

Reinterpretation of Cross-Sectional Images in Patients with Head and Neck Cancer in the Setting of a Multidisciplinary Cancer Center

Laurie A. Loevner, Adina I. Sonners, Brian J. Schulman, Kerstin Slawek, Randal S. Weber, David I. Rosenthal, Gul Moonis, and Ara A. Chalian

BACKGROUND AND PURPOSE: Patients referred to tertiary care centers frequently arrive with images obtained at outside institutions; these images require reinterpretation. We assessed the clinical value of reinterpreting cross-sectional imaging studies of patients with head and neck cancer, in the setting of a multidisciplinary cancer center.

METHODS: Outside CT and MR images of 136 patients with known or presumed head and neck cancer were reinterpreted by a neuroradiologist. Clinical history and findings on physical examination were available. Reinterpretation was performed before review of outside reports, which were subsequently compared with those generated at the cancer center. Changes in interpretation were noted, and their effects on TNM staging, patient care, and prognosis were assessed. Reliability and statistical significance of rates of change in diagnosis were analyzed with 95% confidence intervals (CIs) and the sign test, respectively. Verification of change in diagnosis was confirmed by pathologic analysis (75%), characteristic radiologic findings (18%), or clinical and imaging follow-up (7%).

RESULTS: Change in interpretation occurred in 56 patients (41%) (95% CI: 33–49%, $P < .001$). Forty-six patients (34%) had a change in T, N, and/or M staging (26–42%, $P < .001$). Change in T stage occurred in 27 cases (20%) (13–27%, $P < .001$) (upstaged in 22, downstaged in five), and a change in N stage in 26 cases (19%) (12–26%, $P < .001$) (upstaged in 20, downstaged in six). Two patients (1.5%) had missed systemic metastases. Three patients with an initial diagnosis of cancer were found to be cancer-free, and six patients had a diagnosis of new second primary cancers that were missed at original interpretation. One patient had a missed middle cerebral artery aneurysm. Changes in image interpretation altered treatment in 55 (98%) of 56 patients and affected prognosis in 53 patients (95%) ($P < .001$).

CONCLUSION: Reinterpretation of cross-sectional images in the setting of a multidisciplinary cancer center has a significant effect on staging, management, and prognosis in patients with head and neck cancer.

Clinicians rely on CT and MR imaging studies for preoperative staging in patients with head and neck

cancer. Although surgeons are accurate at assessing the mucosal surface of the aerodigestive tract, they depend on imaging to determine tumor extent into the adjacent submucosal spaces, which they are unable to assess on physical and endoscopic examinations. Mapping the extent of disease on images allows the clinician to stage the tumor and assess prognosis; it also assists in determining surgical versus nonsurgical disease, as well as the type of procedure that will be performed in surgical candidates. All of these factors ultimately affect how patients will be counseled regarding their disease.

Frequently, patients referred to tertiary care medical centers arrive with imaging studies obtained at outside institutions. For patients who elect to be treated at our institution's multidisciplinary cancer center, clinicians are required to request an official inside interpretation of these outside imaging studies.

Received November 14, 2001; accepted after revision June 4, 2002.

From the Departments of Radiology (L.A.L., A.I.S., K.S., G.M.), Otorhinolaryngology: Head and Neck Surgery (L.A.L., R.S.W., A.A.C.), and Radiation Therapy (D.I.R.), University of Pennsylvania School of Medicine, University of Pennsylvania Medical Center (A.I.S.), and the University of Pennsylvania College of Arts and Sciences (B.J.S.), Philadelphia.

Supported by the RSNA Research and Education Fund Scholar Grant.

Recipient of the Radiologist-in-Training Award for Best Scientific Paper at the annual meeting of the American Society of Head and Neck Radiology.

Address reprint requests to Laurie A. Loevner, M.D., Hospital of the University of Pennsylvania, Department of Radiology, 3400 Spruce St, Philadelphia, PA 19104.

Reinterpretations are performed by a neuroradiologist experienced in head and neck imaging. It is important to note that because such interpretations are performed in a multidisciplinary cancer center, detailed clinical histories and physical examination findings are available. Often, at the outside institutions, such information probably is not available to the radiologist at the time of initial image interpretation.

Because we frequently have noted discrepancies between the original reports and those generated at the cancer center, we decided to prospectively assess the clinical relevance of these changes in interpretation. In this study, we evaluated how often there was a change in image interpretation, and how such changes altered TNM staging, management, and prognosis in patients with cancers affecting the head and neck. The accuracy of the reinterpretation was evaluated by using histopathologic analysis, characteristic radiologic findings such as necrotic nodes, and clinical and imaging follow-up.

Methods

From January 1998 through February 2000, 136 consecutive patients with known or presumed cancer (83 men, 53 women) were referred to, treated at, and followed up at our multidisciplinary cancer center and had outside CT or MR images of the head and neck submitted for reinterpretation at our institution. In this prospective study, the mean age of the patients was 60 years (range, 21–85 years). A total of 73 CT scans and 63 MR images were reinterpreted. Among the 136 patients, the images were initially read in a private practice setting in 111 (82%) and at an academic medical center in 25 (18%). For the purposes of this study, we defined academic medical centers as hospital-based practices with established residency and fellowship programs that were also affiliated with a medical school.

In 106 patients, the following primary aerodigestive tract cancers were diagnosed: pharyngeal ($n = 62$), oral cavity ($n = 25$), and laryngeal ($n = 19$). Twenty-seven patients had other neoplasms: thyroid ($n = 7$), parotid ($n = 5$), cervical nodal metastases in the setting of unknown primary cancers ($n = 5$), sinonasal ($n = 4$), skin ($n = 2$), salivary gland ($n = 1$), lymphoma ($n = 1$), glomus tumor ($n = 1$), and soft-tissue sarcoma ($n = 1$). In addition, three patients with suspected head and neck cancer were found to be tumor-free on image reinterpretation.

To eliminate reader bias, all images were reinterpreted by the same neuroradiologist (L.A.L.) with expertise in head and neck imaging before review of the outside reports. At our multidisciplinary cancer center, in addition to the neuroradiologist, multiple specialists are present, including surgeons, medical oncologists, radiation therapists, and dentists who also design oral prosthetics. At the time of image reinterpretation, the neuroradiologist had available the pertinent patient history and physical examination findings, which was perhaps frequently not the case for the radiologist initially interpreting the images. Subsequently, the interpretations generated at our cancer center were compared with the initial reports, and when the two differed, the nature of the discordance was recorded. The effects of a change in interpretation on TNM staging (proposed by the American Joint Committee on Cancer Staging), management, and prognosis were assessed by the clinical specialists.

In those cases in which there was a change in image interpretation, we assessed the accuracy of the interpretation generated at the cancer center. Patient information was obtained from the head and neck cancer database. Pathologic proof was considered the reference standard and was available in 75% of cases. This was obtained by intraoperative assessment and

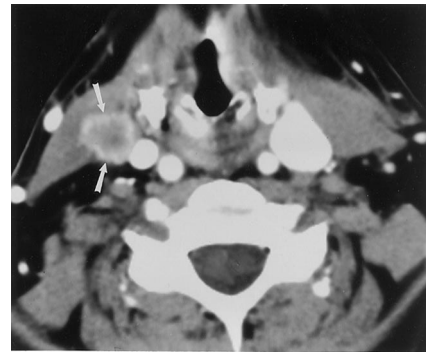


FIG 1. 57-year-old woman with a history of squamous cell carcinoma of the left side of the tongue. Enhanced CT image shows a necrotic regional nodal metastasis (arrows) in the contralateral neck that was detected on reinterpretation in the cancer center, but missed on the initial read. This was pathologically proved following neck dissection.

histologic analysis of tissue obtained from biopsy or surgery. In cases for which surgery was not part of patient care, proof in the form of characteristic radiologic findings was available in 18% of cases. Findings thought to be radiologically characteristic and therefore an accurate measurement of true disease consisted of the presence of necrotic cervical lymph nodes or multiple lung nodules indicating regional and systemic metastatic disease, respectively. Finally, for cases in which neither pathologic proof nor characteristic radiologic findings were available, other means to validate the results were used (7%). This consisted of findings on follow-up physical examination and other radiologic findings that were accepted by the surgeon and radiation therapist treating the patient. Such radiologic findings were normal anatomy that was originally interpreted as an abnormality (such as the submandibular gland being misinterpreted as an enlarged node) and missed extension of tumor outside the confines of the primary site.

To evaluate the reliability of the rates of change, 95% confidence intervals (CIs) were used. To determine statistical significance of the changes in interpretation, the sign test was used.

Results

When the reports generated at the cancer center were compared with the initial outside reports, a change in interpretation occurred in 56 (41%) of 136 cases (95% CI: 0.33–0.49, $P < .001$). There was a change in interpretation in 16 (64%) of 25 images initially read at an academic medical center and 40 (36%) of 111 images initially interpreted at a private practice. These changes in interpretation occurred in 31 (42%) of 73 CT scans and 25 (40%) of 63 MR images. The most common changes were those related to nodes, particularly missed nodal necrosis (Fig 1); changes related to the primary neoplasm, especially under- or overestimation of the extent of the primary tumor; and missed second primary neoplasms (Fig 2) (Table 1). Other nodal changes included mistaking normal anatomic structures (such as the submandibular gland) for nodes, or benign nodes (such as fat-replaced nodes) for pathologic nodes (Fig 3).

Forty-six (34%) of the 136 patients had a change in T, N, and/or M staging (95% CI: 0.26–0.42, $P < .001$) (Table 2). The T stage changed in 27 (20%) of 136 cases (95% CI: 0.13–0.27, $P < .001$). In 22 (81%) of

FIG 2. 49-year-old woman with known cancer of the right side of the tongue and a second primary cancer of the nasopharynx detected at the time of image reinterpretation.

A, Nonenhanced axial CT image shows asymmetry of the nasopharynx, with increased tissue on the left (arrow) and obliteration of the fat along the deep musculature (levator and tensor veli palatini muscles).

B, Nonenhanced coronal CT image again shows increased tissue at the left nasopharynx (arrows). Subsequent biopsy revealed carcinoma.

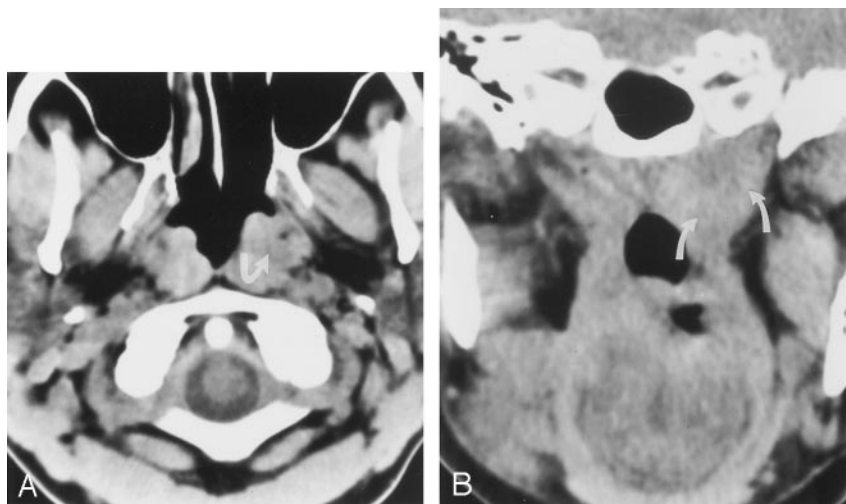


TABLE 1: Categorized changes in image interpretation

Type of Change	No. of Cases
Nodal	
Missed pathologic nodes	18
Cervical	14
Retropharyngeal	4
Nodes misinterpreted as pathologic	3
Submandibular gland mistaken for nodes	2
Changes related to tumor extension	
Missed submucosal extension	5
Parapharyngeal	3
Preepiglottic	2
Presence or absence of cartilage invasion	5
Presence or absence of perineural spread	4
Underestimation of size or extent of tumor	4
Overestimation of size or extent of tumor	3
Changes related to primary neoplasm	
Primary cancer missed on imaging	8
Oral cavity	6
Pharynx	2
Normal anatomy mistaken for primary neoplasm	3
Missed second primary neoplasm	6
Missed metastasis	2
Missed middle cerebral artery aneurysm	1

the 27 cases, there was an upstage, and in five (19%) there was a downstage. The tendency to upstage was statistically significant ($P < .01$). The most common change was identifying advanced disease (upstaging to T3 or T4), which included spread of disease outside the anatomic confines of the origin of the primary tumor, or an underestimation of tumor size. In addition, failure to initially identify invasion of the laryngeal cartilage was common, resulting in a change in the T stage. The N stage changed in 26 (19%) of 136 cases (95% CI: 0.12–0.26, $P < .001$). In 20 (77%) of the 26 cases, there was an upstage and in six cases (23%) there was a downstage ($P < .01$). Nine (7%) of 136 patients had a change in both T and N staging. Two (1.5%) of the 136 patients had missed systemic metastases (skin [$n = 1$], lung [$n = 1$]). Three patients with an initial diagnosis of cancer were subsequently found to be cancer-free. In addition, a new second primary cancer was detected in six patients (Table 2).

Changes in imaging interpretation altered management in 55 (98%) of 56 cases ($P < .001$). Surgical treatment was changed in 39 (70%) of 56 patients ($P < .001$). This included rendering a patient inoperable, altering the type of surgery performed (ie, partial versus total laryngectomy), changing surgical approach or the extent of surgical resection, and/or the need for neck dissection(s) for management of regional nodal metastases. In 29 (52%) of 56 patients, the need for radiation therapy or the radiation portals were altered ($P < .001$). Three patients needed no treatment as they were found to be cancer-free on reinterpretation. Patient prognosis was affected in 53 (95%) of 56 cases ($P < .001$) (worsened in 42 patients and improved in 11).

In 42 (75%) of 56 patients, the change in image interpretation was confirmed pathologically by intraoperative assessment and histologic analysis of either biopsy or surgical specimens. In 10 (18%) of the 56 patients, the change in interpretation was confirmed by characteristic radiologic findings. In the remaining four patients (7%), accuracy of the reinterpretation was confirmed by clinical and radiologic follow-up.

Discussion

In tertiary care referral centers, curbside consults or official reinterpretations of outside imaging studies by a radiologist are common. In patients with head and neck cancer referred to our institution for treatment, reinterpretation of these outside CT and MR studies at the cancer center is always done in conjunction with the full history and physical examination findings. In our study, we found that the report generated at our cancer center differed from that generated on the outside in 41% of cases. In many instances, the initial readings were likely reported without the benefit of the complete clinical history and the findings on physical examination. In addition, in many cases outside interpretations were performed by radiologists with less experience in head and neck imaging. The neuroradiologist at our institution is subspecialized in head and neck imaging, interpreting a high volume (2000) of cancer cases each year.

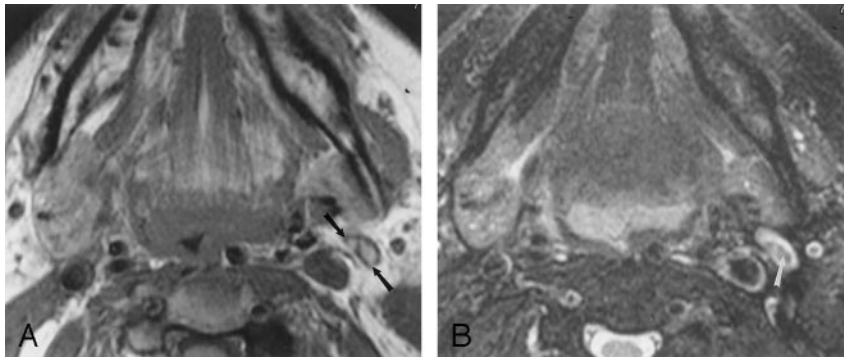


FIG 3. 80-year-old woman with primary pharyngeal cancer. A normal-sized, fat-replaced left jugulogastric lymph node was interpreted as abnormal because of inhomogeneous signal intensity.

A, Axial nonenhanced T1-weighted (600/17/1 [repetition time/echo time/excitations]) MR image shows intrinsic high signal intensity in the lymph node (arrows) consistent with fat.

B, Axial fat-suppressed T2-weighted (4000/80/1) MR image obtained at the same level as that in A shows hypointensity in the hilum of this node (arrow) consistent with suppressed fat.

TABLE 2: Change in tumor staging and means of verification of the change in image interpretation

Patient No.	Change in TNM	Verification of Change in Interpretation	Patient No.	Change in TNM	Verification of Change in Interpretation
1	T ↑, N ↑	Pathologic	24	T ↑	Pathologic
2	T ↑	Pathologic	25	N ↑	Pathologic
3	T ↓, N ↓	Pathologic	26	T ↑, N ↑	Pathologic
4	T ↑	Pathologic	27	N ↑	Pathologic
5	T ↑	Pathologic	28	T ↑	Pathologic
6	N ↑	Pathologic	29	T ↑	Pathologic
7	N ↑	Pathologic	30	N ↑	Pathologic
8	T ↓	Pathologic	31	N ↑	Pathologic
9	T ↑, N ↑	Pathologic	32	T ↑	Pathologic
10	T ↑	Pathologic	33	N ↑	Pathologic
11	T ↑	Pathologic	34	N ↑	Radiologic
12	M ↑	Pathologic	35	T ↑	Radiologic
13	N ↓	Pathologic	36	T ↑, N ↑	Radiologic
14	T ↓	Pathologic	37	N ↓	Radiologic
15	T ↑	Pathologic	38	T ↑, N ↑	Radiologic
16	N ↑	Pathologic	39	T ↓	Radiologic
17	T ↓, N ↓	Pathologic	40	N ↑	Radiologic
18	T ↑	Pathologic	41	M ↑	Radiologic
19	N ↑	Pathologic	42	N ↑	Radiologic
20	N ↑	Pathologic	43	N ↑	Radiologic
21	T ↑	Pathologic	44	N ↓	Clinical/imaging
22	T ↑	Pathologic	45	T ↑	Clinical/imaging
23	T ↑, N ↓	Pathologic	46	T ↑, N ↑	Clinical/imaging

Note.—↑ indicates upstaged; ↓, downstaged; Radiologic, characteristic radiologic findings; Clinical/imaging, clinical and imaging follow-up.

Our study shows the value of reinterpreting images in the context of a multidisciplinary cancer center. Frequently, the change in interpretation in this setting results in a significant change in TNM classification, which directly affects patient management, prognosis, and counseling. Our study found that 34% of patients had a change in TNM staging ($P < .001$) (Table 2). These changes altered patient management and affected prognosis in over 95% and 90% of cases, respectively.

Our results reinforce those of a handful of other studies in which investigators showed that reinterpretation of imaging studies by subspecialized radiologists can affect staging, management, and prognosis in cancer patients (1–7). In a prospective study of patients with lung cancer, investigators found that reinterpretations by a specialized chest radiologist changed TNM staging in one-third of cases and changed the status of surgical versus nonsurgical disease in approximately one-half of these cases (1). In a retrospective study of body CT scans for patients with

biopsy-proved cancer, Gollub et al (2) found a 37% rate of discordance between outside readings and the reinterpretations made in a tertiary care center. This altered treatment in 9% of the discordant interpretations. Since it was a retrospective study, knowledge of the outside readings was available before reinterpretation at the referral center, which biased them from overlooking findings noted in these outside reports. Our study was prospective and differed in that the reinterpretations were performed without knowledge of the findings reported in the initial outside readings. Furthermore, we were able to document the accuracy of our reinterpretations with pathologic proof in most cases, and radiologic or clinical proof in the remaining cases (Table 2).

Other studies have also looked at the accuracy of reinterpretations. Kalbhen et al (3) found clinically significant discrepancies in one-quarter of patients with lung cancer when comparing the reinterpretations performed by an experienced chest radiologist with the initial reports. Furthermore, when assessing

the accuracy between the two reads by using pathologic, imaging, and/or clinical follow-up, these investigators found the reinterpretations to be correct in 95% of cases. In another study that looked at reinterpretation of abdominal and pelvic CT and MR studies in oncology patients, discordant reads were found in 41% of cases. Of these, the second interpretation was found to be correct in 92% of cases (4). Subspecialty expertise by the radiologist was believed to be at least in part responsible for the improved accuracy of the reinterpretations. Bechtold et al (5) analyzed the error rate in the interpretations of abdominal CT scans. They found no statistically significant correlation between error rates and the number of cases read per day by an individual radiologist, the presence of a resident at the time of reading, inpatient versus outpatient cases, or the organ system involved. Therefore, they also believed that a primary determinant of misinterpretations was the skill of the interpreting radiologist.

Another study looked at interobserver variability between radiologists in interpreting MR and CT studies. Bollen et al (8) found a significant interobserver variability among four radiologists in assessing nodal status in lung cancer. In an attempt to explain the inconsistencies, the radiologists' findings were held against histologic analysis of the lymph nodes. These authors suggested that the experience of the reader was the most important factor in determining accuracy. However, they also found other contributing factors such as the volume of cases read in a day, with the error rate doubling when the number of cases read each day was more than 20. From this study, other radiologists concluded that variation among radiologists' observations might arise because margins of nodes may be poorly defined, making it difficult at times to differentiate between adjacent normal-sized nodes and a single enlarged node. Also, nodes in certain areas may be difficult to distinguish from normal structures (9). However, reviewers of this work also concluded that interobserver variability at least in part reflected differences in training among readers.

Similar to these studies, we have shown significant results regarding the value of image reinterpretation. We found a change in interpretation in 41% of our cases, which affected management and prognosis of patients in over 95% of these cases. The most common changes in image interpretation included identification of regional cervical nodal metastases (Fig 1), under- or overestimation of tumor extent, and missed second primary neoplasms (Fig 2) (Table 1). We propose the following factors to explain the greater accuracy of the reinterpretations compared with the initial readings. It is valuable to have a subspecialized radiologist who sees a high volume of cases within an area of interest compared with radiologists in a more generalized practice. Also critical is

the availability of clinical history and physical examination findings at the time of image interpretation. Furthermore, in the setting of a multidisciplinary cancer center, the radiologists have available to them the expertise of their clinical colleagues. We want to stress that the purpose of this study was not to compare academic medical centers with private practice settings. In fact, our results (despite differences in sample size) showed significant rates of misinterpretation for both. Thus, it is not so much the differences between these two settings, but rather the benefits found in the setting of a multidisciplinary cancer center that may lead to more accurate interpretations of imaging studies. Finally, radiologists initially interpreting the images may suspect that certain patients (especially those with significant abnormalities) will be referred to a more specialized medical center, at which time their readings will be expounded upon.

Conclusion

Our study found a change in interpretation in 41% of patients with head and neck cancer when their images were re-read in the cancer center. There was a statistically significant change in TNM staging in 34% of patients ($P < .001$), with management and prognosis affected in over 95% of these cases. These changes in interpretation were found to be more accurate, showing that reinterpretation of cross-sectional imaging studies in the setting of a multidisciplinary cancer center is invaluable.

References

1. Kazerooni NL, Kazerooni EA, Quint LE, Orringer MB. **Added value of thoracic radiology specialist interpretation of CT scans for lung cancer consultations in a thoracic surgery clinic (abstr).** *Radiology* 1998;209(P):171
2. Gollub MJ, Panicek DM, Bach AM, Penalver BS, Castellino RA. **Clinical importance of reinterpretation of body CT scans obtained elsewhere in patients referred for care at a tertiary cancer center.** *Radiology* 1999;210:109-112
3. Kalbhen CL, Yetter EM, Love L, Moncada R, Lawson TL, Albain KA. **Outside film reviews of thoracic imaging studies in oncology patients (abstr).** *AJR Am J Roentgenol* 1998;170(suppl):77
4. Hricak H, Kalbhen CL, Scheidler JE, Schwartz LH, Yu KK, Adams D. **Value of expert interpretation in abdominal oncologic imaging: a multicenter study (abstr).** *Radiology* 1997;205(P):225
5. Bechtold RE, Chen MYM, Ott DJ, et al. **Interpretation of abdominal CT: analysis of errors and their causes.** *J Comput Assist Tomogr* 1997;21:681-685
6. Picolli CW, Schweitzer ME, Mitchell DG, Ricci JA, Levin DC, Harford RJ. **Peer-review quality assessment method for MR imaging: lower-extremity results from 31 centers (abstr).** *Radiology* 1993;189(P):136
7. Wakeley CJ, Jones AM, Kabala JE, Prince D, Goddard PR. **Audit of the value of double reading magnetic resonance imaging films.** *Br J Radiol* 1995;68:358-360
8. Bollen ECM, Goei R, v. 't Hof-Grootenboer BE, Versteeg CW, Engelshove HA, Lamers RJS. **Interobserver variability and accuracy of computed tomographic assessment of nodal status in lung cancer.** *Ann Thorac Surg* 1994;58:158-162
9. Muller NL, Evans KG. **Interobserver variability and accuracy of computed tomographic assessment of nodal status in lung cancer (invited commentary).** *Ann Thorac Surg* 1994;58:162