EINDHOVEN UNIVERSITY OF TECHNOLOGY

Liquid biopsy-based decision support algorithms for diagnosis and subtyping of lung cancer

Citation for published version (APA):
Visser, E., Genet, S. A. A. M., de Kock, R. P. P. A., van den Borne, B. E. E. M., Youssef-El Soud, M., Belderbos, H. N. A., Stege, G., de Saegher, M. E. A., van 't Westeinde, S. C., Brunsveld, L., Broeren, M. A. C., van de Kerkhof, D., Deiman, B. A. L. M., Eduati, F., & Scharnhorst, V. (2023). Liquid biopsy-based decision support algorithms for diagnosis and subtyping of lung cancer. *Lung Cancer*, *178*, 28-36. https://doi.org/10.1016/j.lungcan.2023.01.014

Document license:
CC BY

DOI:
10.1016/j.lungcan.2023.01.014

Document status and date:
Published: 01/04/2023

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

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

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

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

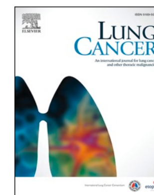
Download date: 05. Oct. 2023

Lung Cancer 178 (2023) 28–36

Contents lists available at ScienceDirect

Lung Cancer

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



Liquid biopsy-based decision support algorithms for diagnosis and subtyping of lung cancer

Esther Visser [a,b,c,d,*,1], Sylvia A.A.M. Genet [a,b,d,e,1], Remco P.P.A. de Kock [a,b,c,d], Ben E.E.M. van den Borne [b], Maggy Youssef-El Soud [c], Huub N.A. Belderbos [f], Gerben Stege [g], Marleen E.A. de Saegher [h], Susan C. van 't Westeinde [i], Luc Brunsveld [a,d,e], Maarten A.C. Broeren [a,c,d], Daan van de Kerkhof [a,b], Birgit A.L.M. Deiman [b,d], Federica Eduati [a,d,e,j,2], Volkher Scharnhorst [a,b,d,e,j,2]

[a] *Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, the Netherlands*
[b] *Catharina Hospital Eindhoven, Eindhoven, the Netherlands*
[c] *Máxima Medical Center, Eindhoven/Veldhoven, the Netherlands*
[d] *Expert Center Clinical Chemistry Eindhoven, Eindhoven, the Netherlands*
[e] *Institute for Complex Molecular Systems, Eindhoven University of Technology, Eindhoven, the Netherlands*
[f] *Amphia Hospital, Breda, the Netherlands*
[g] *Anna Hospital, Geldrop, the Netherlands*
[h] *Sint Jans Gasthuis, Weert, the Netherlands*
[i] *Maasstad Hospital, Rotterdam, the Netherlands*
[j] *Eindhoven Artificial Intelligence Systems Institute, Eindhoven University of Technology, Eindhoven, the Netherlands*

A R T I C L E I N F O

Keywords:
Protein tumor markers
circulating tumor DNA
Lung cancer
Decision–support algorithm
Liquid biopsy

A B S T R A C T

Objectives: Pathologic subtyping of tissue biopsies is the gold standard for the diagnosis of lung cancer (LC), which could be complicated in cases of e.g. inconclusive tissue biopsies or unreachable tumors. The diagnosis of LC could be supported in a minimally invasive manner using protein tumor markers (TMs) and circulating tumor DNA (ctDNA) measured in liquid biopsies (LBx). This study evaluates the performance of LBx-based decision-support algorithms for the diagnosis of LC and subtyping into small– and non-small-cell lung cancer (SCLC and NSCLC) aiming to directly impact clinical practice.
Materials and Methods: In this multicenter prospective study (NL9146), eight protein TMs (CA125, CA15.3, CEA, CYFRA 21-1, HE4, NSE, proGRP and SCCA) and ctDNA mutations in *EGFR, KRAS* and *BRAF* were analyzed in blood of 1096 patients suspected of LC. The performance of individual and combined TMs to identify LC, NSCLC or SCLC was established by evaluating logistic regression models at pre-specified positive predictive values (PPV) of ≥95% or ≥98%. The most informative protein TMs included in the multi-parametric models were selected by recursive feature elimination.
Results: Single TMs could identify LC, NSCLC and SCLC patients with 46%, 25% and 40% sensitivity, respectively, at pre-specified PPVs. Multi-parametric models combining TMs and ctDNA significantly improved sensitivities to 65%, 67% and 50%, respectively.
Conclusion: In patients suspected of LC, the LBx–based decision–support algorithms allowed identification of about two-thirds of all LC and NSCLC patients and half of SCLC patients. These models therefore show clinical value and may support LC diagnostics, especially in patients for whom pathologic subtyping is impossible or incomplete.

* Corresponding author at: Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, the Netherlands.
E-mail address: e.visser@tue.nl (E. Visser).
[1] Shared first author.
[2] Shared last author.

https://doi.org/10.1016/j.lungcan.2023.01.014
Received 4 November 2022; Received in revised form 11 January 2023; Accepted 29 January 2023
Available online 1 February 2023
0169-5002/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

In the past two decades, the development of new lung cancer (LC) therapies based on precision medicine strategies, i.e. targeted therapy and immunotherapy, has led to a significantly improved progression-free and overall survival compared to conventional therapies [1,2]. Currently, treatment is chosen based on the histological subtype (non-small-cell lung cancer (NSCLC) or small-cell lung cancer (SCLC)), tumor stage, presence of oncogenic driver mutations and expression of immune checkpoint proteins [3]. Radiological imaging techniques such as positron emission tomography (PET) scans or computed tomography (CT) scans are mainly used to define the tumor stage. Definitive diagnosis including subtype and genotype of LC is based on examination of tissue or cytology specimens collected by invasive procedures such as bronchoscopy, endobronchial ultrasound-guided needle aspirations (EBUS/TBNA) or CT/ultrasound–guided transthoracic biopsy [3–5]. Unfortunately, these biopsies are not always adequate for histological, cytological and/or molecular profiling, necessitating repeated biopsies [6–8]. Moreover, obtaining tissue *via* invasive procedures may be challenging or impossible in fragile patients or patients with small or unreachable tumors [6].

Additional information on the presence and characteristics of LC can be obtained *via* liquid biopsy (LBx), in which tumor markers (TMs) such as circulating tumor DNA (ctDNA) and proteins in the blood are analyzed [9]. Analysis of plasma-derived ctDNA was shown to enable detection of driver mutations in e.g. *EGFR, KRAS* and *BRAF,* which are frequently occurring and targetable mutations in LC, allowing for the use of targeted therapies [10–12]. Serum protein TMs such as carcinoembryonic antigen (CEA), cytokeratin-19 fragment 21-1 (CYFRA 21-1) and human epididymis protein 4 (HE4) were shown to aid in the identification of LC [13,14]. Moreover, protein TMs were shown to be promising for differentiation between the histological subtypes SCLC and NSCLC, as neuron-specific enolase (NSE) and progastrin-releasing peptide (proGRP) were shown to be elevated in SCLC, whereas carbohydrate antigen 15.3 (CA15.3), cancer antigen 125 (CA125), CEA, CYFRA 21-1 and squamous cell carcinoma antigen (SCCA) were shown to be associated with NSCLC [14–16].

Compared to individual protein TMs, combining multiple TMs has previously been shown to improve the performance to identify and subtype LC [15–19]. However, these studies often combined TMs using standard cut-off values or determined optimal performance using a trade-off between sensitivity and specificity, often leading to insufficient performance for clinical use. Moreover, a few studies suggested that combined assessment of protein and ctDNA TMs might further improve performance to diagnose LC [20,21]. Therefore, this study describes the performance of protein– and ctDNA–based decision support algorithms for the identification of LC and subtyping of SCLC and NSCLC, focusing on direct clinical applicability.

2. Materials and methods

2.1. Study design

The data for this investigation was obtained in a multicenter study (Longmerker studie, NL9146) approved by the Medical Research Ethics Committees United (NL58985.100.16). In six hospitals in the Netherlands, 1096 patients suspected of LC were prospectively included by their lung physician between June 2017 and February 2022. After obtaining written informed consent, blood samples were retrieved before a final diagnosis had been given and any treatment had been initiated. Diagnosis of LC and subtyping was performed according to Dutch guidelines [22,23]. Staging was performed according to the 8th edition of TNM Classification for LC [24].

Patients treated for LC before enrollment (9%), with exception of patients previously treated with curative intent, and patients with another primary tumor (15%) were excluded from analysis (Fig. 1). Also, patients with incomplete laboratory data (3%), without pathological diagnosis (9%) or other LC subtypes than NSCLC or SCLC (2%)

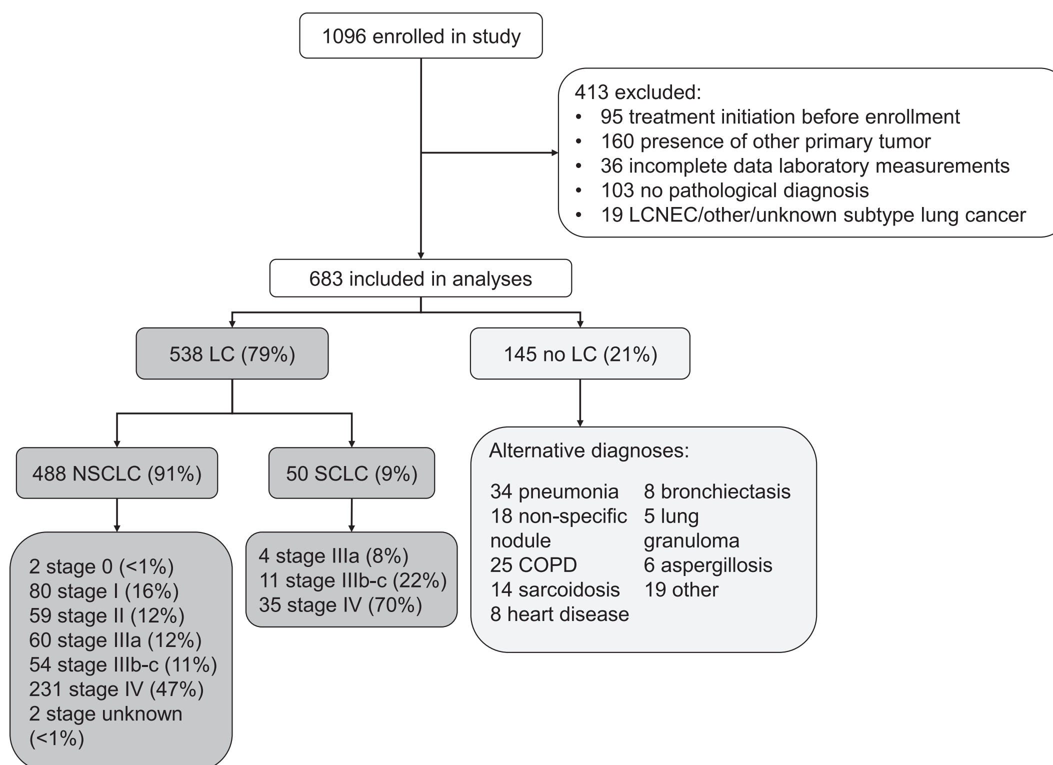


Fig. 1. Overview of the study population. LC = lung cancer, NSCLC = non-small-cell lung cancer; SCLC = small-cell lung cancer; COPD = Chronic Obstructive Pulmonary Disease, LCNEC = Large cell neuroendocrine carcinoma of the lung.

were excluded.

2.2. Sample processing and analysis

Whole blood was collected and processed to obtain plasma for analysis of cfDNA and serum for analysis of protein TMs, as described previously [25,26]. All samples were analyzed in the clinical laboratory of the Catharina Hospital Eindhoven. Wild-type *KRAS*, used as a measure for cell-free DNA (cfDNA) concentrations, and driver mutations in *EGFR*, *KRAS* and *BRAF* were determined on ctDNA using droplet digital PCR (ddPCR), as described previously (QX200 system, Bio-Rad Laboratories, Hercules, CA) [25,27,28]. The serum levels of CA15.3, CA125, CEA, CYFRA 21-1, HE4, NSE, proGRP and SCCA were measured with commercially available electrochemiluminescent assays (Cobas e602 or Cobas Pro e801, Roche Diagnostics, Rotkreuz, Switzerland). The NSE concentrations were corrected for hemolysis [26].

2.3. Model fitting and evaluation

To determine the performance of liquid-biopsy for diagnosis of LC and identification of NSCLC and SCLC, multivariate logistic regression was applied. The models were trained and evaluated using patients with pathology-proven NSCLC or SCLC and patients for whom LC was excluded (Fig. 1). For diagnosis of LC, a model to distinguish LC from no LC was fitted. For identification of NSCLC, another model was fitted to separate NSCLC from no NSCLC, i.e. patients with SCLC and no LC. For identification of SCLC, a model was trained to separate SCLC from no SCLC, comprising patients with NSCLC and no LC.

Models used either individual or multiple protein and DNA (cfDNA concentration and presence of mutation (ctDNA)) TMs. The models were all adjusted for the covariates age and sex by using these covariates as independent variables in the models. The protein TMs and cfDNA concentrations were log10-transformed. The logistic regression models were trained and validated using 200 repetitions of 5-fold stratified cross-validation, resulting in 1000 models where 80 % of the data was used for training and 20 % for validation of the model. Each continuous variable was standardized by subtracting the mean and dividing by the standard deviation of that variable in the training set.

The most informative combination of protein TMs, unadjusted for covariates, was determined by recursive feature elimination (RFE) using repeated stratified K-fold cross-validation. For each cross-validation fold, a logistic regression model was trained starting with the total of eight protein TMs. The least informative TM, based on the coefficient in the logistic regression model was removed and the model was fitted again on the remaining seven TMs. This procedure was repeated up until only one TM remained. The most informative combination of protein TMs was chosen based on the maximum performance (based on the validation data) while using a minimum number of TMs. This combination of protein TMs was thereafter adjusted for the covariates age and sex and combined with the DNA TMs.

2.4. Model evaluation

Overall performance was evaluated using the Area Under the Curve (AUC) of Receiver Operating Characteristic (ROC) curves. The average ROC curves of the cross-validation folds per model were computed by vertical averaging, hence average true positive rates are computed at fixed false positive rates [29]. Additionally, the performance metrics sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were computed using standard formulas. These performance metrics were evaluated using the validation sets, thus patient data that was not used for training of the model in that cross-validation fold. In the end, the median performances of the 1000 validation sets were computed, together with the interquartile range (IQR; 25th-75th percentile) as a measure for the distribution.

To determine potential clinical applicability of the TMs in identification of LC, NSCLC or SCLC with high certainty, the performance metrics of the validation sets were evaluated at pre-specified PPVs for the training set: 98% PPV for LC and 95% PPV for confirmation of NSCLC or SCLC. In addition, the performance was evaluated at 100% NPV for potential exclusion of LC, since it is important to not misclassify a LC patient who would then not get proper treatment. The final classifications of individual patients were determined by the majority vote (>50%) of classifications considering all the cross-validation models with the patient in the validation set.

2.5. Statistical analysis

TM concentrations and age were compared by non-parametric tests (Kruskal-Wallis or two-sided Mann-Whitney-U tests ($\alpha = 0.05$)). Categorical variables were compared by chi-square tests ($\alpha = 0.05$). The performances of the individual and combined TMs models were compared by Mann-Whitney-U-tests (one-sided, $\alpha = 0.05$). The data were represented as number (frequency (%)) or median (IQR; 25th percentile – 75th percentile). The analyses were performed using Python (version 3.8.8), using scikit-learn (version 0.24.1) and SciPy (version 1.8.0).

2.6. Data availability

The data generated in this study are not publicly available due information that could compromise study participant privacy. Data may be made available for substantiated research projects. The code is available at: https://github.com/SysBioOncology/Lungmarkerstudy_diagnosticalgorithm.

3. Results

3.1. Patient characteristics

In total, 1096 patients were included in this study, of whom 683 patients met the inclusion criteria for analysis. 21% of the included subjects did not have LC and 79% were diagnosed with LC (Fig. 1). 91% of LC patients were diagnosed with NSCLC, of whom 58% patients had advanced stages (IIIb-IV) and 41% patients had earlier stages (I-IIIa). Moreover, 9% patients were diagnosed with SCLC, all with stage III-IV.

Some different characteristics were observed between the patient group without LC and with LC. A higher frequency of males was present in the group without LC (63%) compared to the group with LC (NSCLC: 51%, SCLC: 54%) ($p = 0.020$) (Supplemental Table 1). In addition, individuals without LC were slightly younger (66, IQR: 59–74) compared to LC patients (NSCLC: 69, IQR: 62–74; SCLC: 70, IQR: 63–73) ($p = 0.018$). Mutations were detected by ctDNA-ddPCR in 2% of SCLC patients and in 19% of NSCLC patients, while no mutations were found in patients without LC. 11% of NSCLC patients had a targetable mutation in *EGFR, BRAF* or *KRAS G12C*.

3.2. Concentrations of individual TMs

TM concentrations were increased in NSCLC and SCLC patients compared to individuals without LC (Fig. 2). CA15.3 and SCCA were significantly increased in NSCLC patients only ($p < 0.001$ and $p = 0.033$, respectively) and proGRP in SCLC patients only ($p < 0.001$). CA125, CEA, CYFRA 21-1, HE4, NSE and cfDNA were increased for both NSCLC (all $p < 0.001$) and SCLC (all $p < 0.001$). However, commonly used diagnostic cut-off values determined in previous studies [13,15,30] could not clearly separate patients with and without LC, since concentrations above the cut-off value are measured in individuals without LC and below the cut-off for patients with NSCLC and SCLC.

Significant concentration differences between NSCLC and SCLC were observed for all TMs, except for CEA ($p = 0.58$) (Fig. 2). CA125, HE4, NSE, proGRP and cfDNA were significantly higher in SCLC patients (p =

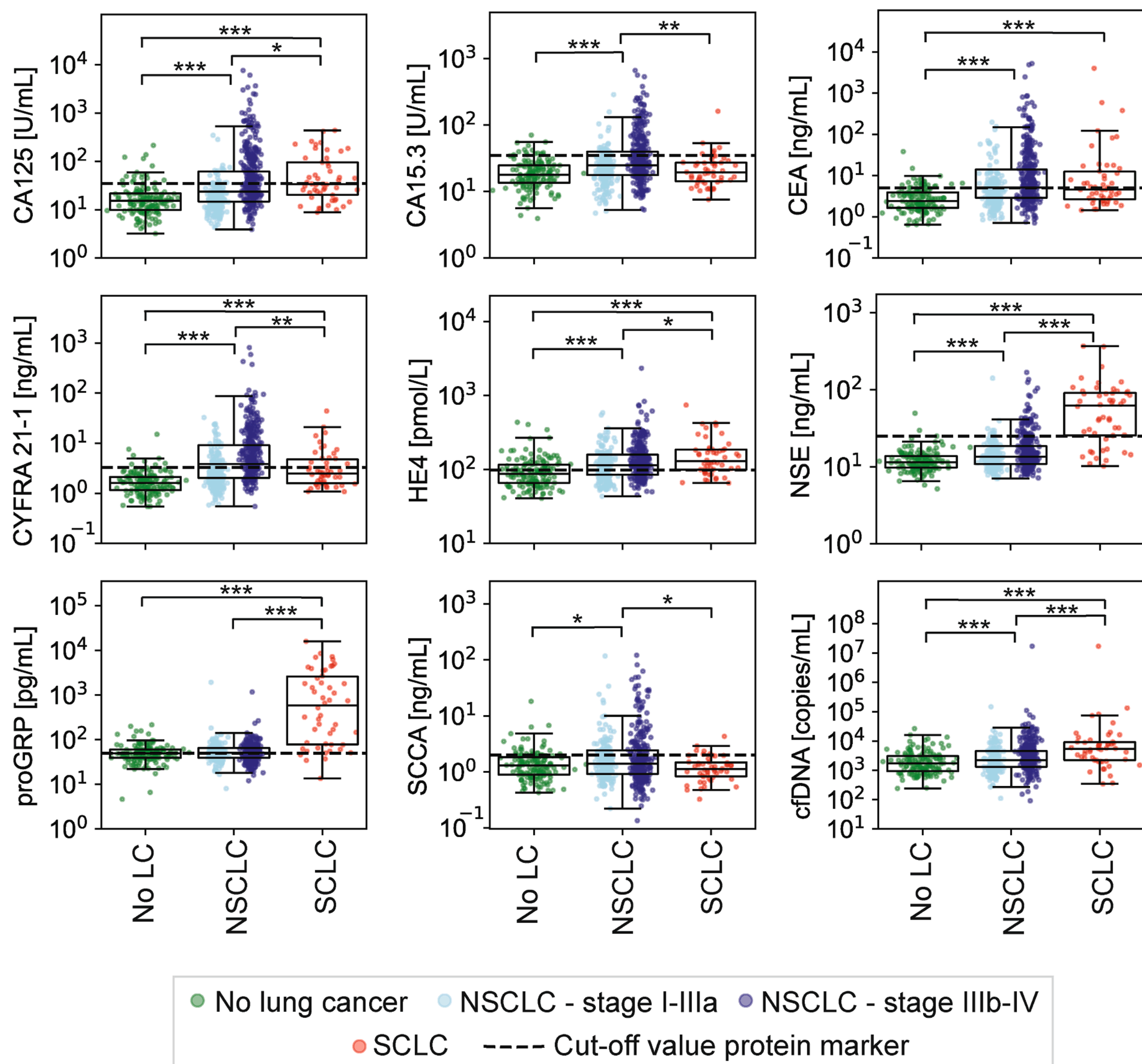


Fig. 2. Concentrations of protein TMs and cfDNA per subgroup. The boxplots and statistical tests include all stage I-IV NSCLC. The cut-off values for the protein TMs shown are: CA125: 35 U/mL; CA15.3: 35 U/mL; CEA: 5 ng/mL; CYFRA 21-1: 3.3 ng/mL; HE4: 97.6 pmol/L; NSE: 25 ng/mL; proGRP: 50 pg/mL; SCCA: 2 ng/mL [13,15,30]. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. NSCLC = non-small-cell lung cancer; SCLC = small-cell lung cancer; cfDNA = cell-free DNA, TMs = tumor markers.

0.046, $p = 0.035$, $p < 0.001$, $p < 0.001$ and $p < 0.001$, respectively) and CA15.3, CYFRA 21-1 and SCCA were elevated in NSCLC patients ($p = 0.0036$, $p = 0.0042$ and $p = 0.011$, respectively). Furthermore, NSCLC patients with early stages (I-IIIa) had lower concentrations than patients with late stage NSCLC (IIIb-IV) for all TMs (all $p < 0.001$), except for proGRP and SCCA ($p = 0.99$ and $p = 0.60$, respectively).

3.3. Detection of primary lung cancer

Of the individual TMs, CYFRA 21-1 had the best overall performance to distinguish LC patients from patients without LC (ROC-AUC = 0.83, IQR: 0.81–0.85), followed by CEA (ROC-AUC = 0.76, IQR: 0.73–0.78) (Fig. 3A, Supplemental Table 2). Moreover, when evaluating the performance at pre-specified PPV ≥ 98% for identification of LC, these two TMs achieved the highest sensitivities (CYFRA 21-1: 46%, IQR: 43–49%; CEA: 31%, IQR: 28–34%) and could therefore identify the largest subset of LC patients (Fig. 3A, Supplemental Table 2). As ctDNA mutations were only found in LC patients, these mutations could identify LC with PPV of 100% with a sensitivity similar to the frequency of mutations in the LC patients (19%, IQR: 17–21%). Evaluating the performance at NPV of 100% showed that exclusion of LC was not possible by individual TMs (Supplemental Table 3).

To determine the most informative combination of the eight protein TMs for identification of LC, the performance of models ranging from one to eight TMs was evaluated using recursive feature elimination (Fig. 4). The best TM combination was found to be CYFRA 21-1 and CEA with a sensitivity of 62% (IQR: 58–65%), specificity of 97% (IQR: 93–100%), NPV of 40% (IQR: 38–42%) and PPV of 98% (IQR: 97–100%) (Fig. 4A, Supplemental Table 4). Addition of more TMs to the model resulted in a significant decrease in sensitivity ($p < 0.001$), but changed the fraction of the patients with NSCLC and SCLC that could be correctly identified (Supplemental Fig. 1A). Here, addition of more SCLC related TMs (NSE and proGRP) resulted in an increase of correctly classified SCLC patients, but also a decrease in correctly classified NSCLC patients. Because the overall fraction of NSCLC patients is about tenfold higher than the SCLC patients, the decline in identified NSCLC patients resulted in lower overall sensitivity.

The combination of the CEA and CYFRA 21-1, corrected by age and

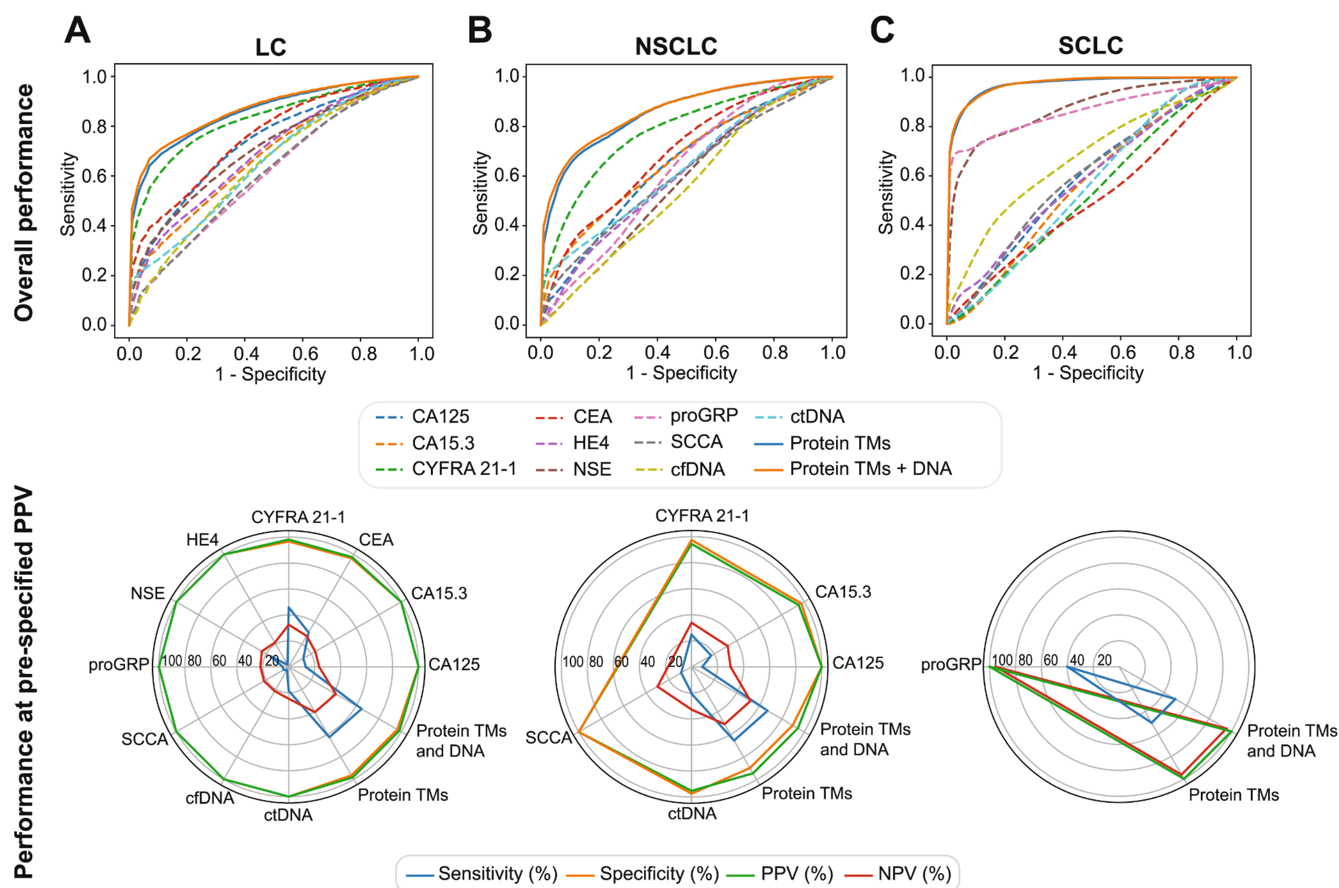


Fig. 3. Performance metrics of the validation sets of individual and multi-TM models, all adjusted for covariates age and sex. Overall performance is shown as ROC curves. Sensitivity, specificity, PPV and NPV were evaluated at pre-specified PPVs for the training set. Performances of individual TMs at pre-specified PPVs are only shown if this PPV could be achieved in all training sets. Median and interquartile ranges (IQR) performance metrics are shown in Supplemental Tables 2, 6 and 8. A) No LC versus LC, evaluated at PPV ≥ 98%. Protein TMs in the multi-marker model are CEA and CYFRA 21-1. B) Absence of NSCLC versus NSCLC, evaluated at PPV ≥ 95%. Protein TMs used in the multi-marker models are CEA, CYFRA 21-1, NSE and proGRP. C) Absence of SCLC versus SCLC, evaluated at PPV ≥ 95%. Protein TMs used in the multi-marker models are CA125, CA15.3, CYFRA 21-1, NSE and proGRP. Due to limited sample size, the multi-marker models of SCLC could not achieve the pre-specified PPV for all training sets. The performances shown include only the validation sets where corresponding training sets met this PPV requirement. LC = lung cancer, TMs = tumor markers, NSCLC = non-small-cell lung cancer; SCLC = small-cell lung cancer; cfDNA = cell-free DNA, ctDNA = circulating tumor DNA, ROC = Receiver Operating Characteristics, AUC = Area under the curve.

sex, increased the ROC-AUC (from 0.83, IQR: 0.81–0.85 to 0.86, IQR: 0.84–0.88, $p < 0.001$) and sensitivity (from 46%, IQR: 43–49% to 63%, IQR: 59–65%, $p < 0.001$) significantly compared to the individual TM CYFRA 21-1, while the PPV did not change significantly (from 98%, IQR: 97–100% to 98%, IQR: 97–99%, $p = 0.74$) (Fig. 3A, Supplemental Table 2). Addition of DNA TMs to the protein TMs model improved ROC-AUC (from 0.86, IQR: 0.84–0.88 to 0.87, IQR: 0.85–0.89) and sensitivity (from 63%, IQR: 59–65% to 65%, IQR: 62–68%) marginally, yet significantly ($p < 0.001$ for both), with insignificant change of PPV (from 98%, IQR: 97–99% to 98%, IQR: 96–99%, $p = 0.37$). The coefficients of the multi-TM models show that higher concentrations of the TMs and a mutation in ctDNA increased the probability of LC (Supplemental Fig. 2). Combinations of the protein TMs did not allow for exclusion of LC with high NPV (Supplemental Table 5).

The multi-parametric model using CEA, CYFRA 21-1 and DNA TMs correctly classified 65% of the NSCLC patients and 48% of the SCLC patients and misclassified 5% of patients without LC as having LC (Fig. 5A). In the NSCLC group, a higher fraction of patients with stage IV disease (80%) was identified compared to stage I (23%), stage II (44%) stage IIIa (57%).

3.4. Identification of NSCLC

Of the individual TMs, CYFRA 21-1 could distinguish NSCLC patients from SCLC and LC-free individuals with the best overall performance (ROC-AUC = 0.78, IQR: 0.76–0.81) and highest sensitivity (25%, IQR: 20–30%) at pre-specified PPV ≥95% (Fig. 3B, Supplemental Table 6). Even though CEA had the second-best overall performance (ROC-AUC = 0.69, IQR: 0.66–0.72), this TM could not identify NSCLC with the required pre-specified PPV of ≥95% for the training sets. Also HE4, NSE, proGRP and cfDNA could not meet this criterium.

CYFRA 21-1 combined with CEA, proGRP and NSE was shown to be the most informative combination of protein TMs to identify NSCLC, due to the highest AUC (0.85, IQR: 0.83–0.87), sensitivity (63%, IQR: 60–66%), NPV (50%, IQR: 48–52%) and PPV (95%, IQR: 93–97%), with a specificity of 92% (IQR: 87–95%) (Fig. 4B, Supplemental Table 7). Addition of other protein TMs did not improve the performance for the overall population and the characteristics of the identified patients in terms of stage of disease and histological subtype were similar (Supplemental Fig. 1B).

Combining CYFRA 21-1, CEA, proGRP and NSE, adjusted for age and sex, significantly improved the ROC-AUC (from 0.78, IQR: 0.76–0.81 to 0.86, IQR: 0.84–0.88, $p < 0.001$) and sensitivity (from 25%, IQR: 20–30% to 65%, IQR: 62–68%, $p < 0.001$) compared to CYFRA 21-1 alone, while the PPV did not change significantly (from 94%, IQR: 92–100% to 94%, IQR: 93–97%, $p = 0.38$) (Fig. 3B, Supplemental Table 6). Adding DNA TMs to the multi-parametric model slightly increased the ROC-AUC (from 0.86, IQR: 0.84–0.88 to 0.87, IQR:

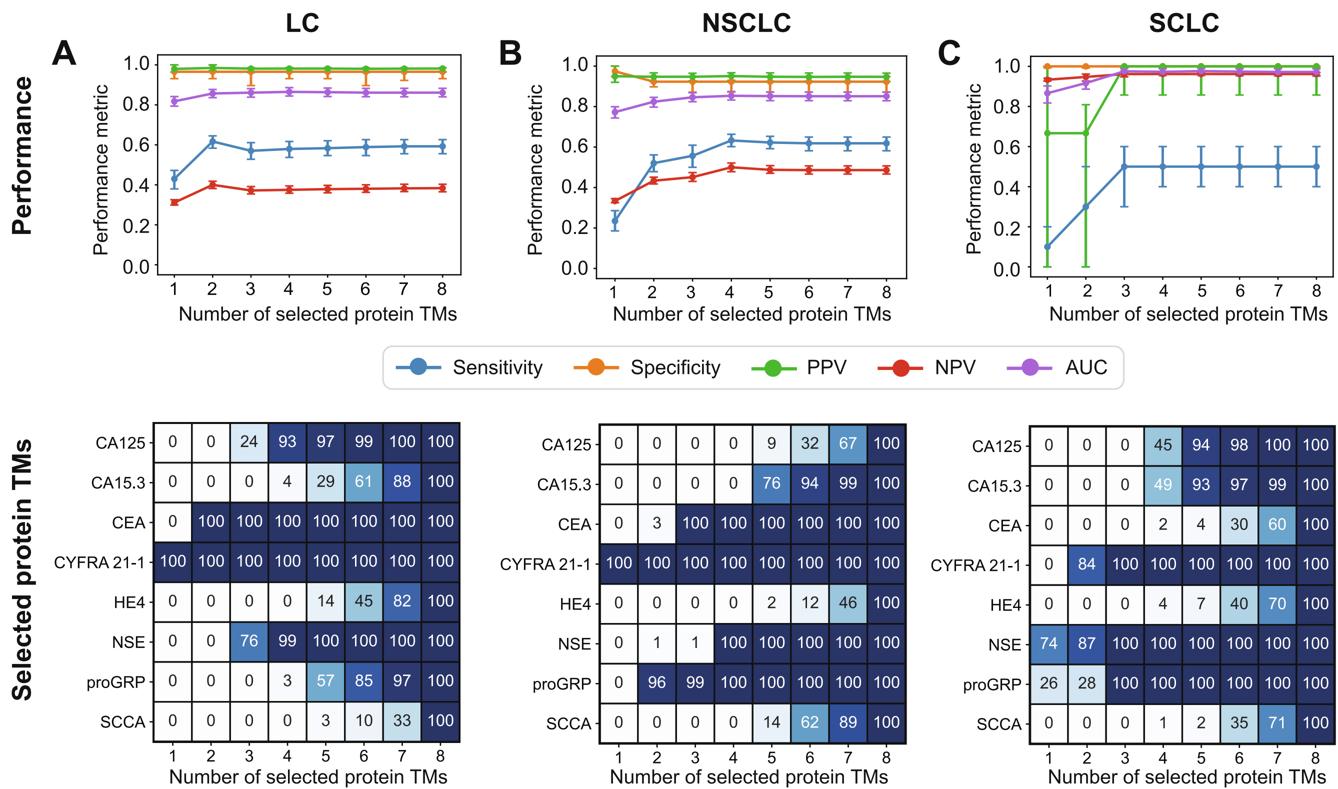


Fig. 4. Selection of most informative combination of protein TMs by recursive feature elimination. The performances per total number of selected protein TMs are shown as median (IQR). The selected protein TMs are shown as percentage of cross-validation folds for which the protein TMs are included in the models. A) No LC versus LC, evaluated at PPV ≥ 98% (Supplemental Table 4). B) Non-NSCLC versus NSCLC, evaluated at PPV ≥ 95% (Supplemental Table 7). C) Non-SCLC versus SCLC, evaluated at PPV ≥ 95%. The performance metrics were computed for validation sets where the corresponding training set could achieve PPV ≥ 95% (Supplemental Table 9). LC = lung cancer, NSCLC = non-small-cell lung cancer; SCLC = small-cell lung cancer; TMs = tumor markers, PPV = positive predictive value, NPV = negative predictive value, AUC = area under the receiver operating characteristics curve.

0.85–0.88, $p < 0.001$) and sensitivity (65%, IQR: 62–68% to 67%, IQR: 64–71%, $p < 0.001$) with insignificantly changed PPV (from 94%, IQR: 93–97% to 94%, IQR: 93–96%, $p = 0.12$). In these models, higher concentrations of CEA and CYFRA 21-1, together with a mutation in ctDNA, increase the probability of having NSCLC (Supplemental Fig. 2). Higher concentrations of NSE, proGRP and cfDNA decrease the probability of having NSCLC.

The model using CYFRA 21-1, CEA, NSE, proGRP and DNA to identify NSCLC correctly classified 67% of all NSCLC patients, with a higher proportion of correctly classified stage IV patients compared to stage I-IIIa (Fig. 5B). The model misclassified 12% of SCLC (n = 6) and 9% of patients without LC (n = 13).

3.5. Identification of SCLC

For distinction of SCLC patients from NSCLC patients and patients without LC, the individual TM NSE had the best overall performance (ROC-AUC = 0.88, IQR: 0.83–0.91), however this TM could not meet the requirement of PPV ≥ 95% for all training sets (Fig. 3C, Supplemental Table 8). ProGRP was the only individual TM able to meet this requirement and had a ROC-AUC of 0.86 (IQR: 0.82–0.91) and sensitivity of 40% (IQR: 30–50%) with a PPV of 100% (IQR: 86–100%).

The most informative combination of protein TMs was again evaluated by recursive feature elimination. However, due to limited sample size of SCLC (n = 50), the pre-specified PPV requirement (≥95%) for the training sets could not always be achieved (Supplemental Table 9). Therefore, the performance metrics were only computed for validation sets for which the corresponding training set could meet this requirement.

The optimal sensitivity (50%, IQR: 40–60%) for identification of SCLC was achieved by models combining four protein TMs (Fig. 4C). CYFRA 21-1, NSE and proGRP were always included in these models, whereas ~50% of the models used either CA125 or CA15.3, thus both TMs contributed to the performance metrics. Therefore, the most informative combination of TMs to identify SCLC was considered to be CA125, CA15.3, CYFRA 21-1, NSE and proGRP.

This multi-TM model, adjusted for age and sex, had an increased ROC-AUC (from 0.86, IQR: 0.82–0.91 to 0.98, IQR: 0.96–0.99, $p < 0.001$) and sensitivity (40%, IQR: 30–50% to 50%, IQR: 30–60%, $p < 0.001$) compared to proGRP, with insignificant change of PPV (both 100%, IQR: 86–100%, $p = 0.85$) (Fig. 3C, Supplemental Table 8). Addition of DNA TMs to the model did not significantly change the performance, except for a slightly decreased PPV (from 100%, IQR: 86–100% to 100%, IQR: 83–100%, $p = 0.021$). In this model, NSE, proGRP, CA125 and cfDNA had positive coefficients, indicating that higher concentrations increased the probability of having SCLC (Supplemental Fig. 2). On the other hand, CYFRA 21-1, CA15.3 and a mutation in ctDNA had negative coefficients and thus decreased the probability of having SCLC.

The CA125, CA15.3, CYFRA 21-1, NSE, proGRP and DNA could correctly classify 50% of SCLC patients, comprising 25% of stage IIIa, 45% of stage IIIb-c and 54% of stage IV patients. Of NSCLC patients, 1% was misclassified (stage IIIa) and no patients without LC were misclassified (Fig. 5C).

4. Discussion

In this study, the added clinical value of protein and DNA TMs in LC diagnostics was evaluated. Compared to the performance of individual TMs, combined assessment of protein and DNA TMs improved the

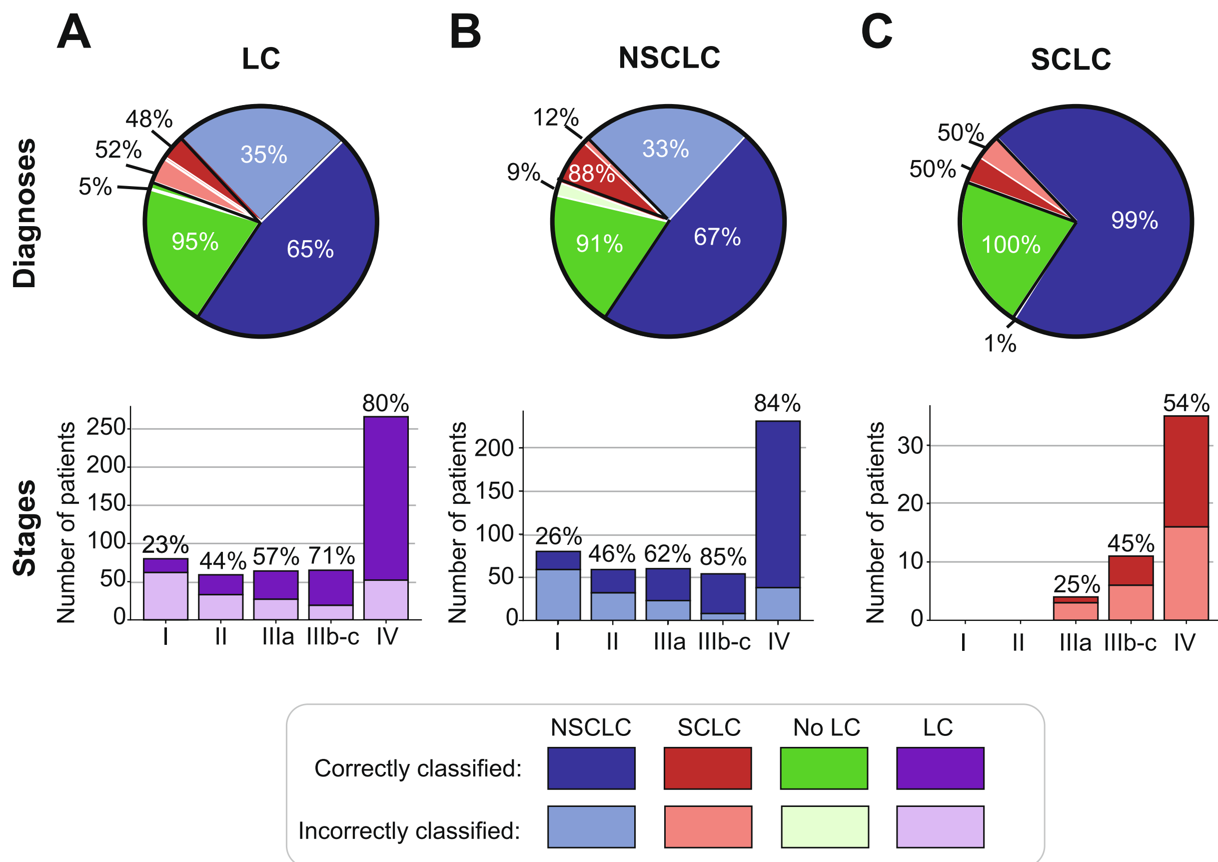


Fig. 5. Characteristics of patients correctly and incorrectly classified by the models combining protein and DNA TMs, specified per subgroup and stage. Stages are shown for the subgroup aimed to identify by the model, i.e. LC, NSCLC or SCLC. A) No LC versus LC, evaluated at PPV $\geq$ 98%. B) Non-NSCLC versus NSCLC, evaluated at PPV $\geq$ 95%. C) Non-SCLC versus SCLC, evaluated at PPV $\geq$ 95%.LC = lung cancer, TMs = tumor markers, NSCLC = non-small-cell lung cancer; SCLC = small-cell lung cancer.

overall performance and the performance at pre-specified PPVs to detect LC and its subtypes NSCLC and SCLC.

Previous research has shown that combined assessment of protein TMs could improve the performance of decision–support algorithms in LC diagnostics [15,16,19,31]. To distinguish LC from absence of LC, studies of Molina et al. [15] and Jiang et al. [31] found comparable ROC-AUCs (0.89 and 0.871, respectively). With the pre-specified NPV of 100% to exclude LC, PPV of $\geq$98% to identify LC and PPVs $\geq$ 95% to identify NSCLC and SCLC, this study however was the first that shifted the focus towards direct clinical impact of a TM–based decision–support algorithm. These pre–specified performance metrics were chosen to ensure minimal overtreatment, while also providing additional information that might lead to faster treatment decisions. The multicenter patient population used in this study consisted of patients suspected of LC and should be representative for the clinical setting in which this model could be used. Moreover, the LC patients showed similar distribution of subtypes and stages compared to the Dutch lung cancer population and therefore might be representative [3].

Exclusion of LC at a pre-specified NPV of 100% could not be achieved by individual nor combined TM assessment due to the overlapping data of patients with and without LC, which makes it difficult to discriminate the LC patients with minimally increased TM levels from the patients without LC that have slightly increased TM levels. For example protein TMs are known to be elevated in benign (lung) diseases, kidney disorders and CEA in smokers [32], while small and/or early stage tumors are known to only induce a minimal increase in TM concentrations [15,18].

Identification of LC, NSCLC and SCLC with PPV $\geq$95-98% was successfully achieved by individual and combined assessment of TMs. Similar to previous studies, CYFRA 21-1 was shown to perform best in identification of LC, followed by CEA [15,17–19] (Fig. 3A). Combining CYFRA 21-1 and CEA resulted in maximum sensitivity (65%) for identification of LC. To identify NSCLC, CYFRA 21-1, CEA, NSE and proGRP enabled maximum sensitivity (67%), as these TMs could improve differentiation between NSCLC and SCLC [16,18]. Lastly, best sensitivity (50%) to identify SCLC was achieved by combining SCLC–associated TMs NSE and proGRP with NSCLC-associated TMs CYFRA 21-1, CA125 and CA15.3. Interestingly, high concentrations of CA125 increased probability of SCLC in these models, a relationship that was to our knowledge not previously described (Supplemental Fig. 2). While previous research showed added value of HE4 and SCCA for LC detection [13,16,17], these TMs were not selected for the optimal marker combinations in this study. However, SCCA might have added value in identifying the NSCLC subtype squamous cell carcinoma [16].

Although some previous studies suggested improved performance to identify LC by combining both protein and DNA TMs [20,21], in our study addition of DNA to protein TMs models only slightly improved the ROC-AUCs and sensitivities. Mutations in ctDNA indicate presence of LC with high specificity, yet only ~20% the LC patients do have a mutation. Moreover, the majority of patients with driver mutations could already be identified based on protein TMs alone. Differences with respect to previous studies could be explained by a lower number of overall mutations identified in our study. Compared to CancerSeek [20], less mutations were identified in our study due to analysis of a smaller number of mutations, whereas CancerSeek used a more extensive panel. Even though the mutations in our study cover frequently occurring and targetable mutations in NSCLC, future research could focus on extension of mutation analyses by performing NGS on blood samples. NGS could overcome some technical limitations of ddPCR to analyze e.g. targetable

alterations in *ALK* and *MET* and could therefore improve performance of DNA TMs in the diagnostic model by identification of more (targetable) ctDNA mutations. In addition, in the study of Yin et al. [21] a higher prevalence of *EGFR* mutations was found (46%), which is common in Asian countries compared to European countries [33] and therefore resulted in improved performance compared to our study. Even though the added value of ctDNA on population level might be limited, information on targetable driver mutations is valuable to identify individual patients who could benefit from targeted therapy [10,11]. Moreover, ctDNA-ddPCR baseline measurements could be useful for monitoring of treatment response [28]. The analysis of ctDNA-ddPCR could therefore still have direct clinical impact for a selected group of patients, but does not necessarily have to be performed to improve the decision–support algorithms.

To adjust for potential effects of age and sex on TM concentrations [34], these characteristics were included in the algorithms. However, since a higher fraction of males occurred in the group without LC compared to LC, inclusion of sex could potentially have influenced model predictions for identification of LC. As only a minimal improvement of performance was found for the multi-TM model adjusted for age and sex, compared to the unadjusted model during the recursive feature elimination, this influence is expected to be limited and could be related to the correction of differences in TM concentrations as well.

The decision–support algorithms presented in this study could aid lung physicians by providing additional information for patients with high probabilities of LC, NSCLC or SCLC, even in an early stage of disease. Here, 23% of stage I and 44% of stage II disease patients were correctly classified, therewith providing important information for these type of patients with small lesions that could be more complicated to take a biopsy of. Although these models will not replace pathological diagnosis, LBx-based models may provide additional information for patients without (conclusive) pathology examination, which in our study was shown to be the case for 9% of the overall population. If the model indicates NSCLC and a mutation has been detected in ctDNA, targeted therapy may be selected, whereas chemotherapy may be the treatment trajectory in case of SCLC. Since these models were evaluated at pre–specified PPVs, a performance metric influenced by prevalence of disease, the models are applicable in patients suspected of LC with high pre–test probability of LC. Further validation should be performed to determine usage of these models for populations with other prevalence of disease, e.g. for screening purposes.

To improve the performance of liquid-biopsy based models and even confirm absence of LC, some future research directions could be considered. Firstly, Molina et al. showed that repetition of TM measurement within 3–4 weeks to determine concentration changes could improve performance [15]. This additional LBx may aid in diagnosis of patients for whom a conservative diagnostic strategy (follow-up with CT) can be applied. Secondly, a combination of the LBx models with radiological information, such as maximal FDG uptake in PET-CT scans [31] and/or nodule size [15], might improve identification of LC. To ensure reproducibility, it is necessary to obtain these input parameters in an objective manner. Thirdly, other protein TMs, such as tumor-associated autoantibodies (TAAbs) and exosomal proteins, were described to enable more specific identification of early-stage patients than currently used protein TMs, reviewed in [35]. However, the actual added value of these markers in a diagnostic setting needs to be further investigated, since reported sensitivities of these individual TAAbs were often limited and the performance was not always evaluated in a clinically relevant population [35]. Lastly, addition of measurement of PD-L1 expression in LBx would be useful to support treatment decision for immunotherapy and therefore further personalize treatment choices based on minimally-invasive information [36].

The models were trained and evaluated after exclusion of patients with other carcinomas. Since the TMs used in this study are not LC-specific, these TMs could also be elevated in patients with other carcinomas [13,32], potentially resulting in false classifications as LC, NSCLC or SCLC. Therefore, lung physicians should be aware that these models would not be applicable to confirm the presence of LC in patients with other carcinomas and could provide false positive results for patients with previously undiagnosed carcinomas.

Model performance was evaluated using cross-validation, allowing to determine performance on patients not used for training of the model. To further verify model performance for application of liquid-biopsy based decision support algorithms in clinical practice, validation on an external patient cohort should be performed.

Further studies focused on SCLC would be needed to further assess model performances on this population. As we could expect, performances of classification of SCLC patients had large uncertainties, due to a low frequency (11%) of SCLC amongst lung cancer patients [3] and therefore also a limited number of SCLC patients in our study (n = 50).

5. Conclusion

In conclusion, this study provides solid evidence that the combined assessment of the protein TMs CYFRA 21-1, CEA, NSE, proGRP, CA125 and CA15.3 and DNA TMs contain sufficient information for the development of three decision–support algorithms that allowed for identification of two–thirds of all LC and NSCLC patients and half of SCLC patients. Our results suggest that, in the future, these models may aid in the reduction of invasive procedures or offer new perspective in the diagnostics of patients for whom pathologic subtyping is impossible or incomplete. To achieve this, further discussions are required with clinicians to determine optimal presentation of model outputs and to define how to optimally integrate model predictions in the clinical decision-making process.

Funding

The study was supported by The Netherlands Organization for Scientific Research (NWO) via LIFT grant 731.017.405. Roche Diagnostics Netherlands partially sponsored the protein TM reagents used in the study. Supported by AstraZeneca grant AZNL201700295 and Catharina Onderzoeksfonds institutional fund grant 2017-02 (R.d.K.).

CRediT authorship contribution statement

Esther Visser: Data curation, Formal analysis, Investigation, Software, Visualization, Writing – original draft. **Sylvia A.A.M. Genet:** Data curation, Formal analysis, Investigation, Visualization, Writing – original draft. **Remco P.P.A. de Kock:** Data curation, Formal analysis, Investigation, Writing – review & editing. **Ben E.E.M. van den Borne:** Resources, Writing – review & editing. **Maggy Youssef-El Soud:** Resources, Writing – review & editing. **Huub N.A. Belderbos:** Resources, Writing – review & editing. **Gerben Stege:** Resources, Writing – review & editing. **Marleen E.A. de Saegher:** Resources, Writing – review & editing. **Susan C. van 't Westeinde:** Resources, Writing – review & editing. **Luc Brunsveld:** Supervision, Writing – review & editing. **Maarten A.C. Broeren:** Writing – review & editing. **Daan van de Kerkhof:** Writing – review & editing. **Birgit A.L.M. Deiman:** Resources, Supervision, Writing – review & editing. **Federica Eduati:** Methodology, Supervision, Writing – review & editing. **Volkher Scharnhorst:** Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.lungcan.2023.01.014.

References

[1] R. Rosell, E. Carcereny, R. Gervais, A. Vergnenegre, B. Massuti, E. Felip, et al., Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial, Lancet Oncol. 13 (2012) 239–246.

[2] M. Reck, D. Rodríguez-Abreu, A.G. Robinson, R. Hui, T. Csőszi, A. Fülöp, et al., Pembrolizumab versus chemotherapy for PD-L1–positive non–small-cell lung cancer, N. Engl. J. Med. 375 (2016) 1823–1833.

[3] L.E.L. Hendriks, A.-M.-C. Dingemans, D.K.M. De Ruysscher, M.J. Aarts, L. Barberio, R. Cornelissen, et al., Lung cancer in the Netherlands, J. Thorac. Oncol. 16 (2021) 355–365.

[4] P.E. Postmus, K.M. Kerr, M. Oudkerk, S. Senan, D.A. Waller, J. Vansteenkiste, et al., Early and locally advanced non-small-cell lung cancer (NSCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up, Ann. Oncol. (2017), 28: iv1–21.

[5] D.J. Sugarbaker, M.C. DaSilva, Diagnostic workup of lung cancer, Surg. Oncol. Clin. N. Am. 20 (2011) 667–679.

[6] A. McLean, D. Barnes, L. Troy, Diagnosing lung cancer: the complexities of obtaining a tissue diagnosis in the era of minimally invasive and personalised medicine, JCM 7 (2018) 163.

[7] R. Nooreldeen, H. Bach, Current and future development in lung cancer diagnosis, IJMS 22 (2021) 8661.

[8] N. Guibert, A. Pradines, G. Favre, J. Mazieres, Current and future applications of liquid biopsy in nonsmall cell lung cancer from early to advanced stages, Eur. Respir. Rev. 29 (2020), 190052.

[9] D. Di Capua, D. Bracken-Clarke, K. Ronan, A.-M. Baird, S. Finn, The liquid biopsy for lung cancer: state of the art, limitations and future developments, Cancers 13 (2021) 3923.

[10] R.D. Schouten, D.C.L. Vessies, L.J.W. Bosch, N.P. Barlo, A.S.R. van Lindert, S.A.G. M. Cillessen, et al., Clinical utility of plasma-based comprehensive molecular profiling in advanced non–small-cell lung cancer, JCO Precision Oncol. (2021) 1112–1121.

[11] M.L. Cheng, E. Pectasides, G.J. Hanna, H.A. Parsons, A.D. Choudhury, G.R. Oxnard, Circulating tumor DNA in advanced solid tumors: clinical relevance and future directions, CA A Cancer J. Clin. 71 (2021) 176–190.

[12] B.A. Chan, B.G.M. Hughes, Targeted therapy for non-small cell lung cancer: current standards and the promise of the future, Transl. Lung Cancer Res. 4 (2015) 19.

[13] B. Nagy, H.P. Bhattoa, Z. Steiber, M. Csobán, M. Szilasi, G. Méhes, et al., Serum human epididymis protein 4 (HE4) as a tumor marker in men with lung cancer [cited 2022 Sep 26];52. Available from: Clin. Chem. Lab. Med. (CCLM) [Internet]. (2014) https://www.degruyter.com/document/doi/10.1515/cclm-2014-0041/html.

[14] T. Zamay, G. Zamay, O. Kolovskaya, R. Zukov, M. Petrova, A. Gargaun, et al., Current and prospective protein biomarkers of lung cancer, Cancers 9 (2017) 155.

[15] R. Molina, R.M. Marrades, J.M. Augé, J.M. Escudero, N. Viñolas, N. Reguart, et al., Assessment of a combined panel of six serum tumor markers for lung cancer, Am. J. Respir. Crit. Care Med. 193 (2016) 427–437.

[16] L. Liu, J. Teng, L. Zhang, P. Cong, Y. Yao, G. Sun, et al., The combination of the tumor markers suggests the histological diagnosis of lung cancer, Biomed. Res. Int. 2017 (2017) 1–9.

[17] E.T. Korkmaz, D. Koksal, F. Aksu, Z.G. Dikmen, D. Icen, E. Maden, et al., Triple test with tumor markers CYFRA 21.1, HE4, and ProGRP might contribute to diagnosis and subtyping of lung cancer, Clin. Biochem. 58 (2018) 15–19.

[18] Z. Chen, X. Liu, X. Shang, K. Qi, S. Zhang, The diagnostic value of the combination of carcinoembryonic antigen, squamous cell carcinoma-related antigen, CYFRA 21-1, neuron-specific enolase, tissue polypeptide antigen, and progastrin-releasing peptide in small cell lung cancer discrimination, Int. J. Biol. Markers 36 (2021) 36–44.

[19] Q. Yang, P. Zhang, R. Wu, K. Lu, H. Zhou, Identifying the best marker combination in CEA, CA125, CY211, NSE, and SCC for lung cancer screening by combining ROC curve and logistic regression analyses: is it feasible? Dis. Markers 2018 (2018) 1–12.

[20] J.D. Cohen, L. Li, Y. Wang, C. Thoburn, B. Afsari, L. Danilova, et al., Detection and localization of surgically resectable cancers with a multi-analyte blood test, Science 359 (2018) 926–930.

[21] J. Yin, W. Hu, H. Gu, J. Fang, Combined assay of circulating tumor DNA and protein biomarkers for early noninvasive detection and prognosis of non-small cell lung cancer, J. Cancer 12 (2021) 1258–1269.

[22] Dutch Federation of Medical Specialists. Guideline Database Non-Small Cell Lung Cancer [Internet]. Available from: https://richtlijnendatabase.nl/richtlijn/niet_kleincellig_longcarcinoom/algemeen.html.

[23] Dutch Federation of Medical Specialists. Guideline Database Small-Cell Lung Cancer [Internet]. Available from: https://richtlijnendatabase.nl/richtlijn/kleincellig_longcarcinoom/kleincellig_longcarcinoom_-_startpagina.html.

[24] F.C. Detterbeck, D.J. Boffa, A.W. Kim, L.T. Tanoue, The eighth edition lung cancer stage classification, Chest 151 (2017) 193–203.

[25] R. de Kock, B. Deiman, R. Kraaijvanger, V. Scharnhorst, Optimized (Pre) analytical conditions and workflow for droplet digital PCR analysis of cell-free DNA from patients with suspected lung carcinoma, J. Mol. Diagn. 21 (2019) 895–902.

[26] S.A.A.M. Genet, E. Visser, B.E.E.M. van den Borne, M.-Y.-E. Soud, H.N. A. Belderbos, G. Stege, et al., Correction of the NSE concentration in hemolyzed serum samples improves its diagnostic accuracy in small-cell lung cancer, Oncotarget 11 (2020) 2660–2668.

[27] R. de Kock, B. van den Borne, M. Youssef-El Soud, H. Belderbos, L. Brunsveld, V. Scharnhorst, et al., Therapy monitoring of EGFR-positive non–small-cell lung cancer patients using ddPCR multiplex assays, J. Mol. Diagn. 23 (2021) 495–505.

[28] R. de Kock, B.van den Borne, M.Y.-E. Soud, H. Belderbos, G. Stege, M. de Saegher, et al., Circulating biomarkers for monitoring therapy response and detection of disease progression in lung cancer patients, Cancer Treat. Res. Commun. 28 (2021) 100410.

[29] T. Fawcett, An introduction to ROC analysis, Pattern Recogn. Lett. 27 (2006) 861–874.

[30] S. Cedrés, I. Nuñez, M. Longo, P. Martinez, E. Checa, D. Torrejón, et al., Serum tumor markers CEA, CYFRA21-1, and CA-125 are associated with worse prognosis in advanced non–small-cell lung cancer (NSCLC), Clin. Lung Cancer 12 (2011) 172–179.

[31] R. Jiang, X. Dong, W. Zhu, Q. Duan, Y. Xue, Y. Shen, et al., Combining PET/CT with serum tumor markers to improve the evaluation of histological type of suspicious lung cancers. Chen C-T, editor, PLoS One 12 (2017) e0184338.

[32] R. Molina, S. Holdenrieder, J.M. Auge, A. Schalhorn, R. Hatz, P. Stieber. Diagnostic relevance of circulating biomarkers in patients with lung cancer. Holdenrieder S, Stieber P, editors. CBM. 6 2010 163–78.

[33] A. Midha, S. Dearden, R. McCormack. EGFR mutation incidence in non-small-cell lung cancer of adenocarcinoma histology: a systematic review and global map by ethnicity (mutMapII). Hong Kong. :25.

[34] Y. Li, M. Li, Y. Zhang, J. Zhou, L. Jiang, C. Yang, et al., Age-stratified and gender-specific reference intervals of six tumor markers panel of lung cancer: a geographic-based multicenter study in China [cited 2022 Sep 29];35. Available from: J. Clin. Lab. Anal. [Internet]. (2021) https://onlinelibrary.wiley.com/doi/10.1002/jcla.23816.

[35] H. Huang, Y. Yang, Y. Zhu, H. Chen, Y. Yang, L. Zhang, et al., Blood protein biomarkers in lung cancer, Cancer Lett. 551 (2022), 215886.

[36] L. Kloten, S. Krahn, Circulating tumor cell PD-L1 expression as biomarker for therapeutic efficacy of immune checkpoint inhibition in NSCLC, Cells 8 (2019) 809.