Diagnostic accuracy of imaging methods for the diagnosis of skeletal malignancies held up against a pathology-proven gold standard reference

Lange, Mine Benedicte

DOI (link to publication from Publisher):
10.54337/aau528831676

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

DIAGNOSTIC ACCURACY OF IMAGING METHODS FOR THE DIAGNOSIS OF SKELETAL MALIGNANCIES HELD UP AGAINST A PATHOLOGY-PROVEN GOLD STANDARD REFERENCE

BY
MINE BENEDICTE LANGE

DISSERTATION SUBMITTED 2022

AALBORG UNIVERSITY
DENMARK

DIAGNOSTIC ACCURACY OF IMAGING METHODS FOR THE DIAGNOSIS OF SKELETAL MALIGNANCIES HELD UP AGAINST A PATHOLOGY-PROVEN GOLD STANDARD REFERENCE

BY
MINE BENEDICTE LANGE

AALBORG UNIVERSITY
DENMARK

DISSERTATION SUBMITTED 2022

Dissertation submitted: December 2022

PhD supervisor: Helle Damgaard Zacho
Professor (chair), MD, PhD, DMSc
Department of Nuclear Medicine,
Clinical Cancer Research Center,
Aalborg University Hospital, Aalborg, Denmark

Assistant PhD supervisors: Lars Jelstrup Petersen
Professor, MD, DMSc, CBA
Department of Nuclear Medicine,
Clinical Cancer Research Center,
Aalborg University Hospital, Aalborg, Denmark

Michael Bachmann Nielsen,
Professor, MD, DMSc, PhD
Department of Biomedical Sciences,
University of Copenhagen, Copenhagen, Denmark
Department of Diagnostic Radiology, Rigshospitalet,
Copenhagen, Denmark

PhD committee: Professor Jens Brøndum Frøkjær (chair)
Aalborg University, Denmark

Professor Nils Vetti
Haukeland University Hospital, Norway

Senior Lecturer John D Thompson
University Hospitals of Morecambe Bay –
NHS Foundation Trust, UK

PhD Series: Faculty of Medicine, Aalborg University
Department: Department of Clinical Medicine
ISSN (online): 2246-1302
ISBN (online): 978-87-7573-762-8

Published by:
Aalborg University Press
Kroghstræde 3
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk

Printed in Denmark by Stibo Complete, 2023

Preface

The discipline of diagnostic imaging was introduced to me late in life. Having spent more than two decades in pharmaceutical and medico-technical industry, the decision of completing an education as a specialist in diagnostic radiology has been an educational journey, especially due to the impressive technological development that has taken place in recent years, driven by devoted radiologists, researchers, and clinicians.

Getting scientifically carried away by such new technological advances is tempting. However, it is indeed also necessary to immerse yourself in solidly documenting what has developed to be indisputable truths. The overall purpose of this Ph.D. is to provide state-of-the-art documentation of diagnostic accuracy of imaging methods for the diagnosis of skeletal malignancies held up against a pathology-proven gold standard reference. We wanted to document what readers of diagnostic imaging consider to be facts, namely that some imaging modalities have higher diagnostic accuracy than others for the detection of bone metastases and that comparing current diagnostic imaging with those previously performed inevitably will improve the diagnostic quality. The uniqueness of our work is the systematic use of a reference gold standard, which we have been able to document in fact can be considered a gold standard.

Radiologists work closely together with other specialties, especially with experts from our sister field, nuclear medicine. Whereas radiologists primarily focus on anatomy and especially pathological anatomy, the focus in nuclear medicine is on the physiology and pathophysiology of organs and tissue. A close collaboration between our specialties is therefore crucial for a complete diagnosis and to me, it has indeed been a great pleasure and honor in this Ph.D. dissertation to work so closely with outstanding experts in both fields.

I hope that this Ph.D. can help provide a solid platform based upon which future research projects and imaging guidelines for early and precise detection of bone metastases can be prepared.

Acknowledgment

Sometimes you meet people of passion, people who constantly inspire and motivate others. During my education as a radiologist at Aalborg University Hospital, I was looking for someone who wanted to start a research project with me, preferably in the musculoskeletal area since that was my main interest. Out of the blue, I got a call from a person at that time unknown to me and that was the beginning of a long and inspiring collaboration with Professor Lars Jelstrup Petersen, Department of Nuclear Medicine, who was bubbling with ideas. Validation and quality assurance were the key questions, what do we know about what we consider to be the truth, is there evidence? Lars came up with the idea of collecting a series of bone biopsies based upon which we should investigate the accuracy and biases of imaging diagnostics and at the same time validate the assumption of bone biopsy as being the gold standard for detection of bone metastases. I was excited and honored.

Lars introduced me to Mads Lausen Nielsen, a student in Medicine with Industrial Specialization, with whom I should experience the great pleasure of collaboration. When I moved back to Copenhagen, Mads continued working on the material to complete a Ph.D. but when his passion for basic research dragged him into another Ph.D. position, I got a second life-changing call from Lars, this time asking if I could be interested in completing the Ph.D. on long-distance alongside my full-time job as a senior consultant at the Radiology Department, Hospital of North Zealand, Hilleroed. As usual, it was not possible to resist Lars´ excitement and enthusiasm. Thank you, Lars, for your trust in me.

As his right-hand Lars introduced me to Professor Helle Damgaard Zacho, who later in the process took over the role of being my principal supervisor with Lars as co-supervisor. A warm thank you to Helle, who handled a not-quite-easy task of having to mentor an elderly, long-term, distant radiologist Ph.D. student. I am more than grateful to you for taking on this task. Without your patience, forbearance, great experience, persistence, and always friendly attitude, this Ph.D. would never have been completed.

Placed in Copenhagen with two distant nuclear medicine supervisors, I was looking for a local co-supervisor, preferable a radiologist, so I called the man on the hill, Professor Michael Bachmann Nielsen, Department of Radiology, National University Hospital of Rigshospitalet, Copenhagen. Thank you so much, Michael, for accepting me and especially for including me in your immense network on so many levels. Your always accurate and sharp feedback to me has been highly appreciated. And also, a warm thank you to my fellow Ph.D. co-students for inspiration and collaboration.

Thank you to Professor Mogens Vyberg, Institute of Pathology, Cancer Research Center, Aalborg University Hospital, for contributing with your great expertise in our work with collecting and analyzing bone biopsies.

Thanks to all the people I have met on my many journeys through life and who without prejudice have encouraged and supported me, when my restlessness and gut feeling have forced me out and away, on new adventures. Thanks to the few, whom I call my good friends.

Last, but not least, thanks to my beloved family who have had to live with my periodical absences and absent-mindedness. Thanks to my husband for his presence, support, companionship, laughter, and love.

Table of Content

Abbreviations

ADC: Apparent Diffusion Coefficient (MRI)

BMD: bone metastatic disease

BS: bone scintigraphy

CT: computed tomography

DWI: diffusion weighted imaging (MRI sequence)

FDG: 18F-fluorodeoxyglucose

GI: Gastrointestinal

HU: Hounsfield Units (linear scale of tissue density measurement mainly used in CT)

MR or MRI: magnetic resonance imaging

MM: Multiple Myeloma

NSCLC: Non-Small Cell Lung Carcinoma

PET/CT: positron emission tomography

PSMA: Prostate-specific membrane antigen

RCC: Renal Cell Carcinoma

SCLC: Small Cell Lung Carcinoma

SNOMED: systematized nomenclature of medicine

SPECT: single-photon emission computed tomography

US: ultrasound

Dansk resume

Introduktion

Metastasering til knogler er den tredje hyppigste lokalisation næst efter lunge og lever. Prostata- og brystkræft udgør op til 70 % af knoglemetastaserne og ca. 75 % af de patienter, der lider af de mest almindelige metastaserende kræftformer, udvikler mindst en knoglemetastase. Tidlig og præcis diagnostik af knoglemetastaser er afgørende for iværksættelse af lindrende og helbredende behandling og dermed for sygelighed og død.

Billeddiagnostiske undersøgelser spiller en afgørende rolle for at kunne opdage mulige knoglemetastaser så tidligt som muligt. Flere billeddiagnostiske modaliteter kan anvendes, herunder konventionel røntgen, CT, MR, knoglescintigrafi-SPECT/CT samt PET/CT (positron emission tomografi) med forskellige radioaktivt mærkede sporstoffer. En række videnskabelige undersøgelser har forsøgt at belyse den diagnostiske træfsikkerhed af disse billeddiagnostiske metoder, men ingen af disse studier har været konsistente i valget af entydig reference, blandt hvilke biopsi fra den mistænkte knogleforandring anses for at være guldstandard.

Formål

Formålet med denne Ph.d. er at tilvejebringe valid dokumentation på den diagnostiske træfsikkerhed for relevante billeddiagnostiske metoder til påvisning af knoglemetastaser udelukkende ved brug af biopsi som guld standard. Samtidig ønskede vi at undersøge hvorvidt denne træfsikkerhed blev påvirket af adgang til tidligere billeddiagnostiske undersøgelser og rapporter samt hvorvidt rækkefølgen af disse kunne have en betydning. Endelig ønskede vi at dokumentere at en knogle bioptisk diagnose rent faktisk kan betragtes som guldstandard. Vi mener, at tilvejebringelsen af sådanne data kan have betydning for planlægning af fremtidige billeddiagnostiske og onkologiske studier samt for udarbejdelsen af evidensbaserede billeddiagnostiske retningslinjer for udredning af knoglemetastaser.

Materiale og metode

Ph.d.-studierne er udført som en retrospektiv konsekutiv kohorteundersøgelse af knoglebioptisk materiale indsamlet via Aalborg Universitetshospitals Patologiske Afdeling ved computersøgning af bioptisk materiale registreret i SNOMED (Systematized Nomenclature of Medicine) som T10* og T11* koder for henholdsvis cytologiske og histologiske knogle biopsier i perioden fra 1. januar 2011 til 31. juli 2013. Hver enkelt biopsi blev identificeret med et unikt dansk cpr-nummer. Med udgangspunkt i denne kohorte har vi udført 3 studier.

Det første studie havde til formål at undersøge den diagnostiske træfsikkerhed af relevante billeddiagnostiske metoder udført indenfor en seks måneders periode før der blev foretaget en knoglebiopsi af den pågældende metastasesuspekte læsion. Vi ønskede samtidig at analysere træfsikkerheden på undergrupper baseret på lokalisation og karakteristika af læsionerne, dvs. om

de var sklerotiske, lytiske eller blandede. Vi inkluderede 409 biopsier fra 395 patienter og diagnosen på de billeddiagnostiske undersøgelser (røntgen, CT, MR, knogle scintigrafi, SPECT/CT og FDG-PET/CT) blev gennemgået og holdt op mod biopsidiagnosen. Resultaterne blev udtrykt som sensitivitet, specificitet, positive og negative prædiktive værdier samt en samlet diagnostisk nøjagtighed.

I det andet studie indgik 216 knoglebiopsier med mindst 2 forskellige billeddiagnostiske undersøgelser i 6 måneders perioden før knoglebiopsien. Den billeddiagnostiske træfsikkerhed af de indgåede undersøgelser med eller uden en anden forudgående billeddiagnostik og den mulige indflydelse af sekvensen af disse blev undersøgt.

Formålet med det tredje studie var at undersøge hvorvidt en biopsi diagnose rent faktisk kan betragtes som guldstandard, dvs. hvorvidt en benign biopsi diagnose reelt er benign. Vi fulgte op på 215 benigne knoglebiopsier fra 207 patienter i to år efter den første benigne biopsi diagnose ved at se på diagnoserne fra samme anatomiske lokalisation fra yderligere biopsier og/eller billeddiagnostiske undersøgelser og/eller klinisk information for derigennem at kunne kategorisere den første biopsi som reelt benign, reelt malign eller tvetydig.

Resultater

Det første studie viste at sensitiviteten af MR og FDG-PET/CT var signifikant bedre end CT, som havde en bedre specificitet; generelt var disse modaliteter signifikant mere træfsikre end røntgen og knoglescintigrafi. Sensitiviteten for osteolytiske og blandede læsioner var signifikant højere for MR og PET/CT end for CT, hvilket ikke var tilfældet for osteosklerotiske læsioner. Ved rygsøjle læsioner viste MR den signifikant bedste sensitivitet, hvilket ikke var tilfældet for læsioner i de øvrige dele af skelettet.

I det andet studie kunne vi ikke dokumentere en signifikant forskel på den diagnostiske træfsikkerhed af de billeddiagnostiske undersøgelser, uanset om de var forudgået af en anden modalitet eller ej. Sekvensanalyserne indikerede, at jo større diagnostisk nøjagtighed en given modalitet antages at have, desto større er risikoen for at påvirke nøjagtigheden ved analyse af efterfølgende billedmodaliteter.

I det tredje studie dokumenterede vi at en benign knoglebiopsi kan betragtes som et validt kriterium for fravær af knoglemetastaser, idet 98 % af de benigne biopsier viste sig at være reelt benigne 2 år efter den første biopsi. To biopsier var falsk negative og tre kunne ikke med sikkerhed rubriceres som maligne eller benigne pga. manglende beskrivelse af de billeddiagnostiske undersøgelser.

Konklusion

Denne Ph.d.-afhandling har for første gang dokumenteret at MR og PET/CT er de bedste billeddiagnostiske undersøgelser til detektion af knoglemetastaser holdt op med guldstandarden, knoglebiopsi, som vi samtidig har dokumenteret rent faktisk kan betragtes som guldstandard. Vi

har ikke kunnet vise at det har nogen signifikant betydning for den diagnostiske træfsikkerhed hvorvidt der er adgang til at jo større diagnostisk nøjagtighed en given modalitet antages at have, desto større er risikoen for at påvirke nøjagtigheden ved analyse af efterfølgende billedmodaliteter.

English summary

Introduction

Bone is the third most common site of metastases after lung and liver. Prostate and breast cancer account for up to 70% of bone metastases and approximately 75% of patients suffering from the most common metastatic cancer diseases develop at least one bone metastasis. Early and accurate detection of bone metastases is essential for palliative and curative treatment and thus for morbidity and mortality.

Correct diagnostic imaging is essential for the early detection of possible bone metastases and several imaging modalities can be used, including conventional X-ray, CT (computed tomography), MRI (magnetic resonance imaging), bone scintigraphy-SPECT/CT, and PET/CT with various radiolabeled ligands. Several scientific studies have tried to document the diagnostic accuracy of these imaging methods, but none of these studies have been consistent in the choice of an unambiguous reference, among which biopsy from the suspected bone lesion is considered to be the gold standard.

Aim

The aim of this Ph.D. was to provide valid evidence on the diagnostic accuracy of relevant imaging methods for detection of bone metastases against pathology proven reference only. At the same time, we wanted to investigate whether diagnostic accuracy was affected by access to and the sequence of previous diagnostic imaging and reports. Finally, we wanted to document that bone biopsy is the true gold standard. We believe that such data are important for planning future imaging and oncological studies as well as for the development of evidence-based imaging guidelines for the detection of bone metastases.

Material and method

The Ph.D. studies were performed as retrospective consecutive cohort studies of bone biopsy material collected via University Hospital of Aalborg, Pathology Department, by a computer search of bioptic material registered in SNOMED (Systematized Nomenclature of Medicine) as T10 * and T11 * codes for cytological and histological bone biopsies, respectively, in the period from 1 January 2011 to 31 July 2013. Each biopsy was identified by a unique Danish civil registration number. Based on this cohort, we performed three studies.

In the first study, we wanted to examine the diagnostic accuracy of imaging methods performed within six months prior to a bone biopsy of the metastasis-suspected lesion. We also wanted to investigate the diagnostic accuracy of subgroups based upon the location and characteristics of the lesions, being either sclerotic, lytic, or mixed. We included 409 biopsies from 395 patients and the diagnosis on the imaging examinations (X-ray, CT, MRI, bone scintigraphy, SPECT/CT, FDG-PET/CT, and ultrasound) were reviewed and held up against the pathology proven reference. The

results were expressed as sensitivity, specificity, positive and negative predictive values as well as overall diagnostic accuracy.

The second study included 216 bone biopsies with at least 2 different imaging studies performed in the 6-month period before the bone biopsy. The imaging accuracy of the modalities with or without another prior different imaging and the possible influence of the sequence of these were examined.

The purpose of the third study was to investigate whether a biopsy diagnosis can be considered a gold standard, i.e., whether a benign biopsy diagnosis is truly benign. We did so by following up on 215 benign bone biopsies from 207 patients for a two-year period after the first benign biopsy diagnosis by looking at the diagnoses of any new biopsies and/or the diagnoses from any new imaging studies performed on the same anatomical location and/or any clinical information to categorize the first biopsy as true benign, true malignant or ambiguous.

Results

The first study documented that the sensitivity of MRI and PET/CT were significantly better than CT, which demonstrated a significantly better specificity, and in general, these modalities were significantly more accurate than X-ray and bone scintigraphy. The sensitivity for osteolytic and mixed lesions was significantly higher for MRI and PET/CT than for CT, which was not the case for osteosclerotic lesions. In spinal lesions, MRI showed the significantly best sensitivity, which was not the case for lesions in other parts of the skeleton.

In the second study, we were unable to document a significant difference in the diagnostic accuracy of the imaging studies, whether the reader had prior access to a different modality or not. The sequence analyzes indicated that the greater the diagnostic accuracy a given modality is assumed to have, the greater the risk of influencing the accuracy when analyzing subsequent imaging modalities.

In the third study, we documented that a benign bone biopsy can be considered a valid criterion for the absence of bone metastases since 98% of the benign biopsies were found to be truly benign two years after the first biopsy. Two biopsies were falsely negative and three could not be classified as malignant or benign due to a lack of description of the diagnostic imaging studies.

Conclusion

This Ph.D. thesis has for the first time documented that MRI and PET/CT are significantly the best imaging modalities for detection of bone metastases when systematically held up against the gold standard, bone biopsy, which we at the same time have documented as a valid gold standard. We have not been able to show any significant difference in diagnostic accuracy whether the reader has access to a previously performed different imaging or not, but it seems to be a source of error if a previously incorrect imaging diagnosis is performed on a modality, considered to have higher accuracy than the one currently performed.

List of papers

This Ph.D. is based on 3 studies all being published. In this Ph.D. thesis, the studies will be referred to as studies (1), (2), and (3) respectively. They can be found reproduced with permission in Appendix 1-3.

(1) Lange MB, Nielsen ML, Andersen JD, Lilholt HJ, Vyberg M, Petersen LJ. Diagnostic accuracy of imaging methods for the diagnosis of skeletal malignancies: A retrospective analysis against a pathology-proven reference. Eur J Radiol 2016 Jan;85(1):61-67.

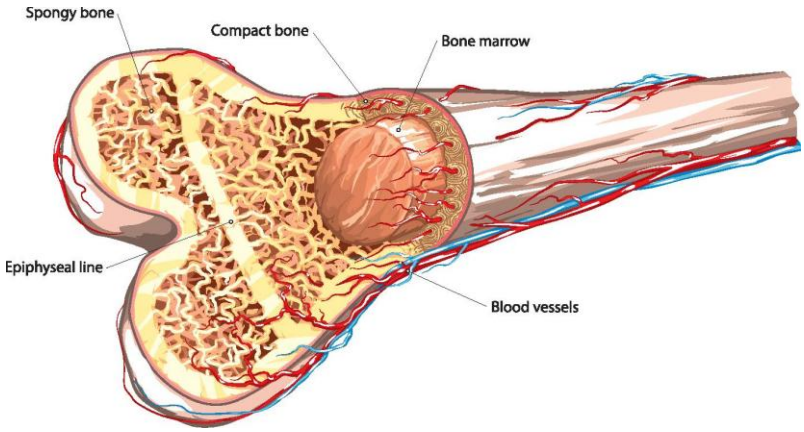
(2) Lange MB, Petersen LJ, Lausen M, Bruun NH, Nielsen MB, Zacho HD. Influence of Prior Imaging Information on Diagnostic Accuracy for Focal Skeletal Processes-A Retrospective Analysis of the Consistency between Biopsy-Verified Imaging Diagnoses. Diagnostics (Basel) 2022 Jul 17;12(7):1735. doi: 10.3390/diagnostics12071735.

(3) Lange MB, Petersen LJ, Nielsen MB, Zacho HD. Validity of negative bone biopsy in suspicious bone lesions. Acta Radiol Open 2021 Jul 27;10(7):20584601211030662.

Chapter 1. Introduction

a. Development and characteristics of bone metastases

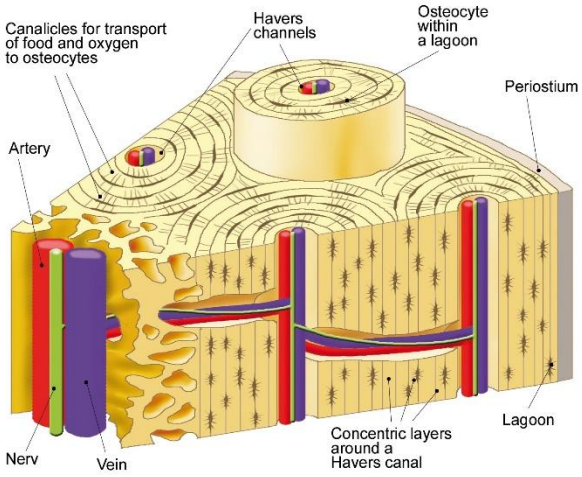
The adult human skeleton consists of 206 bones, out of which 126 are appendicular and 80 axial bones. They are made of mineralized connective tissue forming two types of bone, the outer cortical (compact) and the inner trabecular (spongy) bone matrix, the latter housing the bone marrow. Bone provides shape, support, and protection for the body, serves as a storage site for



minerals, and most also contains and nourishes bone marrow cells (Figure 1).

Figure 1. Schematic drawing of bone structure (figure purchased from Shutterstock)

Cortical bone is organized in so-called Haversian systems with a central channel surrounded by concentric layers of bone cells related to the break-down and formation of new bone (Figure 2).



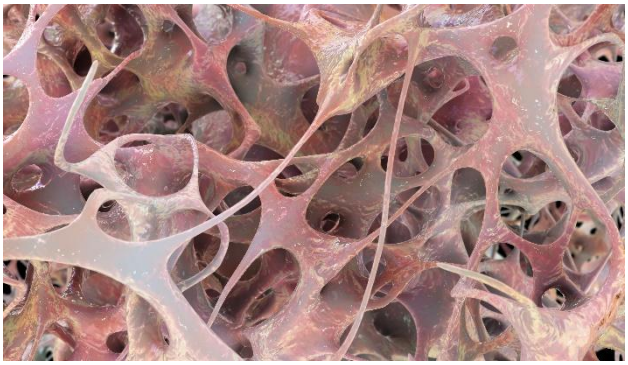


Figure 2. Structure of cortical and trabecular bone structure (figures purchased from Shutterstock)

The trabecular bone consists of solid bridges of bone, forming a complex 3-dimensional net with spaces housing the bone marrow (Figure 2). The numerous stages of hematopoiesis, osteogenesis, osteolysis, and diverse immunological responses are carefully managed in the bone marrow, which is a dynamic organ (1,2). The bone matrix represents a very active and dynamic type of connective tissue with many functions including protection to vital organs, mechanical support, and locomotion with a high regeneration potential in cases of fractures and milder infections (1,2).

Bone tumors can be mainly divided into primary bone tumors (sarcomas), comprising less than 0.2% of overall cancer diagnoses, and secondary bone tumors (metastases), being much more common than primary tumors, especially in adults (3,4). The most common malignancy of the bones is bone metastases. Recent advancements in cancer therapies, like the creation of molecularly targeted drugs and immune checkpoint inhibitors, have increased the survival rate of cancer patients, prolonging their clinical course and reducing morbidity due to bone metastatic disease (BMD) (5).

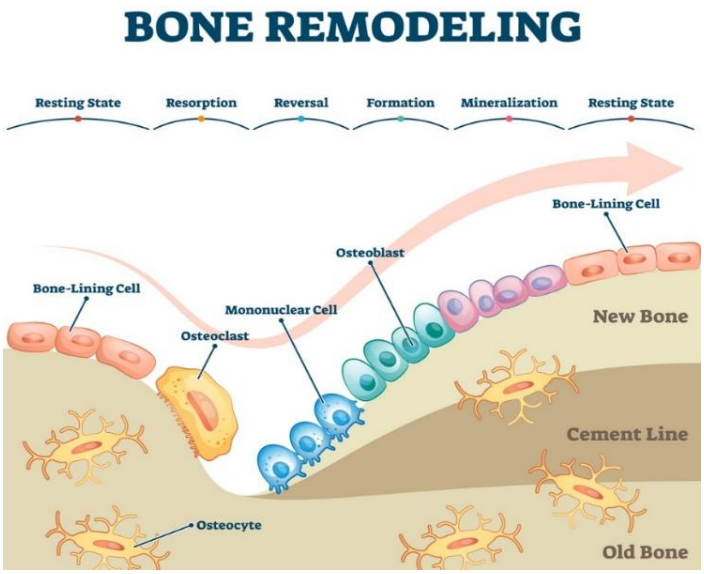
Globally, more than 18 million cancers are registered each year and of those, more than 50% of cases will develop bone metastases, since bone is one of the most common metastatic sites for solid malignancies and in 25-30% of cases, bone metastases are the first manifestation of malignancy (6-8). A large recently published American population study showed that prostate cancer, breast cancer, and kidney cancer have an incidence of bone metastases within the metastatic subset of 88.7%, 53.7%, and 38.7%, respectively (6). Other studies document that 10% of all newly diagnosed prostate cancers present with bone metastasis, increasing to 80% at advanced stages of the disease, and that prostate, breast, and lung cancers among the primary cancers most often tending to metastasize to bone with approximately 75% developing at least one bone metastasis during their course of disease (5,9-12). Approximately, 90% of prostate cancer and 70% of breast cancer patients eventually develop metastases, which represents the most frequent metastatic site behind lymph nodes, lungs, and liver, and nearly all patients who die of prostate cancer have bone metastases (4,11,13-20). Despite diagnostic tests and autopsies, the primary cancer location of bone metastases remains sometimes still unidentified (21).

The spine and pelvis are the most frequent anatomies for bone metastases, and only sporadically the bones distal to the knee and elbow are included (11,22). The most frequent sites of metastases in the spine are the lumbar vertebrae (5). Bone metastases occur either via a direct invasion of bone tissue or much more commonly via cancers that have spread (metastasized) from other tissues in the body through the blood or lymphatic systems to the bone marrow, which in turn spreads cancer to the firm bone matrix (1).

Despite several years of research, the complex biology of metastases is not yet fully understood (8). The bone marrow has been widely recognized to host a unique microenvironment that facilitates tumor colonization, and in the bone marrow, metastatic cancer cells take advantage of

the normal marrow physiology to survive distant from the primary site (5,8). The biology of bone metastasis is determined by tumor cell traits and their interaction with the microenvironment (1,2,8,22,23).

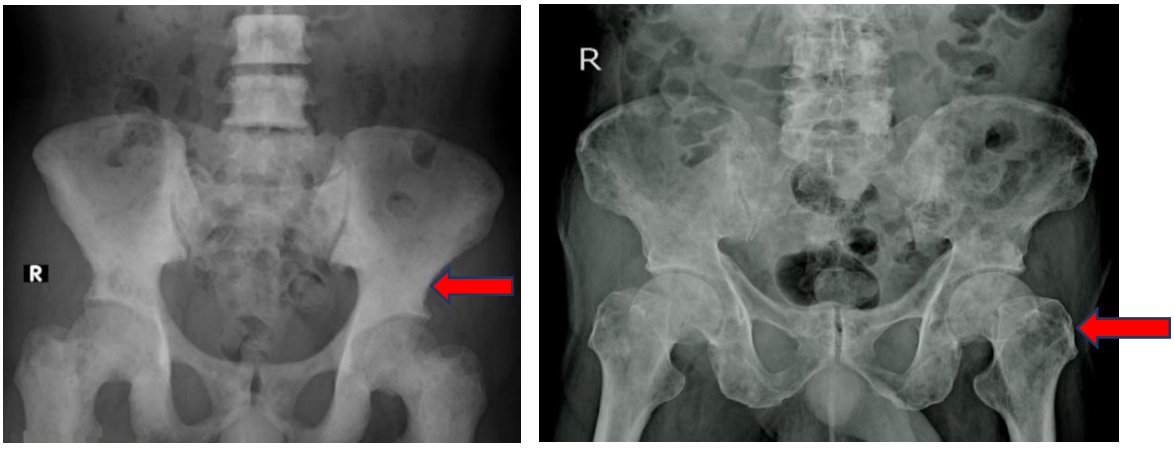
Tumor cells of bone metastases are not capable of destroying the bone matrix directly, but they hijack the physiological mechanisms controlling normal bone homeostasis thereby creating an



imbalance between osteoblasts and osteoclasts (Figure 3).

Figure 3. Schematic drawing of the normal bone remodeling process (figure purchased from Shutterstock)

Malignant cells act differently on the bone microenvironment according to the primary tumor type and can be classified as osteolytic (osteoclastic), sclerotic (osteoblastic), or mixed (5,8,24). For instance, prostate metastatic cancer cells most commonly produce osteoblast-promoting factors, causing sclerotic metastases, whereas metastatic cells from renal cell cancer most commonly overexpress osteoclast-inducing factors, causing lytic metastases; a mixture of the two types can be seen in for instance breast cancer metastases (2). An example of sclerotic and lytic metastases in the hip can be seen in figure 4 and an overview of the types and characteristics of bone



metastases can be seen in table 1.

Figure 4. Hip with osteoblastic (left) and osteolytic (right) metastases (figure purchased from Shutterstock)

Table 1. Overview of types and characteristics of bone metastases (adapted from Bădilă et al (2))

Type of bone metastasis	Manifestation	Symptoms	Primary cancer	Radiographic appearance
Osteolytic (osteoclastic)	Destruction of bone causing bone pain, fracture, hypercalcemia and nerve compression syndromes	Bone pain Fractures Hypercalcemia Nerve compression Syndromes	Breast MM RCC Melanoma NSCLC Non-Hodgkin lymphoma Thyroid cancer	Radiolucent areas
Osteosclerotic (osteoblastic)	Deposition of new bone with dysregulated bone resorption and bone formation	Bone pain Fractures	Prostate Carcinoid SCLC Hodgkin lymphoma Medulloblastoma	Hyperdense areas
Mixed	Osteolytic and osteoblastic components in the same metastasis		Breast cancer GI-cancers Squamous cancers	Fuzzy areas, a sclerotic rim with a center of osteolytic lesions can be seen

Abbreviations MM: Multiple Myeloma, RCC: Renal Cell Carcinoma, NSCLC: Non-Small Cell Lung Carcinoma, SCLC: Small Cell Lung Carcinoma, GI: Gastrointestinal

Bone metastasis frequently occurs in the late stages of malignant diseases, and evidence suggests that metastatic cancer cells can remain in a dormant state for decades before they proliferate and destruct the bone matrix. They are capable in the dormant phase of evading the immune system, which then can be reactivated later. Once developed, it indicates a short prognosis with 5-year survival rates reaching 20% with multiple bone metastases and 40% when presenting with a solitary lesion (2,11,22,23). Some have tested statistical models for the codification of complex clinical data into prognostic models, which can be of value for therapeutic decisions (25). New treatment targets have emerged due to result of recent developments in the knowledge of the molecular mechanisms underlying the progression of bone metastases (1,25).

b. Treatment strategies for bone metastases

Even though significant progress has been made in understanding the nature of BMD, the disease remains mostly incurable once tumor cells gain foothold in bone and start proliferating (8,22). Overall, treatment options may intend to be generally or locally curative by the use of surgery, chemo-, radiotherapy, or ablation or palliative by the use of radiotherapy, chemotherapy, arterial embolization and stabilization with forms of cementoplasty such as kyphoplasty, and/or

vertebroplasty (8,11). These can be combined with diofrequency ablation, cryotherapy, internal fixation, or systemic drug therapy, which can be necessary for the treatment of advanced BMD

(8,11). Surgery remains in many instances indicated for bone metastasis treatment, mostly providing fixation and/or replacement of the affected bone with a bioimplant as opposed to primary tumors, where the target is to excise the lesion (2,11,23). Angiography allows confirmation and mapping of the vascularity of the neoplasm and can be proposed in highly specialized centers along with embolization before performing surgical resection in highly vascular tumors (26). The future development of potential therapeutic approaches for the clinical treatment of bone metastasis might be based upon targeting specific stages of tumor progression and crosstalk with bone niches (22). Stromal cells may also be therapeutic targets by disrupting the interaction between the altered bone niche and cancer cells (22). However, since treatment is at present rarely curative, palliative therapies to alleviate symptoms or morbidity associated with bone metastasis are most frequently initiated (8,22).

Despite intense research and development, BMD is still associated with significantly increased mortality and morbidity, frequently causing pain, fractures, spinal cord compression, hypercalcemia, and bone metastatic disease is also associated with considerable use of health care resources (2,4,5,8,9,11,13,27-30). Only 20% of breast cancer patients remain alive 5 years after the discovery of bone metastasis and the median survival after a diagnosis of bone metastases ranges from around 1 year for patients with lung cancer to 3–5 years for those with breast cancer, prostate cancer or multiple myeloma (8,29). The choice of treatment remains a challenge for clinicians and unfortunately, current therapeutic options are as mentioned mainly palliative. Since the bone microenvironment and its interaction with metastatic tumor cells differs from one cancer type to another, a precise and early diagnosis is of utmost importance for the initiation of targeted therapeutic strategies for BMD to prevent complications, such as pathological fractures and spinal cord compression, causing impairment of QOL and survival once they occur (2,5,11).

In summary, treatment strategies for BMD should be holistically planned by taking the patient's general health status and prognosis into consideration. Multidisciplinary teams focusing on the management of BMD consisting of specialists, with the radiologist playing a central role, will support comprehensive healthcare and treatment of patients (5).

c. Diagnosis of bone metastases

The diagnosis of BMD includes clinical history, diagnostic imaging, and eventually a biopsy (8,31). Most patients with BMD present with either pain, spinal cord compression, hypercalcemia, or pathological fractures and for some, BMD might be the first manifestation of a malignant disease (7). Unexplained bone pain in patients with newly diagnosed prostate cancer has been shown to have a significant predictive value for bone metastases independent of other known clinical variables such as PSA, Gleason stage and T-score (32).

Diagnostic imaging plays a central role in the detection of BMD and the purposes are many, including early detection of the lesions, assessment of possible differential diagnoses, local spread, cortical breach, risk of pending fracture, impact on performance, function, and on the surrounding structures, quantification of the extent of disease load, planning of biopsy site, evaluation of type (osteoblastic or -lytic), planning of surgery and other treatment options and to assess treatment monitoring through interval imaging (11,33).

Because the pattern of bone metastases is so variable, it is essential to understand the capabilities and restrictions of each imaging method (11). It is crucial to keep in mind that, unless proven differently, a bone lesion in a patient with any known underlying cancer should be considered a metastasis (11).

Several studies, systematic reviews, and meta-analyses have investigated diagnostic characteristics of the different imaging modalities and ligands in various tumor types (8,14,34-48). Since all cancer diseases are potentially systemic diseases, multi-modality imaging is frequently employed for disease staging (33), and whole-body imaging techniques are being used more regularly to reflect this (33).

X-ray is most often the primary screening technique used for bone tumors and tumor-like lesions, but due to its low sensitivity, additional imaging is most often necessary (49). When lesions are visible on radiographs, they should be assessed for aggressive or non-aggressive features including bone destruction, lesion margin, tumor matrix, and periosteal reaction, and in such cases, the specificity is high (50). Unless the lesion is sufficiently calcified, as in osteoblastic lesions, or eliminated, as in osteolytic lesions, being apparent when more than 50% of the bone substance has disappeared, metastases are rarely visible (4) . Small lesions are easily missed on radiographs, particularly in patients with low bone density, for instance, osteoporosis, which is sometimes caused by systemic BMD therapy (4,5). The trabecular bone density is higher in the epiphysis and metaphysis than in the diaphysis which makes lesions in those areas easier to detect due to the better contrast against the adjacent normal trabeculae (11).

CT has a higher sensitivity than conventional X-ray and most often, CT is performed as part of the routine staging protocol during the initial workup of any cancer diagnosis, and thus it is commonly the first imaging modality to detect the bone lesion suspected of representing MBD (11). Many of the radiographic features described above can also be assessed on CT, where you more precisely can determine bone lesions by measurement of the density expressed in Hounsfield Units (HU), a relative quantitative measurement of radiodensity on a linear scale, where bone can reach up to 1000 HU (50,51). A lucent lesion can be defined as a lesion where more than 90% has lower attenuation than normal trabecular bone, which typically has a density around 200 HU, whereas osteoporotic trabecular bone density is around 120 HU (52). There is currently no specific attenuation range that clearly defines sclerotic (53), and even if more recent literature define a sclerotic lesion as one where ≥ 50% of the volume is denser than the surrounding normal trabecular bone, the description of a sclerotic lesion is most often subjectively applied by the

interpreter (53). The term mixed lesion is commonly used to describe a lesion that contains a combination of sclerotic and lytic areas relative to adjacent trabecular bone on CT, and a recently published white paper defined a mixed lesion as one with equivalent or near-equivalent amount of sclerosis and lucency (53). Recently, the application of dual-energy CT virtual non-calcium algorithms has proven to be a valuable tool in the assessment of MBD (11).

MRI is superior to the other imaging modalities in detecting bone marrow lesions and is currently considered the best imaging modality to depict diffuse bone marrow involvement, having superior sensitivity and specificity for tissue characterization of BMD (49). MRI is very helpful in local staging and surgical planning and is also used in assessing response to neoadjuvant therapy, restaging and post-therapeutic follow-up (49). Compared to normal tissue, many malignant disorders have low to intermediate signal intensity on T1 and high signal intensity on T2. When discovered, primary benign and malignant bone tumors, as well as metastatic lesions, should be taken into account in the context of additional factors like a history of malignancy known to cause bone metastases, pain associated with the lesion, involvement of the cortical and soft tissues, pathologic fractures, surrounding bone marrow edema, solid mass-like enhancement, and paraclinical results such as PSA (53).

Accurate assessment of the T1 signal intensity is of primary importance (53). Enostoses and osteoblastic metastases are examples of T1 iso- or hypointense bone lesions, which are defined as lesions that exhibit a T1 signal that is the same as or lower than that of nearby skeletal muscle or the intervertebral disc (53). The majority of bone cancers and metastases are isointense/hypointense T1 lesions, hence it is necessary to evaluate their T2 characteristics in order to further distinguish between lesions that are clearly benign and those that need additional investigation (53). When using T2-weighted imaging, many benign and malignant bone lesions exhibit higher signal intensities than the surrounding normal marrow fat because they contain more free water. These bone lesions are frequently accentuated using protocols suppressing the T2 signal of the fat (fat saturated or short tau inversion recovery (STIR) sequences) (53,54). Diffusion-weighted MRI (DWI) is a sequence sensitive to water molecule movement into the tissue, and since malignant tumors are generally more cellular than normal tissue, extracellular water molecule movement is lower (26,55). DWI can quantify pathological states using the Apparent Diffusion Coefficient (ADC), although validation of its use in musculoskeletal tumors is still debated (26). Since MRI allows for a precise assessment of the extent of the disease and the impact of the tumor on the surrounding tissue, it is regarded as the best imaging technology for evaluating locoregional disease (11). Whole-body MRI is promising as it could help deliver the promise of precision oncology for patients with BMD (11,33,53,55).

Radionuclide bone scan detects metabolic bone activity and especially osteoblastic (bone deposition) response (11). The ability to perform whole-body imaging for the same radiation dose allows us to detect polyostotic disease and planar bone scintigraphy (BS) is clinically easy to read, widely available, and cost-effective (33). The extent of osteoblastic activity can indicate disease activity (11). The sensitivity of the bone scan for lytic lesions depends on the magnitude of the osteoblastic reaction. Pitfalls include post-chemotherapy lytic lesions (particularly in breast and lung cancer), bone infarcts, and mucinous cystic lesions (11). Relatively low osteoblastic activity

and high osteoclastic activity limit the detection of lesions on bone scans and bone scintigraphy is associated with relatively poor spatial resolution, limited diagnostic specificity and reduced sensitivity for bone marrow disease (56). It also shows limited diagnostic accuracy in assessing response to therapy in a clinically useful time period (11). An overall improvement is observed when BS is augmented using single-photon emission computerized tomography (SPECT) and combining SPECT with CT furthermore improves the detection of metastatic bone disease (33). Lately, the use of deep neural network models to automatically extract high-level features from the BS data has shown promising results for recognizing the absence or presence of bone metastasis on WBS images (57,58).

PET-CT techniques using bone-specific tracers such as NAF ([18F] sodium fluoride) or tumor-specific tracers such as FDG ([18F] fluorodeoxyglucose) or Prostate-Specific Membrane Antigen (PSMA)–ligands have shown high diagnostic accuracy for the detection of skeletal metastases (59). NAF is chemisorbed into bone crystals with the formation of fluorapatite, a process that mainly takes place in areas of active mineralizing bone and NAF PET/CT has the advantage faster blood clearance, high spatial resolution and attenuation correction (60). The mechanism of FDG uptake into bone metastases differs from NAF in that FDG accumulates in viable, metabolically active tumor cells rather than in reactive bone . As the uptake of FDG is not limited to bone tumors, it has the advantage of showing both skeletal and soft tissue metastases in patients with cancer (56). Prostate-specific membrane antigen (PSMA) is a transmembrane glycoprotein overexpressed in prostate cancer cells, either in prostate cancer tissue or at other metastatic sites, and more than 90% of primary prostate cancer displayed moderate-to-high PSMA expression (61-63). PSMA is therefore considered as a suitable target for prostate (PET) imaging, but since PSMA functions as a folate hydrolase, it is expressed in a range of normal tissues and in other benign and malignant processes, making knowledge of other causes of uptake and distribution essential to minimize false-positive imaging findings (61-63).

Recent research has investigated the ability of specific tracers to distinguish different types of cancer forms from the same organ by using tracers in PET/CT imaging that allow the detection of fibroblast activation protein in the tumor microenvironment, providing insight into the biological characteristics of tumors (64).

Other whole body multiplanar radionuclide imaging techniques utilize existing applications of MRI and CT, combined with nuclear medicine studies with different tracers to provide an overview of MBD load and to ultimately prioritize treatment strategies. As mentioned, PET-CT is commonly used with a proven high sensitivity and PET-MRI seems to be promising but is now only available in a few places (11,33).

Even if one chooses the right imaging strategy for detection of BMD, misinterpretations cannot be avoided. Errors in radiology are known to be complex and often multifactorial and several studies have investigated the incidence and causes of medical errors (65,66). Radiology is known to be a field very liable to claims of diagnostic negligence with an average error rate among radiologists

around 30% referring to images (chest x-rays) as part of a set of unknowns with proven pathology, and the causes are many, most being perceptual (67-73). Radiologists often disagree with others and even with themselves, and comparison with previously obtained examinations is considered standard in order to get as few errors as possible (74-84). Provision of adequate clinical information is considered important for accurate radiological diagnoses and subsequent clinical management, and even if there have been discussions as to whether it should be recommended to review the images before reading the history in order to avoid bias, the common practice of reading diagnostic tests with clinical information seems justified (76,85-88).

Bone biopsy is generally considered the gold standard, i.e., error-free reference standard, for verification of cancer in cases of malignant cells in a bioptic specimen, and an imaging-guided biopsy is often ultimately required to establish a specific diagnosis (4,7,9,26,29,37). It is most often performed in cases of an unknown primary neoplasm, heterogeneous disordered matrix, distinct signal decrease in T1-weighted MRI (many cells), blurred border, perilesional edema, cortex erosion, and with a large soft tissue component (fast growth) when MRI does not lead to a diagnosis or in connection with surgery, but despite being considered the gold standard, it is only rarely performed in the clinical setting (49,89). In diagnostic imaging accuracy scientific trials, biopsy is only rarely used as a reference and an insufficient reference standard has been identified as a major error source in biomedical research (14,34,35,37,90,91). Furthermore, a negative bone biopsy can represent a false negative and this can have serious consequences, not only for the patient, but also causing misleading results in diagnostic accuracy studies (92-97).

To summarize, BMD is a significant healthcare concern, and orthopedic surgeons, pathologists, oncologists, radiotherapists, nuclear specialists, other clinicians and radiologists are involved in the multidisciplinary management (5,11). Since early and accurate diagnosis of BMD is essential and the pattern is highly heterogenous, the radiologist and nuclear specialists play a crucial role in selecting the best imaging strategy, based upon well-documented knowledge of the limitations and benefits of each imaging modality (11).

d. Challenges in imaging diagnostic research and guidelines

Several studies have been conducted to validate the diagnostic accuracy of various imaging modalities, being of crucial importance for diagnosis and thus treatment. However, only a few can live up to the requirements for a correctly performed diagnostic test accuracy study as outlined in STARD (The Standards for Reporting of Diagnostic Accuracy) and recent studies have failed to show any significant improvement in reporting quality over the last 10 years (90,91,98,99). In addition, methodological errors such as sampling bias, failure to blind readers to the results of the reference test as well as other index tests and verification bias are frequent and furthermore, reliable imaging parameters to predict therapy response in cases of bone metastases have not yet been elucidated in large randomized controlled clinical trials (11).

The ability of any imaging diagnostic procedure to distinguish between disease and health can be expressed in different ways and such measures are extremely sensitive to the design of the study. Studies suffering from methodological shortcomings can severely over- or underestimate the

indicators of test performance as well as they can severely limit the possible applicability of the results of the study (91). STARD guidelines and Food and Drug Administration for diagnostic accuracy tests recommend the use of sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) (91,100).

The sensitivity of a test (also called the true positive RATE) is defined as the proportion of people with a disease who will have a positive result (TP/ (TP + FN)), whereas the specificity of a test (also called the true negative rate) is the proportion of people without the disease who will have a negative result (TN/ (TN + FP)) (101). Positive predictive value (PPV) calculated as (TP/ (TP + FP)) and negative predictive value (NPV) calculated as (TN / (TN+FN)) will tell you the odds of you having a disease if you have a positive result. Whereas sensitivity and specificity are based on the patient, PPV and NPV are based on the test. Unlike sensitivity and specificity, predictive values are largely dependent on disease prevalence in examined population (101). Prevalence affects PPV and NPV differently. PPV is increasing, while NPV decreases with the increase of the prevalence of the disease in a population. Whereas the change in PPV is more substantial, NPV is somewhat weaker influenced by the disease prevalence (101). Accuracy (also called diagnostic effectiveness) is expressed as a proportion of correctly classified subjects (TP+TN) among all subjects (TP+TN+FP+FN). Just as with PPV and NPV, diagnostic accuracy is affected by the disease prevalence. With the same sensitivity and specificity, diagnostic accuracy of a particular test increases as the disease prevalence decreases (101). This does not mean that the test is better in a population with low disease prevalence, it only means that the test gives more correctly classified subjects in absolute number. This percentage of correctly classified subjects should always be weighed considering other measures of diagnostic accuracy, especially predictive values (101).

Most importantly, however, the choice of reference test is of utmost importance and the lack of consistent use of a valid reference test, i.e., gold standard, for the presence or absence of malignancy at the lesion site is common, meaning that histological diagnosis is only rarely systematically used as a reference standard (37). In many cases, diagnostic data from other imaging diagnostics or clinical follow-up are used and the true diagnostic accuracy of imaging methods is therefore uncertain.

As earlier described, histology sample from the suspected bone lesion is generally considered gold standard, i.e., a valid test for exclusion of false negative samples. However, this validity has to our knowledge never been tested and proven. Such data are of utmost importance for future imaging diagnosis and research.

Finally, imaging guidelines states it mandatory to compare imaging examinations and reports with those previously obtained. However, to our knowledge, no studies have been published to substantiate this demand, so no data are available to document neither the influence, nor the sequence, of prior imaging on diagnostic accuracy using biopsy as the reference standard, making guideline recommendations less trustworthy.

Chapter 2. Aims and overview of the studies

As described, there are several challenges in existing imaging diagnostic research and guidelines, one of the most important being the lack of consistent use of a valid reference standard for investigation of diagnostic accuracy for detection of skeletal metastases. The overall aim of this Ph.D. was to provide such data based upon a consecutive cohort of skeletal biopsies, based upon which we could not only provide systematically valid evidence on diagnostic accuracy of relevant imaging methods for detection of bone metastases, but we could also investigate whether diagnostic accuracy was affected by access to and the sequence of previous diagnostic imaging and reports. Finally, we wanted to document that bone biopsy is the true gold standard. We therefore designed three studies, for which the specific aims are listed below. An overview of the studies is provided in table 2.

Study I: Having unsuccessfully performed a comprehensive computer literature search to identify studies comparing the diagnostic characteristics of the various imaging modalities with respect to the diagnosis of the suspected skeletal malignancy systematically using bone biopsy as the gold standard for the majority of patients, our aim of this study was to provide such data by retrospectively examining the diagnostic accuracy of standard skeletal imaging modalities versus pathology reports in a large consecutive population of patients (102).

Study 2: The aim of this study was to investigate whether the diagnostic accuracy of the detection of skeletal malignancies, proven malign or benign by subsequent biopsy, is affected by prior imaging examinations and their mutual sequences (103).

Study 3: The aim of this study was to investigate whether targeted bone biopsies described as non-malignant or benign identified in a population with a suspicious focal bone lesion are in fact truly benign after two years of follow-up (104).

Table 2. Overview of the 3 studies

	Study 1	**Study 2**	**Study 3**
Study purpose	Diagnostic accuracy of skeletal imaging modalities against pathology-proven reference	Possible influence of previous imaging on diagnostic accuracy of skeletal imaging modalities	Prove bone biopsy as a true gold standard
Patient cohort	Retrospective consecutive cohort studies based upon a large sample of bone biopsies		
No. of patients	395	207	207
No. of biopsies	409	216	215
Method	Comparing diagnostic imaging reports (X-ray, CT, MRI, BS, and FDG-PET/CT) with pathology reports	Comparing diagnostic imaging reports in cases of 2 or more with pathology reports	2-year follow-up on benign skeletal biopsies to examine any additional biopsy, imaging and clinical follow-up information
Reference method	Bone biopsy pathology report	Bone biopsy pathology report	Additional biopsies +imaging + clinical follow-up
Applied statistics	Sensitivity, Specificity, Prevalence, Accuracy, PPV, NPV with 95% exact confidence intervals, Chi square tests, Fisher's exact test	Sensitivity, Specificity, Prevalence, Accuracy, PPV, NPV with 95% exact confidence intervals, Fisher's exact test	Descriptive statistics: calculation of standard deviations, 95 % confidence intervals
Main findings	MRI and FDG-PET/CT performed best in most patient subgroups, followed by CT with some distance to X-ray and BS	The sequence of the imaging modalities seems to influence the diagnostic accuracy	Negative bone biopsy is documented to be a valid criterion for the absence of bone metastasis

Chapter 3. Material and methods study 1, 2 and 3

Material

All 3 studies were conducted as retrospective consecutive cohort studies based upon bioptic samples gathered via a computer search at The Department of Pathology at Aalborg University Hospital between January 1, 2011, and July 31, 2013, identifying bone biopsy material registered by SNOMED (Systematized Nomenclature of Medicine) as T10* and T11* codes for skeletal cytology and histology biopsies, resulting in 745 samples. Each biopsy was identified by a distinct social security number, based upon which a manual check on each patient's hospital electronic imaging file registration (EasyViz, Karos Health Inc., Waterloo, Ontario, Canada) for any prior imaging procedures carried out six months or less before the biopsy date was performed, resulting in 690 (102). Excluded were samples with no conclusive pathology results, samples from fetuses and provoked abortions, no written description of the imaging findings by a specialist in imaging (radiologist or nuclear medicine physician) and no anatomical match between the site of biopsy and the imaging field (102). If both a cytological and a histological sample were available from the same biopsy, the cytological sample was excluded. Please refer to Figure 5, Consort Diagram.

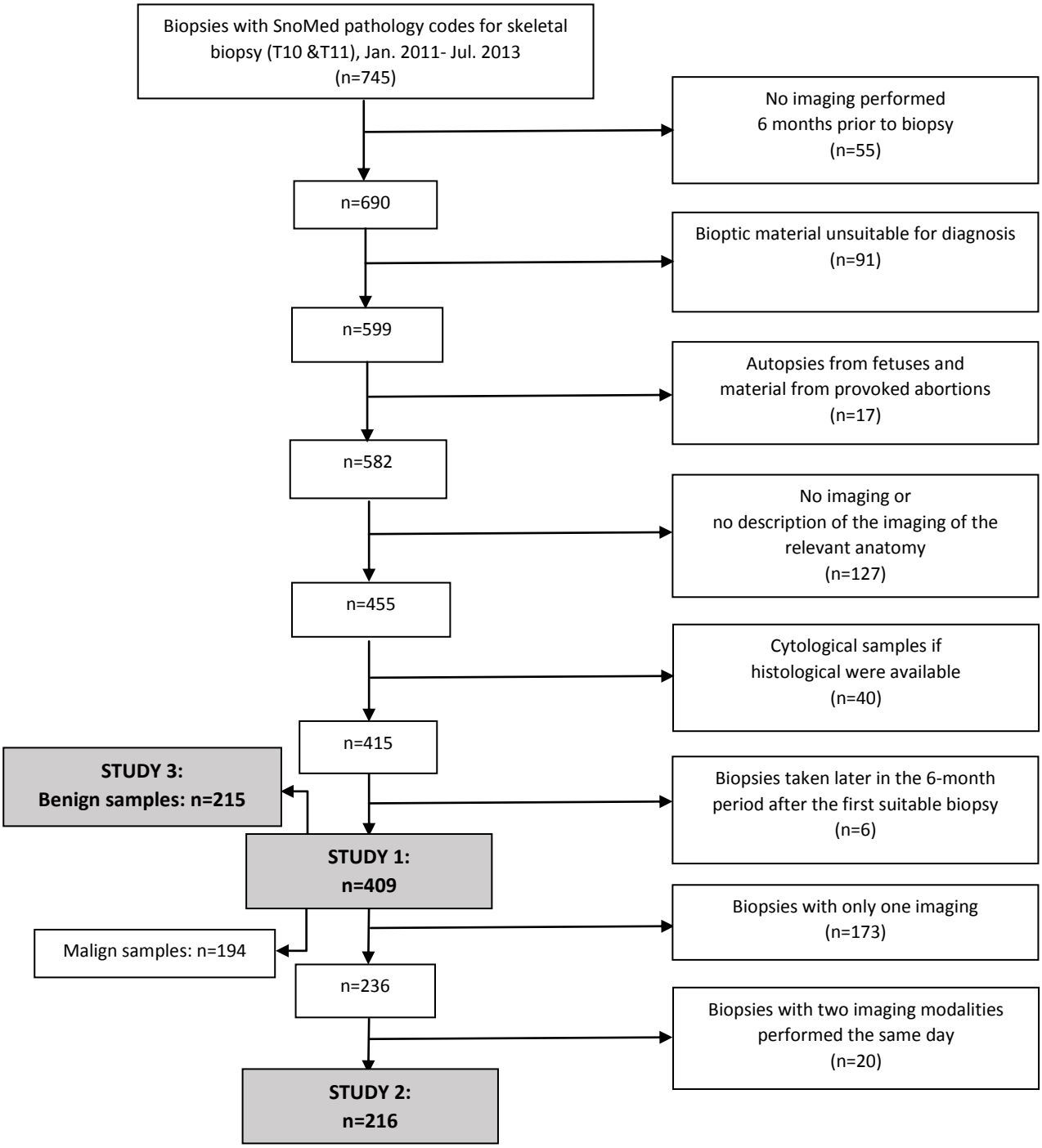


Figure 5. Consort Diagram. The number of specimens ended up being 409 in study 1. In study 2 we excluded biopsies with only one diagnostic imaging performed in the 6 months period prior to the biopsy and samples where two imaging were performed the same day, leaving us with 216 samples. Study 3 were based upon the samples collected for study I described as benign, representing 215 samples.

Methods

As explained in figure 5, 409 biopsies were included in study 1. The diagnostic accuracy was determined by comparing pathology reports from bone biopsies to diagnostic imaging in terms of X-ray, CT, MRI, BS, and FDG-PET/CT conducted within 6 months of the biopsy (102). A subgroup analysis for spine and non-spine anatomical sites as well as for tumor types were performed to see if the diagnostic accuracy of the various imaging modalities differed depending on the anatomical site under investigation. Finally, diagnostic properties of imaging modalities were examined based on whether the bone reaction was expected to be mostly osteosclerotic, osteolytic, or mixed according to Roodman (24). Prostate cancer bone metastases were defined as osteosclerotic, squamous cell adenocarcinoma of the lung bone metastases as mixed, and small cell anaplastic carcinoma of the lung bone metastases as osteolytic. The bone responses for additional tumor types were classified according to documentation available in the public (102).

To analyzing the possible influence of prior imaging reports and their mutual sequences in study 2, we included biopsies with 2 or more imaging examinations (X-ray, CT and MRI) performed in the 6 months period prior to the biopsy, except if the two imaging examinations were performed on the same day, which left us with 216 samples (please refer to figure 5) (103). BS, SPECT/CT and FDG-PET/CT were excluded due the low number of combinations of those with another.

In study 3, where we wanted to investigate whether bone biopsies described as non-malignant or benign identified in a population with a suspicious focal bone lesion are in fact truly benign after two years of follow-up, we selected the samples collected for study I described as benign, representing 215 out of the 409 samples (please refer to figure 5) (104). Based on the unique Danish Central Personal Registration system, supplying inhabitant with a personal ID number, combined with databases that include diagnostic codes for each ID number, we performed a 2-year follow-up for the non-malignant bone biopsies including a careful computer search on each patient in the pathology database for any additional biopsy, in the imaging system EasyViz for any imaging of the relevant structure, and finally in the Electronical Patient Journal charts (EPJ - Clinical Suite, CSC Scandihealth A/S) for any relevant journal notes in order to categorize the original biopsy as truly benign, malignant or equivocal (104). This was done by two independent readers who reached consensus. A combination of a negative biopsy from the same anatomy as the first biopsy taken or no biopsy from the same anatomy available AND no imaging with suspicion of cancer from the same anatomy AND no clinical suspicion of malignancy from the same anatomy were used to characterize a biopsy as truly benign after two years of follow-up (104). If one of the following criteria applied to a biopsy, it was termed malignant: 1) a positive biopsy of the same structure or neighboring soft tissue, 2) any positive imaging of the structure, or 3) clinical suspicion of malignancy based on the relevant anatomy, such as persistent symptoms or blood tests that lead to additional diagnostic testing. Follow-up biopsies that were not diagnosed as true benign or

true malignant were labeled as equivocal. Patients with post-biopsy imaging that was inconclusive for malignancy were also included in this category (104).

Imaging

All radiological imaging procedures were carried out in conformity with the hospital's policies (no experimental imaging investigations were included in the analysis). Digital radiography was used for X-ray imaging, and CT scans were done on a GE (GE Lightspeed VCT, 64 slice, GE LightSpeed Pro, 32 slice, GE Discovery 750 HD, General Electrics, Milwaukee, WI, USA) or Siemens (SIEMENS Definition Flash Siemens AG, 128 slice) scanner (102-104). A 1.5 T MR scanner was used for the MRI scans (Discovery MR450, General Electrics, Milwaukee, WI, USA). T1, T2, and STIR were the MRI image sequences used, with at least one of them being axial on the bone in question; contrast was only used in cases of soft tissue involvement, which was determined by a physician in each case (102-104).

For nuclear medicine procedures, we used BS with 750 mega Becquerel (MBq) 99 mTc-labeled bisphosphonate and whole-body scanning or regional imaging on a dual-headed gamma camera, depending on the indication (Symbia T16 or E.CAM, both Siemens AG, Berlin and Munich, Germany). Only three individuals received supplemental single-photon emission computerized tomography (SPECT/CT). Because of the small sample size, the results from these patients were not analyzed in our study. All PET/CT scans were performed with FDG. One hour after injecting 370 MBq of 18F-FDG intravenously, FDG-PET/CT was conducted. In compliance with institutional standards, whole-body pictures were collected using a 64-slice Discovery VCT PET/CT scanner (General Electric Medical Systems, Milwaukee, WI, USA). All BS were double read by at least one board-certified nuclear medicine physician as per institutional policy. PET/CT scans were always read by a physician and a radiologist, irrespective of the use of diagnostic or low dose CT scans (102-104).

Two independent reviewers, sometimes working in pairs, reviewed all the imaging descriptions. Based upon the imaging report, the lesions were classified as malignant, benign, or inconclusive or equivocal. Without the use of a third-party arbiter, the readers established a consensus for each imaging report after each individual reading. In a small number of instances, the imaging outcome was deemed ambiguous and hence classified as inconclusive. This was most frequently the case with X-ray reports of fractures, where it was not specified whether the fracture was pathological or not (102-104).

Biopsies

Most bone biopsies were taken during surgical or image-guided procedures (CT or ultrasound). In a few cases, biopsies from post-mortem exams were obtained. The biopsies were handled and analyzed according to standard procedures at the hospital. When immunohistochemistry was necessary, it was used. All biopsies were diagnosed by a board-certified pathologist (102-104).

The lesion was classed as malignant, if several biopsies were taken from the same anatomical site over a six-month period and just one of those revealed malignancy. The first biopsy diagnosis was used if repeated biopsies were described as benign. If the eligibility criteria were met, lesions from different anatomical locations in a patient could be added. Each pathology report was evaluated as benign, malignant, or inconclusive by two reviewers. A board-certified pathologist aided with the final decision in cases of inconsistency (102-104).

Statistics

Diagnostic accuracy was in study 1 and 2 calculated as sensitivity, specificity, positive predictive value (PPV), negative predictive values (NPV) and accuracy and our reference test was pathological diagnoses in all cases (103). The cumulative proportion of correct outcome (accuracy) was also calculated. In study 1, we used GraphPad Prism to calculate Pearson´s Chi square tests for independent data for 5 x 2 comparisons and Fisher's exact test for 2 x 2 comparisons, whereas in study 2 (where we used Stata 17 as mentioned below), we only used Fischer´s exact test, because we considered the datasets to be rather small, for which Fischer's exact test is considered better. Recently some articles showed it better to use Pearson's Chi square test in all contexts, possibly with a minor modification for small datasets (105,106). However, in most cases, the difference is marginal. All p-values were two-sided and values <0.05 were considered significant. Data were reported with 95% confidence intervals (103). In some subgroup analyses in study 1, only sensitivity was calculated and thus used for statistical analysis.

In study 2, we used Stata 17 (StataCorp LLC 2021) and the Stata package matrix tools (107) for calculations of diagnostic accuracy for each modality. Then it was calculated for pairs of imaging modalities, such as X-ray/CT and CT/X-ray, while excluding all stand-alone data to reduce the bias that only one imaging was performed rather than two. The accuracy of one imaging modality (CT, MRI, and X-ray) when employed as the initial imaging modality were compared to the diagnostic properties of the modality when preceded by another modality. Finally, the impact of the imaging sequence was studied in patients with malignant and benign biopsy diagnoses using descriptive statistics due to the small number of patients in each group. It should be emphasized that the numbers in some of the subgroups may not be large enough to detect meaningful differences (103).

In study 3, descriptive statistics included calculation of standard deviations and 95% confidence intervals were performed (104).

Approvals

In conformity with national legislation, this retrospective investigation did not require ethical approval or informed consent. The study was approved by the Danish Data Protection Agency, which gave authority to examine medical files for the purpose of the study.

Chapter 4. Results

Study 1

Lange MB, Nielsen ML, Andersen JD, Lilholt HJ, Vyberg M, Petersen LJ. Diagnostic accuracy of imaging methods for the diagnosis of skeletal malignancies: A retrospective analysis against a pathology-proven reference. Eur J Radiol 2016 Jan;85(1):61-67.

Table 3 provides an overview of patient demographics and baseline information.

Out of the 409 biopsies included, 44 were cytological specimens, 379 were regular histological biopsies from patients, and 6 were post-mortem biopsies. There was a slight overweight of men, and the mean age was 64 (1-95). There were slightly more benign biopsies, 215 in total, which were given 28 different pathology diagnoses, with osteochondroma, fibrosis, and inflammation being the most frequent (n = 20 or more). Of the malignant samples, the most prevalent cancer types were lung cancer, breast cancer, and multiple myeloma, representing almost 60 % of the cancer types. According to histology, 82% of lung cancers were non-small-cell lung cancer and 82% of breast cancers were ductal carcinomas. The skeletal lesions were most frequently found in the spine followed by extremities and pelvis.

In total, 758 imaging procedures were completed with 62% of the patients having two or more diagnostic imaging procedures performed prior to the biopsy. The majority of pre-biopsy imaging modalities were radiological investigations, making up 88%, with the number of X-ray, MR, and CT procedures being almost evenly distributed (all >200 each). Out of the 209 MRI scans, contrast-enhancement (gadolinium) was used in 58. The number of nuclear medicine procedures was rather limited, only approximately 100 in total. Diagnostic CT was performed for 53% (31 of 58) of the PET/CT scans, while low-dose CT was used for the rest. An example of different imaging modalities used for detection of possible bone metastasis in the spine is shown in figure 6.

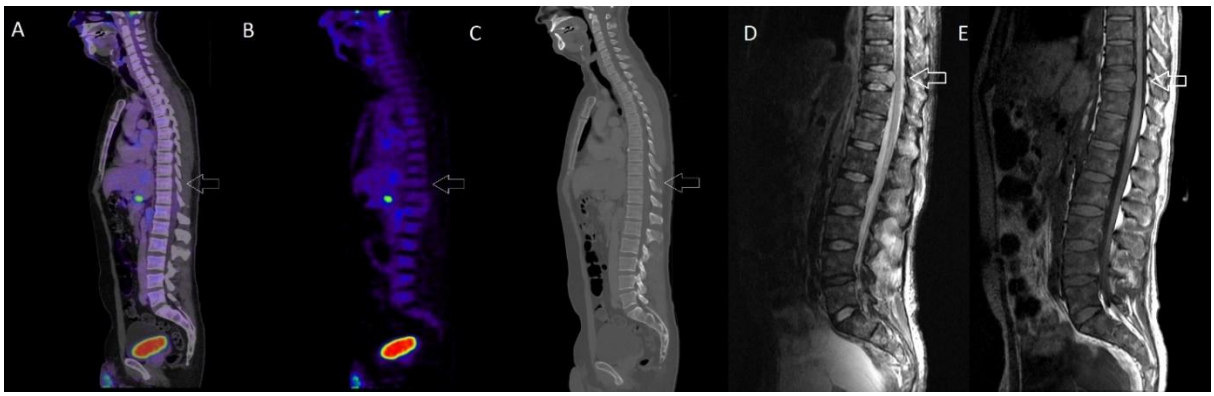


Figure 6. A 46-year-old man was diagnosed with cardia cancer (subcardial grade III). An initial FDG-PET/CT scan showed no bone metastases in the spine (A: PET/CT sagittal fused image; B PET sagittal view; C: CT sagittal view). 3

months later, persisting back pain gave rise to an MRI scan showing disseminated cancer with a biopsy-proven bone metastasis in Th10 (arrow) with posterior soft tissue bulging and compression of the medulla (D: Sagittal T2 Stir TR/TE 3500/38,496 msM E: sagittal T1 TR/TE 350/12,832 ms). Reproduced with permission from European Journal of Radiology (102)

Table 3. Patient demographics and baseline information.

Patients (n)	395
Male n (%)	216 (55%)
Female n (%)	179 (45%)
Age median (range)	64 (1-95)
Bioptic samples (n)	**409**
Benign biopsies n (%)	215 (53%)
Malign biopsies n (%)	**194 (47%)**
Lung n (%)	50 (26%)
Breast n (%)	39 (20%)
Myeloma n (%)	27 (14%)
Lymphoma n (%)	18 (9%)
Prostate n (%)	13 (7%)
Colorectal n (%)	7 (4%)
Kidney n (%)	7 (4%)
Other n (%)	26 (13%)
Unknown n (%)	7 (4%)
Localization of bone lesions	**409 (100%)**
Spine n (%)	169 (41%)
Extremities n (%)	122 (30%)
Pelvis n (%)	59 (14%)
Thorax n (%)	31 (8%)
Head n (%)	28 (7%)
Total n (%)	**758 (100%)**
X-ray n (%)	223 (29%)
CT n (%)	233 (31%)
MRI n (%)	209 (28%)
BS n (%)	35 (5%)
PET/CT n (%)	58 (8%)

Reproduced with permission from European Journal of Radiology (102)

Table 4 shows the diagnostic characteristics of the imaging modalities. Overall, the nominal rank order of accuracy was MRI, PET/CT, CT, BS, and X-ray. The sensitivity for MRI and PET/CT was shown to be superior, both around 90%, whereas X-ray proved to have a very low sensitivity. X-

ray, on the other hand, proved to have the highest specificity and CT the highest PPV. The sensitivity of CT and BS was both around 75%. The specificity was in the 80-90% range for CT and MRI versus approximately 60% for BS and PET/CT. The PPV was acceptable (80% or greater) for most modalities, and so was NPV for MRI and PET/CT. In contrast, NPV was 66-71% for X-ray and CT, and only 41% for BS. MRI showed the best nominal combination of sensitivity, specificity and PPV, and NPV (102).

Statistical analysis proved a significant difference in sensitivity and specificity across the five modalities ($p<0.0001$). Subsequent pairwise comparisons found no difference between MRI and PET/CT ($p=1.00$), both being significantly more sensitive than CT ($p=0.0027$ and $p=0.025$, respectively) with only MRI being significantly better than bone scan ($p=0.047$). There was no difference between MRI and PET/CT ($p=0.067$) or MR vs. CT ($p=0.29$), but CT was more specific than PET/CT ($p=0.0091$). X-ray was significantly more specific than MRI ($p=0.001$), PET/CT ($p=0.0001$) and BS ($p=0.0065$), but not CT ($p=0.058$). It must be noted that the included numbers of BS were relatively low resulting in large confidence intervals (102).

Table 4. Diagnostic characteristics of the imaging modalities X-ray, CT, MRI, BS and PET-CT (reported with 95% confidence intervals)

	X-ray	CT	MRI	BS	PET-CT
	n=223	n=233	n=209	n=35	n=58
Sensitivity	33.0 (23.6 - 43.6)	75.6 (67.8 - 82.6)	90.5 (83.7 - 95.2)	74.1 (53.7 - 88.8)	92.3 (79.1 - 98.3)
Specificity	96.1 (91.2 - 98.7)	89.2 (81.1 - 94.7)	81.1 (73.6 - 89.8)	62.5 (24.7 - 91.0)	63.2 (38.4 - 83.7)
Accuracy	69.5 (63.0 - 75.5)	81.1 (75.5 – 85.9)	87.1 (81.8 – 91.5)	71.4 (53.7 – 85.4)	82.7 (70.6 – 91.4)
PPV	86.1 (70.5 - 95.3)	91.4 (84.7 - 95.8)	86.8 (79.4 - 92.2)	87.0 (66.4 - 97.1)	83.7 (69.3 - 93.2)
NPV	66.3 (59.1 - 73.0)	70.9 (61.8 - 79.0)	87.5 (78.7 - 93.6)	41.7 (15.3 - 72.3)	80.0 (51.9 - 95.4)

Abbreviations: positive predictive value (PPV); negative predictive value (NPV). Reproduced with permission from European Journal of Radiology (102)

Three subgroup analyses' possible influence on imaging characteristics were performed, one on the localization of the lesion (spine versus non-spine), one on the primary tumor type and one on the bone matrix response (osteolytic, osteosclerotic and mixed).

The influence of the localization of the lesion (spine versus non-spine) is displayed in table 5. In terms of spine lesions, MRI is better than the other modalities with a sensitivity on 92% and a specificity above 80%, whereas PET-CT was superior in non-spine lesions with a sensitivity on 96 % which for MRI was only 86% (102). There was a significant difference in sensitivity for the five imaging modalities ($p<0.0001$) among spine lesions as opposed to non-spine lesions ($p=0.15$), whereas there was a significant difference in specificity among the five modalities for both spine ($p=0.028$) and non-spine lesions ($p=0.0007$) (102). MRI, but not PET/CT, was significantly more

sensitive than CT (p=0.0005), but there was no difference in sensitivity between MRI and PET/CT (p=0.63) when pair-wised compared. The nominal rank order for accuracy was MRI, CT, PET/CT, BS and X-ray for spine lesions, and PET/CT, MRI, CT, X-ray, and BS for non-spine lesions (102).

Table 5. Diagnostic accuracy (%) of imaging modalities in study 1 on spine/non-spine lesions (reported with 95% confidence intervals).

Location	**Diagnostic outcome**	**Modality**				
Spine		X-ray (n=77)	CT (n=115)	MRI (n=133)	BS (n=14)	PET/CT (n=22)
	Sensitivity	26.7 (14.6 – 42.0)	70.5 (59.1 – 80.3)	92.0 (84.1 – 96.7)	72.7 (39.1 – 93.7)	87.5 (61.6 – 98.1)
	Specificity	90.6 (75.0 – 97.1)	94.6 (81.8 – 99.2)	80.4 (66.1 – 90.6)	66.7 (11.6 – 94.5)	50.0 (12.4 – 87.6)
	Accuracy	53.2 (41.5 – 64.7)	78.3 (69.6 – 85.4)	88.0 (81.2 – 93.0)	71.4 (41.9 – 91.6)	77.3 (54.6 – 92.2)
	PPV	80.0 (51.9 – 95.4)	96.5 (87.9 – 99.5)	89.8 (81.7 – 95.3)	88.9 (51.7 – 98.2)	82.4 (56.6 – 96.0)
	NPV	46.8 (34.0 - 59,9)	60.3 (46.6 – 73.0)	84.1 (69.9 – 93.3)	40.0 (6.5 – 84.6)	60.0 (15.4 – 93.1)
Non-spine		X-ray (n=146)	CT (n=118)	MRI (n=76)	BS (n=21)	PET/CT (n=36)
	Sensitivity	38.8 (25.2 – 53.8)	82.3 (70.5 – 90.8)	86.2 (68.3 – 96.0)	75.0 (47.6 – 92.6)	95.7 (78.0 – 99.3)
	Specificity	97.9 (92.7 – 99.7)	85.7 (73.8 – 93.6)	85.1 (71.7 – 93.8)	60.0 (15.4 – 93.5)	69.2 (38.6 – 90.7)
	Accuracy	78.1 (70.5 – 84.5)	83.9 (76.0 – 90.0)	85.5 (75.6 – 92.6)	71.4 (47.8 - 88.7)	86.1 (70.5 – 95.3)
	PPV	90.5 (69.6 – 98.6)	86.4 (75.0 – 93.9)	78.1 (60.0 – 90.7)	85.7 (57.2 – 97.8)	84.6 (65.1 – 95.6)
	NPV	76.0 (67.5 – 83.2)	81.4 (69.1 – 90.3)	90.9 (78.3 – 97.4)	42.9 (10.4 – 81.3)	90.0 (55.5 – 98.3)

Abbreviations: positive predictive value (PPV); negative predictive value (NPV). Reproduced with permission from European Journal of Radiology (102)

The sensitivity of PET/CT and MRI were 86% or more for all tumor types; however, some of the subgroups were very small. Likewise, CT proves 86% sensitivity for breast cancer, but only around 70% for other cancers. In general, X-ray had extremely low sensitivity. For lung cancer (p=0.0001), breast cancer (p=0.0013), myeloma (p=0.022), lymphoma (p=0.0004), and prostate cancer (0=0.023), there were substantial differences in sensitivity between modalities. Because of the small number of observations and consequently high probability of type II error, no post-test pairwise comparisons were conducted.

Table 6. Sensitivity (%) of imaging modalities on bone lesions classified by the most frequent primary tumors (reported with 95% confidence intervals).

	X-ray	CT	MRI	BS	PET/CT
Lung cancer	24.2 (11.1 – 42.3; n=33)	75.6 (59.7 – 87.6, n=41)	86.2 (68.33 – 96.0, n=29)	40.0 (6.5 – 84.6, n=5)	100.0 (71.3 – 100.0, n=11)
Breast cancer	47.1 (23.0 -72.1, n=17)	86.7 (69.3 – 96.2, n=30)	95.2 (76.1 – 99.2, n=21)	84.5 (54.5 – 97.6, n=13)	100.0 (19.3 – 100.0, n=2)
Multiple myeloma	46.2 (19.3 – 74.8, n=13)	73.7 (48.8 – 90.8, n=19)	93.8 (69.7 – 99.0, n=16)	100.0 (16.6 – 100.0, n=1)	100.0 (54.1 – 100.0, n=6)
Lymphoma	0.0 (0.0 - 37,1, n=8)	70.0 (34.8 – 93.00, n=10)	90.9 (61.5 – 98.6, n=11)	100.0 (16.6 – 100.0, n=1)	87.5 (51.7 – 98.2, n=8)
Prostate cancer	66.7 (30.1 – 92.1, n=9)	75.0 (20.3 – 95.9, n=4)	90.0 (55.5 – 98.3, n=10)	100.0 (40.2 – 100.0, n=4)	No data

Reproduced with permission from European Journal of Radiology (102)

The final subgroup analysis performed was on diagnostic characteristics of the imaging modalities distributed by bone response, illustrated in figure 7. The primary tumor's location of origin could be identified in 187 of 194 tumor biopsies, out of which imaging reports described osteolytic bone pattern in 105 tumors, osteoblastic bone pattern in 23 tumors and mixed lesions in 59 (102). In osteolytic lesions, all imaging modalities had at least 30 observations, with the exception of BS (n=6). With CT and MRI, there were at least 30 observations of mixed lesions, although fewer observations were obtained with X-ray (n=22), BS (n=15), and PET/CT (n=7). The number of images of osteosclerotic lesions was limited (10-18 with the radiological modalities), especially for the nuclear medicine methods (BS, n=6; PET/CT, n=1). No confidence intervals are reported in Figure 7 (102).

The sensitivity for the detection of osteolytic lesions showed a highly significant difference ($p<0.0001$) between modalities with MRI and PET-CT showing a significantly higher sensitivity than CT ($p=0.0079$ and 0.033, respectively). A significant difference across modalities was also shown in mixed lesions ($p=0.0002$), but no significant differences in sensitivity of CT versus MRI ($p=0.54$) (102). The amount of PET/CT data was too low to perform comparisons with radiological examinations. There were no significant differences in sensitivity among the modalities in osteosclerotic lesions ($p=0.26$). Pairwise comparisons of each modality within the three different bone patterns showed a significant difference in sensitivity for X-ray ($p=0.017$), a trend for BS ($p=0.098$), but no differences for CT, MRI, or PET/CT (data not shown) (102).

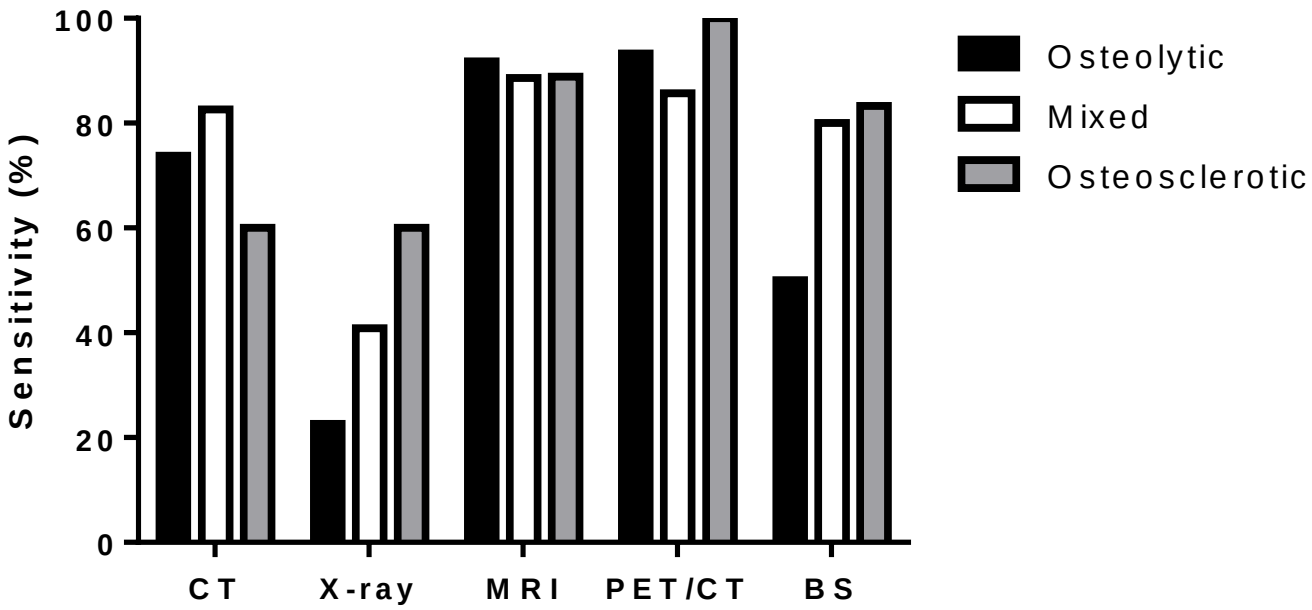


Figure 7. Sensitivity of imaging modalities for the demonstration of malignant bone lesions for which the tumor is known to induce predominantly osteolytic (n = 105), predominantly osteoblastic (n = 23), or mixed-type (n = 59) bone matrix response. There were significant differences among modalities in sensitivity for the detection of osteolytic ($p < 0.0001$) and mixed lesions ($p = 0.0002$), but not for osteosclerotic lesions. Due to the low number of observations in some groups (see text), the data are shown without confidence intervals. Reproduced with permission from European Journal of Radiology (102)

Study 2

Lange MB, Petersen LJ, Lausen M, Bruun NH, Nielsen MB, Zacho HD. Influence of Prior Imaging Information on Diagnostic Accuracy for Focal Skeletal Processes-A Retrospective Analysis of the Consistency between Biopsy-Verified Imaging Diagnoses. Diagnostics (Basel) 2022 Jul 17;12(7):1735. doi: 10.3390/diagnostics12071735

Study 2 included all patients having more than one imaging modality performed within 6 months of the biopsy and table 7 demonstrates patient demographics and baseline information. There were slightly more men than women. Most of the biopsies were described as malignant with lung cancer (31%), breast cancer (19%), multiple myeloma (12%) and lymphoma (11%) being the most frequent types of cancer. Histology specimens accounted for 90 %. The spine accounted for more than 50% of the localizations. In terms of imaging, X-ray, CT and MRI were almost equally represented. The majority of the biopsies (67%) had two imaging modalities performed in the 6-month period, three imaging modalities were performed in 30% of cases, and four imaging modalities were used in 3% of cases (103).

Table 7 Patient demographics and baseline information

Patients (n)	**207**
Male n (%)	116 (56%)
Female n (%)	91 (44%)
Age median (range)	67 (1-93)
Bioptic samples (n)	**216**
Benign biopsies n (%)	84 (39%)
Malignant biopsies n (%)	132 (61%)
Biopsy specimen (n)	**216**
Cytological n (%)	16 (8%)
Histological n (%)	195 (90%)
Dissection (%)	5 (2%)
Imaging modalities performed (n)	**464**
X-ray n (% of biopsies)	143 (31%)
CT n (% of biopsies)	169 (36%)
MRI n (% of biopsies)	152 (33%)
Localization of bone lesion (n)	**216**
Spine n (%)	119 (55%)
Extremities n (%)	39 (18%)
Pelvis n (%)	36 (17%)
Thorax and shoulders n (%)	19 (9%)
Head n (%)	3 (1%)

Reproduced from Diagnostics, Open Access (103)

When the imaging sequence was ignored, MRI was found to have the highest accuracy, followed by CT and X-ray as can be seen in table 8. X-ray and CT were shown to have the highest specificity and PPV, respectively, whereas MRI demonstrated the highest sensitivity and NPV (103).

Table 8. Diagnostic characteristics of the imaging modalities X-ray, CT, MRI (reported with 95% confidence intervals without taking the imaging sequence into consideration)

	X-ray	**CT**	**MRI**
	n=143	n=169	n=152
Sensitivity	31.3 (21.4-42.6)	73.5 (64.3-81.3)	92.1 (84.5-96.8)
Specificity	95.2 (86.7-99.0)	85.7 (73.8-93.6)	81.0 (69.1-89.8)
Accuracy	59.4 (50.9-67.6)	77.5 (70.5-83.6)	87.5 (81.2-92.3)
PPV	89.3 (71.8-97.7)	91.2 (83.4-96.1)	87.2 (78.8-93.2)
NPV	52.2 (42.7-61.6)	61.5 (49.8-72.3)	87.9 (76.7-95.01)

Reproduced from Diagnostics, Open Access (103)

In terms of sequencing, table 9-11 shows that there is no significant difference in accuracy among imaging modalities, whether preceded by another imaging modality or not, except for CT specificity and PPV, which decreased when preceded by MRI (103).

Table 9. Sensitivity, specificity, accuracy, PPV and NPV estimates (reported with 95% confidence intervals) for X-ray and CT without or with access to a preceding MRI

	X-ray			**CT**		
	Not preceded by MRI (n = 122)	Preceded by MRI (n = 21)	*p* value	Not preceded by MRI (n = 140)	Preceded by MRI (n = 29)	*p* value
Sensitivity	28.2 (18.1-40.1)	55.6 (21.2-86.3)	0.13	72.0 (61.8-80.9)	80.0 (56.3-94.3)	0.58
Specificity	98.0 (89.6-100.0)	83.3 (51.6-97.9)	0.09	93.6 (82.5-98.7)	44.4 (13.7-78.8)	0.00
Accuracy	70.0 (63.1-76.3)	65.2 (42.7-83.6)	0.34	79.3 (71.6-85.7)	69.0 (49.2-84.7)	0.23
PPV	95.2 (76.2-99.9)	71.4 (29.0-96.3)	0.15	95.7 (88.0-99.1)	76.2 (52.8-91.8)	0.01
NPV	49.5 (39.4-59.6)	71.4 (41.9-91.6)	0.16	62.9 (50.5-74.1)	50.0 (15.7-84.3)	0.48

Reproduced from Diagnostics, Open Access (103)

Table 10. Sensitivity, specificity, accuracy, PPV and NPV estimates (reported with 95% confidence intervals) for X-ray and MRI without or with access to a preceding CT

	X-ray			**MRI**		
	Not preceded by CT (n = 111)	Preceded by CT (n = 32)	*p* value	Not preceded by CT (n = 70)	Preceded by CT (n = 82)	*p* value
Sensitivity	28.6 (17.9-41.3)	41.2 (18.4-67.1)	0.38	91.7 (77.5-98.2)	92.5 (81.8-97.9)	1.00
Specificity	93.8 (82.8-98.7)	100.0 (78.2-100.0)	1.00	79.4 (62.1-91.3)	82.8 (64.2-94.2)	1.00
Accuracy	56.8 (47.0-66.1)	68.8 (50.0-83.9)	0.84	85.7 (75.3-92.9)	89.0 (80.2-94.9)	0.63
PPV	85.7 (63.7-97.0)	100.0 (59.0-100.0)	0.55	82.5 (67.2-92.7)	90.7 (79.7-96.9)	0.35
NPV	50.0 (39.9-60.7)	60.0 (38.7-78.9)	0.50	90.0 (73.5-97.9)	85.7 (67.3-96.0)	0.70

Reproduced from Diagnostics, Open Access (103)

Table 11. Sensitivity, specificity, accuracy, PPV and NPV estimates (reported with 95% confidence intervals) for CT and MRI without or with access to a preceding X-ray

	CT			MRI		
	Not preceded by X-ray (n = 98)	Preceded by X-ray (n = 71)	*p* value	Not preceded by X-ray (n = 83)	Preceded by X-ray (n = 69)	*p* value
Sensitivity	70.3 (57.6-88.1)	77.6 (63.4-88.2)	0.52	90.2 (78.6-96.7)	94.7 (82.3-99.4)	0.69
Specificity	91.2 (76.3-98.1)	77.3 (54.6-92.2)	0.24	71.9 (53.3-86.3)	90.3 (74.2-98.0)	0.11
Accuracy	77.6 (68.0-65.4)	77.5 (66.0-86.5)	1.00	83.1 (73.3-90.5)	92.8 (83.9-97.6)	0.09
PPV	93.8 (82.8-98.7)	88.4 (74.9-96.1)	0.47	83.6 (71.2-92.2)	92.3 (79.1-98.4)	0.35
NPV	62.0 (47.2-75.3)	60.7 (40.6-78.5)	1.00	82.1 (63.1-93.9)	93.3 (77.9-99.2)	0.25

Reproduced from Diagnostics, Open Access (103)

However, an interesting pattern was revealed when examining imaging sequences split between malignant and benign biopsies. Figure 8A shows that among malignant (positive) biopsies, if X-ray was false negative (75%) and used as the first imaging modality, only 7% of the subsequent MRI and 30% of the subsequent CT imaging were false negative, whereas if MRI was false negative (17%) and used as the first imaging modality, 100% of the subsequent CT scans were false negative as well (Figure 8E) (103). Similarly, in samples with benign histology, 100% of the subsequent CT imaging was false-positive if the MRI was false-positive (33%) (Figure 8F). Figure 8C shows that 100% of the subsequent X-ray exams were false negative when CT scans (30%) were false negative, while only 14% of the subsequent MRIs were false negative. 100 percent of the follow-up scans (CT/MRI and X-ray/MRI) for the few false-positive X-ray and CT exams were also false-positive (Figure 8B and D) (103).

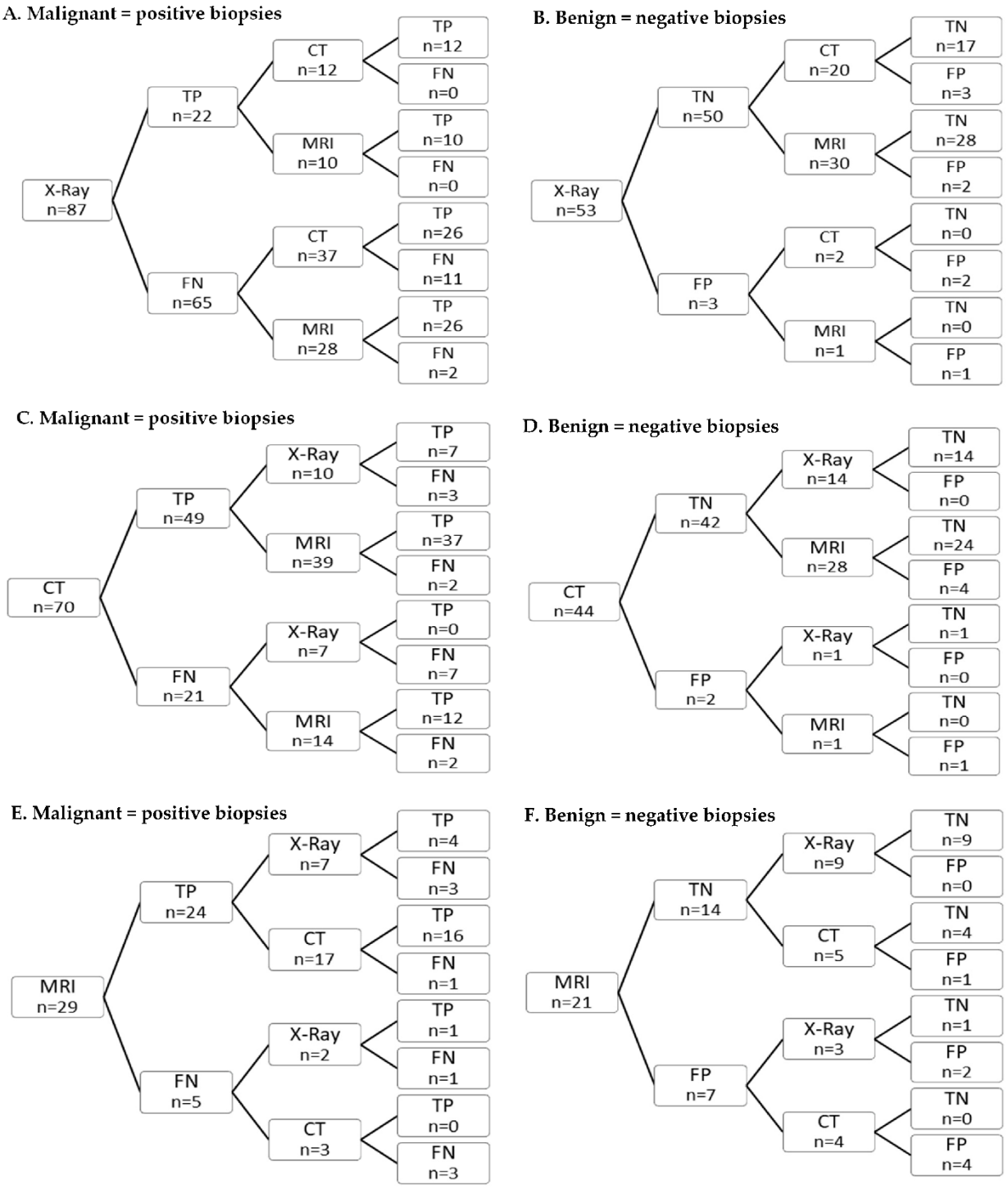


Figure 8. Sequence analyses. (A) Diagnostic results for malignant (positive) biopsies when X-ray is performed as the first modality (TP = true positive, FN = false negative). (B) Diagnostic results for benign (negative) biopsies when X-ray is performed as the first modality (TN = true negative; FP = false positive). (C) Diagnostic results for malignant (positive) biopsies when CT is performed as the first modality (TP = true positive, FN = false negative). (D) Diagnostic results for benign (negative) biopsies when CT is performed as the first modality (TN = true negative; FP = false positive). (E) Diagnostic results for malignant (positive) biopsies when MRI is performed as the first modality (TP = true positive, FN = false negative). (F) Diagnostic results for benign (negative) biopsies when MRI is performed as the first modality (TN = true negative; FP = false positive). Reproduced from Diagnostics, Open Access (103)

Study 3

Lange MB, Petersen LJ, Nielsen MB, Zacho HD. Validity of negative bone biopsy in suspicious bone lesions. Acta Radiol Open 2021 Jul 27;10(7):20584601211030662

As shown in figure 5, this study comprised a total of 215 benign bone samples from 207 patients. Most of the patients were men (57%) and the median age was 64, range 1-94. Eight patients underwent two further biopsies, 6 of which were repeat procedures from the same anatomy. In total, 57 patients (28%) had a cancer diagnosis before having their first bone biopsy, as shown in table 12.

Table 12 Types of prior history of cancer

Cancer types n= 59 (57 patients, 2 with 2 types)	**59 (100%)**
Breast	12 (20%)
Pulmonary (SCLC 9, NSCLC 2)	11 (19%)
Colorectal	9 (15%)
Prostate	5 (8%)
Sarcoma	4 (7%)
Oral cavity	4 (7%)
Non-Hodgins Lymphoma	4 (7%)
Urine bladder	3 (5%)
Malign melanoma	2 (3%)
Other (multiple myeloma, thyroid, cervix, esophagus, pancreas)	5 (9%)

Reproduced from Acta Radiologica, Open Access (104)

Most of the bone biopsies came from the extremities, as shown in table 13.

Table 13 Localization of the bone lesion

Total n (%)	**215 (100%)**
Spine n (%)	62 (29%)
Extremities n (%)	98 (46%)
Pelvis n (%)	17 (8%)
Thorax n (%)	12 (6%)
Head n (%)	26 (12%)

Reproduced from Acta Radiologica, Open Access (104)

Surgical interventions accounted for 163 of 215 (76%) biopsies. Most of these procedures (75/163, 46%) involved surgical resection from the affected anatomy, whereas 35, 24, and 21 biopsies (21, 15, and 13% respectively) were samples obtained during alloplastic surgery, osteosynthesis, and spondylodesis. CT-guided biopsies constituted only 4% of all biopsies, whereas

fluoroscopy-assisted biopsies accounted for 40% of non-surgical biopsies and 10% of all biopsies. The remaining samples included autopsies, arthroscopically collected material, and vertebroplasty.

A probably benign lesion was mentioned as indication for 126 (59%) of the 215 biopsies, 84 (39%) had no indication, and 5 (2%) had suspected malignancy as indication. In cases of discomfort, impending fracture risk, or cosmetic issues, surgical removal of benign lesions like cysts, enchondromas, non-ossifying fibromas, osteochondromas, and osteoid osteomas might be indicated. In conjunction with removal, samples might be sent for a pathological evaluation as well as biopsies might be taken in cases of slight uncertainty. As shown in table 14, the primary histological diagnoses were fibrosis, no malignancy and inflammation.

Table 14 Primary pathological diagnosis of the included 215 benign biopsies

Diagnoses	**n = 215 (100%)**
Inflammation	28 (13%)
No malignancy	24 (11%)
Fibrosis	22 (10%)
Osteochondroma	20 (9%)
Degenerative changes	14 (7%)
Unspecific reactive change	14 (7%)
Necrosis	14 (7%)
Fracture	13 (6%)
Other (cyst, exostosis, hemangioma, Paget, granuloma, hemorrhage a.o.)	66 (30%)

Reproduced from Acta Radiologica, Open Access (104)

According to our criteria, 210 of 215 biopsies were truly benign two years after the initial biopsy, as described in table 15. Three cases were questionably benign due to equivocal imaging (no description) (104).

Table 15. 2-year follow up upon validity of negative bone biopsy

Biopsies (n=215)	**n (%)**
Truly benign Negative biopsy or no biopsy from the same anatomy and No imaging with suspicion of malignancy from the same anatomy and Negative or no clinical suspicion of malignancy from the same anatomy	210 (98%) (95% CI 0.94-0.99)
Questionably benign Negative biopsy or no biopsy from the same anatomy and Equivocal imaging of the same anatomy and Negative or no clinical suspicion of malignancy from the same location	3 (1%) (95% CI 0.001-0.03)
Truly malignant Positive biopsy from the same anatomy or from adjacent soft tissue or Positive imaging or Clinical suspicion of malignancy from the relevant anatomy	2 (1%) (95% CI 0.001-0.03)

Reproduced from Acta Radiologica, Open Access (104)

Table 16 displays the pathology follow-up. Thirty-nine percent of the cases underwent further biopsies from the same, adjacent, or different structure during the two-year follow-up period. Two of them revealed the initial benign diagnosis to be incorrect.

Table 16 Pathological samples during the 2-year follow-up

Biopsies (n=226) *	n (%)	Benign	Malignant
No biopsies performed at all n (%)	138 (61%)		
Biopsies performed n (%)	88 (39%) *	n=65 (29%)	n=23 (10%)
Biopsy from exact same structure, n (%) *	6 (3%)	5 (2%)	1 (1%)
Biopsy from adjacent or other structure, n (%) *	82 (36%)	60 (27%)	22 (9%)

*Selected cases had more than one biopsy performed in the follow-up period. Reproduced from Acta Radiologica, Open Access (104)

In the first case, a percutaneous spinal decompression of L4 was performed in March 11. In conjunction with this, a biopsy was taken and described as benign. One month later, a second biopsy was performed while the patient was under general anesthesia, and this time the biopsy was classified as malignant (multiple myeloma). The localized change in L4 was consistently described as potentially malignant in imaging as well as in journal entries. The patient died three years later (104). In the second case, a right iliac bone biopsy under CT guidance was carried out in March 2013. Despite the sample being deemed benign, the following imaging descriptions consistently described the lesion as potentially malignant. Even though a second CT-guided biopsy was performed 6 months later and again described as benign, concerns regarding the sample's representativeness based on imaging and symptoms persisted. A third CT-guided biopsy from surrounding soft tissue was completed 8 months later, and this sample contained malignant cells from urinary bladder cancer. The patient died within the 2-year follow-up period (104).

Table 17. Diagnosis of malignant biopsies from same, adjacent, or different structures in the 2-year period

Diagnosis of malignant biopsies including (n=23)	n (%)
Skin	7 (31%)
Urinary bladder	4 (17%)
Oral cavity	2 (9%)
Pancreas	2 (9%)
SCLC	2 (9%)
Breast	2 (9%)
Small intestine	1 (4%)
Lymphoma	1 (4%)
Colorectal	1 (4%)
Multiple myeloma (L4)	1 (4%)

Reproduced from Acta Radiologica, Open Access (104)

The most typical locations for benign follow-up biopsies were the cervix (21%), skin (18%), and GI tract (18%), followed by the extremities and urinary system (7%, respectively). The outcomes of the malignant follow-up biopsies are shown in Table 17.

In the two-year follow-up period, 160 (74%) out of the 215 biopsies were followed by imaging of the same anatomic location and was in 97% (n=155) described as benign. In 2 cases from 2 patients the lesions were consistently described as malignant, and 3 cases had no description and were therefore classified as equivocal. In total, 205 imaging procedures were performed with 60 cases undergoing 2 or more. The most frequent imaging follow-up procedure was X-ray, followed by a CT and MRI (104). Please refer to Table 18.

Table 18. Imaging at the 2-year follow-up

Imaging results (n = 215)	n (%)
Imaging of same anatomy negative	155 (72%)
Imaging of same anatomy positive	2 (1%)
No imaging performed	55 (26%)
Equivocal imaging (not described)	3 (1%)
Types of Imaging performed at the 2-year follow-up (n=205; 60 had two or more imaging performed)	
X-ray	126 (62%)
Computed tomography (CT)	44 (21%)
Magnetic resonance imaging MRI	27 (13%)
Bone scintigraphy (BS)	4 (2%)
18F-fluorodeoxyglucose positron emission tomography (FDG-PET/CT)	4 (2%)

Reproduced from Acta Radiologica, Open Access (104)

Of the 59 biopsies taken from the 57 patients who had been diagnosed with cancer previously, 9 (15%) had no imaging performed in the follow-up period at all. Seventy % of patients received at least one diagnostic imaging modality excluding X-ray, with 21 receiving just one imaging modality, 22 receiving two different imaging modalities, and 7 receiving three. Nine (15%) patients had only X-rays done (104).

In 189 (88%) of the biopsies that underwent clinical follow-up, no malignancy was suspected in the same anatomical region as the initial bone biopsy. In 24 (11%) of the biopsies there were no journal entries identified. The 2 lesions later described as malignant were initially noted in the journal comments as suggestive of malignancy, and in subsequent reports the diagnosis of malignancy was described as supported by imaging and pathology (104).

Chapter 5. Discussion

Precise and early detection of bone metastases is essential for accurate disease staging, enabling treatment selection and prognosis estimation. Numerous studies on accuracy of standard diagnostic imaging have been performed, but there are several challenges with these studies and subsequent imaging guidelines, the most important being a valid reference standard that is consistently used. As stated earlier, only few can live up to the requirements as outlined in STARD (The Standards for Reporting of Diagnostic Accuracy) and no significant improvement in reporting quality over the last 10 years is seen (11,90,91,98-100,108).

The uniqueness of this PhD is that the starting point is a collection of consecutive samples of bone biopsy reports, based upon which we retrospectively identified diagnostic imaging in terms of X-ray, BS, CT, MRI, and FDG-PET/CT describing the lesions, from which the biopsies were taken. By doing so, we were able to determine a valid accuracy for each modality since we had a true gold standard to compare with. Therefore, the main achievements of this thesis is that not only have we provided valid imaging accuracy data, but we have also been able to demonstrate that bone biopsy can be used as a valid reference standard for future research and clinical decisions (102,104). Finally, we have been able to demonstrate that access to and the sequence of earlier diagnostic imaging and reports may serve as a bias to the accuracy of diagnosis (103).

Most biopsy material was histological biopsies, with only 6% of samples being cytology specimens and the reading and reporting of the pathology samples were considered appropriate with minimal bias in classifying malignancy (INDSÆT REF). Study 3 eliminated speculations whether potential false negative biopsies could serve as a bias, since 98% of the benign samples proved truly negative. The rate of malignant biopsies included in study 2 was higher than compared to study 1, which might be explained by the fact that only biopsies with more than one imaging was included in study 2, and the probability of malignant changes having more than one imaging performed is greater than for benign. In study 3, 59% of cases specified a benign lesion as a reason for biopsy. This outcome is consistent with those of Scheitza et al., who found that only 21% of their biopsies were taken to confirm a benign diagnosis; in all other cases, biopsies were carried out as a standard part of surgery, which may be necessary in cases of discomfort, impending fracture risk, or cosmetic issues and these samples may be sent for pathological analysis (109).

In study 3, we looked into the nature of the bioptic procedures as part of our investigation into potential sources of error for false negative samples. 76% of the biopsies were surgical, while only 4% and 10%, respectively, were CT-guided biopsies and fluoroscopy-assisted biopsies. Open biopsy has been the traditional "gold standard" approach for acquiring sufficient and representative tissue samples for the diagnosis of musculoskeletal diseases with a claimed accuracy rate of 98%, but with most tumor centers performing CT guided biopsies, which have a diagnostic yield between 70% and 89%, a reported accuracy between 61 and 98%, and fewer biopsy complications than open surgical biopsy, even in sclerotic bone lesions, recent results show that the results from

percutaneous biopsy can be highly effective and accurate (94,110-112). According to other studies, CT-guided biopsy has a drawback in that metabolically active lesions without distinct morphologies may not be reliably assessable, and the false-negative rate of such lesions may be significantly higher; one series shows that 18% of metabolically active lesions required open biopsy after needle biopsy (92,110).

Most pre-biopsy imaging modalities were radiological procedures, equally divided between X-ray, MR and CT (all >200 each) in studies 1 and 2, whereas the proportion of X-ray investigations performed in study 3 was greater, as would be expected for benign lesions, since the diagnosis of these frequently can be made purely using standard X-ray without the need for further imaging.

Our research found notable differences in diagnostic accuracy in general, in selected cancer types, in spine-versus non-spine lesions, and in bone lesions with predicted osteolytic, osteoblastic, or mixed bone matrix responses (102,103) . As recommended in STARD guidelines and Food and Drug Administration for diagnostic accuracy tests we used sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and accuracy as an expression for imaging characteristics (91,100). MRI and FDG PET/CT were found to be the most accurate modalities in general in studies 1 and 2 without taking the imaging sequence into consideration and the same results were shown in most of the subgroup analyses in study 1 (102,103). MRI was in general more specific, while PET/CT was slightly more sensitive. MRI performed significantly better for spine lesions than PET/CT. In study 1 and 2, the spine was the most frequent localization, in contrast to study 3, where it was the extremities, which can be explained by the fact that metastases are most frequently found in the spine, whereas benign lesions more often are localized in the extremities (113-115). X-ray was relatively insensitive in detecting bone metastases as compared to other imaging methods, being explained in part by the fact that bone lesions require a loss of more than 50% of the bone mineral content in order to be visible on plain radiographs (10,116) .

Due to the lack of any systematic studies using pathology as the only reference, it is challenging to compare our data with the findings of earlier research, where MRI is increasingly used as a reference test (38-40). According to our statistics, MRI has an overall sensitivity of 91% and specificity of 81% in study 1 (102) and 92% and 81 % in study 2 (103) and there should therefore be some doubt about the reliability of using MRI as a reference in imaging research. For instance, fat-suppressed sequences also exhibit hyperintense bone marrow oedema, which can cause misdiagnosis and with kappa values below 0.7, observer agreement for malignancy with MRI is not perfect (10,41). However, the information from recent systematic reviews and meta-analyses is consistent with our findings. In a comparison of FDG PET/CT, CT, MRI, and bone scans, Yang et al. found that MRI and PET/CT provided equivalent results, with these techniques being much more accurate than CT and BS for the diagnosis of bone metastases, and later updates have described accuracy supporting our findings (4,8,16,36,48). In a 2013 meta-analysis comparing FDG-PET/CT and gadolinium-enhanced MRI for the detection of bone metastases, Duo et al found that both imaging modalities had excellent diagnostic performance for the detection of bone metastases in

cancer patients and that their diagnostic accuracy was nearly equal (14). However, as is explained under limitations, only one in every four patients with an MRI in our study had a contrast-enhanced MRI.

When evaluating the diagnostic properties of imaging techniques, it is also important to take the primary tumor origin and especially the bone matrix reaction into account. The bone matrix response is known to depend on the primary cancer type (4). Multiple myeloma, renal cell carcinoma, malignant melanoma, non-small cell lung cancer (NSCLC), non-Hodgkin lymphoma, thyroid cancer, and most breast malignancies all exhibit osteolytic matrix response, which is characterized by destruction of normal bone (4). Small cell lung cancer (SCLC), Hodgkin lymphoma, and prostate cancer exhibit osteoblastic metastases, which are characterized by the deposition of new bone matrix. Mixed metastases, in which the patient has both osteolytic and osteoblastic lesions, are found in gastrointestinal cancers and 15-20% of breast cancers (4).

Bone scintigraphy exclusively reflects bone metabolism at areas with active bone mineralization, i.e., osteosclerotic components, making it of value for detection of prostate cancer and of limited use in predominant osteolytic lesions such as renal metastases and myeloma (33,34,42,102) . Our results gave some evidence of a similar tendency for CT and supported this observation with BS and X-ray (102). MRI and PET/CT performed equally well across bone matrix responses. This finding is probably due to the fact that MRI and FDG-PET/CT directly represent the extent of the tumor, whereas BS and, to a lesser extent, CT, reflect the response of the tumor-to-bone interface. Our conclusion that PET/CT is superior to BS are consistent with data from a systematic review (43). Regarding grading of malignancy, tumor-specific diagnostic criteria should be considered, and it is well-established that the avidity of FDG-uptake is related to tumor characteristics. The composition of the tumor types as well as the disease condition (staging or relapse/recurrent disease) should be considered when drawing conclusions from a meta-analysis made by Wu et al, comparing whole-body MRI versus BS for detection of bone metastatic tumors in mixed cancers, showing similar patient-based sensitivity and specificity for MRI and BS, respectively (34).

The spine, pelvis, ribs, and ends of long bones are preferred locations for metastases because of their high red marrow content, whereas extremities are more frequently the site of benign lesions (16,117,118). This corresponds to our findings, that in studies 1 and 2, the spine was the most frequent location representing respectively 41% and 55%, whereas in study 3 it was only 29% with extremities representing 46%. It is generally accepted that MRI is superior to CT when characterizing osseous and soft tissue features, but that CT is superior to MRI when characterizing cortical bone (44). According to our findings in study 1, CT has a 12% higher sensitivity for non-spine lesions than for spine lesions, whereas MRI accuracy is superior to the other modalities for spine lesions with a near to 90% accuracy, with no other reaching 80 %. Apart from this, the methods were largely comparable. It is unclear how much the differences are based upon tumor characteristics (soft tissue component) or anatomical mapping (tomographic versus planar imaging). While x-ray clearly performed worse in cases of spine lesions, BS revealed comparable

diagnostic findings in cases of spine and non-spine lesions. It has been demonstrated that whole-body BS using SPECT/CT improves specificity without compromising sensitivity in BS (45). Bone metastases most commonly affect the axial skeleton (9). As a result, the best techniques must be chosen for each patient and the routine use of X-ray examination for detection of possible metastases in the spine should be interpreted with caution, always leading to further imaging in cases of persistent clinical suspicion. The prevalence in our bioptic sample material of prostate cancer, the most common malignancy in men, is rather low. However, it is known from clinical experience that these patients rarely get a bone biopsy. Our results showed that MRI was superior to BS in staging skeletal metastases in prostate cancer, calling into question the widespread recommendation of BS (46,119,120).

In study 2 we did not find any significant difference on the accuracy of imaging whether preceded by another imaging report or not. These finding challenges current guidelines and earlier studies that stress the significance of consistently comparing actual imaging with previous tests and reports (70,72,75-79,81-84). Our findings might be explained by the small subgroups, however, there was an intriguing pattern that suggested that the reader might give more weight to earlier imaging that was believed to be more accurate than the one under study. When X-ray is used as the first modality, 75% of cases were expectedly false negative, but only 7% of subsequent MRI scans and 30% of subsequent CT scans also resulted in false negative results, in contrast to when MRI was used as the first modality, where only 17% were false negatives, all subsequent CT scans were also false negative. When CT was used as the first modality with 30% of the results being false negative, then all 7 subsequent X-rays were also false negative (103), but this was only the case for 14% of the subsequent MRI scans (103). This might represent another explanation as to why we were not able to document a significant difference in accuracy, since if the MRI is accurate, there is a trend towards the subsequent CT or X-ray is more likely to be accurate, and when the MRI is inaccurate, the subsequent CT or X-ray is inaccurate in more than 80% of cases. These two situations might counteract one another, so that the accuracy does not change significantly compared to whether a modality is not preceded by MRI. It has been shown previously that if one looks at a prior negative report before looking at imaging studies, there is a greater chance of missing a significant abnormality than by looking at the imaging studies first, but in these studies imaging included only X-ray, so no data are to best of our knowledge available to support our results (70,77,80,81,84).

Based upon our findings, it might be reasonable for new guidelines to warn radiologists against being overly influenced by prior imaging, especially if those modalities are typically thought to be more accurate than the current one and that biopsy should be considered the gold standard for a valid diagnosis.

In study 3, we were indeed able to document that biopsy can be considered a true gold standard with 98 % of the bone biopsies described as benign being true negative after 2 years, thereby proving a high validity of negative bone biopsies as an expression of the absence of skeletal

metastases, which to the best of our knowledge has not been previously documented. Three cases were questionably benign due to no description of the imaging, and only two were actual false negatives.

Two biopsies (1%) turned out to be false negatives in repeated biopsies performed due to imaging (CT and MRI) persistently describing them as malignant. According to Monfardini et al., a negative biopsy should in cases of suspicious PET-CT and/or MR findings be carefully evaluated and considered for a second sampling, since they found that 8 out of 10 false negative CT-guided biopsies had positive PET scans and 6 out of 10 had positive MRI scans (94). Inflammation, no malignancy, fibrosis, osteochondroma, and degenerative changes were the most common pathological diagnoses in study 3, which is consistent with prior findings and suggests that our material is typical (121).

Study 3 documents that a negative sample can be truly negative, thereby supporting the validity of biopsy as the gold standard for identifying the absence of metastases. This study has two major implications. Primarily, since bone biopsy from a suspicious lesion in patients with a known primary cancer has been described as benign in 21% of the cases, it is of utmost importance that these do not represent false negative samples, resulting in delayed diagnosis and consequently increased morbidity and mortality (4,9,13,94,121). Furthermore, as highlighted by STARD (The Standards for Reporting of Diagnostic Accuracy), considering biopsy as the gold standard for demonstrating or excluding bone metastases, our result also has significant implications for the validity of scientific studies of treatment efficacy and diagnostic accuracy using biopsy as a reference (91).

There are limitations to our studies. The main being the retrospective nature of the data, including selection of patients, and the lack of uniform and systematized follow-up including standardized re-biopsy from the same anatomy. Second, the pathologists did not perform a blind evaluation of the specimen and were aware of the patient’s history, as in a normal clinical setting, which might have a possible effect on the final diagnosis (104). Radiology reports were often prepared by one radiologist alone or by supervising a radiologist in training, but nuclear medicine reports were always double read. Even though data from the reports were collected by two independent readers with little variation in reporting, each report might have been prejudiced, according to Study 2's findings that access to prior imaging and reports may function as a bias for subsequent investigations (103). Furthermore, the readers had possible access to clinical information via Clinical Suite, which in case they used this opportunity might have influenced the accuracy (85-88).

In study 3, the high percentage of surgical biopsies might have contributed to the high quality of samples and therefore to the low incidence of false negative results. Furthermore, spinal biopsies accounted for 29% and non-spinal biopsies for 71% (104). The accuracy of CT-guided biopsy is significantly influenced by the anatomical site, according to research by Hau et al., with non-spinal sites showing higher accuracy (75%) than spinal sites (61%). Hau further demonstrated that gender, lesion size, and kind of margin did not affect the success or failure rates of the biopsies

(22). Because we only encountered 2 false negative and 3 equivocal lesions, our data set is too small to support those findings, but the high number of non-spinal lesions might have contributed to our high bioptic diagnostic accuracy.

It can also be debated whether technical settings for the imaging methods studied could influence the findings. For example, only one in every four MRI imaging received contrast (102). Some authors have advocated for wide routine use of gadolinium contrast in the MRI evaluation of skeletal malignancies, but a recent meta-analysis showed no substantial differences in diagnostic value of the contrast enhancement (36,47). No diffusion-weighed MRI was used and studies have shown that the sensitivity for using DWI sequences is significantly higher than that for not using DWI sequences in detection of bone metastases, with lower specificity on a per-patient basis (36,122).

Our data may reflect diagnostic characteristics of different populations because all imaging techniques were not used on all individuals. A patient who has had only one X-ray examination performed, probably does not represent the same patient type as one who has had three or more imaging modalities performed. Demonstrating similarities or differences within patient groups is challenging as most patients underwent multiple imaging modalities. Furthermore, some of the subgroups were relatively small and it would be recommendable to clarify a possible relationship between imaging modalities and clinical factors in a separate study.

In study 3, only 28 % had a confirmed diagnosis of prior cancer, out of which approximately 50% would cause mainly osteolytic, 25% mainly osteoblastic and 25% mixed metastatic bone response, so our results might not be applicable to all cancer patients with suspicious bone lesions. The most frequently performed imaging modality in the follow-up period was X-ray (62 %) (104). A rather low X-ray sensitivity (33 %) for diagnosis of skeletal malignancies has been demonstrated, and thus this type of examination might have missed possible positive lesions and a subsequent repeated biopsy (20). However, all samples were described as negative and 88% of the clinical follow-up did not raise any suspicion of malignancy, and thus this impact might be limited.

The statistical power may have been harmed by the low number of nuclear medicine exams overall and radiological investigations in some of the subgroup analyses. However, the conclusions of the statistical analyses were largely in line with the quantitative/numerical estimates of diagnostic characteristics.

In conclusion, we believe that our PhD has provided state-of-the-art documentation on diagnostic characteristics among the most frequently used imaging modalities for detection of skeletal metastases with MRI and FDG-PET/CT proved superior to CT followed by x-ray and BS with some distance. However, even MRI and PET/CT only proved a sensitivity around 90% indicating that imaging reports should be interpreted with caution and always be seen as part of the overall patient evaluation keeping in mind that bone matrix response and localization of skeletal lesions may influence the performance of different imaging of skeletal metastases. Secondly, we believe

that we have raised awareness around the possible influence of previous imaging and reporting on diagnostic accuracy. Therefore, guidelines for diagnostic strategies for skeletal metastases should be prepared baring type, location and primary tumor in mind and also address the issue of access to prior reports and clinical information in mind. New prospective studies on this topic are needed. Finally, we believe that our results prove it is reasonable to assess a negative bone biopsy as an indication of the absence of bone metastasis.

These results offer value not only to diagnosis, morbidity, and mortality of metastatic bone disease but also to the accuracy of future treatment and diagnostic scientific studies.

References

(1) Schmid-Alliana A, Schmid-Antomarchi H, Al-Sahlanee R, Lagadec P, Scimeca JC, Verron E. Understanding the Progression of Bone Metastases to Identify Novel Therapeutic Targets. Int J Mol Sci 2018 Jan 4;19(1):148. doi: 10.3390/ijms19010148.

(2) Bădilă AE, Rădulescu DM, Niculescu AG, Grumezescu AM, Rădulescu M, Rădulescu AR. Recent Advances in the Treatment of Bone Metastases and Primary Bone Tumors: An Up-to-Date Review. Cancers (Basel) 2021 Aug 23;13(16):4229. doi: 10.3390/cancers13164229.

(3) Bernard S, Walker E, Raghavan M. An Approach to the Evaluation of Incidentally Identified Bone Lesions Encountered on Imaging Studies. AJR Am J Roentgenol 2017 May;208(5):960-970.

(4) Macedo F, Ladeira K, Pinho F, Saraiva N, Bonito N, Pinto L, et al. Bone Metastases: An Overview. Oncol Rev 2017 May 9;11(1):321.

(5) Kimura T. Multidisciplinary Approach for Bone Metastasis: A Review. Cancers (Basel) 2018 May 24;10(6):156. doi: 10.3390/cancers10060156.

(6) Huang JF, Shen J, Li X, Rengan R, Silvestris N, Wang M, et al. Incidence of patients with bone metastases at diagnosis of solid tumors in adults: a large population-based study. Ann Transl Med 2020 Apr;8(7):482.

(7) Piccioli A, Maccauro G, Spinelli MS, Biagini R, Rossi B. Bone metastases of unknown origin: epidemiology and principles of management. J Orthop Traumatol 2015 Jun;16(2):81-86.

(8) Coleman RE, Croucher PI, Padhani AR, Clézardin P, Chow E, Fallon M, et al. Bone metastases. Nat Rev Dis Primers 2020 Oct 15;6(1):83-020-00216-3.

(9) Coleman RE. Clinical features of metastatic bone disease and risk of skeletal morbidity. Clin Cancer Res 2006 Oct 15;12(20 Pt 2):6243s-6249s.

(10) Balliu E, Boada M, Pelaez I, Vilanova JC, Barcelo-Vidal C, Rubio A, et al. Comparative study of whole-body MRI and bone scintigraphy for the detection of bone metastases. Clin Radiol 2010 Dec;65(12):989-996.

(11) Isaac A, Dalili D, Dalili D, Weber MA. State-of-the-art imaging for diagnosis of metastatic bone disease. Radiologe 2020 Nov;60(Suppl 1):1-16.

(12) Liu D, Kuai Y, Zhu R, Zhou C, Tao Y, Han W, et al. Prognosis of prostate cancer and bone metastasis pattern of patients: a SEER-based study and a local hospital based study from China. Sci Rep 2020 Jun 4;10(1):9104-020-64073-6.

(13) Disibio G, French SW. Metastatic patterns of cancers: results from a large autopsy study. Arch Pathol Lab Med 2008 Jun;132(6):931-939.

(14) Duo J, Han X, Zhang L, Wang G, Ma Y, Yang Y. Comparison of FDG PET/CT and gadolinium-enhanced MRI for the detection of bone metastases in patients with cancer: a meta-analysis. Clin Nucl Med 2013 May;38(5):343-348.

(15) Filippiadis D, Mazioti A, Kelekis A. Percutaneous, Imaging-Guided Biopsy of Bone Metastases. Diagnostics (Basel) 2018 Apr 18;8(2):25. doi: 10.3390/diagnostics8020025.

(16) O'Sullivan GJ, Carty FL, Cronin CG. Imaging of bone metastasis: An update. World J Radiol 2015 Aug 28;7(8):202-211.

(17) Riihimäki M, Thomsen H, Sundquist K, Sundquist J, Hemminki K. Clinical landscape of cancer metastases. Cancer Med 2018 Nov;7(11):5534-5542.

(18) Fornetti J, Welm AL, Stewart SA. Understanding the Bone in Cancer Metastasis. J Bone Miner Res 2018 Dec;33(12):2099-2113.

(19) Sailer V, Schiffman MH, Kossai M, Cyrta J, Beg S, Sullivan B, et al. Bone biopsy protocol for advanced prostate cancer in the era of precision medicine. Cancer 2018 Mar 1;124(5):1008-1015.

(20) Shah LM, Salzman KL. Imaging of spinal metastatic disease. Int J Surg Oncol 2011;2011:769753.

(21) Argentiero A, Solimando AG, Brunetti O, Calabrese A, Pantano F, Iuliani M, et al. Skeletal Metastases of Unknown Primary: Biological Landscape and Clinical Overview. Cancers (Basel) 2019 Aug 29;11(9):1270. doi: 10.3390/cancers11091270.

(22) Chen F, Han Y, Kang Y. Bone marrow niches in the regulation of bone metastasis. Br J Cancer 2021 Jun;124(12):1912-1920.

(23) Guerrieri AN, Montesi M, Sprio S, Laranga R, Mercatali L, Tampieri A, et al. Innovative Options for Bone Metastasis Treatment: An Extensive Analysis on Biomaterials-Based Strategies for Orthopedic Surgeons. Front Bioeng Biotechnol 2020 Oct 6;8:589964.

(24) Roodman DG. Mechanisms of bone metastases. In UpToDate. 2015; . Accessed 01, 2015.

(25) Forsberg JA, Eberhardt J, Boland PJ, Wedin R, Healey JH. Estimating survival in patients with operable skeletal metastases: an application of a bayesian belief network. PLoS One 2011;6(5):e19956.

(26) Sambri A, Fiore M, Giannini C, Pipola V, Zucchini R, Aparisi Gomez MP, et al. Primary Tumors of the Sacrum: Imaging Findings. Curr Med Imaging 2022;18(2):170-186.

(27) Schulman KL, Kohles J. Economic burden of metastatic bone disease in the U.S. Cancer 2007 Jun 1;109(11):2334-2342.

(28) Svendsen ML, Gammelager H, Svaerke C, Yong M, Chia VM, Christiansen CF, et al. Hospital visits among women with skeletal-related events secondary to breast cancer and bone metastases: a nationwide population-based cohort study in Denmark. Clin Epidemiol 2013;5:97-103.

(29) Pockett RD, Castellano D, McEwan P, Oglesby A, Barber BL, Chung K. The hospital burden of disease associated with bone metastases and skeletal-related events in patients with breast cancer, lung cancer, or prostate cancer in Spain. Eur J Cancer Care (Engl) 2010 Nov;19(6):755-760.

(30) Parkes A, Warneke CL, Clifton K, Al-Awadhi A, Oke O, Pestana RC, et al. Prognostic Factors in Patients with Metastatic Breast Cancer with Bone-Only Metastases. Oncologist 2018 Nov;23(11):1282-1288.

(31) von Moos R, Costa L, Gonzalez-Suarez E, Terpos E, Niepel D, Body JJ. Management of bone health in solid tumours: From bisphosphonates to a monoclonal antibody. Cancer Treat Rev 2019 Jun;76:57-67.

(32) Zacho HD, Mørch CD, Barsi T, Mortensen JC, Bertelsen H, Petersen LJ. Unexplained Bone Pain Is an Independent Risk Factor for Bone Metastases in Newly Diagnosed Prostate Cancer: A Prospective Study. Urology 2017 Jan;99:148-154.

(33) Pesapane F, Downey K, Rotili A, Cassano E, Koh DM. Imaging diagnosis of metastatic breast cancer. Insights Imaging 2020 Jun 16;11(1):79-020-00885-4.

(34) Wu Q, Yang R, Zhou F, Hu Y. Comparison of whole-body MRI and skeletal scintigraphy for detection of bone metastatic tumors: a meta-analysis. Surg Oncol 2013 Dec;22(4):261-266.

(35) Qu X, Huang X, Yan W, Wu L, Dai K. A meta-analysis of ^{18}FDG-PET-CT, ^{18}FDG-PET, MRI and bone scintigraphy for diagnosis of bone metastases in patients with lung cancer. Eur J Radiol 2012 May;81(5):1007-1015.

(36) Yang HL, Liu T, Wang XM, Xu Y, Deng SM. Diagnosis of bone metastases: a meta-analysis comparing (1)(8)FDG PET, CT, MRI and bone scintigraphy. Eur Radiol 2011 Dec;21(12):2604-2617.

(37) Wondergem M, van der Zant, F. M., van der Ploeg T, Knol RJ. A literature review of 18F-fluoride PET/CT and 18F-choline or 11C-choline PET/CT for detection of bone metastases in patients with prostate cancer. Nucl Med Commun 2013 Oct;34(10):935-945.

(38) Kruger S, Buck AK, Mottaghy FM, Hasenkamp E, Pauls S, Schumann C, et al. Detection of bone metastases in patients with lung cancer: 99mTc-MDP planar bone scintigraphy, 18F-fluoride PET or 18F-FDG PET/CT. Eur J Nucl Med Mol Imaging 2009 Nov;36(11):1807-1812.

(39) Nozaki T, Yasuda K, Akashi T, Fuse H. Usefulness of single photon emission computed tomography imaging in the detection of lumbar vertebral metastases from prostate cancer. Int J Urol 2008 Jun;15(6):516-519.

(40) Poulsen MH, Petersen H, Hoilund-Carlsen P, Jakobsen JS, Gerke O, Karstoft J, et al. Spine metastases in prostate cancer: comparison of technetium-99m-MDP whole-body bone scintigraphy, [(18) F]choline positron emission tomography(PET)/computed tomography (CT) and [(18) F]NaF PET/CT. BJU Int 2014 Dec;114(6):818-823.

(41) Stecco A, Lombardi M, Leva L, Brambilla M, Negru E, Delli Passeri S, et al. Diagnostic accuracy and agreement between whole-body diffusion MRI and bone scintigraphy in detecting bone metastases. Radiol Med 2013 Apr;118(3):465-475.

(42) Collins CD. Multiple myeloma. Cancer Imaging 2010 Feb 11;10:20-31.

(43) van Lammeren-Venema D, Regelink JC, Riphagen II, Zweegman S, Hoekstra OS, Zijlstra JM. (1)(8)F-fluoro-deoxyglucose positron emission tomography in assessment of myeloma-related bone disease: a systematic review. Cancer 2012 Apr 15;118(8):1971-1981.

(44) Pettersson H, Gillespy T,3, Hamlin DJ, Enneking WF, Springfield DS, Andrew ER, et al. Primary musculoskeletal tumors: examination with MR imaging compared with conventional modalities. Radiology 1987 Jul;164(1):237-241.

(45) Palmedo H, Marx C, Ebert A, Kreft B, Ko Y, Turler A, et al. Whole-body SPECT/CT for bone scintigraphy: diagnostic value and effect on patient management in oncological patients. Eur J Nucl Med Mol Imaging 2014 Jan;41(1):59-67.

(46) Lecouvet FE, Geukens D, Stainier A, Jamar F, Jamart J, d'Othee BJ, et al. Magnetic resonance imaging of the axial skeleton for detecting bone metastases in patients with high-risk prostate cancer: diagnostic and cost-effectiveness and comparison with current detection strategies. J Clin Oncol 2007 Aug 1;25(22):3281-3287.

(47) May DA, Good RB, Smith DK, Parsons TW. MR imaging of musculoskeletal tumors and tumor mimickers with intravenous gadolinium: experience with 242 patients. Skeletal Radiol 1997 Jan;26(1):2-15.

(48) Heindel W, Gübitz R, Vieth V, Weckesser M, Schober O, Schäfers M. The diagnostic imaging of bone metastases. Dtsch Arztebl Int 2014 Oct 31;111(44):741-747.

(49) Nascimento D, Suchard G, Hatem M, de Abreu A. The role of magnetic resonance imaging in the evaluation of bone tumours and tumour-like lesions. Insights Imaging 2014 Aug;5(4):419-440.

(50) Miller TT. Bone tumors and tumorlike conditions: analysis with conventional radiography. Radiology 2008 Mar;246(3):662-674.

(51) DenOtter TD, Schubert J. Hounsfield Unit. StatPearls Treasure Island (FL): StatPearls Publishing LLC; 2022.

(52) Alawi M, Begum A, Harraz M, Alawi H, Bamagos S, Yaghmour A, et al. Dual-Energy X-Ray Absorptiometry (DEXA) Scan Versus Computed Tomography for Bone Density Assessment. Cureus 2021 Feb 10;13(2):e13261.

(53) Chang CY, Garner HW, Ahlawat S, Amini B, Bucknor MD, Flug JA, et al. Society of Skeletal Radiology- white paper. Guidelines for the diagnostic management of incidental solitary bone lesions on CT and MRI in adults: bone reporting and data system (Bone-RADS). Skeletal Radiol 2022 Mar 28.

(54) Del Grande F, Farahani SJ, Carrino JA, Chhabra A. Bone marrow lesions: A systematic diagnostic approach. Indian J Radiol Imaging 2014 Jul;24(3):279-287.

(55) Stecco A, Trisoglio A, Soligo E, Berardo S, Sukhovei L, Carriero A. Whole-Body MRI with Diffusion-Weighted Imaging in Bone Metastases: A Narrative Review. Diagnostics (Basel) 2018 Jul 9;8(3):45. doi: 10.3390/diagnostics8030045.

(56) Cook GJ. PET and PET/CT imaging of skeletal metastases. Cancer Imaging 2010 Jul 19;10(1):1-8.

(57) Pi Y, Zhao Z, Xiang Y, Li Y, Cai H, Yi Z. Automated diagnosis of bone metastasis based on multi-view bone scans using attention-augmented deep neural networks. Med Image Anal 2020 Oct;65:101784.

(58) Koizumi M, Motegi K, Koyama M, Terauchi T, Yuasa T, Yonese J. Diagnostic performance of a computer-assisted diagnosis system for bone scintigraphy of newly developed skeletal metastasis in prostate cancer patients: search for low-sensitivity subgroups. Ann Nucl Med 2017 Aug;31(7):521-528.

(59) Chang CY, Gill CM, Joseph Simeone F, Taneja AK, Huang AJ, Torriani M, et al. Comparison of the diagnostic accuracy of 99 m-Tc-MDP bone scintigraphy and 18 F-FDG PET/CT for the detection of skeletal metastases. Acta Radiol 2016 Jan;57(1):58-65.

(60) Fonager RF, Zacho HD, Langkilde NC, Fledelius J, Ejlersen JA, Haarmark C, et al. Diagnostic test accuracy study of (18)F-sodium fluoride PET/CT, (99m)Tc-labelled diphosphonate SPECT/CT, and planar bone scintigraphy for diagnosis of bone metastases in newly diagnosed, high-risk prostate cancer. Am J Nucl Med Mol Imaging 2017 Nov 1;7(5):218-227.

(61) Liu FY, Sheng TW, Tseng JR, Yu KJ, Tsui KH, Pang ST, et al. Prostate-specific membrane antigen (PSMA) fusion imaging in prostate cancer: PET-CT vs PET-MRI. Br J Radiol 2022 Mar 1;95(1131):20210728.

(62) Hofman MS, Hicks RJ, Maurer T, Eiber M. Prostate-specific Membrane Antigen PET: Clinical Utility in Prostate Cancer, Normal Patterns, Pearls, and Pitfalls. Radiographics 2018 Jan-Feb;38(1):200-217.

(63) Michalski K, Jilg CA, Engelhardt M, Meyer PT, Ruf J. Diffuse Bone Marrow Involvement of Multiple Myeloma on [18F]PSMA-1007 PET/CT: Is There a Theranostic Potential? Clin Nucl Med 2022 May 27.

(64) Wei Y, Cheng K, Fu Z, Zheng J, Mu Z, Zhao C, et al. (18)F]AlF-NOTA-FAPI-04 PET/CT uptake in metastatic lesions on PET/CT imaging might distinguish different pathological types of lung cancer. Eur J Nucl Med Mol Imaging 2022 Apr;49(5):1671-1681.

(65) Jenkins P, MacCormick A, Stokes J, Lyall F, Rogers A, Gafoor N. Learning from mistakes when reporting urgent and emergency vascular studies. Clin Radiol 2022 Mar;77(3):159-166.

(66) Ahn Y, Hong GS, Park KJ, Lee CW, Lee JH, Kim SO. Impact of diagnostic errors on adverse outcomes: learning from emergency department revisits with repeat CT or MRI. Insights Imaging 2021 Nov 3;12(1):160-021-01108-0.

(67) Busby LP, Courtier JL, Glastonbury CM. Bias in Radiology: The How and Why of Misses and Misinterpretations. Radiographics 2018 Jan-Feb;38(1):236-247.

(68) Pinto A, Brunese L, Pinto F, Reali R, Daniele S, Romano L. The concept of error and malpractice in radiology. Semin Ultrasound CT MR 2012 Aug;33(4):275-279.

(69) Robinson JD, Hippe DS, Deconde RP, Zecevic M, Mehta N. Emergency Radiology: An Underappreciated Source of Liability Risk. J Am Coll Radiol 2020 Jan;17(1):42-45.

(70) Raskin MM. Survival strategies for radiology: some practical tips on how to reduce the risk of being sued and losing. J Am Coll Radiol 2006 Sep;3(9):689-693.

(71) Pinto A, Brunese L. Spectrum of diagnostic errors in radiology. World J Radiol 2010 Oct 28;2(10):377-383.

(72) Bruno MA, Walker EA, Abujudeh HH. Understanding and Confronting Our Mistakes: The Epidemiology of Error in Radiology and Strategies for Error Reduction. Radiographics 2015 Oct;35(6):1668-1676.

(73) Degnan AJ, Ghobadi EH, Hardy P, Krupinski E, Scali EP, Stratchko L, et al. Perceptual and Interpretive Error in Diagnostic Radiology-Causes and Potential Solutions. Acad Radiol 2019 Jun;26(6):833-845.

(74) Abujudeh HH, Boland GW, Kaewlai R, Rabiner P, Halpern EF, Gazelle GS, et al. Abdominal and pelvic computed tomography (CT) interpretation: discrepancy rates among experienced radiologists. Eur Radiol 2010 Aug;20(8):1952-1957.

(75) Kahn CE,J, Langlotz CP, Burnside ES, Carrino JA, Channin DS, Hovsepian DM, et al. Toward best practices in radiology reporting. Radiology 2009 Sep;252(3):852-856.

(76) Standards for interpretation and reporting of imaging investigations; second edition. The Royal College of Radiologists. 2018; . Accessed 10, 2021.

(77) Berbaum KS, Smith WL. Use of reports of previous radiologic studies. Acad Radiol 1998 Feb;5(2):111-114.

(78) Berlin L. Comparing new radiographs with those obtained previously. Am J Roentgenol 1999 01/01; 2021;172(1):3-6.

(79) European Society of Radiology (ESR). Good practice for radiological reporting. Guidelines from the European Society of Radiology (ESR). Insights Imaging 2011 Apr;2(2):93-96.

(80) Brady A, Laoide RÓ, McCarthy P, McDermott R. Discrepancy and error in radiology: concepts, causes and consequences. Ulster Med J 2012 Jan;81(1):3-9.

(81) Berlin L. Malpractice issues in radiology. Alliterative errors. AJR Am J Roentgenol 2000 Apr;174(4):925-931.

(82) White K, Berbaum K, Smith WL. The role of previous radiographs and reports in the interpretation of current radiographs. Invest Radiol 1994 Mar;29(3):263-265.

(83) Aideyan UO, Berbaum K, Smith WL. Influence of prior radiologic information on the interpretation of radiographic examinations. Acad Radiol 1995 Mar;2(3):205-208.

(84) Hunter TB, Boyle RR,J. The value of reading the previous radiology report. AJR Am J Roentgenol 1988 Mar;150(3):697-698.

(85) Hattori S, Yokota H, Takada T, Horikoshi T, Takishima H, Mikami W, et al. Impact of clinical information on CT diagnosis by radiologist and subsequent clinical management by physician in acute abdominal pain. Eur Radiol 2021 Aug;31(8):5454-5463.

(86) Berbaum KS, Franken EA,J. Commentary does clinical history affect perception? Acad Radiol 2006 Mar;13(3):402-403.

(87) Doubilet P, Herman PG. Interpretation of radiographs: effect of clinical history. AJR Am J Roentgenol 1981 Nov;137(5):1055-1058.

(88) Loy CT, Irwig L. Accuracy of diagnostic tests read with and without clinical information: a systematic review. JAMA 2004 Oct 6;292(13):1602-1609.

(89) Weber MA, Bazzocchi A, Nöbauer-Huhmann IM. Tumors of the Spine: When Can Biopsy Be Avoided? Semin Musculoskelet Radiol 2022 Aug;26(4):453-468.

(90) Korevaar DA, van Enst WA, Spijker R, Bossuyt PM, Hooft L. Reporting quality of diagnostic accuracy studies: a systematic review and meta-analysis of investigations on adherence to STARD. Evid Based Med 2014 Apr;19(2):47-54.

(91) Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 2015 Oct 28;351:h5527.

(92) Hau A, Kim I, Kattapuram S, Hornicek FJ, Rosenberg AE, Gebhardt MC, et al. Accuracy of CT-guided biopsies in 359 patients with musculoskeletal lesions. Skeletal Radiol 2002 Jun;31(6):349-353.

(93) Rimondi E, Rossi G, Bartalena T, Ciminari R, Alberghini M, Ruggieri P, et al. Percutaneous CT-guided biopsy of the musculoskeletal system: results of 2027 cases. Eur J Radiol 2011 Jan;77(1):34-42.

(94) Monfardini L, Preda L, Aurilio G, Rizzo S, Bagnardi V, Renne G, et al. CT-guided bone biopsy in cancer patients with suspected bone metastases: retrospective review of 308 procedures. Radiol Med 2014 Nov;119(11):852-860.

(95) Park VY, Kim EK, Moon HJ, Yoon JH, Kim MJ. Evaluating imaging-pathology concordance and discordance after ultrasound-guided breast biopsy. Ultrasonography 2018 Apr;37(2):107-120.

(96) Kamei Y, Nishida J, Mimata Y, Shiraishi H, Ehara S, Satoh T, et al. Core Needle Percutaneous Transpedicular Vertebral Body Biopsy: A Study of 128 Cases. J Spinal Disord Tech 2015 Aug;28(7):E394-9.

(97) Kubo T, Furuta T, Johan MP, Sakuda T, Ochi M, Adachi N. A meta-analysis supports core needle biopsy by radiologists for better histological diagnosis in soft tissue and bone sarcomas. Medicine (Baltimore) 2018 Jul;97(29):e11567.

(98) Smidt N, Rutjes AW, van der Windt, D. A., Ostelo RW, Reitsma JB, Bossuyt PM, et al. Quality of reporting of diagnostic accuracy studies. Radiology 2005 May;235(2):347-353.

(99) Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. Fam Pract 2004 Feb;21(1):4-10.

(100) U.S. Department of Health and Human Services Food and Drug Administration Center for Devices and Radiological Health Diagnostic Devices Branch Division of Biostatistics Office of Surveillance and Biometrics. Guidance for Industry and FDA Staff Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests. 2014; Available at: http://www.fda.gov/RegulatoryInformation/Guidances/ucm071148.htm. Accessed 04/16, 2015.

(101) Šimundić AM. Measures of Diagnostic Accuracy: Basic Definitions. EJIFCC 2009 Jan 20;19(4):203-211.

(102) Lange MB, Nielsen ML, Andersen JD, Lilholt HJ, Vyberg M, Petersen LJ. Diagnostic accuracy of imaging methods for the diagnosis of skeletal malignancies: A retrospective analysis against a pathology-proven reference. Eur J Radiol 2016 Jan;85(1):61-67.

(103) Lange MB, Petersen LJ, Lausen M, Bruun NH, Nielsen MB, Zacho HD. Influence of Prior Imaging Information on Diagnostic Accuracy for Focal Skeletal Processes-A Retrospective Analysis of the Consistency between Biopsy-Verified Imaging Diagnoses. Diagnostics (Basel) 2022 Jul 17;12(7):1735. doi: 10.3390/diagnostics12071735.

(104) Lange MB, Petersen LJ, Nielsen MB, Zacho HD. Validity of negative bone biopsy in suspicious bone lesions. Acta Radiol Open 2021 Jul 27;10(7):20584601211030662.

(105) Lydersen S, Fagerland MW, Laake P. Recommended tests for association in 2 x 2 tables. Stat Med 2009 Mar 30;28(7):1159-1175.

(106) Campbell I. Chi-squared and Fisher-Irwin tests of two-by-two tables with small sample recommendations. Stat Med 2007 Aug 30;26(19):3661-3675.

(107) Bruun NH. Stata module to build, present and style Stata matrices," Statistical Software Components S458201. 2021; . Accessed 10, 2021.

(108) Blomqvist L, Carlsson S, Gjertsson P, Heintz E, Hultcrantz M, Mejare I, et al. Limited evidence for the use of imaging to detect prostate cancer: a systematic review. Eur J Radiol 2014 Sep;83(9):1601-1606.

(109) Scheitza P, Uhl M, Hauschild O, Zwingmann J, Bannasch H, Kayser C, et al. Interobserver Variability in the Differential Diagnosis of Benign Bone Tumors and Tumor-like Lesions. Rofo 2016 May;188(5):479-487.

(110) Wu MH, Xiao LF, Liu HW, Yang ZQ, Liang XX, Chen Y, et al. PET/CT-guided versus CT-guided percutaneous core biopsies in the diagnosis of bone tumors and tumor-like lesions: which is the better choice? Cancer Imaging 2019 Oct 29;19(1):69-019-0253-1.

(111) Hegde V, Burke ZDC, Park HY, Zoller SD, Johansen D, Kelley BV, et al. Is Core Needle Biopsy Reliable in Differentiating Between Aggressive Benign and Malignant Radiolucent Bone Tumors? Clin Orthop Relat Res 2018 Mar;476(3):568-577.

(112) Suh CH, Yun SJ. Diagnostic Outcome of Image-Guided Percutaneous Core Needle Biopsy of Sclerotic Bone Lesions: A Meta-Analysis. AJR Am J Roentgenol 2019 Mar;212(3):625-631.

(113) Umer M, Hasan OHA, Khan D, Uddin N, Noordin S. Systematic approach to musculoskeletal benign tumors. Int J Surg Oncol (N Y) 2017 Dec;2(11):e46.

(114) Hakim DN, Pelly T, Kulendran M, Caris JA. Benign tumours of the bone: A review. J Bone Oncol 2015 Mar 2;4(2):37-41.

(115) Sugiyama H, Omonishi K, Yonehara S, Ozasa K, Kajihara H, Tsuya T, et al. Characteristics of Benign and Malignant Bone Tumors Registered in the Hiroshima Tumor Tissue Registry, 1973-2012. JB JS Open Access 2018 May 29;3(2):e0064.

(116) Maccauro G, Spinelli MS, Mauro S, Perisano C, Graci C, Rosa MA. Physiopathology of spine metastasis. Int J Surg Oncol 2011;2011:107969.

(117) Ren G, Esposito M, Kang Y. Bone metastasis and the metastatic niche. J Mol Med (Berl) 2015 Nov;93(11):1203-1212.

(118) Amelot A, Terrier LM, Cristini J, LeNail LR, Buffenoir K, Pascal-Moussellard H, et al. Approaching spinal metastases spread profile. Surg Oncol 2019 Dec;31:61-66.

(119) Heidenreich A, Bellmunt J, Bolla M, Joniau S, Mason M, Matveev V, et al. EAU guidelines on prostate cancer. Part 1: screening, diagnosis, and treatment of clinically localised disease. Eur Urol 2011 Jan;59(1):61-71.

(120) Mohler JL, Kantoff PW, Armstrong AJ, Bahnson RR, Cohen M, D'Amico AV, et al. Prostate cancer, version 2.2014. J Natl Compr Canc Netw 2014 May;12(5):686-718.

(121) Raphael B, Hwang S, Lefkowitz RA, Landa J, Sohn M, Panicek DM. Biopsy of suspicious bone lesions in patients with a single known malignancy: prevalence of a second malignancy. AJR Am J Roentgenol 2013 Dec;201(6):1309-1314.

(122) Wang T, Wu X, Cui Y, Chu C, Ren G, Li W. Role of apparent diffusion coefficients with diffusion-weighted magnetic resonance imaging in differentiating between benign and malignant bone tumors. World J Surg Oncol 2014 Nov 29;12:365-7819-12-365.

European Journal of Radiology 85 (2016) 61–67

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Diagnostic accuracy of imaging methods for the diagnosis of skeletal malignancies: A retrospective analysis against a pathology-proven reference

Mine Benedicte Lange[a,*], Mads Lausen Nielsen[b], Jeannette Demant Andersen[b], Hanna Järnum Lilholt[b], Mogens Vyberg[c,d], Lars Jelstrup Petersen[b,d]

[a] Department of Radiology, Aalborg University Hospital, Aalborg, Denmark
[b] Department of Nuclear Medicine, Clinical Cancer Research Center, Aalborg University Hospital, Aalborg, Denmark
[c] Institute of Pathology, Cancer Research Center, Aalborg University Hospital, Aalborg, Denmark
[d] Department of Clinical Medicine, Aalborg University, Aalborg, Denmark

ARTICLE INFO

Article history:
Received 9 July 2015
Received in revised form 26 August 2015
Accepted 16 October 2015

Keywords:
Bone metastasis/neoplasms/tumours
Diagnostic accuracy
Sensitivity and specificity
Computed tomography Imaging
Positron emission tomography with computed tomography
Magnetic resonance imaging

ABSTRACT

Objectives: To examine the diagnostic accuracy of imaging modalities in skeletal tumours versus pathology reports.
Materials and methods: Pathology reports of bone biopsies were compared to diagnostic imaging with X-ray, computed tomography (CT), magnetic resonance imaging (MRI), bone scintigraphy (BS), and ^{18}F-fluorodeoxyglucose positron emission tomography (FDG-PET/CT) performed within 6 months of biopsy.
Results: A total of 409 biopsies were included. Sensitivity and specificity were significantly different among the five modalities ($p < 0.0001$). The sensitivity of MRI and PET/CT was better than CT, but CT had a better specificity than PET/CT. In general, these methods outperformed BS and X-ray. The sensitivity for osteolytic lesions varied significantly between modalities ($p < 0.0001$), with MRI and PET/CT being more sensitive than CT. Differences in sensitivity were also observed in mixed lesions ($p = 0.0002$) but not in osteosclerotic lesions. In spine lesions, MRI showed the best sensitivity followed by PET/CT and CT ($p < 0.0005$ vs. MRI). There was no significant differences among non-spine lesions.
Conclusions: MRI and FDG-PET/CT showed comparable diagnostic characteristics in general, in individual tumour types, and in different bone lesions and locations. Nominally, they outperformed CT in most situations. The diagnostic accuracy of X-ray and BS were notably inferior to other modalities.

1. Introduction

Diagnostic imaging plays a major role in the evaluation of cancer patients with skeletal involvement. A large numbers of tumours tend to metastasize to the bone, such as prostate, breast, and lung cancers [1]. Approximately 75% of patients with these tumours develop at least one bone metastasis during their disease [2]. Next to lymph nodes, liver and lung, the skeleton is the most frequent metastatic site across all tumours [3]. Correct staging ensures the appropriate choice of curative or palliative treatment. Early and accurate detection of metastatic disease to the bones enables estimation of prognosis and adequate introduction of relevant therapy to minimize morbidities such as skeletal-related events.

Diagnostic imaging of skeletal malignancies can be performed using a variety of methods, such as conventional X-ray, computerized tomography (CT), magnetic resonance imaging (MRI), bone scintigraphy (BS), and positron emission computerized tomography (PET/CT) with various PET ligands, e.g. ^{18}F-fluorodeoxyglucose (FDG). A number of systematic reviews and meta-analyses have investigated diagnostic characteristics of imaging modalities in various tumour types [4–6]. However, there are some outstanding issues of existing data that deserve some discussion. First, a large proportion of the original diagnostic test trials have methodological flaws such as sampling bias, lack of blinding of readers to the results of the reference test and other index tests, and verification bias. It has been shown that a minority of trials fulfil the requirements for a properly designed diagnostic test accuracy study as outlined by the STARD (The Standards for Reporting of Diagnostic

* Corresponding author at: Department of Imaging Diagnostics,University Hospital of North Zealand, Hilleroed, 29 Dyrehavevej, 3400 Hilleroed, Denmark. Fax: number: +45 4829 4520.
E-mail addresses: bl@benedictelange.dk, mine.benedicte.lange@regionh.dk (M.B. Lange).

http://dx.doi.org/10.1016/j.ejrad.2015.10.012
0720-048X/

Accuracy) recommendations [7,8]. Recent studies have indicated no major improvement in reporting quality over the last 10 years [9]. Most importantly, the true reference (presence or absence) of malignancy at the lesion site is seldom adequately confirmed [10], i.e. histology is rarely used as a reference test. In many cases, diagnostic data from the index tests (the imaging test under study) are used for classification of the presence or absence of bone metastasis. Thus, the true diagnostic accuracy of imaging methods is uncertain but vital for appropriate imaging strategies. Finally, due to the functional principle of different imaging modalities as well as different measures of pathology, some imaging modalities may be of particular importance in some conditions, e.g. BS for the diagnose of predominantly osteosclerotic bone metastasis and MRI for soft tissue lesions in the spine.

Biopsy is the gold standard for the demonstration or exclusion of bone tumours. We have unsuccessfully made a comprehensive computer literature search to identify studies comparing the diagnostic characteristics of the various imaging modalities with respect to the diagnosis of the suspected skeletal malignancy using bone biopsy as the gold standard for the majority of patients. The aim of this study was to provide such data by retrospectively examining the diagnostic accuracy of standard skeletal imaging modalities versus pathology reports in a large consecutive population of patients.

2. Materials and methods

2.1. Collection of data

A computer search of pathology samples (hereafter named biopsies) representing bone material registered by SNOMED (Systematized Nomenclature of Medicine) T10* and T11* codes for skeletal cytology and histology biopsies was performed at our institution from January 1, 2011 to July 31, 2013. Each biopsy was identified by unique Danish social security number. The hospital electronic file register for imaging (EasyViz, Karos Health Inc., Waterloo, Ontario, Canada) was then manually reviewed by one reader for any prior imaging procedures for each patient. The eligibility criteria for a biopsy to be included in the analysis were: (1) No more than 6 months between the dates for imaging and biopsy, (2) conclusive pathology results, (3) no biopsies from foetuses and provoked abortions, (4) written description of the imaging findings by a specialist in imaging (radiologist or nuclear medicine physician), and (5) anatomical match between the site of biopsy and the imaging field of view. Diagnostic investigations performed at the time of the biopsy solely for the purpose of biopsy guidance were excluded. If a patient had histological biopsy as well as cytological sample from the same anatomical site, the cytological biopsy was disregarded. If several biopsies were obtained from the same anatomical region within a period of 6 months, and one of these biopsies showed malignancy, the lesion was classified as malignant. If repeated biopsies showed benign conditions, the first biopsy was used. Lesions from separate anatomical regions in a patient could be included provided the eligibility criteria were fulfilled. Each pathology report was reviewed by two readers and classified as benign, malignant or inconclusive. In the case of inconsistency, a board certified pathologist assisted with a final conclusion.

2.2. Review of imaging reports

Descriptions of all relevant imaging procedures were reviewed by two independent reviewers. They classified the description of the lesion as malignant, benign or inconclusive based upon the description and conclusions in the original text. A number of readers participated in pairs of two. After individual reading, they reached consensus for each imaging report without the need of a third party arbitrator. In a minority of the reports, the conclusion of imaging was considered equivocal. This was predominantly the case with X-ray reports of fractures where it could not be read if the fracture was pathological or not.

2.3. Bone biopsy procedure

Most bone biopsies were acquired image-guided (by CT or ultrasound) or sampled during surgical procedures. Biopsies from post-mortem examinations were acquired in a few instances. The biopsies were processed and analysed in accordance with institutional practice. Immunohistochemical examination was applied when relevant. All biopsies were diagnosed by a board-certified pathologist.

2.4. Imaging procedures

All radiology imaging procedures were performed in accordance with institutional guidelines (no experimental imaging investigations were included in the analysis). X-ray imaging were performed by digital radiography and the CT scans were performed on either a GE (GE Lightspeed VCT, 64 slice, GE LightSpeed Pro, 32 slice, GE Discovery 750HD, General Electrics, Milwaukee, WI, USA) or a Siemens (SIEMENS Definition Flash Siemens AG, 128 slice) scanner. MRI scans were performed on a 1.5 T MR scanner (Discovery MR450, General Electrics, Milwaukee, WI, USA). The MRI image sequences were T1, T2 and STIR, out of which at least one sequence was axial on the bone involved; contrast was only given in cases of soft tissue involvement, which was decided in each case by a radiology specialist. All X-rays, CT scans and MRI scans were reviewed by at least one board certified radiology specialist, all images were reviewed, not only key images.

For nuclear medicine procedures, we included BS with approximately 750 mega Becquerel (MBq) 99 mTc-labelled bisphosphonate and whole body scanning acquisitions or regional images depending on the indication on dual-headed gamma camera (Symbia T16 or E.CAM, both Siemens AG, Berlin and Munich, Germany). Supplementary single-photon emission computerized tomography (SPECT/CT) was applied in three patients only. The results from these patients were not analysed due to low sample size. FDG-PET/CT was performed one hour after intravenous injection of 370 MBq of ^{18}F FDG. Whole-body images were acquired using a64-slice Discovery VCT (General Electric Medical Systems, Milwaukee,WI, USA) PET/CTscanner in accordance with institutional procedures. It was institutional practice that all BS were double read by at least one board-certified nuclear medicine physician. PET/CT was always read by a nuclear medicine physician and a radiologist, irrespective of the use of diagnostic or low dose CT.

2.5. Subgroup analysis

To clarify if the diagnostic characteristics of the various imaging modalities depended on the anatomical site under investigation, a subgroup analysis was performed for spine and non-spine anatomical sites, in addition to a subgroup analysis by tumour type (combining all histological subtypes). Finally, diagnostic characteristics of the imaging modalities were analysed with regard to the pattern of bone response. Tumours were classified based upon the literature into three groups depending on whether the bone response was predominantly osteosclerotic, osteolytic or mixed. For example, bone metastasis from prostate cancer was classified as osteosclerotic, bone metastasis from squamous cell adenocarcinoma of the lung as mixed, and bone metastasis from small cell anaplastic carcinoma of the lung was osteolytic. Most tumours were classified in accordance with Roodman [11]. The bone responses

for additional tumour types were classified according to documentation available in the public domain as identified by the corresponding author.

2.6. Statistical analysis

Diagnostic accuracy was calculated as sensitivity, specificity and positive predictive value (PPV) and negative predictive values (NPV) as recommended by the Food and Drug Administration for diagnostic accuracy tests [12]. The cumulative proportion of correct outcome (accuracy) was also calculated. The reference test was pathology in all cases. Statistical analyses were planned for sensitivity and specificity, but not for accuracy or the predictive values. In some subgroup analyses, only sensitivity was calculated and thus used for statistical analysis. Analytical tests included Chi square tests for independent data for 5×2 comparisons and Fisher's exact test for 2×2 comparisons. All p-values were two-sided. P-values < 0.05 were considered significant. Data with equivocal imaging findings and/or pathology conclusions were excluded (3.8% of all investigations). Due to the low number of equivocal findings, no sensitivity analyses (analysis of equivocal findings as benign or malign) were performed. Data were reported with 95% confidence intervals.

2.7. Approvals

This retrospective study did not require ethical approval or informed consent in accordance with national legislation. The Danish Data Protection Agency approved the study and provided permission to access medical files for the purpose of the study.

3. Results

3.1. Study population

The initial pathology search identified 745 biopsies from 605 patients. The final study population consisted of 409 biopsies from 395 patients. The main reasons for exclusion of biopsies were, in descending order, representing groups of 10 samples or more, lack of the description of imaging by an imaging expert (mostly working images acquired during surgery or outpatient visits and reviewed by surgeons only, $n = 101$), biopsy material declared not suitable for diagnostic use ($n = 91$), duplicate of cytology and biopsy material ($n = 40$), and no anatomical match of the site of biopsy and the region of interest for imaging ($n = 26$). Out of the 409 biopsies, 44 were cytological specimens, 379 were regular biopsies from patients, and 6 were post-mortem biopsies.

A description of the final study population is presented in Table 1. Approximately 50% of the biopsies were classified as malignant. Lung cancer, breast cancer and multiple myeloma were the most frequent types of cancer, accounting for 60% of all tumours. Histology showed that non-small-cell lung cancer accounted for 82% of the lung cancers, and 82% of the breast cancers were ductal carcinomas. The most predominant localization of skeletal lesions was the spine. The 215 benign lesions were given 28 different pathology diagnoses, the most frequent conditions ($n = 20$ or greater) being inflammation, fibrosis and osteochondroma.

A total of 758 imaging procedures were performed. Sixty-two percent of the patients had two or more pre-biopsy diagnostic imaging procedures. The majority of pre-biopsy imaging modalities were radiological procedures, equally distributed by X-ray, MR and CT (all >200 each), and a limited number of nuclear medicine procedures (<100 each). The radiological investigations accounted for 88% of all procedures. Contrast-enhancement (gadolinium) was used in 58 of 209 MRI scans. All PET/CT scans were performed with FDG. Fifty-three percent (31 of 58) of the PET/CT were conducted with diagnostic CT, the remainder with low-dose CT.

Table 1
Patient demographics and baseline information.

Patients (n)	395
Females	179 (45%)
Males	216 (55%)
Biopsies (n)	409
Benign	215 (53%)
Malignant	194 (47%)
Malignant tumours by origin	194
Lung	50 (26%)
Breast	39 (20%)
Myeloma	27 (14%)
Lymphoma	18 (9%)
Prostate	13 (7%)
Colorectal	7 (4%)
Kidney	7 (4%)
Other	26 (13%)
Unknown	7 (4%)
Localization of bone lesions	409
Spine	169 (41%)
Extremities	122 (30%)
Pelvis	59 (14 %)
Thorax	31 (8%)
Head	28 (7%)
Imaging modalities	758
X-ray	223 (29%)
CT	233 (31%)
MRI	209 (28%)
BS	35 (5%)
PET/CT	58 (8%)

Abbreviations: CT, computerized tomography; MRI, magnetic resonance tomography; PET/CT, positron emission tomography/computerized tomography; BS, bone scintigraphy.

3.2. Overall diagnostic characteristics

The diagnostic characteristics of the imaging modalities are shown in Table 2. PET/CT demonstrated the highest nominal sensitivity followed by MRI (both better than 90%). The sensitivity of CT and BS was below 80%. The specificity was in the 80–90 % range for CT and MRI versus approximately 60% for BS and PET/CT. X-ray showed excellent specificity but low sensitivity. The PPV was acceptable (80% or greater) for most modalities, and so was NPV for MRI and PET/CT. In contrast, NPV was 66–71% for X-ray and CT, and only 41% for BS. MRI showed the best nominal combination of sensitivity, specificity and PPV, and NPV. The order of accuracy was MRI, PET/CT, CT, BS, and X-ray.

Statistical analysis showed a significant difference in sensitivity across the five modalities ($p < 0.0001$). It has to be noted that the included numbers of BS were relatively low resulting in large confidence intervals. Subsequent pairwise comparisons found no difference between MRI and PET ($p = 1.00$). MRI and PET/CT were significantly more sensitive than CT ($p = 0.0027$ and $p = 0.025$, respectively). Only MRI was significantly better than bone scan ($p = 0.047$). MRI, PET/CT, CT, and BS were all significantly better than X-ray ($p < 0.05$). There was also a significant difference in specificity among the five modalities ($p < 0.0001$). There was no difference between MRI and PET/CT ($p = 0.067$) or MR vs. CT ($p = 0.29$), but CT was more specific than PET/CT ($p = 0.0091$). X-ray was significantly more specific than MRI ($p = 0.001$), PET/CT ($p = 0.0001$) and BS ($p = 0.0065$), but not CT ($p = 0.058$).

3.3. Diagnostics characteristics per tumour type

The diagnostic characteristics of each imaging modality in the most frequent tumour types (those with 10 or more patients) are shown in Table 3. The confidence of estimates was hampered by

Table 2
Diagnostic characteristics of the imaging modalities (reported with 95% confidence intervals).

	X-ray (n = 223)	CT (n = 233)	MRI (n = 209)	BS (n = 35)	PET/CT (n = 58)
Sensitivity	33.0(23.6–43.6)	75.6(67.8–82.6)	90.5 (83.7–95.2)	74.1 (53.7–88.8)	92.3 (79.1–98.3)
Specificity	96.1 (91.2–98.7)	89.2 (81.1–94.7)	81.1 (73.6–89.8)	62.5 (24.7–91.0)	63.2 (38.4–83.7)
Accuracy	69.5 (63.0–75.5)	81.1 (75.5–85.9)	87.1 (81.8–91.5)	71.4 (53.7–85.4)	82.7 (70.6–91.4)
PPV	86.1 (70.5–95.3)	91.4 (84.7–95.8)	86.8 (79.4–92.2)	87.0 (66.4–97.1)	83.7 (69.3–93.2)
NPV	66.3 (59.1–73.0)	70.9 (61.8–79.0)	87.5 (78.7–93.6)	41.7 (15.3–72.3)	80.0 (51.9–95.4)

Abbreviations: negative predictive value (NPV); positive predictive value (PPV).

the limited number of observations with each individual combination of modality and tumour type. Across tumours, PET/CT and MRI showed satisfactory sensitivities. CT showed acceptable sensitivity in breast cancer but sensitivities in the 70% range in other tumours. The sensitivity of X-ray in the specified tumour types was very low. There were generally few applications of BS in most tumour types. There were significant differences in sensitivity across modalities for lung cancer ($p < 0.0001$), breast cancer ($p = 0.0013$), myeloma ($p = 0.022$), lymphoma ($p = 0.0004$), and prostate cancer (0 = 0.023). Due to the low number of observations (and thus high risk if type II error), no post-test pairwise comparisons were made.

3.4. Bone matrix response

The diagnostic characteristics of the imaging modalities distributed by bone response are shown in Fig. 1. The site of origin of the primary tumour could be classified in 187 of 194 tumour biopsies. Among those 187 tumours, the reported bone pattern of skeletal involvement was osteolytic in 105 tumours, osteoblastic in 23 tumours, and mixed lesions in 59 tumours. There were at least 30 observations from each imaging modality in osteolytic lesions except BS ($n = 6$). In mixed lesions, the number of observations was at least 30 with CT and MRI, whereas a limited number of observations were available with X-ray ($n = 22$), BS ($n = 15$), and PET/CT ($n = 7$). The number of images of osteolytic lesions was limited (10–18 with the radiological modalities), in particular with the nuclear medicine methods (BS, $n = 6$; PET/CT, $n = 1$). For clarity, no confidence intervals are reported in Fig. 1.

The sensitivity for the detection of osteolytic lesions showed a highly significant difference ($p < 0.0001$) between modalities. MRI has a significantly higher sensitivity than CT ($p = 0.0079$), and so did PET/CT ($p = 0.033$). A significant difference across modalities was also shown in mixed lesions ($p = 0.0002$), with no significant differences in sensitivity of CT versus MRI ($p = 0.54$). The amount of PET data was too low to allow meaningful comparisons with radiological examinations. There were no significant differences in sensitivity among the modalities in osteosclerotic lesions ($p = 0.26$). Pairwise comparisons of each modality within the three different bone patterns showed a significant difference in sensitivity for X-ray ($p = 0.017$), a trend for BS ($p = 0.098$), but no differences for CT, MRI, or PET/CT (data not shown).

3.5. Lesion localization

The importance of localization of the lesion (spine versus non-spine) is shown in Table 4. MRI seemed to outperform the other modalities with regard to spine lesions. MRI showed sensitivity as well as specificity above 80% and accuracy close to 90%. The sensitivity of PET/CT paralleled that of MRI but the notably lower specificity reduced accuracy to below 80%.

There was a significant difference in sensitivity among spine lesions for the five imaging modalities ($p < 0.0001$) but not for non-spine lesions ($p = 0.15$). Pairwise comparisons of spine lesions showed no difference in sensitivity between MRI and PET/CT ($p = 0.63$). MRI, but not PET/CT, was significantly more sensitive than CT ($p = 0.0005$). There was a significant difference in specificity among the five modalities for spine lesions ($p = 0.028$) as well as for non-spine lesions ($p = 0.0007$). Due to the low number of observations in the majority of correct and incorrect options for specificity among the five modalities, no post-hoc statistical analyses were performed. The nominal rank order for accuracy was MRI, CT, PET/CT, BS and X-ray for spine lesions, and PET/CT, MRI, CT, X-ray, and BS for non-spine lesions. An illustrative example of bone metastasis localized to the spine is shown in Fig. 2.

4. Discussion

Identification of bone metastasis is critical for correct staging, allowing for treatment decisions and estimate outcome in a large numbers of malignant tumours. Many studies have been performed to investigate diagnostic accuracy in skeletal malignancies. In these conditions, in particular in metastases, the reference is very seldom biopsy but varying combinations of imaging and clinical follow-up criteria [10,13]. In this study, we started with bone biopsies and retrospectively identified recent diagnostic X-ray, BS, CT, MRI, and FDG-PET/CT investigations on skeletal malignancies. We showed notable differences in diagnostic characteristics in general, in selected tumours, in bone lesions with expected osteolytic, osteoblastic or mixed bone matrix responses, and in spine versus non-spine lesions.

Our general evaluation and most subgroup analyses showed that MRI and PET/CT were the most accurate methods. FDG-PET/CT has a generally slightly higher sensitivity than MRI, but MRI was more specific. MRI was significantly better than FDG PET/CT for spine lesions. Compared with other imaging techniques, radiography was relatively insensitive in detecting bone metastases. These observations are in line with studies indicating that bone lesions become apparent on radiographs only after the loss of more than 50% of the bone mineral content [2]. We are aware that the low number of nuclear medicine examinations in general, as well as radiological studies in subgroup analyses, may have impaired the statistical power of the statistical analyses. However, the con-

Table 3
Sensitivity of imaging modalities on bone lesions as classified by the most frequent primary tumours (with 95% confidence intervals).

	X-ray	CT	MRI	BS	PET/CT
Lung cancer	24.2 (11.1–42.3; $n = 33$)	75.6 (59.7–87.6, $n = 41$)	86.2 (68.33–96.0, $n = 29$)	40.0 (6.5–84.6, $n = 5$)	100.0 (71.3–100.0, $n = 11$)
Breast cancer	47.1(23.0–72.1, $n = 17$)	86.7 (69.3–96.2, $n = 30$)	95.2 (76.1–99.2, $n = 21$)	84.5 (54.5–97.6, $n = 13$)	100.0 (19.3–100.0, $n = 2$)
Multiple myeloma	46.2 (19.3–74.8, $n = 13$)	73.7 (48.8–90.8, $n = 19$)	93.8 (69.7–99.0, $n = 16$)	100.0 (16.6–100.0, $n = 1$)	100.0 (54.1–100.0, $n = 6$)
Lymphoma	0.0 (0.0–37,1, $n = 8$)	70.0 (34.8–93.00, $n = 10$)	90.9 (61.5–98.6, $n = 11$)	100.0 (16.6–100.0, $n = 1$)	87.5 (51.7–98.2, $n = 8$)
Prostate cancer	66.7 (30.1–92.1, $n = 9$)	75.0 (20.3–95.9, $n = 4$)	90.0 (55.5–98.3, $n = 10$)	100.0 (40.2–100.0, $n = 4$)	No data

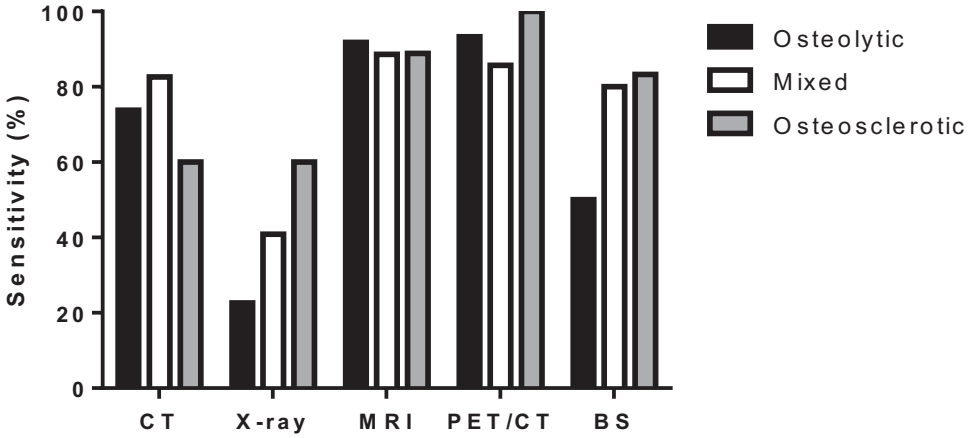


Fig. 1. Sensitivity of imaging modalities for the demonstration of malignant bone lesions for which the tumour is known to induce predominantly osteolytic ($n = 105$), predominantly osteoblastic ($n = 23$), or mixed-type ($n = 59$) bone matrix response. There were significant differences among modalities in sensitivity for the detection of osteolytic ($p < 0.0001$) and mixed lesions ($p = 0.0002$),but not for osteosclerotic lesions.Due to the low number of observations in some groups (see text), the data are shown without confidence intervals.

Table 4
Diagnostic characteristics of imaging modalities in spine and non-spine bone lesions (with 95% confidence intervals).

Location	Diagnostic outcome	Modality				
Spine		**X-ray ($n = 77$)**	**CT ($n = 115$)**	**MRI ($n = 133$)**	**BS ($n = 14$)**	**PET/CT ($n = 22$)**
	Sensitivity	26.7 (14.6–42.0)	70.5 (59.1–80.3)	92.0 (84.1–96.7)	72.7 (39.1–93.7)	87.5 (61.6–98.1)
	Specificity	90.6 (75.0–97.1)	94.6 (81.8–99.2)	80.4 (66.1–90.6)	66.7 (11.6–94.5)	50.0 (12.4–87.6)
	Accuracy	53.2 (41.5–64.7)	78.3 (69.6–85.4)	88.0 (81.2–93.0)	71.4 (41.9–91.6)	77.3 (54.6–92.2)
	PPV	80.0 (51.9–95.4)	96.5 (87.9–99.5)	89.8 (81.7–95.3)	88.9 (51.7–98.2)	82.4 (56.6–96.0)
	NPV	46.8 (34.0–59,9)	60.3 (46.6–73.0)	84.1 (69.9–93.3)	40.0 (6.5–84.6)	60.0 (15.4–93.1)
Non-spine		**X-ray ($n = 146$)**	**CT ($n = 118$)**	**MRI ($n = 76$)**	**BS ($n = 21$)**	**PET/CT ($n = 36$)**
	Sensitivity	38.8 (25.2–53.8)	82.3 (70.5–90.8)	86.2 (68.3–96.0)	75.0 (47.6–92.6)	95.7 (78.0–99.3)
	Specificity	97.9 (92.7–99.7)	85.7 (73.8–93.6)	85.1 (71.7–93.8)	60.0 (15.4–93.5)	69.2 (38.6–90.7)
	Accuracy	78.1 (70.5–84.5)	83.9 (76.0–90.0)	85.5 (75.6–92.6)	71.4 (47.8–88.7)	86.1 (70.5–95.3)
	PPV	90.5 (69.6–98.6)	86.4 (75.0–93.9)	78.1 (60.0–90.7)	85.7 (57.2–97.8)	84.6 (65.1–95.6)
	NPV	76.0 (67.5–83.2)	81.4 (69.1–90.3)	90.9 (78.3–97.4)	42.9 (10.4–81.3)	90.0 (55.5–98.3)

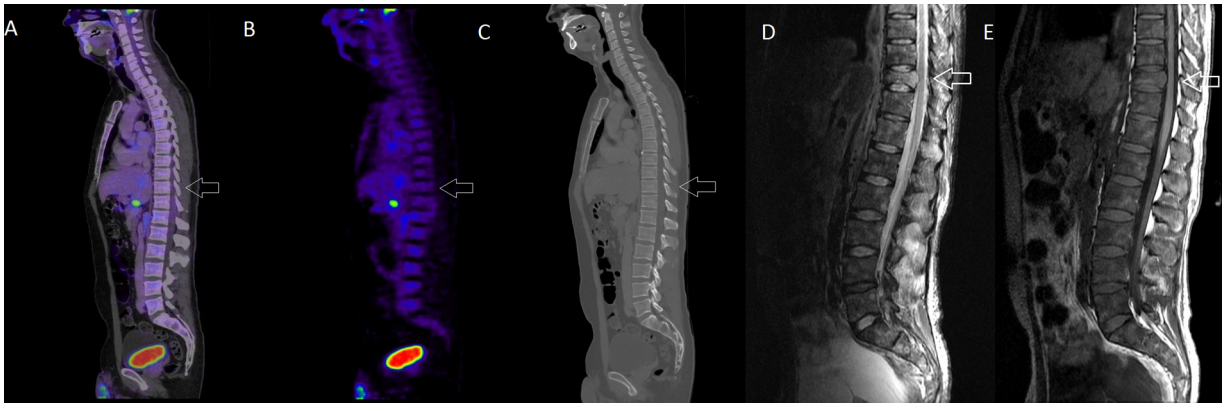


Fig. 2. A 46 year old man was diagnosed with cardia cancer (subcardial grade III). An initial FDG-PET/CT scan showed no bone metastases in the spine (A: PET/CT sagittal fused image; B PET sagittal view; C: CT sagittal view). 3 months later, persisting back paingave rise to an MRI scan showing disseminated cancer with a biopsy-proven bone metastasis in Th10 (arrow) with posterior soft tissue bulging and compression of the medulla (D: Sagittal T2 Stir TR/TE 3500/38,496 msM E: sagittal T1 TR/TE 350/12,832 ms).

clusions of the statistical analyses were largely in line with the quantitative/numerical estimates of diagnostic characteristics. It is difficult to compare our data with the results of prior studies because there are no systematic studies with pathology as the only reference. However, our findings are in line with data from recent systematic reviews and meta-analyses. Yang et al. compared FDG PET, CT, MRI and BS and showed comparable data with MRI and PET/CT with these methods being significantly more accurate than CT and BS for the diagnosis of bone metastases [14]. Duo et al. concluded in a meta-analysis from 2013 comparing FDG-PET/CT and gadolinium-enhanced MRI for detection of bone metastases that FDG PET/CT and gadolinium-enhanced MRI were almost equal in terms of diagnostic accuracy and that both imaging methods have excellent diagnostic performance for the detection of bone metastases in patients with cancer [4].

MRI is increasingly used as a reference test for skeletal malignancies [15–17]. Our data show an overall sensitivity of 91% and specificity of 81% versus pathology. Thus, the validity for MRI as

a reference in imaging trials warrants some concern. For example, bone marrow oedema is also hyperintense in fat-suppressed sequences, which can lead to diagnostic errors [2]. Observer agreement for malignancy with MRI is not perfect, with kappa values below 0.7 [18].

The role of the primary tumour origin (or, perhaps more appropriate for bone metastasis, the bone matrix response), should be considered when comparing diagnostic characteristics of imaging methods. A method such as BS solely reflects bone metabolism at sites with active bone mineralization, such as osteosclerotic components, making it of particular importance in prostate cancer and of limited use in predominant osteolytic lesions such as myeloma [19]. Our data supported this observation with BS and X-ray, and provided some indications of a similar trend for CT. Both MRI and PET/CT performed equally well across bone matrix responses. This observation is likely because MRI and FDG-PET/CT directly reflect the extent of the tumour, whereas BS and, to some extent, CT mirror the reaction of the tumour-to-bone interface. Our findings of the superiority of PET/CT over BS are in line with data from a recent meta-analysis [20]. Tumour-specific diagnostic criteria should be considered in terms of grade of malignancy. It is well-established that the avidity of FDG-uptake is related to tumour characteristics. Similarly, choline PET/CT, a biomarker of cell membrane synthesis, may prove useful in advanced, recurrent prostate cancer but may be of limited value in staging.

The conclusions of a meta-analysis comparing whole-body MRI versus BS for detection of bone metastatic tumours in mixed cancers showing similar patient-based sensitivity and specificity for MRI and BS [6] should be considered in light of the composition of the tumour types as well as disease condition (staging or relapse/recurrent disease).

It is generally believed that CT is superior to MRI with regard to characterization of cortical bone, whereas MRI is superior to CT with regard to characterization of intra-osseous and soft tissue details [21]. Our data showed 12% better sensitivity with CT in non-spine versus spine lesions. It was also revealed that MRI showed excellent diagnostic properties in spine lesions with an accuracy close to 90%, while no other methods reached 80% (with X-ray below 40%). In non-spine lesions, the methods were largely comparable. The extent to which the difference in location is related to anatomical mapping (tomographic methods versus planar imaging) or tumour characteristics (soft tissue component) cannot be clearly defined. BS showed similar diagnostic data in spine and non-spine lesions, whereas X-ray was definitely poorer in spine lesions. With BS, it has been shown that the addition of SPECT/CT to whole-body BS improved specificity without affecting sensitivity [22]. Bone metastases most commonly affect the axial skeleton [1]. Thus, appropriate methods must be selected for the individual patient. The use of X-ray should be performed with caution for detection of malignancy in the spine.

This study was retrospective in design and comprised data from daily clinical life. It may be argued that the imaging reports were not blindly reread without clinical information for the purpose of the study. However, the conclusions of the imaging reports were those presented to the clinicians. Nuclear medicine reports were all double-read, whereas radiology reports were generally made by only one radiologist alone or by supervising a radiologist in training. It remains speculative whether reading biases due to access to clinical information or due to prior imaging were present. Data from the reports were extracted by two independent readers with minimal disagreement in reporting.

Most biopsy material was histological biopsies, with only 6% of samples being cytology specimens. The reading and reporting of the pathology samples were considered appropriate with minimal bias in classifying malignancy. We are aware of a potential false negative biopsy, and prolonged follow up could be considered to improve specificity in this project. Similarly, it could be considered whether a time window of 6 months between imaging and biopsy is adequate. In our opinion, this period is an adequate choice for balancing recruitment and any false negative imaging.

The method of collecting biopsy and imaging data warrants some discussion. The findings are based upon lesions which ended up undergoing biopsy. Thus, the biopsies may not be representative sample of patients presenting with suspicious skeletal lesions. Many lesions may not be biopsied for several reasons, such as perceived definitive imaging results, multiplicity, and a lack of consequences. More than 60% of the patients in this study had more than one imaging modality, indicating some uncertainty in the classification of the bone lesion. The relatively low proportion of prostate cancer patients, the most frequent cancer among males, is apparent. However, it is known from clinical practice that these patients seldom undergo bone biopsy due to limited therapeutic impact. Our data indicated superiority of MRI over BS in prostate cancer [23], and thus question the uniform recommendation of BS for staging of skeletal metastasis in prostate cancer [24,25].

Because all imaging methods were not applied in all patients, the data may reflect diagnostic characteristics in separate populations. A patient undergoing one X-ray may be different from a patient with three or more imaging modalities. Because most patients had more than one imaging modality, demonstration of similarities or differences in patient groups may prove difficult, also taking into consideration the limited samples sizes in some subgroup analyses. The relatedness of imaging modalities and clinical variables may be defined as a separate study.

Finally, methodological issues can be raised. It can be debated whether alternative technical settings for the imaging methods studied could influence the findings. For example, only one in every four patients with an MRI in our study had a contrast-enhanced MRI. Some authors have advocated for wide routine use of gadolinium in the MRI evaluation of skeletal malignancies [26], but a recent meta-analysis showed no substantial differences in diagnostic value of the contrast enhancement [14]. No diffusion-weighed MRI was used. Recent studies have shown that the sensitivity for using DWI sequences is significantly higher than that for not using DWI sequences in detection of bone metastases, with lower specificity on a per-patient basis [14,27].

In conclusion, we observed notable differences in diagnostic characteristics among frequently used imaging modalities for detection of skeletal malignancies. MRI and FDG-PET/CT performed well in most patient subgroups, followed by CT with some distance to X-ray and BS with reservations to the fact that the number of BS was low ($n = 38$, 5% of all examinations). Our findings indicate that the results of imaging investigations should be interpreted with caution. Tumour characteristics/bone matrix response and localization of skeletal lesions may influence the performance of tests and guidelines for a diagnostic strategy for skeletal metastases based upon type, location and primary tumour that could be developed based upon our findings.

Conflict of interest

None for any of the authors.

References

[1] R.E. Coleman, Clinical features of metastatic bone disease and risk of skeletal morbidity, Clin. Cancer Res. 12 (October (20 Pt.2)) (2006) 6243–6249.
[2] E. Balliu, M. Boada, I. Pelaez, J.C. Vilanova, C. Barcelo-Vidal, A. Rubio, et al., Comparative study of whole-body MRI and bone scintigraphy for the detection of bone metastases, Clin. Radiol. 65 (December (12)) (2010) 989–996.
[3] G. Disibio, S.W. French, Metastatic patterns of cancers: results from a large autopsy study, Arch. Pathol. Lab. Med. 132 (June (6)) (2008) 931–939.

[4] J. Duo, X. Han, L. Zhang, G. Wang, Y. Ma, Y. Yang, Comparison of FDG PET/CT and gadolinium-enhanced MRI for the detection of bone metastases in patients with cancer: a meta-analysis, Clin. Nucl. Med. 38 (May (5)) (2013) 343–348.
[5] X. Qu, X. Huang, W. Yan, L. Wu, K. Dai, A meta-analysis of (1)(8)FDG-PET-CT, (1)(8)FDG-PET, MRI and bone scintigraphy for diagnosis of bone metastases in patients with lung cancer, Eur. J. Radiol. 81 (May (5)) (2012) 1007–1015.
[6] Q. Wu, R. Yang, F. Zhou, Y. Hu, Comparison of whole-body MRI and skeletal scintigraphy for detection of bone metastatic tumors: a meta-analysis, Surg. Oncol. 22 (December (4)) (2013) 261–266.
[7] N. Smidt, A.W. Rutjes, D.A. van der Windt, R.W. Ostelo, J.B. Reitsma, P.M. Bossuyt, et al., Quality of reporting of diagnostic accuracy studies, Radiology 235 (May (2)) (2005) 347–353.
[8] P.M. Bossuyt, J.B. Reitsma, D.E. Bruns, C.A. Gatsonis, P.P. Glasziou, L.M. Irwig, et al., Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative, Fam. Pract. 21 (Feburary (1)) (2004) 4–10.
[9] D.A. Korevaar, W.A. van Enst, R. Spijker, P.M. Bossuyt, L. Hooft, Reporting quality of diagnostic accuracy studies: a systematic review and meta-analysis of investigations on adherence to STARD, Evid. Based. Med. 19 (April (2)) (2014) 47–54.
[10] M. Wondergem, F.M. van der Zant, T. van der Ploeg, R.J. Knol, A literature review of 18F-fluoride PET/CT and 18F-choline or 11C-choline PET/CT for detection of bone metastases in patients with prostate cancer, Nucl. Med. Commun. 34 (October (10)) (2013) 935–945.
[11] D. Roodman, Mechanisms of bone metastases. In UpToDate. 2015 http://www.uptodate.com/contents/mechanisms-of-bone-metastases?source=machineLearning&search=bone%C2%B1metastasis.%C2%B1osteolytic&selectedTitle=1~150&sectionRank=1&anchor=H2#H2. (accessed 01.30.15).
[12] U.S. Department of Health and Human Services Food and Drug Administration Center for Devices and Radiological Health Diagnostic Devices Branch Division of Biostatistics Office of Surveillance and Biometrics. Guidance for Industry and FDA Staff Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests. 2014. Available at: http://www.fda.gov/RegulatoryInformation/Guidances/ucm071148.htm. (accessed 04.16.15).
[13] L. Blomqvist, S. Carlsson, P. Gjertsson, E. Heintz, M. Hultcrantz, I. Mejare, et al., Limited evidence for the use of imaging to detect prostate cancer: a systematic review, Eur. J. Radiol. 83 (September (9)) (2014) 1601–1606.
[14] H.L. Yang, T. Liu, X.M. Wang, Y. Xu, S.M. Deng, Diagnosis of bone metastases: a meta-analysis comparing (1)(8)FDG PET, CT, MRI and bone scintigraphy, Eur. Radiol. 21 (December (12)) (2011) 2604–2617.
[15] S. Kruger, A.K. Buck, F.M. Mottaghy, E. Hasenkamp, S. Pauls, C. Schumann, et al., Detection of bone metastases in patients with lung cancer: 99mTc-MDP planar bone scintigraphy, 18F-fluoride PET or 18F-FDG PET/CT, Eur. J. Nucl. Med. Mol. Imaging 36 (November (11)) (2009) 1807–1812.
[16] T. Nozaki, K. Yasuda, T. Akashi, H. Fuse, Usefulness of single photon emission computed tomography imaging in the detection of lumbar vertebral metastases from prostate cancer, Int. J. Urol. 15 (June (6)) (2008) 516–519.
[17] M.H. Poulsen, H. Petersen, P.F. Hoilund-Carlsen, J.S. Jakobsen, O. Gerke, J. Karstoft, et al., Spine metastases in prostate cancer: comparison of technetium-99m-MDP whole-body bone scintigraphy, [(18) F]choline positron emission tomography(PET)/computed tomography (CT) and [(18) F]NaF PET/CT, BJU Int. 114 (December (6)) (2014) 818–823.
[18] A. Stecco, M. Lombardi, L. Leva, M. Brambilla, E. Negru, S. Delli Passeri, et al., Diagnostic accuracy and agreement between whole-body diffusion MRI and bone scintigraphy in detecting bone metastases, Radiol. Med. 118 (April (3)) (2013) 465–475.
[19] Collins CD. Multiple myeloma. Cancer Imaging 2010 (Feb 11), 10 20–31.
[20] D. van Lammeren-Venema, J.C. Regelink, I.I. Riphagen, S. Zweegman, O.S. Hoekstra, J.M. Zijlstra, (1)(8)F-fluoro-deoxyglucose positron emission tomography in assessment of myeloma-related bone disease: a systematic review, Cancer 118 (April (8)) (2012) 1971–1981.
[21] H. Pettersson, T. Gillespy 3rd, D.J. Hamlin, W.F. Enneking, D.S. Springfield, E.R. Andrew, et al., Primary musculoskeletal tumors: examination with MR imaging compared with conventional modalities, Radiology 164 (July (1)) (1987) 237–241.
[22] H. Palmedo, C. Marx, A. Ebert, B. Kreft, Y. Ko, A. Turler, et al., Whole-body SPECT/CT for bone scintigraphy: diagnostic value and effect on patient management in oncological patients, Eur. J. Nucl. Med. Mol. Imaging 41 (January (1)) (2014) 59–67.
[23] F.E. Lecouvet, D. Geukens, A. Stainier, F. Jamar, J. Jamart, B.J. d'Othee, et al., Magnetic resonance imaging of the axial skeleton for detecting bone metastases in patients with high-risk prostate cancer: diagnostic and cost-effectiveness and comparison with current detection strategies, J. Clin. Oncol. 25 (August (22)) (2007) 3281–3287.
[24] A. Heidenreich, J. Bellmunt, M. Bolla, S. Joniau, M. Mason, V. Matveev, et al., EAU guidelines on prostate cancer. Part 1: screening, diagnosis, and treatment of clinically localised disease, Eur. Urol. 59 (January (1)) (2011) 61–71.
[25] J.L. Mohler, P.W. Kantoff, A.J. Armstrong, R.R. Bahnson, M. Cohen, A.V. D'Amico, et al., Prostate cancer, version 2. 2014, J. Natl. Compr. Canc. Netw. 12 (May (5)) (2014) 686–718.
[26] D.A. May, R.B. Good, D.K. Smith, T.W. Parsons, MR imaging of musculoskeletal tumors and tumor mimickers with intravenous gadolinium: experience with 242 patients, Skeletal Radiol. 26 (January (1)) (1997) 2–15.
[27] T. Wang, X. Wu, Y. Cui, C. Chu, G. Ren, W. Li, Role of apparent diffusion coefficients with diffusion-weighted magnetic resonance imaging in differentiating between benign and malignant bone tumors, World J. Surg. Oncol. 29 (November (12)) (2014) 365–7819.

diagnostics

MDPI

Article

Influence of Prior Imaging Information on Diagnostic Accuracy for Focal Skeletal Processes—A Retrospective Analysis of the Consistency between Biopsy-Verified Imaging Diagnoses

Mine Benedicte Lange [1,2,3,*], Lars J. Petersen [3,4], Mads Lausen [5], Niels Henrik Bruun [6], Michael Bachmann Nielsen [7] and Helle D. Zacho [3,4]

1 Department of Radiology, North Zealand Hospital, 3400 Hilleroed, Denmark
2 Department of Diagnostic Imaging, North Zealand Hospital, 3400 Hilleroed, Denmark
3 Department of Clinical Medicine, Aalborg University, 9000 Aalborg, Denmark; ljp@dadlnet.dk (L.J.P.); h.zacho@rn.dk (H.D.Z.)
4 Department of Nuclear Medicine, Clinical Cancer Research Center, Aalborg University Hospital, 9000 Aalborg, Denmark
5 Department of Clinical Microbiology, Copenhagen University Hospital, Rigshospitalet, 2100 Copenhagen, Denmark; madsln@biosustain.dtu.dk
6 Department of Research and Statistics, Aalborg University Hospital, 9000 Aalborg, Denmark; nbru@rn.dk
7 Department of Radiology, Copenhagen University Hospital, Rigshospitalet, 2100 Copenhagen, Denmark; michael.bachmann.nielsen@regionh.dk
* Correspondence: bl@benedictelange.dk; Tel.: +45-2041-1756

Abstract: Introduction: Comparing imaging examinations with those previously obtained is considered mandatory in imaging guidelines. To our knowledge, no studies are available on neither the influence, nor the sequence, of prior imaging and reports on diagnostic accuracy using biopsy as the reference standard. Such data are important to minimize diagnostic errors and to improve the preparation of diagnostic imaging guidelines. The aim of our study was to provide such data. Materials and methods: A retrospective cohort of 216 consecutive skeletal biopsies from patients with at least 2 different imaging modalities (X-ray, CT and MRI) performed within 6 months of biopsy was identified. The diagnostic accuracy of the individual imaging modality was assessed. Finally, the possible influence of the sequence of imaging modalities was investigated. Results: No significant difference in the accuracy of the imaging modalities was shown, being preceded by another imaging modality or not. However, the sequence analyses indicate sequential biases, particularly if MRI was the first imaging modality. Conclusion: The sequence of the imaging modalities seems to influence the diagnostic accuracy against a pathology reference standard. Further studies are needed to establish evidence-based guidelines for the strategy of using previous imaging and reports to improve diagnostic accuracy.

Keywords: bone; cancer; metastasis; tumor; biopsy; diagnostic accuracy; medical imaging; reports; prior imaging

Citation: Lange, M.B.; Petersen, L.J.; Lausen, M.; Bruun, N.H.; Nielsen, M.B.; Zacho, H.D. Influence of Prior Imaging Information on Diagnostic Accuracy for Focal Skeletal Processes—A Retrospective Analysis of the Consistency between Biopsy-Verified Imaging Diagnoses. *Diagnostics* **2022**, *12*, 1735. https://doi.org/10.3390/diagnostics12071735

Academic Editors: Rute Santos and Antonio Barile

Received: 9 May 2022
Accepted: 15 July 2022
Published: 17 July 2022

1. Introduction

Radiology is one of the specialties most liable to claims of diagnostic negligence, which can be defined as errors resulting in incorrect, delayed, or missed diagnoses [1–3]. Several studies have investigated the incidence and causes of medical errors, but such analyses remain challenging due to the lack of effective methods for measurement and limited sources of reliable data [4].

A diagnostic report consists of the complete detection and accurate diagnosis of all abnormalities in an imaging examination and at the same time as accurately as possible to distinguish which lesions can be safely ignored from those requiring additional workup or biopsy, most often described as either benign or possible malignant. The average error rate

among radiologists has been shown to be approximately 30%, referring to images as part of a set of unknowns with proven pathology, a prevalence that has remained unchanged since it was first estimated in the 1960s [5–7]. The etiology of radiological error is multifactorial, including failure to compare with prior imaging and reports, bias, poor technique, failures of perception, lack of knowledge, fatigue, noise, and misjudgments [8]. More than 70% of errors are perceptual, whereas fewer than 30% are cognitive [5]. One study showed that radiologists disagreed with each other more than 30% of the time and with themselves more than 25% of the time [9]. It is considered without debate to be the standard of care by the radiology and the non-radiology medical communities that radiologists must compare new imaging examinations with those obtained previously [10–16]. Failure to consult prior radiologic studies has been shown to represent 5% of the explanation for missed findings [5,7,10,17]. Previous images are subjectively judged to be more valuable than imaging reports for documenting disease progression on conventional X-ray images [18,19]. Studies have shown that if one looks at a prior negative report before looking at imaging studies, there is a greater chance of missing a significant abnormality than by looking at the imaging studies first [5]. It has also been shown that radiological diagnoses made with adequate clinical information are more accurate than those made without clinical information [20–23]. However, to the best of our knowledge, no studies have investigated the influence, or the sequence, of prior imaging and reports on diagnostic accuracy using biopsy as the reference standard. Such data are of great importance not only to minimize diagnostic errors but also to improve the preparation of diagnostic imaging guidelines based upon diagnostic accuracy and cost-effectiveness.

The purpose of our study was to investigate whether the diagnostic accuracy of the detection of skeletal malignancies, proven malign or benign by subsequent biopsy, is affected by prior imaging examinations and their mutual sequences.

2. Materials and Methods

2.1. Collection of Skeletal Biopsies

The study was conducted as a retrospective consecutive cohort study. Bone biopsies were identified by performing a computer search of pathology samples representing bone material registered by SNOMED (Systematized Nomenclature of Medicine) T10* and T11* codes for skeletal cytology and histology biopsies from 1 January 2011 to 31 July 2013, at the Department of Pathology, and each biopsy was identified by a unique social security number [24]. The eligibility criteria for a biopsy to be included in the analysis were conclusive pathology results performed by a board-certified pathologist. The biopsies were processed and analyzed in accordance with institutional practice, and immunohistochemical examination was applied when relevant. If several biopsies were obtained from the same anatomical region within a period of 6 months and one of these biopsies showed malignancy, the lesion was classified as malignant. If repeated biopsies showed a benign condition, the first biopsy was used.

Each pathology report was reviewed by two readers and classified as benign, malignant, or inconclusive. In the case of inconsistency, a board-certified pathologist assisted with a conclusion.

The baseline dataset was used for two previously published articles, and the exclusion criterion for the present study was biopsies performed with less than two different imaging modalities six months prior to the biopsy (Figure 1) [24,25].

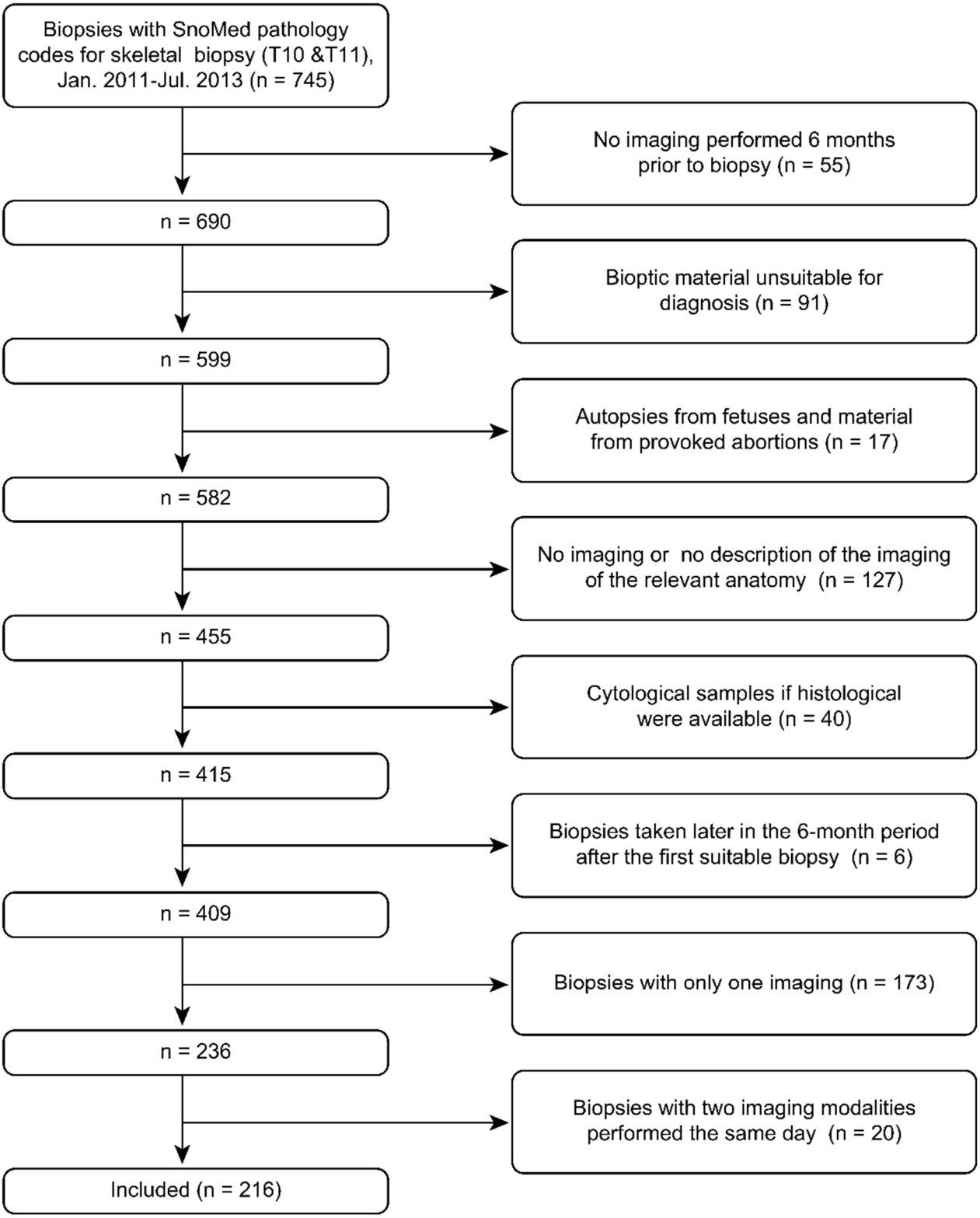


Figure 1. Flow chart of study material.

2.2. Imaging

Diagnostic imaging included X-ray, computed tomography (CT) and magnetic resonance imaging (MRI). X-ray imaging was performed by digital radiography, and the CT scans were performed on either a GE (GE Lightspeed VCT, 64 slice, GE LightSpeed Pro, 32 slice, GE Discovery 750HD, General Electrics, Milwaukee, WI, USA) or a Siemens

(SIEMENS Definition Flash Siemens AG, 128 slice) scanner. MRI scans were performed on a 1.5 T MR scanner (Discovery MR450, General Electrics, Milwaukee, WI, USA). The MRI image sequences were T1, T2 and STIR, of which at least one sequence was axial on the bone involved; contrast was only given in cases of soft tissue involvement, which was decided in each case by a radiology specialist. Bone scintigraphy (BS), single photon emission computed tomography CT (SPECT/CT), 18F-fluorodeoxyglucose positron emission tomography (FDG-PET/CT) and ultrasound (US) were excluded due the low number of combinations of those with another.

All radiology imaging procedures were performed in accordance with institutional guidelines (no experimental imaging investigations were included in the analysis), and the written reports were reviewed by two independent reviewers who, based upon the description and conclusions in the original text, classified the described lesion as malignant, benign, or inconclusive. In cases of disagreement after individual reading, the readers reached consensus for each imaging report without the need for a third-party arbitrator. The radiologists had access to an Electronical Patient Journal charts (EPJ—Clinical Suite, CSC Scandihealth A/S) for any relevant journal notes in case they needed more information than was stated in the referral.

2.3. Statistics

Statistical analysis was performed by using Stata 17 (StataCorp LLC 2021) and the Stata package matrix tools [26]. Sensitivity, specificity, prevalence, accuracy, positive predictive value (PPV) and negative predictive value (NPV) with 95% confidence intervals were calculated for each imaging modality without taking the imaging sequence into consideration. Then, it was calculated for pairs of imaging modalities, such as X-ray/CT and CT/X-ray, and by doing so, not all X-ray stand-alone values were included to minimize the bias that only one imaging was performed as opposed to two. The diagnostic properties of one modality (CT, MRI, and X-ray) when used as the first imaging modality were compared with the diagnostic properties of the modality when it was preceded by another modality using Fisher's exact test. It should be noted that the numbers in some of the subgroups may be too low to detect significant differences. Finally, the effect of the imaging sequence was examined among patients with a malignant biopsy diagnosis and with a benign biopsy diagnosis; due to the small number in each group, only descriptive statistics were used.

2.4. Approval

This retrospective study did not require ethical approval or informed consent in accordance with national legislation. The Danish Data Protection Agency approved the study and provided permission to access medical files for the purpose of the study.

3. Results

3.1. Baseline Data

Most of the biopsies were malignant (Table 1), with lung cancer (31%), breast cancer (19%), multiple myeloma (12%) and lymphoma (11%) being the most frequent types of cancer. The benign lesions were mainly characterized as inflammation, fibrosis, osteochondroma, degenerative changes, nonspecific reactive changes, necrosis, and fracture. There was a slight predominance of males over females, and the spine was the most common anatomical localization of bone biopsy. The three included imaging modalities were almost equally represented (Table 1). Most biopsies (67%) had two imaging modalities performed 6 months prior to biopsy, 30% had three imaging modalities performed and 3% had four imaging modalities performed (details are provided in Table S1).

Table 1. Baseline demographics.

Variable	Value
Patients (n = 207)	
Male, n (%)	116 (56%)
Female, n (%)	91 (44%)
Age, median (range)	67 (1–93)
Biopsies (n = 216)	
Malignant, n (%)	132 (61%)
Benign, n (%)	84 (39%)
Biopsy specimen (n = 216)	
Cytological, n (%)	16 (8%)
Histological, n (%)	195 (90%)
Dissection, n (%)	5 (2%)
Imaging modalities performed (n = 464)	
X-ray, n (% of biopsies)	143 (66%)
CT, n (% of biopsies)	169 (78%)
MRI, n (% of biopsies)	152 (70%)
Localization of bone lesion (n = 216)	
Spine, n (%)	119 (55%)
Extremities, n (%)	39 (18%)
Pelvis, n (%)	36 (17%)
Thorax and shoulders, n (%)	19 (9%)
Head, n (%)	3 (1%)

MRI was shown to have the highest accuracy, followed by CT and X-ray when the sequence of imaging was not taken into consideration (Table 2). MRI also showed the highest sensitivity and NPV, whereas X-ray proved to have the highest specificity and CT the highest PPV (Table 2).

Table 2. Sensitivity, specificity, accuracy, positive and negative predictive values (PPV, NPV) estimates of imaging techniques for detection of focal skeletal lesions.

	X-ray (n = 143)	CT (n = 169)	MRI (n = 152)
Sensitivity	31.3 (21.4–42.6)	73.5 (64.3–81.3)	92.1 (84.5–96.8)
Specificity	95.2 (86.7–99.0)	85.7 (73.8–93.6)	81.0 (69.1–89.8)
Accuracy	59.4 (50.9–67.6)	77.5 (70.5–83.6)	87.5 (81.2–92.3)
PPV	89.3 (71.8–97.7)	91.2 (83.4–96.1)	87.2 (78.8–93.2)
NPV	52.2 (42.7–61.6)	61.5 (49.8–72.3)	87.9 (76.7–95.01)

Note—95% exact confidence intervals for each imaging modality without taking the imaging sequence into consideration.

3.2. Sequence Analysis

Taking the sequence of imaging modalities into account, no significant difference in accuracy within each imaging modality was seen when preceded by another imaging modality or not (Tables 3–5), except for a decrease in CT specificity and PPV when preceded by MRI (Table 3. Despite the lack of difference in overall accuracy, an interesting pattern of observations was seen when examining the sequences for imaging divided by malignant and benign biopsies.

Among malignant (positive) biopsies, it was seen that if X-ray was false negative (75%) and used as the first imaging modality, only 7% of the subsequent MRI and 30% of the subsequent CT imaging were false negative (Figure 2A), whereas if MRI was false negative (17%) and conducted as the first imaging modality, 100% of the following CT scans were false negative as well (Figure 2E). Likewise, among biopsies with a benign (negative) histology, if MRI was false-positive (33%), 100% of the subsequent CT imaging was also false-positive (Figure 2F). Figure 2C demonstrates that when CT scans were false negative

(30%), 100% of the subsequent X-ray examinations were false negative, whereas this was only the case for 14% of the subsequent MRI. For the few false-positive X-ray and CT ex.

Table 3. Sensitivity, specificity, accuracy, PPV and NPV estimates (reported with 95% confidence intervals) for X-ray and CT without or with access to a preceding MRI.

	X-ray			CT		
	Not Preceded by MRI (n = 122)	**Preceded by MRI (n = 21)**	***p* Value**	**Not Preceded by MRI (n = 140)**	**Preceded by MRI (n = 29)**	***p* Value**
Sensitivity	28.2 (18.1–40.1)	55.6 (21.2–86.3)	0.13	72.0 (61.8–80.9)	80.0 (56.3–94.3)	0.58
Specificity	98.0 (89.6–100.0)	83.3 (51.6–97.9)	0.09	93.6 (82.5–98.7)	44.4 (13.7–78.8)	0.00
Accuracy	70.0 (63.1–76.3)	65.2 (42.7–83.6)	0.34	79.3 (71.6–85.7)	69.0 (49.2–84.7)	0.23
PPV	95.2 (76.2–99.9)	71.4 (29.0–96.3)	0.15	95.7 (88.0–99.1)	76.2 (52.8–91.8)	0.01
NPV	49.5 (39.4–59.6)	71.4 (41.9–91.6)	0.16	62.9 (50.5–74.1)	50.0 (15.7–84.3)	0.48

Table 4. Sensitivity, specificity, accuracy, PPV and NPV estimates (reported with 95% confidence intervals) for X-ray and MRI without or with access to a preceding CT.

	X-ray			MRI		
	Not Preceded by CT (n = 111)	**Preceded by CT (n = 32)**	***p* Value**	**Not Preceded by CT (n = 70)**	**Preceded by CT (n = 82)**	***p* Value**
Sensitivity	28.6 (17.9–41.3)	41.2 (18.4–67.1)	0.38	91.7 (77.5–98.2)	92.5 (81.8–97.9)	1.00
Specificity	93.8 (82.8–98.7)	100.0 (78.2–100.0)	1.00	79.4 (62.1–91.3)	82.8 (64.2–94.2)	1.00
Accuracy	56.8 (47.0–66.1)	68.8 (50.0–83.9)	0.84	85.7 (75.3–92.9)	89.0 (80.2–94.9)	0.63
PPV	85.7 (63.7–97.0)	100.0 (59.0–100.0)	0.55	82.5 (67.2–92.7)	90.7 (79.7–96.9)	0.35
NPV	50.0 (39.9–60.7)	60.0 (38.7–78.9)	0.50	90.0 (73.5–97.9)	85.7 (67.3–96.0)	0.70

Table 5. Sensitivity, specificity, accuracy, PPV and NPV estimates (reported with 95% confidence intervals) for CT and MRI without or with access to a preceding X-ray.

	CT			MRI		
	Not Preceded by X-ray (n = 98)	**Preceded by X-ray (n = 71)**	***p* Value**	**Not Preceded by X-ray (n = 83)**	**Preceded by X-ray (n = 69)**	***p* Value**
Sensitivity	70.3 (57.6–88.1)	77.6 (63.4–88.2)	0.52	90.2 (78.6–96.7)	94.7 (82.3–99.4)	0.69
Specificity	91.2 (76.3–98.1)	77.3 (54.6–92.2)	0.24	71.9 (53.3–86.3)	90.3 (74.2–98.0)	0.11
Accuracy	77.6 (68.0–65.4)	77.5 (66.0–86.5)	1.00	83.1 (73.3–90.5)	92.8 (83.9–97.6)	0.09
PPV	93.8 (82.8–98.7)	88.4 (74.9–96.1)	0.47	83.6 (71.2–92.2)	92.3 (79.1–98.4)	0.35
NPV	62.0 (47.2–75.3)	60.7 (40.6–78.5)	1.00	82.1 (63.1–93.9)	93.3 (77.9–99.2)	0.25

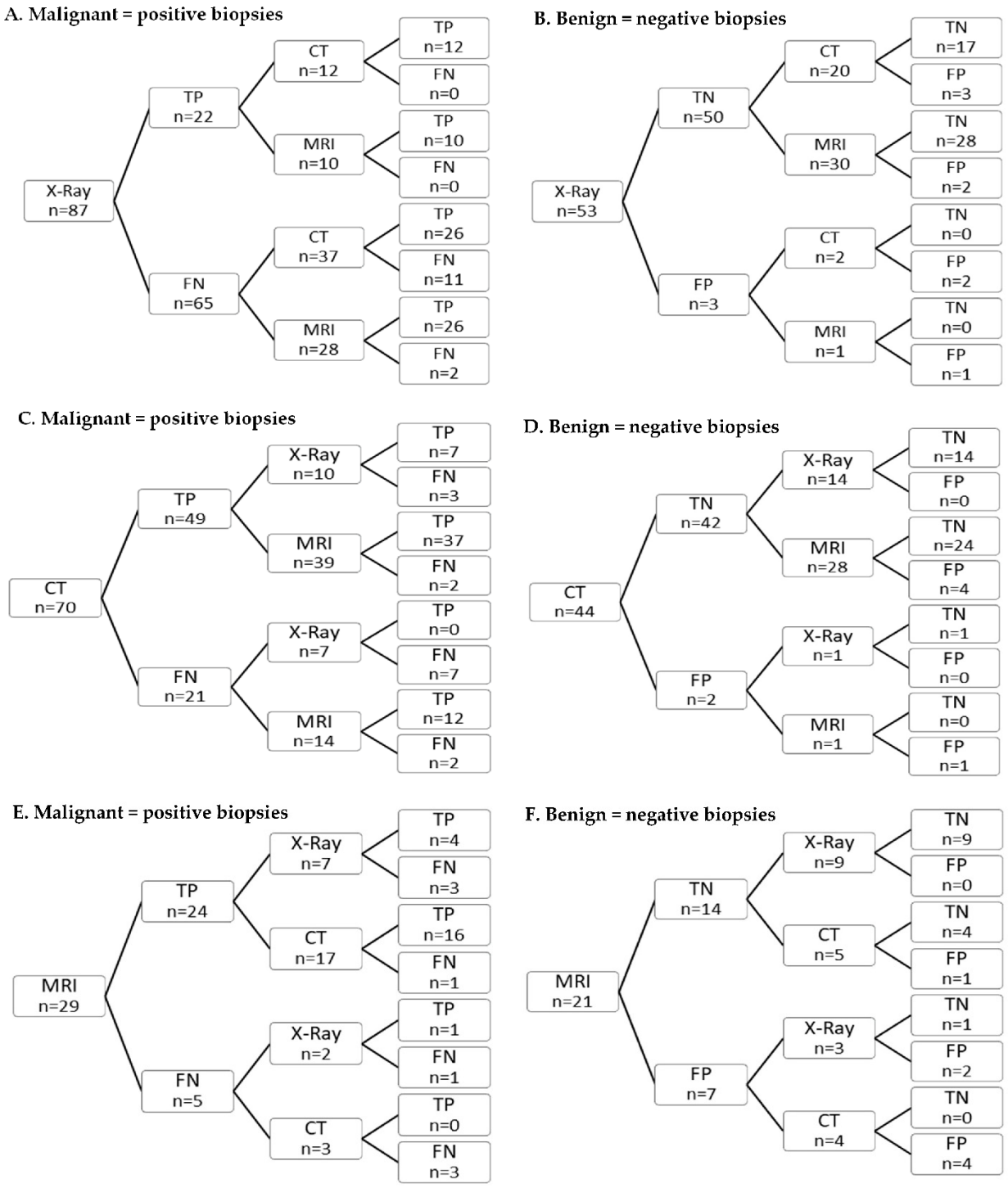


Figure 2. Sequence analyses. (**A**) Diagnostic results for malignant (positive) biopsies when X-ray is performed as the first modality (TP = true positive, FN = false negative). (**B**) Diagnostic results for benign (negative) biopsies when X-ray is performed as the first modality (TN = true negative; FP = false positive). (**C**) Diagnostic results for malignant (positive) biopsies when CT is performed as the first modality (TP = true positive, FN = false negative). (**D**) Diagnostic results for benign (negative) biopsies when CT is performed as the first modality (TN = true negative; FP = false positive). (**E**) Diagnostic results for malignant (positive) biopsies when MRI is performed as the first modality (TP = true positive, FN = false negative). (**F**) Diagnostic results for benign (negative) biopsies when MRI is performed as the first modality (TN = true negative; FP = false positive).

4. Discussion

Without taking the imaging sequence into consideration, MRI was shown to have the highest accuracy, followed by CT and X-ray, and MRI also showed the highest sensitivity and NPV, whereas X-ray proved to have the highest specificity and CT had the highest PPV (Table 2). These findings are consistent with previously published data, out of which one study is against a pathology proven reference [24,27–30]. These imaging characteristics are generally well recognized by radiologists.

Taking image sequence into consideration, our results show that there is no significant difference to prove that the diagnostic accuracy of X-ray, CT or MRI is influenced by access to prior imaging examinations and reports of one of the other modalities. This finding is controversial because it is not in accordance with previous studies and present guidelines, describing the importance of always comparing actual imaging with previous examinations and reports [5,7,10–15,17–19]. There might be several explanations for our findings.

Primarily, the lack of significance may be caused by the small subgroups. Second, the lack of difference in accuracy could cover the two opposing situations, as when MRI is the first imaging modality, it can either be correct or incorrect. According to our sequence analysis, when the MRI is correct, then the subsequent CT or X-ray is more likely to be correct, and when MRI is incorrect, then the subsequent CT or X-ray is incorrect in more than 80% of the situations. These two situations might balance each other so that the accuracy does not change significantly compared to whether a modality is preceded by MRI.

When X-ray is the first modality, 75% are expectedly false negatives, but only 7% of the subsequent MRI and 30% of the subsequent CT examinations are false negatives as well, which could indicate that X-ray results are rightfully not considered to have a high sensitivity and therefore do not influence the reader's evaluation of the second imaging much. When CT is the first modality, 30% are false negatives, and then all the following 7 X-ray examinations are negative, whereas only 14% of the subsequent MRIs are negative as well. CT has a higher accuracy than X-ray, and therefore, the reader might tend to attach greater value to the results from CT than those from the X-ray itself, whereas this is not the case for MRI compared to CT. When MRI is the first modality, only 17% are false negatives, with all subsequent CT scans being false negatives as well. Again, the reader might put more value on the previous MRI.

The specificities of X-ray (98.0) and CT (93.6) as stand-alone are high and decrease when preceded by MRI. On the contrary, the specificity increases for MRI (from 71.9 to 90.3) when preceded by an X-ray. Since X-ray specificity is known to be high, it might influence the reader of the consequent MRI scan.

One might conclude that the higher the diagnostic accuracy a given modality is known to have, the higher the bias of the diagnostic accuracy of the subsequent modalities will be and therefore that the sequence of the imaging modalities is important, especially if the diagnosis of the first modality is proven false. It has been shown previously that if one looks at a prior negative report before looking at imaging studies, there is a greater chance of missing a significant abnormality than by looking at the imaging studies first, but in these studies all imaging involved was X-ray and no other modality was included [5,17].

A direct comparison of the different imaging sequences to evaluate which sequence would be interesting for diagnosis and follow-up should be made with caution. It was not the purpose of our study; some groups are small, and we have not been able to prove any significant differences. MRI preceded by X-ray showed a sensitivity of 94.7 and a PPV of 92.3, slightly higher than MRI preceded by CT, showing a sensitivity of 92.5 and a PPV of 90.7. Since CT gives a higher radiation dosage and is more expensive than X-ray, you could speculate if X-ray followed by MRI would be the best strategy. This could make sense if you consider the bone lesion to be an isolated lesion, but since the malignant lesions represent metastases, you will most often need a CT scan to identify a primary tumor and/or to see if the skeletal lesion is the only metastasis present. The benefit of CT is that it is a whole-body examination, which is more readily available and inexpensive than whole-body MRI or

whole-body fusion imaging techniques such as PET-CT or PET-MRI. Therefore, it would be impossible to avoid performing a CT scan in most cases. Further prospective research is necessary to clarify this topic.

To the best of our knowledge, no direct comparison of pathology-proven diagnostic accuracy, including X-ray, CT, or MRI, with or without previous imaging examinations and the sequence of those has been conducted. Such knowledge should be considered quite important, not only in everyday imaging reporting but also in cases of claims of medical negligence. We identified four studies investigating whether access to prior examinations was valuable. All studies compared plain radiographs with prior plain radiographs and were based on questionnaires completed by the interpreting radiologists on whether they found access to prior examinations to be valuable or not [12,13]. Nevertheless, all present guidelines emphasize the importance of comparison with prior diagnostic examinations and reports of any modality available; however, these recommendations do not seem to be evidence-based.

Our findings could indicate that guidelines for good practice of radiological imaging reading and reporting should point out the importance of the readers not being influenced too much by previous imaging, especially not if these are modalities that are usually considered to be more accurate than the current one and that biopsy should be considered the gold standard for a valid diagnosis [25]. In clinical practice, one should consider evaluating the present study without a prior review of previously available imaging studies. When an independent evaluation has been formed, you can look at the available previous studies. If these conflict with your assessment, you should consider whether you want to be influenced and if so, you could note this in the description.

In addition to the small number in some of our subgroups, there are other limitations to our study. Table S1 in the Supplement shows that 33% of the biopsies had 3 or 4 imaging scans performed, which is a bias to the results since modalities other than the one analyzed could influence the diagnostic accuracy. However, there was no significant difference between the diagnostic accuracy regardless of whether the imaging investigated was preceded by other modalities. Furthermore, the readers had access to clinical information via Clinical Suite, and we do not know how many actually received this clinical information, which is known to influence the diagnostic reports [20–23]. Finally, it has been shown that the localization of the lesion has an influence on the diagnostic accuracy, with MRI showing superior diagnostic properties in spine lesions, whereas in non-spine lesions, the accuracy of the imaging modalities is largely comparable [24]. In our study, the spine accounted for 55% of the localizations, extremities for 18 % and pelvis for 17 %. The limited sample size does not allow for subgroup analysis on localization, which might represent a limitation. In conclusion, our study demonstrates the contribution to the discussion of the possible influence of previous imaging and reporting on diagnostic accuracy and how this possible influence should be addressed in future guidelines for the interpretation and reporting of diagnostic imaging. New prospective studies on this topic are needed for this purpose.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/diagnostics12071735/s1, Table S1: Modality sequence.

Author Contributions: Conceptualization: M.B.L., L.J.P. and H.D.Z.; methodology: M.B.L., L.J.P., H.D.Z., N.H.B. and M.L.; software: M.L., N.H.B. and M.B.L.; formal analysis: M.B.L. and M.L.; investigation: M.B.L. and M.L.; writing—original draft preparation: M.B.L.; writing—review and editing: M.B.L., H.D.Z., L.J.P. and M.B.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This retrospective study did not require ethical approval or informed consent in accordance with national legislation. The Danish Data Protection Agency approved the study and provided permission to access medical files for the purpose of the study.

Informed Consent Statement: Patient consent was waived due to approval from The Danish Data Protection Agency to access medical files for the purpose of the study.

Data Availability Statement: Data supporting reported results can be found in a special locked folder with an excel sheet within our institution and can be provided if necessary.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Busby, L.P.; Courtier, J.L.; Glastonbury, C.M. Bias in radiology: The how and why of misses and misinterpretations. *Radiographics* **2018**, *38*, 236–247. [CrossRef] [PubMed]
2. Pinto, A.; Brunese, L.; Pinto, F.; Reali, R.; Daniele, S.; Romano, L. The concept of error and malpractice in radiology. *Semin. Ultrasound CT MRI* **2012**, *33*, 275–279. [CrossRef]
3. Robinson, J.D.; Hippe, D.S.; Deconde, R.P.; Zecevic, M.; Mehta, N. Emergency radiology: An underappreciated source of liability risk. *J. Am. Coll. Radiol.* **2020**, *17*, 42–45. [CrossRef] [PubMed]
4. Ahn, Y.; Hong, G.-S.; Park, K.J.; Lee, C.W.; Lee, J.H.; Kim, S.-O. Impact of diagnostic errors on adverse outcomes: Learning from emergency department revisits with repeat CT or MRI. *Insights Imaging* **2021**, *12*, 160. [CrossRef] [PubMed]
5. Raskin, M.M. Survival strategies for radiology: Some practical tips on how to reduce the risk of being sued and losing. *J. Am. Coll. Radiol.* **2006**, *3*, 689–693. [CrossRef] [PubMed]
6. Pinto, L.B.A.; Brunese, L. Spectrum of diagnostic errors in radiology. *World J. Radiol.* **2010**, *2*, 377–383. [CrossRef] [PubMed]
7. Bruno, M.A.; Walker, E.A.; Abujudeh, H.H. Understanding and confronting our mistakes: The epidemiology of error in radiology and strategies for error reduction. *Radiographics* **2015**, *35*, 1668–1676. [CrossRef]
8. Degnan, A.J.; Ghobadi, E.H.; Hardy, P.; Krupinski, E.; Scali, E.P.; Stratchko, L.; Ulano, A.; Walker, E.; Wasnik, A.P.; Auffermann, W.F. Perceptual and interpretive error in diagnostic radiology—Causes and potential solutions. *Acad. Radiol.* **2019**, *26*, 833–845. [CrossRef]
9. Abujudeh, H.H.; Boland, G.W.; Kaewlai, R.; Rabiner, P.; Halpern, E.F.; Gazelle, G.S.; Thrall, J.H. Abdominal and pelvic computed tomography (CT) interpretation: Discrepancy rates among experienced radiologists. *Eur. Radiol.* **2010**, *20*, 1952–1957. [CrossRef]
10. Kahn, C.E.; Langlotz, C.P.; Burnside, E.S.; Carrino, J.A.; Channin, D.S.; Hovsepian, D.M.; Rubin, D.L. Toward best practices in radiology reporting. *Radiology* **2009**, *252*, 852–856. [CrossRef]
11. *Standards for Interpretation and Reporting of Imaging Investigations*, 2nd ed.; The Royal College of Radiologists: London, UK, 2008; Available online: https://www.rcr.ac.uk/publication/standards-interpretation-and-reporting-imaging-investigations-second-edition (accessed on 22 October 2021).
12. Hunter, T.B.; Boyle, R.R. The value of reading the previous radiology report. *AJR Am. J. Roentgenol.* **1988**, *150*, 697–698. [CrossRef] [PubMed]
13. Berbaum, K.S.; Smith, W.L. Use of reports of previous radiologic studies. *Acad. Radiol.* **1998**, *5*, 111–114. [CrossRef]
14. Berlin, L. Comparing new radiographs with those obtained previously. *AJR Am. J. Roentgenol.* **1999**, *172*, 3–6. [CrossRef] [PubMed]
15. European Society of Radiology. Good practice for radiological reporting. Guidelines from the European Society of Radiology (ESR). *Insights Imaging* **2011**, *2*, 93–96. [CrossRef]
16. Brady, A.; Laoide, R.; McCarthy, P.; McDermott, R. Discrepancy and error in radiology: Concepts, causes and consequences. *Ulst. Med. J.* **2012**, *81*, 3–9.
17. Berlin, L. Malpractice issues in radiology. Alliterative errors. *AJR Am. J. Roentgenol.* **2000**, *174*, 925–931. [CrossRef]
18. White, K.; Berbaum, K.; Smith, W.L. The role of previous radiographs and reports in the interpretation of current radiographs. *Investig. Radiol.* **1994**, *29*, 263–265. [CrossRef]
19. Aideyan, U.O.; Berbaum, K.; Smith, W.L. Influence of prior radiologic information on the interpretation of radiographic examinations. *Acad. Radiol.* **1995**, *2*, 205–208. [CrossRef]
20. Hattori, S.; Yokota, H.; Takada, T.; Horikoshi, T.; Takishima, H.; Mikami, W.; Mukai, H.; Funakoshi, H.; Naganuma, T.; Yano, T.; et al. Impact of clinical information on CT diagnosis by radiologist and subsequent clinical management by physician in acute abdominal pain. *Eur. Radiol.* **2021**, *31*, 5454–5463. [CrossRef]
21. Berbaum, K.S.; Franken, E.A. Commentary: Does clinical history affect perception? *Acad. Radiol.* **2006**, *13*, 402–403. [CrossRef]
22. Doubilet, P.; Herman, P.G. Interpretation of radiographs: Effect of clinical history. *AJR Am. J. Roentgenol.* **1981**, *137*, 1055–1058. [CrossRef] [PubMed]
23. Loy, C.T.; Irwig, L. Accuracy of diagnostic tests read with and without clinical information: A systematic review. *JAMA* **2004**, *292*, 1602–1609. [CrossRef] [PubMed]
24. Lange, M.B.; Nielsen, M.L.; Andersen, J.D.; Lilholt, H.J.; Vyberg, M.; Petersen, L.J. Diagnostic accuracy of imaging methods for the diagnosis of skeletal malignancies: A retrospective analysis against a pathology-proven reference. *Eur. J. Radiol.* **2016**, *85*, 61–67. [CrossRef] [PubMed]
25. Lange, M.B.; Petersen, L.J.; Nielsen, M.B.; Zacho, H.D. Validity of negative bone biopsy in suspicious bone lesions. *Acta Radiol. Open* **2021**, *10*, 20584601211030662. [CrossRef] [PubMed]
26. Bruun, N. Stata Module to Build, Present and Style Stata Matrices, Statistical Software Components S458201. Available online: https://ideas.repec.org/c/boc/bocode/s458201.html (accessed on 21 October 2021).
27. Coleman, R.E.; Croucher, P.I.; Padhani, A.R.; Clézardin, P.; Chow, E.; Fallon, M.; Guise, T.; Colangeli, S.; Capanna, R.; Costa, L. Bone metastases. *Nat. Rev. Dis. Prim.* **2020**, *6*, 83. [CrossRef] [PubMed]

28. Macedo, F.; Ladeira, K.; Pinho, F.; Saraiva, N.; Bonito, N.; Pinto, L.; Gonçalves, F. Bone metastases: An overview. *Oncol. Rev.* **2017**, *11*, 321. [CrossRef]
29. O'Sullivan, G.J.; Carty, F.L.; Cronin, C.G. Imaging of bone metastasis: An update. *World J. Radiol.* **2015**, *7*, 202–211. [CrossRef]
30. Heindel, W.; Gübitz, R.; Vieth, V.; Weckesser, M.; Schober, O.; Schäfers, M. The diagnostic imaging of bone metastases. *Dtsch. Arztebl. Int.* **2014**, *111*, 741–747. [CrossRef]

Original Article

Acta Radiologica Open
10(7) 1–7

DOI: 10.1177/20584601211030662
journals.sagepub.com/home/arr
SAGE

Validity of negative bone biopsy in suspicious bone lesions

Mine B Lange[1,2], Lars J Petersen[2,3], Michael B Nielsen[4] and Helle D Zacho[2,3]

Abstract

Background: The presence of malignant cells in bone biopsies is considered gold standard to verify occurrence of cancer, whereas a negative bone biopsy can represent a false negative, with a risk of increasing patient morbidity and mortality and creating misleading conclusions in cancer research. However, a paucity of literature documents the validity of negative bone biopsy as an exclusion criterion for the presence of skeletal malignancies.

Purpose: To investigate the validity of a negative bone biopsy in bone lesions suspicious of malignancy.

Material and Method: A retrospective cohort of 215 consecutive targeted non-malignant skeletal biopsies from 207 patients (43% women, 57% men, median age 64, and range 94) representing suspicious focal bone lesions, collected from January 1, 2011, to July 31, 2013, was followed over a 2-year period to examine any additional biopsy, imaging, and clinical follow-up information to categorize the original biopsy as truly benign, malignant, or equivocal. Standard deviations and 95% confidence intervals were calculated.

Results: 210 of 215 biopsies (98%; 95% CI 0.94–0.99) showed to be truly benign 2 years after initial biopsy. Two biopsies were false negatives (1%; 95% CI 0.001–0.03), and three were equivocal (lack of imaging description).

Conclusion: Our study documents negative bone biopsy as a valid criterion for the absence of bone metastasis. Since only 28% had a confirmed diagnosis of prior cancer and not all patients received adequately sensitive imaging, our results might not be applicable to all cancer patients with suspicious bone lesions.

Keywords

Bone, cancer, metastasis, tumor, biopsy, diagnostic accuracy

Received 1 December 2020; Accepted 15 June 2021

Introduction

Metastases account for the majority of malignant bone lesions.[1] Of the metastatic lesions, 85% originate from the breast, lung, prostate, kidney, and thyroid, and in 25–30% of cases, they are the first manifestation of malignancy.[2] Approximately 70% of breast and 90% of prostate cancer patients eventually develop skeletal metastases, which represents the third most frequent metastatic site behind the lung and liver, the most frequent metastatic site among men, and the second most frequent metastatic site among women.[3–8]

[1]Department of Radiology, Department of Diagnostic Imaging, University Hospital of North Zealand, Hilleroed, Denmark
[2]Department of Nuclear Medicine, Clinical Cancer Research Center, Aalborg University Hospital, Aalborg, Denmark
[3]Department of Clinical Medicine, Aalborg University, Aalborg, Denmark
[4]Department of Radiology, National University Hospital, University of Copenhagen, Copenhagen, Denmark

Corresponding author:
Mine B Lange, Department of Radiology, University Hospital of North Zealand, Dyrehavevej 29, Hilleroed 3400, Denmark.
Email: bl@benedictelange.dk

Early diagnosis and treatment of skeletal metastases are crucial because the impact on patient morbidity, including bone pain, fractures, hypercalcemia, and spinal cord compression, is significant and associated with considerable use of healthcare resources.[3,9–13] Only 20% of breast cancer patients remain alive 5 years after the discovery of bone metastasis.[13]

A prompt multimodal management approach depends on early diagnosis, which is most often based on a combination of imaging, clinical information, blood samples, and, to a lesser extent, bone biopsies.[14] Up to 30% of patients have skeletal metastases from an unknown primary neoplasm despite a thorough history, physical examination, appropriate laboratory testing, and advanced imaging technology.[2] In such situations, only extensive histopathological investigations of bone specimens from biopsies can reveal the primary malignancy, and bone biopsies are generally considered the gold standard, that is, error-free reference standard, for verification of the presence or absence of skeletal malignancy.[2,3,9,13,15] Nevertheless, bone biopsies are rarely performed, even in diagnostic test accuracy trials, and an insufficient reference standard has been identified as a major error source in biomedical research.[15–21]

When a bone biopsy documents presence of malignant cells, it is considered the gold standard for verification of cancer.[15] However, a negative biopsy can be a false negative, as observed with unrepresentative tissue sampling, especially in cases of benign bone lesions.[22–26] Such diagnostic errors can have serious consequences and can also supply misleading results in diagnostic accuracy studies.[27] To the best of our knowledge, the validity of a negative or benign bone biopsy for exclusion of skeletal metastases remains to be documented.

The aim of our study was to investigate whether targeted bone biopsies described as non-malignant or benign identified in a population with a suspicious focal bone lesion are in fact truly benign after 2 years of follow-up.

Material and methods

Subjects

A computer search of pathology samples (hereafter named biopsies) representing consecutive bone material registered by SNOMED (Systematized Nomenclature of Medicine) T10* and T11* codes for skeletal cytology and histology biopsies from January 1, 2011, to July 31, 2013, was performed, providing a retrospective cohort of 409 consecutive targeted bone biopsies from 395 patients, who had undergone imaging (X-ray, CT, MRI, bone scintigraphy, or PET/CT) within 6 months of the biopsy. From these data, we extracted all skeletal biopsies categorized as benign or non-malignant, resulting in a total of 215 biopsies from 207 patients (43% women, 57% men, median age 64, and range 94).

Definitions and data collection

Based on the unique Danish Central Personal Registration system, which supplies each inhabitant with a personal ID number, in combination with databases that include diagnostic codes for each ID number, it is possible for researchers to follow patient groups with selected diagnoses for a longer period.

A 2-year follow-up for each of the 207 patients representing 215 non-malignant bone biopsies was conducted by two independent readers who reached consensus. The follow-up included a careful computer search on each patient identified by the unique Danish social security number in the pathology database for any additional biopsy, in the imaging system (EasyViz, Karos Health Inc., Waterloo, ON, Canada) for any imaging of the relevant structure, and finally, in the Electronical Patient Journal charts (EPJ - Clinical Suite, CSC Scandihealth A/S) for any relevant journal notes in order to categorize the original biopsy as truly benign, malignant, or equivocal.

The criteria for a biopsy defined as truly benign after 2 years of follow-up were (1) negative biopsy from the same anatomy or no biopsy from the same anatomy, (2) no imaging with suspicion of malignancy from the same anatomy, and (3) no clinical suspicion of malignancy from the same anatomy. A biopsy was considered malignant if one of the following criteria applied: (1) positive biopsy from the same structure or from adjacent soft tissue, (2) any positive imaging of the structure, or (3) clinical suspicion of malignancy from the relevant anatomy, for example, persistent symptoms or blood tests leading to additional diagnostic tests. Biopsies not classified as true benign or true malignant based on follow-up were categorized as equivocal. This category also included patients with post-biopsy imaging that was indeterminate for malignancy.

Statistical tests

Descriptive statistics included calculation of standard deviations and 95% confidence intervals (95% CI).

Approval

This retrospective observational study did not require ethical approval or informed consent in accordance with national legislation. The Danish Data Protection Agency approved the study and gave permission to access medical files for the purpose of the study.

Results

Characteristics of bone biopsies at the time of inclusion in the study

As mentioned, a total of 215 benign bone biopsies representing 207 patients were included in the study. Eight

Table 1. Types of prior history of cancer.

Cancer types n = 59 (57 patients, 2 patients with 2 types)	n (%)
Breast	12 (20)
Pulmonary (SCLC 9, NSCLC 2)	11 (19)
Colorectal	9 (15)
Prostate	5 (8)
Sarcoma	4 (7)
Oral cavity	4 (7)
Non-Hodgin's lymphoma	4 (7)
Urine bladder	3 (5)
Malign melanoma	2 (3)
Other (multiple myeloma, thyroid, cervix, esophagus, pancreas)	5 (9)

patients had two additional biopsies performed, out of which six were re-biopsies from the exact same anatomy. Overall, 57 patients (28%) had a diagnosis of cancer prior to the initial bone biopsy as described in Table 1. Forty-six percent of the bone biopsies were taken from the extremities followed by 29% from the spine, 12% from the cranium, 8% from the thorax, and 5% from the thoracic skeleton.

Surgical interventions accounted for 163 of 215 (76%) biopsies. The majority of these interventions (75/163, 46%) represented surgical resection from the anatomy in question, whereas samples acquired during alloplastic surgery, osteosynthesis, and spondylodesis accounted for 35, 24, and 21 biopsies (21, 15, and 13%), respectively. Fluoroscopy-assisted biopsy accounted for 40% of the nonsurgical biopsies and only 10% of the total biopsies, whereas CT-guided biopsies only accounted for 4% of the total biopsies. The remaining data represented vertebroplasty, arthroscopically acquired material and autopsies.

Among the 215 biopsies, the indication was suspected malignancy for 84 (39%), no indication was given for 5 lesions (2%), and a possible benign lesion was specified as an indication for 126 (59%) of the lesions. Removal of benign lesions such as cysts, enchondromas, non-ossifying fibromas, osteochondromas, and osteoid osteomas can be indicated in cases of discomfort, imminent fracture risk, or cosmetic problems, and in connection with removal, samples can be sent for pathological evaluation as well as samples taken in cases of slight uncertainty.

The most frequent pathological diagnoses are listed in Table 2. Inflammation was the main finding reported, followed by no malignancy and fibrosis.

True and false negative findings

According to our criteria, 210 of 215 biopsies were truly benign 2 years after the initial biopsy, as described in Table 3.

Table 2. Primary pathological diagnosis of the included 215 benign biopsies.

Diagnoses	n (%)
Inflammation	28 (13)
No malignancy	24 (11)
Fibrosis	22 (10)
Osteochondroma	20 (9)
Degenerative changes	14 (7)
Unspecific reactive change	14 (7)
Necrosis	14 (7)
Fracture	13 (6)
Others (cyst, exostosis, hemangioma, Paget's, granuloma, hemorrhage a.o.)	66 (30)

Three cases were questionably benign due to equivocal imaging (no description).

Follow-up

Pathology. Additional biopsies from the same, adjacent, or different structure were performed in 39% of the cases during the 2-year follow-up period, of which two demonstrated the primary benign diagnosis wrong (Table 4).

In the first case, the primary biopsy was performed in March 2011 in connection with a percutaneous spinal decompression of L4 and was described as benign. A second biopsy was performed in April of the same year under general anesthesia, and this time, the biopsy was described as malignant (multiple myeloma). Imaging and notes in the journal persistently described the focal change in L4 as suspicious for malignant process. The patient died 3 years later in 2014. In the second case, the primary biopsy was acquired via CT-guided imaging in March 2013 from the ala of the right iliac bone and was described as benign, but doubt based on imaging was raised with respect to the representativeness of the tissue. A second CT-guided biopsy was performed in September 2013 and was once again described as benign, but doubt based on imaging and symptoms was persistently raised with respect to the representativeness of the sample. A third CT-guided biopsy including adjacent soft tissue was performed in May 2014, and this soft tissue sample demonstrated malignant cells from urinary bladder cancer. The patient died within the 2-year follow-up period.

The main location of the benign follow-up biopsies was the cervix (21%), skin (18%), and the GI tract (18%) followed by extremities and urinary system (7%, respectively). Diagnoses of the malignant follow-up biopsies can be found in Table 5.

Imaging. Among the original 215 biopsies, 160 (74%) were followed by imaging of the same anatomic region as the previous benign biopsy within the following 2 years. In the

Table 3. 2-year follow-up upon validity of negative bone biopsy.

Biopsies (n = 215)	n (%)
Truly benign Negative biopsy or no biopsy from the same anatomy, no imaging with suspicion of malignancy from the same anatomy, and negative or no clinical suspicion of malignancy from the same anatomy	210 (98%) (95% CI 0.94–0.99)
Questionably benign Negative biopsy or no biopsy from the same anatomy, equivocal imaging of the same anatomy, and negative or no clinical suspicion of malignancy from the same location	3 (1%) (95% CI 0.001–0.03)
Truly malignant Positive biopsy from the same anatomy or from adjacent soft tissue or positive imaging or clinical suspicion of malignancy from the relevant anatomy	2 (1%) (95% CI 0.001–0.03)

Table 4. Pathological samples during the 2-year follow-up.

Biopsies (n = 226)[a]	n (%)	Benign	Malignant
No biopsies performed at all, n (%)	138 (61)	—	—
Biopsies performed, n (%)	88 (39)[a]	n = 65 (29)	n = 23 (10)
Biopsy from exact same structure, n (%)[a]	6 (3)	5 (2)	1 (1)
Biopsy from adjacent or other structure, n (%)[a]	82 (36)	60 (27)	22 (9)

[a]Selected cases had more than one biopsy performed in the follow-up period.

Table 5. Diagnosis of malignant biopsies from same, adjacent, or different structures in the 2-year period.

Diagnosis of malignant biopsies including (n = 23)	n (%)
Skin	7 (31)
Urinary bladder	4 (17)
Oral cavity	2 (9)
Pancreas	2 (9)
SCLC	2 (9)
Breast	2 (9)
Small intestine	1 (4)
Lymphoma	1 (4)
Colorectal	1 (4)
Multiple myeloma (L4)	1 (4)

great majority of cases (n = 155), the follow-up imaging was described as benign, but in two lesions, imaging was consistently described as malignant. These two lesions, representing two patients, resulted in new biopsies, one from the same anatomy and one from adjacent soft tissue, and both were described as malignant. Three imaging cases were equivocal (not described). Sixty patients had two or more imaging performed, giving a total of 205 imaging performed. X-ray was the most frequently used imaging follow-up procedure followed by CT scan and MRI. Please refer to Table 6.

Out of the 59 biopsies taken from 57 patients with prior cancer diagnosis, 9 (15%) did not have any imaging performed in the follow-up period at all. Twenty-one had only one imaging performed, out of which X-ray represented 9 (15%). Twenty-two received 2 different imaging modalities and 7 received 3, meaning that 70% received at least one sensitive diagnostic imaging modality.

Clinical. According to clinical follow-up, 189 biopsies (88%) were not suspected to harbor a malignant condition in the same anatomy as the initial bone biopsy. In 24 (11%) biopsies, no journal notes were recorded in the period data. In the two lesions for which additional imaging and biopsies were described as malignant, the journal notes initially described the lesions as suspicious for malignancy. Later notes described the diagnosis of malignancy as confirmed by imaging and pathology.

Discussion

Based on a 2-year follow-up examination of available additional biopsies, imaging, and clinical information on 215 consecutive negative bone biopsies from 207 patients, we documented that 98% (210 biopsies) were indeed true negative after 2 years, proving a high validity of negative bone biopsies as an expression of the absence of skeletal metastases. Three cases were questionably benign due to equivocal imaging (no description), and only two were actual false negatives. We believe that we have proven a negative biopsy to be a valid marker of exclusion of the presence of skeletal metastases, which has not been previously documented, to the best of our knowledge.

The most frequent pathological inclusion diagnoses in our study were inflammation, no malignancy, fibrosis, osteochondroma, and degenerative changes, in accordance with other findings indicating that our material is representative.[28]

Table 6. Imaging at the 2-year follow-up.

Imaging results (n = 215)	n (%)
Imaging of same anatomy negative	155 (72)
Imaging of same anatomy positive	2 (1)
No imaging performed	55 (26)
Equivocal imaging (not described)	3 (1)
Types of imaging performed at the 2-year follow-up (n = 205; 60 had two or more imaging performed)	
X-ray	126 (62)
Computed tomography (CT)	44 (21)
Magnetic resonance imaging MRI	27 (13)
Bone scintigraphy (BS)	4 (2)
18F-fluorodeoxyglucose positron emission tomography (FDG-PET/CT)	4 (2)

Forty-six percent of the biopsies in our study were taken from extremities, whereas only 29% came from the spine, which is to be expected based on the knowledge that the spine, pelvis, ribs, and ends of long bones are preferred destinations of metastases because of their high red marrow content, whereas extremities are more often the seat of benign lesions.[5,29,30]

In many instances, benign bone tumors and tumor-like lesions of the bone can be diagnosed solely through conventional X-rays and require no biopsies for clarification. In our study, a benign lesion was specified as an indication for biopsy in 59% of cases, the majority of which were performed as a routine component of surgery. This result is in accordance with the findings of Scheitza et al., who demonstrated that only 21% of their biopsies were performed for actual confirmation of a benign diagnosis. In all other cases, biopsies were performed as a routine component of surgery, which can be indicated in cases of discomfort, imminent fracture risk, or cosmetic problems, and in connection with such surgery, samples can be sent for pathological evaluation.[31]

Surgical biopsies accounted for 76% of our biopsy material, whereas CT-guided biopsies and fluoroscopy-assisted biopsy accounted for 4 and 10%, respectively. Open biopsy has been the conventional "gold standard" procedure for obtaining adequate and representative samples of tissue for diagnosis of musculoskeletal lesions, with a reported accuracy rate of 98%.[24] Recent results demonstrate that in carefully controlled situations in which the musculoskeletal radiologist works in a team approach with the orthopedic oncologist and orthopedic pathologist, the results from percutaneous biopsy can be highly effective and accurate. Additionally, most tumor treatment centers advocate for core biopsy performed under CT guidance, with a measured diagnostic yield ranging between 70 and 89%, a reported accuracy between 61 and 98%, and fewer biopsy complications than open surgical biopsy, even in sclerotic bone lesions.[32–34] Other studies claim that the disadvantage of the CT-guided biopsy method is that metabolically active lesions without distinctive morphologies might not be reliably assessable by CT-guided biopsy and the false negative biopsy rate of such lesions might be substantially higher, with one series documenting that 18% required open biopsy after needle biopsy.[22,32] The high percentage of surgical biopsies in our study might have contributed to the quality of samples and therefore to the low incidence of false negative results.

Spinal biopsies accounted for 29% of our material, and non-spinal biopsies accounted for 71%. Hau et al. demonstrated that the anatomical site has a significant effect on the accuracy of CT-guided biopsy, with non-spinal sites exhibiting greater accuracy (75%) than spinal sites (61%). Hau also showed that lesion size, type of margin, and gender did not influence the success or failure rates of the biopsies.[22] Because we only encountered two false negative and three equivocal lesions, our data set is too small to support those findings, but the high number of non-spinal lesions might have contributed to our high bioptic diagnostic accuracy even though only 4% of our material was obtained in CT-guided biopsies.

Two biopsies (1%) showed to be false negatives in repeated biopsies performed due to imaging (CT and MRI) that persistently described the lesions as positive. Monfardini et al. showed that 8 out of 10 false negatives, CT-guided biopsies had positive PET scans, and 6 out of 10 had positive MR scans, leading to the conclusion that a negative biopsy result in cases of suspicious PET and/or MR findings should be carefully evaluated and considered for a second sampling.[24]

A meta-analysis made by Cheng and Alavi concludes that 18F-FDG PET significantly outperforms iliac bone marrow biopsy in the detection of bone marrow infiltration in the initial staging of patients with Hodgkin's lymphoma and therefore should be used as a first-line study.[35] It has been recognized that bone marrow biopsy is associated with a high false-negative rate in early cases of Hodgkin's lymphoma[36] probably because the biopsy is taken from a standard anatomy without any prior suspicion of a focal lesion as opposed to the lesions identified in our study. Furthermore, the meta-analysis defines the tests under investigation as their own reference standard, as do the studies included, which might explain the significant interstudy heterogeneity in the sensitivity data of PET or iliac bone biopsy. This study underlines our statement of the importance of an error-free reference standard.

The bone matrix response to metastatic deposits is known to depend on the primary cancer type.[3] Osteolytic matrix response, characterized by destruction of normal bone, is present in multiple myeloma, renal cell carcinoma, malignant melanoma, non-small cell lung cancer, non-Hodgkin's lymphoma, thyroid cancer, and the great majority of breast cancers. Osteoblastic metastases, characterized by deposition of new bone matrix, are present in Hodgkin's

lymphoma, prostate cancer and small cell lung cancer and mixed metastases, where the patient has both osteolytic and osteoblastic lesions, is present in gastrointestinal cancers and 15–20% of breast cancers.[3] In our cohort, 59 biopsies were taken from patients with prior cancers, out of which approximately 50% would cause mainly osteolytic, 25% mainly osteoblastic, and 25% mixed metastatic bone response in case of bone metastases. Seventy percent received at least one diagnostic imaging modality sensitive for all three types.[20]

Our study has two main important implications. Primarily, we consider our findings important for future diagnostic, prognostic, and treatment purposes because false negative samples can lead to delayed diagnosis and consequently increased morbidity and mortality; bone biopsy from a suspicious lesion in patients with a known primary cancer has shown benign pathology in 21%.[3,9,10,24,28] Our result also has important implications for the validity of scientific studies of treatment efficacy and diagnostic accuracy using biopsy as reference, as underlined by STARD (The Standards for Reporting of Diagnostic Accuracy), which considers biopsy as the gold standard for demonstration or exclusion of bone metastases.[21] Our study documents that a negative sample can be considered truly negative and proves biopsy as a valid gold standard for the absence of metastases.

Our study also contains limitations. The first limitation is the retrospective nature of the data, including selection of patients and the lack of uniform and systematized follow-up including standardized re-biopsy from the same anatomy. Second, the pathologists did not perform a blind evaluation of the specimen and were aware of the patient's history, as in a normal clinical setting, which is known to have a possible effect on the final diagnosis.[24] Since only 28% had a confirmed diagnosis of prior cancer, our results might not be applicable to all cancer patients with suspicious bone lesions. Finally, the most frequently performed imaging modality in the follow-up period was X-ray (62%). A rather low X-ray sensitivity (33%) for diagnosis of skeletal malignancies has been demonstrated, and thus, this type of examination might have missed possible positive lesions and a subsequent repeated biopsy.[20] However, all samples were described as negative and 88% of the clinical follow-up did not raise any suspicion of malignancy, and thus, this impact might be limited.

In conclusion, we believe that our results show that it is reasonable to assess a negative bone biopsy as an indication of the absence of bone metastasis in the structure in question. These results offer value not only to diagnosis, morbidity, and mortality of metastatic bone disease but also to the accuracy of future treatment and diagnostic scientific studies. Prior cancer type, biopsy method, and site of the lesion should be taken into consideration, and possible repeated biopsy should be considered in cases of imaging that persistently describes the lesion as positive.

Acknowledgments

The authors acknowledge Michael Gade, Department of Nuclear Medicine, Clinical Cancer Research Center, Aalborg University Hospital, Aalborg, Denmark, for providing data collection assistance.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Mine B Lange https://orcid.org/0000-0003-3328-2633
Lars J Petersen https://orcid.org/0000-0003-2519-983X

References

1. Bernard S., Walker E., Raghavan M. An approach to the evaluation of incidentally identified bone lesions encountered on imaging studies. Am J Roentgenol 2017; 208: 960–970.
2. Piccioli A., Maccauro G., Spinelli M. S., et al. Bone metastases of unknown origin: epidemiology and principles of management. J Orthop Traumatol 2015; 16: 81–86.
3. Macedo F, Ladeira K, Pinho F, et al. Bone metastases: an overview. Oncol Rev 2017; 11: 321.
4. Filippiadis D, Mazioti A, Kelekis A. Percutaneous, imaging-guided biopsy of bone metastases. Diagnostics 2018; 8: 25. doi:10.3390/diagnostics8020025.
5. O'Sullivan GJ, Carty FL, Cronin CG. Imaging of bone metastasis: an update. World J Radiol 2015; 7: 202–211.
6. Riihimäki M, Thomsen H, Sundquist K, et al. Clinical landscape of cancer metastases. Cancer Med 2018; 7: 5534–5542.
7. Fornetti J, Welm AL, Stewart SA. Understanding the bone in cancer metastasis. J Bone Miner Res 2018; 33: 2099–2113.
8. Sailer V, Schiffman MH, Kossai M, et al. Bone biopsy protocol for advanced prostate cancer in the era of precision medicine. Cancer 2018; 124: 1008–1015.
9. Coleman RE. Clinical features of metastatic bone disease and risk of skeletal morbidity. Clin Cancer Res 2006; 12: 6243s–6249s.
10. Disibio G, French SW. Metastatic patterns of cancers: results from a large autopsy study. Arch Pathol Lab Med 2008; 132: 931–939.
11. Schulman KL, Kohles J. Economic burden of metastatic bone disease in the U.S. Cancer 2007; 109: 2334–2342.
12. Svendsen ML, Gammelager H, Svaerke C, et al. Hospital visits among women with skeletal-related events secondary to breast cancer and bone metastases: a nationwide population-based cohort study in Denmark. Clin Epidemiol 2013; 5: 97–103.

13. Pockett RD, Castellano D, McEwan P, et al. The hospital burden of disease associated with bone metastases and skeletal-related events in patients with breast cancer, lung cancer, or prostate cancer in Spain. Eur J Cancer Care 2010; 19: 755–760.
14. von Moos R, Costa L, Gonzalez-Suarez E, et al. Management of bone health in solid tumours: from bisphosphonates to a monoclonal antibody. Cancer Treat Rev 2019; 76: 57–67.
15. Wondergem M, van der Zant FM, van der Ploeg T, et al. A literature review of 18F-fluoride PET/CT and 18F-choline or 11C-choline PET/CT for detection of bone metastases in patients with prostate cancer. Nucl Med Commun 2013; 34: 935–945.
16. Duo J, Han X, Zhang L, et al. Comparison of FDG PET/CT and gadolinium-enhanced MRI for the detection of bone metastases in patients with cancer: a meta-analysis. Clin Nucl Med 2013; 38: 343–348.
17. Qu X, Huang X, Yan W, et al. A meta-analysis of (1)(8)FDG-PET-CT, (1)(8)FDG-PET, MRI and bone scintigraphy for diagnosis of bone metastases in patients with lung cancer. Eur J Radiol 2012; 81: 1007–1015.
18. Wu Q, Yang R, Zhou F, et al. Comparison of whole-body MRI and skeletal scintigraphy for detection of bone metastatic tumors: a meta-analysis. Surg Oncol 2013; 22: 261–266.
19. Korevaar DA, van Enst WA, Spijker R, et al. Reporting quality of diagnostic accuracy studies: a systematic review and meta-analysis of investigations on adherence to STARD. Evid Based Med 2014; 19: 47–54.
20. Lange MB, Nielsen ML, Andersen JD, et al. Diagnostic accuracy of imaging methods for the diagnosis of skeletal malignancies: A retrospective analysis against a pathology-proven reference. Eur J Radiol 2016; 85: 61–67.
21. Bossuyt PM, Reitsma JB, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. Radiology 2015; 277: 826–832.
22. Hau A, Kim I, Gebhardt MC, et al. Accuracy of CT-guided biopsies in 359 patients with musculoskeletal lesions. Skeletal Radiol 2002; 31: 349–353.
23. Rimondi E, Rossi G, Ruggieri P, et al. Percutaneous CT-guided biopsy of the musculoskeletal system: results of 2027 cases. Eur J Radiol 2011; 77: 34–42.
24. Monfardini L, Preda L, Renne G, et al. CT-guided bone biopsy in cancer patients with suspected bone metastases: retrospective review of 308 procedures. Radiol Med 2014; 119: 852–860.
25. Park VY, Kim EK, Kim MJ, et al. Evaluating imaging-pathology concordance and discordance after ultrasound-guided breast biopsy. Ultrasonography 2018; 37: 107–120.
26. Kamei Y, Nishida J, Satoh T, et al. Core needle percutaneous transpedicular vertebral body biopsy: a study of 128 cases. J Spinal Disord Tech 2015; 28: E394–E399.
27. Kubo T, Furuta T, Adachi N, et al. A meta-analysis supports core needle biopsy by radiologists for better histological diagnosis in soft tissue and bone sarcomas. Medicine 2018; 97: e11567.
28. Raphael B, Hwang S, Panicek DM, et al. Biopsy of suspicious bone lesions in patients with a single known malignancy: prevalence of a second malignancy. AJR Am J Roentgenol 2013; 201: 1309–1314.
29. Ren G, Esposito M, Kang Y. Bone metastasis and the metastatic niche. J Mol Med 2015; 93: 1203–1212.
30. Amelot A, Terrier LM, Pascal-Moussellard H, et al. Approaching spinal metastases spread profile. Surg Oncol 2019; 31: 61–66.
31. Scheitza P, Uhl M, Kayser C, et al. Interobserver variability in the differential diagnosis of benign bone tumors and tumor-like lesions. Rofo 2016; 188: 479–487.
32. Wu MH, Xiao LF, Chen Y, et al. PET/CT-guided versus CT-guided percutaneous core biopsies in the diagnosis of bone tumors and tumor-like lesions: which is the better choice? Cancer Imaging 2019; 19: 69.
33. Hegde V, Burke ZDC, Kelley BV, et al. Is core needle biopsy reliable in differentiating between aggressive benign and malignant radiolucent bone tumors? Clin Orthop Relat Res 2018; 476: 568–577.
34. Suh CH, Yun SJ. Diagnostic outcome of image-guided percutaneous core needle biopsy of sclerotic bone lesions: a meta-analysis. AJR Am J Roentgenol 2019; 212: 625–631.
35. Cheng G, Alavi A. Value of 18F-FDG PET versus iliac biopsy in the initial evaluation of bone marrow infiltration in the case of Hodgkin's disease: a meta-analysis. Nucl Med Commun 2013; 34: 25–31.
36. Hao B, Zhao L, Luo NN et al. Is it sufficient to evaluate bone marrow involvement in newly diagnosed lymphomas using (18)F-FDG PET/CT and/or routine iliac crest biopsy? A new approach of PET/CT-guided targeted bone marrow biopsy. BMC Cancer 2018; 18: 1192.

ISSN (online): 2246-1302
ISBN (online): 978-87-7573-762-8

AALBORG UNIVERSITY PRESS