Can Tumour Antigens Act as Biomarkers for the Early Detection of Non-Small Cell Lung Cancer?

Eithar Mohamed 1, Daniel Fletcher 1*, Simon Hart 2 and Barbara-ann Guinn 1,*

1 Centre for Biomedicine, Hull York Medical School, University of Hull, Hull HU6 7RX, UK; e.a.mohamed-2019@hull.ac.uk (E.M.); d.m.fletcher-2018@hull.ac.uk (D.F.)
2 Respiratory Medicine, Hull York Medical School, University of Hull, Hull HU6 7RX, UK; simon.hart@hyms.ac.uk
* Correspondence: barbara.guinn@hyms.ac.uk; Tel.: +44-1482-466543

Simple Summary: Lung cancer (LC) is one of the most common and serious types of cancer. Unfortunately, it is not easy to detect in the early stages of the disease due to the absence of symptoms. Many patients have late-stage LC when they are diagnosed, and this is associated with limited treatment options and poor survival rates. To try to improve this, we have assessed which proteins in LC patients are recognised by the immune response and could be used to screen at-risk patients for LC before symptoms appear. We have shown that panels of blood and sputum biomarkers may offer the most effective way to improve early LC detection.

Abstract: Lung cancer (LC) is one of the leading causes of cancer-related deaths. Pulmonary nodules are one of the risk factors, and their discovery rate has been increasing due to enhanced performance of chest CT scans, but more than 90% are non-malignant, causing unnecessary stress to patients and costs to healthcare providers. Early diagnosis of LC is associated with a 5-year survival rate of up to 75% following surgical resection, but LC is often diagnosed late due to a lack of symptoms and poor 5-year survival rates as low as 10%. The cost of LC diagnosis is high, with 40% of it associated with benign lesions, which are difficult to differentiate from malignant lesions. Tumour-associated antigens (TAAs) may provide one way in which LC could be diagnosed early using minimally-invasive techniques, under their association with immune responses and specificity for disease. Here we discuss the potential of cancer-testis antigens (CTAs) to act as non-invasive biomarkers for the early detection of non-small cell lung cancer.

Keywords: lung cancer; nodules; CT scan; tumour-associated antigens (TAAs); early diagnosis; sensitivity; specificity; cancer testis antigens (CTA)

1. Introduction

Patients with lung cancer (LC) have poor survival rates predominantly due to late detection (Figure 1) by which time the disease has often spread to other organs of the body and treatment success is limited [1]. About three-quarters of patients present with advanced LC stages which eventually result in high mortality rates within three months of diagnosis [2]. Records have also shown that around 35% of LC patients are diagnosed immediately after emergency admission, while more than 90% of patients with LC are diagnosed at stage III or IV. In comparison, a high 5-year survival rate of up to 75% is seen amongst patients diagnosed in the early stages of LC (stages I and II), following surgical resection [3]. Early diagnosis of LC has the potential to significantly improve survival rates even with the use of existing treatment options. LC is divided into small cell lung carcinomas (SCLC) (10–15%) and non-small cell lung carcinomas (NSCLC) (80–85%) based on histopathology. SCLC grows quickly [4] and responds well to chemotherapy but patients often relapse.
Figure 1. Five-year survival and incidence of LC by stage. A total of 15 and 7% of LC patients are diagnosed at the early stages, i.e., stage I and II, respectively (pink line), with a high 5-year survival rate of 62% for females (orange bar) and 51% for males (dark blue bar) at stage I, while 19–48% of patients with a known stage are diagnosed at the later disease stage (stage III or IV); survival rates are conversely poor at 2–3% at stage IV. Data taken from 2016 to 2020. https://www.cancerresearchuk.org/about-cancer/lung-cancer/survival, accessed on 14 June 2021.

2. Molecular Pathology of NSCLC

NSCLC develops due to a variety of distinct somatic mutations occurring in a heterogeneous population of tumour progenitor cells. NSCLC can be further sub-classified and significant differences in the frequency of common mutations in those sub-classes have been reported (Figure 2).

Adenocarcinoma (ADC) arises from epithelial cells in the alveoli and bronchioles, while squamous cell carcinoma (SCC) develops from epithelial cells in the larger airways of the bronchi. ADC expresses biomarkers such as thyroid transcription factor 1 (TTF1) and keratin 7 (KRT7) [5], while SCC has an increase in expression of cytokeratin 5 and 6 (CTK5/6), SRY-box 2 (SOX2), and phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha (PIK3CA) mutations. ADC also differs between smokers and non-smokers, with the latter carrying a higher frequency of EGFR, anaplastic lymphoma kinase (ALK), and ROS proto-oncogene 1 (ROS1) mutations [6], while smokers show a high frequency of KRAS mutations. This suggests that LC in smokers and non-smokers follows different pathogenetic pathways of tumour development [7].
Figure 2. Common mutations in adenocarcinoma (ADC) and squamous cell carcinoma (SCC). KRAS, LKB1, and EGFR mutations are predominantly found in ADC while FGFR, PIK3CA, and PTEN are common in SCC. TP53 is common in both with 52% and 79% in ADC and SCC, respectively. Data taken from [5,8–10].

3. Tumour Antigens as Biomarkers for LC

At the moment, the diagnosis of LC is widely based on imaging techniques and biopsy/histopathology. The cost of LC diagnosis is high, and 40% of the cost is associated with the diagnosis of benign lesions [11] due to overlapping clinical and radiological features that make it difficult to differentiate benign from malignant disease. Thus, a high level of expertise is required for both imaging and histopathology. Non-invasive diagnostic tools such as biomarkers may aid in the diagnosis of LC [11]. The most commonly investigated serological markers (Table 1) in LC include CA125, carcinoembryonic antigen (CEA), and cytokeratin 19 fragments (CYFRA21-1) [12–14].

The most studied tumour-associated antigen (TAA) in LC is CYFRA21-1, which is overexpressed in NSCLC, mainly SCC. High expression is associated with a negative prognosis, as low levels correlate with both longer overall survival and failure-free survival \( (p < 0.0001 \text{ and } p = 0.0003) \), respectively [15]. Combined CYFRA21-1 with NSE and CEA have high specificity (96%) for LC diagnosis but low sensitivity for early LC (31%) depending on sample size; the sensitivity and specificity varied [16]. Thus, these panels are not suggested for early LC in clinical practice. Overexpression of these TAAs may act as prognostic biomarkers for therapy monitoring and relapses; however, their utilisation in clinical settings remains to be established. Biomarkers can be useful tools to evaluate effective treatment, monitor for disease recurrence after therapy, and enable prognostic prediction.
miRNAs combined with autoantibodies/TAAAs may aid in detecting early NSCLC [17], as the early stages have better survival rates and assist in decreasing the associated costs of the healthcare system.

Tumour-associated antigens (TAA) are highly expressed in LC, but they also have elevated expression in other benign lung diseases and therefore have low specificity [18,19]. Due to the increase in TAA expression in benign lung diseases, the standard cut-off of biomarkers has to be doubled [20] to maximise the diagnostic yield of LC when differentiating between malignant and benign disease. Such cut-off levels have no benefit in imaging studies when using methods such as computer tomography (CT) scans in patients suspected of LC. The positive predictive value (PPV) of the CT scan has a greater value compared to the PPV of biomarkers at standard cut-off levels. The PPV of biomarkers depends on the prevalence rate of lung carcinoma. High-score PPVs for tumour markers have been observed in patient populations with high prevalence rates [21].

4. Cancer-Testis Antigen Expression in NSCLC

Cancer-testis antigens (CTAs) are not expressed in healthy tissues except in immunologically protected sites that lack MHC class I expression, such as the placenta and testes [22]. However, CTAs are often aberrantly expressed in cancers such as melanoma, ovarian and oesophageal cancer, and LC [23] (Figure 3; Supplementary Table S1). Their reactivated expression in cancer cells is thought to be due to epigenetic mechanisms such as DNA demethylation and histone acetylation. This aligns with the observation that changes in CpG methylation patterns, which are established during embryonic development, can be altered in cancer cells [24]. Specifically, hypomethylation of DNA in tumours has been associated with the activation of genes typically only expressed in the germline, known as “cancer-germline” genes.

Tumours expressing CTAs can be divided into three groups based on the number of CTAs and the frequency of their expression. Tumours with high CTA expression (>50% of the CT antigen expression is at >20% frequency) include melanoma and NSCLC. Breast and prostate cancers are examples of tumours with moderate CTA expression, while leukaemia could be considered to have a low CTA expression. Over 90 CTAs have been found in LC [25], and the majority of NSCLC patient samples (79%) expressed at least one of the analysed CTAs, with CTA protein levels corresponding with gene transcription [20]. CTA transcription was significantly related to the pathological N and TNM stages. Patients with stage II–III lymph node metastasis exhibited a greater CTA expression rate than patients with stage I and no lymph node metastases [26]. Expression of CTAs is associated with the most advanced tumour stages and poor outcome in LC [27] (Table 2), suggesting that the overexpression of CTAs promotes LC progression and drives metastasis. However, recently, one CTA, AKAP4, was shown to serve as a valuable biomarker for the early detection of LC. AKAP4 belongs to the A-kinase anchor proteins that bind the Protein kinase A (PKA) regulatory subunit and functions to anchor PKA to specific cellular locations [28]. AKAP4 is highly accurate in distinguishing between NSCLC patients and controls. When comparing all 264 LC cases with all 135 controls, the area under the curve-receiver operating characteristic (AUC-ROC) was found to be 0.97. Moreover, when comparing 136 stage I NSCLC cases with the controls, the AUC increased to 0.98.

Additionally, when comparing all LC patients with 27 controls who had histologically confirmed benign lung nodules—a comparison of significant clinical importance—the AUC reached an even higher value of 0.98. Furthermore, AKAP4 expression levels were found to significantly increase with tumour stage but were independent of age, gender, smoking history, or cancer subtype [29]. Further studies are required to verify whether CTAs could aid in early LC detection.
Table 1. Antigens that are known to be expressed in LC.

<table>
<thead>
<tr>
<th>Gene Name (Symbol)</th>
<th>Function</th>
<th>Healthy Tissue</th>
<th>Expression in LC(s)</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Carcinoembryonic antigen (CEA)</strong></td>
<td>Glycoprotein involved in cell adhesion and signal transduction</td>
<td>Low expression in colon, appendix</td>
<td>High expression in all types at advanced stages</td>
<td>[30,31]</td>
</tr>
<tr>
<td><strong>Osteopontin (OPN)</strong></td>
<td>Cell survival and angiogenesis</td>
<td>Gall bladder, placenta, brain</td>
<td>High expression associated with poor prognosis</td>
<td>[30,32]</td>
</tr>
<tr>
<td><strong>Cytokeratin 19 fragments (CYFRA 21-1)</strong></td>
<td>Part of the cytoskeleton of epithelial cells</td>
<td>All epithelial cells</td>
<td>NSCLC mainly SCC. High expression is associated with a poor prognosis</td>
<td>[33]</td>
</tr>
<tr>
<td><strong>Neuron-specific enolase (NSE)</strong></td>
<td>Glycolytic enzyme involved in inflammatory and neurotrophic activity regulating neuronal growth, differentiation, survival and death</td>
<td>Brain, adrenal, lung</td>
<td>Preferred for SCLC but also NSCLC and a marker of metastasis</td>
<td>[34]</td>
</tr>
<tr>
<td><strong>Serum amyloid A (SAA)</strong></td>
<td>Secreted during acute inflammation, transports cholesterol to the liver, recruits immune cells to inflammatory sites</td>
<td>Housekeeping role in normal human tissues</td>
<td>All types. High expression in late stages</td>
<td>[35,36]</td>
</tr>
</tbody>
</table>
Figure 3. Expression of CTAs in NSCLC. NSCLC is considered to be a tumour type with high expression of CTAs. CAGE/CT26 and NANOS3 are widely expressed in most samples from patients with NSCLC (100% and 94.7%, respectively) while BAGE1, TDRD1 and SYCP1 had limited expression in less than 10% of NSCLC samples. Data taken from Gure et al. [27], image author's own.
5. The Potential of Biomarker Panels for LC Detection

We recently used the Preferred Reporting Items for Systematic Review (PRISMA) guidelines [37,38] to perform a systematic review of the literature [17] to address the research question—what are the most promising biomarkers for an early diagnosis of LC?

The protocol was registered with PROSPERO (CRD42022336488), and we screened seven literature databases (CINAHL, MEDLINE, PubMed, Scopus, Web of Science, Cochrane Library and Clinicaltrial.gov) from 1 January 1970 until 21st May 2023 using the following MeSH terms (cancer* or tumor* or tumour* or neoplasm* or carcinoma* or malignancy*) AND (lung* or pulmonary) AND (antigen* OR protein* OR RNA* OR ctDNA* OR miRNA* OR cell surface marker* OR inflammatory cell*) AND (early detection OR early diagnosis OR early biomarker OR early marker). The initial search, removal of duplicates, title and abstract screening, and full-text reviews were performed by two independent reviewers.

Table 2. CTAs expression in LC using the Kaplan–Meier plot website to determine the association between high and low CTA expression and survival. Data taken from https://kmplot.com/analysis/ accessed on 21 August 2023.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Probe Set</th>
<th>Survival (mo)</th>
<th>p-Value</th>
<th>Gene</th>
<th>Probe Set</th>
<th>Survival (mo)</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low ⌧</td>
<td>High ⌧</td>
<td></td>
<td></td>
<td>Low ⌧</td>
<td>High ⌧</td>
<td></td>
</tr>
<tr>
<td>TPX2</td>
<td>210052_s_at</td>
<td>96.2</td>
<td>42</td>
<td>&lt;1 × 10^{-16}</td>
<td>TSP50</td>
<td>220126_at</td>
<td>81</td>
</tr>
<tr>
<td>DNAJB11</td>
<td>223054_at</td>
<td>119.87</td>
<td>52</td>
<td>8.4 × 10^{-12}</td>
<td>CTAGX10-1</td>
<td>220957_at</td>
<td>79.27</td>
</tr>
<tr>
<td>MAGEA1</td>
<td>207325_x_at</td>
<td>86.27</td>
<td>48.6</td>
<td>1.4 × 10^{-11}</td>
<td>PAGX10-4</td>
<td>205564_at</td>
<td>76</td>
</tr>
<tr>
<td>SSX2IP</td>
<td>203015_s_at</td>
<td>91</td>
<td>52</td>
<td>2.8 × 10^{-11}</td>
<td>SSX3</td>
<td>211670_x_at</td>
<td>78.5</td>
</tr>
<tr>
<td>DDX12</td>
<td>213378_s_at</td>
<td>89</td>
<td>52</td>
<td>1.7 × 10^{-10}</td>
<td>SYCP1</td>
<td>206740_x_at</td>
<td>79.87</td>
</tr>
<tr>
<td>(DNAJB14)</td>
<td>222850_s_at</td>
<td>52</td>
<td>111</td>
<td>1.2 × 10^{-9}</td>
<td>NXF2/CT39</td>
<td>220257_x_at</td>
<td>79.87</td>
</tr>
<tr>
<td>MAGEA3</td>
<td>209942_x_at</td>
<td>86.27</td>
<td>49.97</td>
<td>2.7 × 10^{-9}</td>
<td>SSX1</td>
<td>206626_x_at</td>
<td>78</td>
</tr>
<tr>
<td>DDX11/KRG2</td>
<td>208149_x_at</td>
<td>88.7</td>
<td>54</td>
<td>1.1 × 10^{-8}</td>
<td>DNAJB4</td>
<td>203811_s_at</td>
<td>75.43</td>
</tr>
<tr>
<td>(GAGE3)/CT4.3</td>
<td>207663_x_at</td>
<td>89</td>
<td>54.2</td>
<td>1.1 × 10^{-7}</td>
<td>SLY-1/CT34</td>
<td>220284_at</td>
<td>76</td>
</tr>
<tr>
<td>MAGEA12</td>
<td>210467_x_at</td>
<td>84</td>
<td>52</td>
<td>2.7 × 10^{-7}</td>
<td>MAGEA2</td>
<td>214603_at</td>
<td>74</td>
</tr>
<tr>
<td>GAGE1/4/7/11</td>
<td>207086_x_at</td>
<td>88</td>
<td>56</td>
<td>6.0 × 10^{-7}</td>
<td>FATE/CT43</td>
<td>231573_at</td>
<td>86.27</td>
</tr>
<tr>
<td>TPTECT44</td>
<td>220205_at</td>
<td>80.03</td>
<td>59</td>
<td>1.3 × 10^{-5}</td>
<td>PATCH2</td>
<td>239768_x_at</td>
<td>69</td>
</tr>
<tr>
<td>SAGE</td>
<td>220793_at</td>
<td>79.5</td>
<td>56.5</td>
<td>2.0 × 10^{-5}</td>
<td>SSX2</td>
<td>216471_x_at</td>
<td>76</td>
</tr>
<tr>
<td>MAGEA10</td>
<td>210295_at</td>
<td>86.27</td>
<td>57.33</td>
<td>2.4 × 10^{-5}</td>
<td>SPO11/CT35</td>
<td>222259_s_at</td>
<td>76</td>
</tr>
<tr>
<td>DDX10/HRH-J8</td>
<td>204977_at</td>
<td>79.54</td>
<td>57</td>
<td>8.7 × 10^{-5}</td>
<td>(DNAJB13)</td>
<td>230936_at</td>
<td>70</td>
</tr>
<tr>
<td>NA88A/VENTXP1</td>
<td>216726_at</td>
<td>81.2</td>
<td>61.2</td>
<td>0.0001</td>
<td>PLI-1/KDM5B</td>
<td>211202_s_at</td>
<td>63</td>
</tr>
<tr>
<td>TEX15/CT42</td>
<td>221448_s_at</td>
<td>79.87</td>
<td>59</td>
<td>0.0001</td>
<td>LAGE1</td>
<td>215733_x_at</td>
<td>73.3</td>
</tr>
<tr>
<td>DNAJB2 (HSPF3)</td>
<td>202500_at</td>
<td>62</td>
<td>74</td>
<td>0.0002</td>
<td>TAF7L</td>
<td>220325_at</td>
<td>76</td>
</tr>
<tr>
<td>MORC1/CT33</td>
<td>220850_at</td>
<td>79.27</td>
<td>57</td>
<td>0.0003</td>
<td>TDRD1/CT41.1</td>
<td>221018_s_at</td>
<td>74</td>
</tr>
<tr>
<td>LDHCACT32/</td>
<td>207022_s_at</td>
<td>78</td>
<td>62.2</td>
<td>0.0004</td>
<td>PAGE-1</td>
<td>206897_at</td>
<td>73.2</td>
</tr>
<tr>
<td>DDX13 (SKIY2L)</td>
<td>203727_at</td>
<td>81</td>
<td>59.11</td>
<td>0.0004</td>
<td>MAGE-C2</td>
<td>215932_at</td>
<td>74</td>
</tr>
<tr>
<td>MAGE-C1</td>
<td>206609_at</td>
<td>79.5</td>
<td>61.2</td>
<td>0.0006</td>
<td>LUZIP4/CT28</td>
<td>220665_at</td>
<td>73.3</td>
</tr>
<tr>
<td>CAGE1</td>
<td>1563787_a_at</td>
<td>91</td>
<td>62</td>
<td>0.0008</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

iosity; mo: months.

We identified 98 articles that focused on the identification and assessment of diagnostic biomarkers and achieved a pooled AUC of 0.85 (95% CI 0.82–0.88), indicating that the diagnostic performance of these biomarkers when combined was excellent. However, the heterogeneity was also considerable (I^2 = 98%, p < 0.00001). Of the studies, 30 focused on single/antigen panels, 22 on autoantibodies, 31 on miRNA and RNA panels, and 15 suggested the use of circulating DNA combined with CEA or NSE for early LC detection.
Verification of blood biomarkers with high sensitivities (Ciz1, exoGCC2, ITGA2B), high specificities (CYFRA21-1, antiHE4, OPNV), or both (HSP90α, CEA) along with miR-15b and miR-27b/miR-21 from sputum was deemed a promising biomarker panel that could improve early LC detection (Table 3; Figure 4).

![Figure 4](image_url)

**Figure 4.** The likelihood ratios (LRs) were calculated for all studies detailed in Table 3, which presented their sensitivity and specificity. Values above 10 were considered to have strong evidence to rule in LC [39,40]. The number marker points indicate study reference. Error bars indicate mean and range.

Although CEA and CYFRA21-1 were antigens identified with high specificities for LC [21], both showed low sensitivity for the early detection of LC as single biomarkers. The systematic review benefited from being able to consider all biomarkers studied and published up to the search date. This provided an opportunity to identify those blood biomarkers with high sensitivity (>90%), high specificity (>90%), or both, and to consider how the incorporation of sputum miRNAs into a diagnostic panel could maximise sensitivity and specificity for the detection of LC in the early stages. However, it was notable that none of these antigens were CTAs, reflecting the fact that CTAs tend to be overexpressed in the advanced stages of cancer [27] rather than the early stages.

### 6. Discussion

The National Lung Screening Trial showed that there was a 20% reduction in mortality associated with low-dose CT screening of people with known risk factors for LC. CT scans show good efficiency in detecting small peripheral lesions, particularly ADC [41]. In most cases, both benign and malignant nodules have a high degree of similarity in the early stages, and scanning errors as well as the enormous false positive rate for CT scanning techniques are ongoing confounders [42]. This leads to the requirement for additional procedures consisting of bronchoscopy, fine needle aspiration, transthoracic needle aspiration, and surgical biopsy for further assessment. As an alternative to CT scanning, FDG-PET and contrast CT were used to assess malignancy risk [43]. While CT scans fail to detect lesions that are centrally located, bronchoscopy and sputum cytology can identify 25% of lung malignancies that cannot be detected by imaging techniques. Positive screening of widespread cancers, including breast, colorectal, cervical, prostate, and skin, can be quickly followed up with a tissue biopsy at minimal extra risk to patients [44–46].
However, this is not the case in LC, as there is a requirement for invasive procedures such as Mediastinoscopy with anaesthesia, as the lung is a fragile organ, and this is associated with an increased risk of tissue damage, including significant rates of pneumothorax [47]. Furthermore, cost-benefit analysis showed that more than 40% of the total cost of LC management is attributed to benign diseases being investigated by invasive approaches [11]. Due to these invasive procedures being associated with morbidity, increased costs, and delays in diagnosis, the development of non-invasive approaches are needed. Experience with biomarkers for the evaluation of symptoms in the fields of endocrine (HbA1C for diabetes) or infectious diseases (HIV viral load) are examples of the successful use of biomarkers in clinical practice [48].

Table 3. Biomarkers for the early detection of LC with >80% sensitivity and/or specificity, based on a recent review by Mohamed et al. [17].

<table>
<thead>
<tr>
<th>Name of Protein(s) Evaluated</th>
<th>Comparison Groups</th>
<th>Sample Size</th>
<th>Sensitivity %</th>
<th>Specificity %</th>
<th>AUC 95% CI</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPNV</td>
<td>NSCLC/nodules</td>
<td>1182</td>
<td>80</td>
<td>88</td>
<td>0.88</td>
<td>[49]</td>
</tr>
<tr>
<td>Secretory phospholipase A2-IIa</td>
<td>NSCLC/BN/HC</td>
<td>145</td>
<td>48–67</td>
<td>86</td>
<td>0.68–0.86</td>
<td>[50]</td>
</tr>
<tr>
<td>NSE + CEA + CYFRA21-1</td>
<td>LC/BLD/HC</td>
<td>132/48/92</td>
<td>75.76</td>
<td>89</td>
<td>0.63</td>
<td>[51]</td>
</tr>
<tr>
<td>CYFRA 21-1</td>
<td>LC/BD</td>
<td>161/97</td>
<td>59</td>
<td>94</td>
<td>0.85</td>
<td>[52]</td>
</tr>
<tr>
<td>HSP90α, CEA</td>
<td>LC/HC</td>
<td>175/160</td>
<td>95.63</td>
<td>99.97</td>
<td>0.996</td>
<td>[53]</td>
</tr>
<tr>
<td>CA-125, CEA, CYFRA21-1, EGFR/HER1/ErBB1, Gro-Pan, HGF, IL-10, IL-12p70, IL-16, IL-2, IL-4, IL-5, IL-7, IL-8, IL-9, Leptin, LIF, MCP-1, MIF, MIG, MMP7, MP9, MPO, NSE, PDGF-BB, Rantes, Resistin, sFasl, SAA, sCD40-ligand, sICAM-1, TNFRI, and sTNFRII.</td>
<td>NSCLC/HC</td>
<td>1479</td>
<td>80</td>
<td>95</td>
<td>0.96</td>
<td>[54]</td>
</tr>
<tr>
<td>Ciz1</td>
<td>LC/inflammatory diseases</td>
<td>35/170/160</td>
<td>95</td>
<td>74</td>
<td>0.96</td>
<td>[55]</td>
</tr>
<tr>
<td>Exosomal GCC2</td>
<td>NSCLC/HC</td>
<td>70/16</td>
<td>90</td>
<td>75</td>
<td>0.84</td>
<td>[56]</td>
</tr>
</tbody>
</table>


Many studies have investigated the potential for a non-invasive diagnosis of LC in patients with indeterminate pulmonary nodules (lesions of unknown malignant status), focusing predominantly on circulating biomarkers. Blood biomarkers have been repurposed to distinguish benign from malignant lung nodules [57]. For example, 552 patients have been studied (113 benign nodules and 339 malignant) for serum C-reactive protein (CRP) levels combined with CEA in the presence or absence of nodule spiculation, calcification, and CT bronchus signals. It has been found that CRP correlates with inflammation, while CEA is one of the glycoproteins that can assist in cellular adhesion and is thought to be upregulated in many epithelial cancers, including LC, due to metastasis [58]. Furthermore, the phospholipid hydrolase enzyme is known as secretory phospholipase A2-IIa (sPA2-IIa) and facilitates several precursors to eicosanoid release, regulating mechanisms including immunity, inflammation, and carcinogenesis. The sPA2-IIa was found to be highly expressed in prostate cancer but could also assist in differentiating LCs from healthy individuals. However, sPA2-IIa failed to discriminate between LCs and benign nodules, with an AUC of 0.68 [50].

Blood tests have been developed to assign clinical significance to indeterminate nodules, including an Early Cancer Detection Test (EarlyCDT®) manufactured by Oncimmune (Freenome Limited MediCity, Nottingham, UK) and a multianalyte serum biomarker panel.
by Bigbee et al. [59–61]. The Oncimmune test includes a panel of autoantibodies against NY-ESO-1, p53, GBU4-5, annexin I and SOX2 that were examined and validated in early-stage LC patients. However, the sensitivity was low, around 39%, although the specificity was 90% [60,61]. A panel of ten markers was validated by Bigbee et al. [59] to predict the likelihood of cancer developing in high-risk individuals with indeterminate lung nodules. The marker panel included prolactin, transthyretin, sE-selectin, thrombospondin-1, C-C motif chemokine 5 (CCL5; RANTES), macrophage migration inhibitory factor (MIF), plasminogen activator inhibitor, tyrosine-protein kinase, erbb-2, CYRA 21-1, and SAA. The sensitivity of the combined markers in this panel was 73.3%, although the specificity was 93.3% [59]. However, this panel has not changed the treatment plan, and patients still undergo invasive procedures, suggesting that so far it has only had a modest clinical impact. On the other hand, this panel had a good negative predictive value of 77.8% in the validation set and could aid in the screening of the population [59]. The PPV needs to be improved through the exploration of some additional biomarker targets. Large-scale validation will allow panels of biomarkers to become highly efficient tools in clinical practice in large prospective clinical trials [62]. In addition to predicting the propensity of nodules to become malignant, LC is often referred to as a single disease; however, it is more likely a heterogeneous group of diseases rather than a single entity. It is obvious that patients with NSCLC respond differently to treatment, and it is a clinically and biologically heterogeneous group of LCs [63]. The role of intra-tumoural heterogeneity and genetic diversity within a single tumour remains unclear, as does their impact on the sensitivity of tumours to immune modulation [64–66].

Blood biomarkers represent an important resource for LC diagnosis due to their ease of access and low risk of secondary effects for patients [57]. From our systematic review, it was evident that tumour antigens may assist LC evaluation, that micro-RNA panels can provide suitable candidates for the early diagnosis of LC, and perhaps surprisingly, that their combination with tumour antigens may be worthy of further investigation. Sputum and liquid biopsies are currently being evaluated for LC diagnosis. For example, the detection of folate receptor (FR)-positive circulating tumour cells could aid in the diagnosis of LC with 70% sensitivity and 79% specificity when combined with CEA [67].

Successful biomarker identification and characterisation are required to provide clinical evidence that informs treatment, as well as reduce the need for invasive procedures, diagnosis time, and rates of false positive results [57]. Discovery studies should focus on study design, different controls, sample sizes that have the power to achieve statistical significance where they exist, and validated analytical tests for biomarker measurements. Training and validating cohorts are required to ensure robust biomarker performance in two independent centres [68].

7. Conclusions

Tumour antigens have been investigated as biomarkers for the early diagnosis of LC [17], but most have low sensitivity and specificity and are more accurate at identifying advanced diseases. Further research will be needed to identify protein signatures associated with each cancer subtype. Most tumour antigens are wild-type proteins that are overexpressed or mutated, so they stimulate the immune response in patients with cancer. This altered expression needs to be detectable by non-invasive means to make tumour antigens good biomarkers and targets for the therapy of the disease. The gold standard procedures for LC diagnosis still require invasive biopsy procedures and imaging techniques that require skilled expertise and hi-tech equipment. Identifying biomarkers that may assist in the early diagnosis of LC is essential and will likely involve panels of proteins that exist in independent body fluids such as sputum and blood. Identification of robust biomarkers that correlate with prognosis would not only assist in an accurate diagnosis but may also offer new targets for treatment.
Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/onco4020008/s1, Table S1: The role of CTAs commonly found to be expressed in LC. Refs. [69–80] are cited in the Supplementary Materials.

Author Contributions: Conceptualization, funding acquisition, supervision, review and editing, project administration, S.H. and B.-a.G.; data curation, E.M. and D.F.; writing—original draft preparation, E.M. All authors have read and agreed to the published version of the manuscript.

Funding: E.M. was funded by a University of Hull PhD studentship awarded to S.H. and B.G.

Acknowledgments: E.M. was funded by the University of Hull PhD cluster studentship awarded to B.G. and S.H.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

ADC adenocarcinoma
CAGE cancer-associated gene
CEA carcinoembryonic antigen
CT computer tomography
CTA cancer testis antigen
CYFRA21-1 cytokeratin 19 fragments
HE4 human epididymis 4
HSP70 heat shock protein 70
LC lung cancer
MAGE melanoma-associated antigen gene
NSCLC non-small cell lung carcinoma
NSE neuron-specific enolase
OPN Osteopontin
PET positron emission tomography
SAA serum amyloid A
SCC squamous cell carcinoma
SCCA squamous cell carcinoma antigen
TA tumour antigens
TAA tumour associated antigens

References


### Disclaimer/Publisher’s Note:

The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.