Bimodality surveillance of high-risk patients for lung cancer

Gordon H Downie

Are new diagnostic strategies providing answers?

Thoracic oncology providers confronted with the task of diagnosing and following patients at risk for cancer of the lung face a number of major dilemmas, some of which directly affect the ability to diagnose. First, the majority of patients with lung cancer are diagnosed at a late stage and <15% survive 5 years, so a degree of nihilism is present in patients, providers and policy makers. Second, risk paradigms are changing, from smoking only to occupational, environmental or home carcinogens to the risk associated with premalignant airway changes. Third, advances in early diagnostic options have the potential to discover lung carcinoma while still in a pre-invasive, minimally invasive stage or diagnose lung cancer while still in a pre-malignant transformation. Fourth, advances in diagnostic imaging have the potential to detect airway changes. Fifth, advances in early diagnostic surveillance have the potential to discover lung carcinoma while still in a pre-invasive, minimally invasive stage or diagnose lung cancer while still in a pre-malignant transformation. Sixth, advances in early diagnostic surveillance have the potential to detect airway changes.

In this issue of Thorax (see page 335) Loewen et al report their initial clinical findings in bimodality surveillance of high risk for lung cancer populations using low dose spiral CT scanning (SCT) and autofluorescence bronchoscopy (AFB). They examined two null hypotheses: (1) AFB was equivalent to conventional sputum cytology (CSC) for the detection of pre-malignant lesions and (2) AFB and SCT would be equivalent to SCT alone for the detection of lung cancer.

The authors conclude that AFB is significantly superior to CSC for the detection of airway pre-malignancy in this cohort of high-risk patients and, in fact, argue that, as a surveillance tool, AFB exceeds the cancer detection rate of colonoscopy in patients with positive fecal occult blood. However, the authors were not able to demonstrate a significant superiority of bimodality surveillance with both AFB and SCT over SCT alone, but question whether a larger sample size would have found bimodality significantly better. Beyond their null hypotheses, the article raises several points that are healthy components of any discussion of the future approach to patients at high risk of lung cancer. These include:

1. Premalignant changes are common (66% of the 169 patients receiving all components of surveillance) in this high-risk cohort.
2. AFB is reasonable in patients with atypia in CSC; however, CSC was inadequate for detection of premalignant cytology when frank carcinoma was not present.
3. Screening and surveillance are very different and surveillance of a select population may be a superior strategy in lung cancer management.
4. Regardless of the histology of the lung cancers detected in this study (>50% were adenocarcinoma), the majority of patients had central airway pre-malignant transformation.
5. Spiral CT scan protocols are not adequate at this time for detecting central airway disease by themselves.
6. Central airway pre-malignant lesions appear to be predictive of the presence of peripheral adenocarcinoma identified by SCT.

Several of these observations or conclusions have not been supported by other articles in the field. Haubinger et al performed a prospective, randomised, multicentre trial comparing white light bronchoscopy (WLB) with or without AFB. The high-risk group defined by chronic obstructive pulmonary disease plus occupational exposure failed to demonstrate severe dysplasia or carcinoma-in-situ (CIS), although it was unclear to what extent metaplasia or mild dysplasia were seen in this cohort. Swensen et al and Bechtel et al in two separate studies used bimodality testing using CSC as one portion of their testing and suggested a more significant contribution for CSC in lung cancer detection than was suggested by Loewen et al. However, because of different study designs including inclusion criteria, biopsy and statistical methods and pathology review variations, it may be nearly impossible to compare findings from one study to another.

Although Loewen et al raise several compelling clinical questions in their paper, the most pivotal may well be management issues of airway cellular transformation including dysplasia and CIS. The diagnosis, progression and treatment of dysplasia and CIS, especially in high-risk populations, are demanding more clinical attention to determine surveillance strategies and may affect overall outcomes of lung cancer in the near future. Intense interest in this topic was indicated when most sessions at the 11th World Congress on Lung Cancer
Barcelona (IASLC) included several abstracts directly or indirectly discussing research or clinical aspects of airway cellular transformation. These presentations ranged from the basic science of biomarkers for cellular transformation to endoscopic detection and surveillance. In clinical thoracic oncology, a discussion of airway transformation usually revolves around three basic questions:

1. What is the natural history of mucosal transformations in the airway?
2. How much cellular transformation needs to be present to warrant treatment: moderate or severe dysplasia, CIS?
3. What is the best approach for detection: routine screening, high risk surveillance or early diagnosis strategies?

Current attempts to address each of these questions are mostly in the form of expert opinion, as seen in a recent review of CIS treatments published by Mathur et al.1

Transformation within the central airways remains an unfavourable diagnosis with obvious grey areas overlapping dysplasia, CIS and frank early carcinoma and, as mentioned earlier, interobserver variability between pathologists on the same specimens is not uncommon. Clinically, there are two important pre-neoplastic non-small cell presentations within the lung: central airway CIS as a pre-neoplasm of squamous cell carcinoma and peripheral lung atypical adenomatous hyperplasia (AAH) as a pre-neoplasm of bronchoalveolar cell carcinoma and adenocarcinoma. By definition, CIS is radiographically occult, has a surface area of less than 2 cm with visible margins, and has no invasion beyond the bronchial cartilage.6 The prevalence of CIS might be as high as 20–30% based on the epidemiology of non-small cell lung cancer, but clinical screening or surveillance studies have not been done on the general population and the declining use of autopsies has further limited datacollection to support this claim.

The issue of whether dysplasia or CIS is a pre-malignant state or an extreme atypia is a critical question. Dysplasia has no established role as a sole modality in the identification of stage 0.15 Progression from metaplasia to invasive carcinoma, although the relative numbers that progress to frank disease are small and variable.10–12 George et al12 have recently reported their experience with a high-risk cohort followed prospectively for several years. They report a cumulative risk of developing lung cancer from high grade lesions in their cohort of 33% and 54% at 1 and 2 years. Pasic et al14 defined “early stage” lung cancer as stage I (T1N0, T2N0) and stage II (T1N1, T2N1) and, as such, dysplasia and CIS (T0N0M0 or stage 0) do not fall within this definition. Other major epithelial-based tumours such as gastro-intestinal and breast cancers rely on screening or surveillance methodology to identify their stage 0 cases and to optimise patient outcomes. CT scanning, positron emission tomography (PET), sputum analysis and bronchoscopy are all used to diagnose and stage early lung cancer, although PET and CT scanning have no established role as a sole modality in the identification of stage 0.16 Since CIS is radiographically occult, diagnosis must come from screening populations at risk for neoplastic transformation, performing surveillance on high-risk populations or early diagnosis in patients with suspicious clinical presentations. Detecting radiographically occult disease relies in part on direct visualisation of the abnormal bronchial mucosa, but the ability to use direct visualisation optimally also depends on why and where CIS develops. Insight into growth patterns and how CIS may progress are important parameters for visualisation strategies. However, despite the ability to assess the entire central airway by standard bronchoscopy and understand its pattern of appearance or association with frank cancer, CIS was infrequently discovered. This has changed with more sophisticated ways to optically analyse airway mucosa, the most significant of which has been AFB.17 AF with white-light bronchoscopy (WLB) has improved both the identification and the extent of airway lesions.11–12–16

AFB has repeatedly shown increased sensitivity when used in combination with WLB over WLB alone. However, the addition of endobronchial ultrasound (EBUS) might improve the specificity of AFB by distinguishing between inflammation and neoplastic penetration of the mucosa. Miyazaki et al11 demonstrated the ability of EBUS to accurately assess tumour depth. In addition to optical technologies, detection of airway CIS will include the molecular genetic changes of the progression from metaplasia to invasive carcinoma and the protein products produced by those changes. Detection of that critical point of irreversible transformation to cancer will inevitably include biomarkers and necessitate multimodality strategies.

The paper by Loewen et al suffers as a stand alone protocol, limiting cohort size and not allowing comparisons to the other clinical trials looking at surveillance in high-risk populations for lung cancer. It does energise this topic by supporting multimodality approaches and defining a population that may benefit more readily from surveillance strategies. It is obvious that thoracic oncology providers have a forest of more questions than answers for managing patients at risk for lung cancer and, although these issues are daunting, Loewen et al have provided us with one potential path through the woods.

**REFERENCES**

Severity assessment in CAP

The CURB65 score displays moderate to good discriminatory value in validation studies involving over 11000 patients

Severity assessment in community-acquired pneumonia: moving on

Wei Shen Lim

Severity assessment is recognised as a pivotal step in the management of community-acquired pneumonia (CAP). Consequently, much effort over the last three decades has gone into developing tools to aid this process. The Pneumonia Severity Index (PSI) was introduced in 1997 following a study in over 50 000 patients, and is well established as a robust severity assessment tool in patients with CAP.1 The CURB65 and CRB65 scores—which take account of the presence of Confusion, raised Urea (in the case of CURB65), raised Respiratory rate, low Blood pressure and age >65 years—were introduced more recently in 2003.2 One of the main benefits of the CURB65 and CRB65 scores is their simplicity in comparison with the PSI which comprises 20 variables. A number of studies over the last two years have therefore sought to confirm the value of these scores in different healthcare settings.

In this issue of Thorax, Man et al3 report a large and well conducted validation study of these three severity assessment tools—the PSI, CURB65 and CRB65 scores (see page 348). They recruited 1016 adults with CAP seen in the emergency department of a teaching hospital in Hong Kong and found that all three severity assessment tools performed equally well at discriminating patients into mortality risk groups. The area under the receiver operating characteristic curve (AUC) is a measure of the ability of a test to correctly classify those with and without the outcome in question, and is widely used to describe the performance of these severity assessment tools. The AUC for the PSI, CURB65 and CRB65 scores were 0.74, 0.73 and 0.69, respectively (a perfect test would have an AUC of 1).

This report raises the current total number of patients studied with respect to the performance of the CURB65 score to over 11 000 patients from nine countries: Australia, England, Hong Kong, New Zealand, Scotland, Spain, Sweden, the Netherlands and the United States.2 4–9 The AUC for the CURB65 score across these validation studies has ranged from 0.73 to 0.87—that is, moderate to good discriminatory value. In comparing the performance of the PSI and CURB65 score, one study from the US found a small but significant difference in favour of the PSI (AUC 0.76 vs 0.81).2 Otherwise, all the other comparative validation studies, including that by Man et al,3 have found no significant difference between these two severity assessment tools.

The performance of the CRB65 score has now been studied in over 5000 patients from seven countries. It appears to be comparable to the CURB65 score with AUC values of 0.69–0.86. The CRB65 score does not require results from any laboratory investigation and is therefore suited to use in the community. However, except for one study from Germany which recruited patients from outpatient clinics,6 most of the work with the CRB65 score has been done either in hospitalised patients or in patients initially seen in emergency departments. Further validation of this score in the primary care or community setting, where it has greatest applicability, is therefore warranted.

Some studies have tested the PSI and CURB65 score against outcome measures such as the need for ICU admission or the combined outcome of mortality and/or need for mechanical ventilation and/or septic shock.10 In these situations they perform less well. This is partly because the PSI and CURB65 scores were developed specifically to predict mortality, and also because these other outcome measures are influenced by centre-specific criteria for ICU admission and/or mechanical ventilation. This is reflected in the varying ICU admission rates in different healthcare settings; for instance, the ICU admission rate in the cohort studied by Man et al3 in Hong Kong was 4% compared with 17% in a study conducted in Spain.11 Importantly, all the validation studies performed in the last few years show that no severity assessment tool, whatever the outcome measure, is perfect (ie, has an AUC of 1), underlining the requirement always to exercise clinical judgement when applying these tools to individual patients.

In last month’s Thorax, Barlow et al4 reported a validation study in 419 patients with clinically diagnosed CAP which compared the CURB65 and CRB65 scores with two generic severity assessment tools—the systemic inflammatory response syndrome (SIRS) score and the standardised early warning score (SEWS). They found that the CURB65 and CRB65 scores performed better than the two generic scores (AUC 0.78 for CURB65, 0.73 for CRB65, 0.68 for SIRS and 0.64 for SEWS).

The value of disease-specific severity scores compared with generic severity scores has been a subject of some debate, particularly in the US where severity adjustment scores have been used alongside managed care. The premise underlying generic scores is that illness severity is a universal concept based on derangements in physiology. Therefore, generic scores allow comparison of patients across different diseases. Conversely, disease-specific scores are based on the