Serum tumor markers in early stage lung cancer patients

Terese Liadal

Student project, Medical Faculty

UNIVERSITY OF OSLO

15.03.2016
Serum tumor markers in early stage lung cancer patients
© Terese Liadal

2016

Serum tumor markers in early stage lung cancer patients

Terese Liadal

http://www.duo.uio.no/

Print: Reprosentralen, Universitetet i Oslo
1 Abstract

Lung cancer is the cancer with the highest incidence and the leading cause of cancer-related mortality in the world. Identifying the subtype of lung cancer a patient suffers from is important for choice of treatment, as different types are treated differently and also differ greatly in prognosis. Biopsy is the gold standard for determining lung cancer subtype, but it is not always possible to obtain biopsies of sufficient quality to ensure diagnosis.

Serum tumor markers are inexpensive and non-invasive, and may be of help determining lung cancer histology where a biopsy is not possible. When diagnosing lung cancer, monitoring treatment and screening for relapse there are several serum tumor markers in clinical use. However, which markers are used, and for what purpose, varies.

In this project we examined pre-operative serum samples from 143 lung cancer patients. The majority of our patients had stadium I and II lung cancer, and all had subsequent surgery, thus enabling definitive histological diagnosis. The primary aim of this study was to investigate if elevated tumor marker values or combinations of elevated tumor marker values could be of use to determine lung cancer histology, especially in deciding if a tumor had neuroendocrine origin or features. Our study included a panel of seven tumor markers; NSE, CEA, proGRP, SCC(A), Chromogranin A, CYFRA 21-1 and HE4.

Our samples were taken from patients with adenocarcinoma, squamous cell carcinoma, large cell carcinoma, carcinoid and small cell lung cancer. A small number of the cancers classified as large cell carcinoma were found to have neuroendocrine features. We decided to distinguish between neuroendocrine tumors (NETs) and non-neuroendocrine tumors (non-NETs), in the first group we find the small cell lung cancers, carcinoids and the large cell carcinoma with neuroendocrine features, and in the second group adenocarcinoma, squamous cell carcinoma and large cell carcinoma without neuroendocrine features.

Our numbers indicate that the serum tumor markers that can be helpful in detecting neuroendocrine tumors are proGRP and possibly NSE. The sensitivity of proGRP when using the upper reference limit provided by the laboratory, in determining neuroendocrine status, is 68%, specificity 84%. The sensitivity of NSE is 32% and specificity 94%. If we use the same upper reference limits, and count the NETs that have either proGRP or NSE elevated (or
both), the sensitivity is 76% and specificity 83%. We discuss what raising the cutoff does for our material with regards to sensitivity and specificity of detecting neuroendocrine tumors.

ProGRP is also available for immunohistochemistry, and helps us detect non-small cell lung cancer with neuroendocrine differentiation. The serum tumor marker level of proGRP seems to correlate to the degree of staining.

This should motivate further research on tumor marker testing of early stage cancer patients, especially utilizing the proGRP tumor marker to detect neuroendocrine tumors.
## Index

1. Abstract ........................................................................................................ 5
2. Introduction ................................................................................................. 9
3. Serum tumor markers ............................................................................... 14
4. Methods ..................................................................................................... 18
5. Results ..................................................................................................... 21
6. Discussion ................................................................................................. 33

References ................................................................................................... 38
Appendix ...................................................................................................... 41
2 Introduction

Lung cancer is the cancer with the highest incidence (1.8 million new cases or 13.0% of the world total) and the leading cause of cancer-related mortality (1.6 million deaths or 19.4% of the total) in the world. (1) Lung cancer is among the most frequent cancers in Norway, second and third in incidence in men and women respectively (1596 and 1423 of 31 651 new cases in 2014, or 9.5% in total) and the highest in mortality (2158 of 10971 deaths in 2014, or 19.7%), with roughly the same number of deaths as colon, breast and prostate cancer combined.(2)

A biopsy from a suspicious lung tumor mass secures the diagnosis of lung cancer and the histological differentiation. In the lungs both biopsies and surgery are not without risks, these

procedures can be difficult depending on the location of tumors and the co-morbidity of the patient. To have a serum tumor marker giving the same information would save much time and great resources, and be significantly safer for the patient. Serum tumor markers have an established practice in cancer health care, especially in monitoring disease and screening for relapse (PSA for prostate cancer, CEA for colorectal cancer etc.), but in lung cancer the use is less established and different treatment facilities have different internal guidelines. The reason for this is that there is not one serum tumor marker that is clearly more sensitive for the presence of malignancy or which helps with the histological stratification, and often a combination of several serum tumor markers gives better results, but the different studies do not agree on which serum tumor markers and what combinations.

We had access to high-quality data from a population of early stage lung cancer patients, the type of population where we would expect the greatest benefit of correct diagnosis and treatment. We primarily wanted to see if we could use the data material to find support for determining whether or not a lung tumor was of neuroendocrine origin. Lung cancer is traditionally classified as either small cell lung cancer (SCLC) or non-small cell lung cancer (NSCLC), as this is where the greatest difference in treatment and prognosis lie. NSCLC includes squamous cell carcinoma, large cell carcinoma and adenocarcinoma. The distribution within lung cancers of small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC) is roughly 15% for SCLC and 85% for NSCLC. A much rarer type of lung cancer is carcinoid of the lung, with a prevalence of about 1%. Carcinoids are, like SCLC, of neuroendocrine origin, and are not a lung specific cancer, carcinoid tumors also arise in gastrointestinal tissue.(3) Our data includes about as many patients with SCLC as carcinoids (Figure 2), so we chose to make the distinction between neuroendocrine tumors (NETs) and non-endocrine tumors (non-NETs) in our analyses.
Our serum tumor marker panel was chosen based on combination of existing clinical practice and published studies. Three of our serum tumor markers are the most commonly used in monitoring non-small cell lung cancer: CEA, SCC(A) and CYFRA 21-1. Three of our serum tumor markers have been suggested as helpful in detecting neuroendocrine tumors: NSE, proGRP and Chromogranin A. HE4 has been suggested as a more sensitive marker for detection of early stage malignancy, particularly adenocarcinomas, than the more commonly used serum tumor markers.

The mortality of SCLC is much higher than for NSCLC, and SCLC has a five year survival rate for extensive disease as low as 0-1%, and about 15% for limited disease. NSCLC has a five-year survival for resectable tumors of 65%, but extensive (metastatic) disease of NSCLC is the same as extensive SCLC, at about 1%.(4) The main treatment choice of NSCLC is surgery, with or without adjuvant chemotherapy and radiation, whereas in SCLC surgery usually has no role, and depends on chemotherapy and radiation alone.(5, 6)

Staging is important when diagnosing cancer, as it greatly affects prognosis and also treatment choice. NSCLC are staged after conventional TNM staging (stage I-IV), where T reflects tumor size, N (lymph) node affection and M the presence and extent of metastasis. However in SCLC, in clinical practice there are only two stages; limited and extensive disease. Limited disease is in one hemithorax only and no metastasis to the other lung or other locations. Limited disease is therefore a target for intense one-sided radiation in addition to
chemotherapy, occasionally surgery plays a role. In extensive disease, usually only chemotherapy is used. (5)

In recent years immune modulation therapy, drugs targeting the immune system to become more aggressive, have entered the stage of cancer treatment. A meta-analysis of phase II trials of metastatic stage IV melanoma treated with conventional chemotherapy by Korn et al estimates the 1-year progression free disease to be around 5%.(7) In a study in 2010 by Hodi et al, 1-year progression free disease in metastatic melanoma when treated in monotherapy with ipilimumab is raised to around 10%. (8) Larkin et al recently presented good results for a combination of two immune modulation therapies, which brought metastatic malignant melanoma (stage III and IV) from a 1-year progression free disease of around 20% when treated with ipilimumab monotherapy to 60% in the group receiving ipilimumab/nivolumab combination therapy. (9) Due to a lot of the drugs still being in clinical trials, and also a non-negligible cost, these drugs have not yet entered the standard treatment programme for lung cancer in most countries, including Norway. The results for treating lung cancer with immune modulation therapy so far are promising, but what role these drugs will play remains to be seen (7, 10-14).

The gold standard for determining histology is a tumor sample (biopsy) and direct observation in microscope. However biopsy is not always possible. Research reports some relationship of various serum tumor markers to the histological subtype of lung cancer, but their clinical utility is not clear (15, 16). The serum tumor markers are also used to monitor the treatment response and to detect relapse. The National Academy of Clinical Biochemistry has released an early version of guidelines for the use of serum tumor markers in lung cancer (Table 1) (15). They argue that although there still are not any tumor markers with the necessary quality to be widely implemented, they recognize that several markers are already in use, and as such wish to aid this use by summarizing the current knowledge. The markers they mention in this guideline are NSE, CEA, CYFRA-21-1, proGRP and SCC(A), all in our test panel.

There are NSCLC tumors that exhibit neuroendocrine features, these are considered a subgroup of SCLC by the WHO due to their behavioral characteristics (17, 18). Neuroendocrine tumors are usually more aggressive and have a higher mortality than tumors with other histologies, with the notable exception of carcinoids, which in general have low mortality rates. (19, 20) Some claim lung cancer tumors with mixed histology behave more like neuroendocrine tumors, than non-neuroendocrine tumors, although the usefulness of
determining neuroendocrine differentiation in NSCLC is controversial. Ionescu et al concludes that “disease specific and overall survival is not influenced by neuroendocrine differentiation and therefore non-small cell lung carcinoma with neuroendocrine differentiation should not be a subclass distinct from the other NSCLC” (21). Conversely Lyda and Weiss (22) conclude that immunohistochemistry has a role, but without specifying what that role is. Immunohistochemistry is recommended in the diagnosis of SCLC and NSCLC per guidelines drafted by the European Society of Medical Oncology (ESMO) (5, 17), especially in cases where the biopsy sample has been crushed or otherwise compromised. The immunohistochemistry they recommend is synaptophysin, chromogranin A, CD56, thyroid transcription factor 1 and MIB-1.

In this study we wanted to investigate the properties of commonly used serum tumor markers in early stage lung cancer patients. We wished to see if there is a serum tumor marker, or a combination of serum tumor markers, which could help us predict with a satisfying level of accuracy whether the lung tumor had neuroendocrine features or not. The patients were expected to have limited disease and were all deemed operable. This perhaps makes the findings here transferrable to the cases we want to be able to detect; early stage lung cancer patients still eligible for curative surgery. We do not have a set of healthy controls, so we have to rely on the upper reference limits of the assays. The biggest disadvantage is that we do not have control over whether the population these reference limits are based upon is comparable to ours, especially with regards to age and smoking status.

We also have available more detailed data about the histology by the means of immunohistochemistry with staining by anti-proGRP antibodies. We wanted to see the occurrence of mixed histologies, if the serum tumor marker levels are affected, and if the degree of staining directly on the tumor mass samples corresponds to the serum tumor marker level.

This project is written as a general introduction to the field of serum tumor markers in lung cancer. With the presentation of the results from our serum tumor marker panel, and what these results might suggest, hopefully it can also be of interest to someone already in the serum tumor marker and/or lung cancer field.
Any cancer has the best prognosis when discovered early, and efforts are focused on finding tests that will detect and correctly classify malignancies in early stages. If the tests are good enough, they could play a role in a screening setting. At the moment the most realistic implementation of lung cancer screening utilizes CT scanning on a high-risk population (smokers, ex-smokers with a high number of pack-years). CT screening of lung cancer is known for its low specificity, not all lesions detected are malignant, in addition CT screening often misses SCLC and centrally localized tumors.(23, 24) A blood test which could help find the true positives is therefore very much sought after.

In our project we primarily wanted to see if our serum tumor marker panel could help us establish if the histological subtype was of neuroendocrine origin. In a review by Harmsma et al (17), they conclude that at this moment none of the currently available markers are optimal for diagnosis and monitoring of SCLC. There is currently no consensus on which serum tumor markers are useful when diagnosing lung cancer. There are several tumor markers that will get pathologically elevated during malignant disease, unfortunately their use in screening is limited, since the required sensitivity and specificity is not met. (25)

The guidelines of major clinical societies as summarized in Tietz Fundamentals of Clinical Chemistry and Molecular Diagnostics (26) is given in Table 1.

Table 1: Current major guidelines for the use of serum tumor markers in lung cancer

<table>
<thead>
<tr>
<th>National Academy of Clinical Biochemistry</th>
<th>American Society of Clinical Oncology</th>
<th>American Cancer Society</th>
<th>European Group on Tumor Markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEA and CYFRA 21-1 in NSCLC and NSE and proGRP in SCLC for differential diagnosis, postoperative surveillance, monitoring of therapy in advanced disease, and detection of recurrence; CYFRA</td>
<td>None</td>
<td>None</td>
<td>NSE in SCLC, CYFRA 21-1 in NSCLC, and SCC(A) in NSCLC for differential diagnosis; CEA, CYFRA 21-1 in NSCLC, and NSE in SCLC for prognosis and for follow-up and monitoring of therapy</td>
</tr>
</tbody>
</table>
The serum tumor markers in our test panel are described briefly under.

**NSE neuron-specific enolase**

NSE is a neuron-specific enzyme, produced in central and peripheral neurons and malignant tumors of neuroectodermal origin (e.g. SCLC, neuroblastomas, intestinal carcinoid). NSE is also found in erythrocytes, plasma cells and platelets, and serum must be separated from red cells within 60 minutes of venipuncture. High levels (Molina suggests 100 ng/ml (25)) of NSE in combination with a suspicious lung tumor indicate the presence of SCLC with high probability. Neuroendocrine tumors of other localisations; liver cancer, lymphoma and seminoma have to be considered. Moderate elevations of NSE are also found in benign lung diseases and pancreatic, gastric, colorectal and breast cancer. In lung cancer, numerous studies support the use of NSE for differential diagnosis, particularly SCLC. Several groups have found a clear additive value when NSE was combined with ProGRP (25, 27-33).

**CEA carcinoembryonic antigen.**

CEA is a large glycoprotein related to immunoglobulins. CEA was one of the first tumor markers to be described (in 1965), and has relatively high sensitivity for many advanced adenocarcinomas. It is produced during embryonal and fetal development. In adult organisms, it is produced in low amounts in the GI tract, pancreas and the liver. The analysis is most relevant in colorectal cancer, where it is elevated in up to 70%, but also in lung (40%), gastric (50%), breast (40%), pancreatic (55%), ovarian (25%), and uterine (40%) carcinoma. In lung cancer, CEA can be helpful in diagnosing non-small cell lung carcinoma (>65% of patients have elevated CEA) and in monitoring treatment and screening for relapse. CEA is not suitable for cancer screening. In lung cancer increased release of CEA can also be observed with other lung cancer histologies, in benign lung pathologies and in smokers. (26, 34)

**ProGRP progastrin-releasing peptide**

ProGRP is a precursor of the gut hormone gastrin releasing peptide (GRP). ProGRP is more stable in serum than GRP, but measures the same parameter. GRP is expressed in the human
nervous system and acts potently as a neurotransmitter regulating body temperature and central homeostatic mechanisms. It is also produced by neuroendocrine tissues in the gastrointestinal and respiratory tract. ProGRP is a relatively new serum tumor marker, and has promising results as a marker for SCLC. In other malignant diseases it is rarely elevated, and if so, only to a small extent, with the exception of neuroendocrine medullary thyroid carcinomas. Benign diseases do not lead to ProGRP release into serum, with the exception of renal disease and possibly heart failure. Molina et al suggests ProGRP concentrations > 200 pg/mL are indicative for lung cancer, concentrations >300 pg/mL for SCLC. They claim proGRP as a single marker is superior to NSE, and that the combination of NSE and proGRP provides additional information. (25, 26, 34)

**SCC(A) squamous cell carcinoma antigen**

SCC(A) is expressed from squamous epithelial cells of the respiratory and gastrointestinal tract and the cervix. SCC(A) is found in cytoplasm, and is considered a structural protein reflecting the differentiation grade of squamous cell cancers. SCC(A) is elevated in a variety of squamous cell carcinomas, including those of the cervix, lung, skin, head, neck, digestive tract, ovaries and genital tract. SCC(A) is metabolized by the kidneys, thus renal insufficiencies might cause elevations of SCC(A) levels. Other benign reasons for elevated serum SCC(A) levels are pulmonary infection, skin disease and liver disease. SCC(A) is regarded as less sensitive in NSCLC than CYFRA 21-1. Molina et al claims SCC(A) has a superior specificity for squamous cell cancer and can be used for histological subtyping, but this might be a controversial claim, and does not reflect clinical practice. A high value is associated with a high probability of NSCLC, mainly squamous tumors. (15, 26, 28, 34)

**Chromogranin A**

Chromogranins are acidic glycoproteins that are widely expressed by neuroendocrine cells. Intracellularly chromogranins are suggested to play a role in the regulation of secretory granules. Chromogranin A (CgA) is the most studied of the chromogranins. CgA is widely expressed by neuroendocrine tissue, and secreted alongside neuropeptides and peptide hormones. CgA is one of the most used tumor markers in gastrointestinal NETs. CgA is not a tumor-specific antigen for neuroendocrine tumors (NETs), and abnormal concentrations have been described in some non-malignant diseases such as renal failure, heart failure,
hypertension and in patients with liver disease. There are several methods available to detect CgA. The method is important, since, as reported in pathological conditions such as NETs, different proteolytic processes may take place generating a variable number of fragments that may not be detected by all these techniques (35-39).

**CYFRA 21-1 detects one of the smallest cytokeratin fragment, cytokeratin 19.**

Cytokeratin 19 fragments are found all over the human organism. It is however particularly expressed in lung tissue. Since cytokeratin 19 are disposed of by the kidney, renal disease might lead to elevated CYFRA 21-1 concentrations in blood. CYFRA 21-1 has been reported as the most sensitive tumor marker in NSCLC, especially in squamous cell carcinoma. Besides lung cancer, CYFRA 21-1 is elevated in sera of patients with urological, gastrointestinal and gynecological cancers. According to Molina et al CYFRA 21-1 is especially used for its prognostic value of NSCLC therapies, and National Academy of Clinical Biochemistry recommends this use. CYFRA 21-1 is also used for treatment monitoring and detection of recurrent disease. (15, 25, 26, 28)

**HE4 human epididymis protein 4**

HE4 is a product of the whey-acidic-protein 4-disulphide core domain 2 (WFDC2) gene which is found to be overexpressed in epididymal and ovarian tissue, and also in lung tissue. There is a high concentration is respiratory epithelia, which suggests it might have a role in immune defense. Its function is not entirely known, it might be a protease inhibitor. HE4 is especially used for ovarian cancer, where it can indicate malignancy early in the disease course. HE4 serum concentrations are influenced by age, menstrual status, renal function and of special concern with regards to lung cancer; smoking status. Several studies suggests that HE4 can be a better serum tumor marker in detecting early stage lung cancer when compared to the more conventional serum tumor markers above, however little or no emphasis is put on smoking. We know smoking elevates HE4, and almost all lung cancer patients smokes or have smoked heavily in the past. Until larger studies where smoking is taken into account are published, HE4 results must be used with caution when considering using it in lung cancer. (26, 34, 40, 41)
4 Methods

Our sera were obtained from lung cancer patients pre-operatively. All the samples were immediately centrifuged and frozen, and kept at – 70 degrees Celsius, until thawed, mixed and analyzed. All the samples have been run at the same machines and in a short time period in order to limit variation. All the surgically removed tumors underwent a histological verification. All patients were deemed operable and the majority had TNM stage I or II. However a few of the patients were later found to have more advanced cancer than originally anticipated. None of the patients had had previous treatment for lung cancer.

Age and stadium was available for 133 of 143 patients. Of these 65 were female and 68 male. The average age was 64 years (SD 7.5 years). The average number of pack years was 28 (SD 12). Nine of the patients had never smoked. Only TNM stages are available, and not limited and extensive disease in SCLC. To simplify a and b substages have not been used. (For instance Ia and Ib both are in the Pstadium 1.0 in Figure 3.)

Figure 3: Number of patients stage I-IV
We have also chosen two machine learning algorithms, C4.5 and Classification And Regression Tree, to see what rules are suggested, both with the whole data set, and the data set with carcinoids excluded (since they are very rare and the distribution in our data with them included does not reflect the distribution in the population). The algorithms were chosen because of their established use in the machine learning field, and also because the output is easy to read and understand.

C4.5 builds decision trees from a set of training data using the concept of information entropy. C4.5 is often referred to as a statistical classifier. The input is a sample with attributes and a class. In our cause the serum tumor marker levels are the attributes and the class is either NET/non-NET or SCLC/NSCLC. The output is a rule tree that helps classify
any sample into a certain class. The algorithm computes the percentage of correctly and
incorrectly classified instances. (42) We used the Weka (Java) implementation J48.

CART (Classification And Regression Tree) is a type of classification tree, meaning a tree
model where the classifier can take a finite set of attributes, and classify according to those
attributes. Since CART also can take continuous numerical variables, it is also a regression
tree. CART has the same input as C4.5, a set of attributes along with a class, and the output is
also here a rule tree suitable for classification.(43) We use the Weka (Java) implementation
SimpleCart. Interestingly Patz used CART to find rules to detect malignancy on a serum
tumor marker panel of CEA, SCC(A), AAT (α1-antitrypsin) and RBP (retinol binding
protein). (44)

The machine learning algorithms were executed in Weka, an open source program for
machine learning, provided by the University of Waikato, New Zealand
(http://www.cs.waikato.ac.nz/ml/weka/).

Statistics and charts are made with Microsoft Office Excel 2010 and SofaStatistics, an open-
source statistics, analysis, and reporting program (http://www.sofastatistics.com/).
5 Results

A count of the positive serum samples according to each serum tumor markers upper reference limit, recommended by the tumor marker laboratory at Oslo University Hospital, Radiumhospitalet, is given in Table 2. The upper reference limits for each marker is given in the table.

Table 2: Numbers of samples above upper reference limit per histology, and NET/non-NET.

<table>
<thead>
<tr>
<th>Patients / Serum tumor marker (upper reference limit)</th>
<th>Squamous cell carcinoma</th>
<th>Adenocarcinoma</th>
<th>Large cell carcinoma</th>
<th>Small cell lung cancer</th>
<th>Carcinoi d</th>
<th>NET</th>
<th>Non-NET</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSE (&lt;10µg/L)</td>
<td>4/40</td>
<td>1/40</td>
<td>1/30</td>
<td>10/15</td>
<td>2/18</td>
<td>12/37</td>
<td>6/106</td>
</tr>
<tr>
<td>CEA (&lt;5µg/L)</td>
<td>2/40</td>
<td>10/40</td>
<td>11/30</td>
<td>2/15</td>
<td>0/18</td>
<td>4/37</td>
<td>21/106</td>
</tr>
<tr>
<td>ProGRP (&lt;80ng/L)</td>
<td>5/40</td>
<td>7/40</td>
<td>8/30</td>
<td>11/15</td>
<td>11/18</td>
<td>25/37</td>
<td>17/106</td>
</tr>
<tr>
<td>SCC(A) (&lt;2µg/L)</td>
<td>34/40</td>
<td>34/40</td>
<td>26/30</td>
<td>4/15</td>
<td>14/18</td>
<td>22/37</td>
<td>90/106</td>
</tr>
<tr>
<td>CgA (&lt;92.9µg/L)</td>
<td>7/40</td>
<td>3/40</td>
<td>3/15</td>
<td>3/15</td>
<td>4/18</td>
<td>7/37</td>
<td>14/106</td>
</tr>
<tr>
<td>CYFRA21-1 (&lt;3.3µg/L)</td>
<td>17/40</td>
<td>7/40</td>
<td>13/30</td>
<td>3/15</td>
<td>2/18</td>
<td>6/37</td>
<td>36/106</td>
</tr>
<tr>
<td>HE4 (&lt;60/70 pmol/L)</td>
<td>37/40</td>
<td>40/40</td>
<td>28/30</td>
<td>13/15</td>
<td>15/18</td>
<td>32/37</td>
<td>101/106</td>
</tr>
</tbody>
</table>

In our material, HE4 seems to be a sensitive marker for early stage lung cancer. However this is likely related to smoking status, as it is known that smoking leads to elevated HE4 values. (34) Interestingly eight (of nine) of those who had never smoked had elevated HE4 as well, but this is most likely caused by a combination of tumor-associated HE4 and false self-reporting about smoking status.

SCC(A) also seems to be a sensitive marker in our material. However, contamination from skin SCC(A) is extremely common and has most likely contributed to the high number of positive samples. This is supported by the observation that the number of positives is equal in the adenocarcinoma group as in the group with squamous cell carcinomas, where we would expect SCC(A) to be a more sensitive marker. For this reason, we do not believe that our data

1 HE4 has two reference limits, according to age above 60 (70pmol/L) or under 60 (60pmol/L). Where we did not have the age (10 of 143 patients) we used the average age of the rest; 64 years (with an average deviation of 7.5 years).
provide a true evaluation of the properties of SCC(A) in lung cancer. For a true evaluation of SCC(A), particular care must be taken during venipuncture to avoid contamination from skin.

ProGRP is higher in neuroendocrine tissue than in non-neuroendocrine tissue, as anticipated. It has been suggested that the combination of elevated proGRP and NSE is indicative for SCLC. In seven of the 15 SCLC patients serum levels of both ProGRP and NSE were elevated. Seven of eight samples with both proGRP and NSE positive were SCLC (the last carcinoid). Again our sample size is too small to conclude anything, and we lack healthy controls to be able to estimate sensitivity and specificity of the serum tumor markers with regards to detecting malignancy.

Figure 6: Box plots for log_{10} proGRP (upper reference limit 1.9) and log_{10} NSE (upper reference limit 1) and HE4 (upper reference limit 60/70), entire population. Outliers displayed. Lower whiskers are 1.5 times the Inter-Quartile Range below the lower quartile, or the minimum value, whichever is closest to the middle. Upper whiskers are calculated using the same approach.
All 143 patients have later biopsy confirmed lung cancer with the corresponding histological subtype, shown in Figure 7. In addition the samples have been examined for neuroendocrine features. Four of the samples classified as large cell carcinoma turned out to have neuroendocrine features. Their NSE values were all under limit, proGRP above limit. The sample size is small and the different samples had different degrees of staining for the neuroendocrine marker proGRP, so a larger study is needed to see if proGRP is a good marker for NET status for large cell carcinoma (LCC). In our data a positive result for proGRP on a LCC verified sample, gives a 50% chance is has neuroendocrine features. Furthermore there seemed to be a positive correlation between degree of staining and serum proGRP level, as illustrated in Figure 8-10 below.
In Figure 8 we see a slide of a large cell carcinoma tumor which has been stained with an anti-proGRP staining. The proGRP serum level in the serum sample from this patient was 146 ng/L (<80ng/L). Below in Figure 9 to the left there is a carcinoid with a proGRP serum level of 34 ng/L, and the staining is negative for proGRP. To the right is a carcinoid with a proGRP serum level of 2060 ng/L, and the staining is intense, the sample is highly positive for the proGRP marker. In Figure 10 we see a SCLC with a proGRP staining below threshold of positive sample, serum proGRP level 37 ng/L.
Figure 9: Left: Carcinoid, negative proGRP staining. Right: carcinoid, positive proGRP staining.

Figure 10: SCLC with proGRP staining below threshold, serum proGRP level 37 ng/L.
Below in Figure 11 we see $\log_{10}$ proGRP serum levels in non-NET and NET. We see the distribution of non-NET patients is approximately normal. The distribution of NET has a greater variance and is skewed to the right.

Figure 11: ProGRP distribution, total and per non-NET/NET. The upper reference limit for serum proGRP of 80 ng/L corresponds to $\log$(proGRP) of 1.9
We used the Mann Whitney U test (reliable to extreme outliers) to test the difference between the serum proGRP values in non-NET and NET tumors. That gives us a p-value of < 0.001 (which means we can assume statistical significance) and a U-value of 814.5 (a value of 1961.0 would indicate no difference between groups). The rest of the relevant statistics is given in Table 3.

Table 3: Serum proGRP statistics

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Median</th>
<th>Avg. rank</th>
<th>Average</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-NET</td>
<td>106</td>
<td>55.0</td>
<td>61.2</td>
<td>63.5</td>
<td>21.7</td>
<td>19.0</td>
<td>314.0</td>
</tr>
<tr>
<td>NET</td>
<td>37</td>
<td>135.0</td>
<td>103.0</td>
<td>2613.1</td>
<td>4092.6</td>
<td>25.0</td>
<td>73450.0</td>
</tr>
</tbody>
</table>
The Mann Whitney U test on NSE gives us a p-value of <0.001 (statistical significant) and a U statistic of 1090 (a value of 1961 would indicate no difference between groups). The rest of the relevant statistics is given in Table 4.

Table 4: Serum NSE statistics

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Median</th>
<th>Avg. rank</th>
<th>Average</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-NET</td>
<td>106</td>
<td>0.8</td>
<td>63.8</td>
<td>6.5</td>
<td>8.5</td>
<td>0</td>
<td>303.5</td>
</tr>
<tr>
<td>NET</td>
<td>37</td>
<td>5.6</td>
<td>95.5</td>
<td>12.0</td>
<td>11.7</td>
<td>0</td>
<td>63.6</td>
</tr>
</tbody>
</table>

Using the Mann Whitney U test on CEA and CYFRA 21-1 shows they were both significantly elevated in non-NETs when compared to NETs. The Mann Whitney U test for serum levels of Chromogranin A and HE4 gave no statistically significant difference between NETs and non-NETs.

Below in Table 5, 6 and 7 sensitivity and specificity with regards to NET status is given, when using the serum tumor markers proGRP and NSE. The cutoffs for the proGRP marker is the upper reference limit and values that has been suggested by machine learning algorithms when using proGRP as the sole marker. The NSE cutoffs are chosen to illustrate, raising the limit does not seem to be valuable, in contrast with proGRP where raising the limit significantly improves specificity. This might indicate that proGRP has an optimal cutoff higher than the current upper reference limit, and that the NSE cutoff might be well suited as it is.

Table 5: NET status (SCLC, carcinoid, LCC with neuroendocrine features) with various proGRP upper reference limits

<table>
<thead>
<tr>
<th>ProGRP cutoff</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>80 ng/L</td>
<td>68% (25/37)</td>
<td>84% (89/106)</td>
</tr>
<tr>
<td>104 ng/L</td>
<td>62% (23/37)</td>
<td>93% (99/106)</td>
</tr>
<tr>
<td>150 ng/L</td>
<td>46% (17/37)</td>
<td>100% (106/106)</td>
</tr>
</tbody>
</table>

Table 6: NET status (SCLC, carcinoid, LCC with neuroendocrine features) with various NSE upper reference limits

<table>
<thead>
<tr>
<th>NSE cutoff</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 µg/L</td>
<td>32% (12/37)</td>
<td>94% (100/106)</td>
</tr>
</tbody>
</table>
Table 7: NET status (SCLC, carcinoid, LCC with neuroendocrine features) with both proGRP and NSE in serum elevated above upper reference limit

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Either proGRP &gt;80ng/L or NSE &gt;10 µg/L (or both)</td>
<td>76% (28/37)</td>
<td>83% (88/106)</td>
</tr>
</tbody>
</table>

If we repeat the tables for the cutoffs, but exclude the rare carcinoids from our material, the corresponding sensitivities and specificities look like this.

Table 8: NET status (SCLC, LCC with neuroendocrine features) with various proGRP upper reference limits, carcinoids excluded

<table>
<thead>
<tr>
<th>ProGRP cutoff</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>80 ng/L</td>
<td>74% (14/19)</td>
<td>84% (89/106)</td>
</tr>
<tr>
<td>104 ng/L</td>
<td>32% (6/19)</td>
<td>93% (99/106)</td>
</tr>
<tr>
<td>150 ng/L</td>
<td>16% (3/19)</td>
<td>98% (104/106)</td>
</tr>
</tbody>
</table>

Table 9: NET status (SCLC, LCC with neuroendocrine features) with various NSE upper reference limits, carcinoids excluded

<table>
<thead>
<tr>
<th>NSE cutoff</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 µg/L</td>
<td>53% (10/19)</td>
<td>94% (100/106)</td>
</tr>
<tr>
<td>20 µg/L</td>
<td>26% (5/19)</td>
<td>97% (103/106)</td>
</tr>
<tr>
<td>50 µg/L</td>
<td>16% (3/19)</td>
<td>97% (103/106)</td>
</tr>
</tbody>
</table>

Table 10: NET status (SCLC, LCC with neuroendocrine features) with both proGRP and NSE in serum elevated above upper reference limit, carcinoids excluded

<table>
<thead>
<tr>
<th>Both markers</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Either proGRP &gt;80ng/L or NSE &gt;10 µg/L (or both)</td>
<td>89% (17/19)</td>
<td>78% (83/106)</td>
</tr>
</tbody>
</table>

We ran the machine learning algorithms on the full data set, first classifying samples as either NET or non-NET, and then a second run classifying then as SCLC or NSCLC. We then removed the carcinoid data and repeated the tasks. Below are the most interesting rules and/or
with the highest number of correctly classified instances. The full printout of the runs are given in the Appendix. Explanation on how to read is given below Figure 16.

**CART on full data set (143 patients), classifying into SCLC and NSCLC**

<table>
<thead>
<tr>
<th>Rule</th>
<th>Classifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSE &lt; 9.735:</td>
<td>NSCLC(120.0/5.0)</td>
</tr>
<tr>
<td>NSE &gt;= 9.735</td>
<td>SCC &lt; 2.625: SCLC(10.0/1.0)</td>
</tr>
<tr>
<td></td>
<td>SCC &gt;= 2.625: NSCLC(7.0/0.0)</td>
</tr>
</tbody>
</table>

Correctly Classified Instances 131 91.6084 %

Figure 13

**CART on full data set (143 patients), classifying into NET and non-NET**

<table>
<thead>
<tr>
<th>Rule</th>
<th>Classifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>ProGRP &lt; 109.0:</td>
<td>Non-NET(100.0/14.0)</td>
</tr>
<tr>
<td>ProGRP &gt;= 109.0:</td>
<td>NET(23.0/6.0)</td>
</tr>
</tbody>
</table>

Correctly Classified Instances 120 83.9161 %

Figure 14

**C4.5 on full data set (143 patients), classifying into SCLC and NSCLC**

<table>
<thead>
<tr>
<th>Rule</th>
<th>Classifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSE &lt;= 8.66</td>
<td>ProGRP &lt;= 130: NSCLC(108.0/1.0)</td>
</tr>
<tr>
<td></td>
<td>ProGRP &gt; 130</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>NSE &gt; 8.66</td>
<td>SCC &lt;= 2.62: SCLC(10.0/1.0)</td>
</tr>
<tr>
<td></td>
<td>SCC &gt; 2.62: NSCLC(7.0/0.0)</td>
</tr>
</tbody>
</table>

Correctly Classified Instances 133 93.007 %

Figure 15
C4.5 on data set with carcinoids excluded (125 patients), classifying into SCLC and NSCLC

ProGRP $\leq 146$
- NSE $\leq 8.04$: NSCLC (102.0/2.0)
- NSE $> 8.04$
  - SCC $\leq 2.62$: SCLC (4.0/1.0)
  - SCC $> 2.62$: NSCLC (5.0/0.0)

ProGRP $> 146$
- NSE $\leq 4.58$: NSCLC (2.0/1.0)
- NSE $> 4.58$: SCLC (8.0/0.0)

Correctly Classified Instances 115 92 %

These rules and results can be read as follows (explained for the top rule):

NSE $< 9.735$: NSCLC (120.0/5.0)
NSE $\geq 9.735$
  - SCC $< 2.625$: SCLC (10.0/1.0)
  - SCC $\geq 2.625$: NSCLC (7.0/0.0)

If NSE is under 9.735 ($\mu$g/L), the sample is classified as NSCLC. (When applied to the full data set, 120+5 samples are classified as NSCLC, 120 correctly and 5 incorrectly). If NSE is over 9.735 ($\mu$g/L), we look at the SCC(A) value. If SCC(A) is under 2.625 ($\mu$g/L), the sample is classified as SCLC. (When applied to the full data set, 10+1 are classified as SCLC, 10 correctly and 1 incorrectly. If SCC(A) is over 2.625 ($\mu$g/L), the sample is classified as NSCLC. (When applied to the full data set, 7 are classified correctly, 0 incorrectly).

The idea of using machine learning as a tool to find rules for the use of serum tumor markers is not new. It can be a useful tool to find rules of thumb that might not be obvious from the data, and should subsequently be tested on a different larger data set for sensitivity and specificity.

The machine learning algorithms have no prior knowledge to the upper reference limits, and merely finds the best fit for the data in order to classify the highest percentage correctly.

Molina states the following rule of thumb; when SCC(A) is absent in a serum containing increased concentrations of NSE and proGRP (SCC(A)$<2$ $\mu$g/L, NSE$>35$ $\mu$g/L, proGRP$>100$ng/L), SCLC is the diagnosis of choice. (16, 17, 45) Whenever SCC(A) is
included in our rules, the lower value indicates SCLC. This is in alignment with suggestions that low SCC(A) serum levels might be an exclusion criteria for SCLC. (17)

The rules given in Figure 15 has the highest percentage of correctly classified samples (93%). However, the part of the rule regarding Chromogranin A looks like it might be over-fitting the data. It would be interesting to run the classification on a larger data set. Without the CgA part of the rule, the percentage of correctly classified instances is 91.6%.

For TNM stage I-III or limited disease Molina et al suggests that NSE>35 µg/L and SCC(A) <2 µg/L gives the diagnosis of SCLC with 100% correct classification/total. (16) For stage IV and extensive disease they suggest that NSE>45 µg/L and SCC(A) <2 µg/L gives the diagnosis of SCLC with 97.7% correct classification/total. In our data material 7/143 patients have NSE serum levels above 35 µg/L, and when we look at the stadium stratification in Molina’s patients, stadium I-III are in one group. Most of our patients are in stadium I or II. We lack stadium for 10 patients, mostly SCLC. We therefore lack sufficient metadata to justify direct comparison. However, it is interesting that our machine learning algorithms find a similar rule, given in Figure 13, with 92% correctly classified samples.

A similar rule as in Table 10, with the added criteria of SCC(A) below upper reference limit is given in Table 11. The sensitivity drops, and the specificity rises.

Table 11: NET status (SCLC, LCC with neuroendocrine features) with both proGRP and NSE in serum elevated above upper reference limit, SCC(A) <µg/L, carcinoids excluded

<table>
<thead>
<tr>
<th>Both markers</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Either proGRP &gt;80ng/L or NSE &gt;10 µg/L (or both)); AND SCC(A)&lt;2.0µg/L</td>
<td>58% (11/19)</td>
<td>98% (104/106)</td>
</tr>
</tbody>
</table>
6 Discussion

Can serum tumor markers be used for histological stratification? Today serum tumor markers in lung cancer are used mainly in treatment monitoring and screening for relapse. The definite histological subtyping is based on biopsy. Our samples are from an histologically verified lung cancer affected population, where the majority of the patients are early stage, still deemed operable. We hoped to see if our panel of serum tumor markers could help us make the distinction of neuroendocrine tumors vs non-neuroendocrine tumors.

We observed that to distinguish between neuroendocrine and non-endocrine tumors, especially proGRP and NSE seem to be useful. NSE also has a role in treatment monitoring. A study in 1997 by Pinson et al cite (29) concludes that NSE, although elevated in serum in the majority of their SCLC patients, the NSE level does not correlate with reduction in tumor size (clinical response) and therefore has doubtful utility in the follow up of SCLC treatment. Another study the following year by Fizazi (46) has the opposite conclusion, that NSE “is a strong, independent early predictor of both complete response to therapy and survival.” More than twenty years later today NSE still has no clear role in the diagnosis and follow-up of SCLC (5).

ProGRP has also been around since the early 2000s, and is markedly elevated in many neuroendocrine tumors (47). The ESMO guidelines for Clinical Practice Guidelines for diagnosis, treatment and follow-up of SCLC (5) of 2013 has no mention of the role of biomarkers neither in diagnosis or follow-up of treatment. Tumor markers are however in use in varying degree in different institutions. The National Academy of Clinical Biochemistry recognizes this, and attempts to help streamline the practice, with their early version of guidelines. (15) Many studies suggest that proGRP is a serum tumor marker that can be useful in detecting neuroendocrine tumors.(31, 33, 48)

Chromogranin A has been suggested as a serum tumor marker for neuroendocrine tumors, but in accordance with for instance Molina et al our data did not support the use of Chromogranin A in subtyping cancers.(35)

SCC(A) is a test that needs great precautions to not get contaminated, and almost all of the patients had elevated value. However the only cancer type that had low values for SCC(A) is
SCLC with 4/15 (27%). This might indicate that an elevated SCC(A) serum value can help exclude SCLC, as some studies suggest.\(^{(16, 17)}\)

CEA is most relevant in the monitoring of NSCLC. It has been reported to be most sensitive for adenocarcinomas (of any type, also lung) and large cell carcinoma. Our results are consistent with this, as only a few NETs and squamous cell carcinoma have elevated value, and a relative five-fold serum level increase in adenocarcinoma and large cell carcinoma. Even so, the highest frequency of elevated samples is found in large cell carcinoma, 11/30 (37%), which indicates that it is not sensitive enough in early stage lung cancer for any histology. It is interesting that none of the carcinoids had elevated CEA serum levels. In an old study by Bishopric positive CEA antibody staining is a poor prognostic factor for carcinoid.\(^{(49)}\). In a study by Divisi, 42 patients with bronchial carcinoids (14 stage I, 20 stage II, 8 stage III) all have negative CEA (the upper reference limits in this study of CEA serum level were < 5 µg/L non-smokers and < 7 µg/L smokers). Our highest CEA serum level of the carcinoid tumors is 3.7 µg/L. CEA can be elevated in a patient with a carcinoid tumor, as in a single case study by Panomreongsak concerning a carcinoid originating in the appendix.\(^{(50)}\) It would be interesting to see if an elevated CEA is a poor prognostic factor for carcinoids of the lung in a larger data material.

CYFRA 21-1 is reported to be a sensitive marker, especially for squamous cell carcinoma. In our material CYFRA 21-1 is elevated in 17/40 (43%) of the squamous cell carcinomas (and also 43% of the large cell carcinomas, and less in the other markers), our results are thus supportive of this claim. However, the sensitivity is also here too low to be of use in histological subtyping. It is reported that the main use of CYFRA 21-1 today is as a prognostic factor for NSCLC. Higher values indicate worse prognosis.\(^{(51, 52)}\)

HE4 has been suggested as a better tumor marker in early stage lung cancer. At first glance, our results support this. HE4 was elevated above upper reference limit in 133/143 cases (91%). The next best serum tumor marker when lung cancer patients seen as a whole is SCC(A) with 112/143 (78%). Both these results are however problematic. HE4 is affected by smoking status, and almost all of our patients smoke or have smoked in the past. The reference values for HE4 are for healthy non-smokers. To evaluate if this is a sensitive marker suited for for instance lung cancer screening we would need age- and smoking status adjusted reference values, but this remains a focus for future research. Also, as mentioned, contamination likely contributes to the proportion of elevated SCC(A)-results. It would be
interesting to repeat a similar project, with samples collected using protocols to minimize contamination from skin, to properly evaluate SCC(A) in early stage lung cancer.

In line with the current academic consensus, we found that the best predictive value of if the serum sample was from a NET patient, was an elevated proGRP (Tables 5 and 8), and to some degree an elevated NSE (Tables 6 and 9). Molina (25) suggests that NSE levels over 100 µg/L in combination with a suspicious lung tumor indicates the presence of SCLC with “high probability”. In our data a single sample has NSE level above 100 µg/L (a squamous cell carcinoma with NSE serum level of $\approx$300 µg/L, probably caused by metastases to the central nervous system), illustrating the difference between patient groups with early and advanced disease. It should also be noted that we have four NET samples (three carcinoids, one SCLC) where the NSE value is 0 µg/L. All the large cell carcinomas with neuroendocrine features had NSE serum levels below upper reference limit. Larger studies are needed to see if this is a general feature for these mixed histologies, and what that might entail. NSE alone did not seem sufficient to distinguishing NET from non-NET in early stage lung cancer, proGRP seems more suited as a sole marker. A high level of proGRP also might give confidence that the tumor is in fact neuroendocrine (high specificity). A larger study trying to find the optimal cutoff for neuroendocrine tumors, where we had a larger number or patients and age and smoking status matching controls are needed to conclude.

There are a few obvious problems regarding our data. One clear disadvantage is the low number of SCLC patients (less than 15%). Also the inclusion of the neuroendocrine carcinoids might show connections which do not exist in the general population, and any rule of thumb derived from patient data containing so many carcinoids might be biased. By excluding them in our machine learning runs, we saw that the limit for proGRP was raised roughly 30-50 units, still giving about the same percentage correctly classified samples. This might be a more correct cutoff for the general lung cancer population, but it does mean we will miss more carcinoids. Another disadvantage is that most of our markers can be elevated in the case of poor renal function, and we have no patient information about renal function.

The machine learning algorithms seem to suggest that high proGRP and high NSE, along with a low SCC(A) is a predictor for SCLC. (What exactly “high” and “low” translates to in serum limits is difficult, and should be a focus for future research.) In our data we can find both SCLC and carcinoids with both normal proGRP and NSE, and elevated SCC(A). This might be because the cancer is so early stage that none of the neuroendocrine markers have yet risen,
and the SCC(A) could be skin cell contamination. Until we have larger studies, with more careful procedures to avoid contamination, this remains speculation. The main point is that it seems like great care should be taken when interpreting serum tumor marker results in early stage lung cancer. The specificity and sensitivity of serum tumor markers in this patient group is not good enough as of today. Our patients are early stage, and serum tumor markers that will help us early in a cancer disease course are more valuable, than in more advanced cancer. A lot of other studies are on patients with predominantly advanced disease, making comparison difficult. We believe that more research should be directed on the group of early stage cancer patients.

ProGRP is of special interest at the laboratory where the tests were run, as they have developed in-house antibody staining of proGRP, for detection of mixed histologies (NSCLC with neuroendocrine features). (53) There is some support to the claim that tumors with mixed histologies behave more like a neuroendocrine tumor than a non-neuroendocrine tumor. To be able to research further on this topic, we must be able to distinguish the NETs from the non-NETs. The most relevant question with regards to immunohistochemistry seems to be if the presence and degree of proGRP positive stained material in NSCLC affects prognosis. It would be interesting to see if the four patients with large cell carcinoma with neuroendocrine features has a disease course more like the other large cell carcinoma patients or the SCLC patients. This could be direction for further research.

In conclusion, more effort should be put into finding optimal reference values for the serum tumor markers, especially against a smoking status adjusted population. ProGRP and possibly NSE seem to be the most valuable serum tumor markers when trying to determine neuroendocrine status, elevated SCC(A) serum levels might aid as an exclusion criteria. Further efforts should go into immunohistochemistry with proGRP antibodies to determine which NSCLC has endocrine features, how this corresponds to serum levels and how this affects prognosis.
Acknowledgements

The author wishes to thank Ingjerd Solvoll, Åslaug Helland and Odd Terje Brustugun at the Department of Oncology, Oslo University Hospital, Radiumhospitalet, for providing serum samples and clinical data and Marius Lund-Iversen at the Department of Pathology, Oslo University Hospital, Radiumhospitalet, for providing immunohistochemistry data and images. I would also like to thank Marianne Nordlund Broughton, Miriam Øijordsbakken and my supervisor Nils Bolstad, all at the Department of Medical Biochemistry, Oslo University Hospital, Radiumhospitalet, for help with project planning, tumor marker measurements and manuscript preparation.
References


54. Nordlund MS. Development, characterization and analytical use of monoclonal antibodies against proGRP with relevance to small cell lung cancer University of Oslo; 2009.
Appendix

Run reports from machine learning algorithms.

CART on full data set, classifying into SCLC and NSCLC

=== Run information ===

Scheme: weka.classifiers.trees.SimpleCart -S 1 -M 2.0 -N 5 -C 1.0
Relation: Full_data_set-weka.filters.unsupervised.attribute.Remove-R1-3,11-23
Instances: 143
Attributes: 8
NSE
CEA
ProGRP
SCC
Chromogranin A
Cyfra-21-1
HE4
NSCLC-SCLC
Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

CART Decision Tree

NSE < 9.735: NSCLC(120.0/5.0)
NSE >= 9.735
  SCC < 2.625: SCLC(10.0/1.0)
  SCC >= 2.625: NSCLC(7.0/0.0)

Number of Leaf Nodes: 3
Size of the Tree: 5
Time taken to build model: 0.01 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 131 91.6084 %
Incorrectly Classified Instances 12 8.3916 %
Kappa statistic 0.3668
Mean absolute error 0.1327
Root mean squared error 0.2791
Relative absolute error 68.8275 %
Root relative squared error 90.9444 %
Total Number of Instances 143

=== Detailed Accuracy By Class ===

<table>
<thead>
<tr>
<th>Class</th>
<th>TP Rate</th>
<th>FP Rate</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Measure</th>
<th>ROC Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSCLC</td>
<td>0.992</td>
<td>0.733</td>
<td>0.92</td>
<td>0.992</td>
<td>0.955</td>
<td>0.618</td>
</tr>
<tr>
<td>SCLC</td>
<td>0.267</td>
<td>0.008</td>
<td>0.8</td>
<td>0.267</td>
<td>0.4</td>
<td>0.618</td>
</tr>
</tbody>
</table>
Weighted Avg. | 0.916 | 0.657 | 0.908 | 0.916 | 0.897 | 0.618

=== Confusion Matrix ===

```
a   b   <-- classified as
127  1 |   a = NSCLC
11   4 |   b = SCLC
```

CART on full data set, classifying into NET and non-NET

=== Run information ===

Scheme: weka.classifiers.trees.SimpleCart -S 1 -M 2.0 -N 5 -C 1.0
Instances: 143
Attributes: 8
NSE
CEA
ProGRP
SCC
Chromogranin A
Cyfra-21-1
HE4
NET
Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

CART Decision Tree

ProGRP < 109.0: Non-NET(100.0/14.0)
ProGRP >= 109.0: NET(23.0/6.0)

Number of Leaf Nodes: 2
Size of the Tree: 3
Time taken to build model: 0.19 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 120 83.9161 %
Incorrectly Classified Instances 23 16.0839 %
Kappa statistic 0.5358
Mean absolute error 0.2492
Root mean squared error 0.3749
Relative absolute error 64.624 %
Root relative squared error 85.5527 %
Total Number of Instances 143

=== Detailed Accuracy By Class ===

<table>
<thead>
<tr>
<th>Class</th>
<th>TP Rate</th>
<th>FP Rate</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Measure</th>
<th>ROC Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-NET</td>
<td>0.943</td>
<td>0.459</td>
<td>0.855</td>
<td>0.943</td>
<td>0.897</td>
<td>0.687</td>
</tr>
<tr>
<td>NET</td>
<td>0.541</td>
<td>0.057</td>
<td>0.769</td>
<td>0.541</td>
<td>0.635</td>
<td>0.687</td>
</tr>
</tbody>
</table>
Weighted Avg.  0.839  0.355  0.833  0.839  0.829  0.687

=== Confusion Matrix ===

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classified as</td>
<td>100</td>
<td>6</td>
</tr>
<tr>
<td>a = Non-NET</td>
<td>17</td>
<td>20</td>
</tr>
<tr>
<td>b = NET</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**C4.5 on full data set, classifying into SCLC and NSCLC**

*** Run information ***

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: Full_data_set-weka.filters.unsupervised.attribute.Remove-R1-3,11-23
Instances: 143
Attributes: 8
NSE
CEA
ProGRP
SCC
Chromogranin A
Cyfra-21-1
HE4
NSCLC-SCLC

Test mode: 10-fold cross-validation

*** Classifier model (full training set) ***

**J48 pruned tree**

------------------

NSE <= 8.66
| ProGRP <= 130: NSCLC (108.0/1.0)
| ProGRP > 130
| | SCC <= 1.63: SCLC (2.0/0.0)
| | SCC > 1.63
| | | Chromogranin A <= 34.67: SCLC (2.0/0.0)
| | | Chromogranin A > 34.67: NSCLC (12.0/0.0)
NSE > 8.66
| SCC <= 2.62: SCLC (10.0/1.0)
| SCC > 2.62: NSCLC (7.0/0.0)

Number of Leaves : 6
Size of the tree : 11

Time taken to build model: 0.02 seconds

*** Stratified cross-validation ***

*** Summary ***

<table>
<thead>
<tr>
<th>Correctly Classified Instances</th>
<th>133</th>
<th>93.007 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorrectly Classified Instances</td>
<td>10</td>
<td>6.993 %</td>
</tr>
<tr>
<td>Kappa statistic</td>
<td>0.5478</td>
<td></td>
</tr>
<tr>
<td>Mean absolute error</td>
<td>0.1024</td>
<td></td>
</tr>
<tr>
<td>Root mean squared error</td>
<td>0.2638</td>
<td></td>
</tr>
<tr>
<td>Relative absolute error</td>
<td>53.1467 %</td>
<td></td>
</tr>
<tr>
<td>Root relative squared error</td>
<td>85.9599 %</td>
<td></td>
</tr>
</tbody>
</table>
Total Number of Instances 143

=== Detailed Accuracy By Class ===

<table>
<thead>
<tr>
<th>Class</th>
<th>TP Rate</th>
<th>FP Rate</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Measure</th>
<th>ROC Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSCLC</td>
<td>0.984</td>
<td>0.533</td>
<td>0.94</td>
<td>0.984</td>
<td>0.962</td>
<td>0.771</td>
</tr>
<tr>
<td>SCLC</td>
<td>0.467</td>
<td>0.016</td>
<td>0.778</td>
<td>0.467</td>
<td>0.583</td>
<td>0.771</td>
</tr>
<tr>
<td>Weighted Avg.</td>
<td>0.93</td>
<td>0.479</td>
<td>0.923</td>
<td>0.93</td>
<td>0.922</td>
<td>0.771</td>
</tr>
</tbody>
</table>

=== Confusion Matrix ===

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>-- classified as</th>
</tr>
</thead>
<tbody>
<tr>
<td>126</td>
<td>2</td>
<td>a = NSCLC</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>b = SCLC</td>
</tr>
</tbody>
</table>

C4.5 on full data set, classifying into NET and non-NET

=== Run information ===

Scheme:weka.classifiers.trees.J48 -C 0.25 -M 2
Instances: 143
Attributes: 8
NSE
CEA
ProGRP
SCC
Chromogranin A
Cyfra-21-1
HE4
NET
Test mode:10-fold cross-validation

=== Classifier model (full training set) ===

J48 pruned tree

------------------
ProGRP <= 106
| Cyfra-21-1 <= 2.47
| | NSE <= 8.04
| | | Chromogranin A <= 35.92
| | | | ProGRP <= 56: NET (6.0/3.0)
| | | | ProGRP > 56: Non-NET (5.0/0.0)
| | | Chromogranin A > 35.92
| | | | Chromogranin A <= 85.91: Non-NET (24.0/0.0)
| | | | Chromogranin A > 85.91
| | | | | SCC <= 3.06: NET (2.0/0.0)
| | | | | SCC > 3.06: Non-NET (3.0/0.0)
| | | | NSE > 8.04: NET (4.0/0.0)
| | Cyfra-21-1 > 2.47: Non-NET (65.0/2.0)

ProGRP > 106
| NSE <= 3.3
| | HE4 <= 100.7: NET (5.0/0.0)
| | HE4 > 100.7
| | | HE4 <= 131.3: Non-NET (6.0/0.0)
|   |   | HE4 > 131.3: NET (2.0/0.0) |
|   | NSE > 3.3: NET (16.0/0.0) |

Number of Leaves : 11
Size of the tree : 21

Time taken to build model: 0.04 seconds

=== Stratified cross-validation ===

Correctly Classified Instances 122 85.3147 %
Incorrectly Classified Instances 21 14.6853 %
Kappa statistic 0.5995
Mean absolute error 0.1697
Root mean squared error 0.3635
Relative absolute error 44.0184 %
Root relative squared error 82.9456 %
Total Number of Instances 143

=== Detailed Accuracy By Class ===

<table>
<thead>
<tr>
<th>Class</th>
<th>TP Rate</th>
<th>FP Rate</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Measure</th>
<th>ROC Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-NET</td>
<td>0.925</td>
<td>0.351</td>
<td>0.883</td>
<td>0.925</td>
<td>0.903</td>
<td>0.779</td>
</tr>
<tr>
<td>NET</td>
<td>0.649</td>
<td>0.075</td>
<td>0.75</td>
<td>0.649</td>
<td>0.696</td>
<td>0.779</td>
</tr>
<tr>
<td>Weighted Avg.</td>
<td>0.853</td>
<td>0.28</td>
<td>0.849</td>
<td>0.853</td>
<td>0.85</td>
<td>0.779</td>
</tr>
</tbody>
</table>

=== Confusion Matrix ===

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>98</td>
<td>8</td>
</tr>
<tr>
<td>b</td>
<td>13</td>
<td>24</td>
</tr>
</tbody>
</table>

CART on data set with carcinoids excluded, classifying into SCLC and NSCLC

=== Run information ===

Scheme: weka.classifiers.trees.SimpleCart -S 1 -M 2.0 -N 5 -C 1.0
Relation: Data_set_carcinoids_excluded-
weka.filters.unsupervised.attribute.Remove-R1-3,11-23
Instances: 125
Attributes: 8
NSE
CEA
ProGRP
SCC
Chromogranin A
Cyfra-21-1
HE4
NSCLC-SCLC
Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

CART Decision Tree
\begin{align*}
\text{ProGRP} < 149.5 \\
&| \quad \text{NSE} < 9.425: \text{NSCLC}(102.0/2.0) \\
&\quad \text{NSE} \geq 9.425 \\
&\quad \quad | \quad \text{SCC} < 2.625: \text{SCLC}(4.0/1.0) \\
&\quad \quad \quad | \quad \text{SCC} \geq 2.625: \text{NSCLC}(5.0/0.0) \\
\text{ProGRP} \geq 149.5 \\
&| \quad \text{NSE} < 4.8: \text{NSCLC}(2.0/1.0) \\
&\quad \text{NSE} \geq 4.8: \text{SCLC}(8.0/0.0)
\end{align*}

Number of Leaf Nodes: 5

Size of the Tree: 9

Time taken to build model: 0.01 seconds

\begin{verbatim}
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances         113               90.4    %
Incorrectly Classified Instances        12                9.6    %
Kappa statistic                          0.4505
Mean absolute error                      0.1232
Root mean squared error                  0.2905
Relative absolute error                 56.9371 %
Root relative squared error             89.2705 %
Total Number of Instances              125

=== Detailed Accuracy By Class ===

\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
Class & TP Rate & FP Rate & Precision & Recall & F-Measure & ROC Area \\
\hline
NSCLC & 0.973 & 0.6 & 0.922 & 0.973 & 0.947 & 0.693 \\
SCLC & 0.4 & 0.027 & 0.667 & 0.4 & 0.5 & 0.693 \\
Weighted Avg. & 0.904 & 0.531 & 0.892 & 0.904 & 0.893 & 0.693 \\
\hline
\end{tabular}

=== Confusion Matrix ===

\begin{tabular}{c|cc}
\hline
 & a & b \\
\hline
107 & a = NSCLC \\
3 & b = SCLC \\
9 & \\
6 & \\
\hline
\end{tabular}
\end{verbatim}

\textbf{CART on data set with carcinoids excluded, classifying into NET and non-NET}

\begin{verbatim}
=== Run information ===
Scheme: weka.classifiers.trees.SimpleCart -S 1 -M 2.0 -N 5 -C 1.0
Relation: Data_set_carcinoids_excluded-
weka.filters.unsupervised.attribute.Remove-R1-3,11-22,24
Instances: 125
Attributes: 8
NSE
CEA
ProGRP
SCC
Chromogranin A
Cyfra-21-1
HE4
\end{verbatim}

NET
Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

CART Decision Tree

ProGRP < 131.0
  | Cyfra-21-1 < 1.27: NET(3.0/0.0)
  | Cyfra-21-1 >= 1.27: Non-NET(103.0/5.0)
ProGRP >= 131.0: NET(11.0/3.0)

Number of Leaf Nodes: 3
Size of the Tree: 5
Time taken to build model: 0.01 seconds

=== Stratified cross-validation ===

Summary

Correctly Classified Instances  110  88  
Incorrectly Classified Instances  15  12  
Kappa statistic  0.4516
Mean absolute error  0.1832
Root mean squared error  0.3244
Relative absolute error  69.8907 %
Root relative squared error  90.2897 %
Total Number of Instances  125

Detailed Accuracy By Class

<table>
<thead>
<tr>
<th>Class</th>
<th>TP Rate</th>
<th>FP Rate</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Measure</th>
<th>ROC Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-NET</td>
<td>0.962</td>
<td>0.579</td>
<td>0.903</td>
<td>0.962</td>
<td>0.932</td>
<td>0.663</td>
</tr>
<tr>
<td>NET</td>
<td>0.421</td>
<td>0.038</td>
<td>0.667</td>
<td>0.421</td>
<td>0.516</td>
<td>0.663</td>
</tr>
<tr>
<td>Weighted Avg.</td>
<td>0.88</td>
<td>0.497</td>
<td>0.867</td>
<td>0.88</td>
<td>0.868</td>
<td>0.663</td>
</tr>
</tbody>
</table>

Confusion Matrix

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>&lt;-- classified as</th>
</tr>
</thead>
<tbody>
<tr>
<td>102</td>
<td>4</td>
<td>a = Non-NET</td>
</tr>
<tr>
<td>11</td>
<td>8</td>
<td>b = NET</td>
</tr>
</tbody>
</table>

C4.5 on data set with carcinoids excluded, classifying into SCLC and NSCLC

Run information

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: Data_set_carcinoids_excluded-
weka.filters.unsupervised.attribute.Remove-R1-3,11-23
Instances: 125
Attributes: 8
NSE
CEA
ProGRP
SCC
Chromogranin A
Cyfra-21-1
HE4
NSCLC-SCLC

Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

J48 pruned tree
-------------------

ProGRP <= 146
|   NSE <= 8.04: NSCLC (102.0/2.0)
|   NSE > 8.04
|   |   SCC <= 2.62: SCLC (4.0/1.0)
|   |   SCC > 2.62: NSCLC (5.0/0.0)
ProGRP > 146
|   NSE <= 4.58: NSCLC (2.0/1.0)
|   NSE > 4.58: SCLC (8.0/0.0)

Number of Leaves : 5
Size of the tree : 9

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances         115               92      %
Incorrectly Classified Instances        10                8      %
Kappa statistic                          0.5421
Mean absolute error                      0.0947
Root mean squared error                  0.2669
Relative absolute error                 43.7595 %
Root relative squared error             82.0012 %
Total Number of Instances              125

=== Detailed Accuracy By Class ===

<table>
<thead>
<tr>
<th>Class</th>
<th>TP Rate</th>
<th>FP Rate</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Measure</th>
<th>ROC Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSCLC</td>
<td>0.982</td>
<td>0.533</td>
<td>0.931</td>
<td>0.982</td>
<td>0.956</td>
<td>0.702</td>
</tr>
<tr>
<td>SCLC</td>
<td>0.467</td>
<td>0.018</td>
<td>0.778</td>
<td>0.467</td>
<td>0.583</td>
<td>0.702</td>
</tr>
<tr>
<td>Weighted Avg.</td>
<td>0.92</td>
<td>0.472</td>
<td>0.913</td>
<td>0.92</td>
<td>0.911</td>
<td>0.702</td>
</tr>
</tbody>
</table>

=== Confusion Matrix ===

a   b   <-- classified as
108   2   |   a = NSCLC
8     7    |   b = SCLC

C4.5 on data set with carcinoids excluded, classifying into NET and non-NET

=== Run information ===

Scheme:weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: Data_set_carcinoids_excluded-
weka.filters.unsupervised.attribute.Remove-R1-3,11-22,24
Instances: 125
Attributes: 8
NSE
CEA
ProGRP
SCC
Chromogranin A
Cyfra-21-1
HE4
NET
Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

J48 pruned tree
------------------

ProGRP <= 130
|   SCC <= 2.85
|       |   NSE <= 8.04
|       |       |   SCC <= 2.23: Non-NET (27.0/0.0)
|       |       |   SCC > 2.23
|       |       |       |   ProGRP <= 93
|       |       |       |       |   NSE <= 2: NET (2.0/0.0)
|       |       |       |       |   NSE > 2: Non-NET (2.0/0.0)
|       |       |       |   ProGRP > 42: Non-NET (11.0/0.0)
|       |       |       |   ProGRP > 93: NET (2.0/0.0)
|       |       |   SCC > 2.85: Non-NET (62.0/0.0)

ProGRP > 130
|   NSE <= 4.58
|       |   CEA <= 1.9: NET (2.0/0.0)
|       |   CEA > 1.9: Non-NET (3.0/1.0)
|   NSE > 4.58: NET (8.0/0.0)

Number of Leaves: 10
Size of the tree: 19

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 109 87.2 %
Incorrectly Classified Instances 16 12.8 %
Kappa statistic 0.4811
Mean absolute error 0.1543
Root mean squared error 0.3464
Relative absolute error 58.8602 %
Root relative squared error 96.4319 %
Total Number of Instances 125

=== Detailed Accuracy By Class ===

TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area

Class
<table>
<thead>
<tr>
<th></th>
<th>0.934</th>
<th>0.474</th>
<th>0.917</th>
<th>0.934</th>
<th>0.925</th>
<th>0.757</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-NET</td>
<td>0.526</td>
<td>0.066</td>
<td>0.588</td>
<td>0.526</td>
<td>0.556</td>
<td>0.757</td>
</tr>
<tr>
<td>NET</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weighted Avg.</td>
<td>0.872</td>
<td>0.412</td>
<td>0.867</td>
<td>0.872</td>
<td>0.869</td>
<td>0.757</td>
</tr>
</tbody>
</table>

=== Confusion Matrix ===

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>classified as</th>
</tr>
</thead>
<tbody>
<tr>
<td>99</td>
<td>7</td>
<td>a = Non-NET</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>b = NET</td>
</tr>
</tbody>
</table>