoing extended pelvic lymph node dissection: the essential importance of percentage of positive cores.* Eur Urol. 2012;61(3):480-7.
136. Trägårdh E, Enqvist O, Ulén J, Jögi J, Bitzén U, Hedeer F, et al. *Freely available, fully automated AI-based analysis of primary tumour and metastases of prostate cancer in whole-body [<sup>18</sup>F]-PSMA-1007 PET-CT.* Diagnostics. 2022;12(9):2101.
137. Jadvar H, Desai B, Conti PS. *Sodium <sup>18</sup>F-fluoride PET/CT of bone, joint, and other disorders.* Semin Nucl Med. 2015;45(1):58-65.
138. Vollnberg B, Alberts I, Genitsch V, Rominger A, Afshar-Oromieh A. *Assessment of malignancy and PSMA expression of uncertain bone foci in [<sup>18</sup>F]PSMA-1007 PET/CT for prostate cancer—a single-centre experience of PET-guided biopsies.* Eur J Nucl Med Mol Imaging. 2022;49(1):3910-16.
139. Aggarwal R, Wei X, Kim W, Small EJ, Ryan CJ, Carroll P, et al. *Heterogeneous Flare in prostate-specific membrane antigen positron emission tomography tracer uptake with initiation of androgen pathway blockade in metastatic prostate cancer.* EUr Urol Oncol. 2018;1(1):78-82.
140. Hope TA, Aggarwal R, Chee B, Tao D, Greene KL, Cooperberg MR, et al. *Impact of <sup>68</sup>Ga-PSMA-11 PET on management in patients with biochemically recurrent prostate cancer.* J Nucl Med. 58(12):1956-61.
141. Afshar-Oromieh A, Debus N, Uhrig M, Hope TA, Evans MJ, Holland-Letz T, et al. *Impact of long-term androgen deprivation therapy on PSMA ligand PET/CT in patients with castration-sensitive prostate cancer.* Eur J Nucl Med Mol Imaging. 2018;45(12):2045-54.
142. Zacho HD, Petersen LJ. *Bone flare to androgen deprivation therapy in metastatic, hormone-sensitive prostate cancer on <sup>68</sup>Ga-prostate-specific membrane antigen PET/CT.* Clin Nucl Med. 2018;43(11):404-6.
143. Liu T, Wu L, Fulton M, Johnson JM, Berkman CE. *Prolonged androgen deprivation leads to downregulation of androgen receptor and prostate-specific membrane antigen in prostate cancer cells.* Int J Oncol. 41(6):2087-92.
144. Chakraborty PS, Tripathi T, Agarwal KK, Kumar R, Vijay MK, Bal C. *Metastatic poorly differentiated prostatic carcinoma with neuroendocrine differentiation: Negative on <sup>68</sup>Ga-PSMA PET/CT.* Clin Nucl Med. 40(2):163-6.
145. Usmani S, Ahmed N, Marafi F, Rasheed R, Amanguno HG, Al Kandari F. *Molecular imaging in neuroendocrine differentiation of prostate cancer: <sup>68</sup>Ga-PSMA versus <sup>68</sup>Ga-DOTA NOC PET-CT.* Clin Nucl Med. 2017;42(5):410-13.









**FACULTY OF  
MEDICINE**

Department of Translational Medicine

Lund University, Faculty of Medicine  
Doctoral Dissertation Series 2022:163  
ISBN 978-91-8021-325-7  
ISSN 1652-8220

