RISK FACTORS AND BIOMARKERS FOR METASTATIC CUTANEOUS SQUAMOUS CELL CARCINOMA

Jaakko Knuutila

RISK FACTORS AND BIOMARKERS FOR METASTATIC CUTANEOUS SQUAMOUS CELL CARCINOMA

Jaakko Knuutila

University of Turku

Faculty of Medicine
Institute of Clinical Medicine
Department of Dermatology and Venereology
Doctoral Programme in Clinical Research (DPCR)

Supervised by

Professor Veli-Matti Kähäri, MD, PhD
Department of Dermatology and Venereology
University of Turku
Turku, Finland

Docent Pilvi Riihilä, MD, PhD
Department of Dermatology and Venereology
University of Turku
Turku, Finland

Docent Liisa Nissinen, PhD
Department of Dermatology and Venereology
University of Turku
Turku, Finland

Reviewed by

Associate Professor
Noora Neittaanmäki, MD, PhD
Department of Clinical Pathology
University of Gothenburg and
Sahlgrenska University Hospital
Gothenburg, Sweden

Professor
Virve Koljonen, MD, PhD
Department of Plastic Surgery
University of Helsinki
Helsinki, Finland

Opponent

Professor Ilkka Harvima, MD, PhD
Department of Dermatology
University of Eastern Finland
Kuopio, Finland

UNIVERSITY OF TURKU
Faculty of Medicine
Institute of Clinical Medicine
Dermatology and Venereology
JAAKKO KNUUTILA: Risk Factors and Biomarkers for Metastatic
Cutaneous Squamous Cell Carcinoma
Doctoral Dissertation, 242 pp.
Doctoral Programme in Clinical Research (DPCR)
October 2022

ABSTRACT

The incidence of cutaneous squamous cell carcinoma (cSCC), the most common skin cancer with metastatic potential, continues to increase. Although proportion of cSCCs metastasize and cause mortality, sufficient means to identify the metastasis-prone tumors are not available.

In this thesis the metastatic cSCCs from the area served by Turku University Hospital were identified and characterized revealing that the rate of metastasis in the study region was 2.3%. Further, it was discovered that metastasis occurs rapidly and that there was no history of cSCC in 85% of patients with metastatic cSCC. Invasion depth, tumor diameter, age and location on lower lip or forehead were associated with increased risk of metastasis. On the other hand, usage of isosorbide mono-/dinitrate and aspirin as well as comorbidity with premalignant lesions or basal cell carcinoma were associated with lower risk of metastasis.

With multiplexed immunohistochemistry, it was demonstrated that the activity and phenotype of cancer-associated fibroblasts (CAFs) evolve during the progression of cSCC. Elevation of α-smooth muscle actin (αSMA), secreted protein acidic and rich in cysteine (SPARC) and fibroblast activating protein (FAP) expression was associated with invasion and expression of FAP and platelet-derived growth factor receptor-β (PDGFRβ) with metastasis. High expression of stromal PDGFRβ and periostin were associated with worse prognosis. CAF107 (PDGFRα-/PDGFRβ+/FAP+) subset was associated with invasion and metastasis, and predicted poor prognosis of cSCC.

A deep learning algorithm was harnessed to distinguish primary tumors that metastasize rapidly from non-metastatic cSCCs with slide level area under the receiver operating characteristic curve (AUROC) of 0.747 on whole slide images representing primary cSCCs. Furthermore, a risk factor model, that utilized prediction by AI, was created and provided staging systems and comparative risk factor models surpassing classification and prognostivity.

These results characterize features associated with the metastasis risk of cSCC and indicate that CAF-markers and AI could provide clinical tools for the metastasis risk assessment and thus improve the prognosis of patient with metastatic cSCC.

KEYWORDS: cutaneous squamous cell carcinoma, metastasis, risk factor, cancer-associated fibroblast, deep learning

TIIVISTELMÄ

Yleisimmän etäpesäkkeitä lähettävän ihosyövän, okasolusyövän, ilmaantuvuus jatkaa kasvuaan. Vaikka osa okasolusyövistä lähettää etäpesäkkeitä ja aiheuttaa kuolleisuutta, ei etäpesäkkeitä lähettämään tulevien okasolusyöpien tunnistamiseksi ole toistaiseksi riittäviä keinoja.

Tässä väitöskirjassa karakterisoitiin Turun yliopistollisen keskussairaalan vastuualueen metastasoituneet okasolusyövät ja osoitettiin että tutkimusalueen okasolusyövistä 2.3% etenee etäpesäkkeitä lähettäväksi. Metastasoituminen tapahtui nopeasti ja valtaosassa tapauksista (85%) etäpesäkkeen lähetti ensimmäinen potilaalla todettu okasolusyöpä. Ikä, kasvaimen invaasiosyvyys, halkaisija ja sijainti alahuulessa tai otsalla yhdistyivät kohonneeseen metastaasiriskiin. Isosorbidinitraatin ja aspiriinin käyttö sekä esiasteiden ja tyvisolusyövän esiintyminen taas liittyivät alentuneeseen metastaasiriskiin.

Multiplex-immunohistokemiaa hyödyntäen osoitettin, että syöpään liittyvien fibroblastien (CAF) aktiviteetti ja ilmiasu muuttuu okasolusyövän edetessä. Kohonnut sileälihasaktiini alfan (αSMA), osteonektiinin ja fibroblastia aktivoivan proteiinin (FAP) ilmentyminen liittyi invaasioon ja FAP:n sekä verihiutaleista johdetun kasvutekijäreseptori β:n (PDGFRβ) etäpesäkkeiden lähettämiseen. PDGFRβ:n ja periostiinin ilmentyminen taas yhdistyi huonoon ennusteeseen. CAF107 (PDGFRα-/PDGFRβ+/FAP+) alatyyppi liittyi invaasioon, metastasointiin ja huonoon ennuste eeseen.

Etäpesäkkeitä lähettämään tulevien okasolusyöpien tunnistamiseen valjastettu syväoppimisalgoritmi erotti okasolusyöpiä edustavista digitalisoiduista mikroskopiakuvista nopeasti etäpesäkkeitä lähettävät okasolusyövät okasolusyövistä, jotka eivät lähetä etäpesäkkeitä, leiketason AUROC-arvolla 0.747. Tekoälyarviota hyödyntävä riskitekijämalli voitti luokittelujärjestelmät ja kilpailevat riskitekijämallit okasolusyöpien luokittelussa ja ennusteen arvioinnissa.

Tulokset antavat lisätietoa metastasoituvan okasolusyövän luonteesta ja osoittavat CAF-markkereiden sekä tekoälyn voivan tarjota kliinisiä työkaluja okasolusyövän metastaasiriskin arviointiin ja täten voivan parantaa etäpesäkkeitä lähettävän okasolusyöpäpotilaan ennustetta tulevaisuudessa.

AVAINSANAT: okasolusyöpä, ihon levyepiteelikarsinooma, etäpesäke, riskitekijä, syöpään liittyvä fibroblasti, syväoppiminen

Table of Contents

Abbreviations

αSMA	α smooth muscle actin
A	Adenine
AI	Artificial intelligence
AJCC	American Joint Committee on Cancer
AK	Actinic keratosis
ANN	Artificial neural network
ANXA5	Annexin A5
apCAF	Antigen presenting-CAF
APOBEC	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like
AuI	Augmented intelligence
AUROC	Area under the receiver operating characteristic curve
BCC	Basal cell carcinoma
BMP4	Bone morphogenetic protein 4
BM-MSC	Bone marrow-derived mesenchymal stem cell
BWH	Brigham and Women's Hospital
C	Cytosine
C3	Complement component 3
C3a	Complement component 3a
CAF	Cancer-associated fibroblast
Cav1	Caveolin 1
CD105	Endoglin
CDKN2A	Cyclin-dependent kinase inhibitor 2A
CFB	Complement factor B
CIS	Carcinoma in situ
CLL	Chronic lymphocytic leukemia
CNN	Convolutional neural network
Col1	Type I collagen
COL1A2	Collagen type I α 2 chain
COL11A1	Collagen type XI α 1 chain
CR	Complete response

cSCC	Cutaneous squamous cell carcinoma
cSCCIS	*In situ* cutaneous squamous cell carcinoma
CXCL	C-X-C chemokine ligand
DAPI	4′,6-diamidino-2-phenylindole
dCAF	Development CAF
DCIS	Ductal carcinoma in situ
DDOST	Dolichyl-diphosphooligosaccharide-protein glycosyl-transferase noncatalytic subunit
DDR2	Discoidin domain-containing receptor 2
DL	Deep learning
DNN	Deep neural network
DSS	Disease-specific survival
EADO	The European Association of Dermato-Oncology
ECM	Extracellular matrix
EDF	The European Dermatology Forum
EGFR	Epidermal growth factor receptor
EMA	European Medicines Agency
EMT	Epithelial-to-mesenchymal transition
EndMT	Endothelial-to-mesenchymal transition
ENE	Extranodal extension
EORTIC	The European Organisation for Research and Treatment of Cancer
FAP	Fibroblast activation protein alpha
FDA	US Food and Drug Administration
FGF2	Fibroblast growth factor 2
FN	Fibronectin
FSP1	Fibroblast specific protein 1
GEP	Gene expression profile
H&E	Hematoxylin & eosin
HGF	Hepatocyte growth factor
HIV	Human immunodeficiency virus
HNSCC	Head and neck squamous cell carcinoma
iCAF	Inflammatory CAF
IGF1	Insulin-like growth factor 1
IHC	Immunohistochemical/immunohistochemistry
IL	Interleukin
ITGA11	Integrin α11
ITGB1	Integrin subunit β1
KA	Keratoacanthoma
KIN	Keratinocytic intraepidermal neoplasia

LIF	Leukaemia inhibitory factor
LRRC15	Leucine-rich repeat containing 15
mCAF	Matrix CAF
mcSCC	Metastatic cutaneous squamous cell carcinoma
MFAP5	Microfibril-associated protein 5
MHC	Major histocompatibility complex
mfIHC	Multiplexed fluorescence immunohistochemistry
MMP	Matrix metalloprotein
MSC	Mesenchymal stem cell
myCAF	Myofibroblastic CAF
NF-κB	Nuclear factor-κB
NG2	Neural/glial antigen 2
NLP	Natural language processing
NO	Nitric oxide
non-mcSCC	Non-metastatic cutaneous squamous cell carcinoma
NSAID	Non-steroidal anti-inflammatory drug
OS	Overall survival
PDGF	Platelet derived growth factor
PDGFR	Platelet-derived growth factor receptor
PD-1	Programmed cell death protein 1
PD-L1	Programmed death-ligand 1
PDPN	Podoplanin
PNI	Perineural invasion
POSTN	Periostin
PY	Person year
RDEB	Recessive dystrophic epidermolysis bullosa
RDEBSCC	Recessive dystrophic epidermolysis bullosa-associated cSCC
ResNet	Residual neural network
RF	Risk factor
RFM	Risk factor model
rMSC	Tissue-resident mesenchymal stem cell
RNN	Residual neural network
RT	Radiotherapy, radiation therapy
S100A4	S100 calcium binding protein A4
scRNA-seq	Single-cell RNA sequencing
SHH	Sonic hedgehog
SK	Seborrheic keratosis
SPARC	Secreted protein acidic and rich in cysteine
T	Thymidine
TAF	Tumor-associated fibroblasts

TCGA	The Cancer Genome Atlas
TGFβ	Transforming growth factor β
TIL	Tumor-infiltrating lymphocyte
TIME	Tumor immune microenvironment
TMA	Tissue microarray
TME	Tumor microenvironment
TNC	Tenascin C
TP53	Tumor suppressor protein p53
UV	Ultraviolet
UV-cSCC	UV-incuded cSCC
vCAF	Vascular CAF
VEGFA	Vascular endothelial growth factor A
VIM	Vimentin
WNT5A	Wnt family member 5A
WSI	Whole slide imaging/whole slide image
YAP	Hippo-Yes-associated protein

List of Original Publications

This thesis is based on the following original publications, which are referred to in the text by their Roman numerals:

I   Knuutila JS, Riihilä P, Kurki S, Nissinen L, Kähäri VM. Risk factors and prognosis for metastatic cutaneous squamous cell carcinoma: a cohort study. *Acta Derm Venereol.* 2020; 100(16): adv00266.

II   Knuutila JS, Riihilä P, Nissinen L, Kallionpää R, Pellinen T, Kähäri VM. Cancer-associated fibroblast activation predicts progression, metastasis and prognosis of cutaneous squamous cell carcinoma. (Submitted manuscript)

III Knuutila JS, Riihilä P, Karlsson A, Tukiainen M, Talve L, Nissinen L, Kähäri VM. Identification of metastatic primary cutaneous squamous cell carcinoma utilizing artificial intelligence analysis of whole slide images. *Sci Rep.* 2022; 12(1): 9876.

The original publications have been reproduced with the permission of the copyright holders.

1   Introduction

Cutaneous squamous cell carcinoma (cSCC) is a keratinocyte carcinoma with metastatic potential and increasing incidence both globally and in Finland. A fraction of cSCCs metastasize but in metastatic case the prognosis is poor. Several clinical and histopathological features have been associated with the risk of metastasis but findings are ambiguous. Moreover, established staging systems utilized in the risk assessment of cSCC are unsatisfactory in predicting the risk of metastasis. In addition, to date, there are no clinically established prognostic or metastasis risk-associated biomarkers for cSCC. Thus there is need for better identification of high-risk cSCCs, which benefit from closer surveillance, more radical treatment, adjuvant therapies and staging studies. On the other hand, due to the increasing incidence and excellent post-surgical prognosis in most cases there is need to safely omit low-risk cases from further follow-up.

Cancer-associated fibroblasts (CAFs) contribute to the progression of solid cancers, however their role in cSCC is unestablished. It is widely acknowledged that CAF activity evolves as cancer progresses and that wide variety of CAF subpopulations with diverse functions exist in cancers. In cutaneous cancers, including cSCC, CAFs constitute a major part of the tumor microenvironment and are divided into myofibroblastic and inflammatory subpopulations. Whether specific CAF subsets associate with metastasis or prognosis in cSCC, remains to be clarified.

Artificial intelligence, especially machine learning, powers many aspects of modern, digitalized society and more advanced deep learning algorithms are increasingly used in medical field to revolutionize our perceptions of what can be achieved. Deep learning is effective in performing visual tasks and thus most widely used within medical specialties dealing with visual data such as pathology and dermatology. Digitalization of pathology and especially whole slide imaging enables more efficient incorporation of artificial intelligence algorithms into the clinical workflow.

This thesis consists of three original clinical and translational studies that unveil the epidemiology, prognosis and risk factors for metastatic cSCC (I), elucidate the role of CAFs during the progression of cSCC (II) and demonstrate the feasibility of deep learning algorithm in the detection of metastatic cSCCs (III).

2 Review of the Literature

2.1 Cutaneous squamous cell carcinoma (cSCC)

2.1.1 Clinical characteristics and precursor lesions

Most cutaneous squamous cell carcinomas (cSCCs) arise in the sun-damaged skin of the elderly white individuals of European ancestry, in the background of pre-existing lesions of actinic keratosis (AK) also known as solar keratosis (Elder, 2008). Clinical presentation of cSCC varies greatly from exophytic tumors, indurated nodules and papules to plaques with smooth, scaly, verrucous or ulcerative surface (Adams et al., 2014). cSCC can be asymptomatic, pruritic or tender and if perineural invasion (PNI) is present local neuropathic symptoms including numbness, burning, paresthesia and paralysis can be present (Adams et al., 2014). Due to the protean clinical phenotype, the diagnosis relies predominantly on histopathological analysis of tissue specimen (Alam et al., 2018). Anatomically, sun-exposed areas including head and neck and the backs of the arms and hands represent most common sites for cSCC, although cSCC can develop on any skin surface (Alam & Ratner, 2001).
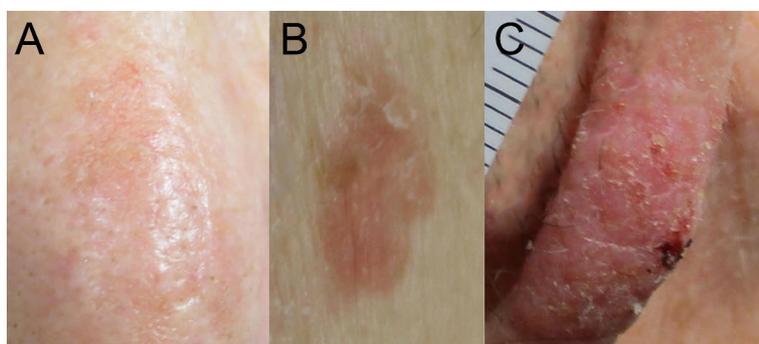
Figure 1. Clinical pictures of actinic keratosis (A), *in situ* cSCC (B) and invasive cSCC (C) demonstrate the difficulty in clinical diagnosis. Modified from Riihilä et al., 2021a.

AKs and cutaneous squamous cell carcinomas *in situ* (cSCCISs) i.e. Bowen's diseases are indicators for increased risk of keratinocyte cancer and considered precursor lesions of cSCC in most instances (Sober & Burstein, 1995). Representative images of AK, cSCCIS and cSCC are shown in Figure 1. Frequently patients present cSCC in association with numerous precursor lesions (Parekh & Seykora, 2017). AKs, cSCCISs and cSCCs usually develop on skin with broad intraepithelial ultraviolet (UV)-induced subclinical damage known as field cancerization and this damage can be seen as the earliest precursor (Fernandez Figueras, 2017). It has been estimated that 60-65% of cSCCs and cSCCISs evolve through AK (Criscione et al., 2009; Marks et al., 1988). AKs and SCCISs typically present as flesh-colored, pink, brown or pigmented scaly patches, papules or plaques on an erythematous base (Parekh & Seykora, 2017). In these precursor lesions, by definition, no dermal invasion is present (Parekh & Seykora, 2017).

The earliest changes in every AK are noted at the basal layer of the interfollicular epidermis and include tumor suppressor protein p53 (TP53) inactivating UVB signature mutations (Fernandez Figueras, 2017). Microscopically there are three main characteristics for AKs and cSCCISs, of which the presence of keratinocytic atypia with nuclear pleomorphism, hyperchromasia and high nucleo-cytoplasmic ratio is the paramount feature followed by presence of compact, orthokeratotic or parakeratotic thick horny layer and presence of actinic elastosis i.e. solar elastosis (Hu et al., 2012). The initial alterations known as basal atypia often found both in the case of subclinical field cancerization and AKs include the presence of hyperchromatic and pleomorphic nuclei with the alteration of the nuclear-cytoplasmic ratio as well as loss of polarity (Fernandez Figueras, 2017). Actinic elastosis is also met in both cases, but the presence of hyperkeratosis and lymphocytic infiltrate differentiate AK from subclinically damaged skin (Fernandez Figueras, 2017). The cornification of the skin represents the earliest feature of AK in the cancerized field and explains why early AKs can clinically be more likely felt than seen (Fernandez Figueras, 2017). AKs can be divided into three grades based on the degree of dysplasia (mild, moderate and severe) (Majores & Bierhoff, 2015). Subpopulation of AKs and cSCCISs progress into invasive cSCC, although the exact risk of progression is ambiguous with estimates of AK to invasive cSCC from less than 0.53% in general to 0.60% at one year and 2.6% at 4 years in a prospective study and 3-16% regarding progression of cSCCIS to invasive cSCC (Criscione et al., 2009; Eimpunth et al., 2017; Fernandez Figueras, 2017; Morton et al., 2014). To further illustrate the obscurity of malignant transformation, rates between 0.025% and 20% per year regarding individual AKs have been reported (Quaedvlieg et al., 2006). However, it seems more likely for AK to spontaneously regress than to progress into invasive cSCC with estimates of regression of one third of AKs within one year, although relapses take place

(Marks et al., 1988, Werner et al., 2013). There are no means to identify the progressible AKs, even though rapid enlargement, diameter over 10mm, induration, ulceration, inflammation and erythema have been postulated to indicate increased risk of progression (Quaedvlieg et al., 2006).
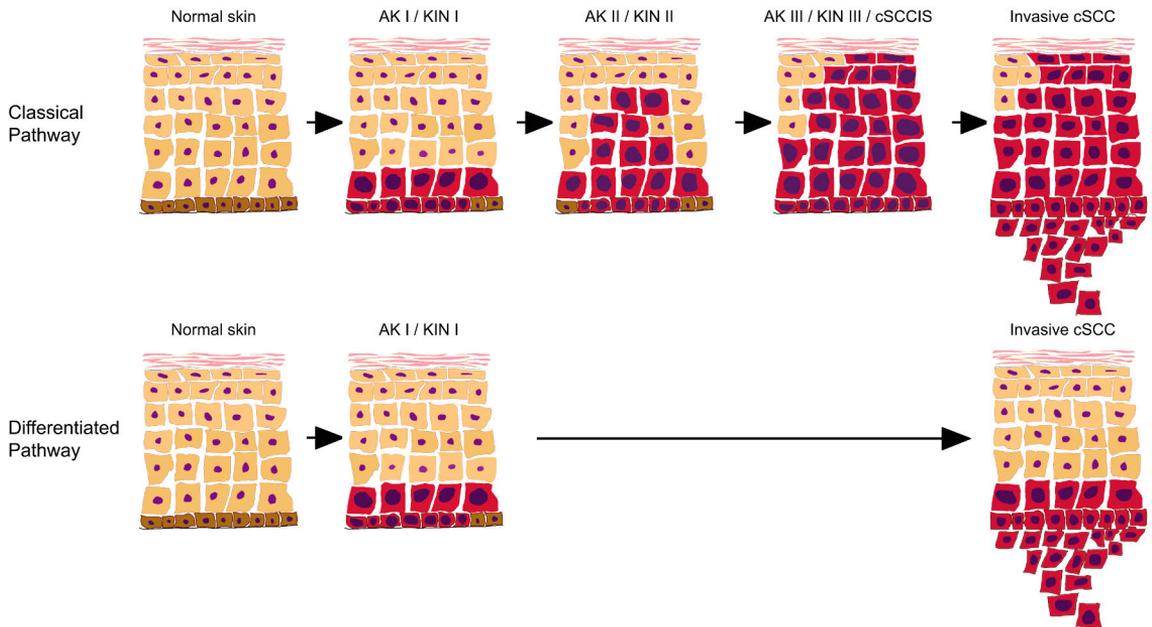
Figure 2. Classical and differentiated pathway of actinic keratosis (AK) to invasive cSCC progression. Modified from Fernandez Figueras, 2017. KIN = keratinocytic intraepidermal neoplasia.

Classically cSCCIS i.e. Bowen's disease has been represented as the intermediate step in the progression of AK to invasive cSCC (Parekh & Seykora, 2017). Histological hallmarks of cSCCIS include laterally sharply demarcated layering disorder of the epidermis with complete loss of architecture, epithelial dysplasia in all levels of epidermis and acanthotic widening (Majores & Bierhoff, 2015). Some authors suggest considering already AKs as *in situ* neoplasms since they derive from clonal DNA modifications in keratinocytes (Heaphy & Ackerman, 2000). Actually, keratinocytic intraepidermal neoplasia (KIN) has been proposed as AK replacing nomenclature due to the ambiguous nature of defining AKs and cSCCISs (Cockerell, 2000). Via the classical or usual pathway (Figure 2), progressive transformation is postulated to occur with progressive stages of KIN from AK or KIN lesions with atypical keratinocytes confined to the lower third of epidermis (KIN I/AK I) through lesions with atypical keratinocytes within two lower thirds of epidermis (KIN II/AK II) to lesions with full thickness epidermal

neoplasia (KIN III/AK III) (Cockerell & Wharton, 2005; Fu & Cockerell, 2003; Yantsos et al., 1999). The differentiated pathway (Figure 2) has been introduced regarding not only other SCCs, but also cSCC, in which it was noted that there is predominance of differentiated pathway in the origin of invasive cSCC with KIN I/AK I being in the highest risk of progression and underlining the importance of treating every AK not only the ones with full epidermal neoplasia and often most suspicious appearance (Arsenic & Kurrer, 2013; Fernández-Figueras et al., 2015; McCluggage, 2013). Notably, regarding the histology, there are distinct differences between KIN III/AK III and Bowen's disease, although the differential diagnosis may occasionally be challenging and some authors consider KIN III AKs synonymous with cSCCISs (Fu & Cockerell, 2003; Majores & Bierhoff, 2015). Perhaps more appropriately AKs and cSCCISs represent distinct variants of intraepithelial neoplasias or SCCs with different clinicohistopathological features and pathogeneses (Fernández-Figueras et al., 2015).

2.1.2   Epidemiology

With an incidence of 30 to 80 per 100000 person years (PY) based on the most comprehensive European nationwide registry and up to 499 per 100000 PY in Australian population, it has been stated that the incidence of cSCC ranges from 5 to 499 per 100000 PY depending on the latitude with up to 40-fold differences between most extreme locations among Caucasian populations (Andersson et al., 2011; Brewster et al., 2007; Green & Olsen, 2017; Nguyen et al., 2014; Staples et al., 2006; Venables et al., 2018). cSCC is commonly regarded as the second most common cancer of the skin after basal cell carcinoma (BCC) and the most common with genuine metastatic potential (Nehal & Bichakjian, 2018). Further, it has even been suggested that the incidence of cSCC would approach and even equate that of BCC (Rogers et al., 2015). In addition, the incidence of cSCC and BCC combined far exceeds the number of all other carcinomas combined (Nehal & Bichakjian, 2018). It has been estimated that lifetime risk of developing cSCC is up to 14-20% in non-Hispanic white population in the United States (Miller & Weinstock, 1994; Stern, 2010). In Finland the incidence was 30 per 100000 PY with 1916 new cSCCs diagnosed in 2020 (Finnish Cancer Registry, 2022).

The incidence of cSCC has been estimated to have increased by 50% to 200% over the past three decades (Waldman & Schmults, 2019). In Finland the incidence has almost doubled during the last decade and nearly quadrupled during past three decades (Finnish Cancer Registry, 2022). The increase is expected to continue in the future due to the aging of the population (Muzic et al., 2017). cSCC is emphasized in the elderly as median age at onset for first primary cSCC was 78

and 80 in British population for males and females correspondingly (Venables et al., 2019a).

2.1.3 Diagnosis and histopathology

cSCC is a keratinocyte carcinoma and comformably to its name originates from keratinocytes that represent the dominant cell type of epidermis (Parekh & Seykora, 2017). cSCC can arise as the result of tumor progression in the sun-damaged skin or occur de novo (Parekh & Seykora, 2017). Characteristically there is invasion of the epidermis by neoplastic squamous epithelial cells (Parekh & Seykora, 2017). Invasive component is formed of infiltrating single cells, sheets or cords, or presents as well-circumscribed nodules, squamous islands or cystic structures composed of atypical keratinocytes (Cassarino et al., 2006b; Quaedvlieg et al., 2006b). In contrast to premalignant lesions, AKs and cSCCISs, the cytomorphology of atypical keratinocytes can vary from very banal to highly anaplastic appearance (Cassarino et al., 2006b).

Diagnosis of cSCC is based on histopathological examination of tissue specimen (Alam et al., 2018). If diagnostic resection is not clinically appropriate expedient, a skin biopsy deep enough should be taken (Alam et al., 2018). It should be, especially depth-wise, extensive enough to allow pathologist to editorialize depth of invasion, presence of perineural and lymphovascular invasion, and histopathological differentiation (Alam et al., 2018).

cSCCs can be histologically divided into three grades of differentiation: good, moderate and poor, based on the degree of keratinization, nuclear atypia and degree of architectural atypia (Parekh & Seykora, 2017). Well-differentiated cSCC is characterized by slightly enlarged keratinocytes with abundant, glassy-pink to eosinophilic cytoplasm and generally visible intercellular bridges (Cassarino et al., 2006a). Well-differentiated cSCCs tend to be well-circumscribed with pushing margins and lobulated appearance, keratinization is usually present and morphologically manifested by "keratin pearls" consisting of central plugs of keratinization within a nest of well-differentiated keratinocytes (Cassarino et al., 2006a). Parakeratosis or retention of keratinocyte nuclei within these keratin pearls can assist in discrimination between well-differentiated cSCC and benign squamoproliferative lesion especially in case with scarce tissue specimen (Cassarino et al., 2006a). In poorly differentiated cSCCs it is difficult to determine a keratinocyte origin as these are characterized by highly infiltrative pattern and composed of very atypical keratinocytes with pleomorphic, hyperchromatic nuclei, numerous atypical mitotic figures and little or no keratinization (Cassarino et al., 2006a). Moderately differentiated cSCCs share characteristics from both well and poorly differentiated cSCCs (Cassarino et al., 2006a).

2.1.4   Histologic variants

Histopathologically, in addition to common cSCC, there are several other cSCC variants with varying prognostic and diagnostic divergence (Parekh & Seykora, 2017; Waldman & Schmults, 2019). These variants include *pigmented*, *signet ring cell*, *clear cell*, *spindle cell*, *desmoplastic* and *acantholytic squamous cell carcinomas*, *adenosquamous* and *verrucous carcinomas* as well as highly disputed *keratoacanthoma* (KA) (Parekh & Seykora, 2017).

KAs are well-differentiated, rapidly growing, solitary squamous proliferations with crateriform appearance (Kwiek & Schwartz, 2016; Parekh & Seykora, 2017; Waldman & Schmults, 2019). These lesions undergo spontaneous resolution which explains the controversy regarding the nature of KAs (Kwiek & Schwartz, 2016). Whether these are actual cSCCs with potential to regress, benign squamoproliferative lesions or represent continuum between benign and malignant proliferation remain debatable, however comprehensive clinical diagnosis is unrealistic and resection not only to exclude cSCC preferable (Hodak et al., 1993; Kwiek & Schwartz, 2016). Furthermore, differential diagnosis to cSCC can be challenging for pathologist especially without the knowledge of rapid growth (Kwiek & Schwartz, 2016). Despite its ambiguous nature, KA should be approached as well differentiated cSCC and treated accordingly (Hernberg et al., 2020; Stratigos et al., 2020a) Other low- to moderate-risk histologic variants of cSCC include *verrucous carcinomas* and *clear cell carcinomas* (Waldman & Schmults, 2019). *Verrucous carcinomas* are well-differentiated, highly distinctive cSCCs with prominent hyperkeratosis and "club-like tongues" of intradermal growth (Schwartz, 1995; Waldman & Schmults, 2019). *Clear cell SCC* or *hydropic SCC* also known as *pale cell SCC* encompasses at least 25% squamous epithelial cells with cytoplasmic vacuolation and can be further divided into three categories comprehending keratinizing, non-keratinizing and pleomorphic tumors (Kuo, 1980).

High risk histologic variants include *acantholytic*, *spindle cell*, *adenosquamous*, *desmoplastic* and *carcinosarcomatous carcinomas* (Parekh & Seykora, 2017). *Acantholytic SCC* is well-differentiated cSCC with acantholysis (Ogawa et al, 2017; Waldman & Schmults, 2019). Histologically acantholysis results in various morphologic patterns and lesional cells show variable degree of desmosomal disruption resulting in rounded cells with centrally placed round nuclei (Ogawa et al., 2017). *Spindle cell* or *sarcomatoid SCCs* are poorly differentiated cSCCs with characteristic closely packed fascicles of pleomorphic spindle-shaped cells in the dermis and high mitotic activity (Evans & Smith, 1980; Silvis et al., 1988). No significant stromal desmoplasia is met and infiltration is often deep extending beyond subcutaneous fat (Evans & Smith, 1980; Silvis et al., 1988). Rare *adenosquamous carcinoma* is characterized by mixed squamous and

true glandular differentiation originating from epidermis with interconnecting nests and anaplastic squamoid cells and desmoplastic stroma with 5% to 80% glandular differentiation (Banks & Cooper, 1991; Waldman & Schmults, 2019). *Desmoplastic SCC* is characterized by infiltrating cords of spindled-squamoud tumor cells surrounded by densely collagenous ergo desmoplastic stroma with at least 30% of tumor volume (Breuninger et al., 1997; Petter & Haustein, 2000). Keratin pearls are usually present and PNI is frequent (Breuninger et al., 1997; Petter & Haustein, 2000). Histogenetically poorly understood *carcinomasarcomatous* or *metaplastic carcinoma* present both epithelial and mesenchymal differentiation (Pazzini et al., 2018).

Uncommon histologic variants include *cSCCs with sarcomatoid differentiation*, *signet ring cell SCC*, *lymphoepithelioma-like carcinoma*, *pseudovascular cSCC* and *cSCC with osteoclast-like giant cells* (Waldman & Schmults, 2019). Regarding uncommon variants and occasionally poorly differentiated common cSCCs immunohistochemical stainings may be needed, although in majority of cSCC cases diagnosis-wise unnecessary (Waldman & Schmults, 2019).

2.1.5   Genomic and molecular landscape of cSCC

On molecular level, it has been established that the genome of cSCC is very heterogeneous and complex (Li et al., 2015; South et al., 2014). The identification of driver mutations and the feasibility of individual gene mutation or copy number variation as biomarker or therapeutic target is hampered by remarkably high background mutation rate in cSCC, which can be 5 to 15 times greater than what is found for non-cutaneous tumors and 4 times higher than the rate in melanoma (Durinck et al., 2011; Inman et al., 2018; Pickering et al., 2014; Tate et al., 2019). Thus, cSCC represents one of the cancers with highest mutation rates with specific mutation signature (Pickering et al., 2014). Closest to cSCC regarding genetics seems to be head and neck squamous cell carcinoma (HNSCC) (Pickering et al., 2014). The large mutational burden of cSCC is due to the life-time exposure to UV-radiation (Riihilä et al., 2019).

Signature 7, that reflects cytosine (C) to thymine (T) transitions of dipyridimines and CC to TT changes, is the classic mutation profile that arises from UV radiation (Alexandrov et al., 2013; Brash, 2015; Chang & Shain, 2021). Majority of mutations found in cSCC contain this UV signature (Inman et al., 2018). Five distinct subtypes of cSCC have been revealed based on DNA sequencings (Chang & Shain, 2021). In patients with rare hereditary disorder, xeroderma pigmentosum, in which patients are extremely sensitive to UV radiation, the mutational burden is high and characterized by high frequency of C to T transitions, but the proportion of *signature 7* mutations is not high (Chang &

Shain, 2021). In cSCCs on patients with recessive dystrophic epidermolysis bullosa (RDEB) the mutational burden is relatively low and characterized by apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like (APOBEC) - mediated mutagenesis (Chang & Shain, 2021). Sporadic cSCCs in patients without known comorbidities comprehend high mutational burden attributable mostly to UV radiation overlapping with mutations shown to be enriched already in sun-exposed normal skin (Chang & Shain, 2021). cSCCs on patients treated with azathioprine have high mutational burden with high proportion of *signature 32* mutations which comprehend predominantly C to T mutations in combination with C to adenine (A), T to A and T to C mutations (Chang & Shain, 2021; Inman et al., 2018). Usage of other immunosuppressive agents is associated with lower mutational burden primarily attributable to UV radiation (Chang & Shain, 2021).

Crucial early event in cSCC progression is the inactivation of *TP53* that further induces the accumulation of UV-induced mutations (Durinck et al., 2011). To date, numerous other driver gene mutations either in tumor suppressing or promoting genes have been identified (Cho et al., 2018; Inman et al., 2018; Li et al., 2015; Pickering et al., 2014). Based on the most comprehensive meta-analysis of exome-sequencing data, 30 cancer genes perturbed in cSCC were nominated and are listed in Table I (Chang & Shain, 2021). These include cyclin-dependent kinase inhibitor 2A (*CDKN2A*) mutations, *Ras* mutations and mutations of *Notch* homologs, which are involved in cSCC carcinogenesis (Chang & Shain, 2021; Pickering et al., 2014; South et al., 2014; Wikonkal & Brash, 1999). However, the prognostic values of suggested driver candidates are ambiguous (Pickering et al., 2014). It has also been stated based on genomic sequencing of metastatic cSCCs (mcSCCs), that epigenetic dysregulation may be a recurrent oncogenic mechanism (Li et al., 2015). Further, while it has been discovered for instance that epidermal growth factor receptor (*EGFR*) is overexpressed persistently in cSCC, *EGFR* activating mutations common in other types of cancers are uncommon in cSCC (Toll et al., 2010; Zandi et al., 2007). *CDKN2A*, instead of inactivating mutations, is frequently inactivated epigenetically in cSCC (Brown et al., 2004). As a proof of the challenges considering the genetic approach in cSCC, a sequencing of 74 cancer genes in biopsies from normal skin was performed with substantially high level of somatic mutations in many driver candidates, demonstrating great tolerance of cancer-causing mutations in normal skin (Martincorena et al., 2015). As long term exposure to UV-radiation causes mutations already in normal skin, alterations in the tumor microenvironment (TME) too are required for precursor lesions to advance into invasive cSCCs and further to metastatic stage (Martincorena et al., 2015; Nissinen et al., 2016). These changes involve extracellular matrix (ECM) components especially collagens and tie cancer-associated fibroblasts (CAFs) to the topic (Karppinen et al., 2016; Nissinen et al., 2016).

Table I. The landscape of driver mutations in cSCC. The percentages of cSCCs harboring pathogenic alterations of the gene are indicated in the column right. Modified from Chang & Shain, 2021.

Pathway	Gene	Percentage
P53	TP53	66,3%
	USP28	12,0%
	MDM2	2,4%
Notch	NOTCH1	55,4%
	NOTCH2	36,1%
	EP300	21,7%
	CREBBP	15,7%
Rb	CDKN2A	34,9%
	CCND1	6,0%
Chromatian Remodel	ARID2	27,7%
	PBRM1	12%
	EZH2	2,4%
Hippo	FAT1	30,1%
	AJUBA	7,2%
	YAP1	1,2%
Cellular stress	CASP8	22,9%
	CHUK	10,8%
	NFE2L2	1,2%
Ras MAPK/PI3K	HRAS	9,6%
	RAP1B	2,4%
	RAC1	2,4%
	KRAS	1,2%
	RRAS2	1,2%
	GNA11	1,2%
	GNAS	1,2%
	ERBB3	1,2%
	PIK3CA	6,0%
	PTEN	6,0%
	MTOR	1,2%

Regarding the inflammation and immunology in cancer, it has been demonstrated that pro-inflammatory cytokines, in addition to possessing anticancer effects, are pivotal for the development, progression and metastasis of cancer (Bai et al., 2007; Esquivel-Velázquez et al., 2015). For instance, interleukin 6 (IL-6) has been shown to promote metastasis and interferon gamma (IFN-γ) to induce the expression of C3, CFH and CFI as well as inhibitory immune checkpoint activating PD-L1, a ligand for programmed cell death protein 1 (PD-1) (Bai et al., 2007;

Liang et al., 2003; Riihilä et al., 2014; Riihilä et al., 2015; Riihilä et al., 2017). PD-1 is a cell surface receptor on activated T cells (Driessens et al., 2009; Thompson et al., 2007). When activated by PD-L1, PD-1 suppresses inflammatory activity of T cells resulting in inhibition of immune system and creation of immune evasion (Dong et al., 2002; Driessens et al., 2009; Francisco et al., 2010). PD-L1 is expressed on various tumors with high mutational burden and correlates with nodal metastasis in cSCC (Dong et al., 2002; García-Pedrero et al., 2017; Zang & Allison, 2007). Regarding complement system and cSCC, overexpressions of complement component 3 (C3) and complement factor B (CFB) as well as complement inhibitors CFI, CFH and factor H-like protein 1 (FHL-1) have been found in cSCC (Riihilä et al., 2014; Riihilä et al., 2015; Riihilä et al., 2017). Furthermore, it has been demonstrated by immunohistochemistry that C3, CFB, CFI and CFH/FHL-1 are *in vivo* specifically expressed by tumor cells with stronger intensity in invasive cSCCs than its precursors and that knockdown of CFI, CFH and FHL-1 as well as C3 and CFB results in diminished proliferation and migration of cSCC cells (Riihilä et al., 2014; Riihilä et al., 2015; Riihilä et al., 2017).

Clinically established panel of biomarkers for the metastasis risk assessment of cSCC is still unavailable, nevertheless there are several proposed molecular markers in addition to above mentioned ones for the progression of cSCC (Kivisaari & Kähäri, 2013). Such proposed biomarkers include matrix metalloproteinases 7 (MMP7) and 13 (MMP13) (Kivisaari et al., 2008; Kivisaari et al., 2010), serine peptidase inhibitor A1 (Farshchian et al., 2011), pS6 (Khandelwal et al., 2016), CD133 (Xu et al., 2016a), Wnt1 and SFRP1 (Halifu et a.l, 2016), but more research is needed in order to evaluate their clinical usability.

2.1.6   Risk factors for the development of cSCC

Due to the cumulative nature of solar UV radiation-related damage, which is the most significant etiologic agent for the development of cSCC, lesions tend to develop to the sun-exposed skin, most frequently to the head and neck region of fair-skinned elderly (English et al., 1998; Fears & Scotto, 1983; Johnson et al., 1992 Liang et al., 1999). The impact of UV exposure is on its part demonstrated by doubling of cSCC incidence with 8º to 10º in latitude (Waldman & Schmults, 2019). UV radiation is followed by male sex, immunosuppression, fair skin and advanced age as risk factors for the development of cSCC (Alam & Ratner, 2001; Lindelöf et al., 2000; Manyam et al., 2017; Venables et al., 2019a). Regarding immunosuppression, in solid organ transplant recipients (SOTRs) the incidence of cSCC is 65- to 250-fold higher than in general population, however the risk varies depending on the number and nature of immunosuppressive agents as well as the amount of UV exposure and skin cancers prior to immunosuppression, but

supposedly 50% of Caucasian SOTRs develop cutaneous neoplasm post-transplantationally (Cheng et al., 2018; Euvrard et al., 2003; Garrett et al., 2017; Manyam et al., 2017; Moloney et al., 2006; Waldman & Schmults, 2019). Azathioprine is an immunosuppressive agent and additionally a potent mutagen that increases the risk of cSCC more than other immunosuppressive agents (Jiyad et al. 2016; O'Donovan et al., 2005). Male sex is another known risk factor as approximately two thirds of patients with cSCC are males (Karia et al., 2012; Venables et al., 2019a).

More uncommon risk factors include environmental exposures and oncogenic human papillomavirus types 16 and 18, which are especially associated with periungual and anogenital cSCC (Faust et al., 2016; Waldman & Schmults, 2019). Environmental exposures include arsenic, polycyclic aromatic hydrocarbons, nitrosamines and alkylating agents (Waldman & Schmults, 2019). Several rare familial syndromes including xeroderma pigmentosum, albinism, epidermolysis bullosa, epidermolysis verruciformis, Ferguson Smith epithelioma, Rothmund-Thomson syndrome and Bloom syndrome predispose individual to multiple cSCCs at young age (Jaju et al., 2016). RDEB is a heritable skin blistering disease, in which mutations in gene encoding type VII collagen lead to epidermal fragility, blistering, chronic hard-to-heal ulcers, fibrosis and molecularly indistinct and extremely aggressive cSCCs (RDEBSCCs) with devastating 5-year survival rate of nearly 0% after the initial diagnosis of first tumor (Fine et al., 2009; Tartaglia et al., 2021). Also patients with for instance chronic lymphocytic leukaemia (CLL) are at 8- to 10- and patients exposed to vismodegib at 8- fold risk for developing cSCC (Waldman & Schmults, 2019).

Risk factors include also smoking, tanning bed usage, long-term UV phototherapy, chronic ulcers, chronic inflammatory cutaneous diseases such as hidradenitis suppurativa, lichen planus and lichen sclerosus et atrophicus, certain pharmacotherapies of cancer such as BRAF inhibitors in addition to above-mentioned vismodegib, certain malignancies such as non-Hodgkin lymphoma and above-mentioned CLL, human immunodeficiency virus (HIV) infection and some familial melanocortin-1 receptor gene variants (Brewer et al., 2015; Czarnecki, 2017; Jensen et al., 2008; Kang & Toland, 2016; Levi et al., 1996; Lindelöf et al., 1999; Manyam 2017; Muzic et al., 2017; Racanelli et al., 2021; Schmults et al., 2013; Velez et al., 2014; Zhang et al., 2012).

The presence of precursor lesions, AKs and/or cSCCISs, is a predictor for the development of invasive cSCC (Criscione et al., 2009; Green & Olsen, 2017; Marks et al., 1988; Werner et al., 2013). In case of field cancerization, manifesting clinically as skin with severe actinic damage, the risk of cSCC development is increased as much as 18-fold (Jiyad et al., 2016). Majority of patients with cSCC present numerous concurrent precursor lesions but on the other hand few precursor

lesions actually develop into invasive cSCC (Jiyad et al., 2016). Both the risk of developing another primary cSCC or nodal metastasis increases significantly with the number of current and prior cSCCs (Levine et al., 2015).

Certain medications can alter the risk for the development of cSCC by their photosensitizing, immunosuppressive or other properties (Pedersen et al., 2018). Medications other than immunosuppressants or cancer therapeutics with most distinct evidence of increased risk for the development of cSCC include hydrochlorothiazides (Pedersen et al., 2018). On the other hand, it has been postulated that non-steroidal anti-inflammatory drugs (NSAIDs) would act in chemoprotective manner, however in another study only weak inverse associations between infrequent aspirin use and cSCC development were discovered (Muranushi et al., 2015; Pandeya et al., 2019).

2.1.7   Rate of metastasis

The reported rate of metastasis of cSCC has varied greatly in publications ranging from 0.1% up to 20.7% probably due to differences in patient and tumor selection, study center, geographical location and patient ethnicity to name but a few (Moore et al., 2005). In a prospective study, a rate of metastasis of 4% was discovered and in a nation-wide British study metastatic rate of 1.1% in women and 2.4% in men was reported during the follow-up of 36 months at most (Brantsch et al., 2008; Venables et al., 2019a). For geographical variety, metastatic rates of 3.0% and 3.7% in US population and 1.9-2.6% in population of New Zealand have been reported (Brougham et al., 2012; Karia et al., 2013; Schmults et al., 2013). Further, in nation-wide Dutch study a metastatic rate of 1.9% was noted with higher rate in men (2.3%) than women (1.4%) (Tokez et al., 2022). Based on the notions, the overall rate of metastasis can be stated to stand approximately 1-4% with significantly higher rates in subpopulations such as and especially SOTRs and others with immunosuppression (Karia et al., 2013; Mourouzis et al., 2009; Nehal & Bichakjian, 2018; Nelson & Ashton, 2017; Rogers et al., 2015; Schmults et al., 2013; Tokez et al., 2022; Venables et al., 2019a).

Metastasis occurs relatively shortly after the diagnosis of primary mcSCC as it has been shown in several studies that approximately 72-90 % of metastases are detected within the first two years (Bobin et al., 2018; Cherpelis et al., 2002; Dinehart & Pollack, 1989; Rowe et al., 1992; Venables et al., 2019a). In the largest studies, detection rate of 85.2% within 2 years and median time to metastasis of 1.5 years have been reported (Tokez et al., 2022; Venables et al., 2019a). Additionally, in a prospective study a rate of 73.1 % within one year has been reported (Brantsch et al., 2008).

2.1.8 Factors associated with the risk of metastasis

First and foremost, it must be stated that the evaluation of features associated with the metastasis risk is hampered by the lack of prospective, nation-wide and all-encompassing studies. As means to identify cSCCs with elevated risk of metastasis are limited, unfortunately in clinical practice local recurrence is often the first identified indicator of the aggressive biologic behavior of the tumor (Levine et al., 2015). In current risk stratification, based on both the eight edition of American Joint Committee on Cancer (AJCC8) and Brigham and Women's Hospital (BWH) staging systems, the primary tumor diameter, invasion depth and presence of PNI are pivotal in primary tumor staging (Roscher et al., 2018). It has been established that tumor diameter greater than 20mm correlates most evidently with disease-specific death, doubles the risk of local recurrence and triples the rate of metastasis (Que et al., 2017; Thompson et al., 2016). Invasion beyond subcutaneous fat of primary tumor was another feature that qualified as risk factor when relative risk of at least 5 and heterogeneity of less than 60% were applied to the data of meta-analysis comprising 23000 tumors (Baum et al., 2018; Thompson et al., 2016). In fact, the risk for nodal metastasis was higher in case of invasion beyond subcutaneous fat, Breslow thickness more than 2mm or 6mm than in case of diameter greater than 20mm and it has been estimated that tumors extending beyond subcutaneous fat possess 11-fold risk of metastasis (Baum et al., 2018; Que et al., 2017). On the other hand, it has been stated that 16% of cSCCs with over 6mm invasion and 30% of tumors with over 20mm diameter metastasize (Alam & Ratner, 2001; Brantsch et al., 2008) PNI involving large calibre nerves increases the risk of nodal metastasis as well as disease specific death (Que et al., 2017). Based on the above-mentioned meta-analysis, invasion depth applied as invasion beyond fat or Breslow thickness followed by tumor diameter, poor differentiation and PNI in this order were the risk factors for metastasis with highest risk ratios (RRs) (Thompson et al., 2016).

Other proposed risk factors associated with local recurrence, metastasis and/or disease specific death include age, lymphovascular invasion, poor histologic differentiation, certain histologic subtypes, tumor budding, histologically positive excision margins, ionizing radiation, prior local recurrence, increasing number of concurrent or prior cSCCs, immunosuppression (especially SOTRs), comorbidities such as CLL and non-Hodgkin's lymphoma as well as certain locations and location properties such as development at the site of chronic inflammation or ulcer (Brewer et al., 2015; Cherpelis et al., 2002; Karayannopoulou et al., 2016; Levine et al., 2015; Manyam et al., 2017; Parekh & Seykora, 2017; Que et al., 2017; Skulsky et al., 2017; Thompson et al., 2016). Notable is also that cSCC arising in location not conventionally exposed to sunlight such as soles of the feet or perineum have proportionally higher rate of metastasis (Motley et al., 2002). It has

been addressed that poor differentiation doubles the risk of metastasis, and in above-mentioned meta-analysis poor differentiation was the most influential metastasis predicting factor after invasion depth and diameter (Brantsch et al., 2008; Thompson et al., 2018). It is indicated that prognosis of locally recurrent tumors is significantly worsened with up to 48 % risk of metastasis (Rowe et al, 1992; Cherpelis et al, 2002). Thus, it seems substantiated that not later than after local recurrence, cSCC should be considered high-risk regardless other features. Location-wise, in the above-mentioned meta-analysis primary tumors located on temple, ear, lip or cheek were significantly associated with heightened risk of metastasis, in this order (Thompson et al., 2018). Peritumoral AK and desmoplasia represent more controversial features associated with the risk of metastasis. (Martorell-Calatayud et al., 2013). Rarer risk factors include exposure to ionizing radiation, which is associated with up to 30% rate of metastasis and development as a complication of hidradenitis suppurativa that is associated with up to 50% rate of metastasis (Racanelli et al., 2021; Waldman & Schmults, 2019).

To date, there is controversy regarding most of the high-risk and especially metastasis-associated factors, as findings are principally based on retrospective single-institution studies with small cohorts especially as far as mcSCCs are concerned and with wide variation regarding study design and the reportage of results. For instance, the association between the metastasis risk and immunosuppression is not that straightforward (Genders et al., 2019). It has been indicated that the metastatic rate among SOTRs would be 7-8%, however in the above-mentioned meta-analysis immunosuppression was only 9th most influential feature associated with the risk of metastasis after invasion depth, diameter, differentiation, PNI and location on temple, ear or lip (Burton et al., 2016; Thompson et al., 2016). On the other hand, in recent nation-wide Dutch study organ transplant recipiency was the most influential factor associated with the risk of metastasis followed by age, male sex and hematologic malignancy (Tokez et al., 2022). Furthermore, in another nation-wide study dealing with English population the results were similar with highest adjusted hazard ratio (HR) for immunosuppression followed by age, male sex, location on ear and lip as well as higher level of deprivation (Venables et al., 2019a). However, these nation-wide studies lack the inclusion of primary tumor characteristics such as diameter, invasion depth, histological differentiation and PNI.

On molecular level, increased expression of annexin A5 (ANXA5) and dolichyl-diphosphooligosaccharide-protein glycosyl-transferase noncatalytic subunit (DDOST) was found to be associated with the development of and reduced time to cSCC metastasis as well as survival (Shapanis et al., 2021). Furthermore, a prediction model utilizing these proteins was shown to possess higher sensitivity and specificity than current staging systems (Shapanis et al., 2021). In prognostic

40- gene expression profile (GEP) test, created to aid in risk stratification of cSCC metastasis risk, 34 discriminant genes were included in the prognostic signature that acted as an independent predictor of metastasis risk (Wysong et al., 2021). These genes include *MMP10* and podoplanin (*PDPN*) to name but a few (Wysong et al., 2021).

2.1.9   Staging systems

There are several staging systems utilized in the stratification of cSCC, but the staging system by AJCC is most widely used (Venables et al., 2021). During past decade, staging systems have developed from the 7th edition of AJCC (AJCC7) with increased emphasis on primary tumor characteristics (Farasat et al., 2011; Roscher et al., 2018; Venables et al., 2021). Notably, until the introduction of AJCC7 all non-melanoma skin cancers were grouped together for staging purposes (Farasat et al., 2011). Due to the poor performance of AJCC7, rival staging systems have been developed, of which systems by BWH and Tübingen University (ergo Breuninger) are apparently most widely acknowledged (Breuninger et al., 2012; Jambusaria-Pahlajani et al., 2013). Further, based on the updated and better performing 8th edition of AJCC tumor staging (AJCC8), a Salamanca classification was developed focusing on refining the AJCC8 T3 stage (Conde-Ferreirós et al., 2021).

Unfortunately, there have been few staging systems validating studies without limitation to single academic center data. In a nation-wide data utilizing study comparing AJCC7, AJCC8, BWH and Breuninger staging systems, it was stated that although staging system by Breuninger gave the best results, all the systems were unsatisfactory in identifying non-selected patients at high risk for metastasis (Roscher et al., 2018). Afterwards, another study with nation-wide data compared AJCC8, BWH, Breuninger and Salamanca classifications showing, that although BWH staging possessed the best overall discriminative ability, each system had their strengths (Venables et al, 2021). It was further deduced that BWH T2b/T3 would be the most appropriate in identification of high-risk patients and AJCC8 T1 for determining patients who can safely be omitted from follow-up visits (Venables et al., 2021). In more detail, 24-38% of patients with BWH stage T2b/T3 and 14-17% of patients with AJCC T3/T4 tumors develop metastasis (Jambusaria-Pahlajani et al., 2013; Karia et al., 2014; Karia et al., 2018; Ruiz et al., 2019). In accordance with other studies dealing staging systems, the need for novel patient and tumor characteristics associated with cSCC progression was expressed (Cañueto et al., 2019; Puebla-Tornero et al., 2021; Roscher et al., 2018, Ruiz et al., 2019; Venables et al., 2021). Probably the most advanced proposition relies fundamentally on BWH staging system with some adjustments considering tumor

diameter and invasion and the inclusion of CLL as comorbidity (Baum et al., 2018). In this approach, the classification is more clinician-friendly and usable as it distributes cases into low, intermediate and high risk groups and provides management and follow-up guidelines based on the classification (Baum et al., 2018). However, it lacks to factor in for instance organ transplant recipiency, which on its own could be construed to characterize cSCC as high risk (Baum et al., 2018; Cheng et al., 2018).

As it has been established that current staging systems for cSCC have suboptimal positive predictive value for identifying metastatic primary tumors, a GEP test was developed to improve the risk stratification (Wysong et al., 2021). This prognostic 40-GEP test divided cSCCs into three classes based on metastasis risk and was found to act as staging systems complementing independent predictor of metastasis risk (Wysong et al., 2021). This 40-GEP test achieved higher risk for metastasis in multivariate Cox regression analysis than AJCC8 or BWH staging systems with adjusted hazard ratio (aHR) of 9.55 for class 2B tumors in comparison to 2.68 for AJCC8 T3/T4 tumors and aHR of 8.72 for class 2B in comparison to BWH T2b/T3 tumors (Wysong et al., 2021). Furthermore, the metastatic rate for class 2B tumors was 60% (Wysong et al., 2021).

Above mentioned classifications are summarized in Table II. Notably, AJCC8 is fundamentally restricted to head and neck sites, although in practice AJCC8 has been noted to perform better when applied to all body sites (Venables et al., 2021).

Table II. Staging systems. BWH risk factors = diameter ≥ 2cm, poor histopathologic differentiation, PNI of at least 0.1mm, invasion beyond subcutaneous fat. Modified from Baum et al., 2018; Karia et al., 2014 and Wysong et al., 2021.

AJCC8	Features
T stage	
T1	Tumor diameter < 2cm
T2	Tumor diameter ≥ 2cm and < 4cm in greatest dimension
T3	Tumor diameter ≥ 4cm or minor bone erosion or PNI or deep invasion
T4	Tumor with gross cortical bone/marrow invasion
BWH	
T stage	
T1	No high risk factors
T2a	1 high risk factor
T2b	2-3 high risk factors
T3	≥ 4 high risk factors or bone invasion
Breuninger et al.	
Stage	
Clinical tumor stage (cT)	Low risk: tumor diameter ≤ 2cm
	High risk: tumor diameter > 2cm
Pathological tumor stage (pT)	No risk: tumor thickness ≤ 2mm
	Low risk: tumor thickness > 2mm and ≤ 6mm
	High risk: tumor thickness > 6mm
Co-risk factors	Immunosuppression
	Desmoplastic type
	Poor differentiation
	Ear location
Baum et al.	
Low-risk cSCC	BWH T1
	BWH T2a
Intermediate-risk cSCC	BWH T2a with diameter > 2cm
	BWH T2a with depth beyond subcutaneous fat
	BWH T2a and CLL with Rai stage III or IV
High-risk cSCC	BWH T2b/T3
	BWH T2a with depth beyond subcutaneous fat
40-GEP test	
Class 1 (low risk)	Low risk of metastasis based on 40-gene expression profile
Class 2A (high risk)	High risk of metastasis based on 40-gene expression profile
Class 2B (highest risk)	Highest risk of metastasis based on 40-gene expression profile

2.1.10 Prognosis

In majority of cSCC cases, due to the relatively low metastatic rate, post-surgical prognosis is excellent (Nehal & Bichakjian, 2018). Nevertheless, in metastatic case the prognosis is poor and due to the high and increasing incidence, cSCC accounts for at least 20% of all skin cancer-related mortality with disease-specific mortality of around 1.5-4% in Caucasian and up to 18% in black population among whom

cSCC is the most common skin cancer (Brantsch et al., 2008; Brougham et al., 2012; Czarnecki, 2017; Karia et al., 2013; Levine et al., 2015; Mora & Perniciaro, 1981; Mourouzis et al., 2009; Nehal & Bichakjian, 2018; Rogers et al., 2015; Rowe et al., 1992; Schmults et al., 2013). It has been suspected that mortality from cSCC would even approach mortality in melanoma at least in some UV radiation abundant geographical regions (Karia et al., 2013; Nehal & Bichakjian, 2018). Mortality is associated predominantly with uncontrollable locoregional metastases instead of distant metastases (Schmults et al., 2013). In addition, the risk for death from any cause and the risk for second primary cancer is increased in patients with cSCC (Wehner et al., 2018).

Survival of patients with mcSCC has generally been regarded poor, but there is variety in survival estimates depending on source. Due to the retrospective nature of the studies overall survivals (OSs) have been commonly reported. Two-, 3-, and 5-year survival rates among studies reporting OS from the metastasis detection have varied between 50-66%, 29-46% and 30-50% respectively (Bobin et al., 2018; Givi et al., 2011; Hirshoren et al., 2017; Venables et al., 2019a). Venables et al. (2019a) reported significantly lower 3-year OS for females (29% vs 46%) and interestingly a remarkably low 3-year survival from first primary cSCC of 65 % for males and 68% for females in unselected study material including only 1.1-2.4 % of mcSCCs. Further, a 5-year disease-specific survival (DSS) of 79.1% for patients with mcSCC has been reported (Tokez et al, 2022).

It has been established that first metastasis typically affects regional lymph node and that most common sites of nodal metastases are head and neck nodes or parotid gland (Brantsch et al., 2008; Venables et al., 2019a). In metastatic case extranodal extension (ENE) and the number of positive lymph nodes have been associated with worse prognosis (Smith et al., 2018).

2.1.11 Treatment

Due to the relatively low rate of metastasis, the treatment of cSCC is conventionally management of the tumor locally and majority of cSCC lesions are successfully managed via surgical excision (Alam et al., 2018). As local recurrence elevates the rate of metastasis, intensive enough locoregional control of primary tumor is cardinal (Baum et al., 2018). When dealing with cSCC the evaluation of local recurrence and metastasis predicting tumor- and patient-related characteristics is fundamental (Alam et al., 2018). The main goal of the surgery is the removal of the tumor, however preservation of function and cosmetics are naturally objectives to take into account (Stratigos et al., 2020b). Selection of treatment modality, evaluation of the need for staging examinations or adjuvant therapies as well as the need for follow-up visits should be based on this as encompassing as possible risk

assessment (Alam et al., 2018). In the lack of comprehensive and universally adopted guidelines, the risk assessment is somewhat heterogeneous and organization dependent. Although there has been significant advance lately in creation of evidence-based risk stratification systems such as AJCC-8 and BWH staging systems, the means to identify tumors with elevated risk of metastasis or local recurrence still lacks (Alam et al., 2018). Main therapeutic indications by European interdisclipinary guideline for cSCC are shown in Figure 3 (Stratigos et al., 2020b).
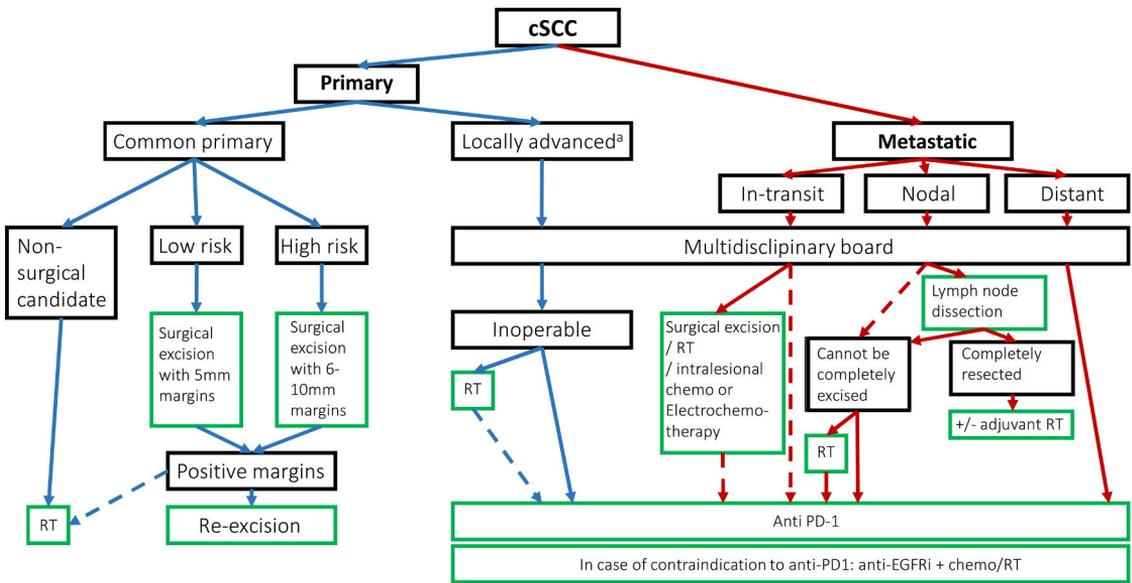
Figure 3. Main therapeutic indications for cSCC. [a] Locally advanced by definition not amenable to curative surgery or curative radiotherapy (RT). Modified from Stratigos et al., 2020b.

When dealing with low-risk primary tumor, conventional surgery with safety margins is the treatment to choose regardless of the anatomic location and age-group (Stratigos et al., 2020b). Excision margins between 4mm and 6mm are consistently proposed by guidelines for tumors without high-risk features and the suggestion by European consensus group is 5mm (Stratigos et al., 2020b). Primary radiation therapy i.e. radiotherapy (RT) is an option if patient is not eligible for surgery (Alam et al., 2018; Stratigos et al., 2020b). There are also indications that curettage and electrodessication, cryosurgery, laser and photodynamic therapy could be used as an alternative for surgical excision in special cases regarding small, low-risk tumors (Stratigos et al., 2020b).

Table III. A list of indicative prognostic high-risk factors for recurrence for cSCC proposed by the European Dermatology Forum (EDF) – the European Association of Dermato-Oncology (EADO) – the European Organisation for Research and Treatment of Cancer (EORTIC). PNI = perineural invasion. Modified from Stratigos et al., 2020a.

Prognostic factors for considering a common primary cSCC as high-risk	
1.	Tumor diameter >20mm
2.	Localization on temple/ear/lip area
3.	Thickness >6mm or invasion beyond subcutaneous fat
4.	Poor grade of differentiation
5.	Desmoplasia
6.	Microscopic, symptomatic or radiological PNI
7.	Bone erosion
8.	Immunosuppression

For high-risk cSCCs, surgical excision with technique offering superior management of margins such as Mohs micrographic surgery is preferred if available as it is associated with lower risk of tumor recurrence (Alam et al., 2018; Motley & Arron, 2019; Stratigos et al., 2020b). The excision margins in conventional surgery should exceed 5mm, however guideline recommendations vary with suggestion by the European consensus group of 6-10mm (Stratigos et al., 2020b). The European consensus group regards features listed in Table III as high-risk factors based on the risk for recurrence (Stratigos et al., 2020a; Stratigos et al., 2020b). Depth-wise, the excision should include the subcutaneous tissue together with galea-aponeurosis in scalp locations but spare tumor-free perichondrium or periosteum (Stratigos et al., 2020b). If curative resection is not feasible such as in case of locally advanced and surgically inoperable or metastatic case, reinforcement of resection with another treatment should be considered by multidisclipinary tumor board (Stratigos et al., 2020b). Most commonly this corresponds to RT that generally is implemented postoperatively (Alam et al., 2018). Adjuvant RT should also be considered in high-risk, prone to local recurrence cases (Carucci et al., 2004). Locally advanced cSCC is defined as non-metastatic cSCC (non-mcSCC), which is not amenable to curative surgery or RT due to multiple recurrences, large extension, bone invasion or deep infiltration beyond subcutaneous fat or along nerves (Stratigos et al., 2020a).

In order to detect subclinical nodal metastases, imaging studies should be considered in patients with high-risk factors (Table III) (Stratigos et al., 2020a). However, the need for staging studies is not well established and specification of the high-risk factors for imaging cannot be given (Stratigos et al., 2020a). As metastases occur predominantly in locoregional lymph nodes, palpation of at least regional lymph nodes including parotid gland should be performed for every

patient (Stratigos et al., 2020a). Ultrasonography, computed tomography scan, positron emission tomography computed scan and magnetic resonance imaging can be used as imaging methods (Stratigos et al., 2020a). Based on limited evidence ultrasonography appear as cost-effective minimally invasive staging modality for lymph node metastases from cSCC (Stratigos et al., 2020a). In case of advanced cSCC consultation in multidisciplinary tumor board for optimal staging studies and treatment is mandatory (Stratigos et al., 2020a). Sentinel lymph node biopsy is currently not recommended in case of cSCC due to the lack of its prognostic impact (Stratigos et al., 2020a).

In case of clinically detected regional node, a fine needle aspiration cytology or ultrasound-guided core biopsy is recommended (Stratigos et al., 2020a). If lymph node involvement is detected clinically or with imaging, therapeutic regional lymph node dissection is the preferred surgical treatment (Stratigos et al., 2020b). The extent of surgical resection should be determined by multidisciplinary tumor board with the aim of radical lymph node dissection of all the affected areas (Stratigos et al., 2020b). Prophylactic lymph node dissections are not recommended, although there are indications that prophylactic neck dissection should be performed in patients with mcSCC to the parotid (Stratigos et al., 2020b). In metastatic case adjuvant chemoradiation seems to offer better recurrence-free survival than radiation therapy alone (Tanvetyanon et al., 2015). Treatment options considering systemic therapies are limited and unestablished, while there has not been any chemotherapy agent, targeted therapy or immune mediator approved by US Food and Drug Administration (FDA) or European Medicines Agency (EMA) until September 2018 when PD-1 blocking monoclonal antibody cemiplimab was approved by FDA and by EMA in 2019 for the treatment of patients with metastatic or locally advanced cSCC who are not candidates for curative surgery or curative radiation (Markham & Duggan, 2018). In phase II study of cemiplimab including patients with both metastatic and locally advanced cSCCs, objective response rate of nigh 50% was obtained (Markham & Duggan, 2018). Traditionally EGFR-inhibitor cetuximab, platinum based chemotherapeutics and fluorouracil in variable combinations have supposedly been the most commonly off-label used drugs, but remission rates and challenging adverse effect profile have left better treatments to be desired (Alam et al., 2018).

In case of cSCC, holistic treatment includes fundamentally also the control of actualization of effective enough protection against further UV radiation, evaluation of the need and the possible execution of chemoprophylaxis considering especially cases with field cancerization as well as the planning and prosecution of risk-based follow-up (Alam et al., 2018).

2.2   Cancer-associated fibroblasts (CAFs)

2.2.1   Fibroblasts in general

Fibroblasts are characterized by their ability to tolerate severe stress usually lethal to other cells and unique survival capability illustrated by successful live-cultures from decaying tissues (Bliss et al., 2012; Kalluri, 2016). Furthermore, fibroblasts are characterized by their contractile as well as ECM remodeling activities and were originally in the 19th century described as collagen synthesizing cells residing in connective tissues (Duvall, 1879; Gabbiani et al., 1971; Virchow, 1858). Fibroblasts in healthy tissues are usually singular, non-epithelial and non-immune cells located in the interstitial space between layers of functional parenchyma or near a capillary without contact to basement membrane but embedded within fibrillary ECM and in general usually quiescent (Bainbridge, 2013; Kalluri, 2016; Tarin & Croft, 1969). Morphologically fibroblasts are thin, elongated cells with front and back extensions exhibiting classic spindle-like or fusiform shape and in indolent or hibernating state manifest minuscule metabolic and transcriptomic activity (Kalluri, 2016; Tarin & Croft, 1969). The ability of quiescent fibroblasts to become active was first described in the process of wound healing followed by acute and chronic inflammation, tissue fibrosis and cancer (Darby et al., 2014; Desmouliere et al., 2003; Micallef et al., 2012).

   In turn, activated fibroblasts, unlike hibernating ones, are highly heterogeneous, morphologically and metabolically active, proliferative, migratory, possess active and dynamic secretomy, epigenetically modified, produce ECM, exhibit various markers not present in hibernating state and act as precursors for chondrocytes, adipocytes, myocytes, endothelial cells and induced pluripotent stem cells (Kalluri, 2016). Activated fibroblasts consequently also coordinate the function of other cell types within the tissue (Van Hove & Hoste, 2021). Due to the ability of activated fibroblasts, under appropriate stimuli, to differentiate into above mentioned mesenchymal lineages, quiescent fibroblast can be considered as tissue-resident mesenchymal stem cells (rMSCs) (Kalluri, 2016). Furthermore, fibroblasts in their activated phenotype can be easily cultured analogously to mesenchymal stem cells (MSCs), which when cultured resemble fibroblasts both morphologically and functionally (Hematti, 2012). Additionally, both activated fibroblasts and MSCs are self-renewing and express similar molecular markers (Kalluri, 2016). Thus, it has been suggested that quiescent mesenchymal cell with ability to become MSC by appropriate stimuli would be more accurate definition of a fibroblast, however whether all activated fibroblasts are MSCs remains to be clarified (Kalluri, 2016). It should also be stated that most properties assigned to fibroblasts are in fact features of activated fibroblasts, myofibroblasts or MSCs (Kalluri,

2016). Myofibroblast in general represents activated fibroblast with contractile properties featuring characteristics of contractile smooth muscle cells and represents key effector cell in fibrosis (Pakshir et al., 2020).

Activation of a fibroblast results in increased activity including ECM synthesis, immune cell recruiting cytokine and chemokine production and physical tissue architecture modifying activities, and this increased activity actually defines fibroblast as activated (Castor et al., 1979; Parsonage et al., 2005; Tomasek et al., 2002). Stroma in most organs contains only a small number of quiescent fibroblasts whereas reactive stroma in tumor or fibrosis presents with increased number of activated, ECM constituents excreting and fibroblast markers expressing activated fibroblasts (Hanahan & Coussens, 2012; Kalluri & Zeisberg, 2006; Marsh et al., 2013; Ronnov-Jessen et al., 1996).

In healthy skin quiescent fibroblasts reside in the dermis and become activated in inflammatory conditions, wound repair and during cancer development (Van Hove & Hoste, 2021). Traditionally cutaneous fibroblasts have been divided into papillary and reticular dermal fibroblasts based on the location and marker expression (Van Hove & Hoste, 2021). Single-cell RNA sequencing (scRNA-seq) of healthy human skin fibroblasts has revealed three distinct fibroblast subtypes (Ascensión et al., 2021).

2.2.2   CAFs and tumor microenvironment

Acknowledging to Paget's "Seed and Soil"-theory in which cancer cells represent the seed and TME the soil, there is increasing evidence of the role of TME or tumor stroma in tumor progression, metastasis, angiogenesis and immune cell modulation (Hanahan & Weinberg, 2011; Joshi et al., 2021; Kalluri, 2016; Paget, 1989). Instead of being a mass comprised of malignant cells, according to current knowledge cancer is more like an organ harboring complex interplay between tumor cells and TME (Joshi et al., 2021). TME consists of ECM which represents major structural component of the TME and cells of which CAFs are pivotal (Chen et al., 2021; Van Hove & Hoste, 2021). TME can be sectioned into four major components including an immune component known as tumor immune microenvironment (TIME), vascular component consisting of vascular and lymphatic endothelial cells as well as pericytes, ECM component comprehending collagens, glycoproteins as well as proteoglycans and stromal component that includes non-immune cells of mesenchymal origin such as CAFs (Eble & Niland, 2019; Gajewski et al., 2013; Kalluri & Zeisberg, 2006; Viallard & Larrivée, 2017). These cells, which were first noted to accelerate growth of epithelial tumors, were later designated as CAFs and are also called tumor-associated fibroblasts (TAFs), activated fibroblasts and activated myofibroblasts (Kalluri, 2016; Louault et al.,

2020). CAFs are by definition activated fibroblasts located within or adjacent to tumor, play role in tumor initiation, progression, invasion, angiogenesis, metastasis as well as therapy resistance and constitute a major component of the TME also in cSCC (Chen et al., 2021; Ramos-Vega et al., 2020; Van Hove & Hoste, 2021). Whether CAFs contribute to lymphoid or hematopoietic cancers remains unclear, but regarding solid tumors there is extensive evidence of involvement across almost every physiologic system, including skin (Bai et al., 2015; Erez et al., 2010; Farmer et al., 2009; Samain & Sanz-Moreno, 2020; Santos et al., 2009; Teng et al., 2016; Vaquero et al., 2020; Vosseler et al., 2009; Wu et al., 2020a; Zhang et al., 2020; Zhu et al., 2020). However, there is significant variability regarding CAF abundance among cancers with pancreatic adenocarcinoma, breast cancer, lung adenocarcinoma and renal clear cell carcinoma representing cancers heavily infiltrated with CAFs whereas leukemia, lymphoma and brain tumors represent cancers mostly devoid of CAFs (Louault et al., 2020). Conclusively, it can be stated that cancer is associated with resilient and plastic fibroblasts at all stages of cancer progression, including metastasis (Kalluri, 2016).

2.2.3   Activation of CAFs

According to one theory, the activation of fibroblasts in cancer context reflects a host defence mechanism that aims to restrain cancer progression and even eliminate the cancer (Dumont et al., 2016; Elenbaas & Weinberg, 2001; Ishii et al., 2016; Kalluri & Zeisberg, 2006). Fibroblasts become activated and proliferative during inflammation, wound repair, fibrosis and malignant progression (Desmouliere et al., 2004; Kalluri & Zeisberg, 2016; Louault et al., 2020). The activation of quiescent fibroblasts is still ambiguous but it has been suggested that there are two activation profiles comprising reversible and irreversible (Zeisberg & Zeisberg, 2013). Key mediators of the activation in acute and chronic tissue damage include transforming growth factor β (TGFβ), platelet derived growth factors (PDGFs) and fibroblast growth factor 2 (FGF2) (Elenbaas & Weinberg, 2001).

In cancer context, the growth factors released by infiltrating immune cells and cancer cells mainly govern the recruitment of stromal fibroblasts, which in many cancers is TGFβ-dependent (Aoyagi, 2004; Löhr et al., 2001). TGFβ1, secreted by stromal and tumor cells is the main factor that activates resident fibroblasts into CAFs (Bhowmick et al., 2004). Further, in TME local CAF proliferation and invasion is stimulated by TGFβ (Kalluri, 2016). Also PDGFs secreted by cancer cells and stromal cells including fibroblasts, activate and induce proliferation of fibroblasts (Bronzent et al., 1987). Other tumor cell-secreted factors also promote the conversion of resident fibroblasts into CAFs including fibroblast growth factor

(FGF), sonic hedgehog (SHH) and IL-1β (Elenbaas & Weinberg, 2001; Erez et al., 2010; Tejada et al., 2006, Tian et al., 2009). Furthermore, hypoxia as well as vitamin A and D promote activation of resident fibroblasts into CAFs (Costa et al., 2014; Liu et al., 2019a; Shany et al., 2019). CAF-secreted TGFβ1 and C-X-C chemokine ligand-12 (CXCL-12) initiate and maintain myofibroblast phenotype and tumor promoting functions in auto- and paracrine fashion (Kojima et al., 2010). In more detail, there are indications that activation by TGFβ leads to generation of alpha smooth muscle actin (αSMA) expressing myofibroblastic CAFs (myCAFs) and FGF2 to generation of inflammatory CAFs (iCAFs) (Van Hove & Hoste, 2021).

2.2.4　Function of CAFs

Functionally, CAFs can be considered as synthetic machines that produce several tumor components, create ECM, reprogram TME both metabolically and immunologically and impact adaptive resistance to cancer treatments (Kalluri, 2016). Originally CAFs were considered homogeneous population of ECM-producing stromal cells responsible for the three dimensional structure of the tumor, but disclosure of their pro- and anti-tumorigenic functions by research have shifted the assumptions into considering CAFs as heterogeneous group of stromal cells with different origins, phenotypes, functions and presences (Louault et al., 2020; Schauer et al., 2011; Öhlund et al., 2014). The functional role of fibroblasts in cancer is not as well-known as in the setting of wound healing as it appears complex and bimodal with cancer-promoting and cancer-restraining functions (De Wever et al., 2014; Dumont et al., 2013; Öhlund et al, 2014). Activated fibroblasts in general produce constituents of the ECM and basement membranes including type I (Col1), III, IV and V collagens, laminins and fibronectin (FN) as well as ECM-degrading proteases such as MMPs that crucially contribute to the ECM turnover (Riihilä et al., 2021; Rodemann & Muller, 1991; Simian et al., 2001; Tomasek et al., 2002). Activated fibroblasts are contractile and in wound healing attributable to contracting the skin (Tarin & Croft, 1969; Tomasek et al., 2002). Furthermore, activated fibroblasts maintain the homeostasis of adjacent epithelia by mesenchymal-epithelial cell interactions and secretion of growth factors (Wiseman & Werb, 2002). Acknowledging the resilient nature, tolerability and survival capabilities of fibroblasts it would not be surprising if CAFs represented the resistant stromal cell type that participates in tumor relapse after chemo- or radiotherapy (Kalluri, 2016). CAFs also directly alter the metabolism of cancer cells by providing energy-rich metabolites such as fatty- and amino acids (Martinez-Outschoorn et al., 2014).

Analogously to tissue fibrosis in which various chronic insults result in chronic wound healing condition described as wound that never heals, in cancer there is genetic insult to the functional parenchyma and these cancerous cells represent tissue injury that results in chronic host repair response known as cancer fibrosis (also known as desmoplastic reaction) (Kalluri, 2016). Although the cancer stroma is generally considered to promote tumor initiation and progression, there is evidence that carcinomas *in situ* (CISs) involve reactive tumor fibrosis, which contingently inhibit invasion and malignant transformation (Dvorak et al., 1984; Folkman & Kalluri, 2004; Ronnov-Jessen et al., 1996). There are theories of mechanisms by which tumor stroma contribute to the progression of CIS to invasive cancer including inflammatory cues that would initiate pro-inflammatory and tumor-promoting functions in fibroblasts and on their part attribute to this progression (Kalluri, 2016). In more detail, it has been shown that IL-1β secretion by immune cells in early lesions may initiate nuclear factor-κB (NF-κB) signaling in fibroblasts, which in turn alters the secretome to pro-tumorigenic direction (Erez et al., 2010). Several growth factors and ILs participate already in the early stages of cancer development to the reshaping of the TME and accumulation of immune cells (Kalluri, 2016). For instance, vascular endothelial growth factor A (VEGFA) produced by cancer cells and CAFs, increases vascular permeability, generates vascular leaks and enables reactive perivascular areas which promote accumulation of immune and endothelial cells as well as fibroblasts leading to enrichment of ECM proteins such as Col1 and initiation of tumor angiogenesis (Brown et al., 1999; Feng et al., 2000; Fukumura et al., 1998; Leung et al., 1989). Stromal content increases as cancer progresses and it is presumed that CAFs regulate this progression by their growth factors and ECM components including secretome (Kalluri, 2016).

CAF-secreted ECM proteases are able to cleave the basement membrane and facilitate invasion as well as to remodel ECM to generate immune or cancer cell migration enabling tracks (Gaggioli et al., 2007; Walker et al., 2018). ECM alterations by CAFs change also the adhesive properties of cancer cells, which drives epithelial-to-mesenchymal transition (EMT) and metastasis as well as increases therapy resistance of cancer cells (Costea et al., 2013; Dongre & Weinberg, 2019; Wang et al., 2018).

2.2.5   Origin of CAFs

Several theories exist regarding the origin of CAFs including differentiation from local cells such as resident fibroblasts existing already in healthy tissues, endothelial and epithelial cells as well as recruitment of circulating progenitors such as bone marrow-derived or adipose stem cells which differentiate into CAFs

within TME (Direkze et al., 2004; Ishii et al., 2003; Jotzu et al., 2010; Kojima et al., 2010; Mezawa & Orimo, 2016; Quante et al, 2011; Xing et al., 2010). Resident fibroblasts of the tissue of origin of the primary tumor represent probably the main origin of CAFs and TGFβ1 the main mobilization and activation promoting factor of these resident fibroblasts (Rønnov-Jessen et al., 1995; Zhang & Liu, 2013).

TGFβ-mediated EMT and TGFβ as well as SMAD-signaling-mediated endothelial-to-mesenchymal transition (EndMT) associates to the differentiation of local, native and mature cells transformed by cancer into functional CAFs and comprehends acquisition of mesenchymal morphology (Forino et al., 2006; Iwano et al., 2002; Kahounova et al., 2018; Potenza et al., 2008; Zeisberg et al., 2007). In addition to this EMT of secondary epithelium following cancer, stellate cells, fibrocytes, pericytes and adipocytes are sources of CAFs, recruited by TGFβ and CXCL-12 and transformed into CAFs by TGFβ and PDGFs (Abe et al., 2001; Hosaka et al., 2016; Jotzu et al., 2010; Yin et al., 2013; Zeisberg et al., 2007). EMT is additionally noted during vertebrate embryogenesis creating primitive mesenchymal cells, which enables cell dispersion and creation of new tissue structures (Kalluri, 2016, Kalluri & Weinberg 2009). Interestingly, there is evidence that also carcinoma cells are able to undergo EMT, which enables cancer cells to dedifferentiate and acquire enhanced migratory and invasive properties permitting them to move and reach distant organs (Kalluri & Weinberg, 2009; Karlsson et al., 2017; Scheel & Weinberg, 2012).

Another hypothesis states that cancer cells, especially cancer stem cells can be the origin of CAFs facilitated by TGFβ (Haviv et al., 2009; Nair et al, 2017). Within the context of MSCs and fibroblasts there is evidence that bone marrow-derived MSCs (BM-MSCs) formulate a significant proportion of CAFs in inflammation induced gastric cancer and that BM-MSCs enhance metastatic capabilities of breast cancer (Quante et al., 2011; Karnoub et al., 2007). This recruitment of MSCs and their activation into CAFs is stimulated by tumor cell-secreted CXCL-12 and TGFβ (Barcellos-de-Souza et al., 2016; Direkze et al., 2004).

2.2.6   Detection of CAFs

To date, there is no specific marker for either fibroblasts in general, activated fibroblasts or CAFs. In the absence of ubiquitous CAF identifying marker, numerous non-specific CAF-related markers have been utilized in order to distinguish CAFs as it is established that various markers can identify activated fibroblasts (Kalluri, 2016; Chen & Song, 2019). Most renowned of these include fibroblast activation protein α (FAP) and α-smooth muscle actin (αSMA), which have been assumed to distinguish CAFs from normal fibroblasts (Nurmik et al.,

2020). FAP is a type II transmembrane glycoprotein and expressed in addition to fibroblasts by a subset of CD45[+] immune cells and αSMA a cytoskeletal protein associated with smooth muscle cells (Arnold et al., 2014; Micallef et al., 2012). It has been estimated that 90% of CAFs express FAP (Huber et al., 2003). Although, αSMA is expressed by most mesenchymal cells including pericytes and myocytes in addition to myofibroblasts and quiescent fibroblasts, it is the main marker used to characterize CAFs (Desmoulière et al., 2004; Hawinkels et al., 2014; Wendling et al., 2009). Due to the expression of αSMA, fibroblasts have been classically called myofibroblasts (Micallef et al., 2012; Ronnov-Jessen & Petersen, 1993). Furthermore, activated fibroblasts in general were first described in the setting of wound healing and identified basically by their expression of αSMA (Micallef et al., 2012).

The absence of epithelial markers, cytokeratin and E-cadherin, endothelial cell marker CD31 and myeloid marker CD45 have been used defining CAFs (Louault et al., 2020). Based on more recent knowledge, it has been stated that the ability to interact with tumorigenic cells in the TME and the ability to persist in hyper-activated, cancer progression enhancing state actually separates CAFs from normal fibroblasts (Joshi et al., 2021). Along with the increasing understanding of the heterogeneous nature of CAFs, it has also been noted that the expression of CAF-related markers is remarkably heterogeneous and in fact it seems on the contrary to previous assumptions and expectations that there is no universal CAF identifying singular marker, at least to date (Chen & Song, 2019; Joshi et al., 2021). These notions have led to the identification of numerous different CAF subpopulations in various cancers with highly varying expression of singular CAF-related markers and to the usage of several markers instead of one in order to identify these subsets (Chen et al., 2021; Ham et al., 2019; Mezheyeuski et al., 2020; Zou et al., 2018). Nonetheless, most commonly used markers comprehend in addition to FAP and αSMA, vimentin (VIM) as well as platelet-derived growth factor receptor-α (PDGFRα) and –β (PDGFRβ) (Chen et al., 2021, Joshi et al., 2021). PDGFRα, for instance, has been reported to be expressed by up to 90% of stromal CAFs in solid tumors (Micke & Östman, 2004). The low specificity of singular markers, limited knowledge and the fact that genetic profile and expressed markers of CAFs in particular may differ greatly from caner to cancer especially in different anatomical locations has led to the urge of identifying novel markers as well as novel methods to identify CAFs (Chen et al., 2021, Joshi et al., 2021). It is also obvious that functionally activated fibroblasts do not express all the markers at same time (Özdemir, 2014). Novel identifying methods include functional approaches such as contraction and hydrogel assays (Nebuloni et al., 2016; Su et al., 2018; Zoetemelk et al., 2019). Further, scRNA-seq has enabled more in detail definition of CAF heterogeneity (Elyada et al., 2019).

Other CAF-associated markers include fibroblast specific protein 1 (FSP1), desmin and discoidin domain-containing receptor 2 (DDR2) (Chen et al., 2021; Quail et al., 2013). Although FSP1 has been used in detection of CAFs, it is reliable in detecting quiescent, non-proliferating fibroblasts, but also identifies macrophages, other immune cells and is expressed by cancer cells (Strutz et al., 1995; Österreicher et al., 2011). FAP is expressed in addition to fibroblasts by a subset of CD45[+] immune cells and mesodermal cells, Desmin and PDGFRβ by pericytes, PDGFRβ by quiescent fibroblasts and majority of markers by cancer cells (Armulik et al., 2011; Arnold et al., 2014; Nurmik et al., 2020; Roberts et al., 2013). Notably, based on messenger RNA (mRNA) expression of CAF markers including *VIM*, *ACTA2* (gene encoding αSMA), S100 calcium binding protein A4 (*S100A4*), collagen type I α 2 chain (*COL1A2*), integrin subunit beta 1 (*ITGB1*; *CD29*), tenascin C (*TNC*), *PDPN*, *POSTN* (gene encoding periostin), *FAP*, microfibril-associated protein 5 (*MFAP5*), *PDGFRβ*, collagen type 11 α chain 1 (*COL11A1*), integrin α 11 (*ITGA11*), and neural/glial antigen 2 (*NG2*), *VIM* was noted to be highly expressed in all types of cancer based on gene expression profiling interactive analysis databases, indicating that it is not specific for CAFs (Louault et al., 2020). On the contrary, in this analysis, *POSTN*, *ACTA2*, *S100A4*, *COL1A1* and *ITGB1* had high expression in cancers known to be abundant with CAFs and low expression in cancers known to be almost CAF free and thus indicating the better specificity of these markers (Louault et al., 2020).

Secreted protein acidic and rich in cysteine (SPARC) or osteonectin and periostin (POSTN) are ECM proteins expressed by fibroblasts that reflect changes in tumor stroma and CAF activation during cancer progression (González-González & Alonso, 2018; Guweidhi et al., 2005). Col1 is a fibroblast-derived collagen and one of the interstitial ECM proteins (Nissen et al., 2019). Additionally, POSTN induces the expression of Col1 (Yang et al., 2012).

Notably, whether CAFs are derived from resident fibroblasts or MSCs they seem to express the same markers including FAP, αSMA, VIM, but MSC-derived CAFs also MSC-related markers (Bhowmick et al., 2004; Borriello et al., 2017; Quante et al., 2011).

2.2.7 CAF heterogeneity

It has been established that CAFs represent a largely heterogeneous population of cells and in numerous cancers subpopulations of CAFs have been identified and characterized based on surface markers, secreted proteins and transcriptome (Louault et al., 2020). Although CAFs have been identified in wide variety of cancers, they embody various and even unique roles in different cancers, which on its part illustrates functional heterogeneity of CAFs (Awaji & Singh, 2019).

Heterogeneity in CAF biomarker expression was first demonstrated by co-staining studies and later reveals by scRNA-seq have led to the identification of variety of biomarker genes defining even wider variety of different CAF subpopulations (Chen et al., 2021; Sugimoto et al., 2006). In 2006, first pro-tumorigenic CAF subpopulations were identified in breast and pancreatic cancer models, followed by numerous other studies identifying various subpopulations with widely varying functions (Chen et al., 2021; Sugimoto et al., 2006). Thus, there is increasing interest in understanding various, even opposing roles of CAFs in cancer, but also CAF-related markers as potential targets for novel therapeutics (Joshi et al., 2021).

For instance, CAF-1-labeled subpopulation, expressing FSP1, has been shown to promote metastatic colonization and CAF-2-labeled and αSMA, NG2 and PDGFRβ expressing subpopulation to function as a source of Col1 and to contribute to the formulation of fibrotic, connective tissue barrier that prevents infiltration by cytotoxic T lymphocytes (Carstens et al., 2017; Li et al., 2017; Sugimoto et al., 2006). In human breast cancer four CAF subtypes, two of which being protumorigenic, were identified based on expression of six markers, integrin β1, FAP, PDGFRβ, FSP1, αSMA and caveolin 1 (Cav1) (Costa et al., 2018). Later on, by scRNA-seq further heterogeneity regarding original classification was revealed with eight new clusters among original CAF-S1 subgroup each expressing specific genes (Kieffer et al., 2020). It has been suggested that some subpopulations of CAFs exert their function by producing ECM proteins and others by acting as inflammatory cells leading to the formation of terms myofibroblastic CAFs (myCAFs) for ECM-producing CAFs and inflammatory CAFs (iCAFs) (Öhlund et al., 2017).

Pancreatic stellate cells, for instance, were shown to differentiate into either αSMA expressing as well as collagens and TGFβ producing myCAFs or into immunomodulatory and IL-6, IL-11 and leukaemia inhibitory factor (LIF) producing iCAFs that facilitate immune escape (Öhlund et al., 2017). By scRNA-seq it has been demonstrated that iCAFs are characterized by the production of inflammatory cytokines and chemokines including IL-6, IL-8, CXCL1, CXCL12, complement factor D, lamin A7C, PDPN as well as dermatopontin and myCAFs by the expression of αSMA, PDPN, tropomyosin 1 and 2 as well as POSTN (Dominguez et al., 2020; Elyada et al., 2019; Hosein et al., 2019). Functionally, iCAFs have been shown to promote angiogenesis and to possess immunomodulatory functions, myCAFs to promote tumor cell proliferation and invasion and both myCAFs and iCAFs to promote metastasis (Louault et al., 2020). Furthermore, myCAF subtype expressing leucine-rich repeat containing 15 (LRRC15) has been shown to be associated with chemoresistance to PD-L1 targeting therapy (Dominguez et al., 2020).

Another recently discovered CAF subtype with immunomodulatory function is antigen presenting-CAF (apCAF) which, in addition to markers expressed by myCAFs and iCAFs, also expresses major histocompatibility complex (MHC) class II family of proteins and inhibits activation and proliferation of lymphocytes (Elyada et al., 2019; Lakins et al., 2018). Another approach defined three different CAF subsets in murine breast cancer model comprehending vascular CAFs (vCAFs), matrix CAFs (mCAFs) and development CAFs (dCAFs) with vCAFs being located in perivascular niche and expressing endothelial cell markers, mCAFs being derived from resident fibroblasts and expressing matricellular proteins and dCAFs expressing stem cell genes (Bartoschek et al., 2018). In oral SCC, genomic-wise close to cSCC, two CAF subsets were identified comprehending CAFs with high expression levels of ECM remodeling proteins and CAFs with upregulation in cytokines and growth factors, demonstrating that relative amounts of CAF subtypes evolve during different stages of cancer progression (Puram et al., 2017). A few subsets characterized in literature are visualized in Figure 4.

There are indications that although for instance loss of Cav1 define fibroblasts with protumorigenic functions, high expression of Cav1 facilitates tumor invasion (Goetz et al., 2011; Simpkins, 2012). To date, it is obvious that there is a wide variety of CAF subpopulations between different cancers and with numerous functions from direct effects on tumor cell survival, proliferation, chemoresistance, invasion and metastasis to effects mediated by changes in ECM, vasculature or immune cells (Louault et al., 2020). Spatial heterogeneity or heterogeneity in CAF distribution within single tumor with pro- and antitumorigenic CAF subsets possibly co-existing adds another dimension of heterogeneity and further challenges interpretation of results as there is convincing indications that for instance myCAFs are located in close proximity to tumor cells while iCAFs are more distantly located (Bartoschek et al., 2018, Öhlund et al., 2017). CAFs can also transition between the myCAF and iCAF state *in vitro* (Öhlund et al., 2017). In contrast to what was previously assumed, CAFs can also harbor genetic mutations adding another dimension of heterogeneity (Patocs et al., 2007).
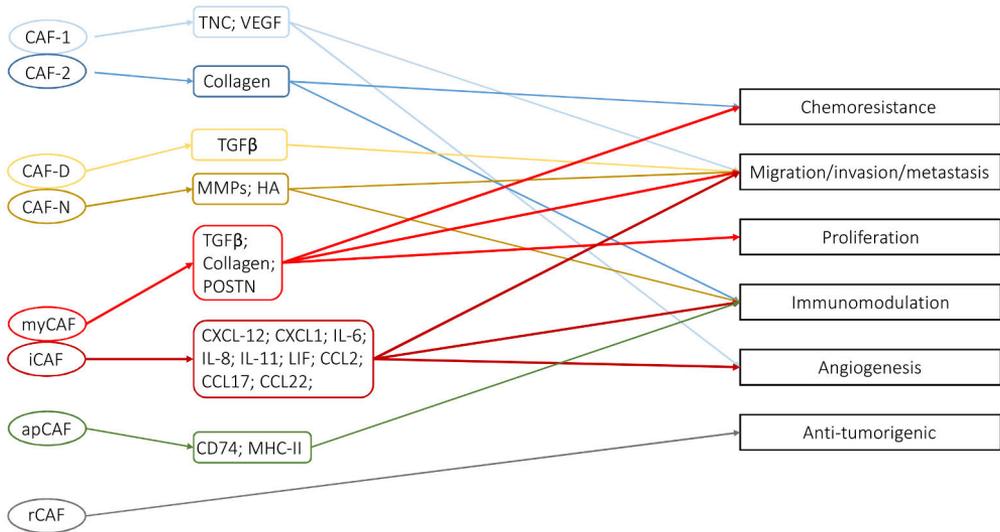
Figure 4. Examples of CAF subsets. Modified from Louault et al., 2020.

2.2.8 Association of CAFs with cancer progression

CAFs dynamically evolve along with cancers and modulate both directly and indirectly cancer progression and tumor immunity via their secretome (Chen et al., 2021). Regulatory functions are mediated by a wide variety of cytokines, chemokines and growths factors (Chen et al., 2021). There is evidence that CAF-derived TGFβ acts in immunosuppressive role and CAF-derived CXCL12 as tumor progression promoter (Orimo et al., 2005; Tauriello et al., 2018). CAF-secreted LIF, insulin-like growth factor 1 (IGF1), hepatocyte growth factor (HGF), IL-6, Wnt family member 5A (WNT5A) and bone morphogenetic protein 4 (BMP4) have been shown to promote tumor growth and progression (Chen et al., 2021). CAFs have been discovered to directly stimulate tumor cell invasion and migration by secreting TGFβ, IL-32, PDGF and FGF, that induce EMT, actin polymerization or depolymerization and cell motility (Wang et al, 2019a, Shu et al, 2020, Wen et al, 2019, Neri et al, 2017). In breast cancer model CAF-secreted IL-32 also promoted metastasis and CAF-secreted complement component 3a (C3a) induced invasion, metastasis and secretion of TGFβ, HGF and PDGF (Shu et al, 2020, Wen et al, 2019). Furthermore, in murine hepatocellular carcinoma-model CAF-induced Hippo-Yes-associated protein (YAP) signaling pathway promoted invasion and metastasis (Qiao et al, 2017). In combination with remodelled ECM by CAF-derived ECM components, ECM-degrading proteases and ECM-crosslinking enzymes, the cytokines and chemokines contribute the creation of immunosuppressive TME enabling cancer progression (Chen et al., 2021). On the other hand, numerous studies have shown context-dependently also tumor-

restraining functions of CAFs presumably due to their promotion of anticancer immunity, pro-inflammatory secretome, tumor-inhibitory signaling and production of barrier forming ECM components that inhibit tumor invasion and dissemination (Chen et al., 2021).

There are indications that FAP-expressing CAFs possess immunosuppressive function (Kraman et al., 2010). Depletion of αSMA positive CAFs have been shown to accelerate tumor progression and decrease survival and high αSMA expression and stromal densities to associate with favourable OS (Torphy et al., 2018; Özdemir et al, 2014). Further, it has been shown that αSMA positive myCAFs can exert tumor-restraining function through the deposition of Col1 and that genetic deletion of *Col1* in CAFs promotes liver colonization of pancreatic and colorectal cancers but does not affect cholangiocarcinoma growth (Affo et al., 2021; Bhattacharjee et al., 2021; Chen et al., 2021). However, there are also studies suggesting tumor-promoting functions of Col1 by increased cancer cell survival, proliferation and invasion (Armstrong et al., 2004). PDPN expression in CAFs has been associated with lymphatic/vascular spread and is expressed by αSMA+ and VIM- CAFs (Kitano et al., 2010). Further PDPN positivity in CAFs within breast cancer positively correlated with poor prognosis, invasion and tumor size as well as poor prognosis in lung carcinoma (Neri et al., 2015; Schoppmann et al., 2012).

Noteworthy, the prognostic value of singular biomarkers is hindered by the expression in other cell types leading to confusing conclusions and it has been noted that many of the CAF biomarkers do not have clear prognostic value when examined as single biomarkers based on analyses of datasets of Human Protein Atlas (Chen et al., 2021; Uhlen et al., 2017). However, αSMA has been shown to characterize peritumoral myofibroblasts and predict unfavourable prognosis in renal cancer, FAP to identify a subset of myCAFs and predict unfavourable prognosis in head and neck cancer and PDGFRβ to identify myCAFs and αSMA-low CAFs and predict unfavourable prognosis in renal and urothelial cancer (Chen et al., 2021). PDGFRα, again, has been shown to identify αSMA-low CAFs and some αSMA positive CAFs and predict unfavourable prognosis in renal cancer and favourable in head and neck cancer and VIM to act as pan-CAF marker and predict unfavourable prognosis in renal cancer and favourable in endometrial cancer (Chen et al., 2021). POSTN, in turn, has been shown to act as pan-CAF marker and predict unfavourable prognosis in renal, lung and stomach cancers and SPARC to act as pan-CAF marker and predict unfavourable prognosis in renal cancer (Chen et al., 2021). Thus, it is not surprising acknowledging the discovered anti- and protumorigenic subpopulations of CAFs in various cancers that there are contradictory results regarding the prognostic value of overall CAF populations (Chen et al., 2021). Further, although there are indications that certain subpopulations of CAFs have similar features and gene-expression

signatures across different cancer types, there are also subpopulations that vary substantially owing to the heterogeneous cellular origin of CAFs between cancers (Chen et al., 2021). These notions underline the need to define whether the prognostic value of biomarker-defined CAF subpopulation is universal or cancer-specific (Chen et al., 2021).

2.2.9   CAFs in cSCC

It has been shown also in skin cancers, including cSCC, that fibroblasts constitute a major component of the TME (Van Hove & Hoste, 2021). Regarding melanoma, it has been shown by scRNA-seq that CAF abundance varies between individual melanoma tumors and that high CAF abundance correlates with drug resistance (Van Hove & Hoste, 2021). It has been demonstrated with murine model that in premalignant, dysplastic skin lesions CAFs express proinflammatory gene signature, whereas in invase cSCCs fibrotic signature dominates indicating that fibroblast activity and phenotype evolves during cSCC progression (Erez et al., 2010; Van Hove et al., 2021). The myCAF – iCAF distinction appears to apply also to cutaneous cancers (Van Hove & Hoste, 2021).

FAP expression has been reported in cSCCs but not in benign cutaneous lesions including trichoepithelioma, pseudoepitheliomatous hyperplasia, poroma, syringoma and chondroid syringoma (El Khoury et al., 2014). It has been suggested that the expression pattern of CAF- and EMT-related proteins plays crucial role in the progression of cSCC and that CAF subset with high PDGFRβ, CD10, S100A4, αSMA, Zeb1, Slug and Twist is associated with cSCC, stating that TME at the invasive front of cutaneous malignancies shows specific expression patterns and vaguely proposing that CAF- and EMT-related markers may be useful in the prediction of lymph node metastasis of cSCC (Sasaki et al., 2018). It has also been noted in monolayers that in comparison to normal fibroblasts isolated from healthy skin, CAFs in cSCC display different morphology, increased proliferation and migration as well as secrete high levels of pro-collagen I and increase invasion (Commandeur et al., 2011). POSTN expression in peritumoral CAFs has been shown to correlate with the aggressiveness of cSCC (Lincoln et al., 2021). It was also noted that approximately 80% of αSMA positive cells express POSTN (Lincoln et al., 2021).

2.3   Artificial intelligence (AI)

2.3.1   AI and machine learning

Artificial intelligence (AI) was defined already in 1955 as the "science and engineering of creating intelligent machines that have the ability to achieve goals like humans via constellation of technologies" (Puri et al., 2020). In the past decades AI has already revolutionized our digitalized society. Within medical field, it has been stated that AI will be a milestone for healthcare in the running decade and that pathology will be right at the focus of this revolution with AI providing the next step towards precision pathology (Acs et al., 2020). AI in general uses algorithms with appropriate data to execute specific tasks and comprehends machine learning and its subset deep learning (DL) (Murphree et al., 2021). Such tasks in medical field include determination of diagnosis, predicting prognosis and outcomes with therapeutic interventions (Murphree et al., 2021). In addition to machine learning, concepts of self-awareness, Bayesian inference and knowledge representation and reasoning can be categorized under the hypernym AI (Sultan et al., 2020).

Machine learning is central discipline of AI and exploits algorithms that detect patterns within existing data, trains itself and makes predictions on new data based on what was detected and learned earlier (Xing & Yang, 2016). Machine learning approaches can be categorized in various ways but generally by the type of data they take into account and by the question they answer (Hastie et al., 2009). Main forms of machine learning include reinforcement leaning, dimensionality reduction and DL (Sultan et al., 2020). Reinforcement learning comprehends continuous learning cycle occurring through positive and negative reinforcement within a computational environment in which the system learns as a consequence of its past actions and is rewarded by successful outcomes (Sultan et al., 2020). Dimensionality reduction bases its learning on reducing the dimension of a large feature-based dataset and aims to improve accuracy and speed of model by reducing the amount of misleading or redundant data (Sultan et al., 2020). Thus, dimensionality reduction is successful in correcting over-fitting and can be achieved for instance by applying feature selection (Sultan et al., 2020).

DL is the most recent evolution of machine learning and can be further divided into discriminative and generative models, both of which can be approached by supervised, self-supervised or unsupervised learning depending on the amount of supervision they need in order to learn from data (Sultan et al., 2020). Discriminative models are effective in tasks such as image classification and segmentation, whereas generative models in visual synthesis and rendering (Sultan et al., 2020). In supervised learning the machine is trained on independent variables

or the data which is associated with dependent variables or known outcomes or responses (Murphree et al., 2021). Supervised learning is used in predictive problems with traditional methods including logistic regression but also in DL utilizing neural networks (NNs) (Murphree et al., 2021). Supervised learning dominates medical applications and is based on training with data with known outcomes with the aim of finding a mathematical function that can map input data into output predictions (Tseng et al, 2020). One example of supervised learning in dermatology field would be classifying skin lesion photographs with histopathologically verified diagnoses as outcomes (Murphree et al., 2021). In self-supervised learning the data is autonomously labeled by the machine without manual labeling by human (Sultan et al., 2020). In unsupervised learning there are no known outcomes, but the algorithm seeks potential relationships between data points (Murphree et al., 2021). Clustering is one example of unsupervised learning, in which data is assigned to groups based on similarity (Murphree et al., 2021).

2.3.2   Deep learning and artificial neural networks

DL, as described above, is a subdiscipline of machine learning, that can handle larger datasets, has more complicated functionality and provides more advanced decision making capabilities (Sultan et al., 2020). DL requires less manual preprocessing and can learn more complex features with higher efficiency enabling more successful pattern recognition and computer vision (Anwar et al., 2018). As DL is effective performing visual tasks, it has been most widely used in medical specialties dealing with visual data such as radiology, pathology and dermatology (Sultan et al., 2020). These pattern-recognition applications have been shown to effectively deal with large medical images in object detection and image classification tasks (Kim et al., 2019). In DL tasks dealing visual material, the input variables (or covariates or predictors) are pixels which are combined to specific outcome (or response) (Murphree et al., 2021). Most common methods in image analysis comprehend classification, detection and segmentation (Murphree et al., 2021). Categorization of entire image is the outcome of classification task, determination of whether there is searched entity and the location of this entity the outcome of detection task and the category of each pixel the outcome of segmentation task which is a subtype of classification classifying images at pixel level instead of entire image level (Murphree et al., 2021). The most promising approaches in medical image analysis utilize algorithms known as artificial neural networks (ANNs), and more specifically deep neural networks (DNNs) referring to the deep instead of shallow construction of the network (Wang et al., 2019b). Most applications in pathology and dermatology field utilize DNNs that have learned in supervised manner (Murphree et al., 2021; Sultan et al., 2020).

ANNs are statistical or computational models that mimic biological neural networks in human brain (Murphree et al., 2021; Sultan et al., 2020). Parameters (or instructions or weights) mathematically combine input variables to calculate predicted output and are adjusted during the training phase of ANNs (Murphree et al., 2021). ANNs incorporate a non-linear special function called activation function which differentiates ANNs from linear or nomogram-based models (Murphree et al., 2021) This function allows ANN to capture significantly broader spectrum of relationships between input variables (Murphree et al., 2021). ML is called DL, when ANN includes multiple (>1) hidden layers between the input and output layers (Duggento et al., 2021).

Architecturally, the most basic form of ANN comprehends three layers or main components, which are input layer, hidden layer and output layer (Sultan et al., 2020). Analogous to biological neural network, input layer equates dendrite receiving the input from initial data source and output layer equals axon producing the final result (Sultan et al., 2020). Number of hidden layers between the input and output layers varies depending on the depth of the architecture and contribute to the solving of more complicated computational tasks (Sultan et al., 2020). In DNNs output from one layer becomes input to the next layer with each subsequent layer learning sequentially more complex features with deepest layers performing the ultimate task such as classification of image into an output category (Murphree et al., 2021). Layers are connected and contain artificial neurons, which are programmed to perform desired tasks (Sultan et al., 2020). Importantly, the size of the network and the performance of ANN is proportional to the size of the data that has been utilized in the training (Sultan et al., 2020). Generally, large ANN trained with massive data perform better than one with less encountered training data, but large ANN trained with limited data produces a model prone to over-fitting and poor generalizability (Sultan et al., 2020).

2.3.2.1 Convolutional and residual neural networks

Numerous specialized DNNs, such as deep convolutional neural networks (CNNs) have been developed for different tasks (Shin et al, 2016). DNNs utilized in medical applications with visual input data are commonly CNNs, which comprise of numerous nodes (or neurons), their connections and parameters demonstrating heightened performance in terms of spatial features (Sultan et al., 2020). Millions of nodes and their parameters are involved in the network that observes positive and negative samples as nodes learn to coordinate their functions to perform desired task requiring high computational power of graphics processing units in training phase (Sultan et al., 2020). Architecturally CNNs incorporate, unlike traditional DNNs, several different layers including hidden convolutional layer(s)

and pooling layers (LeCun et al, 2015; Rawat & Wang, 2017). Convolutional layers function as feature extractors identifying essential features from input signals and pooling layers maintain system steady minimizing the size of the network (Sultan et al, 2020).

Designing DNNs is an active subfield of computer science and many medical applications leverages architectures designed by computer vision groups and major technology corporations such as Google (Murphree et al., 2021). Thus, there are several publicly available architectures to harness into more specific medical task via for instance transfer learning (Murphree et al., 2021). Such classification acrhitectures include Inception V3, VGG16 and Xception as well as segmentation and detection architectures such as YOLO and U-Net (Murphree et al, 2021). Residual neural network-18 (ResNet-18) is an 18-layer deep and ResNet-50 a 50-layer deep CNN, which add residual learning to traditional CNN (He et al., 2016). ResNets resemble pyramidal cells in cerebral cortex and utilize skip connections to jump over layers in order to solve the problem of gradient dispersion and accuracy degradation (He et al., 2016).

2.3.2.2   Operational models and execution

Regarding technical execution of creating and training a DNN, the most important step is undoubtedly to define the clinical aim and to formulate a suitable clinical question, which will quide the collection of data, labeling, choice of the algorithm and training (Murphree et al., 2021).

To specify an ANN or DNN for specific task, architecture of the network, loss function and optimization route need to be determined (Murphree et al., 2021). Architecture comprehends the number and type of layers and further the number and type of nodes (or neurons) in selected layers (Murphree et al., 2021). The loss function, which can for instance be categorical cross-entropy or binary cross-entropy, evaluates the performance of the algorithm on the training data and is used and actively improved during the training of the network (Murphree et al., 2021). Algorithms that aim to minimize the loss function by adjusting parameters of the network are called optimizers (Murphree et al., 2021). In practice, the selection of an optimizer is determined experimentally as also learning rate, a crucial parameter in the optimizer, is usually problem dependent and determined through trial and error (Murphree et al., 2021).

In order to use any ANN, it must first be trained with data (Murphree et al., 2021). In this training phase the parameters of the network are updated to facilitate the network to execute the desired task (Murphree et al., 2021). In supervised learning, the dataset consists of observations and labels, that represent the ground truth description of the corresponding observation (Murphree et al., 2021).

Unsurprisingly, the quality of labeling is crucial and must be correct in order to the network to perform successfully (Murphree et al., 2021). Furthermore, the dataset size is crucial, however, there is no universal answer to what is enough (Murphree et al., 2021). There have been vague suggestions from for instance 1000 to 5000 examples per output category for acceptable performance to propositions of over 10 million examples to exceed human performance (Goodfellow et al., 2016; Krizhevsky et al., 2017). Undoubtedly, success of AI depends on big data as large number of training examples are needed to generate reliable predictive models (Granter et al., 2017). Thus, the establishment of large enough databases of precisely and unanimously characterized input variables is crucial (Sultan et al., 2021). However, in oncologic pathology many cancers, such as mcSCC, are quite rare and large databases do not exist, which hinders the development of DL utilizing applications. Not surprisingly majority of studies to date have focused on the most common cancers with most comprehensive data archives including prostate and breast cancer (Chang et al., 2019).

There are several ways to improve the performance of DNN including dropout, data augmentation and transfer learning (Murphree et al., 2021). In dropout, certain parts of the network are randomly turned off to reduce the probability of overfitting (Murphree et al., 2021). Data augmentation takes advantage of prior knowledge to generate new data and increase the size of the dataset by incorporating for instance rotated versions of the same image multiplying the number of examples within limited data set (Murphree et al., 2021). Transfer learning enables transfer of an existing successful algorithm to a newly created model which reduces the training time and requirement for computational power and enables the use of smaller datasets (Rawat & Wang, 2017). In more detail, a pre-built and pre-trained on large dataset classifier is transferred to a new model by for instance fine-tuning, in which earlier convolutional layers, tasked to extract low-level features, of the network are frozen and deeper layers left for modification to extract specific features relevant to the task related to new dataset (Sultan et al, 2020).

Computationally, high-performance graphical processing units are necessary to train DNNs as numerous calculations are performed during the training (Murphree et al., 2021). Even with state of the art computer designed for DL, training may take weeks depending naturally on the size of the dataset (Murphree et al., 2021). Using a DL algorithm, in turn, does require only modest computational resources (Murphree et al., 2021).

2.3.3 Whole slide imaging

AI-driven solutions have already demonstrated to enhance diagnostic accuracy within the field of oncologic pathology, which comprehends various time

consuming manual tasks (Sultan et al., 2020). The main motive for the urge to incorporate machine learning to histopathological image analysis is to reduce intra- and inter-observer variability and to improve objectivity and reproducibility (Xing & Yang, 2016). Advances in computing power, networks and storage has facilitated the management of digital slide images with more ease and flexibility bringing the technology to clinical use (Niazi et al., 2019). Digital pathology in general can be described as computer technology enabled, image-based information environment which allows management of information generated from digital slides (Niazi et al., 2019). Whole slide imaging (WSI) on its part allows entire slides to be imaged, digitalized and stored at high resolution and acts as an enabling platform for the application of AI (Niazi et al., 2019; Zarella et al., 2019). WSI, with the parallel advancements in AI, have enabled image-based and computer-aided diagnostic possibilities that utilize and integrate knowledge beyond human boundaries (Niazi et al., 2019).

WSI is widely used in teaching and sharing purposes and integration of AI further improves also this aspect (Niazi et al., 2019; Zarella et al., 2019). WSI has had great impact on quality assurance by allowing fast teleconsultations, proficiency testing and archiving as well as gauging of inter- and intra-observer variance (Niazi et al., 2019). AI can further ameliorate this by providing error preventive and diagnostic quality improving quality checks and digital reviews (Niazi et al., 2019). In clinical practice there are already slide-less pathology laboratories, as based on numerous studies comparing digital slides and glass slides in diagnostic point of view it has been concluded that WSI can be considered as equivalent golden standard to conventional microscopy (Mukhopadhyay et al., 2018; Niazi et al., 2019). AI can improve the maintaining of digital pathology work flow by detecting out-of-focus areas and improving color standardization as the quality of digitalized images by WSI scanners influences the performance of pathologist and the reliability of diagnosis (Shrestha et al., 2016).

Most modern scanners are equipped with autofocus optics, which enables accurate capturing of three-dimensional tissue morphology into the two-dimensional image, however due to the varying thickness of tissue sections focus points in different focal planes are necessary but can predispose to production of out-of-focus areas in the digital image due to erroneous selection of focus point in incorrect plane (Moles Lopez et al., 2013; Senaras et al., 2018). One approach to overcome this issue is to apply AI such as DeepFocus to the task, which utilizes representation learning to recognize out-of-focus regions and due to its generalized learning abilities can be applied to images representing various tissues and different staining modalities (Senaras et al., 2018). One major challenge for diagnosis carried out by both human and machine is color variations on digital slides, that can result from numerous reasons including variations in the thickness

of section, disparity in scanning characteristics and differences in staining protocols or staining reagents (Komura & Ishikawa, 2018). These color variations also hinder the generalization and universal application of ML algorithms as for instance outstandingly performing AI-algorithm may fail when applied to external material scanned with different scanner or with different adjustments due to the lack of color standardization (Komura & Ishikawa, 2018). Further, the prediction performance of diagnostic AI-models can suffer in performance due to non-biological experimental variations ergo batch effects comprehending the differences in slide preparation, specification and adjustments of imaging device and post-processing software as well as the age of the tissue specimen (Kothari et al., 2014). Thus, there are several AI utilizing approaches to improve the color standardization (Bentaieb & Hamarneh, 2018; Gatys et al., 2016; Zanjani et al, 2018). Even a virtual histological staining method that utilizes CNN has been introduced (Rivenson et al., 2019).

2.3.4 AI in pathology

There are already several image analysis tools in clinical practice including computer aided diagnosis for mitosis counting, HER2/neu assessment in breast cancer and Ki67 assessment in carcinoid tumors (Niazi et al., 2019). AI utilizing applications can be divided into diagnostic, prognostic and predictive, integrating, workflow efficiency improving and educative classes (Couture et al., 2018). Diagnostic applications further comprehend independent reporting algorithms that provide diagnosis without the input from pathologist, diagnosis-aided tools that assess features on the slides such as tumor grade, automatic quantification tools of specific features such as immunohistochemistry stainings and tools that predict biomarkers on hematoxylin and eosin (H&E) stained slides (Couture et al., 2018). Prognostic and predictive tools aim to provide prognosis and response to specific therapy based on myriad of morphological features including tumor cell morphology, architecture of tissue and the presence of lymphovascular invasion without the need to specify isolated features (Ferroni et al., 2019; Wulczyn et al, 2020). Integrative applications aim to find associations between morphological features and tumor genetic profiles or predict the underlying molecular alterations based on morphological features (Ninomiya et al., 2009).

Applying DL to the analysis of digitalized tumor biopsies has been under interest focusing first on smaller tasks such as mitosis detection dealing small portion of the WSI and later on utilizing whole-slide information in order to develop fully automated methods to detect carcinoma on biopsies (Albarqouni et al., 2016; Bejnordi et al., 2017a; Litjens et al., 2016). In these tumor detection applications, the WSIs are classified as malignant or benign based on the presence

of tumor regions (Bejnordi et al., 2017a; Cruz-Roa et al., 2018; Niazi et al, 2018). Such diagnostic application was developed to detect breast cancer lymph node metastasis, BCC on skin and prostate cancer on prostate biopsy with an area under the receiver operating characteristic curve (AUROC) of >0.98 allowing pathologists to exclude 65-75% of slides with 100% sensitivity (Campanella et al., 2019). Further, in CAMELYON16 challenge an application developed to identify breast cancer metastasis reached comparable to pathologist detection of micrometastasis with no time-constraint, demonstrating applicability to act as a screening tool that eliminates metastasis-free samples with 100% sensitivity on cytokeratin-stained lymph node sections (Bejnordi et al., 2017b; Holten-Rossing et al., 2017). Another study further demonstrated AI-algorithm as productivity improving and false negative reducing tool in breast cancer metastasis detection (Liu et al., 2019b).

Immunohistochemical (IHC) stainings are time consuming and prone to observer variability, but widely performed and provide crucial information for diagnosis, prognosis and treatment selection (Das et al., 2013; Zaha, 2014). It has even been implied that machine is superior to pathologist in Ki67 scoring indicating that AI could settle the problem of insufficient reproducibility in Ki67 scoring across laboratories and pathologists which have prevented the usage in clinical practice, although Ki67 is known to possess prognostic and predictive capabilities (Acs et al., 2019a; Klauschen et al., 2015; Rimm et al., 2019; Stålhammar et al., 2018; Zhong et al., 2016). Reproducibility for pathologists is also an issue regarding scoring of tumor-infiltrating lymphocytes (TILs) demonstrated by results of a study in which automated TIL scoring represented with independent prognostic potential in melanoma unlike scoring by pathologist (Acs et al., 2019b). Regarding oral SCC, a CNN with an accuracy of 96% in quantifying TILs from WSIs was developed and consequently revealed that TILs represent strong prognostic indicator of disease-free survival (Shaban et al, 2019).

As an example of integrative application that sheds light on IHC status and histopathological pattern, a DL algorithm was harnessed to predict PD-L1 status from H&E stained WSIs of non-small cell lung cancer (NSCLC) with an AUROC of 0.80 indicating that PD-L1 expression correlates with morphological features of the TME (Sha et al., 2019). Additionally, there are numerous integrative applications that illustrate the correlations between mutational status and morphological features of the tumor (Coudray et al., 2018; Kather et al., 2019; Schaumberg et al., 2018).

Several prognostic applications have revealed features possessing prognostic abilities and indicated that prediction by AI acts as an independent prognostic factor (Bychkov et al., 2018; Lu et al., 2017; Kulkarni et al., 2020; Mobadersany et al., 2018; Wang et al., 2017). Moreover, AI-assisted utilization of image data with

comprehensive clinical and outcome data, that can be gathered AI-assistedly, enables high dimensional analysis and transforms pathology towards informatics science in which actual tissue would act only as one of the data sources (Niazi et al, 2019).

As comprehensive histopathological analysis of cancer ultimately requires several factors to be accounted, also AI applications have been developed to complete multiple tasks (Couture et al., 2018). Such algorithm identified histologic subtype with 94%, histological grade with 82%, estrogen receptor status with 84%, molecular subtype with 77% and stratified recurrence risk with 76% accuracy on H&E stained tissue microarray (TMA) images of breast cancer (Couture et al., 2018).

2.3.5   AI in dermatology

Dermatology represents another field, in addition to pathology, radiology and cardiology, in which images play key part and that consequently is susceptible for DL utilizing medical image analysis (Puri et al., 2020). DNNs dominate in dermatological applications and usually utilize transfer learning (Puri et al., 2020).

For instance, a CNN was developed to distinguish cSCCs from seborrheic keratoses (SK) and melanomas from benign nevi on images of lesions with dermatologist-level performance (Esteva et al., 2017). Another CNN was harnessed to identify melanoma from dermoscopic images with dermatologist surpassing classification efficiency illustrated by an AUROC of 0.86, eventhough dermatologists were allowed to incorporate clinical information to decision making (Haenssle et al., 2018). A CNN harnessed to classify 12 different skin lesions on clinical non-dermoscopy images reached AUROCs of 0.96, 0.83 and 0.96 for BCC, cSCC and melanoma respectively (Han et al., 2018). It was also noted that algorithm trained on Asian patients performed poorly on Caucasians and vice-versa, underlining an aspect to note during training (Han et al., 2018).

Regarding dermatopathology, CNN was trained on whole slide images (WSIs) to identify BCCs, dermal nevi and SKs with AUROCs of 0.99, 0.97 and 0.99 respectively (Olsen et al., 2018). Another CNN distinguished Spitz nevi from conventional melanocytic lesions on WSIs with an accuracy of 92% (Hart et al., 2019). An AUROC of 0.998 was reach for automated detection of aggressive malignant melanoma from benign nevus on WSIs (Wang et al., 2020). Even more interestingly a CNN was harnessed to predict the prognosis of early stage melanoma on standard H&E images reaching AUROCs of 0.880 and 0.905 for DSS (Kulkarni et al., 2019).

Further regarding the development of the prognostic 40-GEP test for cSCC, DL was utilized to train the algorithm on gene expression data (Wysong et al., 2021).

3   Aims

With current knowledge, there is controversy regarding most of the high-risk and especially metastasis-associated clinicohistopathological features of cSCC. Furthermore, in the absence of clinically established biomarkers, the assessment of metastasis risk is challenging. Thus, the aim of this thesis was to create comprehensively characterized cohorts of non-mcSCCs and mcSCCs and to examine the feasibility of clinicohistopathological features, CAF-associated markers and AI in the risk stratification of cSCC. The specific aims of the substudies were:

Study I: To determine the epidemiology, risk factors, and prognosis of metastatic cSCC (I, the cohort study)

Study II: To explore the activation of cancer-associated fibroblasts in cSCC (II, the CAF study)

Study III: To harness AI algorithm to detect primary metastatic cSCCs on whole slide images (III, the AI study)

4 Materials and Methods

4.1 Ethical issues (I-III)

The studies were approved by The Ethics Committee of the Hospital District of Southwest Finland (187/2006) and Auria Biobank's Scientific Steering Committee (AB15-9721). The research conformed to the Declaration of Helsinki. Registry study approval for collection and use of clinicohistopathological data was received from the Turku University Hospital Clinical Research Centre (18.11.2018; TO5/042/18).

4.2 Patient and tumor cohorts (I-III)

Study region comprehends Hospital District of Southwest Finland, Satakunta Hospital District, Vaasa Hospital District and Ahvenanmaa Hospital District which is not administratively included into Turku University Hospital Catchment Area, unlike the rest. Patients demanding academic tertiary care center-level treatment from above-mentioned hospital districts are principally all treated at Turku University Hospital as is the case regarding mcSCC.

Three sequential automated screenings were executed by Auria Biobank utilizing topographical codes and keywords. First screening with topographical code C44 (ICD-10) and keyword "squamous" resulted in 136 patients with probable cSCC and formed the basis for non-metastatic patient cohort. Second screening with topographical code C44 (ICD-10) and keywords "squamous" and "metastatic" resulted in 77 patients potentially with mcSCC. After manual review of patient records and pathology reports there were 35 patients with genuine mcSCC and 42 patients with non-mcSCC accompanied by another metastatic malignancy. These 42 patients were translocated into non-metastatic patient cohort. Patients in original non-metastatic patient cohort were reviewed leading to the exclusion of patients with only BCC and/or cSCCIS and revelation of 8 patients with mcSCC, who were translocated into metastatic patient cohort if they were not already included. This provided 4 additional patients with mcSCC. Due to the low number of patients with mcSCC (n=39) a third screening covering wider time frame was carried out in the beginning of 2019 with topographical

code C44 and keyword "metastatic" resulting 772 patients potentially with mcSCC. Patient record and pathology report review revealed 71 patients with genuine mcSCC. 28 of these were already found in previous screenings and 43 patients were novel to the study and were included into metastatic patient cohort. No patient was included into non-metastatic patient cohort from third screening. Duplicates were removed on multiple occasions. Patients with mcSCC were identified by review of pathology reports and medical records regarding whole patient history of every patient. Based on this as comprehensive as possible review it was determined whether patient had genuine mcSCC. Every cSCC of included patients was registered and scrutinized. Notably, keratoacanthomas were not classified as cSCCs. Formulation of patient cohorts is visualized in flow chart (Figure 5).
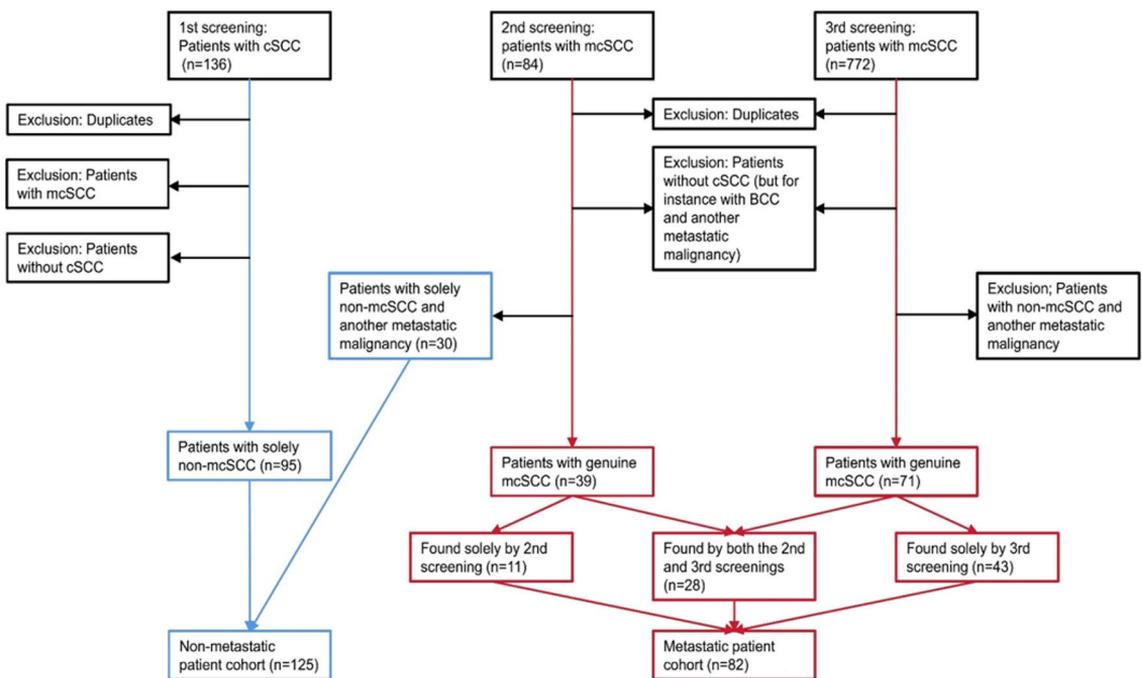
Figure 5. Flow chart showing formation of patient cohorts from raw screening data to characterized final patient cohorts. Red shows metastatic and blue non-metastatic patient cohort. From original publication I.

Eventually, 207 patients were included in the study and divided into a patient cohort comprising of 82 patients with at least one mcSCC and a control cohort consisting of 125 patients with solely non-mcSCC(s) (I, Table 1).

4.2.1   Patient-level variables (I-III)

Clinical data was gathered from the patient record system of Turku University Hospital utilizing every specialty sheet available. Additionally, regional patient data service Altti, which provides data over organizational boundaries, was reviewed. Age, sex, every medication and comorbidity as well as smoking, occupational and survival information were patient-level variables under interest. In respect of patients with mcSCC, also data associated with treatment and diagnostics was collected.

Sex was registered regarding every patient and every tumor and evaluated in both patient and tumor analyses. Age was determined regarding every tumor. Age at the initial diagnosis of first cSCC was applied in patient characteristics-related analyses and age at the initial diagnosis of the exact tumor in tumor characteristics-related analyses.

Inclusion criteria considering medication for patients with non-metastatic disease was long term (at least 6 months) regular usage prior to the diagnosis of first cSCC and respectively for patients with metastatic disease prior to the diagnosis of first primary mcSCC. Low-dose aspirin usage in cardioprotective manner was registered on its own and if used on higher doses as analgesic long term it was registered under NSAIDs. In collection phase, isosorbide mononitrate and isosorbide dinitrate were combined under same variable i.e. for inclusion patient had to have history of either using isosorbide mononitrate and/or isosorbide dinitrate long term.

In similar manner comorbidities were included if they were diagnosed prior to the diagnosis of first cSCC for patients with non-metastatic disease and prior to the first primary mcSCC for patient with metastatic disease. However, co-malignancies and keratinocyte premalignancies were included whether diagnosed prior to or after the diagnosis of first cSCC or mcSCC respectively and time relation was determined.

SOTRs, patients with HIV infection, CLL and Non-Hodgkin lymphoma as well as patients with certain medications were regarded immunocompromised. Notable is that there was great overlap regarding the cause of immunosuppression as majority of patients had history of several immunosuppressive medications and/or other above-mentioned immunosuppressive characteristics. Following medications were classified as immunosuppressive: mycophenolate, interferone, chlorambusil, cyclophosphamide, azathioprine, systemic steroid, ciclosporin, thalidomide, methotrexate, sulfasalazine, chloroquine, hydroxychloroquine.

In survival analyses death was the clinical endpoint. If death was not encountered, patient was interpreted as censored. Cause of death was collected from the patient records if available. Unfortunately, the exact underlying cause of death was not reported regarding majority of cases and autopsies were rarely

performed. Thus, unambiguous OS was chosen as an endpoint for substudy I. However, DSS was deduced from patient records by clinical judgment and utilized in substudies II and III. Last contact to health care provider represented the end of follow-up. Follow-up time was specified regarding both censored and dead patients. Patient records for the patients that were alive at primary data collection were revisited during August 2019 in order to get as up to date as possible survival information and to extend the follow-up times. Date of the first tissue specimen representing named cSCC represented the beginning of follow-up in all substudies regarding primary tumor analyses.

In substudy I, for mcSCCs, it was determined whether complete response (CR) was or was not achieved. CR was achieved if both the primary tumor and all metastases had been successfully treated and there was no disease activity during the rest of the follow-up. Respectively if disease continued to progress despite interventions it was classified as "CR not achieved". CR was not achieved either if patient developed distant metastasis(-es). Primary tumor or metastasis was considered locally advanced if the lesion was clinically regarded unresectable based on tumor properties.

4.2.2   Tumor-level variables (I-III)

Tumor-related characteristics were collected primarily from pathology reports but also patient records of the patient record system of Turku University Hospital as well as from Altti. Pathology reports were provided by database application, QPAWEB. Archived records in paper form were reviewed when necessary. Tumors were biopsied and/or surgically excised by dermatologists, otorhinolaryngologists, plastic surgeons, ophthalmologists, general surgeons and hand surgeons in Turku University Hospital as well as primary care physicians and specialists in smaller hospitals and private clinics in study region. Age at exact tumor diagnosis, sex and every histopathologic characteristic available considering both primary tumors and metastases were collected.

Date of the first histopathologic specimen (biopsy or resection) of named cSCC was interpreted as the date of diagnosis regarding primary tumors. Primary tumors included into the tumor analyses were diagnosed between years 1993-2018. Local recurrences were regarded as continuum to the primary tumor and not counted as novel individual tumors. Local recurrence was reckoned when cSCC on the same exact location had been previously treated and new clinically detectable tumor had thereafter formed.

Histopathologic variables were manually derived from pathology reports which were not structured apart from some of the latest cases, 2017 onwards. In tumor analyses, considering non-mcSCCs presence of characteristic in one point of tumor

continuum was enough for inclusion and the highest met value of variable was chosen taking into account every biopsy and resection including primary tumor and its local recurrences. Regarding mcSCCs, only characteristics present and highest value of variable met prior to or on day of metastasis detection were included into the tumor analyses. In this manner for instance tumor diameter was registered. Histological diameter was utilized primarily if available. When exact histological diameter was not reported (tumor n=144 (47.5%) in substudy I) clinical diameter was registered instead with clinical judgment in concordance with available histopathological information. Due to this approach, diameter regarding every tumor except for one could be defined, but instead of scale measure nominal measure was regarded appropriate expedient and thus applied. Smallest tumor-free resection margin was extracted from the reports with provided accuracy. Quantitative invasion depth was infrequently (25.7% in substudy I) expressed in pathology reports but anatomical invasion depth i.e. Clark's level could be deduced from the pathology reports considering 96,3% of tumors included in substudy I with accuracy segregating Clark 2-4 from Clark 5.

Assumption regarding nominal values such as PNI, vascular and lymphatic invasion or desmoplasia was that if it was detected during the diagnostic inspection of the tissue sample by pathologist it would have been expressed in pathology report. Presence of above-mentioned variables was extremely rare in pathology reports especially prior to metastasis detection and no sensible statistical analysis could be executed. In addition, histologic subtype was rarely encountered in pathology reports and thus excluded from the analyses. Histological differentiation was expressed in pathology reports from grade 1 through 4. As there was only one tumor with grade 4 differentiation in the material of substudy I, it was combined with grade 3, which together formed the group of poor differentiation. Similarly, grade 1 tumors formed well differentiated and grade 2 tumors moderately differentiated groups.

Regarding metastases, the date of diagnosis was either the date of the imaging study or the date of histopathologic sample depending on which one was performed first. Concerning metastases, all mcSCCs had histopathologically confirmed metastasis(-es), except for 2 primary tumors which had solely radiologically confirmed metastases. Topography, tumor characteristics, lymphatic drainage of the primary tumor, time association and the lack of other potential primary cancer were endorsing the establishment of cSCC being metastatic and in addition grounded the determination of primary mcSCC tumor. Regarding every mcSCC, the exact primary tumor could be determined. In most challenging cases with multiple potential primary tumors, the metastatic one was chosen based on clinical judgment utilizing all the available clinical and histopathological information.

Interpretation of cutaneous tumor as cutaneous metastasis was made with great caution and only tumors that were unambiguously cutaneous metastases of cSCC by pathology report and additionally by clinical evaluation, were considered actual cutaneous metastases. In numerous cases despite pathology report suggesting cutaneous metastasis it was by clinical judgment regarded instead as primary tumor, most frequently as local recurrence. As far as nodal metastases are concerned clinical judgment was conducted especially in respect of metastases located in parotid gland as there were several occasions when it was not obvious based on pathology report that metastasis was nodal in origin. In these cases, metastases were interpreted as nodal as was the case regarding some cervical metastases where the origin by pathology report was not evident but clinically referred to nodal rather than cutaneous or other structure in origin. Presence of ENE of nodal metastasis was registered from clinical records and pathology reports. If ENE was evident by clinical record it was registered whether mentioned apart in pathology report and vice versa.

Both AJCC8 and BWH primary tumor staging was executed regarding every applicable primary tumor to visualize the association between these established staging systems and metastasis as well as prognosis. Both tumor stagings were determined in above-mentioned manner ergo utilizing characteristics present prior to or on day of metastasis detection as far as mcSCCs are concerned. Acknowledging that AJCC8 is in practice reserved exclusively for the head and neck cSCCs, the AJCC8 was applied only to tumors located at head and neck region in substudy I but in substudies II and III AJCC8 was applied to tumors regardless the location due to the notion by Venables et al. (2021). Primary tumors that could not be assessed were classified as TX and AJCC8 T4a and T4b were combined in analyses as there was only one T4b tumor (mcSCC) prior to or on day of metastasis detection in substudy I.

4.2.3   Research material utilized in CAF study (II)

In substudy II, TMAs representing study cohorts consisting of formalin-fixed, paraffin-embedded human tissue specimens obtained by biopsy or resection were constructed. TMAs are comprised of spots representing normal skin (n=73), benign epidermal tumors: seborrheic keratoses (SKs) (n=17), premalignant lesions: AKs (n=67), cSCCISs (n=59), UV-cSCCs (n=217, 143 individual tumors): non-mcSCCs (n=146, 110 individual tumors) and mcSCCs, (n=71, 33 individual tumors), and metastases of UV-cSCCs (n=16, 11 individual metastases). In addition, RDEBSCCs (n=77, 12 individual tumors) were included (Kivisaari et al., 2008). Normal skin, SK, AK and cSCCIS specimens were acquired from patients treated in Turku University Hospital and are characterized in earlier study by Riihilä et al.

(2015). Above-mentioned numbers represent the number of spots with approved high-quality multiplexed fluorescence immunohistochemistry (mfIHC) stainings. Normal skin specimens include specimens from both sun-exposed and sun-protected skin. Tissue material was obtained from patients treated in Turku University Hospital except RDEBSCCs. In analyses, mean values of the expression in spots representing individual UV-cSCCs, RDEBSCCs and metastases were calculated and utilized.

UV-cSCC cohorts and metastases utilized in substudy II represent part of the cohorts described in substudy I. All mcSCCs included in substudy II developed at least one nodal metastasis and part extranodal metastases.

4.2.4   Research material utilized in AI study (III)

In substudy III, H&E stained, archived formalin-fixed and paraffin-embedded tissue specimens representing UV-induced primary non-mcSCC and mcSCC tumors characterized in substudy I were scanned and digitalized into WSIs. Ultimately, one tumor per patient and one WSI representing each cSCC was used to train and test the AI algorithm, and in the analyses by dermatopathologist. All in all, 45 mcSCC from 45 and 59 non-mcSCCs from 59 individual patients, who did not develop cSCC metastasis during at least 5-year follow-up were included in the analyses. Thus, only part of the tumors and patients characterized in substudy I were included in the substudy III. As the substudy III continued, mcSCCs were divided into two subcohorts comprehending subcohort with primary mcSCCs that developed metastasis within 180 days from the initial diagnosis of primary tumor (rapid metastasis cohort) (n=22) and another subcohort with primary mcSCCs that developed metastasis after 180 days (slow metastasis cohort) (n=23). Thus the research material utilized in the final rapid metastasis AI –model comprehended in total 81 patients and tumors with 59 non-mcSCCs and 22 mcSCC and formed the base for further analyses including the creation of risk factor models (RFMs) and survival analyses.

4.3   Rate of metastasis (I)

In order to determine the rate of metastasis in the study region statistical information of overall cSCC incidence was gathered from the database of Finnish Cancer Registry, which by the time of the drafting of the manuscript of substudy I covered years between 1953 and 2016 (Finnish Cancer Registry, 2022). The mcSCCs were distributed based on the year of the initial primary mcSCC diagnosis. The number of mcSCCs was then divided by the total number of cSCCs in the study region concerning each year. Due to the transition into electronic

health record system during 2001-2003, it is possible that there are mcSCCs undetected by our screenings prior to this transition. Thus we ended up taking into account cSCCs that occurred during ten-year period between years 2004 and 2013 in the determination of the rate of metastasis. This allowed cSCCs with more than five years for metastasis development and we can assume that even the latest tumors, diagnosed 2013, would have metastasized by the time of the third screening and thus would be included into this study as metastatic.

4.4 Gene expression and bioinformatics analyses (II)

For RNA analyses, surgically removed primary non-metastatic (n=3) and metastatic (n=2) human cSCC specimens were collected from patients in Turku University Hospital (Viiklepp et al., 2022). Total RNA was isolated from the tissue specimens as described previously (Viiklepp et al., 2022). RNA (100ng) was hybridized overnight at 65°C with the Human Fibrosis Panel, Cancer Progression Panel and Cancer Pathways Panel (NanoString Technologies, Seattle, WA). For purification and binding of the hybridized probes to the cartridge nCounter Prep Station was used. nCounter Digital Analyzer (Nanostring Technologies) was utilized for the preparation of the scanning of the cartridge. nSolver 4.0 (NanoString Technologies, Seattle, WA) was used for the preparation of data analysis. The default QC settings was used for the normalization and confirmation of the quality of the data.

GSE111582 (Cho et al., 2018) dataset representing RDEBSCC gene expression profile was downloaded from the publicly available GEO and used to analyze the expression of CAF markers in RDEBSCC samples (n=8).

The online Gene Expression Profiling Interactive Analysis tool (GEPIA; http://gepia.cancer-pku.cn/) was used to analyze the relationship between *PDGFRβ, FAP*, and *PDGFRα* mRNA–expression and prognosis of HNSCC and lung SCC (LUSCC) in The Cancer Genome Atlas (TCGA) data (Uhlen et al., 2017; Weinstein et al., 2013).

4.5 Multiplexed fluorescence immunohistochemistry (II)

Two mfIHC panels comprehending following eight CAF markers (II, Supplementary Table SI): PDGFRα, PDGFRβ, αSMA and FAP in panel 1, and SPARC, PDGFRβ, POSTN, Col1 and VIM in panel 2 were utilized to define the CAF markers and their distribution in stromal and epithelial compartments of tissue specimens represented in created TMAs. Used antibodies are listed in II,

Supplementary Table SI. For both panels, PanEpi cocktail was used to detect epithelial cells (II, Supplementary Table SI). Notably, these markers are expressed in tumor tissue in UV-cSCCs and in RDEBSCCs by mRNA analysis (II, Supplementary Figure S1A and S1B).

4.5.1   Imaging and image analysis (II)

The whole-slide TMA imaging (to achieve 5-channel fluorescence images) was executed using Axio Scan.Z1 Digital Slide Scanner (The Zeiss™, Germany) equipped with 20X (0.8NA) Plan-Apochromat objective (The Zeiss™, Germany), ORCA-Flash 4.0 V2 Digital CMOS camera (Hamamatsu Photonics K.K., Japan) and Colibri.7 LED light source (The Zeiss™, Germany). DAPI, FITC, Cy3, Cy5 and Cy7 filters were applied.

The created 5-channel fluorescence images were exported as single-channel grayscale images (64 Bit, Big Tiff Format) and resized to a quarter of the original image size. ImageJ (version 1.53 M for Windows) with Roi1 1-Click Tools was used to crop the images to individual TMA spots. The second staining round TMA spot images were overlaid with the first-round images using nuclei (DAPI) as reference. MATLAB and Statistics Toolbox Release 2020b (The MathWorks Inc., Natick, Massachusetts, US) was used for registering.

For autofluorescence (e.g. red blood cells and empty areas (background)) detection from the images Ilastik (version 1.3.3post1 for Windows) was utilized, generating final tissue mask (Berg et al., 2019). CellProfiler (version 4.2.1) was used for final image analysis pipeline (Stirling et al., 2021). Tissues were classified into epithelium and stroma compartments by thresholding the epithelial channel (epithelial compartment) and subtracting the epithelium from the total final tissue (stroma compartment). By diligent visual determination of the positivity threshold each marker was set as negative or positive for every single pixel. "Mask image" module was used to define different marker combinations. Regarding both the epithelial and stromal compartments, the total number of marker classified pixels were measured. For the generation of relative cell areas, proportions of marker and marker combination positive pixels were divided by either total epithelial or stromal pixels. Mean intensities of channels were also measured in both stromal and epithelial compartments.

The quality of TMA cores was evaluated by visual inspection. Every TMA core with low quality (e.g. ruptured or folded tissue or staining artefact) was excluded from further analyses.

Expression level of every marker in both panels was analyzed independently in both stromal and epithelial compartments for both the relative cell area and intensity (II, Supplementary Table SII). Additionally, regarding panel 1, 18 two-

marker combination CAF subsets (CAFa-r), 28 three-marker combination CAF subsets (CAF101-128) and 15 four-marker combination CAF subsets (CAF1-15) based on relative cell area analysis were created (II, Supplementary Table SII). Regarding panel 2, further nine two-marker combination CAF subsets (CAF1a-1i) based on relative cell area analysis were created (II, Supplementary Table SII).

4.6 Whole slide image processing (III)

Tissue specimens (slides) utilized in substudy III were scanned into WSIs with 3DHistech scanner (3DHistech, Budapest, Hungary) and annotated using CaseViewer (3DHistech, Budapest, Hungary).

All the WSIs in substudy III represented cSCCs and were manually annotated. Annotated area comprehended whole tumor on the WSI and included tumor cells, both intra- and peritumoral stroma and intratumoral inflammatory cells (III, Figure S3). Manual exclusions were made in case of artefacts on tumor area. DNNs were trained and validated by these annotated areas of the WSIs.

As WSIs are large, the images were cut into 1024×1024 and 512×512 pixel tiles at the largest zoom level (20×) from the annotated tumor areas. Next, the tissue content of the tiles was analyzed utilizing a combination of binary and Otsu thresholding. Tile was valid if it contained more than 50% of tissue pixels as defined by the thresholding algorithm. The tile was considered as tumor tile if it contained over 50% pixels from the tumor annotation mask. The cohort label of the original slide was inherited by all the tiles. Before training the algorithm, every tile was resized into 299×299 pixels.

4.7 Training and validation of AI algorithm (III)

The task of distinguishing mcSCCs from non-mcSCCs was approached as a binary classification problem on the level of single tumor tile. Every tumor tile was assigned to a single cohort and the binary classification problem was to classify the tiles into either non-mcSCC or mcSCC cohort. Due to the relatively small dataset, we approached the task with cross-validation. In cross-validation the data is sectioned into subsets with one subset at a time used for testing and rest for training. As a result, several models are trained and performance of each model is tested with alternating validation subset.

In final rapid metastasis -AI-model ResNet18 architecture with a custom head consisting of average pooling and two dense layers with heavy dropout to combat overfitting were utilized. More complex architecture based on ResNet50 was also tested but was more prone to overfitting, which lead to the ultimate choice of a simplier model. All models were trained using 3-fold cross-validation (3-CV) with

different extraction tile sizes until the rapid metastasis -AI-model in which 4-fold cross-validation (4-CV) was used instead. Binary cross-entropy loss function was used. In order to prevent the leakage of information between the training and validation sets, during training it was made sure that the tiles from a given patient and slide were exclusively sampled into one of the cross-validation-folds. The model was trained 10 or 20 epochs with every fold.

Out-of-the-fold (OOF) tile level receiver operating characteristic (ROC) curves and calculated area under the curve (AUC) values were created in addition to analyzing the cross-validation accuracy and loss. The OOF tile level predictions were mapped to slide level. Each tile and the corresponding location on the WSI was assigned a prediction probability between 0 and 1 for being mcSCC. This probability was scaled to [0,100] and then shifted to the interval [-50,50] for each tile. The scores on WSIs were then spatially smoothed using a median filter with window size 2. Meaning that a window of 2x2 tiles was moved along the slide and the tile value was replaced with the median of the tile values within the window. This was done to remove noise based on our hypothesis, that the features representing metastasis risk would vary smoothly across the tissue slide. However, there is no way to scientifically select a correct size for the smoothing window. Our selection was based on visualizations of the results. The metastasis scores were visualized by probability maps of the annotated tumors (III, Figure 5).

After the OOF tile level predictions were accumulated to slide level, a simple majority vote of the scaled predictions was performed to determine the predicted label of the WSI and the tumor by assessing the mean slide level score. We took 0 as the decision threshold to discriminate between the low/high metastasis risk tumors or non-mcSCCs and mcSCC and visualized the slide level results in ROC curves with AUC scores and summary confusion matrices. The workflow from WSI input to tile level result is visualized in III, Figure 6.

4.8   AI-models (III)

Single tile -AI-model represented our initial approach, in which individual tiles served as input. This approach is backed by the hypothesis that as most of the cells are the offspring of the cancerous cells and inherit the genetic alterations of the first generation of tumor cells, most tiles should represent the possible differences between primary mcSCCs and non-mcSCC. Further, this classification algorithm is easy to implement and it is easy to aggregate and visualize the single tile predictions at slide level, however this kind of model is prone to label noise. Next approach was to focus on the invasive edge of the tumor (invasive front -AI-model). Metastatic characteristics of the tumor are hypothetically more likely visualized in the invasive front and thus focusing on this area was expected to lead

to less noisy training labels. Using stack of tiles (multi-tile -AI-model) as input data represented the third approach. In this model, instead of just one tile, randomly selected sample of stack of n tiles from the same WSI was used as the input of the classifier. In above mentioned approaches ResNet50 architecture was used with 3-CV and single tiles from annotated WSIs representing the two cohorts (primary mcSCCs and non-mcSCCs) were used as input data. In multi-tile -AI-model instead of single tile a stack of tiles and in invasive front -AI-model only tiles adjacent to the edge of the annotation (either inside and/or outside of the annotated area) from annotated WSIs were used as input data.

Due to the notion of rapid development of metastasis in substudy I, we ended up subdividing the mcSCC cohort into cases that metastasize rapidly and slowly. We also hypothesized that features characteristic for the metastatic primary tumor would be more prominent or more probably already present at the date of tissue specimen in primary tumors, which metastasize rapidly. Therefore, only tumors, which metastasized rapidly were utilized in the rapid metastasis -AI-model. Both stack of tiles and single tiles from annotated WSIs representing primary non-mcSCCs and rapid mcSCCs were used as input data. ResNet18 architecture with 4-CV was ultimately used. In order to further prevent overfitting in this more exclusive dataset, we used a "zoomed-in" approach in which 512x512 pixels instead of 1024x1024 pixels tiles were used.

4.9   Blinded assessment by pathologist (III)

For comparison purposes in substudy III, every WSI included in the final rapid metastasis -AI-model was analyzed by experienced dermatopathologist. Only tissue sample ID and information, whether the specimen represented biopsy or resection was provided to the pathologist in accordance with access to CaseCenter folder including 81 WSIs representing non-mcSCCs and rapid metastasis mcSCCs without the knowledge of the proportion of cases. The pathologist also had the option not to classify the specimen into the two cohorts (non-metastatic/metastatic), if assessment was not possible, as was the case regarding biopsies.

4.10   Statistical analyses (I-III)

4.10.1   Cohort study (I)

The variables were collected from the database of Turku University Hospital and Altti between June 1, 2018 and August 22, 2019. The data was initially collected and organized in Microsoft Excel 2016 (*Microsoft* Corp., Redmond, WA, *USA)*

and then extracted to IBM SPSS Statistics for Windows, version 25.0 (IBM Corp., Armonk, N.Y., USA) in order to carry out statistical analyses. Bidirectional p-value less than 0.05 and 95% confidence intervals (CIs) of odds ratios (ORs) not including 1.00 were considered to show statistical significance. All analyses were conducted using IBM SPSS Statistics for Windows, version 25.0 (IBM Corp., Armonk, N.Y., USA). In tables, count of missing variables was expressed but statistical tests performed excluding these.

Regarding analyses of patient-related characteristic, every patient was counted once concerning both metastatic and non-metastatic patient cohorts. Respectively in tumor analyses every cSCC including locally recurrent tumors were counted once regardless patient cohort.

Baseline patient and tumor characteristics were analyzed using descriptive statistics mainly crosstabs and frequency tabulation. Statistical analyses were conducted using principally Pearson χ^2 test. Fisher's Exact Test was executed if > 20% of cells had expected count less than 5 or the minimum expected count was < 1. For polytomous variables exact tests were executed and column proportions compared via Bonferroni method except in case of detailed location, which was executed via Monte Carlo method. Regarding scale variables such as age, Mann-Whitney U test or independent sample T-test was utilized based on the normality of distribution with interquartile ranges (IQR) and standard deviations (SD) respectively.

Binary logistic regression analyses with 95% CIs were executed in order to determine the ORs regarding the development of metastasis and prognosis. In respect of tumor characteristics logistic regressions were performed using generalized estimating equations in order to factor in the impact of multiple tumors on same patient. Quasi likelihoods were examined regarding every logistic regression on tumor characteristics and working correlation matrix structure with lowest corrected likelihood selected, which happened to be "independent" regarding every analysis. Both unadjusted or crude (cOR) and adjusted ORs (aOR) were conducted. Adjustments were executed as shown in tables. Staging systems (AJCC-8 and BWH) were excluded from adjusted model due to multicollinearity between independent variables. After staging systems were removed from the model, variance inflation factor (VIF) between the rest 11 independent variables were < 3.0. References were selected based on statistical strength and clinical judgment. Variables were selected into logistic regression analyses based on clinical and histological ponderability as well as the results of descriptive statistical analyses.

Kaplan-Meier method was applied to construct survival curves and to define survival probabilities. Statistical analysis was executed by Log Rank (Mantel-Cox)

test. In similar manner one minus survival curve was created to visualize time to metastasis.

4.10.2   CAF study (II)

Applicable raw data of the substudy I was utilized in substudy II. All statistical analyses were conducted using IBM SPSS Statistics for Windows, version 25.0 (IBM Corp., Armonk, NY, USA) and/or R statistical Software 4.1.1 (Foundation for Statistical Computing, Vienna, Austria). Two different analysis paths were followed: one for all different tissue types (normal skin, SKs, AKs, cSCCISs, UV-cSCCs, metastases and RDEBSCCs) and another for UV-induced non-mcSCCs and mcSCCs. Statistical significance was based on the 95% confidence level and all the tests were two-tailed. Independent samples Kruskal-Wallis H-test with Dunn-Bonferroni post hoc method was used to test differences of ordinal and continuous variables between the seven cohorts. Mann-Whitney U test was used in comparison of non-metastatic and metastatic UV-cSCC tumor cohorts. Binary logistic regression analyses were performed in order to determine crude ORs (cORs) with 95% CIs for the risk of metastasis. The Kaplan–Meier method with Log-Rank (Mantel-Cox) test was applied to generate survival curves and define survival estimates. Proportional hazards assumption was tested and univariate and multivariate Cox regression analyses were conducted in order to further visualize the prognostic influences with test for PH assumption. Statistically significant clinical binary variables in univariate Cox regression analysis were included into multivariate analysis except invasion beyond fat and staging systems (AJCC-8, BWH), which were excluded due to multicollinearity. Individual CAF marker or marker combination was included one at a time to the model. Median, quartile and highest quartile versus rest distributions were tested in the Kaplan-Meier and Cox regression analyses but without exception highest quartile versus rest approach resulted in best results. The reported Kaplan-Meier survivals and hazard ratios (HRs) are based on stratification between the highest quartile of values (denoted below as high and +) versus the lowest three quartiles (denoted below as low and -) with respect to markers and marker combinations. In logistic regression analyses the numerical marker values were multiplied by 100 to represent the percentage of marker or marker combination positive cells within named compartment and to make ORs more relevant.

4.10.3   AI study (III)

Applicable raw data of the substudy I was utilized in substudy III. All statistical analyses were conducted using IBM SPSS Statistics for Windows, version 25.0

(IBM Corp., Armonk, NY, USA). Bidirectional *p*-values <0.05 and 95% CIs of ORs not including 1.00 were considered statistically significant. Baseline tumour characteristics were analyzed using descriptive statistics mainly crosstabs and frequency tabulation. Statistical analyses were conducted with Pearson χ^2 test and Fisher's exact test. Binary logistic regression analyses with 95% CIs were performed in order to determine ORs regarding the risk of metastasis. For every risk factor and risk factor combination an AUROC was calculated in order to further visualize results in relation to AI prediction. Pearson correlation was conducted in order to visualize multicollinearity and to examine whether predictions correlated with some of the clinicopathological variables. The Kaplan–Meier method was applied to generate survival curves and define survival probabilities.

5 Results

5.1 Patient-level characteristics of cSCCs (I)

In substudy I, 207 patients were included into patient-level analyses. Non-metastatic patient cohort comprised 125 patients with in total 184 non-mcSCCs and metastatic cohort 82 patients with 85 mcSCCs and 68 non-mcSCCs (I, Table I). In the analyses every patient was counted once regardless the number of invasive cSCCs and non-metastatic and metastatic patient cohorts were compared. Number of primary cSCCs per patient varied between 1 and 26 with median of 1.0 in both cohorts (p=0.742) and with slightly higher mean of 1.9 in metastatic cohort compared to 1.5 in non-metastatic patient cohort (I, Table SI).

Median age at the time of the initial diagnosis of first primary cSCC in non-metastatic and metastatic cohort was 77 (range 47-92) and 75 years (range 27-95) respectively for males and 81.5 (range 55-102) and 78 years (range 60-93) for females (I, Table SI). Men were younger in metastatic cohort and cohorts combined with statistical significance (p=0.037 and p=0.003), but there was no significant difference between cohorts sexes combined (p=0.084). Neither was there significant difference between the cohorts regarding sex (p=0.139).

Occupational data was available concerning 70.5% of patients. Farming, masonry, construction, seafaring, gardening, charging and forestry were classified as high UV-exposure occupations. 11.2% and 15.9% of patients in non-metastatic and metastatic cohorts respectively represented above-mentioned high UV-exposure occupations (p=0.322). Smoking information was available regarding only 50.3% of patients and there was no statistically significant difference in smoking between the cohorts (p=0.549). (I, Table SI)

Six (4.8%) patients in non-metastatic and 2 (2.4%) patients in metastatic cohort were SOTRs (p=0.483) and one patient in metastatic cohort HIV positive (p=0.396). In non-metastatic cohort 16.8% of patients and in metastatic cohort 13.4% of patients had history of usage of at least one medication classified as immunosuppressive (p=0.510). There was no significant difference between cohorts regarding immunosuppression in general (p=0.822) or any of the immunocompromising features separately. (I, Table SI)

Cohorts were relatively similar regarding the number of cSCCs per patient and comorbidities despite acne rosacea which was not reported in any metastatic cohort patient but in 7 patients in non-metastatic cohort. There were no statistically significant differences between cohorts regarding co-malignancies excluding AK and BCC. When AK diagnosed regardless time association to cSCC was analyzed, there was statistically significant difference between the cohorts (p=0.016) as 36.6% of patients in metastatic cohort and 53.6% of patients in non-metastatic cohort had AK diagnosed by the time of data collection. Additionally, BCC was more commonly met in patients with non-metastatic disease (p=0.001). (I, Table SI)

The number of medications per patient was both as mean and median value almost identical in the cohorts. There was statistically significant difference between the cohorts regarding previous aspirin (p=0.012), isosorbide mono-/dinitrate (p=0.009) and 5α-reductase inhibitor usage (p=0.024) and close to statistically significant difference concerning the use of thiazide diuretics (p=0.076) (I, Table SI).

There were 40 pharmaceutical agents and 36 comorbidities that were not included in the analyses because the frequency cohorts combined was less than 3.

5.2   Tumor-level characteristics of cSCCs (I)

In total 303 individual cSCCs from 206 patients were included into tumor analyses (I, Table I) in substudy I. Non-metastatic tumor cohort consists of 218 cSCC, of which 173 (79.4%) were on patients with solely non-mcSCC(s) and 45 (20.6%) on patients with mcSCC in addition (I, Table I). Metastatic tumor cohort in turn consists of 85 individual mcSCC from 82 patients (I, Table I). In these tumor level analyses cohorts of non-mcSCCs and mcSCCs were compared. Median age at the time of exact tumor diagnosis was 77.0 years in metastatic and 78.0 years in non-metastatic tumor cohort (p=0.821) (I, Table SII). Majority of the tumors in both cohorts were on males with no statistical significance between the cohorts (p=0.696) (I, Table SII). However, men were younger than women at exact tumor diagnosis in metastatic tumor cohort (p=0.018), non-metastatic tumor cohort (p<0.001) and both cohorts combined (p<0.001).

Majority of the primary tumors were located on head and neck region in both the non-metastatic (71.6%) and the metastatic (82.4%) tumor cohorts. Every primary tumor located on orbital region was metastatic. In addition, every tumor located on lower lip in males and 62.5 % of the tumors in females were metastatic. Representing the other end, only 8.9 % of the primary tumors on cheek excluding preauricular region were metastatic. There was statistically significant difference between cohorts regarding location on lower lip, orbita and cheek excluding

preauricular region when every detailed location was analyzed. Another interesting notion was prominent male predominance regarding both metastatic and non-metastatic tumors located on auricle with 8.7 male:female count-adjusted ratio as there were only 2 tumors on auricle in females compared to 33 tumors in males. However, metastatic tumor percentage in males (36.4%) was in line with the percentage regarding head and neck region in general (33.8%). Other main regions (trunk, upper limb and lower limb) had significantly less statistical strength with lower metastatic tumor percentages. (I, Table SII)

Percentage of metastatic tumors increased in linear manner from good to poor histological differentiation and there was statistical significance between cohorts regarding well, moderately and poorly differentiated tumors (p<0.001). As far as primary tumor diameter is concerned there was statistical difference between cohorts in diameter groups of <10mm, 20-29.9mm, 30-39.9mm and ≥40mm (p<0.001). Interestingly co-location of precursor lesion and/or previous histopathologically diagnosed precursor lesion in the same exact location was more common regarding non-mcSCCs than mcSCCs (p=0.006). Necrosis among primary tumor was more common feature in mcSCCs than non-mcSCCs with statistical significance (p=0.002). (I, Table SII)

Regarding invasiveness, 58.8% of Clark 5 tumors and only 4.5% of Clark 2-4 tumors were metastatic with p-value of <0.001 between the tumor cohorts. In similar manner 70.1% of tumors invading beyond subcutaneous fat and 12.8% of tumors without invasion beyond subcutaneous fat were metastatic (p<0.001). Interestingly also diffuse growth pattern associated with metastasis (p<0.001). Tumor was registered to possess diffuse growth pattern if based on pathology report it was clear that tumor formed distinct unattached nests in addition to actual primary tumor focus. (I, Table SII)

Local recurrence prior to metastasis detection elevated the risk of metastasis as 53.8% of the locally recurrent tumors metastasized in comparison to 22.7% of the tumors that did not develop local recurrence prior to metastasis detection (p<0.001). In addition, there were local recurrences regarding metastatic tumors after metastasis detection. Further, when every local recurrence was taken into account the local recurrence seemed to occur sooner for metastatic than non-metastatic tumors with median times of 224 and 127 days respectively, but no statistical significance was reached (p=0.060). Mentionable is also that when exact tumor-free margins were examined there was only slight decrease in locally recurrent tumor percentage whether margin was >1mm (11.0%), >2mm (10.1%) or >3mm (9.5%). (I, Table SII)

5.3   Characteristics of metastatic cSCCs (I)

In substudy I, 82 patients had in total 85 individual mcSCCs. Seventy-nine (96.3%) patients had singular mcSCC and 3 (3.7%) patients 2 independent mcSCCs. Interestingly, regarding 72 out of 85 (84.7%) mcSCCs, the primary metastatic tumor was the first cSCC ever diagnosed on named patient. Only 29.4% had AK or cSCCIS diagnosed prior to or on same day as the first diagnosis of primary mcSCC. If in addition BCC, metatypical carcinoma and prior cSCC were taken into account the percentage was slightly higher (36.5%). (I, Table SIII)

For 62 (72.9%) primary mcSCCs the original detection of first metastasis was made clinically, which led to further examinations. In 22 (35.5%) of above 62 cases staging-oriented imaging and in 2 (3.2%) cases sentinel node biopsy was performed with negative result prior to clinical suspicion of metastasis rose. In turn, subclinical imaging and sentinel node biopsy revealed the first metastasis in 19 (22.4%) and 4 (4.7%) cases correspondingly. (I, Table SIII)

For 82 (96.5%) mcSCCs the first detected metastasis was nodal and for 3 (3.5%) mcSCCs cutaneous. Not surprisingly, the first detected metastasis for primary mcSCCs located on head and neck region was most likely located in ipsilateral parotid gland (44.3%) or on ipsilateral neck (35.9%). Overall, 83 (97.6%) of the mcSCCs were accompanied with nodal metastasis/-es. One of the two cases without nodal metastasis had cutaneous and pulmonary metastases and the other only cutaneous metastasis. For 42 out of 83 (50.6%) mcSCCs with nodal metastases the number of nodes involved was three or higher and for 17 (20.5%) the diameter of largest nodal metastasis was >60mm. ENE was met in at least one involved nodal metastasis regarding 48 (57.8%) mcSCCs and contralaterally to primary tumor located nodal metastasis(-es) in 10 (12.0%) cases. (I, Table SIII)

Twelve (14.1%) mcSCCs were accompanied with cutaneous metastasis/-es and 13 (15.3%) with distant metastasis/-es, pulmonary metastases representing the most common distant metastases. In 15 (17.6%) cases the primary tumor and in 23 (27.1%) cases the metastasis could be classified as locally advanced. Further, in 2 (2.4%) cases both primary tumor and metastasis evolved into locally advanced phase. (I, Table SIII)

5.4   Rate of metastasis and time to metastasis (I)

During the 10-year period between 2004 and 2013 there were 2097 cSCCs diagnosed based on the statistics of Finnish Cancer Registry and 46 mcSCCs based on substudy I in the study region. During above-mentioned study period the annual rate of metastasis varied between 0.82% and 4.46% with mean metastatic rate of 2.28%. When viewed sex-wise the numbers for females were 1034 cSCCs, 18

mcSCCs and a mean rate of metastasis of 1.89% and for males 1063 cSCCs, 28 mcSCCs and a mean rate of metastasis of 2.71%. (I, Figure 1)

Median time between the initial diagnosis of primary mcSCC and its metastasis was 198 days (IQR 62-527 days) and the metastasis was detected within six months in 42 (49.4%) cases and during two years in 72 (84.7%) cases. Notably, in cases with longest intervals there were several local recurrences during multiple years before the development of metastasis. Concerning 7 (8.2%) mcSCCs first metastasis was diagnosed prior to or on day of initial diagnosis of primary mcSCC. (I, Table SIII and Figure 2)

5.5   Association between CAFs and the invasion of cSCC (II)

In substudy II, the expression of CAF markers *in vivo* was examined by mfIHC in TMAs containing tissue samples representing different stages of UV-induced cSCC progression, *i.e.* AKs (n=67), cSCCISs (n=59), sporadic UV-cSCCs (n=146) and metastases of UV-cSCCs (n=11). The expression levels were also compared with those in RDEBSCCs (n=12), normal skin (n=73), and SK samples (n=17). mfIHC stained CAF markers and their relative frequency distributions in different tissue sample types are shown in representative images (II, Figure 1 and Figure 2A). In tissue samples, the heterogeneity in the expression of CAF markers was evident (II, Figure 1). Cells with strong expression of PDGFRβ, FAP, POSTN, and SPARC were more abundant in the stroma adjacent to invasive tumor islets in comparison to stroma adjacent to normal skin (II, Figure 1).

FAP, αSMA and SPARC positive stromal cells were more prevalent in invasive UV-cSCCs than in AKs or cSCCISs, indicating association of these markers with invasion (II, Figure 2B). Furthermore, proportion of POSTN positive stromal cells was higher in UV-cSCCs than in cSCCISs (II, Figure 2B). In turn, proportion of PDGFRβ positive stromal cells was similar in AKs, cSCCISs and UV-cSCCs, whereas stromal PDGFRα positivity was more prevalent in normal skin than in UV-cSCCs (II, Figure 2B).

Interestingly, the stromal expression of markers was as strong in metastases as in UV-cSCCs with the exception of POSTN, which was more prevalent in metastases, and VIM, which was less frequent in metastases (II, Figure 2B).

Unsurprisingly, the expression of CAF markers in the epithelial compartments of the tissue samples was lower than in stroma. The frequencies of CAF marker positive cells in the epithelial compartments were in accordance with those of stromal compartments. (II, Supplementary Figure S2)

The results show prominent stromal staining in RDEBSCCs for FAP, αSMA, SPARC, POSTN, and VIM. Interestingly, the proportion of stromal cells positive

for PDGFRβ and Col1 was significantly lower in RDEBSCCs than in UV-cSCCs providing evidence for the presence of distinct CAF subset in this type of cSCC. (II, Figure 2B)

5.6 Factors associated with the risk of metastasis (I-III)

5.6.1 Clinical and histopathological (I)

In substudy I, 85 mcSCCs and 218 non-mcSCCs were eligible for logistic regression analysis regarding tumor-related characteristics and 82 and 121 patients for logistic regression regarding patient-related characteristics when metastasis risk was analyzed.

Tumor-level factors associated with the risk of metastasis with statistically significant aORs included, from highest to lowest: primary tumor location on lower lip (aOR 38.64; 95% CI, 5.41-275.75), location on forehead (aOR 15.55; 95% CI 2.13-113.49), Clark's level 5 (aOR 10.18; 95% CI 2.35-44.08), primary tumor diameter 20-29.9mm (aOR 6.31; 95% CI, 1.31-30.49) and age 70-79 years at the diagnosis of exact tumor (aOR 5.87; 95% CI, 1.66-20.83). (I, Table III)

In addition to above mentioned features with statistically significant metastasis risk elevating aORs, local recurrence prior to metastasis detection (cOR 3.97; 95% CI, 2.03-7.77), diameter at least 30mm (cOR 13.35; 95% CI, 5.33-33.45), moderate (cOR 4.13; 95% CI, 2.24-7.64) and poor histologic differentiation (cOR 7.89; 95% CI, 3.60-17.33), necrosis among primary tumor (cOR 3.78; 95% CI, 1.54-9.32), invasion beyond fat (cOR 14.54; 95% CI, 7.78-27.16), diffuse growth pattern (cOR 7.07; 95% CI, 2.60-19.20) and location on upper lip (cOR 10.25; 95% CI, 1.11-94.36), auricle (cOR 5.35; 95% CI, 1.55-18.45), neck (cOR 6.15; 95% CI, 1.06-35.74) or in retro- (cOR 10.25; 95% CI, 1.23-85.52) or preauricular (cOR 6.15; 95% CI, 1.69-22.38) region were associated with increased risk of metastasis with cORs. On the other hand, AK or cSCCIS among primary tumor or confirmedly preceding primary tumor was associated with lower risk of metastasis (cOR 0.40; 95% CI, 0.21-0.76). Regarding staging systems, that were not included in adjusted model due to multicollinearity, AJCC8 applied for head and neck tumors provided unlinear elevation in metastasis risk; T2 (cOR 2.05; 95% CI, 0.52-7.99), T3 (cOR 14.32; 95% CI, 6.88-29.79), T4a-T4b (cOR 10.00; 95% CI, 2.12-47.18). Findings were coaxial although with higher cORs regarding BWH staging; T2a (cOR 6.53; 95% CI, 2.98-14.29), T2b (cOR 25.50; 95% CI, 11.30-57.52), T3 (cOR 24.15; 95% CI, 5.47-106.59). (I, Table III)

There were no statistically significant patient-related metastasis risk elevating factors except for 5α-reductase inhibitor usage when applied to whole patient

material (cOR 3.73; 95% CI, 1.11-12.55), but in more appropriate approach taking only males into account no statistical significance was reached (cOR 3.28; 95% CI, 0.95-11.28). On the contrary, prior use of aspirin (cOR, 0.45; 95 % CI, 0.24-0.84) as well as isosorbide mono-/dinitrate (cOR, 0.32; 95% CI, 0.13-0.78) were associated with lower risk of metastasis with statistical significance. Aspirin and isosorbide mono-/dinitrate usage in combination lowered the risk of metastasis even more drastically (cOR, 0.08; 95 % CI, 0.01-0.60). Further aspirin and isosorbide were examined in the manner of four classes (aspirin alone, isosorbide alone, aspirin and isosorbide in combination, neither aspirin or isosorbide) with "neither aspirin or isosorbide" as reference class the cORs were following: aspirin alone (cOR, 0.6; 95% CI, 0.3-1.2), isosorbide alone (cOR, 0.6; 95% CI, 0.2-1.7) and aspirin and isosorbide in combination (cOR, 0.1; 95% CI, 0.0-0.5). (I, Table II)

5.6.2   Cancer-associated fibroblasts (CAFs) (II)

Invasive UV-induced primary non-mcSCCs (n=110) and mcSCCs (n=33) were compared on tumor level, in order to elucidate the feasibility of CAF markers in the assessment of metastasis risk of cSCC. The proportion and distribution of CAF marker positive cells in non-mcSCCs and mcSCCs is demonstrated in example mfIHC images (II, Figure 3A). PDGFRβ positive stromal cells were more prevalent in mcSCCs than in non-mcSCCs in both panel 1 (p=0.004) (II, Figure 3B) and panel 2 (p=0.006) (II, Supplementary Figure S3C) stainings. Also FAP positive stromal cells were more commonly detected in mcSCCs than in non-mcSCCs (p=0.026) (II, Figure 3B). Regarding stromal PDGFRα, αSMA, SPARC, POSTN, VIM and Col1 expression, no statistically significant differences were detected between the two tumor cohorts (II, Supplementary Figure S3A). Findings were similar in epithelial compartments with PDGFRβ and FAP positive cells being more prevalent in mcSCCs than non-mcSCC (II, Supplementary Figure S3B, D). POSTN positive cells were more frequent in mcSCCs than in non-mcSCCs in epithelial compartment, unlike in stromal compartment (II, Supplementary Figure S3B).

Binary logistic regression analysis was carried out to elucidate the value of different CAF markers in comparison to clinical and histopathological parameters in the prediction of metastasis risk (II, Table I and Supplementary Table SIII). This approach supported the observation, that increased risk of metastasis is associated with higher frequency of stromal PDGFRβ positive fibroblasts and was validated with both mfIHC panels (II, Table I). Respectively, elevated FAP was associated with metastasis (II, Table I). In epithelial compartments, FAP and POSTN expression was associated with increased risk of metastasis (II, Table I).

Clark's level positively correlated with FAP, stromal PDGFRβ and epithelial POSTN (II, Supplementary Table SIV). On the other hand, PDGFRα expression correlated negatively with tumor diameter, invasion beyond fat and more advanced BWH tumor stage (II, Supplementary Table SIV).

5.6.3   Artificial intelligence (AI) (III)

In substudy III, we utilized input data from 45 whole slide images (WSIs) representing mcSCCs, except for rapid metastasis -AI-model, in which rapid metastasis subcohort of 22 WSIs was used, and 59 WSIs representing non-mcSCCs. Tumor characteristics of the rapid metastasis -AI-model are shown in table (III, Table I).

In single tile -AI-model, a slide level AUROC of 0.689 was reached at best. However, the model could not reliably reproduce the results between different sampling of folds due to dramatical overfitting to the training data despite heavy regulation and data augmentation. Invasive front -AI-model produced inferior results to the single tile -AI-model with tile level AUROC of 0.629 at best whether tiles inside the annotation, outside the annotation or both were taken into account. On stack tile level, multi-tile -AI-model did not produce more convincing results (AUROC 0.672 at best).

From the beginning, rapid metastasis -AI-model generated most convincing results and tile level AUROCs of 0.754 – 0.814 were ultimately reached depending on the fold (III, Figure 1A). An average AUROC of 0.747 on slide level was achieved (III, Figure 1B). Slide level results visualizing summary confusion matrices show that sensitivity of the rapid metastasis -AI-model was 64%, specificity 76% and accuracy 73% (III, Supplementary Figure S1). Thus, the rapid metastasis -AI-model was utilized in the creation of RFMs and in survival analyses.

Next, we evaluated whether traditional histopathological features would predict the risk of metastasis better than the rapid metastasis -AI-model and whether the predictive power of the model could be explained by histopathological features. Pearson correlation was conducted taking into account every variable in III, Table I. Regarding AI prediction, the highest correlation was 0.329 with AJCC-8 tumor staging and second highest 0.256 with BWH tumor staging (III, Supplementary Table SI). In comparison, Pearson correlations regarding prediction by pathologist were higher with highest correlation of 0.494 with BWH staging system and second highest of 0.415 with Clark's level (III, Supplementary Table SI). Thus, it can be concluded that AI prediction did not strongly rely on any of the classical clinicopathological variables but seemed to be based on other morphological features of the tumor.

Logistic regression analysis was conducted and AUROCs created to evaluate the metastasis risk by variable and to visualize the classification power of the rapid-metastasis -AI-model. Rapid metastasis -AI-model provided the slide level AUROC of 0.747 as described above and further an cOR of 5.63. In comparison, pathologist was able to reach an AUROC of 0.694 and a cOR of 5.71. Clark's level provided higher cOR (13.75) and AUROC (0.788) than invasion beyond fat, but was inferior to diameter, which provided highest AUROC of 0.804 within the single clinicopathological variables. AJCC-8 and BWH tumor staging systems provided slightly higher AUROCs (0.816 and 0.818, respectively), but in logistic regression analysis the increase of risk was non-linear. (III, Table II)

5.7 Prognosis (I-III)

5.7.1 Non-metastatic and metastatic cSCCs (I-III)

Regarding the research material utilized in substudy I, median follow-up time was 64.0 (IQR 25-119) months in non-metastatic patient cohort and 32.0 (IQR 17-67) months in metastatic patient cohort calculated from the initial diagnosis of first cSCC (I, Table SVI). Follow-up ended for 101 (80.8%) patients in non-metastatic and for 60 (73.2%) patients in metastatic cohort due to exitus (I, Table SVI). For 43 (71.7% of the dead) patients in metastatic cohort the underlying cause of death was deduced to be cSCC (I, Table SIII). For the patients alive, the follow-up ended at the time of last contact to health care provider.

When calculated from the time of the initial diagnosis of first cSCC, 1-, 2-, 3-, 4- and 5-year OS estimates were 92.8%, 76.8%, 66.4%, 60.0% and 52.0% for the patients in non-metastatic and respectively 92.7%, 66.5%, 54.8%, 43.6% and 37.2% for the patients in metastatic cohort. At 5-year the difference between cohorts was statistically significant (p=0.028). Regarding metastatic cohort when calculated from the time of the initial diagnosis of first primary mcSCC the survival estimates were 92.7%, 63.8%, 50.5%, 38.9% and 30.7%. Finally, when survival for metastatic cohort patients was calculated from the initial diagnosis of first metastasis estimates were 68.0%, 43.4%, 31.6%, 31.6% and 31.6% respectively. Regarding the 3 patients with two mcSCCs the calculations were performed from the first diagnosed primary mcSCC and first diagnosed metastasis. (I, Table SVI, Figure 3)

In the analysis of poor prognosis-associated factors in mcSCC there were 81 out of 85 (95.3%) mcSCCs eligible for the analysis based on the determinability of CR. Poor prognosis (that is CR not achieved) was associated with the number of nodal metastases of 3 or more (aOR, 10.16; 95 % CI 2.19-47.14) and ENE (aOR,

8.19; 95 % CI, 1.79-37.52). None of the patients aged 90 or older at the time of the diagnosis of primary mcSCC achieved complete response. (I, Table IV)

In substudy II, prognostivity of clinicohistopathological features of primary cSCCs were analyzed in comparison with CAF-associated markers with uni- and multivariate Cox regression analyses. Notably, these analyses apply only on part of the original cSCC cohorts characterized in substudy I and utilize DSS. Every feature was processed as binary. The prognostivity in univariate analysis was highest regarding diameter, followed by Clark's level, diffuse growth pattern, BWH primary tumor staging and invasion beyond fat (II, Table II). In multivariate analysis, primary tumor diameter was the feature associated with worse prognosis with statititical significance in addition to either CAF107 (PDGFRα-/PDGFRβ+/FAP+) or CAF6 (PDGFRβ+/FAP+/PDGFRα-/αSMA-) CAF phenotype (II, Table II).

Furthermore, in substudy III survivals in slow and rapid metastasis cohorts, as well as in non-metastatic cohort were analyzed taking both OS and DSS into account with Kaplan-Meier method. Notably, these analyses apply only to part of original cohorts characterized in substudy I. Survival was better in slow metastasis cohort up to approximately 4 years until slow metastasis cohort reached the level of rapid metastasis cohort with 50% alive at 1.2 years for rapid metastasis and 3.4 years for slow metastasis cohort when OS was considered and 1.3 and 4.1 years, respectively, when DSS was considered (III, Figure S2). Furthermore, there were distinct differences between cohorts representing patients with non-mcSCCs and mcSCCs with 5-year DSS estimate of 98% in non-mcSCC patient cohort, 39% in rapid-metastasis cohort and 38% in slow metastasis cohort respectively (III, Figure S2B). Regarding associations between clinicohistopathological features and survival, the prognostivity of diameter, Clark's level and histopathological grade were visualized (III, Figure 3A-B). As far as DSS was considered, the prognostivity of diameter and Clark's level were superior to grade with 5-year DSS estimate of 98% for patients with primary cSCC less than 10mm in diameter, 40% with primary cSCC at least 30mm in diameter, 97% with primary cSCCs with Clark 2-4 invasion, 64% with primary cSCC with Clark 5 invasion, 86% with grade 1 primary cSCC and 72% with grade 3 primary cSCC (III, Figure 3B).

5.7.2 CAFs and prognosis (II)

In substudy II, the prognostic power of CAF markers as well as clinical and histopathological parameters was evaluated with Cox regression analysis and Kaplan-Meier survival estimates regarding patients with UV-cSCCs (non-mcSCCs and mcSCCs) on tumor level. In univariate Cox regression analysis, high frequencies of either PDGFRβ positive or POSTN positive stromal cells correlated

with poor prognosis (II, Table II and Supplementary Table SV). Correspondingly, the prognosis of patients with UV-cSCC with high stromal PDGFRβ or high stromal POSTN expression was worse with log-rank test (II, Figure 3C and Supplementary Figure S4A). Whereas PDGFRβ expression, solely in stromal compartment, was associated with worse prognosis, POSTN expression also in epithelial compartment had prognostic significance (II, Table II).

5.7.3   Prediction by AI and prognosis (III)

In substudy III, survival analyses were conducted to further evaluate the discriminative power of AI from prognostic point of view. The prognostic power of discrimination by both AI and dermatopathologist was analyzed. AI and pathologist provided nearly similar OS and DSS prediction with 5-year survival estimate of 92% for patients with cSCC categorized by AI as non-metastatic, 93% for patients with cSCC categorized by pathologist as non-metastatic, 61% for patients with cSCC categorized by AI as metastatic and 62% for patients with cSCC categorized by pathologist as metastastic (III, Figure 2A-B).

5.7.4   Treatment modalities and prognosis (I)

Regarding the research material utilized in substudy I, 81 (95.3%) primary mcSCCs were surgically excised and 31 (36.5%) were treated with RT in combination of excision and/or systemic treatment or alone (I, Table SIV). Respectively, considering 78 (91.8%) mcSCCs, at least one of its metastases was surgically excised and 55 (64.9%) cases treated with RT (I, Table SIV). In 11 (12.9%) cases of mcSCC, patient received systemic treatment (I, Table SIV). Pharmaceuticals used comprehended cisplatin, fluorouracil, paclitaxel, cetuximab and cemiplimab, which were used either alone or in different combinations (I, Table SIV). Regarding treatment modalities there were no statistically significant differences between the cohorts of CR achieved in logistic regression analysis of prognosis. (I, Table SV, Figure S3)

5.8   CAF subsets in the progression, metastasis and prognosis of cSCC (II)

In substudy II, also multimarker combinations representing distinct CAF subsets (II, Supplementary Table SII) were analyzed as markers for invasion, metastasis and prognosis.   CAF107   (PDGFRα-/PDGFRβ+/FAP+)   and   CAF6 (PDGFRβ+/FAP+/PDGFRα-/αSMA-) subsets among others were associated with invasion (II, Figure 4A). Further, there were 16 subsets associated with metastasis

including CAF107 and CAF6 (II, Table 1 and Figure 4B). CAF107 and CAF6 in addition to seven subsets were associated with worse prognosis in univariate Cox regression analysis (II, Table 2). Each of these subsets was included separately in multivariate analysis with clinical features associated with metastasis, but only CAF107 and CAF6 subsets were associated with prognosis with statistically significant adjusted HRs (II, Table 1 and Figure 4C). The prognostic power of these CAF subsets is also visualized with Kaplan-Meier survival estimate curves (II, Figure 4D).

Positive correlation between primary tumor diameter, poorer differentiation, higher Clark's level and CAF107 phenotype was noted (II, Supplementary Table SIV). CAF6 phenotype, in turn, correlated negatively with advanced age and positively with poorer differentiation and higher Clark's level (II, Supplementary Table SIV).

Gene expression data from TCGA was analyzed for validation to determine whether individual CAF markers PDGFRα, PDGFRβ and FAP, included in CAF107 subset, have prognostic value in other SCCs. PDGFRβ expression was observed to associate with worse prognosis in lung SCC (II, Supplementary Figure S4D). Furthermore, FAP expression was associated with worse prognosis and PDGFRα expression with better prognosis in HNSCC (II, Table 2 and Supplementary Figure S4D).

5.9 Risk factor models (II & III)

### 5.9.1 CAF subset utilizing risk factor model (II)

Regarding substudy II, a RFM was created to further visualize the usability of CAF phenotype in the risk stratification of metastasis and poor prognosis. Created RFM regarded fibroblast phenotype (highest quartile of proportion of positive stromal cells), Clark's level 5 and tumor diameter at least 30mm as risk factors based on findings in substudy I. In logistic regression analysis CAF107 (PDGFRα-/PDGFRβ+/FAP+) RFM provided an OR of 61.00 (95% CI 9.59-388.13) for cSCCs with 3 RFs surpassing comparable ORs of AJCC-8 T4a-T4b (OR 26.33; 95% CI 4.85-143.12) and BWH T3 (OR 24.00; 95% CI 4.41-130.62) (Table IV). The association with prognosis was also visualized with Kaplan-Meier curves, in which after 5 years of the initial diagnosis of primary tumor, patients with cSCC with 0 risk factors (RFs) had excellent survival estimate (DSS 100%) and patients with cSCC with 3 RFs poor prognosis (DSS 31.8%) (Figure 6). Findings were coaxial regarding CAF6 (PDGFRβ+/FAP+/PDGFRα-/αSMA-) RFM with corresponding OR of 62.00 (95% CI 10.55-364.22) and excellent prognosis for patients with cSCC with 0 RFs (DSS 100%) and poor for patients with cSCC with 3 RFs (DSS 37.0%) (Table IV and Figure 6).

Table IV. Logistic regression analysis of conventional characteristics and CAF subset risk factors models in association with the risk of metastasis.

Variable	Non-mcSCC pos/total	mcSCC pos/total	Risk of metastasis by variable cOR (95% CI)	P value
Age				
1st quartile (27-69y), n (%)	27/110 (24.5)	9/33 (27.3)	1.67 (0.52-5.30)	0.387
2nd quartile (69-79y), n (%)	23/110 (20.9)	13/33 (39.4)	2.83 (0.93-8.57)	0.066
3rd quartile (79-84y), n (%)	30/110 (27.3)	5/33 (15.2)	0.83 (0.23-3.03)	0.782
4th quartile (85-102y), n (%)	30/110 (27.3)	6/33 (18.2)	1 (ref)	
Sex				
Male, n (%)	69/110 (67.2)	23/33 (69.7)	1.37 (0.59-3.16)	0.464
Female, n (%)	41/110 (37.3)	10/33 (30.3)	1 (ref)	
Location				
Head and neck, n (%)	92/110 (83.6)	24/33 (72.7)	1 (ref)	
Upper limb, n (%)	8/110 (7.3)	4/33 (12.1)	1.92 (0.53-6.90)	0.320
Lower limb, n (%)	6/110 (5.5)	5/33 (15.2)	3.19 (0.90-11.36)	0.073
Torso, n (%)	3/110 (2.7)	0/33 (0.0)	0.00 (0.00-NA)	0.999
Missing, n (%)	1/110 (0.9)	0/33 (0.0)		
Diameter				
<10mm, n (%)	36/110 (32.7)	2/33 (6.1)	1 (ref)	
10-19.9mm, n (%)	44/110 (40.0)	7/33 (21.2)	2.86 (0.56-14.65)	0.206
20-29.9mm, n (%)	11/110 (10.0)	2/33 (6.1)	3.27 (0.41-26.01)	0.262
≥30mm, n (%)	19/110 (17.3)	22/33 (66.7)	20.84 (4.42-98.25)	**<.001**
Differentiation				
Good, n (%)	59/110 (53.6)	5/33 (15.2)	1 (ref)	
Moderate, n (%)	39/110 (35.5)	20/33 (60.6)	6.05 (2.10-17.47)	**0.001**
Poor, n (%)	12/110 (10.9)	8/33 (24.2)	7.87 (2.19-28.24)	**0.002**
Necrosis among primary tumor				
No, n (%)	101 /110 (91.8)	26/33 (78.8)	1 (ref)	
Yes, n (%)	7/110 (6.4)	6/33 (18.2)	3.33 (1.03-10.76)	**0.044**
Missing, n (%)	2/110 (1.8)	1/33 (3.0)		
Clark's level				
2-4, n (%)	78/110 (70.9)	5/33 (15.2)	1 (ref)	
5, n (%)	30/110 (27.3)	27/33 (81.8)	14.04 (4.95-39.84)	**<.001**
Missing, n (%)	2/110 (1.8)	1/33 (3.0)		
Invasion beyond fat				
No, n (%)	96/110 (87.3)	16/33 (48.5)	1 (ref)	
Yes, n (%)	14/110 (12.7)	16/33 (48.5)	6.86 (2.81-16.72)	**<.001**
Missing, n (%)	0/110 (0.0)	1/33 (3.0)		
Diffuse growth pattern				
No, n (%)	104/110 (94.5)	28/33 (84.8)	1 (ref)	
Yes, n (%)	5/110 (4.5)	5/33 (15.2)	3.71 (1.00-13.74)	**0.049**
Missing, n (%)	1/110 (0.9)	0/33 (0.0)		
AJCC-8				
T1, n (%)	79/110 (71.8)	5/33 (15.2)	1 (ref)	
T2, n (%)	10/110 (9.1)	3/33 (9.1)	4.74 (0.98-22.90)	0.053
T3, n (%)	18/110 (16.4)	19/33 (57.6)	16.68 (5.50-50.62)	**<.001**
T4a-T4b, n (%)	3/110 (2.7)	5/33 (15.2)	26.33 (4.85-143.12)	**<.001**
TX, n (%)	0/110 (0.0)	1/33 (3.0)		
BWH				
T1, n (%)	72/110 (65.5)	5/33 (15.2)	1 (ref)	
T2a, n (%)	23/110 (20.9)	9/33 (27.3)	5.64 (1.72-18.52)	**0.004**
T2b, n (%)	12/110 (10.9)	13/33 (39.4)	15.60 (4.70-51.74)	**<.001**
T3, n (%)	3/110 (2.7)	5/33 (15.2)	24.00 (4.41-130.62)	**<.001**
TX, n (%)	0/110 (0.0)	1/33 (3.0)		

Variable	Non-mcSCC pos/total	mcSCC pos/total	Risk of metastasis by variable cOR (95% CI)	P value
CAF107 risk factor model				
0 RF, n (%)	64/110 (58.2)	2/33 (6.1)	1 (ref)	
1 RF, n (%)	23/110 (20.9)	5/33 (15.2)	6.78 (1.29-35.76)	**0.024**
2 RF, n (%)	18/110 (16.4)	17/33 (51.5)	30.50 (6.35-146.56)	**<.001**
3 RF, n (%)	3/110 (2.7)	8/33 (24.2)	61.00 (9.59-388.13)	**<.001**
Missing, n (%)	2/110 (1.8)	1/33 (3.0)		
CAF6 risk factor model				
0 RF, n (%)	62/110 (56.4)	2/33 (6.1)	1 (ref)	
1 RF, n (%)	18/110 (16.4)	5/33 (15.2)	8.61 (1.54-48.17)	**0.014**
2 RF, n (%)	23/110 (20.9)	15/33 (45.5)	20.22 (4.29-95.36)	**<.001**
3 RF, n (%)	5/110 (4.5)	10/33 (30.3)	62.00 (10.55-364.22)	**<.001**
Missing, n (%)	2/110 (1.8)	1/33 (3.0)		

AJCC-8 = 8th edition of American Joint Committee on Cancer tumor staging, BWH = Brigham and Women's Hospital tumor staging, CAF107/CAF6 risk factor model = risk factor model considering high proportion of CAF107/CAF6 fibroblasts (Q4), Clark's level 5 and primary tumor diameter at least 30mm as risk factors, CI = confidence interval, cOR = crude odds ratio, mcSCC = metastatic primary cutaneous squamous cell carcinoma, non-mcSCC = non-metastatic cutaneous squamous cell carcinoma, pos = positive, RF = risk factor.
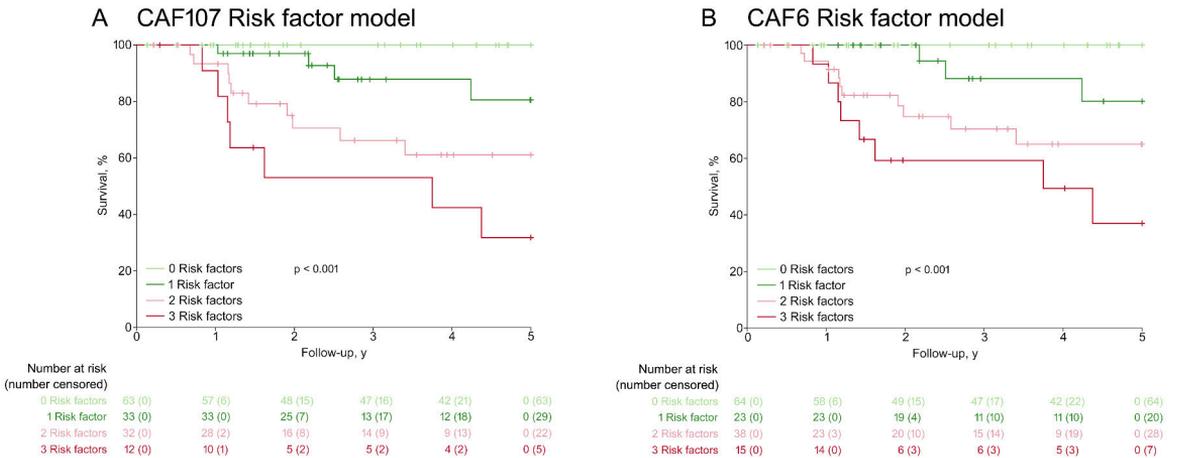
Figure 6. Prognostic power of CAF107 and CAF6 risk factor models. A risk factor model regarding high proportion of CAF107 (PDGFRα-/PDGFRβ+/FAP+) (A) or CAF6 (PDGFRβ+/FAP+/PDGFRα-/αSMA-) (B) positive CAFs (Q4 (n=36)), Clark's level 5 (n=56) and primary tumor diameter at least 30mm (n=39) as risk factors.

5.9.2 AI prediction utilizing risk factor model (III)

In substudy III, it was evaluated whether prediction by AI as metastatic could serve as an individual risk factor in multifactorial RFM and improve the clinical risk assessment. A RFM taking into account Clark's level 5 and tumor diameter ≥30mm as risk factors (conventional-RFM) was created for comparison and

already provided higher AUROC (0.862) than BWH or AJCC-8 staging systems. Next, a RFM comprehending prediction by AI as metastatic and Clark's level 5 as risk factors was created and provided even higher AUROC of 0.872. Finally, a RFM including prediction by AI as metastatic, Clark's level 5 and diameter ≥30mm as risk factors (AI-RFM) produced an AUROC of 0.917. Comparative RFMs taking prediction by pathologist into account instead of AI were created but resulted with lower AUROCs of 0.807 and 0.841 respectively. (III, Table II)

Different thresholds in AI-RFM regarding primary tumor diameter were tested. However, lower AUROC of 0.913 was reached if 20mm instead of 30mm was used as a threshold for the risk factor. In addition, if histological grade was included into the AI-RFM, the discriminative power of the RFM was decreased (AUROC 0.903) (Data not shown).

The prognostic power of conventional histopathologic parameters was evaluated and visualized in order to elucidate the feasibility of AI-RFM. Especially when DSS was considered, the survival prediction by histological grade was inferior to diameter and Clark's level (III, Figure 3A-B). This notion supports the inclusion of Clark's level 5 and diameter ≥30mm into the AI-RFM. The AI-RFM provided survival prediction with poor prognosis for patients with cSCC with 2 (5-year DSS estimate of 41.7%) or 3 (5-year DSS estimate of 40.0%) RFs and excellent prognosis for patients with cSCC with 0 (5-year DSS estimate of 100%) or 1 (5-year DSS estimate of 95.7%) RFs (III, Figure 3D and Figure 4B, D). In respect of DSS, the discriminative power of AI-RFM was superior to "conventional-RFM" (III, Figure 3D). Furthermore, the AI-RFM was compared to BWH tumor staging (III, Figure 4A-B) and to the comparative RFM utilizing prediction by pathologist instead of AI (pathologist-RFM) (III, Figure 4C-D). The discriminative power of AI-RFM was superior to both BWH tumor staging (III, Figure 4B) and pathologist-RFM (III, Figure 4D) as far as DSS was considered. The superiority is based on the discrimination by AI-RFM into good (0-1 RFs) and poor prognosis (2-3 RFs) cases and the lack of "grey zone" which turns out to be the problem related to comparative RFMs and BWH staging.

6   Discussion

6.1   Epidemiology and prognosis of cSCC (I)

The mean rate of metastasis in the study region during the 10-year period between 2004 and 2013 was 2.3%. In previous publications the rate of metastasis of cSCC has varied greatly from 0.1% up to 20.7% probably due to differences in patient and tumor selection, study center, geographical location and patient ethnicity to name but a few (Moore et al., 2005). In a prospective study by Brantsch et al. (2008) the rate of metastasis of 4% was discovered and in nation-wide English and Dutch studies rates of 1.1%-2.4% and 1.9% were reported (Tokez et al., 2022; Venables et al., 2019a). For geographical variety Karia et al. (2013) expressed a rate of metastasis of 3.0% and Schmults et al. (2013) of 3.7% in US population as well as Brougham et al. (2012) 1.9-2.6% in the population of New Zealand. It is likely that the metastatic rate in unselected Caucasian populations lies between 1-4%.

Results in substudy I substantiate that cSCCs in general population occur at advanced age with median age of 78 years at the diagnosis of first cSCC. This is in accordance with literature in which it has been established that the incidence of cSCC increases with age (Xiang et al., 2014). In a nation-wide English study, median age of 80 years at first registered cSCC was reported and higher age at mcSCC diagnosis adduced (Venables et al., 2019a). Although, there was slight indication in our study that patients who developed metastases would have been younger at the diagnosis of first cSCC, there was no difference when age at exact tumor diagnosis of each cSCC was scrutinized. Mentionable is also that 46.1% of patients under 60 years of age at the diagnosis of first cSCC were immunocompromised. The association between age and the risk of metastasis was non-linear and did not provide any univocal conclusions. However, the prognosis of mcSCC seemed to worsen almost in linear manner as CR was not achieved regarding mcSCCs diagnosed at the age of 90 or higher. This is on the contrary to what for instance Smith et al. (2018) has reported.

The results in substudy I represent male preponderance within both the patient cohorts, which is in accordance with previous observations (Que et al., 2018). When viewed on tumor level the male predominance was remarkably homologous

in both cohorts (66.4% and 67.1%). Male sex did not appear as an independent risk factor for metastasis. Although there was no statistically significant difference in male predominance between the non-metastatic and metastatic cohorts either examined on patient (p=0.139) or tumor level (p=0.910), the mean rate of metastasis during 2004-2013 was higher regarding males (2.7%) than females (1.9%). Due to the nature of this study, the last-mentioned represents the greatest of significance and is in line with nation-wide English and Dutch studies, in which 1.1% and 1.4% for women and 2.4% and 2.3% for men were reported respectively (Tokez et al., 2022; Venables et al., 2019a). It seems plausible that male sex not only increases the risk of cSCC but also the risk of metastasis development.

Furthermore, the results of substudy I show that metastasis occurs relatively shortly after the diagnosis of primary mcSCC. It has been expressed by several studies including substudy I of this thesis (84.7%) that approximately 72-90% of metastases are detected within the first two years (Bobin et al., 2018 (83%); Cherpelis et al., 2002 (72%); Dinehart & Pollack, 1989 (90%); Rowe et al., 1992 (75%); Venables et al., 2019a (85.2%)). Our study is the largest after Venables et al. (2019a) (1566 metastatic cases), who reported a metastasis detection rate of 85.2% within 2 years. Additionally, for example Brantsch et al. (2008) reported a rate of 73.1% within one year already and Tokez et al. (2022) a median time to metastasis of 1.5 years calculated from the diagnosis of first cSCC of the patient.

Survival of patients with mcSCC has generally been regarded poor, but there is variety in survival estimates depending on source. We reported the OS in substudy I due to its unambiguous nature. Two-, 3-, and 5-year survival rates among studies reporting OS from the metastasis detection have varied between 50-66% (Bobin et al., 2018 (50%); Hirshoren et al., 2017 (66%)), 29-46% (Givi et al., 2011 (33%); Venables et al., 2019a (29% female, 46% male)) and 30-50% (Bobin et al., 2018 (35%); Givi et al., 2011 (30%), Hirshoren et al., 2017 (50%)) respectively. Venables et al. (2019a) reported significantly lower 3-year OS for females (29% vs 46%) and interestingly a remarkably low 3-year survival from first primary cSCC of 65% for males and 68% for females in unselected study material including only 1.1-2.4% of mcSCCs. Findings in substudy I are in line but represent the poorer end with OSs of 43%, 31% and 31% at 2, 3 and 5 years after the metastasis detection. Regarding the poor prognosis of the patients with non-metastatic disease in our study the 3-year survival of 62.1% is actually not that far from what Venables et al. (2019a) reported and on part explained by patient selection and on the other hand can be seen strengthening the assumption of cSCC as a general poor prognosis marker. It can also be speculated whether this is related to the number of other malignant diseases. On sidenote, there was no statistically significant difference between sexes in OS regarding either patients with mcSCC (p = 0.897) or non-mcSCC (p = 0.228) (not shown in tables).

On the other hand, in substudies II and III DSSs were determined and in substudy III 5-year DSS estimate of 98% in non-mcSCC patient cohort, 39% in rapid-metastasis cohort and 38% in slow metastasis cohort were noted respectively. The DSS regarding metastatic patient cohorts is poor in comparison to for instance Tokez et al. (2022), that reported a 5-year DSS of 79.1% for patients with mcSCC. This could partly be explained by later study period between years 2007-2008 in comparison to ours and by the requirement of C44 death certificate for the definition of disease specific death (Tokez et al., 2022). If this would have been the criteria in our study, the DSS would have undoubtedly been better.

6.2 Factors associated with the risk of metastasis (I-III)

6.2.1 Clinical and histopathological (I-III)

Based on several points of view explored in the substudy I, it is indicated that the presence of precursor lesions does not increase the risk of mcSCC but instead lowers it. This was evident when diagnosis of AK at any time point was examined as only 36.6% of patients with metastatic compared to 53.6% of patients with non-metastatic disease had AK diagnosed (p=0.016). Additionally, regarding mcSCCs, only 29.4% had AK or cSCCIS diagnosed prior to or on the day of primary mcSCC diagnosis. We also investigated the presence of precursor lesions among the primary cSCC and whether there was preceding histopathologically confirmed precursor on the same exact location. In this analysis, too, there was significant difference between the tumor cohorts as 30.3% of non-mcSCCs and only 14.1% of mcSCCs had preceding precursor or precursor among primary tumor (p=0.006). When examined from another standpoint 15.4% of tumors with preceding precursor or precursor among primary tumor were metastatic compared to 31.5% of tumors with no precursor among or preceding.

Onwards, we discovered that in 84.7% of mcSCCs the metastatic primary tumor was the first cSCC diagnosed for named patient. This could be hypothesized for one's part to be due to enhanced surveillance of patients with history of cSCC. On the other hand, for instance Levine et al. (2015) has earlier indicated that patients with more than one cSCC are in elevated risk for nodal metastasis. Our finding scrutinizes this from slightly different standpoint and indicates that in remarkably high percentage of cases the first diagnosed cSCC is metastatic. Further, as far as metastatic primary tumor diameter prior to or on day of metastasis detection is concerned 8.5% of mcSCCs first in ordinal number were <10mm compared to 7.7% of second or later in ordinal number. Respectively 36.6% of first and 7.7% of second or later were 10-19.9mm in diameter and 16.9 %

of first and 46.2% of second or later were at least 40mm in diameter. This in fact indicates that the finding should not be simply explained by larger tumors in patients with mcSCC first in ordinal number, which could be speculated to result from inferior surveillance and longer delay for medical attention. Based on above-mentioned notions the history of previous cSCCs and precursor lesions should not be reckoned as major high-risk factor for metastasis when individual cSCC is evaluated. From another standpoint 39.3% patients with only one cSCC, 39.6 % of patients with 2-9 cSCCs and 50 % patients with ten or more cSCCs developed cSCC metastasis.

It has been stated that tumor diameter more than 20mm triples the rate of metastasis compared to lesions less than 20mm in diameter (Rowe et al., 1992). In more recent meta-analysis by Thompson et al. (2016) it was evidenced that tumor diameter exceeding 20mm was statistically significant risk factor for metastasis but with lower RR than Breslow thickness 2mm or 6mm or invasion beyond subcutaneous fat which was the factor reported with highest RR. On the other hand, diameter exceeding 20mm was the factor with highest RR for disease-specific death. In our study it was also obvious that increasing horizontal growth elevates the risk of metastasis but not in linear manner when examined in four diameter groups. If distribution between tumors was made at 20mm (<20mm vs. ≥20mm) the unadjusted OR would have been 5.4 (95% CI, 3.2-9.2). Further, in adjusted model only diameter between 20 and 29.9mm was associated with elevated risk of metastasis.

In substudy I, Clark's level 5 was associated with increased risk of metastasis with higher aOR than diameter 20-29.9mm or invasion beyond subcutaneous fat. Anatomical invasion depth interpreted as Clark's level in cSCC has been studied only in a few published papers. Already Breuninger et al. (1990) in their study of 673 cSCCs including 22 mcSCCs took into account both anatomical and quantitative invasion depth and discovered that all mcSCCs except for one represented Clark 5 invasion (9 with invasion into subcutaneous fat and 12 with invasion beyond fat). However, the Clark's level 5 invasion was more often met also in non-mcSCCs than in our study. This may partly be due to the fact that in that study only tumors treated in named tertiary care center were included instead of every applicable tumor of every patient as was the case in this study. Concordantly, Cherpelis et al. (2002) indicated in their study dealing 25 mcSCCs and 175 non-mcSCCs that tumors with Clark's level 5 invasion were statistically significantly more likely to metastasize. They observed in addition that primary tumors with small tumor nests and infiltrating strands were more prone to metastasis. This finding is comparable to our finding of diffuse growth pattern as metastasis risk elevating factor with statistically significant cOR and has not been examined or at least published in elsewhere and could be worth further

investigation. Moreover, in respect of invasion Brantsch et al. (2008) discovered that only tumors with Breslow depth greater than 2mm were associated with significant risk of metastasis as there were no tumors 2mm or less in thickness that metastasized. In similar manner Breuninger et al. (1990) reported metastasis rate of 0% for tumors with 2mm or less in depth, 4.5% for tumors between 2 and 6mm and 15% for tumors with more than 6mm in depth. In our study metastasis rate for tumors with Clark 2-4 invasion was 4.5% and for tumors with Clark 5 invasion 58.8%, however the proportions of cohorts do vary between studies. In conclusion, our findings strengthen the perception that invasion depth is more predictive in risk stratification of cSCC than horizontal tumor diameter. However, more evidence is needed in order to evaluate whether Clark's level or Breslow depth is the most prognostic indicator of invasion in the case of cSCC.

From a slightly different point of view, but underlining the notions above, in substudy II features associated with poor prognosis in univariate analysis comprehended diameter, Clark's level, diffuse growth pattern, BWH primary tumor staging and invasion beyond fat. In multivariate analysis, primary tumor diameter was the feature associated with worse prognosis with statistitical significance. In substudy III, the prognostivity of diameter and Clark's level were superior to histopathological grade.

Necrosis among primary tumor seemed to predict metastasis development with statistically significant cOR. There are no publications scrutinizing necrosis from the metastasis risk elevating point of view in cSCC, but the finding is concordant with notions of necrosis in cancer in general as for instance it has been evidenced that primary tumor necrosis is a hallmark of aggressive endometrial cancer and is associated with poor prognosis and via activated angiogenesis also metastatic spread (Stefansson et al., 2006; Bredholt et al., 2015).

Moderate and poor histologic differentiation appeared as metastasis risk elevating factors in this study with statistically significant cORs. Previously it has been addressed that poor differentiation doubles the risk of metastasis (Brantsch et al., 2008) and in meta-analysis by Thompson et al. (2018) poor differentiation was the most influential metastasis predicting factor after invasion depth and diameter. Local recurrence was another statistically significant risk elevating factor for metastasis in substudy I with cOR of 3.97 and metastatic rate of 53.8 %. Previously it has been stated that the prognosis of locally recurrent tumors is significantly worsened with up to 48% risk of metastasis (Rowe et al., 1992; Cherpelis et al., 2002). It seems substantiated that after local recurrence cSCC should be considered high risk regardless of other features.

Every (4/4) primary tumor developed in orbital region (practically upper or lower lid) was metastatic in substudy I. Further in logistic regression analysis lower lip was by far the location associated with highest risk of metastasis and in adjusted

logistic regression analysis in addition to forehead represented location associated with increased risk of metastasis. Interestingly, for instance Thompson et al (2018) in their meta-analysis reported lower RR for lip than ear or temple, although there have been studies indicating clearer metastasis risk elevation regarding lip (Rowe et al., 1992; Brougham et al., 2012). We distinguished preauricular region from the rest of the cheek as it seemed already in the beginning of data collection to differ from the rest of the cheek in the rate of metastasis. As a result, cheek excluding preauricular region was the only primary tumor location associated with significantly reduced risk of metastasis with cOR. It seems substantiated that preauricular region resembles risk-wise auricle and retroauricular region rather than the rest of the cheek. Count-adjusted male to female ratio of every primary tumor location was calculated revealing clear female predominance concerning forehead (0.3), lower lip (0.3) and cheek (0.3) as well as male predominance concerning especially auricle (9.0) and trunk (6.0). Similar findings regarding location on ear has been discovered for instance by Venables et al. (2019a) with absolute male to female ratio of 23.0.

SOTRs represented only a small fraction in substudy I (3.9%) as did immunocompromised individuals in general (20.3%). There was no significant difference between non-metastatic and metastatic cohorts regarding immunosuppression in general or in any specific factors causing immunosuppression. Additionally, in logistic regression analyses immunocompromised individuals in general or SOTRs on their own did not have higher risk of metastasis or poor prognosis. This is in contradiction with what has been reported in studies including larger cohorts and reviews of immunocompromised study subjects (Lam et al., 2018; Madeleine et al., 2017; Martinez et al., 2004; Sahovaler et al., 2019), although the association between the metastasis risk and immunosuppression is not that straightforward as Genders et al. (2019) have stated. However, if we view the small subcohort of immunocompromised on their own (n=42), 38.1% of patients with immunosuppression developed mcSCC in this study. Similarly, 25% of SOTRs developed metastatic disease. In studies with more applicable research design it has been represented that the metastatic rate among SOTRs would be 7-8% (Burton et al., 2016). On note, immunosuppression was not by any means main focus in substudy I, however it can be concluded that mcSCC is far from being exclusively limited to immunocompromised patients.

The finding in substudy I indicating that aspirin and isosorbide mono-/dinitrate especially in combination lower the risk of metastasis is intriguing and novel concerning cSCC. Interestingly, in relation to comorbidities, there was practically no difference between patient cohorts in respect of indicative coronary artery disease, stroke or even hypertension. It is substantiated in clinical studies that

aspirin use enhances the survival of patients with at least colorectal cancer (Albandar et al., 2018; Dubé et al., 2007; Rothwell et al., 2010; Rothwell et al., 2012). With respect to isosorbide mono- and dinitrate, Pipili-Synetos et al. already in 1995 indicated in murine Lewis Lung carcinoma-model that nitric oxide (NO)-releasing isosorbide mono- and dinitrate inhibit angiogenesis, tumor growth and most interestingly also metastasis. Further, Wang et al. (2015) evidenced on human colon cancer cells that aspirin enhances NO release from isosorbide mononitrate and that aspirin and isosorbide mononitrate synergistically inhibit the growth of colon cancer cells and induce apoptosis. In xenograft model above mentioned drug combination possessed potent anti-tumor effect (Wang et al., 2015). Underlying mechanisms of action included the induction of cell apoptosis, the activation of the NO pathway and the inhibition of the Wnt pathway (Wang et al., 2015). In respect of cSCC it has been postulated that NSAIDs would act in chemoprotective manner (Muranushi et al., 2015), however Pandeya et al. (2019) did find at most weak inverse associations between infrequent aspirin use and cSCC development. Indicated metastasis risk decreasing impact of both low-dose aspirin and isosorbide mono- and dinitrate especially in combination constitute a basis for further studies.

In addition to NSAIDs, nicotinamide and retinoids (acitretin and isotretinoin) have been used as systemic agents for chemoprevention i.e. to prevent and reduce the risk of developing another cSCC for patients at risk of developing numerous and/or aggressive cSCC (Stratigos et al., 2020a). However, the evidence of clinical chemoprotective effectiveness of retinoids and nicotinamide is limited (Stratigos et al., 2020a). Furthermore, the usage of retinoids is limited by teratogenicity and dose-related toxicity (Stratigos et al., 2020a). Antioxidants, phytochemicals, vitamin D and selenium have also been studied as chemoprotective agents but the evidence is inconclusive (Stratigos et al., 2020a). Topical 5-fluorouracil especially in combination with calcipotriol has been shown with limited evidence to reduce the risk of cSCC requiring surgery by 75% (Rosenberg et al., 2019). On the other hand, topical tretinoin has been shown to be ineffective at reducing the risk of cSCC, BCC and AK (Weinstock et al., 2012).

There were no surprises in regard of actual metastases. It has been established that first metastasis virtually always affects regional lymph node (Brantsch et al., 2008), and that most common sites of nodal metastases are head and neck nodes or parotid gland (Venables et al., 2019a). Findings are additionally in concordance with for instance Smith et al. (2018) who expressed in their study that ENE and the number of positive lymph nodes are statistically significant factors associated with prognosis. Our notion that remarkably high percentage of metastases is initially detected clinically underlines the difficulties in risk stratification and projection of staging studies.

6.2.2   Cancer-associated fibroblasts (CAFs) (II)

The notion of protective tumor fibrosis in CISs in general and the remarks of the development of cancers following wound generation in cancer-prone animals as well as the association between fibrosis and increased risk of cancer are hypothetically interesting bearing in mind the development and risk factors of cSCC (Dolberg et al., 1985; Dvorak et al., 1984; Folkman & Kalluri, 2004; Li et al., 2014; Park et al., 2001; Rønnov-Jessen et al., 1996; Schuh et al., 1990; Wang et al., 2013). In substudy II, we analyzed a large panel of human tissue specimens with mfIHC to elucidate the role of CAFs in tumorigenesis and progression of cSCC. Elevated expression of FAP, αSMA, and SPARC in stromal fibroblasts appear as an early event in the tumorigenesis and associate with the invasion of cSCC.

The main purpose of this study was to investigate whether distinct changes in marker expression of CAFs or certain CAF phenotypes associate with the metastasis of cSCC. It has been demonstrated that CAFs support metastasis by paracrine effect on adjacent malignant cells and that PDGF-mediated activation contributes to this (Karnoub et al., 2007; Shinagawa et al., 2013). In general, cancer cells express PDGFs, whereas PDGFRs are mainly expressed by CAFs (Kitadai et al., 2006). In addition to CAFs, PDGFRβ is expressed by perivascular cells (Östman, 2017). In colon cancer, especially CAFs derived from MSCs are associated with high PDGFRβ expression and stromal PDGFRβ is associated with vascularity, tumor stage and metastatic potential (Kitadai et al., 2006; Shinagawa et al., 2013). Further, in breast and prostate cancers PDGFRβ expression in stromal CAFs is associated with worse prognosis (Hägglöf et al., 2010; Paulsson et al., 2009). We observed that there are higher frequencies of PDGFRβ positive CAFs in mcSCCs than in non-mcSCCs. This is in accordance with observations in other cancers (Östman, 2017). Furthermore, correlations between high PDGFRβ expression and clinicohistopathological poor prognosis associated markers have been previously observed (Frings et al., 2013; Paulsson et al., 2009). Regarding different CAF subsets and the association with metastasis and poor prognosis the conjunctive feature in our study was almost invariably the PDGFRβ positivity, which underlines the pivotal role of PDGFRβ as metastasis- and poor prognosis-associated marker in cSCC.

PDGFRα is less studied, but it has been shown that PDGFRα is down-regulated upon activation of fibroblasts by TGFβ and postulated that PDGFRα expression would mark resting and growth restraining fibroblast population (Crowley et al., 2005; Östman, 2017). Interestingly, in respect of invasion, it has been discovered that during the conversion from ductal carcinoma *in situ* to invasive cancer, the activation of periglandular fibroblasts takes place and is associated with the disruption of basement membrane and downregulation of PDGFRα and

upregulation of PDGFRβ in affected CAFs (Östman, 2017). Our observation that PDGFRα expression is higher in normal skin than in UV-cSCCs is in agreement of its tumor suppressive association in different tumor types. In line with our findings, high stromal PDGFRα expression has been shown to associate with better and high stromal PDGFRβ expression with worse prognosis in non-small cell lung cancer (Kilvaer et al., 2019).

We demonstrated that FAP, a type-II transmembrane serine protease, is associated with invasion and metastasis of cSCC (Fitzgerald & Weiner, 2020). Earlier studies have shown that FAP expression is not limited to activated fibroblasts, but can be expressed also in endothelial, immunologic and malignant epithelial cells (Fitzgerald & Weiner, 2020). In fact, as it has been stated that FAP expression is associated with worse prognosis in solid tumors, the association is pronounced if FAP overexpression is found in tumor cells (Liu et al., 2015). Associations between high FAP expression and metastasis as well as worse survival have been reported regarding colorectal, pancreatic, and gastric cancers as well as melanoma (Fitzgerald & Weiner, 2020). Further, regarding cutaneous cancers, it has been noted that FAP expression is present in malignant lesions including cSCC but not in benign tumors (El Khoury et al., 2014).

High POSTN expression has been linked to poor prognosis in many cancer types including cSCC (González-González & Alonso, 2018; Lincoln et al., 2021; Xu et al., 2016b). It has also been implied that POSTN is not expressed by cSCC cells (Lincoln et al., 2021), although our findings implicate also infrequent epithelial expression in cSCC. Further, in our study high POSTN expression both in epithelial and stromal compartments was associated with worse prognosis, as has been noted for instance regarding ovarian and breast cancers (Kim et al., 2017; Sung et al., 2016).

Based on our results regarding singular markers, CAF107 (PDGFRα-/PDGFRβ+/FAP+) appears as justifiable CAF phenotype associated with both metastasis and poor prognosis in cSCC. First, PDGFRβ and FAP were clearly associated with metastasis and PDGFRβ with worse prognosis. Furthermore, there were indications of associations between FAP and worse as well as PDGFRα and better prognosis. Additionally, majority of CAF subsets associated with elevated risk of metastasis or worse prognosis included PDGFRβ positivity and to lesser extent FAP positivity and PDGFRα negativity. The justification of CAF6, although resulting with similar prognosis and metastasis risk associations as CAF107, is not as straightforward due to the ambiguous nature of αSMA in cSCC context. Regarding its association with the risk of metastasis or prognosis, it seems that αSMA could be either positive or negative without impact on metastasis risk or prognosis. Notably, utilization of the four-marker combination in CAF6, instead of the three-marker combination in CAF107, did neither provide clear added

metastasis risk nor prognosis associated value, underlining the ambiguous nature of αSMA in cSCC context. Based on the results, we propose CAF107 fibroblast subset as the most potential metastasis risk- and prognosis-associated CAF phenotype in cSCC. Reinforcing our findings, we observed that the single markers of the CAF107 subset showed also prognostic value in the TCGA data of other squamous cell carcinomas.

It has been established that RDEBSCCs differ from UV-cSCCs with respect to genomic alterations, and based on our results also in relation to CAFs. Our findings demonstrate that RDEBSCCs have low PDGFRβ and Col1 expression in comparison to UV-cSCCs. In addition to mutations in COL7A1 gene that encodes type VII collagen and causes RDEB, it has been demonstrated that TGFβ pathways are involved in the modulation of disease severity (Odorisio et al., 2014). It has been postulated that in RDEB, the activated fibroblasts resemble CAFs prior to the development of cancer (Condorelli et al., 2019). Due to the chronic inflammation and fibrosis in RDEB, the transcriptome of fibroblasts from non-tumor RDEB already resembles that of CAFs from UV-cSCC rather than normal fibroblasts, indicating that there may be stromal predisposition to cSCC development (Ng et al., 2012). It has also been demonstrated that in several cancer-prone genodermatoses, including RDEB, the transcriptional profile of primary skin fibroblasts is similar and unrelated to primary genetic defect (Chacón-Solano et al., 2019). Little is known about the PDGFR signaling in the pathogenesis of RDEB, however it has been demonstrated that PDGFRβ signaling is weakened already in RDEB fibroblasts (Martínez-Martínez et al., 2021). Our results conclude that the CAF phenotype in RDEBSCCs differs from UV-cSCCs and that low PDGFRβ expression represents the hallmark of these differences.

Poor prognosis in cSCC is almost exclusively due to metastasis and thus metastasis risk-associated markers tend to have prognostic power (Schmults et al., 2013). Notably part of the normal skin specimens utilized in substudy II were from sun exposed skin and thus these specimens present some of the features of the cSCC especially in relation to UV-radiation induced damage. With respect to epithelial expression of biomarkers investigated in this study, there are similar associations with many of the markers also in epithelial compartments, although the expression levels are significantly lower. This could at least partially be explained by the expression of these markers by cells of epithelial origin (González-González & Alonso, 2018; Kahounová et al., 2018). Notably, associations between epithelial or stromal expressions and prognosis or metastasis risk were almost invariably in accordance, which reinforces the impact of named markers. The notions regarding differences in CAF marker expressions between UV-cSCCs and metastases include elevation of POSTN as well as decrease of VIM

and provoke origin for further studies with larger study cohorts especially in respect of specimens representing metastases.

Undoubtedly, extensive evidence of the protumorigenic contribution of CAFs in variety of cancers encourages to target CAFs in therapeutic manner, although challenges are well established (Louault et al., 2020). Expectations have arisen as more precise targeting has become enabled following more in detail characterization of CAF subtypes and their specific functions (Louault et al., 2020). For instance, targeting immunosuppressive CAFs in combination with cell-mediated immunotherapies could at least hypothetically open new opportunities to fight cancer (Louault et al., 2020). Not surprisingly, there are numerous strategies already utilized to target CAFs in preclinical and clinical studies (Louault et al., 2020). CAF targeting approaches can be divided into three categories comprehending direct targeting of CAFs by eliminating them or preventing their activation, targeting CAF activity by inhibiting factors they produce and taking advantage of the tumor-tropism of MSCs to deliver anti-neoplastic molecules to tumors (Louault et al., 2020).

Direct targeting of CAFs has been hindered by the lack of specific markers and limited knowledge of CAF heterogeneity, however there have been several approaches, of which targeting FAP has been most widely used with promising efficacy and without toxicity in preclinical studies (Louault et al., 2020). For instance, sibrotuzumab, a monoclonal antibody against FAP, has been tested already 20 years ago in phase II trials without severe toxicity but with limited efficacy as have other molecules with same target (Hofheinz et al., 2003; Louault et al., 2020). A monoclonal antibody against endoglin (CD105) expressed on the surface of ECs, CAFs and MSCs has been discovered to enhance immunotherapy and tested even in phase III trials as anti-angiogenic agent (Rosen et al., 2014). CAF activation reversing or preventing strategies include targeting of PDGFs and TGFβ, the main CAF activators (Louault et al., 2020). Notably, PDGF receptors, PDGFRα and PDGFRβ, and the associated downstream signaling can be pharmalogically targeted (Abdollahi et al., 2005; Paulsson et al., 2014; Pietras et al., 2008). There are also several clinical trials with molecules targeting CXCL12 including NOX-A12 which is trialed in combination with PD-1 inhibitor, pembrolizumab (Suarez-Carmona et al., 2021). Additionally, due to the indications that CAF activation involves regulatory miRNAs and epigenetic mechanisms, targeting these provide another approach to reverse CAF activation (Louault et al, 2020).

Regarding CAF activity targeting approaches, there are several molecules in clinical trials that target proteins secreted by CAFs including ECM molecules, GFs and cytokines (Louault et al., 2020). These drugs include angiotensin receptor blockers such as losartan, which has in trials been used as adjuvant therapy and

shown to improve bioavailability of for instance chemotherapeutics by decreasing the amount of αSMA positive cells and the amount of collagen in tumor stroma (Diop-Frimpong et al., 2011). Targeting factors produced by CAFs enables also more specific blocking of CAF subpopulations such as iCAFs by inhibiting cytokine-mediated signaling and myCAFs by inhibiting TGFβ (Louault et al., 2020). For instance, suppression of IL-1 induced LIF signaling in iCAFs by JAK inhibitors have been shown to skew iCAFs towards ECM-producing myCAF phenotype which facilitates ECM deposition and results in reduction of tumor growth (Biffi et al., 2019).

Taking advantage of the tumor tropism of CAFs, in particular MSCs, includes three strategies comprehending use of *in vitro* grown MSCs to produce large amounts of extracellular vesicles to be used to deliver drugs to tumor *in vivo*, use of ex vivo amplified MSCs to act as vehicles of anti-tumor drugs and use of ex vivo genetically modified MSCs to produce tumor progression reducing molecules (Hmadcha et al., 2020; Mendt et al., 2019).

Noteworthy, the discoveries of tumor restraining functions of CAFs at least partly explains the unsuccess of clinical studies targeting CAFs and stromal components, underlining the importance of in-detail classification of CAF subsets with various functions (Chen et al., 2021). Future studies will also enlighten the complex origin of CAFs but to date there is strong evidence that heterogeneous lineages exist from which CAFs can be derived including resident fibroblasts, differentiating precursor cells, de- or transdifferentiating mature cells and tumor cells (Haviv et al., 2009). Evidence also indicates that in tumors there are CAFs with various origins with majority arising from resident fibroblasts and up to 20% of CAF populations arising from MSCs (Quante et al., 2011).

The results of substudy II show that stromal changes take place and fibroblast activity evolves during the progression of cSCC, as has been noted in other cancers (Elwakeel et al., 2019). Early changes in tumor continuum include elevation in FAP, αSMA and SPARC expression which contributes to the achievement of invasive properties and the development of invasive UV-cSCC. The predominance of PDGFRα-/PDGFRβ+/FAP+ CAF subset increases during the continuum and significant elevations in the proportion of this subset is associated with invasion and metastasis. In summary, we demonstrate that singular CAF marker, PDGFRβ, and especially the PDGFRα-/PDGFRβ+/FAP+ CAF subset (CAF107) are potential metastasis risk-associated and prognostic biomarkers in cSCC. In addition, CAFs provide potential therapeutic targets also in cSCC.

6.2.3 Artificial intelligence (AI) (III)

Despite the clear benefits of AI algorithms, there are clinical, legal and regulatory issues still to be solved (Niazi et al., 2019). There is also need to make algorithms easier to interpret and to provide more insight to the execution process in order to create more transparency, which on the other hand easily deteriorates the performance (Niazi et al., 2019). One major criticism regarding AI and DL in particular is called the black box problem, which criticizes the fact that it cannot be explained how the algorithm arrives at its decision (Niazi et al., 2019). Also, the generalization of results is major concern, meaning in practice that network trained with one population does not generalize to another or images scanned with one scanner do not generalize to images scanned with another (Murphree et al., 2021). Further, regarding training itself, in addition to learning the intended task, it is possible that algorithm learns an unrelated situational feature instead, such as added mark on part of the input data, used scanner or classifies lesions based on location instead of morphology (Murphree et al., 2021). There should neither be examples of the same case in both training and testing datasets as this produces positive bias (Murphree et al., 2021).

There have long been speculations whether AI will replace clinicians such as pathologists or dermatologists, however at least to date decisions of AI are bound to narrow fields of information, whereas humans continue to be superior in accounting numerous factors and synthesizing information that leads to ultimate decisions (Niazi et al., 2019). Undoubtedly pathology will continue to become more digital and AI will advance further bringing new insights and contribution to the clinical workflow, but there are numerous technical, ethical and legal obstacles in this evolution (Niazi et al., 2019). Further, The American Academy of Dermatology in addition to recognizing the potential of AI in the field of dermatology, suggested augmented intelligence (AuI) as an alternative term for AI highlighting the role of AI or AuI as a patient care enhancing assistive tool (Kovarik et al., 2019). Large multinational validation studies with wide generalization such as the PANDA challenge are essential for future adoption of AI into the clinical practice (Bulten et al., 2022). Regarding cSCC for instance the analysis of the degree of differentiation is somewhat subjective and prone to inter-observer variability which could be overcome by incorporation of AI algorithm to the workflow.

In the substudy III of this thesis, the ability of AI to recognize primary cSCCs that will develop metastasis was examined. As to date there are no established biomarkers to predict the risk of metastasis or prognosis of primary cSCC, this task of identification is challenging which was also demonstrated by the results in substudy I.

Our results show that AI with limited training and testing data and an experienced dermatopathologist perform quite similarly in order to distinguish rapidly metastatic primary cSCCs from non-metastatic cSCCs. Here dermatopathologist was asked to evaluate the risk of metastasis of primary cSCC, although this is not directly done in clinical practice. Both AI and pathologist had only one WSI available for analysis with no access to whole tissue material of the tumor or to clinical information from which the actual diameter or invasion depth of the whole tumor could have been deduced. By Pearson correlations, it appeared that prediction by pathologist is based more on conventional histopathological features, such as invasion depth than the prediction by AI. This notion was further reinforced by similar survival prediction by AI and pathologist alone, whereas AI-RFM was superior to pathologist-RFM. Thus, the prediction by AI seems to be based on yet unestablished morphological features or feature combinations that appear in primary tumors temporally close to the time of metastasis.

The results of substudy III show that the inclusion of AI algorithm into a RFM can improve the risk assessment of cSCC metastasis. The AI-RFM created in this study was superior to other tested RFMs and staging systems and clearly differentiated cases with 0 or 1 risk factors from cases with 2 or 3 risk factors in respect of DSS prediction. It was indicated that cases with ≤1 risk factor should be considered as low risk for metastasis and cases with ≥2 risk factors as high risk for metastasis of cSCC.

So far, there are few published studies with similar study design in cancer field and no studies regarding metastasis risk of cSCC. In the prediction of the risk of metastasis in pancreatic neuroendocrine tumors, a leave-one-out cross-validation accuracy of 80.77% was reported (Klimov et al., 2021). With a CNN harnessed to predict the lymph node metastasis in prostate cancer from analysis of primary tumor tissues an AUROC of 0.68 was achieved (Wessels et al., 2021). An AUROC of 0.79, in turn, was reached in third study in which CNN was harnessed to predict lung cancer recurrence and metastasis from histopathological images (Wu et al., 2020b). Our results are in concordace but indicate more augmentative role for AI prediction by incorporating it into multifactorial RFM that provides more encouraging results.

In summary, the results of substudy III provide proof of concept, that there are certain yet unknown morphological features or feature combinations associated with the risk of cSCC metastasis that can be recognized by AI. Further studies are warranted in order to unveil these and to further develop AI algorithm as prognostic tool along with potential biomarkers and clinicopathological variables to the challenging clinical assessment of cSCC metastasis risk.

6.3   Limitations

As any retrospective study, substudy I is limited by data availability and the inherent difficulty of control cohort selection. It is notable that the non-metastatic cohort both on patient- and tumor-level is relatively small and does represent only a fraction of patients with non-mcSCC under the study period in the study region. However, every effort was made to create as comparable as possible cohorts and to eliminate bias related to treatment at tertiary care center. Thus every applicable primary cSCC of every patient included in the study was registered and processed in tumor analyses in order to resemble unselected variety of cSCCs among whole range of tumors treated on different levels of healthcare as is the case on population-level. This encompasses also the inclusion of all the applicable non-mcSCCs of the patients with metastatic disease into the non-metastatic tumor cohort.

Cohorts were remarkably similar regarding comorbidities and medications but there is some bias concerning co-malignancies based on the transfer of patients from the second screening, who did not have mcSCC, into non-metastatic patient cohort. This transfer resulted with patients who due to the screening criteria did have another metastatic cancer in addition to cSCC. This at least on its part accounts for the relatively poor prognosis in non-metastatic patient cohort and elevates the proportion of patients with co-malignancy.

Limitation in terms of prognosis comprehends the fact that rather many metastatic primary mcSCCs and even more metastases were diagnosed 2014 or later resulting in relatively high percentage of censored patients in survival analyses.

Retrospectively collected information about medications is based on medical record markings and vulnerable to bias as extent of data varies patient to patient and lacks the concrete knowledge about the actual use of prescribed medications. Another limitation in this study is the lack of sufficient information of several features such as Breslow depth, cancer subtype and presumably perineural invasion, which could not be processed in the analyses. This is due to the fact that structural histopathology reports (Table V) were not used until the very end of the study period and that pathologist did not re-evaluate the tissue material for this study. Whether pathologists reported features such as presence of necrosis or perilesional AK is also debatable. It may also be speculated whether alive patients in non-metastatic patient cohort develop metastasis as time passes. Furthermore, selected study design regarding several factors is not optimal as is the case concerning immunosuppression for instance.

Table V. Basic features included in the histopathological report of cSCC diagnosis. [a] Tumor thickness measured from the granular layer of adjacent normal epidermis to the base of the tumor. KA = keratoacanthoma, PNI = perineural invasion. Modified from Stratigos et al., 2020a.

Histopathologic report of cSCC	
Histologic subtype	Common/KA/Acantholytic/Spindle cell SCC/Verrucous/Adenosquamous/Clear cell SCC/Desmoplastic/other
Degree of differentiation	Well/Moderately/Poorly differentiated
Tumor histological thickness[a]	mm
Invasion beyond subcutaneous fat	Yes/No
PNI	Yes/No
Lymphatic/vascular invasion	Yes/No
Complete excision	Yes/No
Minimum lateral margin	mm
Minimum deep margin	mm

Considering the association between CAFs and cSCCs it would have been beneficial if there would have been more antibodies for CAF markers in the same panel, for instance the association between PDGFRα negative, PDGFRβ positive, FAP positive and POSTN positive CAF subset and metastasis as well as prognosis would have been interesting to investigate. Additionally, the number of actual metastases should have been higher. Also the examination of CAF marker expression in primary mcSCC and its actual metastasis would have been intriguing to investigate.

Regarding technical execution of the substudy III, during repeated runs of 3× or 4× CV analysis on all of the available slides for training, it seemed that depending on the choice of the CV split, the results varied considerably. Often it seemed that one fold was much more difficult for the model to analyze than the others, leading to considerably lower AUROC values. We took this as a sign that the dataset contained some examples that were difficult for the model to analyze, for example by being very atypical cases of rapidly metastatic tumors. More careful analysis of the metadata revealed, however, that the difficult cases appeared in folds where the slide data had been scanned in 2020 instead of 2016. Most of the rapid metastasis cases were scanned earlier because various research projects have been conducted with the samples in the past. The scanner used was the same, but it was hypothesized that the scanning software, image packing algorithms or the physical components of the scanner itself could have been updated affecting the image colors, noise patterns or other qualities. Heavy color augmentations were used in

training of the algorithm, but to study the possible effect of slides scanned at different times we tried and restricted the data to old samples only. This reduced the size of the dataset available, but was useful in ruling out at least one more possible source of error. This did affect the AUROC results somewhat and brought the average AUROC value up.

For further studies regarding AI and metastasis risk of primary cSCC, larger cohorts are required. Due to our notion that small subcohorts scanned with different scanner or on another date can generate bias, it is recommended to use large enough, preferably equally sized cohorts of samples scanned on different dates or tissue specimens scanned on same occasion. Same applies to the usage of different scanners. Scanning settings should naturally be identical. Larger cohorts would enable further analysis of the ability of AI algorithm to recognize both rapid and slow metastasis cases as well as biopsies and resections. Additionally, the analysis of tissue samples from different points in tumor continuum would be interesting to reveal the point in time when features associated with metastasis appear.

7 Summary/Conclusions

The metastatic rate of cSCC in the area served by Turku University Hospital is 2.3%. Metastasis occurs rapidly and in majority of cases patients do not have history of earlier cSCCs. Projection of staging studies is suboptimal.

Activity and phenotype of cancer-associated fibroblasts (CAFs) evolve during the progression of cSCC. Early change in the tumor continuum of cSCC comprehends the elevation of stromal FAP, αSMA and SPARC expression that is associated with invasion.

Invasion depth is a crucial clinicohistopathological feature associated with the risk of metastasis. Usage of isosorbide mono- or dinitrate and ASA, especially in combination, lowers the risk of metastasis. Further, the presence of precursor lesions and BCC is associated with lower risk of metastasis. On whole slide images representing cSCC, AI is able to recognize unknown morphological features or feature combinations associated with metastasis. On molecular level, elevation in stromal FAP and PDGFRβ expression as well as the predominance of CAF107 phenotype are associated with metastasis.

In case of metastasis the prognosis of cSCC is poor. Poor prognosis is associated with ENE and the number of nodal metastases. CAF107 subset independently associates with poor prognosis of cSCC and AI provides added prognostic value.

The findings of this thesis provide new insight to the nature and risk stratification of metastatic cSCC which with yet increasing incidence undoubtedly will not be less significant or less acknowledged medical entity in the future. CAF-markers and AI were proven to possess potential to act as novel metastasis risk stratification tools that could enhance the prognosis of patients with metastatic cSCC.

Acknowledgements

Turku, October 2022
Jaakko Knuutila

References

Abe, R., Donnelly, S. C., Peng, T., Bucala, R. & Metz, C. N. (2001). Peripheral Blood Fibrocytes: Differentiation Pathway and Migration to Wound Sites. *J Immunol,* 166(12), 7556-7562.

Abdollahi, A., Li, M., Ping, G., Plathow, C., Domhan, S., Kiessling, F., Lee, L. B., McMahon, G., Gröne, H. J., Lipson, K. E. & Huber, P. E. (2005). Inhibition of platelet-derived growth factor signaling attenuates pulmonary fibrosis. *J Exp Med*, 201(6), 925-935.

Acs, B., Rantalainen, M. & Hartman, J. (2020). Artifiial intelligence as the next step towards precision pathology. *J Intern Med*, 288(1), 62-81.

Acs, B., Pelekanou, V., Bai, Y., Martinez-Morilla, S., Toki, M., Leung, S. C. Y., Nielsen, T. O. & Rimm, D. L. (2019a). Ki67 reproducibility using digital image analysis: an inter-platform and inter-operator study. *Lab Invest*, 99(1), 107-117.

Acs, B., Ahmed, F. S., Gupta, S., Wong, P. F., Gartrell, R. D., Sarin Pradhan, J., Rizk, E. M., Gould Rothberg, B., Saenger, Y. M. & Rimm, D. L. (2019b). An open source automated tumor infitrating lymphocyte algorithm for prognosis in melanoma. *Nat Commun*, 10(1), 5440.

Adams, C. C., Thomas, B. & Bingham, J. L. (2014). Cutaneous squamous cell carcinoma with perineural invasion: a case report and review of the literature. *Cutis*, 93(3), 141-144.

Affo, S., Nair, A., Brundu, F., Ravichandra, A., Bhattacharjee, S., Matsuda, M., Chin, L., Filliol, A., Wen, W., Song, X., Decker, A., Worley, J., Caviglia, J. M., Yu, L., Yin, D., Saito, Y., Savage, T., Wells, R. G., Mack, M., Zender, L., Arpaia, N., Remotti, H. E., Rabadan, R., Sims, P., Leblond, A. L., Weber, A., Riener, M. O., Stockwell, B. R., Gaublomme, J., Llovet, J. M., Kalluri, R., Michalopoulos, G. K., Seki, E., Sia, D., Chen, X., Califano, A. & Schwabe, R. F. (2021). Promotion of cholangiocarcinoma growth by diverse cancer- associated fibroblast subpopulations. *Cancer Cell*, 39(6), 866-882.

Alam, M., Armstrong, A., Baum, C., Bordeaux, J. S., Brown, M., Busam, K. J., Eisen, D. B., Iyengar, V., Lober, C., Margolis, D. J., Messina, J., Miller, A., Miller, S., Mostow, E., Mowad, C., Nehal, K., Schmitt-Burr, K., Sekulic, A., Storrs, P., Teng, J., Yu, S., Huang, C., Boyer, K., Begolka, W. S., Bichakjian, C., Kim, J. Y. S., Kozlow, J. H., Mittal, B., Moyer, J., Olenecki, T. & Rodgers, P. (2018). Guidelines of care for the management of cutaneous squamous cell carcinoma. *J Am Acad Dermatol*, 78(3), 560-578.

Alam, M. & Ratner, D. (2001). Cutaneous squamous-cell carcinoma. *N Engl J Med*, 344(13), 975-983.

Albandar, H. J., Markert, R. & Agrawal, S. (2018). The relationship between aspirin use and mortality in colorectal cancer. *J Gastrointest Oncol*, 9(6), 1133-1137.

Albarqouni, S., Baur, C., Achilles, F., Belagiannis, V., Demirci, S. & Navab, N. (2016). AggNet: deep learning from crowds for mitosis detection in breast cancer histology images. *IEEE Trans Med Imaging*. 35(5), 1313–1321.

Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A., Behjati, S., Biankin, A. V., Bignell, G. R., Bolli, N., Borg, A., Børresen-Dale, A. L., Boyault, S., Burkhardt, B., Butler, A. P., Caldas, C., Davies, H. R., Desmedt, C., Eils, R., Eyfjörd, J. E., Foekens, J. A., Greaves, M., Hosoda, F., Hutter, B., Ilicic, T., Imbeaud, S., Imielinski, M., Jäger, N., Jones, D. T., Jones, D., Knappskog, S., Kool, M., Lakhani, S. R., López-Otín, C., Martin, S., Munshi, N. C., Nakamura, H.,

Northcott, P.A., Pajic, M., Papaemmanuil, E., Paradiso, A., Pearson, J. V., Puente, X. S., Raine, K., Ramakrishna, M., Richardson, A. L., Richter, J., Rosenstiel, P., Schlesner, M., Schumacher, T. N., Span, P. N., Teague, J. W., Totoki, Y., Tutt, A. N., Valdés-Mas, R., van Buuren, M. M., van 't Veer, L., Vincent-Salomon, A., Waddell, N., Yates, L. R.; Australian Pancreatic Cancer Genome Initiative; ICGC Breast Cancer Consortium; ICGC MMML-Seq Consortium; ICGC PedBrain, Zucman-Rossi, J., Futreal, P. A., McDermott, U., Lichter, P., Meyerson, M., Grimmond, S. M., Siebert, R., Campo, E., Shibata, T., Pfister, S. M., Campbell, P. J. & Stratton, M. R. (2013). Signatures of mutational processes in human cancer. *Nature*, **500**(7463), 415-421.

Amin, M. B., Edge, S. B., Greene, F. L., Byrd, D. R., Brookland, R. K., Washington, M. K., Gershenwald, J. E., Compton, C. C., Hess, K. R., Sullivan, D. C., Jessup, L. M. Brierley, J. D., Gaspar, L. E., Schilsky, R. L., Balch, C. M., Winchester, D. P., Asare, E. A., Madera, M., Gress, D. M. & Meyer, L. R. (2017). AJCC Cancer Staging Manual, 8th ed. New York, Springer.

Andersson, E. M., Paoli, J. & Wastensson, G. (2011). Incidence of cutaneous squamous cell carcinoma in coastal and inland areas of Western Sweden. *Cancer Epidemiol*, 35(6), e69-74.

Anwar, S. M., Majid, M., Qayyum, A., Awais, M., Alnowami, M. & Khan, M. K. (2018). Medical image analysis using convolutional neural networks: a review. *J Med Syst*, 42(11), 226.

Aoyagi, Y., Oda, T., Kinoshita, T., Nakahashi, C., Hasebe, T., Ohkohchi, N. & Ochiai, A. (2004). Overexpression of TGF-β by infiltrated granulocytes correlates with the expression of collagen mRNA in pancreatic cancer. Br J Cancer, 91(7), 1316-1326.

Armstrong, T., Packham, G., Murphy, L. B., Bateman, A. C., Conti, J. A., Fine, D. R., Johnson, C. D., Benyon, R. C. & Iredale, J. P. (2004). Type I collagen promotes the malignant phenotype of pancreatic ductal adenocarcinoma. *Clin Cancer Res*, 10(21), 7427-7437.

Armulik, A., Genove, G. & Betsholtz, C. (2011). Pericytes: developmental, physiological, and pathological perspectives, problems, and promises. *Dev Cell*, 21(2), 193-215.

Arnold, J. N., Magiera, L., Kraman, M. & Fearon, D. T. (2014). Tumoral immune suppression by macrophages expressing fibroblast activation protein-α and heme oxygenase-1. Cancer Immunol Res, 2(2), 121-126.

Arsenic, R. & Kurrer, M. O. (2013). Differentiated dysplasia is a frequent precursor or associated lesion in invasive squamous cell carcinoma of the oral cavity and pharynx. *Virchows Arch*, 462(6), 609-17.

Ascensión, A. M., Fuertes-Álvarez, S., Ibañez-Solé, O., Izeta, A. & Araúzo-Bravo, M. J. (2021). Human Dermal Fibroblast Subpopulations Are Conserved across Single-Cell RNA Sequencing Studies. *J Invest Dermatol*, 141(7), 1735-1744.

Awaji, M. & Singh, R. K. (2019). Cancer-associated fibroblasts' functional heterogeneity in pancreatic ductal adenocarcinoma. *Cancers (Basel)*, 11(3), 290.

Bai, L., Mao, G. P., Cao, C. P. (2007). Effects of inflammatory cytokines on the recurrence of liver cancer after an apparently curative operation. *J Dig Dis,* 8(3), 154-159.

Bai, Y. P., Shang, K., Chen, H., Ding, F., Wang, Z., Liang, C., Xu, Y., Sun, M. H. & Li, Y. Y. (2015). FGF-1/-3/FGFR4 signaling in cancer-associated fibroblasts promotes tumor progression in colon cancer through Erk and MMP-7. *Cancer Sci,* 106(10), 1278-1287.

Bainbridge, P. (2013). Wound healing and the role of fibroblasts. *J Wound Care,* 22(8), 407-408.

Banks, E. R. & Cooper, P. H. (1991). Adenosquamous carcinoma of the skin: a report of 10 cases. *J Cutan Pathol*, 18(4), 227-234.

Barcellos-de-Souza, P., Comito, G., Pons-Segura, C., Taddei, M. L., Gori, V., Becherucci, V., Bambi, F., Margheri, F., Laurenzana, A., Del Rosso, M. & Chiarugi P. (2016). Mesenchymal Stem Cells are Recruited and Activated into Carcinoma-Associated Fibroblasts by Prostate Cancer Microenvironment-Derived TGF-β1. *Stem Cells*, 34(10), 2536-2547.

Bartoschek, M., Oskolkov, N., Bocci, M., Lövrot, J., Larsson, C., Sommarin, M., Madsen, C. D., Lindgren, D., Pekar, G., Karlsson, G., Ringnér, M., Bergh, J., Björklund, Å. & Pietras, K. (2018). Spatially and Functionally Distinct Subclasses of Breast Cancer-Associated Fibroblasts Revealed by Single Cell RNA Sequencing. *Nat Commun*, 9(1), 5150.

Baum, C. L., Wright, A. C., Martinez, J. C., Arpey, C. J., Brewer, J. D., Roenigk, R. K. & Otley, C. C. (2018). A new evidence-based risk stratification system for cutaneous squamous cell carcinoma into low, intermediate, and high risk groups with implications for management. *J Am Acad Dermatol*, 78(1), 141-147.

Bejnordi, B. E., Lin, J., Glass, B., Mullooly, M., Gierach, G. L., Sherman, M. E., Karssemeijer, N., van der Laak, J. & Beck, A. H. (2017a). Learning-based assessment of tumor-associated stroma for diagnosing breast cancer in histopathology images. *Proc IEEE Int Symp Biomed Imaging*, 2017, 929-932.

Bejnordi, B. E., Veta, M., Johannes van Diest, P., van Ginneken, B., Karssemeijer, N., Litjens, G., van der Laak, J. A. W. M.; the CAMELYON16 Consortium, Hermsen, M., Manson, Q. F., Balkenhol, M., Geessink, O., Stathonikos, N., van Dijk, M. C., Bult, P., Beca, F., Beck, A. H., Wang, D., Khosla, A., Gargeya, R., Irshad, H., Zhong, A., Dou, Q., Li, Q., Chen, H., Lin, H. J., Heng, P. A., Haß, C., Bruni, E., Wong, Q., Halici, U., Öner, M. Ü., Cetin-Atalay, R., Berseth, M., Khvatkov, V., Vylegzhanin, A., Kraus, O., Shaban, M., Rajpoot, N., Awan, R., Sirinukunwattana, K., Qaiser, T., Tsang, Y. W., Tellez, D., Annuscheit, J., Hufnagl, P., Valkonen, M., Kartasalo, K., Latonen, L., Ruusuvuori, P., Liimatainen, K., Albarqouni, S., Mungal, B., George, A., Demirci, S., Navab, N., Watanabe, S., Seno, S., Takenaka, Y., Matsuda, H., Ahmady Phoulady, H., Kovalev, V., Kalinovsky, A., Liauchuk, V., Bueno, G., Fernandez-Carrobles, M. M., Serrano, I., Deniz, O., Racoceanu, D. & Venâncio, R. (2017b). Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA*, 318(22), 2199-2210.

Bentaieb, A. & Hamarneh, G. (2018). Adversarial stain transfer for histopathology image analysis. *IEEE Trans Med Imaging*, 37(3), 792-802.

Berg, S., Kutra, D., Kroeger, T., Straehle, C. N., Kausler, B. X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M., Eren, K., Cervantes, J. I., Xu, B., Beuttenmueller, F., Wolny, A., Zhang, C., Koethe, U., Hamprecht, F. A. & Kreshuk, A. (2019). ilastik: interactive machine learning for (bio)image analysis. *Nat Methods*, 16(12), 1226-1232.

Bhattacharjee, S., Hamberger, F., Ravichandra, A., Miller, M., Nair, A., Affo, S., Filliol, A., Chin, L., Savage, T. M., Yin, D., Wirsik, N. M., Mehal, A., Arpaia, N., Seki, E., Mack, M., Zhu, D., Sims, P. A., Kalluri, R., Stanger, B. Z., Olive, K. P., Schmidt, T., Wells, R. G., Mederacke, I. & Schwabe, R. F. (2021). Tumor restriction by type I collagen opposes tumor- promoting effects of cancer-associated fibroblasts. *J Clin Invest*, 131(11), e146987.

Bhowmick, N. A., Chytil, A., Plieth, D., Gorska, A. E., Dumont, N., Shappell, S., Washington, M. K., Neilson, E. G. & Moses, H. L. (2004). TGF- Beta Signaling in Fibroblasts Modulates the Oncogenic Potential of Adjacent Epithelia. *Science,* 303(5659), 848-851.

Biffi, G., Oni, T. E., Spielman, B., Hao, Y., Elyada, E., Park, Y., Preall, J. & Tuveson, D. A. (2019). IL1-induced JAK/STAT signaling is antagonized by TGFβ to shape CAF heterogeneity in pancreatic ductal adenocarcinoma. *Cancer Discov*, 9(2):282-301.

Bliss, L. A., Sams, M. R., Deep-Soboslay, A., Ren-Patterson, R., Jaffe, A. E., Chenoweth, J. G., Jaishankar, A., Kleinman, J. E. & Hyde, T. M. (2012). Use of postmortem human dura mater and scalp for deriving human fibroblast cultures. *PLoS One*, 7(9), e45282.

Blom, S., Paavolainen, L., Bychkov, D., Turkki, R., Mäki-Teeri, P., Hemmes, A., Välimäki, K., Lundin, J., Kallioniemi, O. & Pellinen, T. (2017). Systems pathology by multiplexed immunohistochemistry and whole-slide digital image analysis. *Sci Rep*, 7(1), 15580.

Bobin, C., Ingrand, P., Dréno, B., Rio, E., Malard, O. & Espitalier, F. (2018). Prognostic factors for parotid metastasis of cutaneous squamous cell carcinoma of the head and neck. *Eur Ann Otorhinolaryngol Head Neck Dis*, 135(2), 99-103.

Borriello, L., Nakata, R., Sheard, M. A., Fernandez, G. E., Sposto, R., Malvar, J., Blavier, L., Shimada, H., Asgharzadeh, S., Seeger, R. C. & DeClerck, Y. A. (2017). Cancer-Associated Fibroblasts Share Characteristics and Protumorigenic Activity with Mesenchymal Stromal Cells. *Cancer Res*, 77(18), 5142-5157.

Brantsch, K. D., Meisner, C., Schönfisch, B., Trilling, B., Wehner-Caroli, J., Röcken, M. & Breuninger, H. (2008). Analysis of risk factors determining prognosis of cutaneous squamous-cell carcinoma: a prospective study. *Lancet Oncol*, 9(8), 713-720.

Brash, D. E. (2015). UV signature mutations. *Photochem Photobiol*, 91(1), 15-26.

Bredholt, G., Mannelqvist, M., Stefansson, I. M., Birkeland, E., Bø, T. H., Øyan, A. M., Trovik, J., Kalland, K. H., Jonassen, I., Salvesen, H. B., Wik, E. & Akslen, L. A. (2015). Tumor necrosis is an important hallmark of aggressive endometrial cancer and associates with hypoxia, angiogenesis and inflammation responses. *Oncotarget*. 6(37), 39676-39691.

Breuninger, H., Black, B. & Rassner, G. (1990). Microstaging of squamous cell carcinomas. *Am J Clin Pathol*, 94(5), 624-627.

Breuninger, H., Brantsch, K., Eigentler, T. & Häfner, H. M. (2012). Comparison and evaluation of the current staging of cutaneous carcinomas. *J Dtsch Dermatol Ges*, 10(8), 579-586.

Breuninger, H., Schaumburg-Lever, G., Holzschuh, J. & Horny, H. P. (1997). Desmoplastic squamous cell carcinoma of skin and vermilion surface: a highly malignant subtype of skin cancer. *Cancer*, 79(5), 915-919.

Brewer, J. D., Shanafelt, T. D., Khezri, F., Sosa Seda, I. M., Zubair, A. S., Baum, C. L., Arpey, C. J., Cerhan, J. R., Call, T. G., Roenigk, R. K., Smith, C. Y., Weaver, A. L. & Otley, C.C. (2015). Increased incidence and recurrence rates of nonmelanoma skin cancer in patients with non-Hodgkin lymphoma: a Rochester Epidemiology Project population-based study in Minnesota. *J Am Acad Dermatol*, 72(2), 302-309.

Brewster, D. H., Bhatti, L. A., Inglis, J. H., Nairn, E. R. & Doherty, V. R. (2007). Recent trends in incidence of nonmelanoma skin cancers in the east of Scotland, 1992-2003. *Br J Dermatol*, 156(6), 1295-1300.

Bronzert, D. A., Pantazis, P., Antoniades, H. N., Kasid, A., Davidson, N., Dickson, R. B. & Lippman, M. E. (1987). Synthesis and secretion of platelet-derived growth factor by human breast cancer cell lines. *Proc Natl Acad Sci U S A*, 84(16), 5763-5767.

Brougham, N. D., Dennett, E. R., Cameron, R. & Tan, S. T. (2012). The incidence of metastasis from cutaneous squamous cell carcinoma and the impact of its risk factors. *J Surg Oncol*, 106(7), 811-815.

Brown, L. F., Guidi, A. J., Schnitt, S. J., Van De Water, L., Iruela-Arispe, M. L., Yeo, T. K., Tognazzi, K. & Dvorak, H. F. (1999). Vascular stroma formation in carcinoma in situ, invasive carcinoma, and metastatic carcinoma of the breast. *Clin Cancer Res*, 5(5), 1041-1056.

Brown, V. L., Harwood, C. A., Crook, T., Cronin, J. G., Kelsell, D. P. & Proby, C. M. (2004). p16INK4a and p14ARF tumor suppressor genes are commonly inactivated in cutaneous squamous cell carcinoma. *J Invest Dermatol,* 122(5), 1284-1292.

Bulten, W., Kartasalo, K., Chen, P. C., Ström, P., Pinckaers, H., Nagpal, K., Cai, Y., Steiner, D. F., van Boven, H., Vink, R., Hulsbergen-van de Kaa, C., van der Laak, J., Amin, M. B., Evans, A. J., van der Kwast, T., Allan, R., Humphrey, P. A., Grönberg, H., Samaratunga, H., Delahunt, B., Tsuzuki, T., Häkkinen, T., Egevad, L., Demkin, M., Dane, S., Tan, F., Valkonen, M., Corrado, G. S., Peng, L., Mermel, C. H., Ruusuvuori, P., Litjens, G. & Eklund, M.; PANDA challenge consortium. (2022). Artificial intelligence for diagnosis and Gleason grading of prostate cancer: the PANDA challenge. *Nat Med*, 28(1), 154-163.

Burton, K. A., Ashack, K. A. & Khachemoune, A. (2016). Cutaneous Squamous Cell Carcinoma: A Review of High-Risk and Metastatic Disease. *Am J Clin Dermatol,* 17(5), 491-508.

Bychkov, D., Linder, N., Turkki, R., Nordling, S., Kovanen, P. E., Verrill, C., Walliander, M., Lundin, M., Haglund, C. & Lundin, J. (2018). Deep learning based tissue analysis predicts outcome in colorectal cancer. *Sci Rep*, 8(1), 3395.

Campanella, G., Hanna, M. G., Geneslaw, L., Miraflor, A., Werneck Krauss Silva, V., Busam, K. J., Brogi, E., Reuter, V. E., Klimstra, D. S. & Fuchs, T. J. (2019). Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med*, 25(8), 1301-1309.

Cañueto, J., Burguillo, J., Moyano-Bueno, D., Viñolas-Cuadros, A., Conde-Ferreirós, A., Corchete-Sánchez, L. A., Pérez-Losada, J. & Román-Curto, C. (2019). Comparing the eighth and the seventh editions of the American Joint Committee on Cancer staging system and the Brigham and Women's Hospital alternative staging system for cutaneous squamous cell carcinoma: Implications for clinical practice. *J Am Acad Dermatol*, 80(1), 106-113.

Carstens, J. L., Correa de Sampaio, P., Yang, D., Barua, S., Wang, H., Rao, A., Allison, J. P., LeBleu, V. S. & Kalluri, R. (2017). Spatial Computation of Intratumoral T Cells Correlates with Survival of Patients with Pancreatic Cancer. *Nat Commun*, 8, 15095.

Carucci, J. A., Martinez, J. C., Zeitouni, N. C., Christenson, L., Coldiron, B., Zweibel, S. & Otley, C. C. (2004). In-transit metastasis from primary cutaneous squamous cell carcinoma in organ transplant recipients and nonimmunosuppressed patients: clinical characteristics, management, and outcome in a series of 21 patients. *Dermatol Surg*, 30(4 Pt 2), 651-655.

Cassarino, D. S., Derienzo, D. P. & Barr, R. J. (2006a). Cutaneous squamous cell carcinoma: a comprehensive clinicopathologic classification--part one. *J Cutan Pathol*, 33(3), 191-206.

Cassarino, D. S., Derienzo, D. P. & Barr, R. J. (2006b). Cutaneous squamous cell carcinoma: a comprehensive clinicopathologic classification--part two. *J Cutan Pathol*, 33(4), 261-279.

Castor, C. W., Wilson, S. M., Heiss, P. R. & Seidman, J. C. (1979). Activation of lung connective tissue cells in vitro. *Am Rev Respir Dis*, 120(1), 101-106.

Chacón-Solano, E., León, C., Díaz, F., García-García, F., García, M., Escámez, M. J., Guerrero-Aspizua, S., Conti, C. J., Mencía, Á., Martínez-Santamaría, L., Llames, S., Pévida, M., Carbonell-Caballero, J., Puig-Butillé, J. A., Maseda, R., Puig, S., de Lucas, R., Baselga, E., Larcher, F., Dopazo, J. & Del Río, M. (2019). Fibroblast activation and abnormal extracellular matrix remodelling as common hallmarks in three cancer-prone genodermatoses. *Br J Dermatol*, 181(3), 512-522.

Chan, A. T., Ogino, S. & Fuchs, C. S. (2009). Aspirin use and survival after diagnosis of colorectal cancer. *JAMA*, 302(6), 649-658.

Chang, D. & Shain, A. H. (2021). The landscape of driver mutations in cutaneous squamous cell carcinoma. *NPJ Genom Med*, 6(1), 61.

Chang, H. Y., Jung, C. K., Woo, J. I., Lee, S., Cho, J., Kim, S. W. & Kwak, T. Y. (2019). Artificial intelligence in pathology. *J Pathol Transl Med*, 53(1), 1-12.

Chen, X. & Song, E. (2019). Turning foes to friends: targeting cancer-associated fibroblasts. *Nat Rev Drug Discov*, 18(2), 99-115.

Chen, Y., McAndrews, K. M. & Kalluri, R. (2021). Clinical and therapeutic relevance of cancer-associated fibroblasts. *Nat Rev Clin Oncol*, 18(12), 792-804.

Cheng, J. Y., Li, F. Y., Ko, C. J. & Colegio, O. R. (2018). Cutaneous Squamous Cell Carcinomas in Solid Organ Transplant Recipients Compared With Immunocompetent Patients. *JAMA Dermatol*, 154(1), 60-66.

Cherpelis, B. S., Marcusen, C. & Lang, P. G. (2002). Prognostic factors for metastasis in squamous cell carcinoma of the skin. *Dermatol Surg*, 28(3), 268-273.

Cho, R. J., Alexandrov, L. B., den Breems, N. Y., Atanasova, V. S., Farshchian, M., Purdom, E., Nguyen, T. N., Coarfa, C., Rajapakshe, K., Prisco, M., Sahu, J., Tassone, P., Greenawalt, E. J., Collisson, E. A., Wu, W., Yao, H., Su, X., Guttmann-Gruber, C., Hofbauer, J. P., Hashmi, R., Fuentes, I., Benz, S. C., Golovato, J., Ehli, E. A., Davis, C. M., Davies, G. E., Covington, K. R., Murrell, D. F., Salas-Alanis, J. C., Palisson, F., Bruckner, A. L., Robinson, W., Has, C., Bruckner-Tuderman, L., Titeux, M., Jonkman, M. F., Rashidghamat, E., Lwin, S. M., Mellerio, J. E., McGrath, J. A., Bauer, J. W., Hovnanian, A., Tsai, K. Y. & South, A. P. (2018). APOBEC mutation drives early-onset squamous cell carcinomas in recessive dystrophic epidermolysis bullosa. *Sci Transl Med*, 10(455), eaas9668.

Cockerell, C. J. (2000). Histopathology of incipient intraepidermal squamous cell carcinoma ("actinic keratosis"). *J Am Acad Dermatol*, 42(1 Pt 2), 11-17.

Cockerell, C. J. & Wharton, J. R. (2005). New histopathological classification of actinic keratosis (incipient intraepidermal squamous cell carcinoma). *J Drugs Dermatol*, 4(4), 462-467.

Commandeur, S., Ho, S. H., de Gruijl, F. R., Willemze, R., Tensen, C. P. & El Ghalbzouri, A. (2011). Functional characterization of cancer-associated fibroblasts of human cutaneous squamous cell carcinoma. *Exp Dermatol*, 20(9), 737-742.

Conde-Ferreirós, A., Corchete, L. A., Puebla-Tornero, L., Corchado-Cobos, R., García-Sancha, N., Román-Curto, C. & Cañueto, J. (2021). Definition of prognostic subgroups in the T3 stage of the eighth edition of the American Joint Committee on Cancer staging system for cutaneous squamous cell carcinoma: Tentative T3 stage subclassification. *J Am Acad Dermatol*, 85(5), 1168-1177.

Condorelli, A. G., Dellambra, E., Logli, E., Zambruno, G. & Castiglia, D. (2019). Epidermolysis Bullosa-Associated Squamous Cell Carcinoma: From Pathogenesis to Therapeutic Perspectives. *Int J Mol Sci*, 20(22), 5707.

Costa, A., Kieffer, Y., Scholer-Dahirel, A., Pelon, F., Bourachot, B., Cardon, M., Sirven, P., Magagna, I., Fuhrmann, L., Bernard, C., Bonneau, C., Kondratova, M., Kuperstein, I., Zinovyev, A., Givel, A. M., Parrini, M. C., Soumelis, V., Vincent-Salomon, A. & Mechta-Grigoriou, F. (2018). Fibroblast Heterogeneity and Immunosuppressive Environment in Human Breast Cancer. *Cancer Cell*, 33(3), 463-479.

Costa, A., Scholer-Dahirel, A. & Mechta-Grigoriou, F. (2014). The Role of Reactive Oxygen Species and Metabolism on Cancer Cells and Their Microenvironment. *Semin Cancer Biol, 25*, 23–32.

Costea, D. E., Hills, A., Osman, A. H., Thurlow, J., Kalna, G., Huang, X., Pena Murillo, C., Parajuli, H., Suliman, S., Kulasekara, K. K., Johannessen, A. C. & Partridge, M. (2013). Identification of two distinct carcinoma-associated fibroblast subtypes with differential tumor-promoting abilities in oral squamous cell carcinoma. *Cancer Res*, 73(13), 3888-3901.

Coudray, N., Ocampo, P. S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., Moreira, A. L., Razavian, N. & Tsirigos, A. (2018). Classifiation and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med*, 24(10), 1559-1567.

Couture, H. D., Williams, L. A., Geradts, J., Nyante, S. J., Butler, E. N., Marron, J. S., Perou, C. M., Troester, M. A. & Niethammer, M. (2018). Image analysis with deep learning to predict breast cancer grade, ER status, histologic subtype, and intrinsic subtype. *NPJ Breast Cancer*, 4, 30.

Criscione, V. D., Weinstock, M. A., Naylor, M. F., Luque, C., Eide, M. J. & Bingham, S. F.; Department of Veteran Affairs Topical Tretinoin Chemoprevention Trial Group. (2009). Actinic keratoses: natural history and risk of malignant transformation in the veterans affairs topical tretinoin chemoprevention trial. *Cancer*, 115(11), 2523-2530.

Crowley, M. R., Bowtell, D. & Serra, R. (2005). TGF-beta, c-Cbl, and PDGFR-alpha the in mammary stroma. *Dev Biol*, 279(1), 58-72.

Cruz-Roa, A., Gilmore, H., Basavanhally, A., Leldman, M., Ganesan, S., Shih, N., Tomaszewski, J., Madabhushi, A. & Gonzalez, L. (2018). High-throughput adaptive sampling for whole-slide histopathology image analysis (HASHI) via convolutional neural networks: Application to invasive breast cancer detection. *PloS One*, 13(5), e0196828.

Czarnecki, D. (2017). Non-melanoma skin cancer mortality rising in susceptible Australians. *J Eur Acad Dermatol Venereol*, 31(6), e286-e287.

Darby, I. A., Laverdet, B., Bonte, F. & Desmouliere, A. (2014). Fibroblasts and myofibroblasts in wound healing. Clin Cosmet Investig Dermatol, 7, 301-311.

Das, H., Wang, Z., Niazi, M. K. K., Aggarwal, R., Lu, J., Kanji, S., Das, M., Joseph, M., Gurcan, M. & Cristini, V. (2013). Impact of diffusion barriers to small cytotoxic molecules on the efficacy of immunotherapy in breast cancer. *PloS One*, 8(4), e61398.

Desmouliere, A., Darby, I. A. & Gabbiani, G. (2003). Normal and pathologic soft tissue remodeling: role of the myofibroblast, with special emphasis on liver and kidney fibrosis. *Lab Invest*, 83(12), 1689-1707.

Desmoulière, A., Guyot, C. & Gabbiani, G. (2004). The Stroma Reaction Myofibroblast: A Key Player in the Control of Tumor Cell Behavior. *Int J Dev Biol*, 48(5-6), 509-517.

De Wever, O., Van Bockstal, M., Mareel, M., Hendrix, A. & Bracke, M. (2014). Carcinoma-associated fibroblasts provide operational flexibility in metastasis. Semin Cancer Biol, 25, 33-46.

Dinehart, S. M. & Pollack, S. V. (1989). Metastases from squamous cell carcinoma of the skin and lip. An analysis of twenty-seven cases. *J Am Acad Dermatol*, 21(2 Pt 1), 241-248.

Diop-Frimpong, B., Chauhan, V. P., Krane, S., Boucher, Y. & Jain, R. K. (2011). Losartan Inhibits Collagen I Synthesis and Improves the Distribution and Efficacy of Nanotherapeutics in Tumors. *Proc Natl Acad Sci U S A*, 108(7), 2909-2914.

Direkze, N. C., Hodivala-Dilke, K., Jeffery, R., Hunt, T., Poulsom, R., Oukrif, D., Alison, M. R. & Wright, N. A. (2004). Bone Marrow Contribution to Tumor-Associated Myofibroblasts and Fibroblasts. *Cancer Res*, 64(23), 8492-8495.

Dolberg, D. S., Hollingsworth, R., Hertle, M. & Bissell, M. J. (1985). Wounding and its role in RSV-mediated tumor formation. Science, 230(4726), 676-678.

Dominguez, C. X., Müller, S., Keerthivasan, S., Koeppen, H., Hung, J., Gierke, S., Breart, B., Foreman, O., Bainbridge, T. W., Castiglioni, A., Senbabaoglu, Y., Modrusan, Z., Liang, Y., Junttila, M. R., Klijn, C., Bourgon, R. & Turley, S. J. (2020). Single-Cell RNA Sequencing Reveals Stromal Evolution into LRRC15+ Myofibroblasts as a Determinant of Patient Response to Cancer Immunotherapy. *Cancer Discov*, 10(2), 232-253.

Dong, H., Strome, S. E., Salomao, D. R., Tamura, H., Hirano, F., Flies, D. B., Roche, P. C., Lu, J., Zhu, G., Tamada, K., Lennon, V. A., Celis, E. & Chen, L. (2002). Tumor-associated B7-H1 promotes T-cell apoptosis: a potential mechanism of immune evasion. *Nat Med,* 8(8), 793-800.

Dongre, A. & Weinberg, R. A. (2019). New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. *Nat Rev Mol Cell Biol*, 20(2), 69-84.

Driessens, G., Kline, J., Gajewski, T. F. (2009). Costimulatory and coinhibitory receptors in antitumor immunity. *Immunol Rev,* 229(1), 126-144.

Dubé, C., Rostom, A., Lewin, G., Tsertsvadze, A., Barrowman, N., Code, C., Sampson, M., Moher, D. & U.S. Preventive Services Task Force. (2007). The use of aspirin for primary prevention of colorectal cancer: a systematic review prepared for the U.S. Preventive Services Task Force. *Ann Intern Med*, 146(5), 365-375.

Duggento, A., Conti, A., Mauriello, A., Guerrisi, M. & Toschi, N. (2021). Deep computational pathology in breast cancer. *Semin Cancer Biol*, 72, 226-237.

Dumont, N., Liu, B., Defilippis, R. A., Chang, H., Rabban, J. T., Karnezis, A. N., Tjoe, J. A., Marx, J., Parvin, B. & Tlsty, T. D. (2013). Breast fibroblasts modulate early dissemination, tumorigenesis, and metastasis through alteration of extracellular matrix characteristics. *Neoplasia*, 15(3), 249-262.

Durinck, S., Ho, C., Wang, N. J., Liao, W., Jakkula, L. R., Collisson, E. A., Pons, J., Chan, S. W., Lam, E. T., Chu, C., Park, K., Hong, S. W., Hur, J. S., Huh, N., Neuhaus, I. M., Yu, S. S., Grekin, R. C., Mauro, T. M., Cleaver, J. E., Kwok, P. Y., LeBoit, P. E., Getz, G., Cibulskis, K., Aster, J. C., Huang, H., Purdom, E., Li, J., Bolund, L., Arron, S. T., Gray, J. W., Spellman, P. T. & Cho, R. J. (2011). Temporal dissection of tumorigenesis in primary cancers. *Cancer Discov*, 1(2), 137–143.

Duvall, M. (1879). Atlas d'Embryologie. (ed. Masson, G.) (Paris, 1879)

Dvorak, H. F., Form, D. M., Manseau, E. J. & Smith, B. D. (1984). Pathogenesis of desmoplasia. I. Immunofluorescence identification and localization of some structural proteins of line 1 and line 10 guinea pig tumors and of healing wounds. *J Natl Cancer Inst*, 73(5), 1195-1205.

Eble, J. A. & Niland, S. (2019). The Extracellular Matrix in Tumor Progression and Metastasis. Clin Exp *Metastasis*, 36(3), 171-198.

Eimpunth, S., Goldenberg, A., Hamman, M. S., Oganesyan, G., Lee, R. A., Hunnangkul, S., Song, S. S., Greywal, T. & Jiang, S. I. B. (2017). Squamous cell carcinoma *in situ* upstaged to invasive

squamous cell carcinoma: a 5-year, single institution retrospective review. *Dermatol Surg*, 43(5), 698-703.

Elder, D. E. (2008). Lever's histopathology of the skin. (10th edition), Wolters Kluwer/Lippincott Williams & Williams, Philadelphia.

Elenbaas, B. & Weinberg, R. A. (2001). Heterotypic signaling between epithelial tumor cells and fibroblasts in carcinoma formation. *Exp Cell Res*, 264(1), 169-184.

El Khoury, J., Kurban, M., Kibbi, A. G. & Abbas, O. (2014). Fibroblast-activation protein: valuable marker of cutaneous epithelial malignancy. *Arch Dermatol Res*, 306(4), 359-365.

Elwakeel, E., Brüggemann, M., Fink, A. F., Schulz, M. H., Schmid, T., Savai, R., Brüne, B., Zarnack, K. & Weigert, A. (2019). Phenotypic Plasticity of Fibroblasts during Mammary Carcinoma Development. *Int J Mol Sci,* 20(18), 4438.

Elyada, E., Bolisetty, M., Laise, P., Flynn, W. F., Courtois, E. T., Burkhart, R. A., Teinor, J. A., Belleau, P., Biffi, G., Lucito, M. S., Sivajothi, S., Armstrong, T. D., Engle, D. D., Yu, K. H., Hao, Y., Wolfgang, C. L., Park, Y., Preall, J., Jaffee, E. M., Califano, A., Robson, P. & Tuveson, D. A. (2019). Cross-Species Single-Cell Analysis of Pancreatic Ductal Adenocarcinoma Reveals Antigen-Presenting Cancer-Associated Fibroblasts. *Cancer Discov*, 9(8), 1102-1123.

English, D. R., Armstrong, B. K., Kricker, A., Winter, M. G., Heenan, P. J. & Randell, P. L. (1998) Demographic characteristics, pigmentary and cutaneous risk factors for squamous cell carcinoma of the skin: a case-control study. *Int J Cancer*, 76(5), 628-634.

Erez, N., Truitt, M., Olson, P., Arron, S. T. & Hanahan, D. (2010). Cancer-Associated Fibroblasts Are Activated in Incipient Neoplasia to Orchestrate Tumor-Promoting Inflammation in an NF-kappaB-Dependent Manner. *Cancer Cell*, 17(2), 135-147.

Esquivel-Velázquez, M., Ostoa-Saloma, P., Palacios-Arreola, M. I., Nava-Castro, K. E., Castro, J. I. & Morales-Montor, J. (2015). The role of cytokines in breast cancer development and progression. *J Interferon Cytokine Res,* 35(1), 1-16.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M. & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115–118.

Euvrard, S., Kanitakis, J. & Claudy, A. (2003). Skin cancers after organ transplantation. *N Engl J Med*, 348(17), 1681-1691.

Evans, H. L. & Smith, J. L. (1980). Spindle cell squamous carcinomas and sarcoma-like tumors of the skin: a comparative study of 38 cases. Cancer, 45(10), 2687-2697.

Farasat, S., Yu, S. S., Neel, V. A., Nehal, K. S., Lardaro, T., Mihm, M. C., Byrd, D. R., Balch, C. M., Califano, J. A., Chuang, A. Y., Sharfman, W. H., Shah, J. P., Nghiem, P., Otley, C. C., Tufaro, A. P., Johnson, T. M., Sober, A. J. & Liégeois, N. J. (2011). A new American Joint Committee on Cancer staging system for cutaneous squamous cell carcinoma: creation and rationale for inclusion of tumor (T) characteristics*. J Am Acad Dermatol*, 64(6), 1051-1059.

Farmer, P., Bonnefoi, H., Anderle, P., Cameron, D., Wirapati, P., Becette, V., André, S., Piccart, M., Campone, M., Brain, E., Macgrogan, G., Petit, T., Jassem, J., Bibeau, F., Blot, E., Bogaerts, J., Aguet, M., Bergh, J., Iggo, R. & Delorenzi, M. (2009). A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer. *Nat Med*, 15(1), 68-74.

Farshchian, M., Kivisaari, A., Ala-Aho, R., Riihilä, P., Kallajoki, M., Grénman, R., Peltonen, J., Pihlajaniemi, T., Heljasvaara, R. & Kähäri, V. M. (2011). Serpin peptidase inhibitor clade A member 1 (SerpinA1) is a novel biomarker for progression of cutaneous squamous cell carcinoma. *Am J Pathol,* 179(3), 1110-1119.

Faust, H., Andersson, K., Luostarinen, T., Gislefoss, R. E. & Dillner J. (2016). Cutaneous Human Papillomaviruses and Squamous Cell Carcinoma of the Skin: Nested Case-Control Study. *Cancer Epidemiol Biomarkers Prev*, 25(4), 721-724.

Fears, T. R. & Scotto, J. (1983). Estimating increases in skin cancer morbidity due to increases in ultraviolet radiation exposure. *Cancer Invest*, 1(2):119-126.

Feng, D., Nagy, J. A., Brekken, R. A., Pettersson, A., Manseau, E. J., Pyne, K., Mulligan, R., Thorpe, P. E., Dvorak, H. F. & Dvorak, A. M. (2000). Ultrastructural localization of the vascular permeability factor/vascular endothelial growth factor (VPF/VEGF) receptor-2 (FLK-1, KDR) in normal mouse kidney and in the hyperpermeable vessels induced by VPF/VEGF-expressing tumors and adenoviral vectors. *J Histochem Cytochem*, 48(4), 545-556.

Fernandez Figueras M. T. (2017). From actinic keratosis to squamous cell carcinoma: pathophysiology revisited. *J Eur Acad Dermatol Venereol*, 31 Suppl 2, 5-7.

Fernández-Figueras, M. T., Carrato, C., Sáenz, X., Puig, L., Musulen, E., Ferrándiz, C. & Ariza, A. (2015). Actinic keratosis with atypical basal cells (AK I) is the most common lesion associated with invasive squamous cell carcinoma of the skin. *J Eur Acad Dermatol Venereol*, 29(5), 991-997.

Ferroni, P., Zanzotto, F. M., Riondino, S., Scarpato, N., Guadagni, F. & Roselli, M. (2019). Breast cancer prognosis using a machine learning approach. *Cancers (Basel)*, 11(3), 328.

Fine, J. D., Johnson, L. B., Weiner, M., Li, K. P. & Suchindran, C. (2009). Epidermolysis bullosa and the risk of life-threatening cancers: the National EB Registry experience, 1986–2006. *J Am Acad Dermatol*, 60(2), 203-211.

Finnish Cancer Registry, Cancer Statistics. https://syoparekisteri.fi/tilastot/tautitilastot/. Accessed February 1, 2022.

Fitzgerald, A. A. & Weiner, L. M. (2020). The role of fibroblast activation protein in health and malignancy. *Cancer Metastasis Rev*, 39(3), 783-803.

Folkman, J. & Kalluri, R. (2004). Cancer without disease. *Nature*, 427(6977), 787.

Forino, M., Torregrossa, R., Ceol, M., Murer, L., Vella, M. D., Prete, D. D., D'Angelo, A. & Anglani, F. (2006). TGFβ1 induces epithelial-mesenchymal transition, but not myofibroblast transdifferentiation of human kidney tubular epithelial cells in primary culture. *Int J Exp Pathol*, 87(3), 197-208.

Francisco, L. M., Sage, P. T. & Sharpe, A. H. (2010). The PD-1 pathway in tolerance and autoimmunity. *Immunol Rev*, 236, 219-42.

Frings, O., Augsten, M., Tobin, N. P., Carlson, J., Paulsson, J., Pena, C., Olsson, E., Veerla, S., Bergh, J., Ostman, A. & Sonnhammer, E. L. (2013). Prognostic significance in breast cancer of a gene signature capturing stromal PDGF signaling. *Am J Pathol*, 182(6), 2037-2047.

Fu, W. & Cockerell, C. J. (2003). The actinic (solar) keratosis: a 21st-century perspective. *Arch Dermatol*, 139(1):66-70.

Fukumura, D., Xavier, R., Sugiura, T., Chen, Y., Park, E. C., Lu, N., Selig, M., Nielsen, G., Taksir, T., Jain, R. K. & Seed, B. (1998). Tumor induction of VEGF promoter activity in stromal cells. *Cell*, 94(6), 715-725.

Gabbiani, G., Ryan, G. B. & Majne, G. (1971). Presence of Modified Fibroblasts in Granulation Tissue and Their Possible Role in Wound Contraction. *Experientia*, 27(5), 549-550.

Gaggioli, C., Hooper, S., Hidalgo-Carcedo, C., Grosse, R., Marshall, J. F., Harrington, K. & Sahai E. (2007). Fibroblast-led collective invasion of carcinoma cells with differing roles for RhoGTPases in leading and following cells. *Nat Cell Biol*, 9(12), 1392-1400.

Gajewski, T. F., Schreiber, H. & Fu, Y. X. (2013). Innate and Adaptive Immune Cells in the Tumor Microenvironment. *Nat Immunol*, 14(10), 1014-1022.

García-Pedrero, J. M., Martínez-Camblor, P., Diaz-Coto, S., Munguia-Calzada, P., Vallina-Alvarez, A., Vazquez-Lopez, F., Rodrigo, J. P. & Santos-Juanes, J. (2017). Tumor programmed cell death ligand 1 expression correlates with nodal metastasis in patients with cutaneous squamous cell carcinoma of the head and neck. *J Am Acad Dermatol*, 77(3), 527-533.

Garrett, G. L., Blanc, P. D., Boscardin, J., Lloyd, A. A., Ahmed, R. L., Anthony, T., Bibee, K., Breithaupt, A., Cannon, J., Chen, A., Cheng, J. Y., Chiesa-Fuxench, Z., Colegio, O. R., Curiel-Lewandrowski, C., Del Guzzo, C. A., Disse, M., Dowd, M., Eilers, R. Jr., Ortiz, A. E., Morris, C., Golden, S. K., Graves, M. S., Griffin, J. R., Hopkins, R. S., Huang, C. C., Bae, G. H., Jambusaria, A., Jennings, T. A., Jiang, S. I., Karia, P. S., Khetarpal, S., Kim, C., Klintmalm, G.,

Konicke, K., Koyfman, S. A., Lam, C., Lee, P., Leitenberger, J. J., Loh, T., Lowenstein, S., Madankumar, R., Moreau, J. F., Nijhawan, R. I., Ochoa, S., Olasz, E. B., Otchere, E., Otley, C., Oulton, J., Patel, P. H., Patel, V. A., Prabhu, A. V., Pugliano-Mauro, M., Schmults, C. D., Schram, S., Shih, A. F., Shin, T., Soon, S., Soriano, T., Srivastava, D., Stein, J. A., Sternhell-Blackwell, K., Taylor, S., Vidimos, A., Wu, P., Zajdel, N., Zelac, D., Arron, S. T. (2017). Incidence of and risk factors for skin cancer in organ transplant recipients in the United States. *JAMA Dermatol*, 153(3), 296-303.

Gatys, L. A., Ecker, A. S. & Bethge M. (2016). Image style transfer using convolutional neural networks. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2414–2423. doi:10.1109/CVPR.2016.265.

Genders, R. E., Weijns, M. E., Dekkers, O. M. & Plasmeijer, E. I. (2019). Metastasis of cutaneous squamous cell carcinoma in organ transplant recipients and the immunocompetent population: is there a difference? a systematic review and meta-analysis. *J Eur Acad Dermatol Venereol*, 33(5), 828-841.

Givi, B., Andersen, P. E., Diggs, B. S., Wax, M. K. & Gross, N. D. (2011). Outcome of patients treated surgically for lymph node metastases from cutaneous squamous cell carcinoma of the head and neck. *Head Neck*, 33(7), 999-1004.

Goetz, J. G., Minguet, S., Navarro-Lérida, I., Lazcano, J. J., Samaniego, R., Calvo, E., Tello, M., Osteso-Ibáñez, T., Pellinen, T., Echarri, A., Cerezo, A., Klein-Szanto, A. J., Garcia, R., Keely, P. J., Sánchez-Mateos, P., Cukierman, E. & Del Pozo, M. A. (2011). Biomechanical remodeling of the microenvironment by stromal caveolin-1 favors tumor invasion and metastasis. *Cell*, 146(1), 148-163.

González-González, L. & Alonso, J. (2018). Periostin: A Matricellular Protein With Multiple Functions in Cancer Development and Progression. *Front Oncol*, 8, 225.

Goodfellow, I., Bengio, Y. & Courville, A. (2016). Deep learning. Cambridge, Massachusetts: The MIT Press.

Granter, S. R., Beck, A. H. & Papke, D. J. (2017). AlphaGo, deep learning, and the future of the human microscopist. *Arch Pathol Lab Med*, 141(5), 619-621.

Green, A. C. & Olsen, C. M. (2017). Cutaneous squamous cell carcinoma: an epidemiological review. *Br J Dermatol*, 177(2), 373-381.

Gurney, B. & Newlands, C. (2014). Management of regional metastatic disease in head and neck cutaneous malignancy. 1. Cutaneous squamous cell carcinoma. *Br J Oral Maxillofac Surg,* 52(4), 294-300.

Guweidhi, A., Kleeff, J., Adwan, H., Giese, N. A., Wente, M. N., Giese, T., Büchler, M. W., Berger, M. R. & Friess, H. (2005). Osteonectin influences growth and invasion of pancreatic cancer cells. *Ann Surg*, 242(2), 224-34.

Haenssle, H. A., Fink, C., Schneiderbauer, R., Toberer, F., Buhl, T., Blum, A., Kalloo, A., Hassen, A. B. H., Thomas, L., Enk, A., Uhlmann, L., Reader study level-I and level-II Groups, Alt, C., Arenbergerova, M., Bakos, R., Baltzer, A., Bertlich, I., Blum, A., Bokor-Billmann, T., Bowling, J., Braghiroli, N., Braun, R., Buder-Bakhaya, K., Buhl, T., Cabo, H., Cabrijan, L., Cevic, N., Classen, A., Deltgen, D., Fink, C., Georgieva, I., Hakim-Meibodi, L. E., Hanner, S., Hartmann, F., Hartmann, J., Haus, G., Hoxha, E., Karls, R., Koga, H., Kreusch, J., Lallas, A., Majenka, P., Marghoob, A., Massone, C., Mekokishvili, L., Mestel, D., Meyer, V., Neuberger, A., Nielsen, K., Oliviero, M., Pampena, R., Paoli, J., Pawlik, E., Rao, B., Rendon, A., Russo, T., Sadek, A., Samhaber, K., Schneiderbauer, R., Schweizer, A., Toberer, F., Trennheuser, L., Vlahova, L., Wald, A., Winkler, J., Wölbing, P. & Zalaudek, I. (2018). Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Ann Oncol*, 29(8), 1836–1842.

Halifu, Y., Liang, J. Q., Zeng, X. W., Ding, Y., Zhang, X. Y., Jin, T. B., Yakeya, B., Abudu, D., Zhou, Y. M., Liu, X. M., Hu, F. X., Chai, L. & Kang, X. J. (2016). Wnt1 and SFRP1 as potential

prognostic factors and therapeutic targets in cutaneous squamous cell carcinoma. *Genet Mol Res,* 15(2), doi: 10.4238/gmr.15028187.

Ham, I. H., Lee, D. & Hur, H. (2019). Role of cancer-associated fibroblast in gastric cancer progression and resistance to treatments. *J Oncol,* 2019, 6270784.

Han, S. S., Kim, M. S., Lim, W., Park, G. H., Park, I. & Chang, S. E. (2018). Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm. *J Invest Dermatol*, 138(7), 1529-1538.

Hanahan, D. & Coussens, L. M. (2012). Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell*, 21(3), 309-322.

Hanahan, D. & Weinberg, R.A. (2011). Hallmarks of Cancer: The Next Generation. *Cell,* 144(5), 646-674.

Hart, S. N., Flotte, W., Norgan, A. P., Shah, K. K., Buchan, Z. R., Mounajjed, T. & Flotte, T. J. (2019). Classification of melanocytic lesions in selected and whole-slide images via convolutional neural networks. *J Pathol Inform*, 10, 5.

Hastie, T., Tibshirani, R. & Friedman, J. H. (2009). The elements of statistical learning : data mining, inference, and prediction. 2nd ed New York, NY: Springer; 2009.

Haviv, I., Polyak, K., Qiu, W., Hu, M. & Campbell, I. (2009). Origin of Carcinoma Associated Fibroblasts. *Cell Cycle,* 8(4), 589–595.

Hawinkels, L. J., Paauwe, M., Verspaget, H. W., Wiercinska, E., van der Zon, J. M., van der Ploeg, K., Koelink, P. J., Lindeman, J. H., Mesker, W., ten Dijke, P. & Sier, C. F. (2014). Interaction With Colon Cancer Cells Hyperactivates Tgf-β Signaling in Cancer-Associated Fibroblasts. *Oncogene*, 33(1), 97-107.

He, K., Zhang, X., Ren, S. & Sun, J. (2016). Deep Residual Learning for Image Recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778.

Heaphry M. R. & Ackerman, A. B. (2000). The nature of solar keratosis: a critical review in historical perspective. J Am Acad Dermatol, 43(1 Pt 1), 138-150.

Hematti, P. (2012). Mesenchymal stromal cells and fibroblasts: a case of mistaken identity? *Cytotherapy*, 14(5), 516-521.

Hernberg, M., Ilmonen, S., Juteau, S., Jääskeläinen, A., Koljonen, V., Koskenmies, S., Leivo, T., Mäkitie, A., Pekkonen, P., Pitkänen, S. & Uusitalo, M. (2020). Kansallinen ei-melanoottisten ihosyöpien hoito-ohjeistus. https://www.terveysportti.fi/apps/ltk/article/hsu00009?toc=23040

Hirshoren, N., Danne, J., Dixon, B. J., Magarey, M., Kleid, S., Webb, A., Tiong, A., Corry, J. & Gyorki, D. (2017). Prognostic markers in metastatic cutaneous squamous cell carcinoma of the head and neck. *Head Neck*, 39(4), 772-778.

Hmadcha, A., Martin-Montalvo, A., Gauthier, B. R., Soria, B. & Capilla-Gonzalez, V. (2020). Therapeutic Potential of Mesenchymal Stem Cells for Cancer Therapy. *Front Bioeng Biotechnol,* 8, 43.

Hodak, E., Jones, R. E. & Ackerman, A. B. (1993). Solitary keratoacanthoma is a squamous-cell carcinoma: three examples with metastases. *Am J Dermatopathol*, 15(4), 332-342.

Hofheinz, R. D., al-Batran, S. E., Hartmann, F., Hartung, G., Jäger, D., Renner, C., Tanswell, P., Kunz, U., Amelsberg, A., Kuthan, H. & Stehle, G. (2003). Stromal Antigen Targeting by a Humanised Monoclonal Antibody: An Early Phase II Trial of Sibrotuzumab in Patients With Metastatic Colorectal Cancer. *Onkologie*, 26(1), 44-48.

Holten-Rossing, H., Talman, M. M., Jylling, A. M. B., Laenkholm, A. V., Kristensson, M. & Vainer, B. (2017). Application of automated image analysis reduces the workload of manual screening of sentinel lymph node biopsies in breast cancer. *Histopathology*, 71(6), 866-873.

Hosaka, K., Yang, Y., Seki, T., Fischer, C., Dubey, O., Fredlund, E., Hartman, J., Religa, P., Morikawa, H., Ishii, Y., Sasahara, M., Larsson, O., Cossu, G., Cao, R., Lim, S. & Cao, Y. (2016). Pericyte-Fibroblast Transition Promotes Tumor Growth and Metastasis. *Proc Natl Acad Sci U S A*, 113(38), E5618-5627.

Hosein, A. N., Huang, H., Wang, Z., Parmar, K., Du, W., Huang, J., Maitra, A., Olson, E., Verma, U. & Brekken, R. A. (2019). Cellular Heterogeneity During Mouse Pancreatic Ductal Adenocarcinoma Progression at Single-Cell Resolution. *JCI Insight*, 5(16), e129212.

Hu, B., Castillo, E., Harewood, L., Ostano, P., Reymond, A., Dummer, R., Raffoul, W., Hoetzenecker, W., Hofbauer, G. F. & Dotto, G. P. (2012). Multifocal epithelial tumors and field cancerization from loss of mesenchymal CSL signaling. Cell, 149(6), 1207–1220.

Huber, M. A., Kraut, N., Park, J. E., Schubert, R. D., Rettig, W. J., Peter, R. U. & Garin-Chesa, P. (2003). Fibroblast Activation Protein: Differential Expression and Serine Protease Activity in Reactive Stromal Fibroblasts of Melanocytic Skin Tumors. *J Invest Dermatol*, 120(2), 182-188.

Hägglöf, C., Hammarsten, P., Josefsson, A., Stattin, P., Paulsson, J., Bergh, A. & Östman, A. (2010). Stromal PDGFRbeta expression in prostate tumors and non-malignant prostate tissue predicts prostate cancer survival. *PLoS One*, 5(5), e10747.

Inman, G. J., Wang, J., Nagano, A., Alexandrov, L. B., Purdie, K. J., Taylor, R. G., Sherwood, V., Thomson, J., Hogan, S., Spender, L. C., South, A. P., Stratton, M., Chelala, C., Harwood, C. A., Proby, C. M. & Leigh, I. M. (2018). The genomic landscape of cutaneous SCC reveals drivers and a novel azathioprine associated mutational signature. *Nat Commun*, 9(1), 3667.

Ishii, G., Ochiai, A. & Neri, S. (2016). Phenotypic and functional heterogeneity of cancer-associated fibroblast within the tumor microenvironment. *Adv Drug Deliv Rev*, 99(Pt B), 186-196.

Ishii, G., Sangai, T., Oda, T., Aoyagi, Y., Hasebe, T., Kanomata, N., Endoh, Y., Okumura, C., Okuhara, Y., Magae, J., Emura, M., Ochiya, T. & Ochiai, A. (2003). Bone-marrow-derived myofibroblasts contribute to the cancer-induced stromal reaction. *Biochem Biophys Res Commun*, 309(1), 232-240.

Iwano, M., Plieth, D., Danoff, T. M., Xue, C., Okada, H. & Neilson, E. G. (2002). Evidence that fibroblasts derive from epithelium during tissue fibrosis. *J Clin Invest*, 110(3), 341-350.

Jaju, P. D., Ransohoff, K. J., Tang, J. Y. & Sarin, K. Y. (2016). Familial skin cancer syndromes: increased risk of nonmelanotic skin cancers and extracutaneous tumors. *J Am Acad Dermatol,* 74(3), 437-451.

Jambusaria-Pahlajani, A., Kanetsky, P. A., Karia, P. S., Hwang, W. T., Gelfand, J. M., Whalen, F. M., Elenitsas, R., Xu, X. & Schmults, C. D. (2013). Evaluation of AJCC tumor staging for cutaneous squamous cell carcinoma and a proposed alternative tumor staging system. *JAMA Dermatol*, 149(4), 402-410.

Jensen, A. Ø., Bautz, A., Olesen, A. B., Karagas, M. R., Sørensen, H. T. & Friis, S. (2008). Mortality in Danish patients with nonmelanoma skin cancer, 1978-2001. *Br J Dermatol*, 159(2), 419-425.

Jiyad, Z., O'Rourke, P., Soyer, H. P. & Green, A. C. (2017). Actinic keratosis-related signs predictive of squamous cell carcinoma in renal transplant recipients: a nested case–control study. *Br J Dermatol*, 176(4), 965-970.

Jiyad, Z., Olsen, C. M., Burke, M. T., Isbel, N. M. & Green, A. C. (2016). Azathioprine and risk of skin cancer in organ transplant recipients: systematic review and meta-analysis. *Am J Transplant*, 16(12), 3490–3503.

Johnson, T. M., Rowe, D. E., Nelson, B. R. & Swanson, N. A. (1992). Squamous cell carcinoma of the skin (excluding lip and oral mucosa). *J Am Acad Dermatol,* 26(3 Pt 2), 467-84.

Joshi, R. S., Kanugula, S. S., Sudhir, S., Pereira, M. P., Jain, S. & Aghi, M. K. (2021). The Role of Cancer-Associated Fibroblasts in Tumor Progression. *Cancers (Basel)*, 13(6), 1399.

Jotzu, C., Alt, E., Welte, G., Li, J., Hennessy, B. T., Devarajan, E., Krishnappa, S., Pinilla, S., Droll, L. & Song, Y. H. (2010). Adipose tissue-derived stem cells differentiate into carcinoma-associated fibroblast-like cells under the influence of tumor-derived factors. *Anal Cell Pathol (Amst)*, 33(2), 61-79.

Kahounová, Z., Kurfürstová, D., Bouchal, J., Kharaishvili, G., Navrátil, J., Remšík, J., Šimečková, Š., Študent, V., Kozubík, A. & Souček, K. (2018). The fibroblast surface markers FAP, anti-fibroblast, and FSP are expressed by cells of epithelial origin and may be altered during epithelial-to-mesenchymal transition. *Cytometry A*, 93(9), 941-951.

Kalluri, R. (2016). The biology and function of fibroblasts in cancer. *Nat. Rev. Cancer*, *16*, 582–598.

Kalluri, R. & Weinberg, R. A. (2009). The basics of epithelialmesenchymal transition. J Clin Invest, 119(6), 1420-1428.

Kalluri, R. & Zeisberg, M. (2006). Fibroblasts in cancer. *Nat Rev Cancer*, 6(5), 392-401.

Kang, S. Y. & Toland, A. E. (2016). High risk cutaneous squamous cell carcinoma of the head and neck. *World J Otorhinolaryngol Head Neck Surg*, 2(2), 136-140.

Karayannopoulou, G., Euvrard, S. & Kanitakis, J. (2016). Tumour Budding Correlates With Aggressiveness of Cutaneous Squamous-cell Carcinoma. *Anticancer Res*, 36(9), 4781-4785.

Karia, P. S., Jambusaria-Pahlajani, A., Harrington, D. P., Murphy, G. F., Qureshi, A. A. & Schmults, C. D. (2014). Evaluation of American Joint Committee on Cancer, International Union Against Cancer, and Brigham and Women's Hospital tumor staging for cutaneous squamous cell carcinoma. *J Clin Oncol*, 32(4), 327-334.

Karia, P. S., Han, J. & Schmults, C. D. (2013) Cutaneous squamous cell carcinoma: estimated incidence of disease, nodal metastasis, and deaths from disease in the United States, 2012. *J Am Acad Dermatol*, 68(6), 957-966.

Karia, P. S., Morgan, F. C., Califano, J. A. & Schmults, C. D. (2018). Comparison of Tumor Classifications for Cutaneous Squamous Cell Carcinoma of the Head and Neck in the 7th vs 8th Edition of the AJCC Cancer Staging Manual. *JAMA Dermatol*, 154(2), 175-181.

Karlsson, M. C., Gonzalez, S. F., Welin, J. & Fuxe, J. (2017). Epithelial-mesenchymal transition in cancer metastasis through the lymphatic system. *Mol Oncol*, 11(7), 781-791.

Karnoub, A. E., Dash, A. B., Vo, A. P., Sullivan, A., Brooks, M. W., Bell, G. W., Richardson, A. L., Polyak, K., Tubo, R. & Weinberg, R. A. (2007). Mesenchymal stem cells within tumour stroma promote breast cancer metastasis. *Nature*, 449(7162), 557-563.

Karppinen, S. M., Honkanen, H. K., Heljasvaara, R., Riihilä, P., Autio-Harmainen, H., Sormunen, R., Harjunen, V., Väisänen, M. R., Väisänen, T., Hurskainen, T., Tasanen, K., Kähäri, V. M. & Pihlajaniemi, T. (2016). Collagens XV and XVIII show different expression and localisation in cutaneous squamous cell carcinoma: type XV appears in tumor stroma while XVIII becomes upregulated in tumor cells and lost from microvessels. *Exp Dermatol*, 25(5), 348–354.

Kather, J. N., Pearson, A. T., Halama, N., Jäger, D., Krause, J., Loosen, S. H., Marx, A., Boor, P., Tacke, F., Neumann, U. P., Grabsch, H. I., Yoshikawa, T., Brenner, H., Chang-Claude, J., Hoffmeister, M., Trautwein, C. & Luedde, T. (2019). Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med*, 25(7), 1054-1056.

Khandelwal, A. R., Ma, X., Egan, P., Kaskas, N. M., Moore-Medlin, T., Caldito, G., Abreo, F., Gu, X., Aubrey, L., Milligan, E., Nathan, C. A. (2016). Biomarker and Pathologic Predictors of Cutaneous Squamous Cell Carcinoma Aggressiveness. *Otolaryngol Head Neck Surg,* 155(2), 281-288.

Kieffer, Y., Hocine, H. R., Gentric, G., Pelon, F., Bernard, C., Bourachot, B., Lameiras, S., Albergante, L., Bonneau, C., Guyard, A., Tarte, K., Zinovyev, A., Baulande, S., Zalcman, G., Vincent-Salomon, A. & Mechta-Grigoriou, F. (2020). Single-Cell Analysis Reveals Fibroblast Clusters Linked to Immunotherapy Resistance in Cancer. *Cancer Discov*, 10(9), 1330-1351.

Kilvaer, T. K., Rakaee, M., Hellevik, T., Vik, J., Petris, L., Donnem, T., Strell, C., Ostman, A., Busund, L. R. & Martinez-Zubiaurre, I. (2019). Differential prognostic impact of platelet-derived growth factor receptor expression in NSCLC. *Sci Rep*, 9(1), 10163.

Kim, G. E., Lee, J. S., Park, M. H. & Yoon, J. H. (2017). Epithelial periostin expression is correlated with poor survival in patients with invasive breast carcinoma. *PLoS One*, 12(11), e0187635.

Kim, M., Yun, J., Cho, Y., Shin, K., Jang, R., Bae, H. J. & Kim, N. (2019). Deep learning in medical imaging. *Neurospine*, 16(4), 657-668.

Kitadai, Y., Sasaki, T., Kuwai, T., Nakamura, T., Bucana, C. D., Hamilton, S. R. & Fidler, I. J. (2006). Expression of activated platelet-derived growth factor receptor in stromal cells of human colon carcinomas is associated with metastatic potential. *Int J Cancer*, 119(11), 2567-2574.

Kitano, H., Kageyama, S., Hewitt, S. M., Hayashi, R., Doki, Y., Ozaki, Y., Fujino, S., Takikita, M., Kubo, H. & Fukuoka, J. (2010). Podoplanin Expression in Cancerous Stroma Induces Lymphangiogenesis and Predicts Lymphatic Spread and Patient Survival. *Arch Pathol Lab Med*, 134(10):1520-1527.

Kivisaari, A. K., Kallajoki, M., Ala-aho, R., McGrath, J. A., Bauer, J. W., Königová, R., Medvecz, M., Beckert, W., Grénman, R. & Kähäri, V. M. (2010). Matrix metalloproteinase-7 activates heparin-binding epidermal growth factor-like growth factor in cutaneous squamous cell carcinoma. *Br J Dermatol*, 163(4), 726-735.

Kivisaari, A. K., Kallajoki, M., Mirtti, T., McGrath, J. A., Bauer, J. W., Weber, F., Königová, R., Sawamura, D., Sato-Matsumura, K. C., Shimizu, H., Csikós, M., Sinemus, K., Beckert, W. & Kähäri, V. M. (2008). Transformation-specific matrix metalloproteinases (MMP)-7 and MMP-13 are expressed by tumour cells in epidermolysis bullosa-associated squamous cell carcinomas. *Br J Dermatol,* 158(4), 778-785.

Kivisaari, A. & Kähäri, V. M. (2013). Squamous cell carcinoma of the skin: Emerging need for novel biomarkers. *World J Clin Oncol,* 4(4), 85-90.

Klauschen, F., Wienert, S., Schmitt, W. D., Loibl, S., Gerber, B., Blohmer, J. U., Huober, J., Rüdiger, T., Erbstößer, E., Mehta, K., Lederer, B., Dietel, M., Denkert, C. & von Minckwitz, G. (2015). Standardized Ki67 diagnostics using automated scoring-clinical validation in the GeparTrio Breast Cancer Study. *Clin Cancer Res*, 21(16), 3651-3657.

Klimov, S., Xue, Y., Gertych, A., Graham, R. P., Jiang, Y., Bhattarai, S., Pandol, S. J., Rakha, E. A., Reid, M. D. & Aneja, R. (2021). Predicting metastasis risk in pancreatic neuroendocrine tumors using deep learning image analysis. *Front Oncol*, 10, 593211.

Kodama, M., Kitadai, Y., Sumida, T., Ohnishi, M., Ohara, E., Tanaka, M., Shinagawa, K., Tanaka, S., Yasui, W. & Chayama, K. (2010). Expression of platelet-derived growth factor (PDGF)-B and PDGF-receptor β is associated with lymphatic metastasis in human gastric carcinoma. *Cancer Sci*, 101(9), 1984-1989.

Kojima, Y., Acar, A., Eaton, E. N., Mellody, K. T., Scheel, C., Ben-Porath, I., Onder, T. T., Wang, Z. C., Richardson, A. L., Weinberg, R. A. & Orimo, A. (2010). Autocrine TGF- and stromal cell-derived factor-1 (SDF-1) signaling drives the evolution of tumor-promoting mammary stromal myofibroblasts. Proc Natl Acad Sci U S A, 107(46), 20009–20014.

Komura, D. & Ishikawa, S. (2018). Machine learning methods for histopathological image analysis. *Comput Struct Biotechnol J*, 16, 34-42.

Korhonen, N., Ylitalo, L., Luukkaala, T., Itkonen, J., Häihälä, H., Jernman, J., Snellman, E. & Palve, J. (2019). Characteristics and Trends of Cutaneous Squamous Cell Carcinoma in a Patient Cohort in Finland 2006-2015. *Acta Derm Venereol,* 99(4), 412-416.

Kothari, S., Phan, J. H., Stokes, T. H., Osunkoya, A. O., Young, A. N. & Wang, M. D. (2014). Removing batch effects from histopathological images for enhanced cancer diagnosis. *IEEE J Biomed Health Inform*, 18(3), 765-772.

Kovarik, C., Lee, I., Ko, J. & Ad Hoc Task Force on Augmented Intelligence. (2019). Commentary: Position statement on augmented intelligence (AuI). *J Am Acad Dermatol*, 81(4), 998-1000.

Kraman, M., Bambrough, P. J., Arnold, J. N., Roberts, E. W., Magiera, L., Jones, J. O., Gopinathan, A., Tuveson, D. A. & Fearon, D. T. (2010). Suppression of antitumor immunity by stromal cells expressing fibroblast activation protein-alpha. *Science*, 330(6005), 827-830.

Krizhevsky, A., Sutskever, I. & Hinton, G. E. (2017). ImageNet Classification with Deep Convolutional Neural Networks. *Commun Acm*, 60, 84–90.

Kulkarni, P. M., Robinson, E. J., Sarin Pradhan, J., Gartrell-Corrado, R. D., Rohr, B. R., Trager, M. H., Geskin, L. J., Kluger, H. M., Wong, P. F., Acs, B., Rizk, E. M., Yang, C., Mondal, M., Moore, M. R., Osman, I., Phelps, R., Horst, B. A., Chen, Z. S., Ferringer, T., Rimm, D. L., Wang, J. & Saenger, Y. M. (2020). Deep learning based on standard H&E images of primary melanoma tumors identifies patients at risk for visceral recurrence and death. *Clin Cancer Res*, 26(5), 1126-1134.

Kuo, T. (1980). Clear cell carcinoma of the skin. A variant of the squamous cell carcinoma that simulates sebaceous carcinoma. *Am J Surg Pathol*, 4(6), 573-583.

Kwiek, B. & Schwartz, R. A. (2016). Keratoacanthoma (KA): An update and review. *J Am Acad Dermatol*, 74(6), 1220-1233.

Lakins, M. A., Ghorani, E., Munir, H., Martins, C. P. & Shields, J. D. (2018). Cancer-Associated Fibroblasts Induce Antigen-Specific Deletion of CD8 + T Cells to Protect Tumour Cells. *Nat Commun,* 9(1), 948.

Lam, J. K. S., Sundaresan, P., Gebski, V. & Veness, M. J. (2018). Immunocompromised patients with metastatic cutaneous nodal squamous cell carcinoma of the head and neck: Poor outcome unrelated to the index lesion. *Head Neck*, 40(5), 985-992.

LeCun, Y., Bengio, Y. & Hinton, G. (2015). Deep learning. *Nature*. 521(7553), 436-444.

Leung, D. W., Cachianes, G., Kuang, W. J., Goeddel, D. V. & Ferrara, N. (1989). Vascular endothelial growth factor is a secreted angiogenic mitogen. *Science,* 246(4935), 1306-1309.

Levi, F., Randimbison, L., Te, V. C. & La Vecchia, C. (1996). Non-Hodgkin's lymphomas, chronic lymphocytic leukaemias and skin cancers. *Br J Cancer*, 74(11), 1847-1850.

Levine, D. E., Karia, P. S. & Schmults, C. D. (2015). Outcomes of patients with multiple cutaneous squamous cell carcinomas: a 10-year single-institution cohort study. *JAMA Dermatol,* 151(11), 1220-1225.

Li, H., Courtois, E. T., Sengupta, D., Tan, Y., Chen, K. H., Goh, J. J. L., Kong, S. L., Chua, C., Hon, L. K., Tan, W. S., Wong, M., Choi, P. J., Wee, L. J. K., Hillmer, A. M., Tan, I. B., Robson, P. & Prabhakar, S. (2017). Reference Component Analysis of Single-Cell Transcriptomes Elucidates Cellular Heterogeneity in Human Colorectal Tumors. *Nat Genet*, 49(5), 708-718.

Li, J., Yang, M., Li, P., Su, Z., Gao, P. & Zhang, J. (2014). Idiopathic pulmonary fibrosis will increase the risk of lung cancer. *Chin Med J (Engl)*, 127(17), 3142-3149.

Li, Y. Y., Hanna, G. J., Laga, A. C., Haddad, R. I., Lorch, J. H. & Hammerman, P. S. (2015). Genomic analysis of metastatic cutaneous squamous cell carcinoma. *Clin Cancer Res,* 21(6), 1447-1456.

Liang, S. B., Ohtsuki, Y., Furihata, M., Takeuchi, T., Iwata, J., Chen, B. K. & Sonobe, H. (1999). Sun-exposure- and aging-dependent p53 protein accumulation results in growth advantage for tumour cells in carcinogenesis of nonmelanocytic skin cancer. *Virchows Arch*, 434(3), 193-199.

Liang, S. C., Latchman, Y. E., Buhlmann, J. E., Tomczak, M. F., Horwitz, B. H., Freeman, G. J., Sharpe & A. H. (2003). Regulation of PD-1, PD-L1, and PD-L2 expression during normal and autoimmune responses. *Eur J Immunol,* 33(10), 2706-2716.

Lindelöf, B., Sigurgeirsson, B., Gäbel, H. & Stern, R. S. (2000). Incidence of skin cancer in 5356 patients following organ transplantation. *Br J Dermatol*, 143(3), 513-519.

Lindelöf, B., Sigurgeirsson, B., Tegner, E., Larkö, O., Johannesson, A., Berne, B., Ljunggren, B., Andersson, T., Molin, L., Nylander-Lundqvist, E. & Emtestam, L. (1999). PUVA and cancer risk: the Swedish follow-up study. *Br J Dermatol*, 141(1), 108-112.

Lincoln, V., Chao, L., Woodley, D. T., Murrell, D., Kim, M., O'Toole, E. A., Ly, A., Cogan, J., Mosallaei, D., Wysong, A. & Chen, M. (2021). Over-expression of stromal periostin correlates with poor prognosis of cutaneous squamous cell carcinomas. *Exp Dermatol*, 30(5), 698-704.

Litjens, G., Sánchez, C. I., Timofeeva, N., Hermsen, M., Nagtegaal, I., Kovacs, I., Hulsbergen-van de Kaa, C., Bult, P., van Ginneken, B. & van der Laak, J. (2016). Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Sci Rep,* 6, 26286.

Liu, F., Qi, L., Liu, B., Liu, J., Zhang, H., Che, D., Cao, J., Shen, J., Geng, J., Bi, Y., Ye, L., Pan, B., Yu, Y. (2015). Fibroblast Activation Protein Overexpression and Clinical Implications in Solid Tumors: A Meta-Analysis. *PLoS One*, 10(3), e0116683.

Liu, T., Han, C., Wang, S., Fang, P., Ma, Z., Xu, L. & Yin, R. (2019a). Cancer-Associated Fibroblasts: An Emerging Target of Anti-Cancer Immunotherapy. *J Hematol Oncol*, 12(1), 86.

Liu, Y., Kohlberger, T., Norouzi, M., Dahl, G. E., Smith, J. L., Mohtashamian, A., Olson, N., Peng, L. H., Hipp, J. D. & Stumpe, M. C. (2019b). Artificial Intelligence-Based Breast Cancer Nodal

Metastasis Detection: Insights Into the Black Box for Pathologists. *Arch Pathol Lab Med*, 143(7), 859-868.

Louault, K., Li, R. R. & Declerck, Y. A. (2020). Cancer-associated fibroblasts: Understanding their heterogeneity. *Cancers (Basel),* 12(11), 3108.

Lu, C., Lewis, J. S., Dupont, W. D., Plummer, W. D., Janowczyk, A. & Madabhushi, A. (2017). An oral cavity squamous cell carcinoma quantitative histomorphometric-based image classifier of nuclear morphology can risk stratify patients for disease-specific survival. *Mod Pathol*, 30(12), 1655-1665.

Löhr, M., Schmidt, C., Ringel, J., Kluth, M., Müller, P., Nizze, H. & Jesnowski, R. (2001). Transforming growth factor-β1 induces desmoplasia in an experimental model of human pancreatic carcinoma. Cancer Res. 61(2), 550-555.

MacFarlane, D., Shah, K., Wysong, A., Wortsman, X. & Humphreys, T. R. (2017). The role of imaging in the management of patients with nonmelanoma skin cancer: diagnostic modalities and applications *J Am Acad Dermatol*, 76(4), 579-588.

Madeleine, M. M., Patel, N. S., Plasmeijer, E. I., Engels, E. A., Bouwes Bavinck, J. N., Toland, A. E. & Green, A. C.; the Keratinocyte Carcinoma Consortium (KeraCon) Immunosuppression Working Group. (2017). Epidemiology of keratinocyte carcinomas after organ transplantation. *Br J Dermatol*. 177(5), 1208-1216.

Majores, M. & Bierhoff, E. (2015). Aktinische Keratose, Morbus Bowen, Keratoakanthom und Plattenepithelkarzinom der Haut. *Pathologe*, 36(1), 16-29.

Manyam, B. V., Garsa, A. A., Chin, R. I., Reddy, C. A., Gastman, B., Thorstad, W., Yom, S. S., Nussenbaum, B., Wang, S. J., Vidimos, A. T. & Koyfman, S. A. (2017) A multi-institutional comparison of outcomes of immunosuppressed and immunocompetent patients treated with surgery and radiation therapy for cutaneous squamous cell carcinoma of the head and neck. *Cancer,* 123(11), 2054-2060.

Markham, A. & Duggan, S. (2018). Cemiplimab: First Global Approval. *Drugs*, 78(17), 1841-1846.

Marks, R., Rennie, G. & Selwood, T. S. (1988). Malignant transformation of solar keratoses to squamous cell carcinoma. *Lancet*, 1(8589), 795-797.

Marsh, T., Pietras, K. & McAllister, S. S. (2013). Fibroblasts as architects of cancer pathogenesis. *Biochim Biophys Acta*, 1832(7), 1070-1078.

Martincorena, I., Roshan, A., Gerstung, M., Ellis, P., Van Loo, P., McLaren, S., Wedge, D. C., Fullam, A., Alexandrov, L. B., Tubio, J. M., Stebbings, L., Menzies, A., Widaa, S., Stratton, M. R., Jones, P. H. & Campbell, P. J. (2015). High burden and pervasive positive selection of somatic mutations in normal human skin. *Science*, 348(6237), 880–886.

Martinez, J. C., Otley, C. C., Okuno, S. H., Foote, R. L. & Kasperbauer, J. L. (2004). Chemotherapy in the management of advanced cutaneous squamous cell carcinoma in organ transplant recipients: theoretical and practical considerations. *Dermatol Surg*, 30(4 Pt 2), 679-686.

Martínez-Martínez, E., Tölle, R., Donauer, J., Gretzmeier, C., Bruckner-Tuderman, L. & Dengjel, J. (2021). Increased abundance of Cbl E3 ligases alters PDGFR signaling in recessive dystrophic epidermolysis bullosa. *Matrix Biol*, 103-104, 58-73.

Martinez-Outschoorn, U. E., Lisanti, M. P. & Sotgia, F. (2014). Catabolic cancer-associated fibroblasts transfer energy and biomass to anabolic cancer cells, fueling tumor growth. *Semin Cancer Biol*, 25, 47-60.

Maubec, E. (2020). Update of the Management of Cutaneous Squamous-cell Carcinoma. *Acta Derm Venereol*, 100(11), adv00143.

McCluggage, W. G. (2013). Premalignant lesions of the lower female genital tract: cervix, vagina and vulva. Pathology, 45(3), 214-228.

Mendt, M., Rezvani, K. & Shpall, E. (2019). Mesenchymal Stem Cell-Derived Exosomes for Clinical Use. *Bone Marrow Transplant*, 54(Suppl 2), 789-792.

Mezawa, Y. & Orimo, A. (2016). The roles of tumor- and metastasis-promoting carcinoma-associated fibroblasts in human carcinomas. *Cell Tissue Res,* 365(3), 675-689.

Mezheyeuski, A., Segersten, U., Leiss, L. W., Malmström, P. U., Hatina, J., Östman, A. & Strell, C. (2020). Fibroblasts in urothelial bladder cancer define stroma phenotypes that are associated with clinical outcome. *Sci Rep,* 10(1), 281.

Micallef, L., Vedrenne, N., Billet, F., Coulomb, B., Darby, I. A. & Desmoulière, A. (2012). The myofibroblast, multiple origins for major roles in normal and pathological tissue repair. *Fibrogenesis Tissue Repair*, 5(Suppl 1), S5.

Micke, P. & Östman, A. (2004). Tumour-stroma interaction: cancer-associated fibroblasts as novel targets in anti-cancer therapy? *Lung Cancer*, 45 Suppl 2, S163-175.

Miller, D. L. & Weinstock, M. A. (1994). Nonmelanoma skin cancer in the united states: incidence. *J Am Acad Dermatol*, 30(5 Pt 1), 774-778.

Mobadersany, P., Yousefi, S., Amgad, M., Gutman, D. A., Barnholtz-Sloan, J. S., Velázquez Vega, J. E., Brat, D. J. & Cooper, L. A. D. (2018). Predicting cancer outcomes from histology and genomics using convolutional networks. *Proc Natl Acad Sci U S A*, 115(13), E2970-E2979.

Moles Lopez, X., D'Andrea, E., Barbot, P., Bridoux, A. S., Rorive, S., Salmon, I., Debeir, O. & Decaestecker, C. (2013). An automated blur detection method for histological whole slide imaging. *PLoS One*, 8(12), e82710.

Moloney, F. J., Comber, H., O'Lorcain, P., O'Kelly, P., Conlon, P. J. & Murphy, G. M. (2006). A population-based study of skin cancer incidence and prevalence in renal transplant recipients. *Br J Dermatol*, 154(3), 498-504.

Moore, B. A., Weber, R. S., Prieto, V., El-Naggar, A., Holsinger, F. C., Zhou, X., Lee, J. J., Lippman, S. & Clayman, G. L. (2005). Lymph node metastases from cutaneous squamous cell carcinoma of the head and neck. *Laryngoscope*, 115(9), 1561-1567.

Mora, R. G. & Perniciaro, C. (1981). Cancer of the skin in blacks. I. A review of 163 black patients with cutaneous squamous cell carcinoma. *J Am Acad Dermatol,* 5(5), 535-543.

Morton, C. A., Birnie, A. J. & Eedy, D. J. (2014). British Association of Dermatologists' guidelines for the management of squamous cell carcinoma in situ (Bowen's disease). *Br J Dermatol*, 170(2), 245-260.

Mourouzis, C., Boynton, A., Grant, J., Umar, T., Wilson, A., Macpheson, D. & Pratt, C. (2009). Cutaneous head and neck SCCs and risk of nodal metastasis—UK experience. *J Craniomaxillofac Surg*, 37(8), 443-447.

Motley, R., Kersey, P. & Lawrence, C; British Association of Dermatologists; British Association of Plastic Surgeons; Royal College of Radiologists, Faculty of Clinical Oncology. (2002). Multiprofessional guidelines for the management of the patient with primary cutaneous squamous cell carcinoma. *Br J Dermatol*, 146(1), 18-25.

Motley, S. & Arron, S. (2019). Mohs micrographic surgery for cutaneous squamous cell carcinoma. *Br J Dermatol*, 181(2), 233-234.

Mukhopadhyay, S., Feldman, M. D., Abels, E., Ashfaq, R., Beltaifa, S., Cacciabeve, N. G., Cathro, H. P., Cheng, L., Cooper, K., Dickey, G. E., Gill, R. M., Heaton, R. P. Jr, Kerstens, R., Lindberg, G. M., Malhotra, R. K., Mandell, J. W., Manlucu, E. D., Mills, A. M., Mills, S. E., Moskaluk, C. A., Nelis, M., Patil, D. T., Przybycin, C. G., Reynolds, J. P., Rubin, B. P., Saboorian, M. H., Salicru, M., Samols, M. A., Sturgis, C. D., Turner, K. O., Wick, M. R., Yoon, J. Y., Zhao, P. & Taylor, C. R. (2018). Whole slide imaging versus microscopy for primary diagnosis in surgical pathology: a multicenter blinded randomized noninferiority study of 1992 cases (Pivotal Study). *Am J Surg Pathol,* 42(1), 39-52.

Muranushi, C., Olsen, C. M., Pandeya, N. & Green, A. C. (2015). Aspirin and nonsteroidal anti-inflammatory drugs can prevent cutaneous squamous cell carcinoma: a systematic review and meta-analysis. *J Invest Dermatol*, 135(4), 975-983.

Murphree D. H., Puri, P., Shamim, H., Bezalel, S. A., Drage, L. A., Wang, M., Pittelkow, M. R., Carter, R. E., Davis, M. D. P., Bridges, A. G., Mangold, A. R., Yiannias, J. A., Tollefson, M. M., Lehman, J. S., Meves, A., Otley, C. C., Sokumbi, O., Hall, M. R. & Comfere, N. (2020). Deep

learning for dermatologists: Part I fundamental concepts. *J Am Acad Dermatol*, S0190-9622(20)30921-X.

Muzic, J. G., Schmitt, A. R., Wright, A. C., Alniemi, D. T., Zubair, A. S., Olazagasti Lourido, J. M., Sosa Seda, I. M., Weaver A. L. & Baum, C. L. (2017). Incidence and trends of basal cell carcinoma and cutaneous squamous cell carcinoma: a population-based study in Olmsted County, Minnesota, 2000 to 2010 *Mayo Clin Proc*, 92(6), 890-898.

Nair, N., Calle, A. S., Zahra, M. H., Prieto-Vila, M., Oo, A. K. K., Hurley, L., Vaidyanath, A., Seno, A., Masuda, J., Iwasaki, Y., Tanaka, H., Kasai, T. & Seno, M. (2017). A cancer stem cell model as the point of origin of cancer-associated fibroblasts in tumor microenvironment. *Sci Rep*, 7(1), 6838.

Nebuloni, M., Albarello, L., Andolfo, A., Magagnotti, C., Genovese, L., Locatelli, I., Tonon, G., Longhi, E., Zerbi, P., Allevi, R., Podestà, A., Puricelli, L., Milani, P., Soldarini, A., Salonia, A. & Alfano, M. (2016). Insight on Colorectal Carcinoma Infiltration by Studying Perilesional Extracellular Matrix. *Sci Rep*, 6, 22522.

Nehal, K. S. & Bichakjian, C. K. (2018). Update on Keratinocyte Carcinomas. *N Engl J Med*, 379(4), 363-374.

Nelson, T. G. & Ashton, R. E. (2017). Low incidence of metastasis and recurrence from cutaneous squamous cell carcinoma found in a UK population: do we need to adjust our thinking on this rare but potentially fatal event?. *J Surg Oncol*, 116(6), 783-788.

Neri, S., Ishii, G., Hashimoto, H., Kuwata, T., Nagai, K., Date, H. & Ochiai A. (2015). Podoplanin-Expressing Cancer-Associated Fibroblasts Lead and Enhance the Local Invasion of Cancer Cells in Lung Adenocarcinoma. *Int J Cancer*, 137(4), 784-796.

Neri, S., Miyashita, T., Hashimoto, H., Suda, Y., Ishibashi, M., Kii, H., Watanabe, H., Kuwata, T., Tsuboi, M., Goto, K., Menju, T., Sonobe, M., Date, H., Ochiai, A. & Ishii, G. (2017). Fibroblast-Led Cancer Cell Invasion Is Activated by Epithelial–Mesenchymal Transition Through Platelet-Derived Growth Factor BB Secretion of Lung Adenocarcinoma. *Cancer Lett*, 395, 20-30.

Ng, Y. Z., Pourreyron, C., Salas-Alanis, J. C., Dayal, J. H., Cepeda-Valdes, R., Yan, W., Wright, S., Chen, M., Fine, J. D., Hogg, F. J., McGrath, J. A., Murrell, D. F., Leigh, I. M., Lane, E. B. & South, A. P. (2012). Fibroblast-derived dermal matrix drives development of aggressive cutaneous squamous cell carcinoma in patients with recessive dystrophic epidermolysis bullosa. *Cancer Res*, 72(14), 3522-3534.

Nguyen, K. D., Han, J., Li, T. & Qureshi, A. A. (2014). Invasive cutaneous squamous cell carcinoma incidence in US health care workers. *Arch Dermatol Res*, 306(6), 555-560.

Niazi, M. K. K., Parwani, A. V. & Gurcan, M. N. (2019). Digital pathology and artificial intelligence. *Lancet Oncol*, 20(5), e253-e261.

Niazi, M. K. K., Tavolara, T. E., Arole, V., Hartman, D. J., Pantanowitz, L. & Gurcan, M. N. (2018). Identifying tumor in pancreatic neuroendocrine neoplasms from Ki67 images using transfer learning. *PLoS One*, 13(4), e0195621.

Ninomiya, H., Hiramatsu, M., Inamura, K., Nomura, K., Okui, M., Miyoshi, T., Okumura, S., Satoh, Y., Nakagawa, K., Nishio, M., Horai, T., Miyata, S., Tsuchiya, E., Fukayama, M. & Ishikawa, Y. (2009). Correlation between morphology and EGFR mutations in lung adenocarcinomas significance of the micropapillary pattern and the hobnail cell type. *Lung Cancer*, 63(2), 235-240.

Nissen, N. I., Karsdal, M. & Willumsen, N. (2019). Collagens and Cancer associated fibroblasts in the reactive stroma and its relation to Cancer biology. *J Exp Clin Cancer Res*, 38(1), 115.

Nissinen, L., Farshchian, M., Riihilä, P. & Kähäri, V. M. (2016). New perspectives on the role of tumor microenvironment in progression of cutaneous squamous cell carcinoma. *Cell Tissue Res*, 365(3), 691–702.

Nordby, Y., Richardsen, E., Rakaee, M., Ness, N., Donnem, T., Patel, H. R., Busund, L. T., Bremnes, R. M. & Andersen, S. (2017). High expression of PDGFR-β in prostate cancer stroma is

independently associated with clinical and biochemical prostate cancer recurrence. *Sci Rep*, 7, 43378.

Nurmik, M., Ullmann, P., Rodriguez, F., Haan, S. & Letellier, E. (2020). In search of definitions: Cancer-associated fibroblasts and their markers. *Int J Cancer,* 146(4), 895-905.

O'Donovan, P., Perrett, C. M., Zhang, X., Montaner, B., Xu, Y. Z., Harwood, C. A., McGregor, J. M., Walker, S. L., Hanaoka, F. & Karran P. (2005). Azathioprine and UVA light generate mutagenic oxidative DNA damage. *Science*, 309(5742), 1871-1874.

Odorisio, T., Di Salvio, M., Orecchia, A., Di Zenzo, G., Piccinni, E., Cianfarani, F., Travaglione, A., Uva, P., Bellei, B., Conti, A., Zambruno, G. & Castiglia, D. (2014). Monozygotic twins discordant for recessive dystrophic epidermolysis bullosa phenotype highlight the role of TGF-β signalling in modifying disease severity. *Hum Mol Genet*, 23(15), 3907-3922.

Ogawa, T., Kiuru, M., Konia, T. H. & Fung, M. A. (2017). Acantholytic squamous cell carcinoma is usually associated with hair follicles, not acantholytic actinic keratosis, and is not "high risk": Diagnosis, management, and clinical outcomes in a series of 115 cases. *J Am Acad Dermatol*, 76(2), 327-333.

Olsen, T. G., Jackson, B. H., Feeser, T. A., Kent, M. N., Moad, J. C., Krishnamurthy, S., Lunsford, D. D. & Soans, R. E. (2018). Diagnostic performance of deep learning algorithms applied to three common diagnoses in dermatopathology. *J Pathol Inform,* 9, 32.

Orimo, A., Gupta, P. B., Sgroi, D. C., Arenzana-Seisdedos, F., Delaunay, T., Naeem, R., Carey, V. J., Richardson, A. L. & Weinberg, R. A. (2005). Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion. *Cell*, 121(3), 335-348.

Paget, S. (1989). The Distribution of Secondary Growths in Cancer of the Breast. *Cancer Metastasis Rev,* 8(2), 98-101.

Pakshir, P., Noskovicova, N., Lodyga, M., Son, D. O., Schuster, R., Goodwin, A., Karvonen, H. & Hinz, B. (2020). The myofibroblast at a glance. *J Cell Sci*, 133(13), jcs227900.

Pandeya, N., Olsen, C. M., Thompson, B. S., Dusingize, J. C., Neale, R. E., Green, A. C., Whiteman, D. C., & QSkin Study. (2019). Aspirin and nonsteroidal anti-inflammatory drug use and keratinocyte cancers: a large population-based cohort study of skin cancer in Australia. *Br J Dermatol*, 181(4), 749-760.

Parekh, V. & Seykora, J. T. (2017). Cutaneous Squamous Cell Carcinoma. *Clin Lab Med*, 37(3), 503-525.

Park, J., Kim, D. S., Shim, T. S., Lim, C. M., Koh, Y., Lee, S. D., Kim, W. S., Kim, W. D., Lee, J. S. & Song, K. S. (2001). Lung cancer in patients with idiopathic pulmonary fibrosis. *Eur Respir J*, 17(6), 1216-1219.

Parsonage, G., Filer, A. D., Haworth, O., Nash, G. B., Rainger, G. E., Salmon, M. & Buckley, C. D. (2005). A stromal address code defined by fibroblasts. *Trends Immunol*, 26(3), 150-156.

Patocs, A., Zhang, L., Xu, Y., Weber, F., Caldes, T., Mutter, G. L., Platzer, P. & Eng, C. (2007). Breast- cancer stromal cells with TP53 mutations and nodal metastases. *N Engl J Med*, 357(25), 2543-2551.

Paulsson, J., Ehnman, M. & Ostman, A. (2014). PDGF receptors in tumor biology: prognostic and predictive potential. *Future Oncol*, 10(9):1695-1708.

Paulsson, J., Sjöblom, T., Micke, P., Pontén, F., Landberg, G., Heldin, C. H., Bergh, J., Brennan, D. J., Jirström, K. & Ostman, A. (2009). Prognostic significance of stromal platelet-derived growth factor beta-receptor expression in human breast cancer. *Am J Pathol*, 175(1), 334-341.

Pazzini, R., Pereira, G. A. & Macarenco, R. S. (2018). Cutaneous Metaplastic Carcinoma: Report of a Case With Sebaceous Differentiation. *Am J Dermatopathol*, 40(7), e100-e103.

Pedersen, S. A., Gaist, D., Schmidt, S. A. J., Hölmich, L. R., Friis, S. & Pottegård, A. (2018). Hydrochlorothiazide use and risk of nonmelanoma skin cancer: A nationwide case-control study from Denmark. *J Am Acad Dermatol*, 78(4), 673-681.

Petter, G. & Haustein, U. F. (2000). Histologic subtyping and malignancy assessment of cutaneous squamous cell carcinoma. *Dermatol Surg*, 26(6), 521-530.

Pickering, C. R., Zhou, J. H., Lee, J. J., Drummond, J. A., Peng, S. A., Saade, R. E., Tsai, K. Y., Curry, J. L., Tetzlaff, M. T., Lai, S. Y., Yu, J., Muzny, D. M., Doddapaneni, H., Shinbrot, E., Covington, K. R., Zhang, J., Seth, S., Caulin, C., Clayman, G. L., El-Naggar, A. K., Gibbs, R. A., Weber, R. S., Myers, J. N., Wheeler, D. A. & Frederick, M. J. (2014). Mutational landscape of aggressive cutaneous squamous cell carcinoma. *Clin Cancer Res*, 20(24), 6582-6592.

Pietras, K., Pahler, J., Bergers, G. & Hanahan, D. (2008). Functions of paracrine PDGF signaling in the proangiogenic tumor stroma revealed by pharmacological targeting. *PLoS Med*, 5(1), e19.

Pipili-Synetos, E., Papageorgiou, A., Sakkoula, E., Sotiropoulou, G., Fotsis, T., Karakiulakis, G. & Maragoudakis, M. E. (1995). Inhibition of angiogenesis, tumour growth and metastasis by the NO-releasing vasodilators, isosorbide mononitrate and dinitrate. *Br J Pharmacol*, 116(2), 1829-1834.

Potenta, S., Zeisberg, E. & Kalluri, R. (2008). The role of endothelial-to-mesenchymal transition in cancer progression. *Br J Cancer*, 99(9), 1375-1379.

Puebla-Tornero, L., Corchete-Sánchez, L. A., Conde-Ferreirós, A., García-Sancha, N., Corchado-Cobos, R., Román-Curto, C. & Cañueto, J. (2021). Performance of Salamanca refinement of the T3-AJCC8 versus the Brigham and Women's Hospital and Tubingen alternative staging systems for high-risk cutaneous squamous cell carcinoma. *J Am Acad Dermatol*, 84(4), 938-945.

Puram, S. V., Tirosh, I., Parikh, A. S., Patel, A. P., Yizhak, K., Gillespie, S., Rodman, C., Luo, C. L., Mroz, E. A., Emerick, K. S., Deschler, D. G., Varvares, M. A., Mylvaganam, R., Rozenblatt-Rosen, O., Rocco, J. W., Faquin, W. C., Lin, D. T., Regev, A. & Bernstein, B. E. (2017). Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. *Cell*, 171, 1611e24.e24.

Puri, P., Comfere, N., Drage, L. A., Shamim, H., Bezalel, S. A., Pittelkow, M. R., Davis, M. D. P., Wang, M., Mangold, A. R., Tollefson, M. M., Lehman, J. S., Meves, A., Yiannias, J. A., Otley, C. C., Carter, R. E., Sokumbi, O., Hall, M. R., Bridges, A. G. & Murphree, D. H. (2020). Deep learning for dermatologists: Part II. Current applications. *J Am Acad Dermatol,* S0190-9622(20)30918-X.

Qiao, Y., Chen, J., Lim, Y. B., Finch-Edmondson, M. L., Seshachalam, V. P., Qin, L., Jiang, T., Low, B. C., Singh, H., Lim, C. T. & Sudol, M. (2017). YAP Regulates Actin Dynamics through ARHGAP29 and Promotes Metastasis. *Cell Rep*, 19(8), 1495-1502.

Quaedvlieg, P. J. F., Creytens, D. H. K. V., Epping, G. G., Peutz-Kootstra, C. J., Nieman, F. H. M., Thissen, M. R. T. M. & Krekels, G. A. (2006b). Histopathological characteristics of metastasizing squamous cell carcinoma of the skin and lips. *Histopathology*, 49(3), 256-264.

Quaedvlieg, P. J. F., Tirsi, E., Thissen, M. R. T. M. & Krekels, G. A. (2006). Actinic keratosis: howto differentiate the good from the bad ones? *Eur J Dermatol*, 16(4), 335-339.

Quail, D. F. & Joyce, J. A. (2013). Microenvironmental regulation of tumor progression and metastasis. *Nat Med*, 19(11), 1423-1437.

Quante, M., Tu, S. P., Tomita, H., Gonda, T., Wang, S. S. W., Takashi, S., Baik, G. H., Shibata, W., DiPrete, B., Betz, K. S., Friedman, R., Varro, A., Tycko, B. & Wang, T. C. (2011). Bone Marrow-Derived Myofibroblasts Contribute to the Mesenchymal Stem Cell Niche and Promote Tumor Growth. *Cancer Cell,* 19(2), 257-272.

Que, S. K. T, Zwald, F. O. & Schmults, C. D. (2018). Cutaneous squamous cell carcinoma: Incidence, risk factors, diagnosis, and staging. *J Am Acad Dermatol,* 78(2), 237-247.

Racanelli, E., Jfri, A., Gefri, A., O'Brien, E., Litvinov, I. V., Zubarev, A., Savin, E. & Netchiporouk, E. (2021). Cutaneous Squamous Cell Carcinoma in Patients with Hidradenitis Suppurativa. *Cancers (Basel)*, 13(5), 1153.

Ramos-Vega, V., Rojas, B. V. & Torres, W. D. (2020). Immunohistochemical analysis of cancer-associated fibroblasts and podoplanin in head and neck cancer. *Med Oral Patol Oral Cir Bucal,* 25(2), e268-e276.

Rawat, W. & Wang, Z. (2017). Deep convolutional neural networks for image classification: a comprehensive review. *Neural Comput*, 29(9), 2352-2449.

Riihilä, P., Knuutila, J. & Kähäri, V. M. (2021a). Aikuisten yleisimmät melanosyyttiperäiset ihokasvaimet. *Lääketieteellinen aikakauskirja Duodecim*, 137(12), 1279-88.

Riihilä, P., Nissinen, L. & Kähäri, V. M. (2021b). Matrix metalloproteinases in keratinocyte carcinomas. *Exp Dermatol*, 30(1), 50-61.

Riihilä, P. M., Nissinen, L. M., Ala-aho, R., Kallajoki, M., Grénman, R., Meri, S., Peltonen, S., Peltonen, J. & Kähäri, V. M. (2014). Complement factor H: a biomarker for progression of cutaneous squamous cell carcinoma. *J Invest Dermatol,* 134(2), 498-506.

Riihilä, P., Nissinen, L., Farshchian, M., Kivisaari, A., Ala-aho, R., Kallajoki, M., Grénman, R., Meri, S., Peltonen, S., Peltonen, J. & Kähäri, V. M. (2015). Complement factor I promotes progression of cutaneous squamous cell carcinoma. *J Invest Dermatol,* 135(2), 579-588.

Riihilä, P., Nissinen, L., Farshchian, M., Kallajoki, M., Kivisaari, A., Meri, S., Grénman, R., Peltonen, S., Peltonen, J., Pihlajaniemi, T., Heljasvaara, R. & Kähäri, V. M. (2017). Complement Component C3 and Complement Factor B Promote Growth of Cutaneous Squamous Cell Carcinoma. *Am J Pathol,* 187(5), 1186-1197.

Riihilä, P., Nissinen, L., Knuutila, J., Rahmati Nezhad, P., Viiklepp, K. & Kähäri, V. M. (2019). Complement system in cutaneous squamous cell carcinoma. *Int J Mol Sci*, 20(14), 3550.

Riihilä P, Viiklepp K, Nissinen L, Farshchian M, Kallajoki M, Kivisaari A, Meri S, Peltonen J, Peltonen S & Kähäri VM. (2020). Tumour-cell-derived complement components C1r and C1s promote growth of cutaneous squamous cell carcinoma. *Br J Dermatol,* 182(3), 658-670.

Rimm, D. L., Leung, S. C. Y., McShane, L. M., Bai, Y., Bane, A. L., Bartlett, J. M. S., Bayani, J., Chang, M. C., Dean, M., Denkert, C., Enwere, E. K., Galderisi, C., Gholap, A., Hugh, J. C., Jadhav, A., Kornaga, E. N., Laurinavicius, A., Levenson, R., Lima, J., Miller, K., Pantanowitz, L., Piper, T., Ruan, J., Srinivasan, M., Virk, S., Wu, Y., Yang, H., Hayes, D. F., Nielsen, T. O. & Dowsett M. (2019). An international multicenter study to evaluate reproducibility of automated scoring for assessment of Ki67 in breast cancer. *Mod Pathol*, 32(1), 59-69.

Rinkevich, Y., Walmsley, G. G., Hu, M. S., Maan, Z. N., Newman, A. M., Drukker, M., Januszyk, M., Krampitz, G. W., Gurtner, G. C., Lorenz, H. P., Weissman, I. L. & Longaker, M. T. (2015). Skin fibrosis. Identification and isolation of a dermal lineage with intrinsic fibrogenic potential. *Science*, 348(6232), aaa2151.

Rivenson, Y., Wang, H., Wei, Z., de Haan, K., Zhang, Y., Wu, Y., Günaydın, H., Zuckerman, J. E., Chong, T., Sisk, A. E., Westbrook, L. M., Wallace, W. D. & Ozcan, A. (2019). Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning. *Nat Biomed Eng*, 3(6), 466-477.

Roberts, E. W., Deonarine, A., Jones, J. O., Denton, A. E., Feig, C., Lyons, S. K., Espeli, M., Kraman, M., McKenna, B., Wells, R. J., Zhao, Q., Caballero, O. L., Larder, R., Coll, A. P., O'Rahilly, S., Brindle, K. M., Teichmann, S. A., Tuveson, D. A. & Fearon, D. T. (2013). Depletion of Stromal Cells Expressing Fibroblast Activation Protein-α From Skeletal Muscle and Bone Marrow Results in Cachexia and Anemia. *J Exp Med*, 210(6), 1137-1151.

Rodemann, H. P. & Muller, G. A. (1991). Characterization of human renal fibroblasts in health and disease: II. In vitro growth, differentiation, and collagen synthesis of fibroblasts from kidneys with interstitial fibrosis. Am J Kidney Dis, 17(6), 684-686.

Rogers, H. W., Weinstock, M. A., Feldman, S. R. & Coldiron, B. M. (2015). Incidence Estimate of Nonmelanoma Skin Cancer (Keratinocyte Carcinomas) in the U.S. Population, 2012. *JAMA Dermatol*, 151(10), 1081-1086.

Rønnov-Jessen, L. & Petersen, O. W. (1993). Induction of α-smooth muscle actin by transforming growth factor-β1 in quiescent human breast gland fibroblasts. Implications for myofibroblast generation in breast neoplasia. *Lab Invest*, 68(6), 696-707.

Rønnov-Jessen, L., Petersen, O. W. & Bissell, M. J. (1996). Cellular changes involved in conversion of normal to malignant breast: importance of the stromal reaction. Physiol Rev, 76(1), 69-125.

Rønnov-Jessen, L., Petersen, O. W., Koteliansky, V. E. & Bissell, M. J. (1995). The Origin of the Myofibroblasts in Breast Cancer. Recapitulation of Tumor Environment in Culture Unravels Diversity and Implicates Converted Fibroblasts and Recruited Smooth Muscle Cells. *J Clin Invest,* 95(2), 859-873.

Roscher, I., Falk, R. S., Vos, L., Clausen, O. P. F., Helsing, P., Gjersvik, P. & Robsahm, T. E. (2018). Validating 4 staging systems for cutaneous squamous cell carcinoma using population-based data: a nested case-control study. *JAMA Dermatol*, 154(4), 428-434.

Rosen, L. S., Gordon, M. S., Robert, F. & Matei, D. E. (2014). Endoglin for Targeted Cancer Treatment. *Curr Oncol Rep,* 16(2), 365.

Rosenberg, A. R., Tabacchi, M., Ngo, K. H., Wallendorf, M., Rosman, I. S., Cornelius, L. A. & Demehri, S. (2019). Skin cancer precursor immunotherapy for squamous cell carcinoma prevention. *JCI Insight*, 4(6), e125476.

Rothwell, P. M., Wilson, M., Elwin, C. E., Norrving, B., Algra, A., Warlow, C. P. & Meade, T. W. (2010). Long-term effect of aspirin on colorectal cancer incidence and mortality: 20-year follow-up of five randomised trials. *Lancet*, 376(9754), 1741-1750.

Rothwell, P. M., Wilson, M., Price, J. F., Belch, J. F., Meade, T. W. & Mehta, Z. (2012). Effect of daily aspirin on risk of cancer metastasis: a study of incident cancers during randomised controlled trials. *Lancet*, 379(9826), 1591-1601.

Rowe, D. E., Carroll, R. J. & Day, C. L. Jr. (1992) Prognostic factors for local recurrence, metastasis, and survival rates in squamous cell carcinoma of the skin, ear, and lip. Implications for treatment modality selection. *J Am Acad Dermatol,* 26(6), 976-990.

Ruiz, E. S., Karia, P. S., Besaw, R. & Schmults, C. D. (2019). Performance of the American Joint Committee on Cancer Staging Manual, 8th Edition vs the Brigham and Women's Hospital Tumor Classification System for Cutaneous Squamous Cell Carcinoma. *JAMA Dermatol*, 155(7), 819-825.

Sahovaler, A., Krishnan, R. J., Yeh, D. H., Zhou, Q., Palma, D., Fung, K., Yoo, J., Nichols, A. & MacNeil, S. D. (2019). Outcomes of Cutaneous Squamous Cell Carcinoma in the Head and Neck Region With Regional Lymph Node Metastasis: A Systematic Review and Meta-analysis. *JAMA Otolaryngol Head Neck Surg*, 145(4), 352-360.

Samain, R. & Sanz-Moreno, V. (2020). Cancer-associated fibroblasts: Activin A adds another string to their bow. *EMBO Mol Med,* 12(4), e12102.

Santos, A. M., Jung, J., Aziz, N., Kissil, J. L. & Puré, E. (2009). Targeting fibroblast activation protein inhibits tumor stromagenesis and growth in mice. *J Clin Invest,* 119(12), 3613-3625.

Sasaki K, Sugai T, Ishida K, Osakabe M, Amano H, Kimura H, Sakuraba M, Kashiwa K, Kobayashi S. (2018). Analysis of cancer-associated fibroblasts and the epithelial-mesenchymal transition in cutaneous basal cell carcinoma, squamous cell carcinoma, and malignant melanoma. *Hum Pathol*, 79, 1-8.

Schauer, I. G., Sood, A. K., Mok, S. & Liu, J. (2011). Cancer-Associated Fibroblasts and Their Putative Role in Potentiating the Initiation and Development of Epithelial Ovarian Cancer. *Neoplasia,* 13(5), 393-405.

Schaumberg, A. J., Rubin, M. A. & Fuchs, T. J. (2018). H&E-stained whole slide image deep learning predicts SPOP mutation state in prostate cancer. *BioRxiv*, 064279.

Scheel, C. & Weinberg, R. A. (2012). Cancer stem cells and epithelial-mesenchymal transition: concepts and molecular links. *Semin Cancer Biol*, 22(5-6), 396-403.

Schmults, C. D., Karia, P. S., Carter, J. B., Han, J. & Qureshi, A. A. (2013). Factors predictive of recurrence and death from cutaneous squamous cell carcinoma: a 10-year, single-institution cohort study. *JAMA Dermatol,* 149(5), 541-547.

Schoppmann, S. F., Berghoff, A., Dinhof, C., Jakesz, R., Gnant, M., Dubsky, P., Jesch, B., Heinzl, H. & Birner, P. (2012). Podoplanin-Expressing Cancer-Associated Fibroblasts Are Associated With Poor Prognosis in Invasive Breast Cancer. *Breast Cancer Res Treat*, 134(1), 237-244.

Schuh, A. C., Keating, S. J., Monteclaro, F. S., Vogt, P. K. & Breitman, M. L. (1990). Obligatory wounding requirement for tumorigenesis in v-jun transgenic mice. *Nature*, 346(6286), 756-760.

Schwartz, R. A. (1995). Verrucous carcinoma of the skin and mucosa. *J Am Acad Dermatol*, 32(1), 1-21.

Senaras, C., Niazi, M. K. K., Lozanski, G. & Gurcan, M. N. (2018). DeepFocus: Detection of out-of-focus regions in whole slide digital images using deep learning. *PLoS One*, 13(10), e0205387.

Sha, L., Osinski, B. L., Ho, I. Y., Tan, T. L., Willis, C., Weiss, H., Beaubier, N., Mahon, B. M., Taxter, T. J. & Yip, S. S. F. (2019). Multi-Field-of-View Deep Learning Model Predicts Nonsmall Cell Lung Cancer Programmed Death-Ligand 1 Status from Whole-Slide Hematoxylin and Eosin Images. *J Pathol Inform*, 10, 24.

Shaban, M., Khurram, S. A., Fraz, M. M., Alsubaie, N., Masood, I., Mushtaq, S., Hassan, M., Loya, A. & Rajpoot, N. M. (2019). A novel digital score for abundance of tumour infiltrating lymphocytes predicts disease free survival in oral squamous cell carcinoma. *Sci Rep*, 9(1), 13341.

Shany, S., Sigal-Batikoff, I. & Lamprecht, S. (2019). Vitamin D and Myofibroblasts in Fibrosis and Cancer: At Cross-purposes with TGF-β/SMAD Signaling. *Anticancer Res, 36(12)*, 6225–6234.

Shapanis, A., Lai, C., Smith, S., Coltart, G., Sommerlad, M., Schofield, J., Parkinson, E., Skipp, P. & Healy, E. (2021). Identification of proteins associated with development of metastasis from cutaneous squamous cell carcinomas (cSCCs) via proteomic analysis of primary cSCCs. *Br J Dermatol*, 184(4), 709-721.

Shin, H. C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D. & Summers, R. M. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans Med Imaging*, 35(5), 1285-1298.

Shinagawa, K., Kitadai, Y., Tanaka, M., Sumida, T., Onoyama, M., Ohnishi, M., Ohara, E., Higashi, Y., Tanaka, S., Yasui, W. & Chayama, K. (2013). Stroma-directed imatinib therapy impairs the tumor-promoting effect of bone marrow-derived mesenchymal stem cells in an orthotopic transplantation model of colon cancer. *Int J Cancer*, 132(4), 813-823.

Shrestha, P., Kneepkens, R., Vrijnsen, J., Vossen, D., Abels, E. & Hulsken, B. (2016). A quantitative approach to evaluate image quality of whole slide imaging scanners. *J Pathol Inform*, 7, 56.

Shu, C., Zha, H., Long, H., Wang, X., Yang, F., Gao, J., Hu, C., Zhou, L., Guo, B. & Zhu, B. (2020). C3a-C3aR Signaling Promotes Breast Cancer Lung Metastasis via Modulating Carcinoma Associated Fibroblasts. *J Exp Clin Cancer Res*, 39(1), 11.

Siljamäki, E., Rappu, P., Riihilä, P., Nissinen, L., Kähäri, V. M. & Heino, J. (2020). H-Ras activation and fibroblast-induced TGF-β signaling promote laminin-332 accumulation and invasion in cutaneous squamous cell carcinoma. *Matrix Biol*, 87,26-47.

Silvis, N. G., Swanson, P. E., Manivel, J. C., Kaye, V. N. & Wick, M. R. (1988). Spindle-cell and pleomorphic neoplasms of the skin. A clinicopathologic and immunohistochemical study of 30 cases, with emphasis on "atypical fibroxanthomas". Am J Dermatopathol, 10(1), 9-19.

Simian, M., Hirai, Y., Navre, M., Werb, Z., Lochter, A. & Bissell, M. J. (2001). The interplay of matrix metalloproteinases, morphogens and growth factors is necessary for branching of mammary epithelial cells. Development, 128(16), 3117-3131.

Simpkins, S. A., Hanby, A. M., Holliday, D. L. & Speirs, V. (2012). Clinical and functional significance of loss of caveolin-1 expression in breast cancer-associated fibroblasts. *J Pathol*. 227(4), 490-498.

Skulsky, S. L., O'Sullivan, B., McArdle, O., Leader, M., Roche, M., Conlon, P. J. & O'Neill, J. P. (2017). Review of high-risk features of cutaneous squamous cell carcinoma and discrepancies between the American Joint Committee on Cancer and NCCN Clinical Practice Guidelines In Oncology. *Head Neck*, 39(3), 578-594.

Smith, J. A., Virk, S., Palme, C. E., Low, T. H., Ch'ng, S., Gupta, R., Gao, K. & Clark, J. (2018). Age is not a predictor of prognosis in metastatic cutaneous squamous cell carcinoma of the head and neck. *ANZ J Surg*, 88(4), E273-E277.

Sober, A. J. & Burstein, J. M. (1995). Precursors to skin cancer. *Cancer*, 75(2 Suppl), 645-650.

South, A. P., Purdie, K. J., Watt, S. A., Haldenby, S., den Breems, N., Dimon, M., Arron, S. T., Kluk, M. J., Aster, J. C., McHugh, A., Xue, D. J., Dayal, J. H., Robinson, K. S., Rizvi, S. H., Proby, C. M., Harwood, C. A. & Leigh, I. M. (2014). NOTCH1 mutations occur early during cutaneous squamous cell carcinogenesis. *J Invest Dermatol,* 134(10), 2630-2638.

Staples, M. P., Elwood, M., Burton, R. C., Williams, J. L., Marks, R. & Giles, G. G. (2006). Non-melanoma skin cancer in Australia: the 2002 national survey and trends since 1985. *Med J Aust*, 184(1), 6-10.

Stefansson, I. M., Salvesen, H. B. & Akslen, L. A. (2006). Loss of p63 and cytokeratin 5/6 expression is associated with more aggressive tumors in endometrial carcinoma patients. *Int J Cancer*, 118(5), 1227-1233.

Stern, R. S. (2010). Prevalence of a history of skin cancer in 2007: results of an incidence-based model. *Arch Dermatol*, 146(3), 279-282.

Stirling, D. R., Swain-Bowden, M. J., Lucas, A. M., Carpenter, A. E., Cimini, B. A. & Goodman, A. (2021). CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics*, 22(1), 433.

Stratigos, A. J., Garbe, C., Dessinioti, C., Lebbe, C., Bataille, V., Bastholt, L., Dreno, B., Fargnoli, M. C., Forsea, A. M., Frenard, C., Harwood, C. A., Hauschild, A., Hoeller, C., Kandolf-Sekulovic, L., Kaufmann, R., Kelleners-Smeets, N. W., Malvehy, J., Del Marmol, V., Middleton, M. R., Moreno-Ramirez, D., Pellecani, G., Peris, K., Saiag, P., van den Beuken-van Everdingen, M. H. J., Vieira, R., Zalaudek, I., Eggermont, A. M. M. & Grob, J. J.; European Dermatology Forum (EDF), the European Association of Dermato-Oncology (EADO) and the European Organization for Research and Treatment of Cancer (EORTC). (2020a). European interdisciplinary guideline on invasive squamous cell carcinoma of the skin: Part 1. epidemiology, diagnostics and prevention. *Eur J Cancer*, 128, 60-82.

Stratigos, A. J., Garbe, C., Dessinioti, C., Lebbe, C., Bataille, V., Bastholt, L., Dreno, B., Concetta Fargnoli, M., Forsea, A. M., Frenard, C., Harwood, C. A., Hauschild, A., Hoeller, C., Kandolf-Sekulovic, L., Kaufmann, R., Kelleners-Smeets, N. W. J., Malvehy, J., Del Marmol, V., Middleton, M. R., Moreno-Ramirez, D., Pellecani, G., Peris, K., Saiag, P., van den Beuken-van Everdingen, M. H. J., Vieira, R., Zalaudek, I., Eggermont, A. M. M. & Grob, J. J.; European Dermatology Forum (EDF), the European Association of Dermato-Oncology (EADO) and the European Organization for Research and Treatment of Cancer (EORTC). (2020b). European interdisciplinary guideline on invasive squamous cell carcinoma of the skin: Part 2. Treatment. *Eur J Cancer*, 128, 83-102.

Strutz, F., Okada, H., Lo, C. W., Danoff, T., Carone, R. L., Tomaszewski, J. E. & Neilson, E. G. (1995). Identification and characterization of a fibroblast marker: FSP1. *J Cell Biol*, 130(2), 393-405.

Stålhammar, G., Robertson, S., Wedlund, L., Lippert, M., Rantalainen, M., Bergh, J. & Hartman J. (2018). Digital image analysis of Ki67 in hot spots is superior to both manual Ki67 and mitotic counts in breast cancer. *Histopathology*, 72(6), 974-989.

Su, S., Chen, J., Yao, H., Liu, J., Yu, S., Lao, L., Wang, M., Luo, M., Xing, Y., Chen, F., Huang, D., Zhao, J., Yang, L., Liao, D., Su, F., Li, M., Liu, Q. & Song, E. (2018). CD10$^+$GPR77$^+$ Cancer-Associated Fibroblasts Promote Cancer Formation and Chemoresistance by Sustaining Cancer Stemness. *Cell*, 172(4), 841-856.

Suarez-Carmona M, Williams A, Schreiber J, Hohmann N, Pruefer U, Krauss J, Jäger D, Frömming A, Beyer D, Eulberg D, Jungelius JU, Baumann M, Mangasarian A, Halama N. (2021). Combined inhibition of CXCL12 and PD-1 in MSS colorectal and pancreatic cancer: modulation of the microenvironment and clinical effects. *J Immunother Cancer*, 9(10), e002505.

Sugimoto, H., Mundel, T. M., Kieran, M. W. & Kalluri, R. (2006). Identification of Fibroblast Heterogeneity in the Tumor Microenvironment. *Cancer Biol Ther,* 5(12), 1640-1646.

Sultan, A. S., Elgharib, M. A., Tavares, T., Jessri, M. & Basile, J. R. (2020). The use of artificial intelligence, machine learning and deep learning in oncologic histopathology. *J Oral Pathol Med,* 49(9), 849-856.

Sun, S., Wang, Y., Wu, Y., Gao, Y., Li, Q., Abdulrahman, A. A., Liu, X. F., Ji, G. Q., Gao, J., Li, L., Wan, F. P., Li, Y. Q. & Gao, D. S. (2018). Identification of COL1A1 as an Invasion-Related Gene in Malignant Astrocytoma. *Int J Oncol*, 53(6), 2542-2554.

Sung, P. L., Jan, Y. H., Lin, S. C., Huang, C. C., Lin, H., Wen, K. C., Chao, K. C., Lai, C. R., Wang, P. H., Chuang, C. M., Wu, H. H., Twu, N. F., Yen, M. S., Hsiao, M. & Huang, C. Y. (2016). Periostin in tumor microenvironment is associated with poor prognosis and platinum resistance in epithelial ovarian carcinoma. *Oncotarget*, 7(4), 4036-4047.

Tanvetyanon, T., Padhya, T., McCaffrey, J., Kish, J. A., Deconti, R. C., Trotti, A. & Rao, N. G. (2015). Postoperative concurrent chemotherapy and radiotherapy for high-risk cutaneous squamous cell carcinoma of the head and neck. *Head Neck*, 37(6), 840-845.

Tarin, D. & Croft, C. B. (1969). Ultrastructural features of wound healing in mouse skin. *J Anat*, 105(Pt 1), 189-190.

Tartaglia, G., Cao, Q., Padron, Z. M. & South, A. P. (2021). Impaired Wound Healing, Fibrosis, and Cancer: The Paradigm of Recessive Dystrophic Epidermolysis Bullosa. *Int J Mol Sci*, 22(10), 5104.

Tate, J. G., Bamford, S., Jubb, H. C., Sondka, Z., Beare, D. M., Bindal, N., Boutselakis, H., Cole, C. G., Creatore, C., Dawson, E., Fish, P., Harsha, B., Hathaway, C., Jupe, S. C., Kok, C. Y., Noble, K., Ponting, L., Ramshaw, C. C., Rye, C. E., Speedy, H. E., Stefancsik, R., Thompson, S. L., Wang, S., Ward, S., Campbell, P. J. & Forbes, S. A. (2019). COSMIC: The catalogue of somatic mutations in cancer. *Nucleic Acids Res*, 47(D1), D941-D947.

Tauriello, D. V. F., Palomo-Ponce, S., Stork, D., Berenguer-Llergo, A., Badia-Ramentol, J., Iglesias, M., Sevillano, M., Ibiza, S., Cañellas, A., Hernando-Momblona, X., Byrom, D., Matarin, J. A., Calon, A., Rivas, E. I., Nebreda, A. R., Riera, A., Attolini, C. S. & Batlle, E. (2018). TGFβ drives immune evasion in genetically reconstituted colon cancer metastasis. *Nature*, 554(7693), 538-543.

Tejada, M. L., Yu, L., Dong, J., Jung, K., Meng, G., Peale, F. V., Frantz, G. D., Hall, L., Liang, X., Gerber, H. P. & Ferrara, N. (2006). Tumor-Driven Paracrine Platelet-Derived Growth Factor Receptor Alpha Signaling Is a Key Determinant of Stromal Cell Recruitment in a Model of Human Lung Carcinoma. Clin Cancer Res, 12(9), 2676-2688.

Teng, F., Tian, W. Y., Wang, Y. M., Zhang, Y. F., Guo, F., Zhao, J., Gao, C. & Xue, F. X. (2016). Cancer-associated fibroblasts promote the progression of endometrial cancer via the SDF-1/CXCR4 axis. *J Hematol Oncol,* 9, 8.

Thompson, A. K., Kelley, B. F., Prokop, L. J., Murad, M. H., Baum, C. L. (2016). Risk factors for cutaneous squamous cell carcinoma recurrence, metastasis, and disease-specific death: a systematic review and meta-analysis. *JAMA Dermatol*, 152(4), 419-428.

Thompson, R. H., Dong, H., Lohse, C. M., Leibovich, B. C., Blute, M. L., Cheville, J. C. & Kwon, E. D. (2007). PD-1 is expressed by tumor-infiltrating immune cells and is associated with poor outcome for patients with renal cell carcinoma. *Clin Cancer Res,* 13(6), 1757-1761.

Tian, H., Callahan, C. A., DuPree, K. J., Darbonne, W. C., Ahn, C. P., Scales, S. J. & de Sauvage, F. J. (2009). Hedgehog Signaling Is Restricted to the Stromal Compartment During Pancreatic Carcinogenesis. *Proc Natl Acad Sci USA,* 106(11), 4254-4259.

Tokez, S., Wakkee, M., Kan, W., Venables, Z. C., Mooyaart, A. L., Louwman, M., Nijsten, T. & Hollestein, L. M. (2022). Cumulative incidence and disease-specific survival of metastatic cutaneous squamous cell carcinoma: A nationwide cancer registry study. *J Am Acad Dermatol*, 86(2), 331-338.

Toll, A., Salgado, R., Yébenes, M., Martín-Ezquerra, G., Gilaberte, M., Baró, T., Solé, F., Alameda, F., Espinet, B. & Pujol, R. M. (2010). Epidermal growth factor receptor gene numerical

aberrations are frequent events in actinic keratosis and invasive squamous cell carcinoma. *Exp Dermatol,* 19(2), 151-153.

Tomasek, J. J., Gabbiani, G., Hinz, B., Chaponnier, C. & Brown, R. A. (2002). Myofibroblasts and mechano-regulation of connective tissue remodelling. *Nat Rev Mol Cell Biol*, 3(5), 349-363.

Torphy, R. J., Wang, Z., True-Yasaki, A., Volmar, K. E., Rashid, N., Yeh, B., Anderson, J. M., Johansen, J. S., Hollingsworth, M. A., Yeh, J. J. & Collisson, E. A. (2018). Stromal content is correlated with tissue site, contrast retention, and survival in pancreatic adenocarcinoma. *JCO Precis Oncol*, 2018, PO.17.00121.

Tseng, H. H., Wei, L., Cui, S., Luo, Y., Ten Haken, R. K. & El Naqa, I. (2020). Machine learning and imaging informatics in oncology. *Oncology*, 98(6), 344-362.

Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhori, G., Benfeitas, R., Arif, M., Liu, Z., Edfors, F., Sanli, K., von Feilitzen, K., Oksvold, P., Lundberg, E., Hober, S., Nilsson, P., Mattsson, J., Schwenk, J. M., Brunnström, H., Glimelius, B., Sjöblom, T., Edqvist, P. H., Djureinovic, D., Micke, P., Lindskog, C., Mardinoglu, A. & Ponten, F. (2017). A pathology atlas of the human cancer transcriptome. *Science*, 357(6352), eaan2507.

Vaquero, J., Aoudjehane, L. & Fouassier, L. (2020). Cancer-associated fibroblasts in cholangiocarcinoma. *Curr Opin Gastroenterol*, 36(2), 63-69.

Van Hove, L. & Hoste, E. (2022). Activation of Fibroblasts in Skin Cancer. *J Invest Dermatol*, 142(4), 1026-1031.

Van Hove, L., Lecomte, K., Roels, J., Vandamme, N., Vikkula, H. K., Hoorens, I., Ongenae, K., Hochepied, T., Donati, G., Saeys, Y., Quist, S. R., Watt, F. M., van Loo, G. & Hoste E. (2021). Fibrotic enzymes modulate wound-induced skin tumorigenesis. *EMBO Rep*, 22(5), e51573.

Velez, N. F., Karia, P. S., Vartanov, A. R., Davids, M. S., Brown, J. R. & Schmults, C. D. (2014). Association of advanced leukemic stage and skin cancer tumor stage with poor skin cancer outcomes in patients with chronic lymphocytic leukemia. *JAMA Dermatol*, 150(3), 280-287.

Venables, Z. C., Autier, P., Nijsten, T., Wong, K. F., Langan, S. M., Rous, B., Broggio, J., Harwood, C., Henson, K., Proby, C. M., Rashbass, J. & Leigh, I. M. (2019a). Nationwide Incidence of Metastatic Cutaneous Squamous Cell Carcinoma in England. *JAMA Dermatol*, 155(3), 298-306.

Venables, Z. C., Nijsten, T., Wong, K. F., Autier, P., Broggio, J., Deas, A., Harwood, C. A., Hollestein, L. M., Langan, S. M., Morgan, E., Proby, C. M., Rashbass, J. & Leigh, I. M. (2019b). Epidemiology of basal and cutaneous squamous cell carcinoma in the U.K. 2013–15: a cohort study. *Br J Dermatol,* 181(3): 474–482. (Tämä on fibroblastijutun viite)

Venables, Z. C., Tokez, S., Hollestein, L. M., Mooyaart, A. L., van den Bos, R. R., Rous, B., Leigh, I. M., Nijsten, T. & Wakkee, M. (2021). Validation of Four Cutaneous Squamous Cell Carcinoma Staging Systems Using Nationwide Data. *Br J Dermatol*, doi: 10.1111/bjd.20909.

Viallard, C. & Larrivée, B. (2017). Tumor Angiogenesis and Vascular Normalization: Alternative Therapeutic Targets. *Angiogenesis,* 20(4), 409-426.

Viiklepp, K., Nissinen, L., Ojalill, M., Riihilä, P., Kallajoki, M., Meri, S., Heino, J. & Kähäri, V. M. (2022). C1r upregulates production of matrix metalloproteinase-13 and promotes invasion of cutaneous squamous cell carcinoma *J Invest Dermatol,* 142(5), 1478-1488.

Virchow, R. (1858). Die Cellularpathologie in lhrer Begruendung auf Physiologische und Pathologische Gewebelehre (ed. Hirschwald, A.) (Berlin,1858)

Vosseler, S., Lederle, W., Airola, K., Obermueller, E., Fusenig, N. E. & Mueller, M. M. (2009). Distinct progression-associated expression of tumor and stromal MMPs in HaCaT skin SCCs correlates with onset of invasion. *Int J Cancer,* 125(10), 2296-2306.

Waldman, A. & Schmults, C. (2019). Cutaneous squamous cell carcinoma. *Hematol Oncol Clin N Am,* 33(1), 1-12.

Walker, C., Mojares, E. & Del Río Hernández, A. (2018). Role of extracellular matrix in development and cancer progression. *Int J Mol Sci,* 19(10), 3028.

Wang, L., Ding, L., Liu, Z., Sun, L., Chen, L., Jia, R., Dai, X., Cao, J. & Ye, J. (2020). Automated identification of malignancy in whole-slide pathological images: identification of eyelid malignant melanoma in gigapixel pathological slides using deep learning. *Br J Ophthalmol*, 104(3), 318-323.

Wang, H. M., Hung, C. H., Lu, S. N., Chen, C. H., Lee, C. M., Hu, T. H. & Wang, J. H. (2013). Liver stiffness measurement as an alternative to fibrotic stage in risk assessment of hepatocellular carcinoma incidence for chronic hepatitis C patients. *Liver Int*, 33(5), 756-761.

Wang, R., Sun, Y., Yu, W., Yan, Y., Qiao, M., Jiang, R., Guan, W. & Wang, L. (2019a). Downregulation of miRNA-214 in Cancer-Associated Fibroblasts Contributes to Migration and Invasion of Gastric Cancer Cells Through Targeting FGF9 and Inducing EMT. *J Exp Clin Cancer Res*, 38(1), 20.

Wang, S., Yang, D. M., Rong, R., Zhan, X. & Xiao, G. (2019b). Pathology Image Analysis Using Segmentation Deep Learning Algorithms. *Am J Pathol*, 189(9), 1686-1698.

Wang, X., Diao, Y., Liu, Y., Gao, N., Gao, D., Wan, Y., Zhong, J. & Jin G. (2015). Synergistic apoptosis-inducing effect of aspirin and isosorbide mononitrate on human colon cancer cells. *Mol Med Rep*, 12(3), 4750-4758.

Wang, X., Zhang, W., Sun, X., Lin, Y. & Chen W. (2018). Cancer-associated fibroblasts induce epithelial-mesenchymal transition through secreted cytokines in endometrial cancer cells. *Oncol Lett*, 15(4), 5694-5702.

Wang, X., Janowczyk, A., Zhou, Y., Thawani, R., Fu, P., Schalper, K., Velcheti, V. & Madabhushi, A. (2017). Prediction of recurrence in early stage non-small cell lung cancer using computer extracted nuclear features from digital H&E images. *Sci Rep*, 7(1), 13543.

Wehner, M. R., Cidre Serrano, W., Nosrati, A., Schoen, P. M., Chren, M. M., Boscardin, J. & Linos, E. (2018). All-cause mortality in patients with basal and squamous cell carcinoma: A systematic review and meta-analysis. *J Am Acad Dermatol*, 78(4), 663-672.

Weinstein, J. N., Collisson, E. A., Mills, G. B., Shaw, K. R., Ozenberger, B. A., Ellrott, K., Shmulevich, I., Sander, C. & Stuart, J. M. (2013). The Cancer Genome Atlas Pan-cancer analysis project. *Nat Genet*, 45(10), 1113-1120.

Weinstock, M. A., Bingham, S. F., Digiovanna, J. J., Rizzo, A. E., Marcolivio, K., Hall, R., Eilers, D., Naylor, M., Kirsner, R., Kalivas, J., Cole, G. & Vertrees, J. E.; Veterans Affairs Topical Tretinoin Chemoprevention Trial Group. (2012). Tretinoin and the prevention of keratinocyte carcinoma (Basal and squamous cell carcinoma of the skin): a veterans affairs randomized chemoprevention trial. *J Invest Dermatol*, 132(6), 1583-1590.

Wen, S., Hou, Y., Fu, L., Xi, L., Yang, D., Zhao, M., Qin, Y., Sun, K., Teng, Y. & Liu, M. (2019). Cancer-Associated Fibroblast (CAF)-Derived il32 Promotes Breast Cancer Cell Invasion and Metastasis via Integrin β3–p38 MAPK signalling. *Cancer Lett*, 442, 320-332.

Wendling, O., Bornert, J. M., Chambon, P. & Metzger, D. (2009). Efficient Temporally-Controlled Targeted Mutagenesis in Smooth Muscle Cells of the Adult Mouse. *Genesis*, 47(1), 14-18.

Werner, R. N., Sammain, A., Erdmann, R., Hartmann, V., Stockfleth, E. & Nast, A. (2013). The natural history of actinic keratosis: a systematic review. *Br J Dermatol*, 169(3), 502-518.

Wessels, F., Schmitt, M., Krieghoff-Henning, E., Jutzi, T., Worst, T. S., Waldbillig, F., Neuberger, M., Maron, R. C., Steeg, M., Gaiser, T., Hekler, A., Utikal, J. S., von Kalle, C., Fröhling, S., Michel, M. S., Nuhn, P. & Brinker, T. J. (2021). Deep learning approach to predict lymph node metastasis directly from primary tumor histology in prostate cancer. *BJU Int,* 128(3), 352-360.

Wikonkal, N. M., Brash, D. E. (1999). Ultraviolet radiation induced signature mutations in photocarcinogenesis. *J Investig Dermatol Symp Proc*, 4(1), 6-10.

Wiseman, B. S. & Werb, Z. (2002). Stromal effects on mammary gland development and breast cancer. *Science*, 296(5570), 1046-1049.

Wu, H. J., Hao, M., Yeo, S. K. & Guan, J. L. (2020a). FAK signaling in cancer-associated fibroblasts promotes breast cancer cell migration and metastasis by exosomal miRNAs-mediated intercellular communication. *Oncogene*, 39(12), 2539-2549.

Wu, Z., Wang, L., Li, C., Cai, Y., Liang, Y., Mo, X., Lu, Q., Dong, L. & Liu, Y. (2020b). DeepLRHE: A deep convolutional neural network framework to evaluate the risk of lung cancer recurrence and metastasis from histopathology images. *Front Genet*, 11, 768.

Wulczyn, E., Steiner, D. F., Xu, Z., Sadhwani, A., Wang, H., Flament-Auvigne, I., Mermel, C. H., Chen, P. C., Liu, Y. & Stumpe, M. C. (2020). Deep learning-based survival prediction for multiple cancer types using histopathology images. *PLoS One*, 15(6), e0233678.

Wysong, A., Newman, J. G., Covington, K. R., Kurley, S. J., Ibrahim, S. F., Farberg, A. S., Bar, A., Cleaver, N. J., Somani, A. K., Panther, D., Brodland, D. G., Zitelli, J., Toyohara, J., Maher, I. A., Xia, Y., Bibee, K., Griego, R., Rigel, D. S., Meldi Plasseraud, K., Estrada, S., Sholl, L. M., Johnson, C., Cook, R. W., Schmults, C. D. & Arron, S. T. (2021). Validation of a 40-gene expression profile test to predict metastatic risk in localized high-risk cutaneous squamous cell carcinoma. *J Am Acad Dermatol*, 84(2), 361-369.

Xiang, F., Lucas, R., Hales, S. & Neale, R. (2014). Incidence of nonmelanoma skin cancer in relation to ambient UV radiation in white populations, 1978-2012: empirical relationships. *JAMA Dermatol*, 150(10), 1063-1071.

Xing, F., Saidou, J. & Watabe, K. (2010). Cancer associated fibroblasts (CAFs) in tumor microenvironment. *Front Biosci (Landmark Ed)*, 15, 166-179.

Xing, F. & Yang, L. (2016). Machine learning and its application in microscopic image analysis. *Machine Learning and Medical Imaging*. Amsterdam, Netherlands: Elsevier Inc, 97-127.

Xu, R., Cai, M. Y., Luo, R. Z., Tian, X., Han, J. D. & Chen, M. K. (2016a). The Expression Status and Prognostic Value of Cancer Stem Cell Biomarker CD133 in Cutaneous Squamous Cell Carcinoma. *JAMA Dermatol*, 152(3), 305-311.

Xu, X., Chang, W., Yuan, J., Han, X., Tan, X., Ding, Y., Luo, Y., Cai, H., Liu, Y., Gao, X., Liu, Q., Yu, Y., Du, Y., Wang, H., Ma, L., Wang, J., Chen, K., Ding, Y., Fu, C. & Cao, G. (2016b). Periostin expression in intra-tumoral stromal cells is prognostic and predictive for colorectal carcinoma via creating a cancer-supportive niche. *Oncotarget*, 7(1), 798-813.

Yang, L., Serada, S., Fujimoto, M., Terao, M., Kotobuki, Y., Kitaba, S., Matsui, S., Kudo, A., Naka, T., Murota, H. & Katayama, I. (2012). Periostin facilitates skin sclerosis via PI3K/Akt dependent mechanism in a mouse model of scleroderma. *PLoS One*, 7(7), e41994.

Yantsos, V. A., Conrad, N., Zabawski, E. & Cockerell, C. J. (1999). Incipient intraepidermal cutaneous squamous cell carcinoma: a proposal for reclassifying andgrading solar (actinic) keratoses. *Semin Cutan Med Surg*, 18(1), 3-14.

Yin, C., Evason, K. J., Asahina, K. & Stainier, D. Y. (2013). Hepatic Stellate Cells in Liver Development, Regeneration, and Cancer. *J Clin Invest*, 123(5), 1902-1910.

Yuzawa, S., Kano, M. R., Einama, T. & Nishihara, H. (2012). PDGFRβ expression in tumor stroma of pancreatic adenocarcinoma as a reliable prognostic marker. *Med Oncol*, 29(4), 2824-2830.

Zaha, D. C. (2014). Significance of immunohistochemistry in breast cancer. *World J Clin Oncol*, 5(3), 382-392.

Zandi, R., Larsen, A. B., Andersen, P., Stockhausen, M. T. & Poulsen, H. S. (2007). Mechanisms for oncogenic activation of the epidermal growth factor receptor. *Cell Signal*, 19(10), 2013-2023.

Zang, X. & Allison, J. P. (2007). The B7 family and cancer therapy: costimulation and coinhibition. *Clin Cancer Res*, 13(18 Pt 1), 5271-5279.

Zanjani, F. G., Zinger, S., Bejnordi, B. E., van der Laak, J. A. & de With PH. (2018). Stain normalization of histopathology images using generative adversarial networks. *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 573-577. doi:10.1109/ISBI.2018.8363641.

Zarella, M. D., Bowman, D., Aeffner, F., Farahani, N., Xthona, A., Absar, S. F., Parwani, A., Bui, M. & Hartman, D. J. (2019). A Practical Guide to Whole Slide Imaging: A White Paper From the Digital Pathology Association. *Arch Pathol Lab Med*, 143(2), 222-234.

Zeisberg, E. M., Potenta, S., Xie, L., Zeisberg, M. & Kalluri, R. (2007). Discovery of Endothelial to Mesenchymal Transition as a Source for Carcinoma-Associated Fibroblasts. *Cancer Res*, 67(21), 10123-10128.

Zeisberg, E. M. & Zeisberg, M. (2013). The role of promoter hypermethylation in fibroblast activation and fibrogenesis. J Pathol, 229(2), 264-273.

Zhang, H., Deng, T., Liu, R., Ning, T., Yang, H., Liu, D., Zhang, Q., Lin, D., Ge, S., Bai, M., Wang, X., Zhang, L., Li, H., Yang, Y., Ji, Z., Wang, H., Ying, G. & Ba, Y. (2020). CAF secreted miR-522 suppresses ferroptosis and promotes acquired chemo-resistance in gastric cancer. *Mol Cancer*, 19(1), 43.

Zhang, J. & Liu, J. (2013). Tumor Stroma as Targets for Cancer Therapy. *Pharmacol Ther,* 137(2), 200-215.

Zhang, M., Qureshi, A. A., Geller, A. C., Frazier, L., Hunter, D. J. & Han, J. (2012). Use of tanning beds and incidence of skin cancer. *J Clin Oncol*, 30(14). 1588-1593.

Zhong, F., Bi, R., Yu, B., Yang, F., Yang, W. & Shui, R. (2016). A comparison of visual assessment and automated digital image analysis of Ki67 labeling index in breast cancer. *PLoS One*, 11(2), e0150505.

Zhou, L., Yang, K., Wickett, R. R., Kadekaro, A. L. & Zhang, Y. (2016). Targeted deactivation of cancer-associated fibroblasts by b-catenin ablation suppresses melanoma growth. *Tumour Biol,* 37(10), 14235-14248.

Zhu, Y., Shi, C., Zeng, L., Liu, G., Jiang, W., Zhang, X., Chen, S., Guo, J., Jian, X., Ouyang, J., Xia, J., Kuang, C., Fan, S., Wu, X., Wu, Y., Zhou, W. & Guan, Y. (2020). High COX-2 expression in cancer-associated fibiroblasts contributes to poor survival and promotes migration and invasiveness in nasopharyngeal carcinoma. *Mol Carcinog*, 59(3), 265-280.

Zoetemelk, M., Rausch, M., Colin, D. J., Dormond, O. & Nowak-Sliwinska, P. (2019). Short-term 3D culture systems of various complexity for treatment optimization of colorectal carcinoma. *Sci Rep,* 9(1), 7103.

Zou, B., Liu, X., Gong, Y., Cai, C., Li, P., Xing, S., Pokhrel, B., Zhang, B. & Li, J. (2018). A novel 12-marker panel of cancer-associated fibroblasts involved in progression of hepatocellular carcinoma. *Cancer Manag Res,* 10, 5303-5311.

Öhlund, D., Elyada, E. & Tuveson, D. (2014). Fibroblast Heterogeneity in the Cancer Wound. *J Exp Med,* 211(8), 1503-1523.

Öhlund, D., Handly-Santana, A., Biffi, G., Elyada, E., Almeida, A. S., Ponz-Sarvise, M., Corbo, V., Oni, T. E., Hearn, S. A., Lee, E. J., Chio, I. I., Hwang, C. I., Tiriac, H., Baker, L. A., Engle, D. D., Feig, C., Kultti, A., Egeblad, M., Fearon, D. T., Crawford, J. M., Clevers, H., Park, Y. & Tuveson, D. A. (2017). Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. *J Exp Med*, 214(3), 579-596.

Österreicher, C. H., Penz-Österreicher, M., Grivennikov, S. I., Guma, M., Koltsova, E. K., Datz, C., Sasik, R., Hardiman, G., Karin, M. & Brenner, D. A. (2011). Fibroblast-specific protein 1 identifies an inflammatory subpopulation of macrophages in the liver. Proc Natl Acad Sci U S A, 108(1), 308-313.

Östman, A. (2017). PDGF receptors in tumor stroma: Biological effects and associations with prognosis and response to treatment. *Adv Drug Deliv Rev*, 121, 117-123.

Özdemir, B. C., Pentcheva-Hoang, T., Carstens, J. L., Zheng, X., Wu, C. C., Simpson, T. R., Laklai, H., Sugimoto, H., Kahlert, C., Novitskiy, S. V., De Jesus-Acosta, A., Sharma, P., Heidari, P., Mahmood, U., Chin, L., Moses, H. L., Weaver, V. M., Maitra, A., Allison, J. P., LeBleu, V. S. & Kalluri R. (2014). Depletion of carcinoma-associated fibroblasts and fibrosis induces immunosuppression and accelerates pancreas cancer with reduced survival. *Cancer Cell*, 25(6), 719-734.

**TURUN
YLIOPISTO**
UNIVERSITY
OF TURKU