Data-driven identification of targets for fluorescence-guided surgery in non-small cell lung cancer

Meijer, R.P.J.; Neijenhuis, L.K.A.; Zeilstra, A.P.; Roerink, S.F.; Bhairosingh, S.S.; Hilling, D.E.; ... ; Hutteman, M.

Molecular Imaging and Biology (2023) 25:228–239
https://doi.org/10.1007/s11307-022-01791-5

RESEARCH ARTICLE

Data-Driven Identification of Targets for Fluorescence-Guided Surgery in Non-Small Cell Lung Cancer

Ruben P. J. Meijer[1,2] · Lisanne K. A. Neijenhuis[1,2] · Annette P. Zeilstra[1] · Sophie F. Roerink[3] · Shadhvi S. Bhairosingh[1] · Denise E. Hilling[1,4] · J. Sven D. Mieog[1] · Peter J. K. Kuppen[1] · Cornelis F. M. Sier[1,5] · Jerry Braun[6] · Jacobus Burggraaf[1,2] · Alexander L. Vahrmeijer[1] · Danielle Cohen[7] · Merlijn Hutteman[1,6]

Received: 30 June 2022 / Revised: 16 November 2022 / Accepted: 17 November 2022 / Published online: 27 December 2022

Abstract

Purpose Intraoperative identification of lung tumors can be challenging. Tumor-targeted fluorescence-guided surgery can provide surgeons with a tool for real-time intraoperative tumor detection. This study evaluated cell surface biomarkers, partially selected via data-driven selection software, as potential targets for fluorescence-guided surgery in non-small cell lung cancers: adenocarcinomas (ADC), adenocarcinomas *in situ* (AIS), and squamous cell carcinomas (SCC).

Procedures Formalin-fixed paraffin-embedded tissue slides of resection specimens from 15 patients with ADC and 15 patients with SCC were used and compared to healthy tissue. Molecular targets were selected based on two strategies: (1) a data-driven selection using > 275 multi-omics databases, literature, and experimental evidence; and (2) the availability of a fluorescent targeting ligand in advanced stages of clinical development. The selected targets were carbonic anhydrase 9 (CAIX), collagen type XVII alpha 1 chain (collagen XVII), glucose transporter 1 (GLUT1), G protein-coupled receptor 87 (GPR87), transmembrane protease serine 4 (TMPRSS4), carcinoembryonic antigen (CEA), epithelial cell adhesion molecule (EpCAM), folate receptor alpha (FRα), integrin αvβ6 (αvβ6), and urokinase-type plasminogen activator receptor (uPAR). Tumor expression of these targets was assessed by immunohistochemical staining. A total immunostaining score (TIS, range 0–12), combining the percentage and intensity of stained cells, was calculated. The most promising targets in ADC were explored in six AIS tissue slides to explore its potential in non-palpable lesions.

Results Statistically significant differences in TIS between healthy lung and tumor tissue for ADC samples were found for CEA, EpCAM, FRα, αvβ6, CAIX, collagen XVII, GLUT-1, and TMPRSS4, and of these, CEA, CAIX, and collagen XVII were also found in AIS. For SCC, EpCAM, uPAR, CAIX, collagen XVII, and GLUT-1 were found to be overexpressed.

Conclusions EpCAM, CAIX, and Collagen XVII were identified using concomitant use of data-driven selection software and clinical evidence as promising targets for intraoperative fluorescence imaging for both major subtypes of non-small cell lung carcinomas.

Key words Lung cancer · Molecular imaging · Near-infrared fluorescence imaging · Target selection · Data-driven selection · Immunohistochemical staining

Ruben P. J. Meijer and Lisanne K. A. Neijenhuis contributed equally to the manuscript.

✉ Merlijn Hutteman
m.hutteman@lumc.nl

1 Department of Surgery, Leiden University Medical Center, Leiden, The Netherlands

2 Centre for Human Drug Research, Leiden, The Netherlands

3 Euretos, Utrecht, The Netherlands

4 Department of Surgical Oncology and Gastrointestinal Surgery, Erasmus MC Cancer Institute, University Medical Center Rotterdam, Rotterdam, The Netherlands

5 Percuros BV, Leiden, The Netherlands

6 Department of Cardiothoracic Surgery, Leiden University Medical Center, 2333 ZA Leiden, The Netherlands

7 Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands

Introduction

Non-small cell lung cancer (NSCLC) is the leading cause of cancer-related deaths worldwide [1]. Most NSCLC patients are diagnosed with disseminated disease, leaving only palliative treatment options. Therefore, much attention is given to screening programs to identify patients with early-stage tumors [2–4] This may change the treatment options for instance by lung parenchyma sparing resections, potentially resulting in more residual lung tissue and better postoperative lung function.

Although lung parenchyma sparing resections (e.g., segmentectomies), as treatment of small-sized pulmonary nodules, have good clinical outcomes in terms of disease-free and overall survival, these may be technically difficult, particularly for small nodules [5–7]. Especially adenocarcinomas *in situ* (AIS) are difficult to identify as they are barely palpable. In addition, more lung resections are performed minimally invasive, using video-assisted thoracoscopic surgery (VATS) or robotic-assisted thoracoscopic surgery (RATS), and the diminished or absent tactile feedback increases the risk of missing synchronous local metastases [8, 9]. Also, in patients with more advanced tumor stages, there may be challenges albeit from a different nature. Namely, tumor-positive resection margins are found at histopathologic assessment in 20% of the patients with T3 or T4 lung cancer [10, 11], which negatively affect prognosis and approximately halves the 5-year survival rate [11–13]. Moreover, up to 60% of patients develop a recurrence despite curatively intended surgeries of early-stage NSCLC [14–16]. It is hypothesized that a significant portion of these local recurrences is explained by occult local metastases that are already present during primary surgery. Tumor-targeted fluorescence-guided surgery has the potential to overcome several of these challenges.

Near-infrared (NIR) fluorescence imaging is a real-time imaging technique, which requires a dedicated camera system and a fluorescent agent [17]. These agents can be divided into two groups: non-targeted and targeted, binding to a specific ligand, which is upregulated in tumor cells or activated by the tumor-specific environment. Light in the NIR spectrum (700–900 nm) has better depth penetration, up to 10 mm, compared to visible light. Moreover, this light is invisible to the human eye and will thereby not interfere with the standard surgical field [18]. The potentially added value of tumor-targeted fluorescence imaging in lung parenchyma sparing resections of NSCLC is threefold: detection of the surgical target, identification of synchronous additional, mostly small malignant lesions, and reduction of tumor-positive resection margins.

Optimal target characteristics for fluorescence-guided surgery include target localization on the cell membrane, evenly distribution throughout the tumor tissue, low expression in surrounding normal tissue, and upregulation in a majority percentage of patients [19]. Potential targets are often identified by reviewing key publications from recent literature. Currently, a data-driven selection method is available that combines multiple databases and interlinks literature. It provides researchers with unique insights and has the potential to dramatically reduce the time required compared to a conventional literature search. This study aims to identify potentially suitable targets for fluorescence-guided surgery in NSCLC, both using conventional strategies, and with the use of a data-driven selection platform, combining public data on gene expression and protein localization.

Material and Methods

Patients

Formalin-fixed paraffin-embedded tissue blocks from 30 patients (15 adenocarcinomas (ADC) and 15 squamous cell carcinomas (SCC)) who underwent surgical resection of a primary NSCLC between 2000 and 2017 and without previous neoadjuvant therapy were randomly selected from the Department of Pathology, Leiden University Medical Center (LUMC), the Netherlands. Medical records and pathology reports were retrospectively reviewed. A board-certified pathologist (DC) reviewed all tissue samples before inclusion in the study. After analysis of the immunohistochemistry staining results of the ADC and SCC slides, additional analysis on a random selection of 6 AIS was performed. AIS tissue slides were stained for the most promising targets, based on the immunostaining results in ADC.

Tissue samples were used in accordance with the code for secondary use of human tissue as prescribed by the Dutch Federation of Medical Scientific Societies, which does not require additional informed consent. The study was approved by the Institutional Ethics Review Board of the Leiden University Medical Center (Leiden, The Netherlands).

Target Selection

In this study, a selection of ten potential targets was chosen based on two different strategies. The first 5 targets were chosen after an unbiased search using data-driven selection software (Euretos, Utrecht, the Netherlands)—integrating more than 275 multi-omics databases along with literature and experimental evidence. This data-driven selection platform compares the RNA expression of genes between different types of tissues. In our search, we compared RNA expression in lung cancer to healthy lung and lymph node tissue. RNA expression data on primary NSCLC were derived from the TCGA-LUAD and TCGA-LUSC datasets

[20, 21]. Expression data on healthy lymph node tissue and healthy lung tissue was derived from the GTEX-project [22]. All expression values were normalized to transcripts per million (TPM), which is the most effective normalization in reducing non-biological variability in transcriptomic data [23]. To predict whether encoded proteins are present on the cell surface, we explored presence in the experimentally derived cell surface protein atlas (CSPA) [24] or bioinformatically predicted sets of cell surface proteins [25, 26]. We performed an exhaustive search for targets by applying the following criteria: (1) median expression in tumor exceeding expression in the healthy lung or lymph nodes and (2) surface expression predicted or experimentally confirmed. The different proteins for ADC (56 results) and SCC (114 results) were presented with multiple parameters per gene. A log2-fold change of tumor vs normal tissue was calculated by converting the TPM-values to log2-values and analyzed with the limma-package in R [27]. Log2-fold changes were then obtained from the topTable-function. From this gene list, only genes that are abundantly expressed in both ADC and SCC but minimally in healthy lung tissue or lymph nodes were included, resulting in a set of 30 genes. The five genes with the highest log2-fold change of tumor in comparison to normal were then selected for immunohistochemical analysis: carbonic anhydrase 9 (CAIX), collagen type XVII alpha 1 chain (collagen XVII), glucose transporter 1 (GLUT1), G protein-coupled receptor 87 (GPR87), and transmembrane protease serine 4 (TMPRSS4).

In addition to the data-driven selection-derived targets, 5 targets were included, based on the availability of a fluorescent targeting ligand in advanced stages of clinical development for the detection of various cancer types: carcinoembryonic antigen (CEA), epithelial cell adhesion molecule (EpCAM), folate receptor alpha (FRα), integrin αvβ6 (αvβ6), and urokinase-type plasminogen activator receptor (uPAR).

Immunohistochemistry Staining

Formalin-fixed and paraffin-embedded (FFPE) tissue blocks were cut in sections with a thickness of 4 µm and mounted on adhesive Starfrost slides (Waldemar Knittel Glasbearbeitungs GmbH, Braunschweig, Germany). Subsequently, the slides were deparaffinized with xylene and rehydrated in decreasing concentrations of diluted ethanol (100%, 70%, and 50%). Then slides were rinsed with demineralized water, and endogenous peroxidase was blocked with 0.3% hydrogen peroxide (Merck Millipore, Darmstadt, Germany) for 20 min at room temperature. The method used for antigen retrieval (AR) was dependent on the protocol of the antibody (Supplementary Table 1). Antigen retrieval for uPAR, CEA, FRα, GLUT-1, CAIX, collagen XVII, and GPR87 was performed by heat induction at 95 °C during 10 min using PT Link (Dako) with low-pH (pH 6.0). For TMPRSS4, PT Link with high-pH (pH 9.0) was used. For αvβ6, antigen retrieval was performed with 0.4% pepsin at 37 °C using a water bath for 10 min, and for EpCAM, trypsin was used during 20 min. The sections were incubated overnight at room temperature with primary antibody (diluted in 1% BSA/PBS). The negative control was incubated with 1% BSA/PBS. Tumor sections known to express the target(s) were used as positive controls. The slides were washed with PBS followed by incubation with HRP-labelled secondary antibody (either anti-mouse or anti-rabbit (both Dako)) for 30 min at room temperature. For staining, DAB substrate was applied for 10 min, and for counterstaining, hematoxylin (VWR international, Amsterdam, the Netherlands) was applied for 15 s. Lastly, the tissue sections were dehydrated at 37 °C for 60 min and mounted in Pertex (Histolab, Askim, Sweden).

Immunohistochemistry Staining Scoring

Two authors (AZ and DC) evaluated the stained slides independently. In case of discrepancy, rescoring was done until consensus was achieved. Total immunostaining score (TIS) was used for tumor and normal tissue. TIS is the mathematical product of a proportion score (PS) and an intensity score (IS) with a maximum score of 12. The PS describes the estimated fraction of positively stained tumor cells (0 = none, 1 = < 10%, 2 = 10–50%, 3 = 51–80%, 4 = > 80%), and the IS represents the estimated staining intensity (0 = no staining, 1 = weak, 2 = moderate, 3 = strong). The calculated TIS was then defined as no expression (0), weak expression (1–4), moderate expression (6–8), or intense (9–12) expression.

Statistical Analysis

For statistical analysis, IBM SPSS Statistics version 25 (IBM Corp., Armonk, NY, USA) and GraphPad Prism 6 (GraphPad Software Inc., La Jolla, CA, USA) were used. Paired differences in expression level between normal and tumor tissue were calculated using the Wilcoxon signed rank test. Test results were considered statistically significant at the level of $p < 0.05$.

Results

Patient and tumor characteristics are summarized in Table 1. Histological subtyping showed heterogeneity within the ADC group; all SCC samples were non-keratinizing tumors. Representative images of tissue samples are presented in Fig. 1 (ADC), Fig. 2 (AIS), and Fig. 3 (SCC). In the case of tumor-positive staining, a predominant cell membrane pattern was seen in both ADC and SCC for CEA, EpCAM,

Table 1 Patient and tumor characteristics

Characteristics	ADC ($n=15$)	AIS ($n=6$)	SCC ($n=15$)
Age at surgery median years (range)	69 (53–79)	65.5 (47–72)	66 (57–78)
Gender			
Male	9	2	11
Female	6	4	4
Tumor size (cm) median (range)	2.9 (1.2–5.2)	1.3 (0.6–2.6)	4.2 (1.8–8.2)
Histological subtype			
Acinair	9		
Lepidic	6	6	
Mucinous	2		
Micropapillary	3		
Papillary	4		
Solid	5		
Keratinized			0
Non-keratinized			15
TNM staging			
Tumour stage (T)			
1	5	5	5
2	7	0	4
3	1	0	3
4	2	1	3
Nodal stage (N)			
0	13	6	10
1	2	0	5
2	0	0	0
3	0	0	0
Metastatic stage (M)			
0	15	6	14
1	0	0	1

Abbreviations: *ADC*, adenocarcinoma; *AIS*, adenocarcinoma *in situ*; *SCC*, squamous cell carcinoma cm centimeter

FRα, αvβ6, CAIX, collagen XVII, GLUT-1, and GPR87. The staining pattern of uPAR was dominant in the extracellular matrix, whereas TMPRSS4 showed positive staining for both the cell membrane as the extracellular matrix.

Adenocarcinoma

Based on the TIS, normal tissue had low (no to weak) expression in all of the samples for CEA, EpCAM, FRα, uPAR, collagen XVII, GLUT-1, and GPR87, whereas 86.7% of the samples were low for CAIX (Fig. 4, Supplementary Table 2). Tumor tissue showed high expression (moderate to intense) in 100% of the samples for TMPRSS4 and in 93.3% for αvβ6, CAIX, and collagen XVII. Matched sampling using the Wilcoxon-signed rank test showed significantly higher expression in the tumor compared to the adjacent normal lung tissue for CEA, EpCAM, FRα, αvβ6, CAIX, collagen XVII, GLUT-1, and TMPRSS4 (respectively $p=0.003$, 0.003, 0.005, 0.006, 0.001, 0.000, 0.007, and 0.035; Fig. 5). No significant difference in staining intensity was found for uPAR and GPR87 (both $p=0.066$). Multiple samples showed CAIX and collagen XVII expression in necrotic tumor fields. The uPAR staining pattern for all samples was dominant in stromal tissue, as was staining GLUT-1 in erythrocytes and nerve sheath branches and for TMPRSS4 in lung epithelium.

Adenocarcinoma *In Situ*

Since AIS in particular are barely palpable, 6 additional AIS were stained for the four most promising targets in ADC. Low target expression in normal lung tissue was seen for all CEA, EpCAM, CAIX, and collagen XVII samples (Fig. 6, Supplementary Table 3). Tumor expression was high in 66.7%, 0%, 83.3%, and 100% of the samples for respectively CEA, EpCAM, CAIX, and collagen XVII. Differences in staining intensity between AIS and healthy lung were significant for CEA, CAIX, and collagen XVII (respectively $p=0.043$, 0.024, and 0.026) (Fig. 7). No statistical difference was found for EpCAM ($p=0.180$).

Squamous Cell Carcinoma

The TIS value for CEA, EpCAM, uPAR, collagen XVII, GLUT-1, and GPR87 was low in all normal tissue samples. Expression of FRα and CAIX was low in 93.3% of normal tissue samples (Fig. 8, Supplementary Table 4). A high TIS value in tumor tissue was seen in 100% of the samples for GLUT-1 and TMPRSS4 and in 86.7% of the samples for collagen XVII. Differences in staining intensity between tumor and healthy lung were significant for EpCAM, uPAR, CAIX, collagen XVII, and GLUT-1 (respectively $p=0.003$, 0.012, 0.003, 0.001, and 0.000; Fig. 9). No statistical difference was found for CEA, FRα, αvβ6, GPR87, and TMPRSS4 (respectively $p=0.059$, 0.317, 0.705, 0.066, and 0.102). Necrotic tumor fields were positive for CEA, αvβ6, uPAR, and GLUT-1 in multiple samples. uPAR showed high expression in the stromal tissue of all samples. A substantial number of samples showed staining for GLUT-1 in erythrocytes and nerve sheath branches and TMPRSS4 in lung epithelium.

Discussion

A suitable molecular target for FGS in NSCLC should be overexpressed in the tumor compared to adjacent normal lung tissue. As NSCLC comprises various subtypes, it would be beneficial if a potential agent is suitable for multiple types of NSCLC, hence overexpressed in ADC, AIC, and SCC.

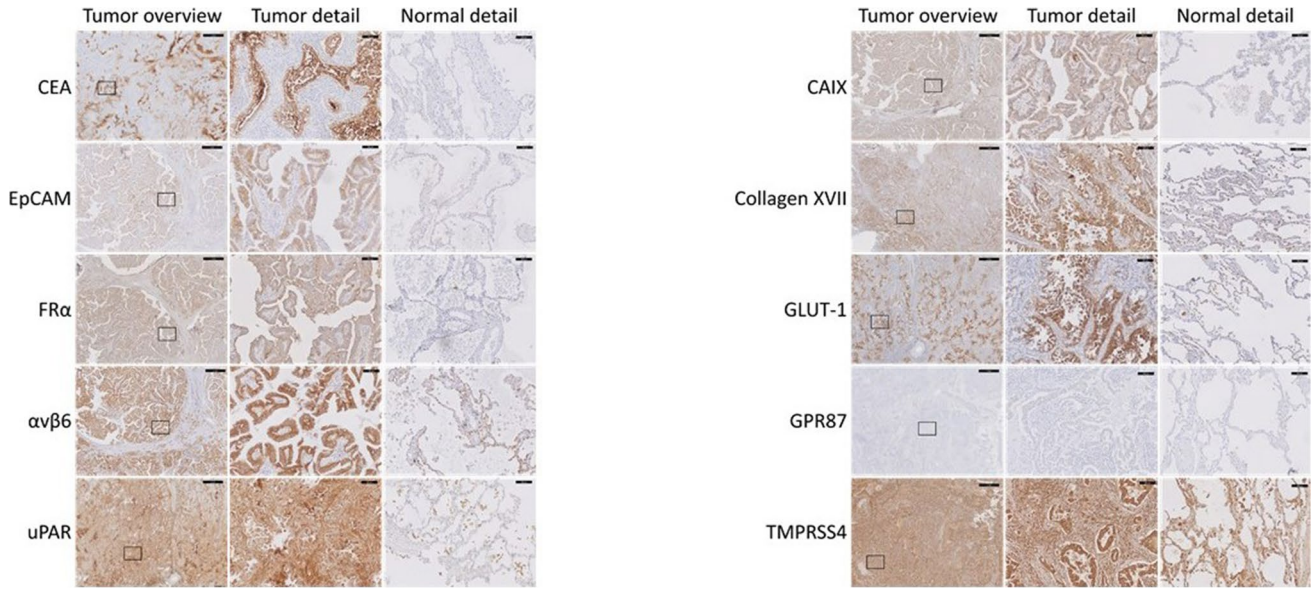

Fig. 1 Representative slides of adenocarcinoma cases stained for different markers. A representative lung tumor sample (first row), detailed tumor sample (second row), and detailed normal lung tissue (third row) are shown for each marker. The area of the detailed tumor staining is indicated with a rectangle in the tumor overview. The black bar represents 1 mm in the overview pictures and 100 µm in the detailed pictures. Abbreviations: *CEA*, carcinoembryonic antigen; *EpCAM*, epithelial cell adhesion molecule; *FRα*, folate receptor alpha; *αvβ6*, integrin αvβ6; *uPAR*, urokinase-type plasminogen activator receptor; *CAIX*, carbonic anhydrase 9; *Collagen XVII*, collagen type XVII alpha 1 chain; *GLUT-1*, glucose transporter 1; *GPR87*, G protein-coupled receptor 87; *TMPRSS4*, transmembrane protease serine 4

In this study, ten cell surface targets in NSCLC were investigated with IHC staining to find a suitable target for tumor-targeted fluorescence imaging. Five targets were selected using an extensive multiomics database search using data-driven selection software, and those were compared with five targets chosen based on the availability of fluorescent targeting ligands in advanced stages of clinical development in other cancer types.

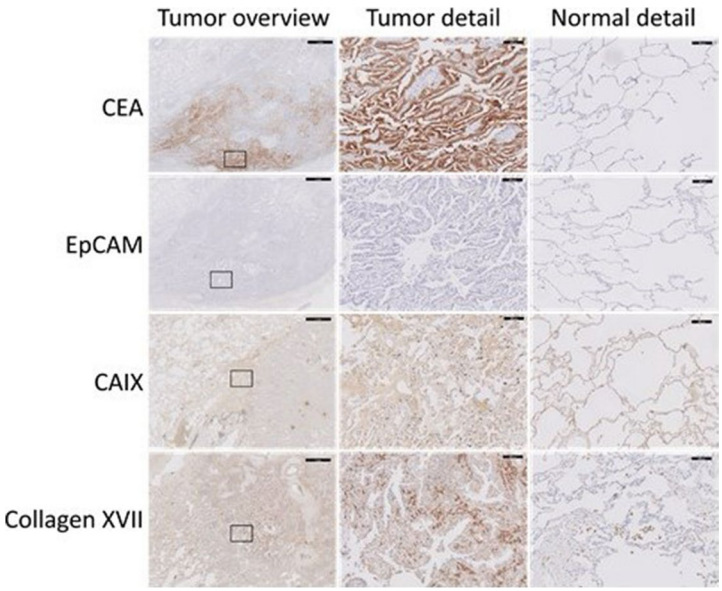

Fig. 2 Representative slides of adenocarcinoma *in situ* cases stained for different markers. A representative lung tumor sample (first row), detailed tumor sample (second row), and detailed normal lung tissue (third row) are shown for each marker. The area of the detailed tumor staining is indicated with a rectangle in the tumor overview. The black bar represents 1 mm in the overview pictures and 100 µm in the detailed pictures. Abbreviations: *CEA*, carcinoembryonic antigen; *EpCAM*, epithelial cell adhesion molecule; *CAIX*, carbonic anhydrase 9; *Collagen XVII*, collagen type XVII alpha 1 chain

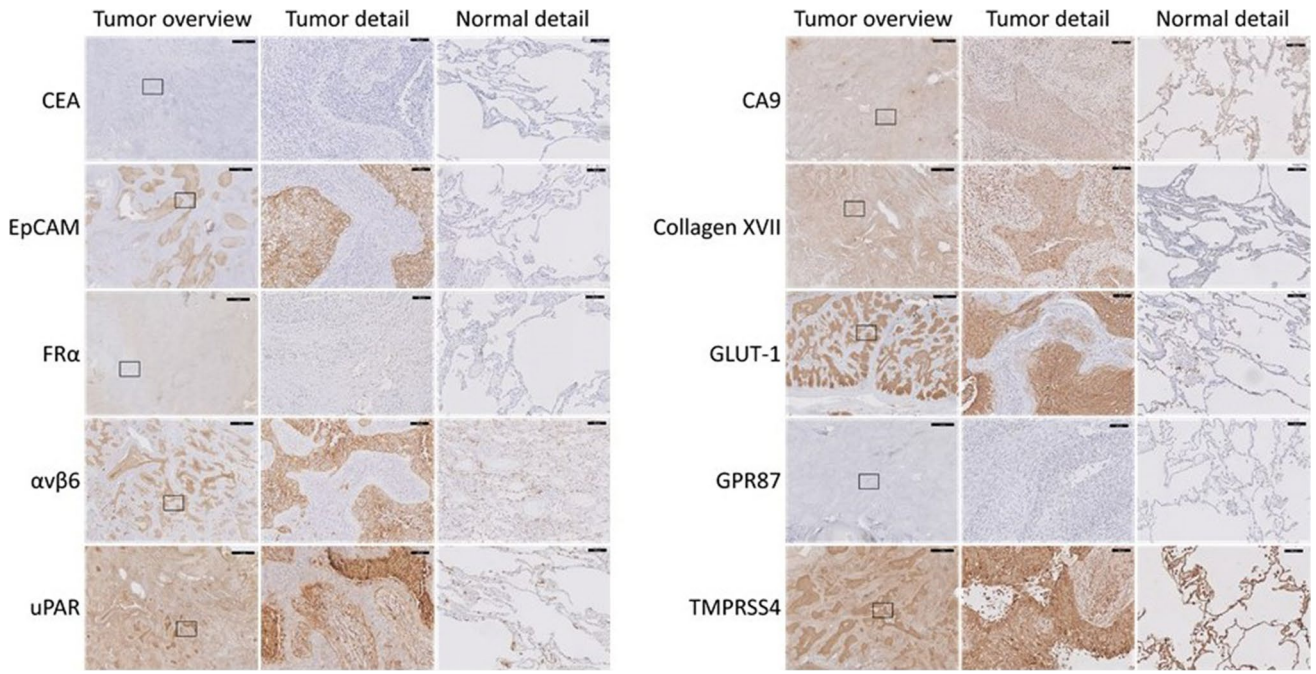

Fig. 3 Representative slides of squamous cell carcinoma cases stained for different markers. A representative lung tumor sample (first row), detailed tumor sample (second row), and detailed normal lung tissue (third row) are shown for each marker. The area of the detailed tumor staining is indicated with a rectangle in the tumor overview. The black bar represents 1 mm in the overview pictures and 100 µm in the detailed pictures. Abbreviations: *CEA*, carcinoembryonic antigen; *EpCAM*, epithelial cell adhesion molecule; *FRα*, folate receptor alpha; *αvβ6*, integrin αvβ6; *uPAR*, urokinase-type plasminogen activator receptor; *CAIX*, carbonic anhydrase 9; *Collagen XVII*, collagen type XVII alpha 1 chain; *GLUT-1*, glucose transporter 1; *GPR87*, G protein-coupled receptor 87; *TMPRSS4*, transmembrane protease serine 4

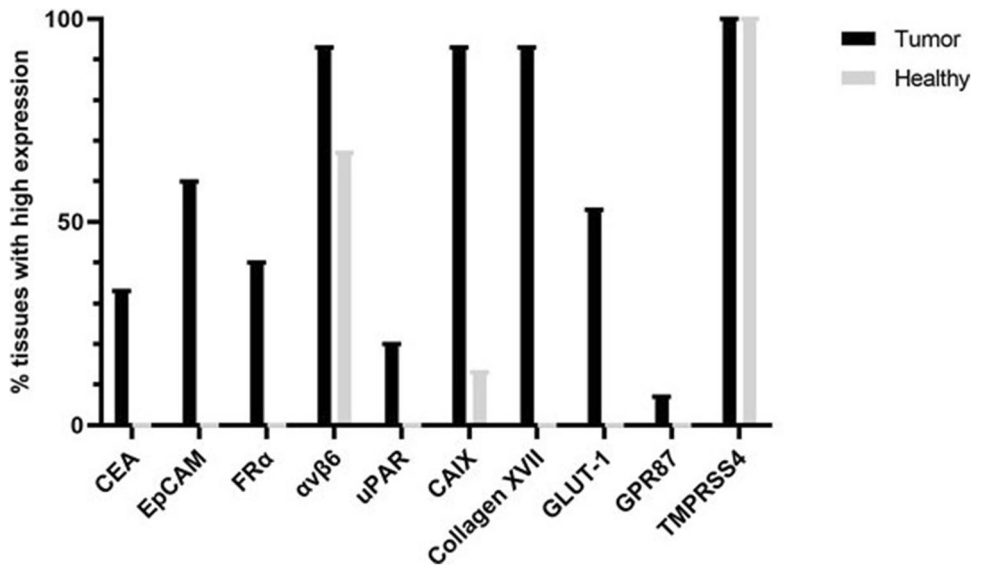

Fig. 4 Target expression in adenocarcinomas. This figure shows per target the percentage of adenocarcinoma tissue slides and healthy tissue slides that had high target expression with immunohistochemistry staining. High expression was defined as a total immunostaining score of 6 or higher. The TIS is a product of a proportion score (0 = none, 1 = < 10%, 2 = 10–50%, 3 = 51–80%, 4 = > 80%) and intensity score (0 = no staining, 1 = weak, 2 = moderate, 3 = strong). Abbreviations: *TIS*, total immunostaining score; *CEA*, carcinoembryonic antigen; *EpCAM*, epithelial cell adhesion molecule; *FRα*, folate receptor alpha; *αvβ6*, integrin αvβ6; *uPAR*, urokinase-type plasminogen activator receptor; *CAIX*, carbonic anhydrase 9; *Collagen XVII*, collagen type XVII alpha 1 chain; *GLUT-1*, glucose transporter 1; *GPR87*, G protein-coupled receptor 87; *TMPRSS4*, transmembrane protease serine 4

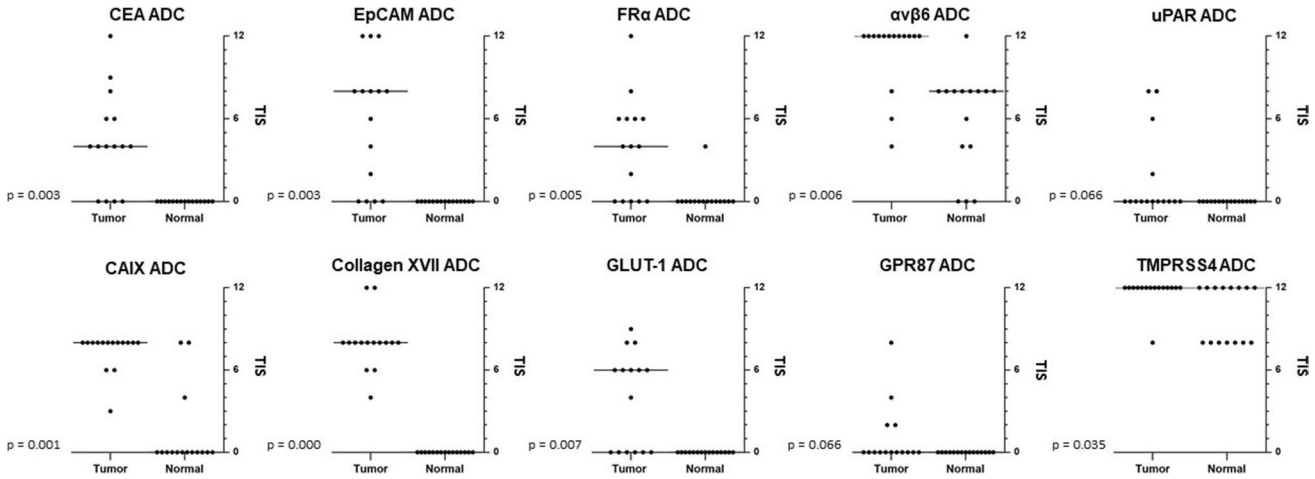

Fig. 5 Target expression of adjacent normal lung compared to adenocarcinoma tissue. Shown are the total immunostaining scores of the normal lung and tumor tissue in adenocarcinomas, per patient and per biomarker. The horizontal line indicates the median TIS. Abbreviations: *CEA*, carcinoembryonic antigen; *EpCAM*, epithelial cell adhesion molecule; *FRα*, folate receptor alpha; *αvβ6*, integrin αvβ6; *uPAR*, urokinase-type plasminogen activator receptor; *CAIX*, carbonic anhydrase 9; *Collagen XVII*, collagen type XVII alpha 1 chain; *GLUT-1*, glucose transporter 1; *GPR87*, G protein-coupled receptor 87; *TMPRSS4*, transmembrane protease serine 4

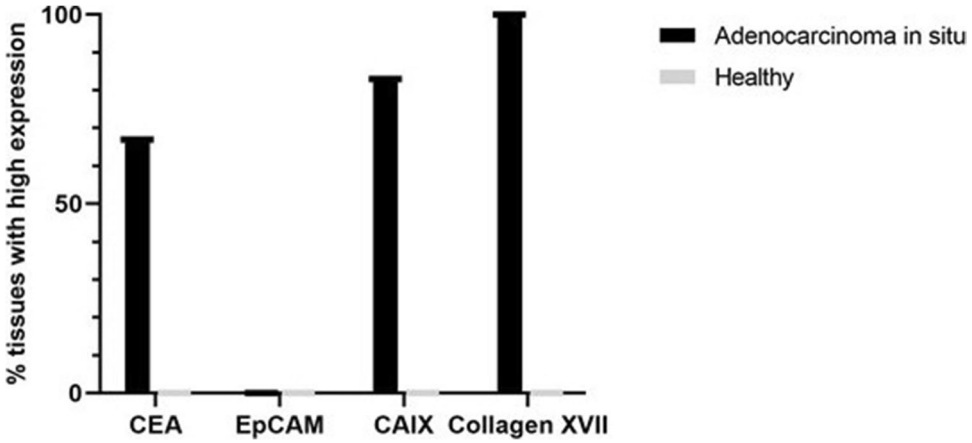

Fig. 6 Target expression in adenocarcinomas in situ. This figure shows per target the percentage of adenocarcinoma in situ tissue slides and healthy tissue slides that had high target expression with immunohistochemistry staining. High expression was defined as a total immunostaining score of 6 or higher. The TIS is a product of a proportion score (0 = no staining, 1 = weak, 2 = moderate, 3 = strong). Abbreviations: *TIS*, total immunostaining score; *CEA*, carcinoembryonic antigen; *EpCAM*, epithelial cell adhesion molecule; *CAIX*, carbonic anhydrase 9; *Collagen XVII*, collagen type XVII alpha 1 chain

Following this strategy, EpCAM, CAIX, collagen XVII, and GLUT-1 are significantly overexpressed in both ADC and SCC. However, in contrast to EpCAM, CAIX, and collagen XVII, GLUT-1's high expression on erythrocytes makes it an unsuitable target for FGS as this will most likely provide high background fluorescence. Resulting in a high false positive rate using fluorescence imaging during surgery. Even though the expression of GLUT-1 on erythrocytes has been described in previous publications, it was reported as a promising option after our data-driven selection [28]. This was not an error of the platform itself, but a consequence of an incomplete search strategy. To perform a search in the platform, you have to decide between which tissue types you want to compare RNA expression in genes. In our research, we chose to compare lung tumor tissue to normal lung and lymph node tissue. Besides tumor tissue and normal lung tissue, this data-driven selection did not include target expression in other cell types present in the healthy lung (e.g., blood). This emphasizes the importance of a relevant and complete research question before starting a search.

Data-driven selection software identified 2 out of the 3 most promising targets (CAIX and collagen XVII).

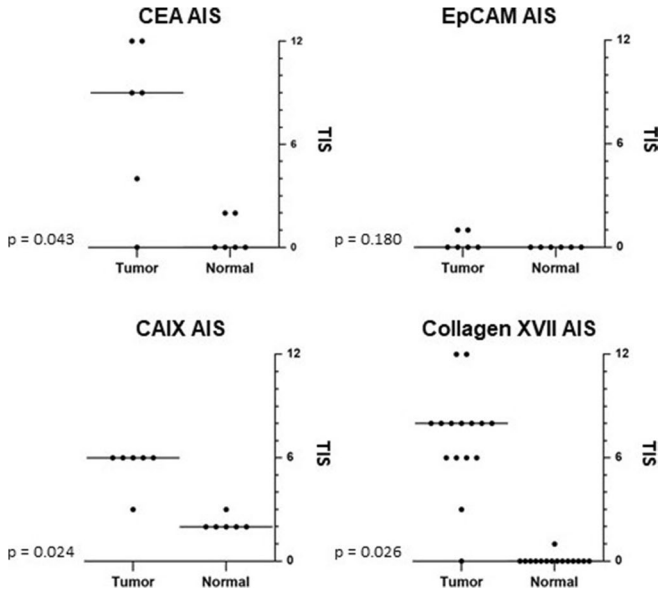

Fig. 7 Target expression of adjacent normal lung compared to adenocarcinoma in situ tissue. Shown are the total immunostaining scores of the normal lung and adenocarcinoma in situ tissue, per patient and per biomarker. The horizontal line indicates the median TIS. Abbreviations: *CEA*, carcinoembryonic antigen; *EpCAM*, epithelial cell adhesion molecule; *CAIX*, carbonic anhydrase 9; *Collagen XVII*, collagen type XVII alpha 1 chain

EpCAM, chosen based on our institution's experience in colorectal cancer, was also identified as a potential target for ADC in the initial results of the data-driven selection software. Due to the absence in the SCC results and, according to the data-driven selection, the relatively high expression on normal lung tissue, it was excluded in the final selection. These conflicting results between the results of the data-driven selection and IHC demonstrate the value of IHC confirmation of the data-driven selection. It must be stated that data-driven selection software is a discovery

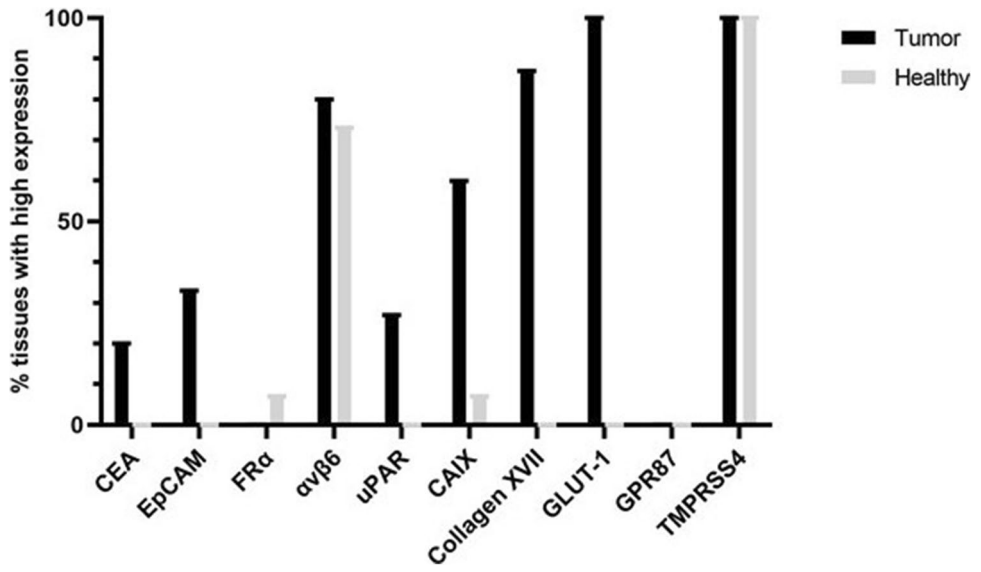

Fig. 8 Target expression in squamous cell carcinomas. This figure shows per target the percentage of squamous cell carcinoma tissue slides and healthy tissue slides that had high target expression with immunohistochemistry staining. High expression was defined as a total immunostaining score of 6 or higher. The TIS is a product of a proportion score (0 = none, 1 = < 10%, 2 = 10–50%, 3 = 51–80%, 4 = > 80%) and intensity score (0 = no staining, 1 = weak, 2 = moderate, 3 = strong). Abbreviations: *TIS*, total immunostaining score; *CEA*, carcinoembryonic antigen; *EpCAM*, epithelial cell adhesion molecule; *FRα*, folate receptor alpha; *αvβ6*, integrin αvβ6; *uPAR*, urokinase-type plasminogen activator receptor; *CAIX*, carbonic anhydrase 9; *Collagen XVII*, collagen type XVII alpha 1 chain; *GLUT-1*, glucose transporter 1; *GPR87*, G protein-coupled receptor 87; *TMPRSS4*, transmembrane protease serine 4

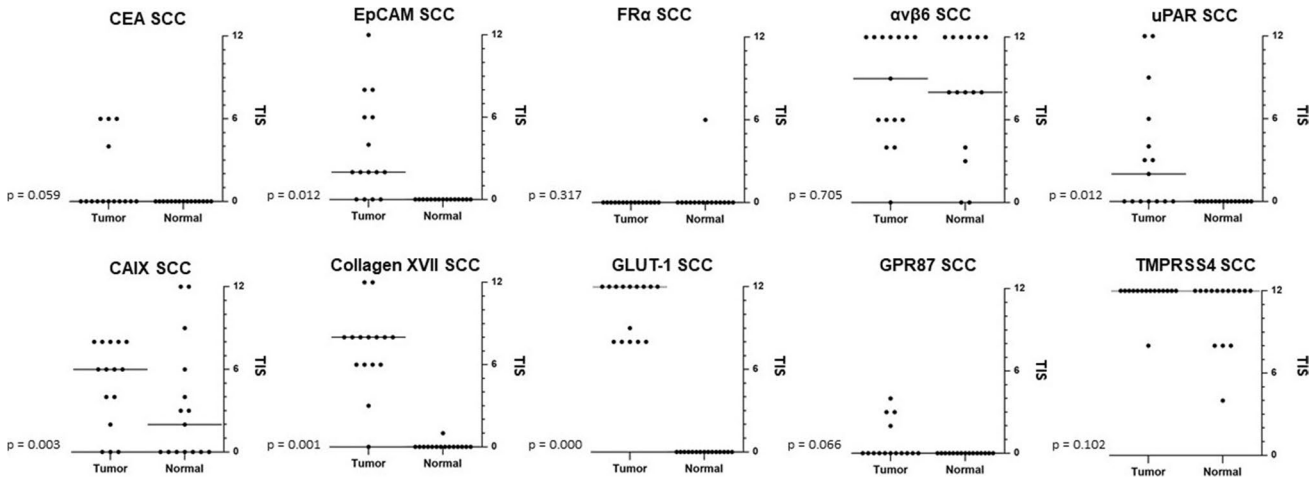

Fig. 9 Target expression of adjacent normal lung compared to squamous cell carcinoma tissue. Shown are the total immunostaining scores of the normal lung and tumor tissue in squamous cell carcinomas, per patient and per biomarker. The horizontal line indicates the median TIS. Abbreviations: *CEA*, carcinoembryonic antigen; *EpCAM*, epithelial cell adhesion molecule; *FRα*, folate receptor alpha; *αvβ6*, integrin αvβ6; *uPAR*, urokinase-type plasminogen activator receptor; *CAIX*, carbonic anhydrase 9; *Collagen XVII*, collagen type XVII alpha 1 chain; *GLUT-1*, glucose transporter 1; *GPR87*, G protein-coupled receptor 87; *TMPRSS4*, transmembrane protease serine 4

tool and can therefore be beneficial to discover new targets, but cannot replace current validation tools as IHC. A possible explanation for the difference between the results of the data-driven selection and IHC could be that the data-driven selection platform uses RNA expression of genes, and this does not always correspond to the protein expression of genes.

Not only suitable targets for fluorescence imaging in invasive ADC and SCC were assessed, but also for AIS. Significantly overexpressed targets in AIS are CEA, CAIX, and collagen XVII. Preferably, a target is used that is overexpressed in all types of lung cancer (ADC, SCC, and AIS); therefore, CAIX and collagen XVII are the most favorable targets, given the results in ADC (including early stage) and SCC. CEA and CAIX had similar expression in ADC and AIS; EpCAM was not expressed in AIS samples, while 60% of ADC samples had high expression of EpCAM. This might be due to the fact that EpCAM is involved in tumor initiation and tumor cell invasion [29]. A difference in target expression between carcinoma *in situ* and invasive carcinoma has been described previously [30].

NSCLC is genetically and phenotypically known to be highly heterogenic, both across individuals as well as within individual tumors [31, 32]. Therefore, it is not remarkable that none of the targets had a perfect score concerning higher expression in the tumor compared to normal lung tissue. To counteract this heterogeneity, we expect the long-term clinical application of this technology to include a tumor-tailored selection of different targets resulting in the use of multiple fluorescent agents in one patient.

OTL-38, targeting FRα, is the only tumor-targeting agent of which clinical results in NSCLC patients have been published. Intraoperative tumor detection using OTL-38 was successful in 81% of the ADC and 54% of the SCC. During back table imaging (ex vivo), fluorescence imaging with OTL-38 was able to visualize 100% of the ADC and 69% of the SCC [33, 34]. Several studies have evaluated the expression of FRα in lung cancer. FRα was found to be expressed in 72–87% of the ADC and in 13–57% of the SCC [35, 36]. Our study report supports these results in ADC and corroborates the finding of high variability in the expression of FRα in SCC tumor tissue. The IHC staining rates for FRα are lower as compared with fluorescence imaging rates with OTL-38 in tumor tissue. There is no explanation for this difference. There has been suggested that IHC is only able to detect large differences in amounts of FRα molecules, while a smaller difference could be enough to discriminate normal from tumor tissue using fluorescence imaging [37]. However, this needs to be further confirmed with subsequent research.

Roughly, there are two methods for fluorescent agent administration: locally (e.g., topically and inhalation) or systemic (intravenously). CT-guided intraparenchymal injection of indocyanine green into lung tumors has been performed successfully in the past [38]. However, this is an invasive method and carries the risk for several complications, such as pulmonary hemorrhage or pneumothorax. In addition, this method can only be used to intraoperatively localize known tumors, as it cannot identify additional lesions. Inhalation of non-targeted contrast agents has been evaluated, but not yet in a clinical setting

[39]. At present, most clinical and preclinical studies use intravenous administration of fluorescent agents. To use this administration route, it is important that there is blood flow to the tumor. Intravenous administration appears to be feasible in lung tumors, since it has been shown that these tumors even have increased blood flow [40]. This is supported by the observation that tumors were visible after the administration of the agent OTL-38 intravenously.

Currently, there are clinical trials recruiting lung cancer patients scheduled for resection aided by fluorescence-guided surgery. Among the fluorescent agents used in these studies are panitumumab-IRDye800 and IRDye800CW-nimotuzumab, both targeting the epidermal growth factor receptor (EGFR) [41, 42]. EGFR is upregulated in most cancer types and is a popular target for fluorescence-guided surgery, because of the clinical availability of EGFR antibodies. According to our search with the data-driven selection platform, the expression of EGFR in NSCLC is lower compared to our selected targets with the platform. However, we did show that IHC scorings in these targets were not always correlated to what was expected with the search. Another study, recruiting patients with different tumor types, including NSCLC, investigates ONM-100, a pH-activable fluorescent agent that does not bind to a target, but exploits the acid microenvironment of solid tumors [43]. Moreover, SGM-101, targeting CEA, is being investigated not only in primary lung cancer, but also in colorectal lung metastases [44, 45]. A common endpoint in all these studies is the tumor-to-background ratio (TBR). This ratio of the mean fluorescence intensity of the tumor and the surrounding tissue (background) is the most used semi-quantitative parameter in fluorescence-guided surgery. Although immunohistochemistry serves as an indicator for potential clinically useful fluorescent agents, its correlation with the TBR is unknown.

Targets suitable for ADC, SCC, and AIS are CAIX and collagen XVII. A fluorescent-labeled contrast agent, ^{111}In-DOTA-girentuximab-IRDye800CW, targeting CAIX, has already been clinically investigated and shown to be safe and able to detect all CAIX-positive clear cell renal cell carcinomas [46]. Although EpCAM does not seem to be an optimal target for AIS, it is still a promising target for NSCLC in general. Firstly, it is suitable for ADC and SCC, and these are the most common subtypes of NSCLC. Moreover, a fluorescence-labeled contrast agent VB5-845D-800CW, a fluorescent agent targeting EpCAM, is currently being tested in healthy volunteers and colorectal cancer patients [47]. The fact that these agents targeting CAIX and EpCAM have already been tested and proven to be safe in patients could facilitate a clinical trial using these agents for the visualization of NSCLC in the short term.

Conclusion

Data-driven selection has proven to be a valuable tool for the identification of novel targets in NSCLC. After histological evaluation, EpCAM, CAIX, and collagen XVII have been shown to be the most promising targets for NSCLC. However, the clinical potential of these targets for fluorescence-guided surgery has to be confirmed in clinical trials.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11307-022-01791-5.

Declarations

Ethical Approval "All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards."

Conflict of Interest The authors declare no competing interests.

References

1. Global Burden of Disease Cancer Collaboration, Fitzmaurice C, Abate D, Abbasi N, Abbastabar H, Abd-Allah F, Abdel-Rahman O, Abdelalim A, Abdoli A, Abdollahpour I et al (2019) Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 cancer groups, 1990 to 2017: a systematic analysis for the global burden of disease study. JAMA Oncol 5(12):1749–1768
2. National Lung Screening Trial Research T, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, Gareen IF, Gatsonis C, Marcus PM et al (2011) Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med 365(5):395–409
3. Walter JE, Heuvelmans MA, de Jong PA, Vliegenthart R, van Ooijen PMA, Peters RB, Ten Haaf K, Yousaf-Khan U, van der Aalst CM, de Bock GH et al (2016) Occurrence and lung cancer probability of new solid nodules at incidence screening with low-dose CT: analysis of data from the randomised, controlled NELSON trial. Lancet Oncol 17(7):907–916
4. de Koning HJ, van der Aalst CM, de Jong PA, Scholten ET, Nackaerts K, Heuvelmans MA, Lammers JJ, Weenink C, Yousaf-Khan U, Horeweg N et al (2020) Reduced lung-cancer mortality with volume CT screening in a randomized trial. N Engl J Med 382(6):503–513
5. Yamato Y, Tsuchida M, Watanabe T, Aoki T, Koizumi N, Umezu H, Hayashi J (2001) Early results of a prospective study of limited resection for bronchioloalveolar adenocarcinoma of the lung. Ann Thorac Surg 71(3):971–974
6. Chen T, Luo J, Wang R, Gu H, Gu Y, Huang Q, Wang Y, Zheng J, Yang Y, Zhao H (2018) Prognosis of limited resection versus lobectomy in elderly patients with invasive lung adenocarcinoma with tumor size less than or equal to 2 cm. J Thorac Dis 10(4):2231–2239
7. Yoshizawa A, Motoi N, Riely GJ, Sima CS, Gerald WL, Kris MG, Park BJ, Rusch VW, Travis WD (2011) Impact of proposed IASLC/ATS/ERS classification of lung adenocarcinoma: prognostic subgroups and implications for further revision of staging based on analysis of 514 stage I cases. Mod Pathol 24(5):653–664

8. Cerfolio RJ, Bryant AS (2008) Is palpation of the nonresected pulmonary lobe(s) required for patients with non-small cell lung cancer? A prospective study. J Thorac Cardiovasc Surg 135(2):261–268
9. Ellis MC, Hessman CJ, Weerasinghe R, Schipper PH, Vetto JT (2011) Comparison of pulmonary nodule detection rates between preoperative CT imaging and intraoperative lung palpation. Am J Surg 201(5):619–622
10. Orosco RK, Tapia VJ, Califano JA, Clary B, Cohen EEW, Kane C, Lippman SM, Messer K, Molinolo A, Murphy JD et al (2018) Positive surgical margins in the 10 most common solid cancers. Sci Rep 8(1):5686
11. Osarogiagbon RU, Lin CC, Smeltzer MP, Jemal A (2016) Prevalence, prognostic implications, and survival modulators of incompletely resected non-small cell lung cancer in the U.S. national cancer data base. J Thorac Oncol 11(1):e5-16
12. Wind J, Smit EJ, Senan S, Eerenberg JP (2007) Residual disease at the bronchial stump after curative resection for lung cancer. Eur J Cardiothorac Surg 32(1):29–34
13. Riquet M, Achour K, Foucault C, Le Pimpec BF, Dujon A, Cazes A (2010) Microscopic residual disease after resection for lung cancer: a multifaceted but poor factor of prognosis. Ann Thorac Surg 89(3):870–875
14. al-Kattan K, Sepsas E, Fountain SW, Townsend ER (1997) Disease recurrence after resection for stage I lung cancer. Eur J Cardiothorac Surg 12(3):380–4
15. Roth K, Nilsen TI, Hatlen E, Sorensen KS, Hole T, Haaverstad R (2008) Predictors of long time survival after lung cancer surgery: a retrospective cohort study. BMC Pulm Med 8:22
16. Uramoto H, Tanaka F (2014) Recurrence after surgery in patients with NSCLC. Transl Lung Cancer Res 3(4):242–249
17. Vahrmeijer AL, Hutteman M, van der Vorst JR, van de Velde CJ, Frangioni JV (2013) Image-guided cancer surgery using near-infrared fluorescence. Nat Rev Clin Oncol 10(9):507–518
18. Frangioni JV (2008) New technologies for human cancer imaging. J Clin Oncol 26(24):4012–4021
19. van Oosten M, Crane LM, Bart J, van Leeuwen FW, van Dam GM (2011) Selecting potential targetable biomarkers for imaging purposes in colorectal cancer using target selection criteria (TASC): a novel target identification tool. Transl Oncol 4(2):71–82
20. Cancer Genome Atlas Research N (2012) Comprehensive genomic characterization of squamous cell lung cancers. Nature 489(7417):519–25
21. Cancer Genome Atlas Research N (2014) Comprehensive molecular profiling of lung adenocarcinoma. Nature 511(7511):543–50
22. Consortium GT (2013) The genotype-tissue expression (GTEx) project. Nat Genet 45(6):580–585
23. Abrams ZB, Johnson TS, Huang K, Payne PRO, Coombes K (2019) A protocol to evaluate RNA sequencing normalization methods. BMC Bioinform 20(Suppl 24):679
24. Bausch-Fluck D, Hofmann A, Bock T, Frei AP, Cerciello F, Jacobs A, Moest H, Omasits U, Gundry RL, Yoon C et al (2015) A mass spectrometric-derived cell surface protein atlas. PLoS One 10(3):e0121314
25. Bausch-Fluck D, Goldmann U, Muller S, van Oostrum M, Muller M, Schubert OT, Wollscheid B (2018) The in silico human surfaceome. Proc Natl Acad Sci U S A 115(46):E10988–E10997
26. Fonseca AL, da Silva VL, da Fonsêca MM, Meira IT, da Silva TE, Kroll JE, Ribeiro-Dos-Santos AM, Freitas CR, Furtado R, de Souza JE et al (2016) Bioinformatics analysis of the human surfaceome reveals new targets for a variety of tumor types. Int J Genomics 8346198
27. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015) Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43(7):e47
28. Zhang JZ, Ismail-Beigi F (1998) Activation of Glut1 glucose transporter in human erythrocytes. Arch Biochem Biophys 356(1):86–92
29. Imrich S, Hachmeister M, Gires O (2012) EpCAM and its potential role in tumor-initiating cells. Cell Adh Migr 6(1):30–38
30. Lonergan KM, Chari R, Coe BP, Wilson IM, Tsao MS, Ng RT, Macaulay C, Lam S, Lam WL (2010) Transcriptome profiles of carcinoma-in-situ and invasive non-small cell lung cancer as revealed by SAGE. PLoS ONE 5(2):e9162
31. Zito Marino F, Bianco R, Accardo M, Ronchi A, Cozzolino I, Morgillo F, Rossi G, Franco R (2019) Molecular heterogeneity in lung cancer: from mechanisms of origin to clinical implications. Int J Med Sci 16(7):981–989
32. Carnio S, Novello S, Papotti M, Loiacono M, Scagliotti GV (2013) Prognostic and predictive biomarkers in early stage non-small cell lung cancer: tumor based approaches including gene signatures. Transl Lung Cancer Res 2(5):372–381
33. Predina JD, Newton AD, Xia L, Corbett C, Connolly C, Shin M, Sulyok LF, Litzky L, Deshpande C, Nie S et al (2018) An open label trial of folate receptor-targeted intraoperative molecular imaging to localize pulmonary squamous cell carcinomas. Oncotarget 9(17):13517–13529
34. Predina JD, Newton AD, Keating J, Dunbar A, Connolly C, Baldassari M, Mizelle J, Xia L, Deshpande C, Kucharczuk J et al (2018) A phase I clinical trial of targeted intraoperative molecular imaging for pulmonary adenocarcinomas. Ann Thorac Surg 105(3):901–908
35. O'Shannessy DJ, Yu G, Smale R, Fu YS, Singhal S, Thiel RP, Somers EB, Vachani A (2012) Folate receptor alpha expression in lung cancer: diagnostic and prognostic significance. Oncotarget 3(4):414–425
36. Cagle PT, Zhai QJ, Murphy L, Low PS (2013) Folate receptor in adenocarcinoma and squamous cell carcinoma of the lung: potential target for folate-linked therapeutic agents. Arch Pathol Lab Med 137(2):241–244
37. Boogerd LS, van der Valk MJ, Boonstra MC, Prevoo HA, Hilling DE, van de Velde CJ, Sier CF, Fariña Sarasqueta A, Vahrmeijer AL (2018) Biomarker expression in rectal cancer tissue before and after neoadjuvant therapy. Onco Targets Ther 11:1655–1664
38. Li X, Xu K, Cen R, Deng J, Hao Z, Liu J, Takizawa H, Ng CSH, Marulli G, Kim MP et al (2021) Preoperative computer tomography-guided indocyanine green injection is associated with successful localization of small pulmonary nodules. Transl Lung Cancer Res 10(5):2229–2236
39. Quan YH, Oh CH, Jung D, Lim JY, Choi BH, Rho J, Choi Y, Han KN, Kim BM, Kim C et al (2020) Evaluation of intraoperative near-infrared fluorescence visualization of the lung tumor margin with indocyanine green inhalation. JAMA Surg 155(8):732–740
40. Huang C, Liang J, Lei X, Xu X, Xiao Z, Luo L (2019) Diagnostic performance of perfusion computed tomography for differentiating lung cancer from benign lesions: a meta-analysis. Med Sci Monit 25:3485–3494
41. Bharadwaj S (n.d.) Evaluation of IRDye800CW-nimotuzumab in lung cancer surgery
42. Rosenthal E (n.d.) A Phase I/II study evaluating the safety and pharmacokinetics of panitumumab-IRDye800 as an optical imaging agent to detect lung cancer during surgical procedures
43. A Phase 2a, single-dose, open-label study to evaluate diagnostic performance, safety & timing of postdose imaging of ONM-100, an intraoperative fluorescence imaging agent for the detection of cancer, in patients with solid tumors undergoing routine surgery [Internet]. Available from: https://www.clinicaltrials.gov/ct2/show/record/NCT03735680

44. Intraoperative imaging of pulmonary nodules by SGM-101 [Internet]. Available from: https://www.clinicaltrials.gov/ct2/show/NCT04315467
45. Intraoperative Molecular Imaging Of Pulmonary Nodules By SGM-101, A Fluorochrome-labeled anti-carcino-embryonic antigen (CEA) monoclonal antibody [Internet]. Available from: https://www.clinicaltrials.gov/ct2/show/NCT04315467
46. Hekman MC, Rijpkema M, Muselaers CH, Oosterwijk E, Hulsbergen-Van de Kaa CA, Boerman OC, Oyen WJ, Langenhuijsen JF, Mulders PF (2018) Tumor-targeted dual-modality imaging to improve intraoperative visualization of clear cell renal cell carcinoma: a first in man study. Theranostics 8(8):2161–2170
47. A phase I study assessing the safety and performance of VB5-845D-800CW, an anti-Epcam fluorescent agent, for the intraoperative detection of gastrointestinal cancer [Internet]. Available from: https://www.trialregister.nl/trial/7363