Statistical approach to mediastinal staging in NSCLC with M.E.S.S.i.a. software. Preliminary data and multicenter prospective validation study framework

Thomas Galasso[1], Lorenzo Corbetta[2], Laura Mancino[3], Lucio Michieletto[3], Loris Ceron[3]

[1]Pulmonology Unit, Santa Croce Hospital, Fano; [2]Unit of Diagnostic and Interventional Pulmonology, Careggi University Hospital, Florence; [3]Pulmonology Unit, dell'Angelo Hospital, Venice-Mestre, Italy

Abstract

The exclusion of pathological involvement of mediastinal lymph nodes in patients affected by NSCLC plays a central role in assessing their prognosis and operability. Ceron *et al*. developed a software - called M.E.S.S.i.a (Mediastinal Evaluation with Statistical Support; instan approach) - that allows the calculation of the residual probability of lymph node involvement after a certain number of tests has been done, by integrating every test result with the pre-test prevalence. M.E.S.S.i.a. bridges a gap of current American College of Chest Physicians (ACCP) guidelines, providing probability values of mediastinal metastasis for a correct clinical decision. We conducted a preliminary retrospective study in a series of 108 patients affected by non small cell lung cancer (NSCLC). Pathological staging was compared to the probability of nodal involvement calculated by M.E.S.S.i.a. software. Forty-two out of 108 subjects (39%) had a calculated post-test probability <8%; none of these had proven N2/N3 metastasis at surgical staging (negative predictive value, NPV: 100%). In 12/41 cases M.E.S.S.i.a. was able to avoid invasive procedures. The remaining 66 (61%) patients did not reach the surgical threshold; among these, 11 displayed N2 positivity at pathological staging. Receiving operator curve (ROC) analysis produced an area under curva (AUC) value of 0.773 ($p<0.001$).

These preliminary data show a high accuracy of M.E.S.S.i.a. software in excluding N2/N3 lymph node involvement in NSCLC. We have therefore promoted a prospective multicenter study in order to to get a validation of the calculator at different levels of probability of lymph node involvement. The recruitable subjects are potentially operable NSCLC patients; the gold standard for detection of mediastinal disease is the surgical lymph node dissection.

Correspondence: Loris Ceron, UOC Pneumologia, Ospedale dell'Angelo, Via Paccagnella 11, 30174 Mestre (VE), Italy.
Mobile +39.348.5500914 – Fax: +39.041.9657290.
E-mail: lorceron@gmail.com

Key words: pre-test probability; post-test probability; Bayes' theorem; mediastinal staging.

Contributions: TG, LCo, manuscript drafting; LCe, revision and amendments on the basis of M.E.S.S.i.a.'s requirements; LMa, LMi, contributed to the discussion of some controversial points, verification of the final draft.

Conflict of interest: the authors declare that there is no potential conflicts of interest in the manuscript, including financial, consultant, institutional and other relationships that might lead to bias or a conflict of interest.

Received for publication: 24 March 2019.
Accepted for publication: 24 July 2019.

Monaldi Archives for Chest Disease 2019; 89:1068
doi: 10.4081/monaldi.2019.1068

Introduction

An accurate clinical staging of lung cancer according to the TNM system is essential to determine the anatomic extent of the disease, to define the best treatment strategy and to establish a correct prognosis [1-3]. In absence of distant metastases, neoplastic involvement of mediastinal lymph nodes is the most important factor affecting prognosis and treatment [3,4].

The aim of mediastinal staging is to identify patients with mediastinal lymph node involvement with the highest degree of certainty, in order to exclude them from surgical treatment. Each investigation is different in terms of sensitivity and specificity, as well as in terms of costs and invasiveness on the patient [5]. Of course, an extensive use of invasive or minimally invasive approach would be associated with overall significant morbidity and costs; besides, a negative aspirate does not rule out metastatic lymphadenopathy.

The most used strategy in N2/N3 assessment begins with imaging study of the mediastinum and then continues with mini-invasive bioptic procedures and potentially more invasive procedures in case of cytological negativity. Each finding, except for a positive biopsy, should be interpreted depending on the positive predictive value (PPV) or the NPV of the test employed; the same can be determined by knowing the intrinsic performance characteristics of the test (*i.e.*, *sensitivity* and *specificity*) and the *prevalence* of the disease in the sample [6]. By expressing the predictive value of a test without informations on the sample characteristics, as if the predictive values were intrinsic and fixed qualities of the test, leads to substantial

differences in the interpretation of a result and it explains, for example, the different importance given to a negative cytology in different papers [7,8].

The most recent guidelines recommend various staging strategies which consider more or less implicitly the *a priori* probability of lymph node involvement, by identifying situations in which the same features (for example negative positron emission tomography, PET) can be conclusive for surgical decision, and others requiring additional investigations [1-3]. Three possible imaging scenarios at computed tomography (CT) and PET scan, which correlate with different probabilities of mediastinal nodal involvement, are described:

1 - peripheral tumor with size ≤3 cm and absence of hilar nodes (cN0) or stage cIA → low risk; 2 - central tumor *or* size >3 cm *or* presence of hilar lymphadenopathy (cN1) → intermediate risk; 3 - tumor with mediastinal lymphadenopathy (cN2) → high risk [5,9,10].

Direct use of surgery is recommended only in scenario 1, due to very low probability of mediastinal lymph node involvement. In the other two scenarios guidelines recommend mediastinal lymph node sampling, starting with minimally invasive techniques (transbronchial needle aspiration by endo bronchial ultrasound (EBUS-TBNA) or fine needle aspiration by endoscopic ultrasound (EUS-FNA)) since they reduce the number of unnecessary thoracotomies [11] and are more cost-effective [12,13] with respect to mediastinoscopy. This approach however does not suggest any objective criteria to quantify the risk of metastases in case of negative cytology, leaving the subsequent choise undetermined. In fact, while NCCN guidelines recommend surgical confirmation whenever negative cytology occurs in a clinically (PET and/or CT) positive mediastinum [2], ACCP ones loosely suggest to proceed to invasive surgical biopsy *"if clinical suspicion of nodal disease remains high"* without however providing any effective measure to specify this statement [5]. Actually, systematic resort to surgical staging when lymph node aspirate is negative, without a leastwise rough evaluation of the "pre-test" risk of nodal involvement appears inappropriate.

The M.E.S.S.i.a. Project

Ceron *et al*. suggested to overcome these limitations with a probabilistic, reasoned and evidence-based approach to mediastinal staging [14], like already applied to solitary pulmonary nodule. Ceron's proposal, based on Bayes' theorem, integrates the prognostic factors for N2/N3 involvement with the performance of each diagnostic test. It is therefore possible to interpret step by step the result of every test by combining it with the pre-test probability as obtained by the previous tests. A calculating software (M.E.S.S.i.a) was built, which allows to determine the residual probability of mediastinal lymph node involvement after every investigation performed, and therefore to assess when a patient should undergo surgery or alternatively further investigations are required [14,15].

A pre-operative strategy should reduce the post-test probability of unexpected mediastinal metastases at surgery below a threshold value, *i.e.* <5% as reported by Dooms [16] or, more realistically, up to 10% as suggested by the Working Group of the European Society of Thoracic Surgery (ESTS) [9]. Ceron *et al*. proposed a threshold around 7-8%, based on a previous virtual economic assessment of different staging strategies [17]. Therefore the possibility to accurately calculate the post-test probability of mediastinal involvement could improve the use of the available tools and resources in a cost-effective way.

The calculating software along with its work process and bibliographic data are published on a dedicated web site (www.messiaproject.com). The M.E.S.S.i.a. calculator interface is shown in Figure 1. By selecting the initial informations concerning the characteristics of the tumor [location, size, pleural contact, histology, carcinoembryonic antigen (CEA), N1 status] it is possible to get a real time numeric value corresponding to the pre-test probability of mediastinal node involvement. Based on the results from the different tests carried out in mediastinal staging (Table 1) this number will change, according to the changing post-test probability of residual lymph node metastasis.

The two most frequent histologic types are considered, adenocarcinoma (ADK) and squamous cell carcinoma (SCC). The software assumes ADK as histologic type in case of unknown histology or different histologic subtypes other than SCC, since ADK is the most frequent tumor and has the highest risk of metastatic spread; this choice seems to be reasonable and prudential as the main purpose is to select subjects with the lowest risk of mediastinal metastasis, although an overestimation of the final probability is possible in case of SCC. Therefore a preoperative histologic definition is desirable, not mandatory, thus avoiding the need for invasive investigations (*e.g.*, percutaneous needle aspiration).

About CEA level, many contributions report a significative correlation with tumor's histology, lymphatic spread, recurrence after surgery and disease free survival [18-21]. However, most works reporting mediastinal metastasis prevalence in lung cancer do not consider CEA value, therefore the reported prevalence represents an "intermediate" value between cases respectively with normal and elevated CEA; this value was chosen in our software when CEA is not avalaible (pre-test value). In case of ADK the prevalences for "CEA <5" (post-test values) were calculated using a likelihood ratio (LR) -0.719 [22] starting from the pre-test values [23-28]; the prevalences for "CEA ≥5" were obtained from literature data [29,30] and compared with the measures obtained using Bayes' theorem [22]. Limited to SCC, the values for "CEA <5" are the same as for "CEA unknown", given the low prevalence of elevated CEA in SCC.

Table 1. Sensitivity, specificity and likelywood ratios used in M.E.S.S.i.a. calculation software.

Exam	Sensitivity (%)	specificity (%)	LR+	LR-
CT	55	80	2.75	0.562
PET (LN <1 cm)	75	93	10,714	0.269
PET (LN ≥1 cm)	91	78	4136	0.115
TBNA	78	99	78	0.222
EBUS/EUS	90	99	90	0.101

M.E.S.S.i.a. system allows a mediastinal evaluation on a "per lymph node" basis. Actually, CT is considered *positive* when even one lymph node is enlarged, along with other small ones ("per patient" judgement); however, M.E.S.S.i.a. approach permits a dedicated prediction of each lymph node involvement, and consequently suggests the correct decision "node per node". For example, in case of large peripheral ADK (>7 cm) with normal CEA and no fluorodeoxyglucose (FDG) activity in the mediastinum, by selecting "CT pos" on the "Staging Pathway" column, the resulting probability is 9%, while by selecting "CT neg" the same decreases to 6%; this means that PET negative enlarged mediastinal lymph nodes should be sampled, while PET negative normal sized ones should not (distinct prediction of mediastinal involvement "node per node"); of course, just one suspicious node (probability ≥8%) is enough to submit a patient to additional investigations, focused on the suspicious target - namely, on the lymph node(s) with probability ≥8%.

Based on previous assessments [17], the surgical threshold is considered to be reached when the calculated value falls below 8%, meaning that the patient could directly undergo surgical intervention without need for any further investigation. Otherwise we consider "surgical threshold not reached" if the post-test probability remains ≥8%.

The main advantage of M.E.S.S.i.a. is to optimize the use of resources in lung cancer staging; in comparison with the guideline recommendations, the staging path could be stopped earlier in some situations, while in others more investigations should be performed to lower the risk under the threshold of 8%. In both cases a cost-saving is expected: in the former, as fewer investigations are performed; in the latter, due to futile thoracotomies sparing. For example, in peripheral cIA stage ACCP guidelines state that negative CT prompts a direct recourse to surgery; conversely, M.E.S.S.i.a. calculates that this is valid only for SCC or subcentrimetric ADK, while in case of larger ADK, PET is mandatory (residual probability of mediastinal involvement 8-15%). On the other hand, in central tumor or cN1 involvement guidelines suggest EBUS regardless of CT and PET results, while M.E.S.S.i.a. suggests possible direct recourse to surgery after negative CT and PET in SCC (probability 4% in central and 6% in cN1 tumor).

Hence, M.E.S.S.i.a. allows a more precise application of current guidelines; furthermore, it bridges the gap that arises when the same recommend to proceed to invasive surgical biopsy *"if clinical suspicion of nodal disease remains high"* after negative cytology, without giving a precise definition and estimate of *"high clinical suspicion"*; conversely, M.E.S.S.i.a. provides probability and threshold values for a correct decision.

Patients and Methods

To get preliminary data on calculator's performance 108 patients (73 men, 35 women; mean age at intervention 69 years) who had undergone surgical resection for non-small cell lung cancer (NSCLC) at the Ospedale dell'Angelo (Mestre, Venice, Italy), between January 2015 and October 2016 were retrospectively analyzed. Before surgery, tumor histology was available in 36 patients (ADK n=23; SCC n=13). Conventional transbronchial needle aspiration (TBNA) was performed in 5 cases. Thirteen patients underwent EBUS; 1 patient underwent cervical mediastinoscopy.

The pathological staging was: N0 in 86 patients (86%), N1 in 11 patients (10%) and N2 in 11 patients (10%). No patients with surgical N3 status were present in our sample. The 11 patients with pathologic N2 were as follows: T1aN2 (n=3); T2aN2 (n=5); T3n2 (n=3) (Table 2).

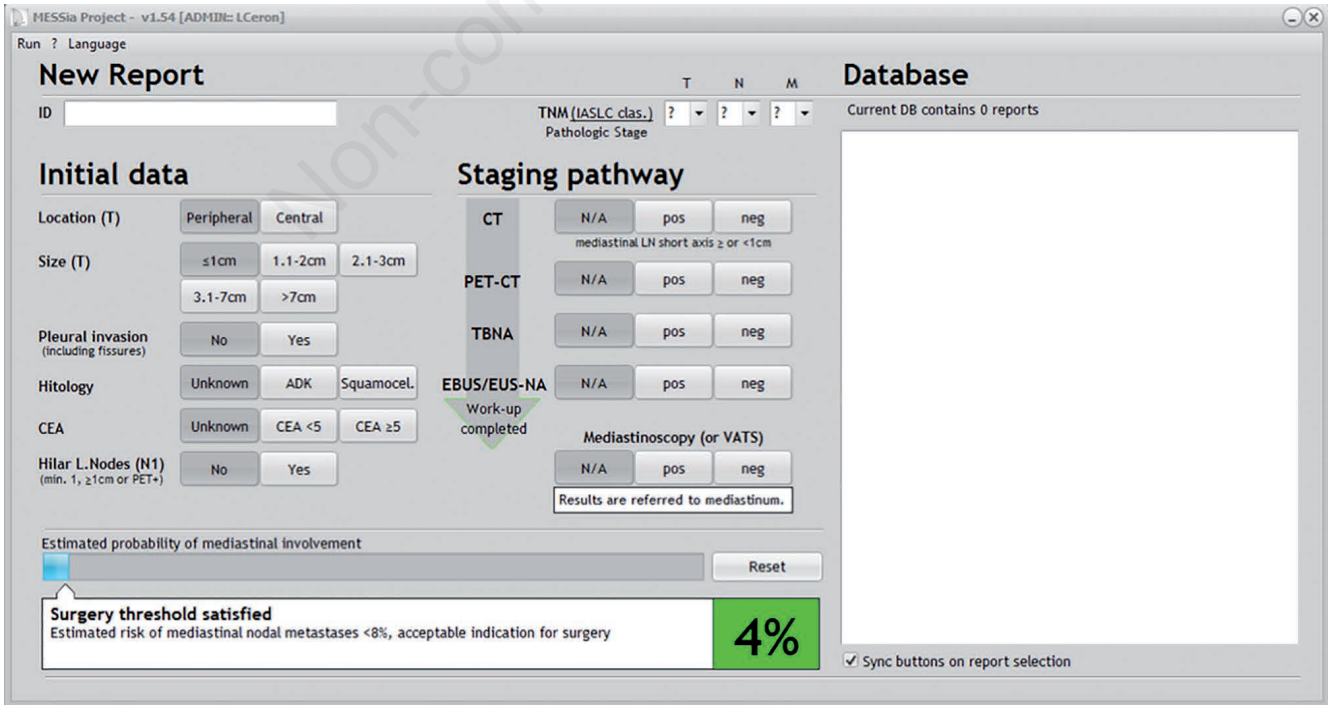

Figure 1. M.E.S.S.i.a. software interface.

Table 2. Clinical staging and post-test probability values as calculated by M.E.S.S.i.a. software *vs* surgical staging of each patient. In bold font the pathological N2 patients.

Patient	Clinical staging	M.E.S.S.i.a. (%)	Pathological staging
1	I a	2	pT1aN0
2	I b	6	pT2aN0
3	I a	13	pT1aN0
4	II a	26	pT1aN0
5	I a	25	pT1bN0
6	III a	2	pT1bN0
7	**III a**	**16**	**pT2aN2**
8	I a	6	pT1bN0
9	**I b**	**8**	**pT2aN2**
10	III a	7	pT2aN0
11	I b	8	pT2aN0
12	II a	26	pT1aN1
13	I a	8	pT1aN0
14	I a	8	pT1aN0
15	I a	1	pT1aN0
16	I b	13	pT1aN0
17	I a	8	pT1aN0
18	I a	2	pT1aN0
19	I b	4	pT2aN0
20	I a	3	pT1bN0
21	I a	2	pT1bN0
22	III a	11	pT1bN0
23	I b	13	pT2aN0
24	I a	8	pT1aN0
25	II a	26	pT2bN1
26	II a	26	pT1bN0
27	I a	8	pT1bN0
28	I a	8	pT1bN0
29	II a	3	pT2aN0
30	II a	26	pT2aN0
31	II a	26	pT2aN1
32	I a	26	pT2aN0
33	II a	26	pT1aN0
34	**II b**	**13**	**pT3N2**
35	**II a**	**26**	**pT3N2**
36	**I a**	**8**	**pT1aN2**
37	II b	26	pT2bN0
38	I a	8	pT3N0
39	II a	3	pT1bN1
40	I a	8	pT1bN0
41	I a	2	pT3N1
42	I b	4	pT4N0
43	I a	25	pT1aN0
44	I b	8	pT2aN0
45	**I b**	**8**	**pT1aN2**
46	I a	2	pT1bN0
47	I a	8	pT2aN0
48	I a	9	pT1bN0
49	I a	6	pT1bN0
50	I b	6	pT2aN0
51	I a	1	pT1aN0
52	I a	19	pT1aN0
53	I a	2	pT1aN0
54	I b	13	pT1bN0
55	II a	26	pT1aN1
56	I b	3	pT2bN0
57	I a	6	pT1aN0
58	I b	13	pT1bN0
59	III a	14	pT3N1
60	III a	16	pT2N0
61	II b	**26**	**pT2aN2**
62	I a	2	pT1aN0
63	I b	2	pT3N1
64	II a	6	pT3N0
65	I a	8	pT1aN0
66	I a	25	pT1bN0
67	III a	21	pT1bN0
68	II a	**26**	**pT2aN2**
69	I a	2	pT1bN0
70	I a	28	pT1bN0
71	I a	11	pT1aN0
72	III a	1	pT3N0
73	II a	3	pT1aN1
74	I b	**25**	**pT3N2**
75	II a	26	pT2aN0
76	II a	26	pT2N0
77	III a	6	pT1aN0
78	II a	8	pT3N0
79	I a	1	pT1aN0
80	III a	**27**	**pT2aN2**
81	III a	6	pT2bN0
82	I a	2	pT2N0
83	I a	8	pT1bN0
84	I b	19	pT1bN0
85	I a	8	pT1aN0
86	I a	2	pT1aN0
87	I a	4	pT2N1
88	III a	72	pT2aN0
89	I b	8	pT2aN0
90	II b	26	pT2bN0
91	I a	2	pT1aN0
92	III a	53	pT2aN0
93	I a	4	pT2bN0
94	I a	9	pT1aN0
95	I a	4	pT1aN0
96	I a	9	pT1bN0
97	II a	8	pT3N0
98	III a	**96**	**pT1aN2**
99	I a	2	pT1aN0
100	III a	16	pT1aN0
101	I a	2	pT2N0
102	I a	26	pT1aN0
103	I a	8	pT2N0
104	III a	1	pT2N1
105	I b	8	pT2bN0
106	III a	4	pT3N0 disease
107	I b	8	pT2N0
108	I a	1	pT1aN0

Results

Forty-two out of 108 subjects (39%) were recognized by the software as "compatible with surgical indication". Forty-one of these 42 patients were considered compatible with surgery after CT and PET; 12 of them (29%) would be candidates to invasive investigations on guideline basis. In the remaining patient, M.E.S.S.i.a. yielded a probability >8% after CT and PET, whereas guidelines did not indicate the need for additional testing; the following negative EBUS reduced the probability below 8%. In this group (which included 70% clinical stage I) no N2 involvement was found, hence in 12/41 cases M.E.S.S.i.a. was able to avoid invasive procedures; just in 1 case (2%) M.E.S.S.i.a. caused an increase in resource consumption (EBUS). The remaining 66 patients (61%) did not reach the surgical threshold; of these, 11 were N2 positive at surgical staging. Therefore sensitivity, specificity, accuracy, negative predictive value and positive predictive value were 100%, 43%, 42%, 100% and 17%, respectively. About the apparently very low positive predictive value, see argumentation in the next section.

We divided the 66 patients whose probability fell above the 8% treshold into sub-groups depending on increasing levels of post-test probability (8-24% n=40 ; 25-50% n=23 and >50% n=3). The prevalence of N2-positivity observed at surgical staging was 12% (5/40) in the 8-24% range, 21% (5/23) in the 25-50% range and 33% (1/3) in the >50% group (Table 3).

ROC analysis produced an AUC value of 0.773 [95% CI 0.683 -0.848] (p<0.001) corresponding to a moderately accurate test according to the Swets classification [31] (Figure 2). The 95% confidence interval is rather small, showing that, in spite of overall accuracy little more than sufficient, M.E.S.S.i.a. method is not affected by important variability.

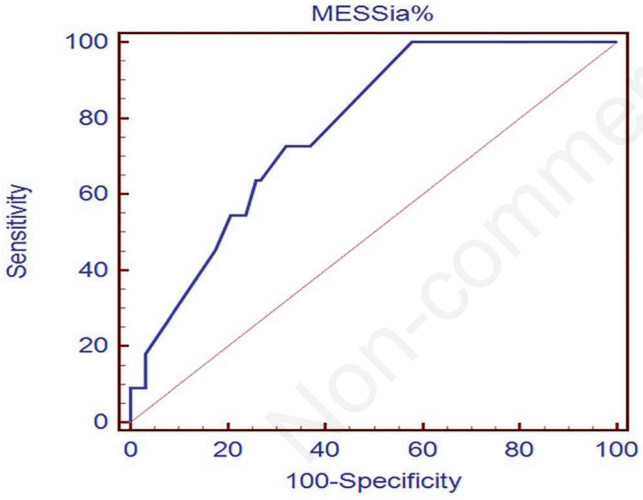

Figure 2. Operating characteristic curve of M.E.S.S.i.a.; AUC 0.773 [95% CI 0.683-0.848] (p<0.001).

Discussion

M.E.S.S.i.a. is a software that calculates the probability of lymph node involvement in NSCLC after a certain number of tests has been done. We used the calculator in a retrospective cohort of patients operated for NSCLC, to obtain preliminary data of accuracy. We showed that a surgical threshold below 8% generates a very high negative predictive value (equal to 100%).

Besides, we noticed a rough correlation between the prevalence of surgical N2 positivity and the expected probability as calculated by the software, despite overall moderate overestimation. In spite of a very high sensitivity, a low specificity (43%) was observed; namely, the majority of patients above the surgical threshold (probability of mediastinal metastasis ≥8%) demonstrated no mediastinal spread at surgical time (false positives). This would produce a very disappointing PPV (17%), if standard statistical formulas were employed (PPV=TP/TP+FP). However, one must bear in mind that the expected PPV of M.E.S.S.i.a. is not 100, but varies according to the value of estimated probability; actually, the optimal rate of surgical confirmation for a given level of probability is not 100%, but a percentage matching the value of probability calculated by the software, *i.e.* 20% if the calculator result is positive with a probability prediction of 20%. In other terms, it means that out of 100 patients considered positive with a 20% probability of mediastinal involvement, not all of them, but ideally 20%, are expected to have a pathological mediastinum; in this situation the best PPV would be therefore 20%, not 100%.

The overestimation produced by the calculator can be due to a selection bias as well. In fact, by excluding patients with a positive cytological assessment (who therefore did not underwent surgery) we obtained a really small rate (16%) of true positive cases, resulting in a lower than expected PPV. This selection bias prevented us from satisfactory conclusions on the overall accuracy of the M.E.S.S.i.a. software in case of positive results (*i.e.*, when the post-test probability value lies above the surgical threshold); in fact in our analysis the overall accuracy is just more than sufficient. Despite that it has a very little statistical variability, and therefore it seems to perform in a stable and reproducible manner. Another

Table 3. Prevalence of pN2 (*i.e.*, positivity at pathological staging) in patient groups divided according with post-test probability calculated by M.E.S.S.i.a. software.

Post-test value interval*	Number of patients	Positive patients pN2(n)	Positive patients pN2 (%)
0-7%	42	0	0
8-12%	27	3	11
13-24%	13	2	15
25-50%	23	5	22
>50%	3	1	33

*Probability values generated by the calculator after all diagnostic tests have been performed.

limitation, although less important, is represented by the low proportion of patients whose preoperative histology was available; hence, some patients with SCC were by default examined as having an ADK, which implicates a higher rate of nodal metastasis. Moreover, one could argue that assuming ADK as histologic type in case of different histologic subtypes other than SCC, potentially generates some bias; however, authors believe that this cannot significantly reduce MESSIA's evaluating power, given the negligible incidence of these histotypes [32].

Finally, an economic evaluation was conducted of different mediastinal staging strategies; the same demonstrated that the main saving determiner in mediastinal staging is reducing the final probability of nodal metastasis through a rational use of the current tests; in particular cost-effectiveness is evident under 7-8%, with further investigations increasing expenses without significant advantage in terms of accuracy [17]; so, a statistical approach which allows an objective estimation of the final probability is crucial in the staging path.

To confirm all the above mentioned assumptions and to evaluate the performance of the calculator at each level of probability above the surgical threshold, a very large sample is therefore needed.

Multicenter validation prospective study

A large multicenter, prospective study (study code: ARC239) endorsed by Italian Association of Hospital Pulmonologists (AIPO) has started in January 2019, which could eliminate the selection bias and have a sufficient statistical power to assess the calculator accuracy; to date, 26 Italian Centres of Interventional Pulmonology have joined the project. The main purpose of the study is to evaluate the accuracy of M.E.S.S.i.a. software in identifying patients with NSCLC at very low risk of mediastinal lymph node involvement. The recruitable subjects are potentially operable patients that need pre-operative mediastinal staging, in full agreement with the current guidelines; almost 1000 patients will be collected to obtain a sufficient statistical strength.

For each patient the probability of N2-N3 involvement will be determined by the M.E.S.S.i.a. calculator, based on the pre-operative CT characteristics of the tumor, N status at CT and PET scan, as well as any cytological result from TBNA, EBUS-TBNA or EUS FNA; furthermore, all patients should have a determination of serum CEA and if possible of tumor histology at the time of staging. Clear definitions and well agreed criteria should be used with regard to tumor location (peripheral *vs* central), pleural contact, T size, CT and/or PET scan positive lymph nodes. A thorough application of the diagnostic tests is recommended, in order to obtain precise data available for statistical analysis, which could finally better define and reset the values of the *a prior*" data adopted in a "feed-back fashion". The cytological techniques (EBUS and EUS) should be used according to the criteria of accuracy suggested by the literature (selection of nodal stations and of individual lymph nodes, sequence of sampling, numbers of aspirates, *etc*.) [33,34]; mandatory requirement is that at least every lymph node considered suspicious by the calculator (probability ≥8%) be sampled. For each test the sensitivity and specificity values are set in the software, based on the latest and more reliable literature data; the gold standard for detection of mediastinal disease is surgical lymph node dissection during anatomical resection of the tumor. As stated by the study protocol, for ethical reasons the step-by-step decision whether to operate the patient or to proceed with further investigations will not be based on calculator's results, but instead on investigators' opinion, according with current guidelines. After any clinical decision and regardless of the same, calculator will be employed in order to get an indipendent, statistics based, node-per-node evaluation of the staging path (see paragraph "The M.E.S.S.i.a. project"); the results of M.E.S.S.i.a. prospective study will be then used to validate the calcutator's performance. Once its efficacy and accuracy is confirmed, the same can be utilized in clinical practice to guide decisions on staging management.

The primary goal of the study is the determination of the negative predictive value of M.E.S.S.i.a. when the pre-determined surgical threshold <8% is reached [17], irrespective of which and how many tests are used; *i.e*., if the surgical threshold is reached after negative CT and PET, the prediction ability of the calculator can be verified at this staging level. Secondary target will be the evaluation of the overall accuracy of the calculator.

Conclusions

The statistical approach to mediastinal staging seems rigorous and promising. The M.E.S.S.i.a. software is the first and practical tool based on statistical approach widely available for mediastinal staging; by our preliminary data, it has shown very high sensitivity and NPV. A multicenter prospective study on a large sample as representative as possible of the different scenarios of nodal disease prevalence is running, to obtain a clinical validation of its accuracy at all levels of probability. The strength of the Bayesian method relies in its dynamic and open essence, namely the ability to update the sensitivity, specificity and likehood ratios of the various tests based on the scientific evidence or on the performance characteristics of the single center/operator. In the future M.E.S.S.i.a. may also include new possible tools for the study of mediastinal lymph nodes, for example the endosonographic and elastosonographic or even the magnetic resonance (MRI) features of the lymph nodes (35,36), as soon as their predictive values have been calculated in homogeneous samples of patients where the prevalence of the nodal disease is well known.

References

1. Rivera MP, Mehta AC, American College of Chest Physicians. Initial diagnosis of lung cancer: ACCP evidence-based clinical practice guidelines (2nd edition). Chest 2007;132:131S-48.
2. National Comprehensive Cancer Network (NCCN). Clinical practice guidelines in oncology (NCCN Guidelines). Non-small cell lung cancer. Ver. 5.2019. Accessed: June 7, 2019.
3. De Leyn P, Lardinois D, Van Schil PE, et al. ESTS guidelines for preoperative lymph node staging for non-small cell lung cancer. Eur J Cardiothor Surg 2007;32:1-8.
4. Andre F, Grunenwald D, Pignon JP, et al. Survival of patients with resected N2 non-small-cell lung cancer: Evidence for a sub-classification and implications. J Clin Oncol 2000;18:2981-9.
5. Silvestri GA, Gonzalez AV, Jantz MA, et al. Methods for staging non-small cell lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest 2013;143:e211S-50.
6. Coulthard MG. Quantifying how tests reduce diagnostic uncertainty. Arch Dis Child 2007;92:404-8.
7. Hwangbo B, Kim SK, Lee HS, et al. Application of endobronchial ultrasound-guided transbronchial needle aspiration

following integrated PET/CT in mediastinal staging of potentially operable non-small cell lung cancer. Chest 2009;135:1280-7.
8. Rintoul RC, Tournoy KG, El Daly H, et al. EBUS-TBNA for the clarification of PET positive intra-thoracic lymph nodesan international multi-centre experience. J Thorac Oncol 2009;4:44-8.
9. De Leyn P, Dooms C, Kuzdzal J, et al. Revised ESTS guidelines for preoperative mediastinal lymph node staging for non-small-cell lung cancer. Eur J Cardiothorac Surg 2014;45:787-98.
10. Vilmann P, Clementsen PF, Colella S, et al. Combined endobronchial and esophageal endosonography for the diagnosis andstaging of lung cancer: European Society of Gastrointestinal Endoscopy (ESGE) Guideline, in cooperation with the European Respiratory Society (ERS) and the European Society of Thoracic Surgeons (ESTS). Endoscopy 2015;47:c1.
11. Annema JT, van Meerbeeck JP, Rintoul RC, et al. Mediastinoscopy vs endosonography formediastinal nodal staging of lung cancer: a randomized trial. JAMA 2010;304:2245-52.
12. Rintoul RC, Glover MJ, Jackson C, et al. Cost effectiveness of endosonography versus surgical staging in potentially resectable lung cancer: a health economics analysis of the ASTER trial from a European perspective. Thorax 2014;69:679-81.
13. Sharples LD, Jackson C, Wheaton E, et al. Clinical effectiveness and costeffectiveness of endobronchial and endoscopic ultrasound relative to surgical staging in potentially resectable lung cancer: results from the ASTER randomised controlled trial. Health Technol Assess 2012;16:1-75, iii-iv.
14. Ceron L, Michieletto L, Zamperlin A. Mediastinal staging in lung cancer: a rational approach. Monaldi Arch Chest Dis 2009;71:170-5.
15. Ceron L, Michieletto L, Zamperlin A, et al. The challenge of mediastinal staging. Eur Oncol Haematol 2011;7:31-5.
16. Dooms C, Decaluwe H, De Leyn P. Mediastinal staging. Eur Respir Monogr 2015;68:159-66.
17. Ceron L, Mancino L, Michieletto L, Zamperlin A. Bayesian approach to mediastinal staging of lung cancer: economic analysis. Rassegna Patologia Apparato Respiratorio 2013;28:258-64.
18. Hanagiri T, Sugaya M, Takenaka M, et al. Preoperative CYFRA 21-1 and CEA as prognostic factors in patients with stage I non-small cell lung cancer. Lung Cancer 2011;74:112-7.
19. Lee S, Lee CY, Kim DJ, et al. Correlation of serum carcinoembryonic antigen and cytokeratin fragment in resected nonsmall cell lung cancer. Korean J Thorac Cardiovasc Surg 2013;46:192-6.
20. Kim JJ, Hyun K, Park JK, Moon SW. The significance of serum carcinoembryonic antigen in lung adenocarcinoma. Korean J Thorac Cardiovasc Surg 2015;48:335-44.
21. Maeda R, Suda T, Hachimaru A, et al. Clinical significance of preoperative carcinoembryonic antigen level in patients with clinical stage IA non-small cell lung cancer. J Thorac Dis 2017;9:176-86.
22. Chen F, Yan CE, Li J, et al. Diagnostic value of CYFRA 21-1 and CEA for predicting lymph node metastasis in operable lung cancer. Int J Clin Exp Med 2015;8:9820-4.
23. Zhou Q, Suzuki K, Anami Y, et al. Takamochi K. Clinicopathologic features in resected subcentimeter lung cancer – status of lymph node metastases. Interact Cardiovasc Thorac Surg 2010;10:53-7.
24. Seok Y, Yang HC, Kim TJ, et al. Frequency of lymph node metastasis according to the size of tumors in resected pulmonary adenocarcinoma with a size of 30 mm or smaller. J Thorac Oncol 2014;9:818-24.
25. Choi YS, Shim YM, Kim Y, Kim K. Mediastinoscopy in patients with clinical stage I non-small cell lung cancer. Ann Thorac Surg 2003;75:364-6.
26. Graham ANJ, Chan KJM, Pastorino U, Goldstra P. Systematic nodal dissection in the intrathoracic staging of patients with non-small cell lung cancer. J Thorac Cardiovasc Surg 1999; 117:246-51.
27. Fibla JJ, Molins L, Simon C, et al. The yield of mediastinoscopy with respect to lymph node size, cell type, and the location of the primary tumor. J Thorac Oncol 2006;1:430-3.
28. Asamura H, Nakayama H, Kondo H, et al. Lymph node involvement, recurrence, and prognosis in resected, small, peripheral, non-small cell lung carcinomas: are these carcinomas candidates for video-assisted lobectomy? J Thorac Cardiovasc Surg 1996;111:1125-34.
29. Suzuki K, Nagai K, Yoshida J, et al. Clinical predictors of N2 disease in the setting of a negative computed tomographic scan in patients with lung cancer. J Thorac Cardiovasc Surg 1999;117:593-8.
30. Takamochi K, Nagai K, Suzuki K, et al. Clinical predictors of N2 disease in non-small cell lung cancer. Chest 2000;117: 1577-82.
31. Swets JA. Measuring the accuracy of diagnostic systems. Science 1998;240:1285-93.
32. Travis WD, Brambilla E, Nicholson AG, et al. The 2015 World Health Organization classification of lung tumors. J Thorac Oncol 2015;10:1243-60.
33. Detterbeck F, Puchalski J, Rubinowitz A, Cheng D. Classification of the thoroughness of mediastinal staging of lung cancer. Chest 2010; 137:436-42.
34. Kinsey CM, Arenberg DA. Endobronchial ultrasound-guided transbronchial needle aspiration for non-small cell lung cancer staging. Am J Respir Crit Care Med 2014;189:640-9.
35. Dietrich CF, Jenssen C, Herth FJ. Endobronchial ultrasound elastography. Endosc Ultrasound 2016;5:233-8.
36. Brea TP, Raviña AR, Villamor JMC, et al. Use of magnetic resonance imaging for N-staging in patients with non-small cell lung cancer. A systematic review. Arch Bronconeumol 2019;55:9-16. doi: 10.1016/j.arbres.2018.03.007.