Exhaled breath to screen for malignant pleural mesothelioma: a validation study

Kevin Lamote[1,2], Matthijs Vynck[3], Olivier Thas[3,4], Joris Van Cleemput[5], Kristiaan Nackaerts[6] and Jan P. van Meerbeeck[2,7]

Affiliations: [1]Dept of Respiratory Medicine, Ghent University Hospital, Ghent, Belgium. [2]Dept of Internal Medicine, Ghent University, Ghent, Belgium. [3]Dept of Mathematical Modelling, Statistics and Bio-informatics, Ghent University, Ghent, Belgium. [4]National Institute for Applied Statistics Research Australia (NIASRA), University of Wollongong, Keiraville, Australia. [5]Occupational Health Service, Eternit N.V., Kapelle-op-den-Bos, Belgium. [6]Dept of Respiratory Diseases, KU Leuven, University Hospitals Leuven, Leuven, Belgium. [7]Thoracic Oncology, Multi-disciplinary Oncological Center Antwerp (MOCA), Antwerp University Hospital, Edegem, Belgium.

Correspondence: K. Lamote, Dept of Respiratory Medicine, Ghent University Hospital, Building 7K12IE, De Pintelaan 185, Ghent 9000, Belgium. E-mail: kevin.lamote@ugent.be

@ERSpublications
Breath analysis can be used to screen for malignant pleural mesothelioma in high-risk asbestos-exposed persons http://ow.ly/GppL30gCOaD

Cite this article as: Lamote K, Vynck M, Thas O, *et al*. Exhaled breath to screen for malignant pleural mesothelioma: a validation study. *Eur Respir J* 2017; 50: 1700919 [https://doi.org/10.1183/13993003.00919-2017].

ABSTRACT Malignant pleural mesothelioma (MPM) is predominantly caused by asbestos exposure and has a poor prognosis. Breath contains volatile organic compounds (VOCs) and can be explored as an early detection tool. Previously, we used multicapillary column/ion mobility spectrometry (MCC/IMS) to discriminate between patients with MPM and asymptomatic high-risk persons with a high rate of accuracy. Here, we aim to validate these findings in different control groups.

Breath and background samples were obtained from 52 patients with MPM, 52 healthy controls without asbestos exposure (HC), 59 asymptomatic former asbestos workers (AEx), 41 patients with benign asbestos-related diseases (ARD), 70 patients with benign non-asbestos-related lung diseases (BLD) and 56 patients with lung cancer (LC).

After background correction, logistic lasso regression and receiver operating characteristic (ROC) analysis, the MPM group was discriminated from the HC, AEx, ARD, BLD and LC groups with 65%, 88%, 82%, 80% and 72% accuracy, respectively. Combining AEx and ARD patients resulted in 94% sensitivity and 96% negative predictive value (NPV). The most important VOCs selected were P1, P3, P7, P9, P21 and P26.

We discriminated MPM patients from at-risk subjects with great accuracy. The high sensitivity and NPV allow breath analysis to be used as a screening tool for ruling out MPM.

This article has supplementary material available from erj.ersjournals.com

Received: May 04 2017 | Accepted after revision: Sept 24 2017

Support statement: This study was funded by the Belgian Foundation against Cancer (grants STK 2010-205 and STK 2012-223) and the Emmanuel Van der Schueren Grant from the Flemish League against Cancer. Funding information for this article has been deposited with the Crossref Funder Registry.

Conflict of interest: None declared.

Introduction

Malignant pleural mesothelioma (MPM) is a tumour that develops in the serosal lining of the thorax, and is predominantly linked to previous asbestos exposure [1, 2]. Despite the European ban on the use of asbestos in 2005 [3], asbestos is still being mined and imported to countries in need of industrial growth and therefore remains an important worldwide health issue [4]. The incidence of MPM is expected to increase further, fuelled by the amount of asbestos used in the past and the long mean latency period of 40–50 years between first exposure and incidence date. MPM is usually diagnosed at an advanced stage and requires a tissue sample of adequate size, often obtained by invasive biopsy [5, 6]. This delayed diagnosis results in a poor outcome with a best-observed median overall survival in highly selected patients of up to 18 months with the standard of care (platinum-based chemotherapy in combination with bevacizumab) [7, 8]. Together, this evidence highlights the importance of screening tools for early detection, where treatment options are believed not to be restricted to a palliative setting [9].

There are currently no uniformly agreed guidelines on MPM screening in asbestos-exposed persons [10] and this has resulted in an inappropriate and "blind" use of different imaging techniques and blood tests. Present research efforts have not yet resulted in a validated diagnostic blood biomarker like soluble mesothelin-related peptide (SMRP) or fibulin-3 [11–13]. A first-pass "rule out" test would help to separate asbestos-exposed persons at low risk (and not needing further screening) from those with early stage MPM requiring further follow-up and treatment. The analysis of volatile organic compounds (VOCs) in breath is a promising tool for this [14]. A proof-of-concept study using gas chromatography–mass spectrometry (GC-MS) identified cyclohexane as the discriminator in breath samples of patients with MPM, compared to those with occupational exposure to asbestos (AEx) and non-exposed healthy controls (HCs), with 97.4% accuracy [15]. Other series using pattern recognition with cross-reactive sensor technology discriminated patients with MPM from AEx subjects with acceptable accuracy [16, 17]. Using ion mobility spectrometry (IMS), patients with benign asbestos-related diseases (ARD) were discriminated from HCs with 99.9% accuracy [18] and we discriminated patients with MPM from HC and AEx with 82% and 87% accuracy, respectively [19]. The goal of the present study was to extend previous research, validate these earlier findings in a larger population and determine the specificity of VOC analysis for the detection of MPM compared to lung cancer.

Materials and methods

Study design and participants

This is a multicentre, cross-sectional, case-control study, approved by the Institutional Review Board of Ghent University Hospital (Belgian registration number B670201111954) and conducted in accordance with the Helsinki Convention. Participants gave written informed consent before inclusion. HCs, ARDs, benign lung diseases unrelated to asbestos exposure (BLD), primary lung cancer (LC) and MPM were randomly recruited from the respiratory medicine departments of the University Hospitals of Ghent, Leuven and Antwerp (Belgium) and the OLV Hospital in Aalst (Belgium). Other patients with ARD and asymptomatic persons with AEx were recruited *via* the occupational health service of a company using asbestos until 1997. MPM cases were confirmed by the Belgian Mesothelioma Pathology Panel. MPM and LC patients had to be treatment-naïve and were included between diagnosis and start of treatment. Benign asbestos-related diseases were not allowed to be present in any of the control groups except for ARD. A recent computed tomography (CT) scan or chest radiograph (<12 months) was mandatory to confirm medical conditions. Participants refrained from eating, drinking and smoking for at least 2 h before breath sampling. Confirmation of inclusion criteria, demographic information and previous asbestos exposure data were obtained using questionnaires. For patients, a detailed record of their medical condition had to be available.

Breath sampling and analysis

Breath samples were obtained between January, 2012 and December, 2014. A SpiroScout breath sampler (Ganshorn Medizin Electronic, Niederlauer, Germany) was connected to the sample loop of a BioScout multicapillary column/ion mobility spectrometer (MCC/IMS; B&S Analytik, Dortmund, Germany) [14, 19]. After resting for at least 10 min, all participants were asked to rinse their mouth with distilled water and put on rubber gloves and a nose clip. While sitting upright and without any forced breathing manoeuvres, they breathed normally through the SpiroScout mouthpiece, which was connected to a bacteria filter and the MCC/IMS sample loop. After 3 min, 10 mL of alveolar air was sampled and sent to the MCC/IMS for analysis. The breath analysis protocol is described in detail elsewhere [14, 19]. In brief, breath analytes were pre-separated by a non-polar OV-5 MCC column (Multichrom Ltd, Novosibirsk, Russia), then ionised by a 95MBq ^{63}Ni β-radiation source. Subsequently, the ionised breath compounds entered a 12 cm drift tube where a second separation took place based upon their ion mobility characteristics under the influence of an electrical field and a counter gas (α_1-nitrogen gas; 99.999% pure;

Air Liquide Medical, Schelle, Belgium). Finally, the VOCs collided on a Faraday plate detector, evoking an electrical current, which resulted in a VOC peak intensity (in volts, V) that correlated with VOC concentration. MCC/IMS is a technique that allows "pseudo-identification" of VOCs based on their retention time and ion mobility and cross-checking with an MCC/IMS database. However, for definite identification, the MCC/IMS data need to be crosschecked with GC-MS analysis of VOCs. After breath sampling, a background sample was taken using the same materials and sampling conditions. To rule out external contamination or sampling artefacts, we used disposable mouthpieces and filters and the MCC/IMS was constantly flushed with α_1-nitrogen gas. All unheated sample lines were made of Teflon (PTFE), which does not retain compounds [20]. The MCC/IMS was flushed with humid air between participants, to remove contaminants and ensure IMS chromatograms were clean.

Statistics

VisualNow v3.7 software (B&S Analytik, Dortmund, Germany) was used for VOC analysis as described previously [19]. In brief, raw IMS chromatograms were de-noised through baseline correction using a 5×3 low-pass filter, and aligned [19, 21]. The data were subsequently normalised to the reactant ion peak (RIP) and RIP-tailing was compensated for by subtracting a median spectrum from each chromatogram within the data set [19, 21]. Next, the data were smoothed and the chromatograms were inspected visually for the presence or absence of VOCs. If a VOC was present in either breath or background sample, these were manually selected and analysed (N=250), resulting in a list of VOC peak intensities (maximum peak height in the selected peak area).

To remove the effect of environmental chemical confounders, the alveolar gradient was calculated for every VOC by subtracting the standardised peak intensity in the background samples from that in the corresponding breath samples [22]. These alveolar gradient intensities (in V) were used as predictors in R v3.3.1 [23]. Because of the high dimensionality setting (large number of variables/low number of samples), penalised logistic regression (lasso) was used to discriminate MPM patients from those with AEx and HC. We used the *glmnet* R-package (v2.0-2) for fitting binomial lasso logistic models [24]. Using the predicted outcomes of all patients, we then constructed a receiver operating characteristic (ROC) curve and estimated sensitivity, specificity, positive (PPV) and negative predictive value (NPV), diagnostic accuracy of the final model, and 95% confidence intervals for all estimates. Furthermore, we looked at the number of times (folds) a VOC was selected by the lasso regressions. We opted to consider variables as important when they were selected in a large proportion of folds (>50%), as previously described [19]. Furthermore, since asbestos-exposed persons are at risk for MPM, we investigated whether patients with MPM could be discriminated from those with ARD and those in the pooled AEx and ARD groups. We also compared patients with MPM to those with BLD or LC, and patients with LC to participants in the HC, AEx and BLD groups.

Summary statistics were calculated for continuous variables. A Fisher's exact test was used to determine whether the categorical outcomes were equally likely. For continuous variables, a Kolmogorov–Smirnov test was performed to assess normality and, subsequently, an ANOVA or Kruskal–Wallis test was performed to assess differences of means or distributions. Bonferroni-adjusted $p < 0.05$ was considered statistically significant.

Results

Patient characteristics

In total, 330 participants were included in the study: 52 HCs, 59 with AEx (asymptomatic), 41 with ARD, 70 with BLD, 56 with primary LC, and 52 with MPM (table 1). The BLD group mainly comprised patients with COPD (40%), cystic fibrosis (21%) and pneumonia (14%). There were significantly more males in the MPM, AEx and ARD groups than in the other groups, and participants in the patient groups were significantly older than those in the HC and AEx groups. There were more current smokers in the AEx, ARD, BLD and LC groups than in the HC group.

Breath analysis

Patients with MPM could be discriminated from HCs with 65% accuracy, 89% sensitivity and 79% NPV (table 2, figure 1). We discriminated MPM from AEx with 88% accuracy, 87% sensitivity, 90% specificity, 88% PPV and 88% NPV. The area under the receiver operator characteristic curve (AUC_{ROC}) was 0.879. Patients with MPM were discriminated from those with ARD with 82% accuracy. The sensitivity, specificity, PPV and NPV were 89%, 73%, 81% and 83%, respectively. The AUC_{ROC} was 0.850. Furthermore, pooling the AEx and ARD groups allowed us to discriminate participants in the MPM group from those in the combined group with 85% accuracy, 94% sensitivity, 80% specificity, 71% PPV and 96% NPV. The AUC_{ROC} was 0.890. Patients with MPM were also discriminated well (80% accuracy) from those with BLD. The MPM patients were discriminated from LC patients with 72% accuracy, 73% sensitivity, 71% specificity, 70% PPV and 74% NPV. The AUC_{ROC} was 0.770 (table 2, figure 1).

TABLE 1 Baseline patient characteristics

Characteristic	Category						
	HC	AEx	ARD	BLD	LC	MPM	p-value
Subjects	52	59	41	70	56	52	
Gender							<0.001
Male	34 (65)	58 (98)	40 (98)	47 (67)	37 (66)	43 (83)	
Female	18 (35)	1 (2)	1 (2)	23 (33)	19 (34)	9 (17)	
Age years	51.2 (34.5–56.7)	53.2 (50.2–55.3)	58.3 (55.3–62.2)	58.8 (40.6–68.0)	69.9 (64.3–72.7)	67.3 (61.6–72.9)	<0.001
Weight kg	78.3±17.1	85.6±11.4	84.5±14.6	71.3±15.1	70.7±14.4	74.5±10.3	<0.001
Length m	1.76±0.09	1.77±0.06	1.74±0.05	1.71±0.09	1.68±0.08	1.72±0.08	<0.001
BMI kg·m^{-2}	25.2 (22.2–27.7)	26.9 (24.9–28.9)	26.8 (24.5–31.5)	24.4 (20.8–25.9)	24.0 (21.6–27.8)	25.4 (23.6–27.2)	<0.001
Smoking status							<0.001
Never smoked	35 (67)	19 (32)	15 (37)	24 (35)	6 (10)	19 (37)	
Current smoker	1 (2)	14 (24)	8 (20)	14 (20)	25 (45)	5 (9)	
Ex-smoker	16 (31)	26 (44)	18 (43)	31 (45)	25 (45)	28 (54)	
Smoking duration pack-years	0.0 (0.0–1.61)	6.0 (0.0–21.5)	5.3 (0.0–24.8)	7.5 (0.0–36.0)	30.0 (14.4–45.0)	2.65 (0.0–14.7)	<0.001
Subgroups							
Lung embolism				3			
COPD				28			
Fibrous tumour				1			
Allergy				1			
Cystic fibrosis				15			
Asthma				6			
Pneumonia				10			
Emphysema				3			
Pleuritis				1			
Pulmonary hypertension				1			
Granulomatosis with polyangiitis				1			
Epithelioid						36	
Sarcomatoid						3	
Biphasic/mixed						3	
Unknown						10	

Data are presented as n, n (%), median (Q1–Q3) or mean±SD unless otherwise stated. BMI: body mass index; COPD: chronic obstructive pulmonary disease; AEx: asymptomatic former asbestos-exposed individual; ARD: patients with benign asbestos-related diseases; BLD: patients with benign non-asbestos-related lung diseases; HC: healthy controls; LC: patients with primary lung cancer; MPM: patients with malignant pleural mesothelioma.

TABLE 2 Model characteristics for mesothelioma discrimination

Results	Comparison					
	MPM *versus* HC	**MPM *versus* AEx**	**MPM *versus* ARD**	**MPM *versus* AEx+ARD**	**MPM *versus* BLD**	**MPM *versus* LC**
Subjects n	52 *versus* 52	52 *versus* 59	52 *versus* 41	52 *versus* 100	52 *versus* 70	52 *versus* 56
Sensitivity %	88.5 (77.6–95.2)	86.5 (75.2–93.9)	88.5 (77.6–95.2)	94.2 (85.1–98.5)	71.2 (57.8–82.2)	73.1 (59.9–83.8)
Specificity %	42.3 (29.5–56.0)	89.8 (80.1–95.8)	73.2 (58.2–85.0)	80.0 (71.3–87.0)	87.1 (77.8–93.5)	71.4 (58.7–82.1)
PPV %	60.5 (49.3–71.0)	88.2 (77.2–95.1)	80.7 (69.0–89.4)	71.0 (59.6–80.8)	80.4 (67.2–90.0)	70.4 (57.3–81.4)
NPV %	78.7 (60.7–90.8)	88.3 (78.3–94.7)	83.3 (68.6–92.9)	96.4 (90.5–99.1)	80.3 (70.2–88.1)	74.1 (61.2–84.4)
Accuracy %	65.4 (55.9–74.0)	88.3 (81.3–93.3)	81.7 (72.9–88.6)	84.9 (78.5–89.9)	80.3 (72.6–86.7)	72.2 (63.3–80.0)
AUC_{ROC}[#]	0.612 (0.502–0.724)	0.879 (0.799–0.948)	0.850 (0.764–0.927)	0.890 (0.832–0.942)	0.837 (0.759–0.907)	0.770 (0.678–0.855)
VOCs[¶]	**P0**, **P4**, **P10**, **P15**, P66, P85, P88, P92, **P99**, **P103**, **P104**, **P108**, **P114**, **P119**, P170, P189, P192, P196, **P203**, **P207**, **P208**, **P212**, **P218**, **P223**	**P1**, **P3**, P7, **P9**, **P15**, **P21**, P22, **P26**, P65, **P66**, P73, P75, **P84**, P99, **P101**, **P110**, **P112**, **P114**, P118, **P120**, **P126**, **P132**, P133, **P137**, P176, **P177**, P184, P186, **P195**, P210, **P212**, **P221**, **P223**, **P225**, P229, **P231**, **P237**, **P243**, P244, **P248**	**P1**, **P9**, P15, **P21**, **P26**, **P34**, **P83**, P88, **P92**, **P94**, **P102**, P108, **P114**, **P119**, **P127**, **P176**, P181, **P185**, **P187**, **P195**, P201, **P207**, **P212**, **P220**	**P1**, **P7**, **P9**, **P15**, **P21**, **P26**, P70, P83, **P84**, **P88**, **P101**, P110, P118, **P122**, P123, P142, **P151**, **P153**, **P159**, **P161**, **P167**, **P173**, **P178**, P222, P235, **P236**, **P240**	**P1**, **P8**, P9, P15, **P42**, P98, P115, **P121**, **P123**, **P130**, P131, P137, **P164**, P220, P237, **P243**, **P245**	**P0**, **P7**, **P8**, **P9**, P15, **P21**, **P28**, **P37**, **P42**, **P43**, **P48**, P64, **P73**, **P78**, **P107**, **P108**, **P115**, **P116**, **P117**, P123, **P129**, **P136**, **P145**, **P150**, **P151**, **P156**, P172, **P181**, **P186**, **P215**, **P216**, **P223**, **P224**, **P225**, P231, **P237**, **P244**

Data are presented with 95% CI in parentheses. AEx: asymptomatic asbestos-exposed controls; ARD: patients with benign asbestos-related diseases; HC: healthy controls; MPM: malignant pleural mesothelioma; BLD: patients with benign non-asbestos-related lung diseases; LC: patients with primary lung cancer; NPV: negative predictive value; PPV: positive predictive value; AUC_{ROC}: area under the receiver operator characteristic curve; VOC: volatile organic compound. [#]: AUC_{ROC} significantly different from 0.5; [¶]: >50% of times selected (VOCs in bold are selected in >80% of folds).

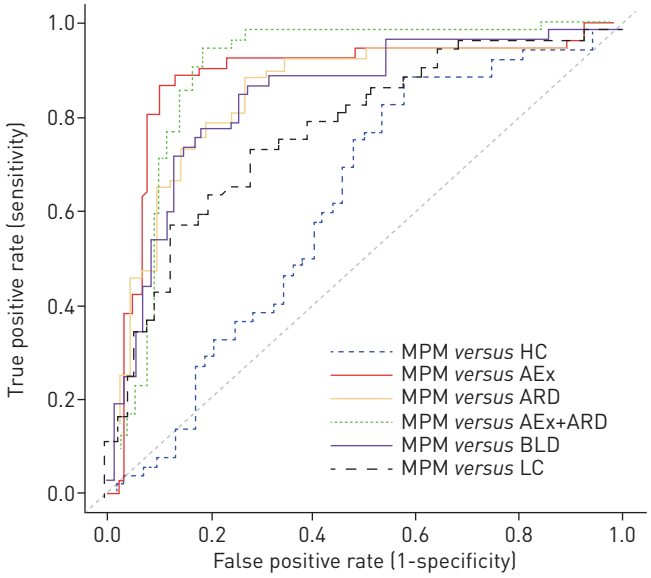

FIGURE 1 Receiver operating characteristic (ROC) curves for malignant pleural mesothelioma (MPM) discrimination. AEx: asymp-tomatic persons with past asbestos exposure; ARD: patients with benign asbestos-related diseases; BLD: patients with benign non-asbestos-related lung diseases; HC:healthy controls without occupationalasbestos exposure; LC: patients with lung cancer

We were not able to discriminate those with AEx from ARD controls, even when many VOCs were included in the models (table 3, figure 2). Participants with BLD were discriminated from those with AEx and ARD with 90% and 85% accuracy, respectively.

Patients with LC were discriminated from HCs with 71% accuracy, 77% sensitivity, 65% specificity, 71% PPV and 72% NPV. They were also discriminated from those with AEx and BLD with 90% and 71% accuracy, respectively (table 3, figure 2).

Lasso regression revealed that the most important VOCs for discriminating participants with MPM from those in the at-risk and BLD groups were P1, P3, P7, P9, P21, and P26 (table 3, table 4, supplementary figure S1). These were not selected when discriminating between participants in the MPM and HC groups (table 2), highlighting the importance of these VOCs as markers for MPM. These VOCs were also used to discriminate patients with MPM from those with LC, but not with great accuracy, suggesting a more common VOC signature.

Discussion

In this multicentre, cross-sectional, case–control study, we showed that breath analysis using MCC/IMS discriminated between participants with MPM and those with AEx, ARD and BLD, with clinically relevant accuracy. The lifetime risk for MPM in persons with occupational asbestos exposure is approximately 10% and can be higher in those who directly processed asbestos fibres [25]. We investigated whether persons with MPM could be discriminated from those with known exposure to asbestos (healthy or with benign asbestos-related stigmata) and confirmed that they could, with 88% and 82% accuracy, respectively. Because we could not discriminate between both asbestos-exposed groups, we examined the screening capability of the breath test after combining the AEx and ARD groups. The resulting accuracy of discrimination between persons in the combined group and those with MPM was 85%, sensitivity 94% and NPV 96%. Although screening studies for diagnostic purposes typically require a large specificity [26], PPV and NPV are more clinically meaningful because their interpretation is more straightforward [27]. Therefore, the high sensitivity and NPV from our comparisons make them a powerful tool for ruling out the disease in a true negative population. This will exclude persons from further investigations and could help diagnose MPM in a more cost-effective manner. Compared to blood biomarkers, our results outperform those of SMRP and fibulin-3 [11, 13]. A meta-analysis showed that, at 95% sensitivity, the specificity of SMRP was only 22%, limiting its use in ruling out disease, and, when using a high specificity of 95% to rule in the diagnosis in high-risk individuals, its sensitivity of 33% fell too short [11]. Furthermore, a recent meta-analysis of fibulin-3 showed only a modest discrimination between patients with MPM and cancer-free controls, with 62% sensitivity and 82% specificity [13].

TABLE 3 Model characteristics for lung cancer discrimination

Results	Comparison					
	LC *versus* HC	**LC *versus* AEx**	**LC *versus* BLD**	**AEx *versus* ARD**	**AEx *versus* BLD**	**ARD *versus* BLD**
Subjects n	56 *versus* 52	56 *versus* 59	56 *versus* 70	59 *versus* 41	59 *versus* 70	41 *versus* 70
Sensitivity %	76.8 (64.5–86.4)	89.3 (79.1–95.5)	64.3 (51.2–76.0)	82.9 (69.2–92.2)	88.6 (79.5–94.5)	88.6 (79.5–94.5)
Specificity %	65.4 (51.8–77.3)	89.8 (80.1–95.8)	77.1 (66.3–85.8)	35.6 (24.2–48.4)	91.5 (82.3–96.8)	78.0 (63.6–88.7)
PPV %	70.5 (58.2–80.9)	89.3 (79.1–95.5)	69.2 (55.8–80.6)	47.2 (36.0–58.7)	92.5 (84.3–97.2)	87.3 (78.1–93.6)
NPV %	72.3 (58.4–83.7)	89.8 (80.1–95.8)	73.0 (62.1–82.1)	75.0 (56.7–88.3)	87.1 (77.0–93.8)	80.0 (65.6–90.2)
Accuracy %	71.3 (62.3–79.2)	89.6 (83.0–94.2)	71.4 (63.1–78.8)	55.0 (45.2–64.5)	89.9 (83.8–94.3)	84.7 (77.1–90.5)
AUC_{ROC}	0.752 (0.659–0.839)[#]	0.936 (0.884–0.976)[#]	0.724 (0.630–0.813)[#]	0.522 (0.366–0.591)	0.957 (0.917–0.988)[#]	0.855 (0.766–0.930)[#]
VOCs[¶]	**P4**, P7, **P8**, **P10**, P23, **P28**, P43, **P55**, P59, **P76**, **P83**, **P107**, **P112**, P115, **P116**, **P118**, P131, **P136**, P151, **P163**, P167, **P184**, **P191**, **P215**, **P220**, P223, **P224**, **P226**, P239, **P244**	P0, **P1**, **P3**, P14, **P21**, **P26**, P43, **P61**, **P65**, **P66**, P72, P84, **P88**, P90, **P101**, P112, **P114**, P115, **P116**, **P118**, P129, **P136**, **P141**, P158, **P176**, **P180**, **P181**, P187, P199, **P203**, P205, **P216**, P227, **P229**, P230, **P231**, **P233**, P244	**P0**, P1, **P42**, P44, P107, **P125**, **P126**, P127, **P168**, P170, P233	P1, **P3**, P20, P23, P26, P34, P37, P44, P66, P69, P70, P80, P83, P84, **P90**, P92, **P99**, P101, **P103**, P120, P123, P126, P134, P137, P144, P166, P169, P170, P180, P183, P184, **P190**, P192, **P199**, P201, **P203**, P223, P226, P234, P237, P244	**P1**, **P3**, **P21**, **P42**, P50, **P84**, **P87**, **P88**, P97, **P101**, P104, **P128**, **P130**, **P132**, **P150**, **P171**, P179, **P213**, **P216**, **P217**, P226, **P230**, P233	**P21**, P25, **P42**, **P87**, **P88**, **P101**, **P110**, **P132**, P136, **P153**, **P198**, **P199**, **P212**, **P221**, **P243**, P247

Data are presented with 95% CI in parentheses. AEx: asymptomatic former asbestos-exposed controls; ARD: patients with benign asbestos-related diseases; HC: healthy controls; LC: patients with primary lung cancer; BLD: patients with benign non-asbestos-related lung diseases; NPV: negative predictive value; PPV: positive predictive value; AUC_{ROC}: area under the receiver operator characteristic curve; VOC: volatile organic compound. [#]: AUC_{ROC} significantly different from 0.5; [¶]: >50% of times selected (VOCs in bold are selected in >80% of folds).

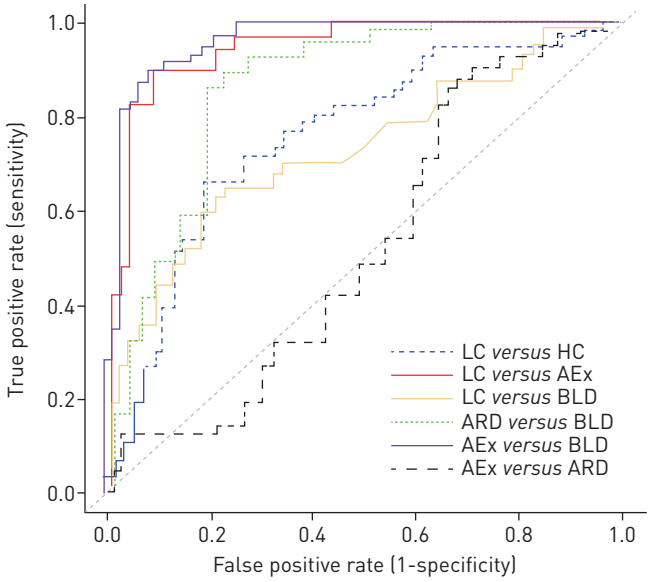

FIGURE 2 Receiver operating characteristic (ROC) curves for lung cancer discrimination. AEx: asymp-tomatic persons with past asbestos exposure; ARD: patients with benign asbestos-related diseases; BLD: patients with benign non-asbestos-related lung diseases; HC: healthy controls without occupational asbestos exposure; LC: patients with lung cancer

The strength of the present study lies in its inclusion and comparison of multiple control groups from a large number of participants. It therefore serves as the final proof-of-concept by confirming and validating all previous research on breath analysis for MPM screening in small sample sizes. DE GENNARO *et al.* used GC-MS to discriminate persons with MPM from those with AEx and from HCs with 97.4% accuracy, and identified cyclohexane as a marker for MPM [15]. Using pattern recognition, DRAGONIERI *et al.* distinguished the same groups with 80.8% and 84.6% accuracy, respectively [16] and CHAPMAN *et al.* discriminated MPM patients from HC subjects and ARD patients with 88% accuracy [17]. However, these studies using pattern recognition did not identify VOCs because of a technical drawback of electronic noses. Using MCC/IMS, CAKIR *et al.* discriminated patients with ARD from HCs with 96% sensitivity and 50% specificity and identified α-pinene and 4-ethyltoluol as markers for asbestos-related diseases [18], where we previously discriminated persons with MPM from those with AEx and from HCs with 87% and 82% accuracy, respectively [19]. We identified P3, P5, P50 and P71 as the most important VOCs in these discriminations, of which P3 was confirmed in the present study. However, there was a modest discrimination between patients with MPM and those with LC, a characteristic also observed with SMRP [11]. This could be due to VOCs reflecting underlying inflammation. Because inflammation is a hallmark of cancer [28], these VOCs could be general markers of cancer rather than tumour-specific markers. However, the modest discrimination between MPM and LC patients, and the difference in VOCs used for this discrimination and those used to discriminate MPM from the AEx and ARD groups, suggests that at least some VOCs might be able to discriminate between different tumour types.

Despite these satisfying results, we acknowledge some limitations. First, this study was not randomised, and the groups were not matched for age, gender or smoking status. Patients with MPM and LC were significantly older than those in the other groups, which could be explained by the latency period between the first exposure to the causal agent and the diagnosis of these diseases, and by the fact that age-matched healthy controls without significant comorbidities are hard to find. Although some studies suggest that aging affects human metabolism and VOCs [29–31], other groups did not find this correlation [32–34]. Furthermore, a higher incidence of males was seen in the groups with asbestos exposure. This may be because the asbestos industry had a male predominance. Similarly, the difference in smoking status could arise from the higher incidence of current smokers in this blue-collar industry [35, 36]. Furthermore, because smoking is the leading cause of lung cancer, we expected the LC group to have the highest incidence of current smokers. However, we do not believe smoking had any impact on the modelling, considering that the pathogenesis of MPM is independent of smoking. This is further strengthened by the fact that we could not satisfactorily discriminate between patients in the MPM and LC groups. Furthermore, we did not include patients with secondary malignant pleural effusions, which can impede the differential diagnosis of MPM. Hence, important compounds that could be used to differentiate

TABLE 4 Peak characteristics

Peak	Peak characteristics		Alveolar gradient V						
	RT s	$1/K_0$ V·cm^{-2}	HC	AEx	ARD	BLD	LC	MPM	p-value[#]
P0	282.1	0.668	0.0019 (−0.0003;0.0066)	0.0002 (−0.0008;0.0012)	0.0000 (−0.0022;0.0017)	−0.0005 (−0.0019;0.0010)	0.0007 (−0.0002;0.0031)	−0.0005 (−0.0018;0.0016)	<0.001
P1	5.9	0.503	0.0354 (0.0070;0.0597)	0.0948 (0.669;0.1277)	0.0904 (0.0405;0.1274)	0.0746 (0.0482;0.1043)	0.0507 (0.0261;0.0665)	0.0350 (0.0179;0.0688)	<0.001
P3	6.7	0.544	0.0333 (0.0003;0.0786)	0.0766 (0.0622;0.1026)	0.0588 (0.0370;0.0922)	0.0416 (0.0073;0.0660)	0.0383 (0.0030;0.0669)	0.0363 (0.0039;0.0672)	<0.001
P7	6.6	0.578	0.0158 (0.0064;0.0265)	0.0168 (0.0093;0.0323)	0.0202 (0.0109;0.0288)	0.0108 (0.0015;0.0275)	0.0068 (−0.0012;0.0272)	0.0083 (−0.0038;0.0297)	NS
P9	1.6	0.601	0.0083 (0.0005;0.0270)	0.0032 (−0.0002;0.0067)	0.0045 (0.0005;0.0068)	−0.0007 (−0.0121;0.0044)	0.0018 (−0.0052;0.0052)	0.0070 (0.0021;0.0192)	<0.001
P15	4.5	0.715	−0.0012 (−0.0036;0.0027)	−0.0027 (−0.0050;0.0015)	−0.0022 (−0.0061;0.007)	0.0020 (−0.0015;0.0040)	−0.0042 (−0.0091;−0.0001)	−0.0053 (−0.0090;−0.0017)	<0.001
P21	1.2	0.514	−0.0250 (−0.0661;−0.032)	−0.0039 (−0.0481;0.0420)	−0.0341 (−0.0540;0.0201)	−0.0518 (−0.0768;−0.0201)	−0.0464 (−0.0679;−0.0175)	−0.0499 (−0.0853;−0.0258)	NS
P26	4.2	0.689	0.0040 (−0.0024;0.0121)	0.0045 (0.0007;0.0186)	0.0063 (0.0016;0.0245)	0.0055 (0.0028;0.0073)	0.0036 (−0.0005;0.0068)	0.0024 (0.0001;0.0117)	<0.001
P42	4.5	0.457	−0.1371 (−0.2189;−0.0841)	−0.2069 (−0.2878;−0.0951)	−0.2364 (−0.3166;−0.0887)	0.0035 (−0.0543;0.0485)	−0.1551 (−0.2243;−0.0387)	−0.2212 (−0.2969;−0.0912)	<0.001
P66	66.5	0.733	−0.0057 (−0.0125;0.0003)	−0.0149 (−0.0206;−0.0080)	−0.0111 (−0.0201;−0.0066)	−0.0052 (−0.0114;−0.0024)	−0.0124 (−0.0178;−0.0070)	−0.0100 (−0.0209;−0.0039)	<0.001
P83	160.6	0.764	−0.0009 (−0.0026;0.0003)	−0.0028 (−0.0044;−0.0016)	−0.0041 (−0.0048;−0.0015)	−0.0011 (−0.0023;−0.0001)	−0.0027 (−0.0047;−0.0010)	−0.0022 (−0.0038;−1.0006)	<0.001
P84	116.1	0.742	0.0001 (−0.0006;0.0008)	−0.0010 (−0.0024;−0.0004)	−0.0002 (−0.0015;0.0004)	0.0003 (−0.0004;0.0009)	−0.0001 (−0.0009;0.0007)	−0.0002 (−0.0011;0.0007)	<0.001
P88	5.5	0.657	0.0013 (−0.0003;0.0037)	0.0011 (−0.0011;0.0027)	0.0014 (−0.0002;0.0032)	−0.0007 (−0.0025;0.0022)	0.0004 (−0.0049;0.0023)	0.0009 (−0.0030;0.0032)	NS
P101	20.0	0.716	−0.0007 (−0.0037;0.0025)	−0.0122 (−0.0216;−0.0028)	−0.0100 (−0.0166;−0.0025)	−0.0033 (−0.0062;−0.0001)	−0.0042 (−0.0072;−0.0010)	−0.0008 (−0.0041;0.0002)	<0.001
P114	99.8	0.728	−0.0001 (−0.0009;0.0007)	−0.0009 (−0.0023;−0.0002)	−0.0004 (−0.0017;0.0001)	−0.0001 (−0.0008;0.0004)	−0.0005 (−0.0011;0.0002)	−0.0003 (−0.0014;0.0004)	0.025
P115	98.7	0.665	0.0004 (−0.0005;0.0012)	−0.0009 (−0.0029;0.0002)	−0.0009 (−0.0031;−0.0002)	0.0000 (−0.0010;0.0007)	−0.0004 (−0.0012;0.0007)	−0.0006 (−0.0015;0.0005)	<0.001
P118	184.6	0.639	−0.0001 (−0.0006;0.0005)	−0.0006 (−0.0014;0.0000)	−0.0005 (−0.0011;0.0003)	−0.0003 (−0.0007;0.0003)	−0.0003 (−0.0010;0.0004)	−0.0001 (−0.0007;0.0007)	NS
P123	67.1	0.759	−0.0011 (−0.0030;0.0005)	−0.0049 (−0.0091;−0.0023)	−0.0044 (−0.0074;−0.0027)	−0.0019 (−0.0039;−0.0008)	−0.0042 (−0.0074;−0.0016)	−0.0033 (−0.0046;−0.0014)	<0.0001
P136	34.6	0.910	0.0002 (−0.0009;0.0009)	−0.0004 (−0.0007;0.0004)	−0.0002 (−0.0006;0.0003)	0.0001 (−0.0004;0.0006)	−0.0003 (−0.0009;0.0004)	−0.0004 (−0.0010;0.0003)	NS
P212	138.1	0.771	−0.0001 (−0.0005;0.0008)	−0.0003 (−0.0016;0.0007)	−0.0004 (−0.0014;0.0001)	−0.0002 (−0.0006;0.0003)	0.0001 (−0.0005;0.0005)	0.0002 (−0.0009;0.0009)	NS
P223	309.1	0.601	−0.0005 (−0.0012;0.0001)	0.0000 (−0.0005;0.0005)	0.0002 (−0.0006;0.0006)	0.0001 (−0.0005;0.0008)	−0.0001 (−0.0005;−0.0005)	−0.0001 (−0.0009;0.0005)	NS
P237	4.5	0.700	0.0020 (−0.0004;0.0046)	0.0018 (−0.0007;0.0038)	0.0025 (−0.0006;0.0055)	0.0012 (−0.0004;0.0034)	0.0019 (−0.0001;0.0043)	0.0015 (−0.0005;0.0061)	NS
P244	13.5	0.917	−0.0001 (−0.0009;0.0008)	−0.0004 (−0.0013;0.0003)	−0.0001 (−0.0008;0.0002)	−0.0004 (−0.0010;0.0002)	−0.0006 (−0.0017;0.0002)	−0.0004 (−0.0012;0.0009)	NS

Data as presented as median (Q1;Q3). RT: retention time; $1/K_0$: inverse reduced ion mobility; AEx: asymptomatic former asbestos workers; ARD: patients with benign asbestos-related diseases; BLD: patients with benign, non-asbestos-related diseases; HC: healthy control participants without asbestos exposure; LC: patients with lung cancer; MPM: patients with malignant pleural mesothelioma; [#]: Kruskal–Wallis test.

patients with MPM from those with secondary malignant effusions could be missed. Future validation research should include this group of patients to optimise screening.

Second, we cannot fully exclude the possibility that external VOCs could have influenced the breath samples, even after correcting for this using background samples. Depending on their kinetics, inhaled VOCs can be stored in the body's fat compartments and slowly released over time [37, 38]. Although we tried to counteract environmental contamination as much as possible by using inert sampling materials and by calculating the alveolar gradient of the VOCs [22], this may not be sufficient to completely remove the impact of environmental confounders. However, as our patients and HCs were randomly recruited (hence background samples were also random), the effects of contamination were excluded as much as possible, and can be expected to be minimal.

Last, our selected VOCs have not yet been identified and are not included in the MCC/IMS VOC library. For definite identification of the VOCs, the MCC/IMS data needs to be crosschecked against additional GC-MS analysis of the suspected VOCs. This has no impact on the accuracy of discrimination, but further identification of the VOCs should allow us to link these to the pathogenesis of MPM and serve as additional proof. Therefore, GC-MS analysis of VOCs in breath and in the headspace of MPM cell lines is advocated.

In summary, we have shown that MCC/IMS allows adequate discrimination of patients with MPM from at-risk individuals. The present study describes an extension cohort of our previous proof-of-principle study. Future research should now focus on the external validation in a prospective, case-control series in independent patient cohorts, with blinding of the investigator to the underlying pathology, and with follow-up of at-risk subjects over time. This will ultimately identify the VOCs that reflect the transition from chronic inflammation towards malignant transformation and allow assessment of the clinical utility of the breath test [39].

Conclusion

Using breath analysis with MCC/IMS, we discriminated patients with MPM from high-risk participants previously exposed to asbestos, and patients with benign asbestos-related and non-related lung diseases. Validating these results in an independent, blinded prospective study will allow assessment of the clinical utility of breath analysis for MPM screening in persons previously exposed to asbestos. The high sensitivity and NPV observed mean this can be used as a step-up tool in MPM screening, whereby only asbestos-exposed individuals with an aberrant VOC signature should be further investigated for the possible presence of MPM.

Acknowledgements

The authors wish to thank Prof. K. Tournoy (Dept of Respiratory Medicine, "Onze-Lieve-Vrouw" Hospital, Aalst, Belgium) for his help in patient recruitment. Part of this study was presented at the ERS International Congress in Milan, September 9–13, 2017.

Kevin Lamote had full access to the data in the study and takes responsibility for its integrity. Kevin Lamote, Matthijs Vynck and Olivier Thas take responsibility for the accuracy of the data analysis. Joris Van Cleemput, Kristiaan Nackaerts and Jan P. van Meerbeeck helped with the inclusion of the participants. Kevin Lamote drafted the manuscript. All authors contributed to the design of the manuscript and gave their approval for final submission. Kevin Lamote is the guarantor of the content of the manuscript.

References

1 Peto J, Decarli A, La Vecchia C, *et al.* The European mesothelioma epidemic. *Br J Cancer* 1999; 79: 666–672.
2 Hiddinga BI, Rolfo C, van Meerbeeck JP. Mesothelioma treatment: are we on target? A review. *J Adv Res* 2015; 6: 319–330.
3 European Commission. Commission Directive 1999/77/EC of July 26, adapting to technical progress for the sixth time Annex I to Council Directive 76/769/EEC on the approximation of the laws, regulations and administrative provisions of the Member States relating to restrictions on the marketing and use of certain dangerous substances and preparations (asbestos) (Text with EEA relevance). *Off J Eur Commun* 1999; L207: 18–20. http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31999L0077&from=EN Date last accessed: November 20, 2017.
4 Ogunseitan OA. The asbestos paradox: global gaps in the translational science of disease prevention. *Bull World Health Organ* 2015; 93: 359–360.
5 Rodriguez Panadero F. Diagnosis and treatment of malignant pleural mesothelioma. *Arch Bronconeumol* 2015; 51: 177–184.
6 Zhang W, Wu X, Wu L, *et al.* Advances in the diagnosis, treatment and prognosis of malignant pleural mesothelioma. *Ann Transl Med* 2015; 3: 182.
7 Robinson BW, Lake RA. Advances in malignant mesothelioma. *N Engl J Med* 2005; 353: 1591–1603.
8 Zalcman G, Mazieres J, Margery J, *et al.* Bevacizumab for newly diagnosed pleural mesothelioma in the Mesothelioma Avastin Cisplatin Pemetrexed Study (MAPS): a randomised, controlled, open-label, phase 3 trial. *Lancet* 2016; 387: 1405–1414.

9 National Lung Screening Trial Research Team, Aberle DR, Adams AM, *et al.* Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011; 365: 395–409.
10 Baas P, Fennell D, Kerr KM, *et al.* Malignant pleural mesothelioma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2015; 26: Suppl 5, v31–v39.
11 Hollevoet K, Reitsma JB, Creaney J, *et al.* Serum mesothelin for diagnosing malignant pleural mesothelioma: an individual patient data meta-analysis. *J Clin Oncol* 2012; 30: 1541–1549.
12 Pass HI, Levin SM, Harbut MR, *et al.* Fibulin-3 as a blood and effusion biomarker for pleural mesothelioma. *N Engl J Med* 2012; 367: 1417–1427.
13 Pei D, Li Y, Liu X, *et al.* Diagnostic and prognostic utilities of humoral fibulin-3 in malignant pleural mesothelioma: evidence from a meta-analysis. *Oncotarget* 2017; 8: 13030–13038.
14 Lamote K, Nackaerts K, van Meerbeeck JP. Strengths, weaknesses, and opportunities of diagnostic breathomics in pleural mesothelioma-a hypothesis. *Cancer Epidemiol Biomarkers Prev* 2014; 23: 898–908.
15 de Gennaro G, Dragonieri S, Longobardi F, *et al.* Chemical characterization of exhaled breath to differentiate between patients with malignant pleural mesothelioma from subjects with similar professional asbestos exposure. *Anal Bioanal Chem* 2010; 398: 3043–3050.
16 Dragonieri S, van der Schee MP, Massaro T, *et al.* An electronic nose distinguishes exhaled breath of patients with malignant pleural mesothelioma from controls. *Lung Cancer* 2012; 75: 326–331.
17 Chapman EA, Thomas PS, Stone E, *et al.* A breath test for malignant mesothelioma using an electronic nose. *Eur Respir J* 2012; 40: 448–454.
18 Cakir Y, Métrailler L, Baumbach JI, *et al.* Signals in asbestos related diseases in human breath – preliminary results. *Int J Ion Mobil Spectrom* 2014; 17: 87–94.
19 Lamote K, Vynck M, Van Cleemput J, *et al.* Detection of malignant pleural mesothelioma in exhaled breath by multicapillary column/ion mobility spectrometry (MCC/IMS). *J Breath Res* 2016; 10: 046001.
20 Ruzsanyi V, Baumbach JI, Sielemann S, *et al.* Detection of human metabolites using multi-capillary columns coupled to ion mobility spectrometers. *J Chromatogr A* 2005; 1084: 145–151.
21 Bader S, Urfer W, Baumbach JI. Preprocessing of ion mobility spectra by lognormal detailing and wavelet transform. *Int J Ion Mobility Spectrom* 2008; 11: 43–49.
22 Phillips M, Greenberg J, Sabas M. Alveolar gradient of pentane in normal human breath. *Free Radic Res* 1994; 20: 333–337.
23 R Core Team. R: a language and environment for statistical computing. Vienna, R Foundation for Statistical Computing, 2014. www.R-project.org Date last accessed: November 20, 2017.
24 Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 2010; 33: 1–22.
25 Pass HI, Carbone M. Current status of screening for malignant pleural mesothelioma. *Semin Thorac Cardiovasc Surg* 2009; 21: 97–104.
26 van Klaveren RJ, Oudkerk M, Prokop M, *et al.* Management of lung nodules detected by volume CT scanning. *N Engl J Med* 2009; 361: 2221–2229.
27 Altman DG, Bland JM. Diagnostic tests 2: predictive values. *BMJ* 1994; 309: 102.
28 Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011; 144: 646–674.
29 Mazzatenta A, Pokorski M, Di Giulio C. Real time analysis of volatile organic compounds (VOCs) in centenarians. *Respir Physiol Neurobiol* 2015; 209: 47–51.
30 Lechner M, Moser B, Niederseer D, *et al.* Gender and age specific differences in exhaled isoprene levels. *Respir Physiol Neurobiol* 2006; 154: 478–483.
31 Blanchet L, Smolinska A, Baranska A, *et al.* Factors that influence the volatile organic compound content in human breath. *J Breath Res* 2017; 11: 016013.
32 Peng G, Hakim M, Broza YY, *et al.* Detection of lung, breast, colorectal, and prostate cancers from exhaled breath using a single array of nanosensors. *Br J Cancer* 2010; 103: 542–551.
33 Poli D, Goldoni M, Corradi M, *et al.* Determination of aldehydes in exhaled breath of patients with lung cancer by means of on-fiber-derivatisation SPME-GC/MS. *J Chromatogr B Analyt Technol Biomed Life Sci* 2010; 878: 2643–2651.
34 Mazzone PJ, Hammel J, Dweik R, *et al.* Diagnosis of lung cancer by the analysis of exhaled breath with a colorimetric sensor array. *Thorax* 2007; 62: 565–568.
35 Vanloon AJM, Goldbohm RA, Vandenbrandt PA. Lung cancer: is there an association with socioeconomic status in The Netherlands. *J Epidemiol Community Health* 1995; 49: 65–69.
36 Smith DR. Tobacco smoking by occupation in Australia and the United States: a review of national surveys conducted between 1970 and 2005. *Ind Health* 2008; 46: 77–89.
37 Amann A, Mochalski P, Ruzsanyi V, *et al.* Assessment of the exhalation kinetics of volatile cancer biomarkers based on their physicochemical properties. *J Breath Res* 2014; 8: 016003.
38 Jia C, Yu X, Masiak W. Blood/air distribution of volatile organic compounds (VOCs) in a nationally representative sample. *Sci Total Environ* 2012; 419: 225–232.
39 Lyman GH, Moses HL. Biomarker tests for molecularly targeted therapies: laying the foundation and fulfilling the dream. *J Clin Oncol* 2016; 34: 2061–2066.