Radioiodine Ablation Outcomes After Imaging with ^{123}I or ^{131}I: Is No News Good News?

TO THE EDITOR: In a recent paper (*1*), Silberstein reported data from his study assessing outcomes of radioiodine ablation in patients with differentiated thyroid carcinoma after imaging with 2 different isotopes. This study analyzed the results from 49 patients, 26 of whom received ^{123}I before ablation and 23 of whom received ^{131}I before ablation. Acknowledging the difficulties of adequately defining successful ablation, Silberstein reported that 81% of the patients receiving ^{123}I had a successful ablation, compared with 74% of the patients receiving ^{131}I, and that this difference was not statistically significant.

However, we would suggest that the author has overextrapolated from this result to the statement that "the same" ablation rate was achieved, irrespective of diagnostic agent. The logical conclusion of such a statement is that either agent could be used for the purpose, with no loss of patient benefit. Even if true, that conclusion is not demonstrated by Silberstein's study, as it is underpowered to detect what may be clinically significant differences between the techniques. What constitutes such a difference is always difficult to judge, but one might argue that a reduction in the ablation failure rate from 26% to 19% (i.e., nearly a 27% reduction in failures) is clinically significant. A simple power calculation (*2*) would have revealed that to detect the difference between 74% and 81% would require 479 patients for each diagnostic agent. Even if Silberstein had powered his study to look for a bigger difference of 15%, which we believe that most in the oncology community would agree represents a clinical improvement, achieving this difference would have required 71 patients for each diagnostic agent. The power calculations assume a 1-sided χ^2 test, 80% power, and a 0.05 significance level. Conversely, for the patient numbers Silberstein reported, the rate of successful ablations would have needed to rise to 100% for ^{123}I (compared with ^{131}I) for the difference between the techniques to reach statistical significance (Fisher exact test, $P = 0.014$).

The danger of interpreting absence of evidence as absence of negative effects has recently been highlighted in this journal by a letter in which Walter et al. (*3*) made a plea for adequately powered trials. We would add our voice to that plea: Silberstein's study set out to answer an important question that was never going to be answered with the number of patients recruited. When studies are limited by the small number of patients referred through a single hospital or unit, a multicenter approach is the option of choice. Small-scale studies not only represent a waste of resources but also can lead to incorrect conclusions.

REFERENCES

1. Silberstein EB. Comparison of outcomes after ^{123}I versus ^{131}I preablation imaging before radioiodine ablation in differentiated thyroid carcinoma. *J Nucl Med.* 2007; 48:1043–1046.
2. Machin D, Campbell M, Fayers P, et al. *Sample Size Tables for Clinical Studies.* Oxford, U.K.: Blackwell Science; 1997:19.
3. Walter MA, Muller-Brand J, Muller B. Antithyroid drugs and radioiodine and the absence of evidence [letter]. *J Nucl Med.* 2007;48:1403.

Maria T. Burniston
Daniel J. Wilson
Leeds Teaching Hospitals NHS Trust
Leeds, United Kingdom

DOI: 10.2967/jnumed.107.047076

REPLY: I appreciate the thoughtful comments directed to my paper (*1*), wherein I attempted to examine patient outcomes when imaging with ^{123}I versus ^{131}I before ^{131}I ablation. As they point out, it is always important to recognize the possibility, and actually the probability, of type I, or alpha, and type II, or beta, errors in any scientific inquiry.

I must first disagree with the authors' statement that I acknowledged any difficulty in adequately defining successful ablation. As stated in my paper, the determination of successful ablation required both negative diagnostic ^{131}I follow-up findings 6 mo after ablation, performed with the serum thyroid-stimulating hormone elevated in excess of 30 μIU/mL, and, simultaneously, an undetectable level of serum thyroglobulin in the absence of antithyroglobulin antibodies. I examined our data with, and without, the serum thyroglobulin requirement, in order to compare our findings to the majority of data on ^{131}I-induced thyroid ablation—data accumulated over many years and acquired without use of the thyroglobulin assay.

I found that complete ablation, assessed without determining the level of serum thyroglobulin, occurred in 88% of patients initially scanned with ^{123}I and 91% of patients initially scanned with ^{131}I. It is a dubious premise that adding 10 times more patients to the number examined, as suggested by Burniston and Wilson, would have led to the conclusion that the 3.4% difference I described between the 2 groups was real or significant. In the context of the question asked, the same comment applies to the 8.6% difference found in ablation rates between the ^{123}I and ^{131}I groups when these were assessed by both scintigraphic and thyroglobulin criteria. The ablation rates we found are well within the range of those noted in the unflawed studies cited in my paper and have been replicated many times, as a trip to PubMed or any textbook on the topic documents.

However, let us suppose, just for the moment, that we had the time and resources to study 10 times as many patients, as Burniston and Wilson would have us do, and discovered that the 8.6% difference we described between outcomes in the 2 groups was significant. What would be the implications? I believe that clinicians who do not have access to ^{123}I would not be surprised to find that even if stunning were to occur at the dosages used (a concept with which our data and those of our cited references disagree), it would hardly be clinically relevant, because the phenomenon would produce such a small decrease in ablation rates. Of course, we believe, in accordance with the data, that this stunning does not occur.

It is impossible, of course, to disagree with the point made by Burniston and Wilson that with 10 times as many patients studied in this, or any, scientific endeavor, there may occasionally be slightly different outcomes and data interpretations.

REFERENCE

1. Silberstein EB. Comparison of outcomes after ^{123}I versus ^{131}I preablation imaging before radioiodine ablation in differentiated thyroid carcinoma. *J Nucl Med.* 2007;48:1043–1046.

Edward B. Silberstein
University of Cincinnati Medical Center
Cincinnati, Ohio

DOI: 10.2967/jnumed.107.047456

Merits of V/Q SPECT Scintigraphy Compared with CTPA in Imaging of Pulmonary Embolism

TO THE EDITOR: We refer to the recent Invited Perspective which proposes that there is now a reduced role for V/Q scintigraphy in the detection of pulmonary embolism (PE) given the emergence of CT pulmonary angiography (CTPA) (*1*). Although the article addressed many of the strengths and limitations of CTPA in the evaluation of suspected PE, we do not agree with the suggestion that there is now a role for ventilation/perfusion (V/Q) scanning only in a very limited number of patient groups. It is surprising that this review article makes no mention of the use of V/Q SPECT, which has been shown to be superior to planar V/Q scintigraphy (*2,3*) and CTPA (*2*). As with other areas of nuclear medicine, the transition from planar techniques to SPECT has led to improvements in sensitivity and diagnostic accuracy. The published literature on V/Q SPECT has consistently shown improvements in sensitivity (*2,3*), specificity (*3,4*), and intraobserver reproducibility (*2–4*), as well as reducing the number of intermediate or inconclusive results to less than 5% (*3,5*). In a direct comparison of CTPA and V/Q SPECT in 83 patients with suspected PE, Reinartz et al. demonstrated a sensitivity of 97% for SPECT, compared with 86% for CTPA (*2*). The CTPA sensitivity in that paper is consistent with the 83% result quoted in the recent prospective multicenter PIOPED 2 study (*6*). These publications suggest that, even with current-generation CT technology, CTPA fails to diagnose PE in approximately 1 in every 6 patients. Given this failure rate, we consider that few clinicians would agree with Dr. Strashun in his statement that CTPA is "very sensitive" in the diagnosis of PE (*1*). As he correctly states, PE is a potentially fatal condition and its accurate diagnosis is essential. The limitations of CTPA in the detection of smaller emboli (particularly at the segmental and subsegmental levels) are well recognized (*7*); however, larger emboli can also be missed (*6*). Although it has been suggested that emboli not detected by CTPA are not clinically significant, this may not be the case in patients with cardiorespiratory disease (*7*), and in these patients in particular, accurate detection is crucial. Furthermore, the PIOPED 2 study demonstrated that the CTPA accuracy deteriorates further if the scan results do not correlate with the clinical likelihood of disease, and in these circumstances, the incidence of false-positive and false-negative results is significant (*6*). In this paper, 40% of negative CTPA results were false-negative if the clinical suspicion was high, and 42% of positive scan results were false-positive if the clinical suspicion was low. Dr. Strashun states that radiation dose and the risks of contrast media are the only 2 reasons why V/Q scintigraphy should be used in preference to CTPA (*1*). We would suggest the suboptimal sensitivity of CTPA is another reason why V/Q scintigraphy (using SPECT) should be used preferentially in most patients. V/Q SPECT using Technegas (Cyclomedica) also has the advantage of an extremely high negative predictive value, reaching 98.5% in a large prospective series (*5*).

As Dr. Strashun states, there are other limitations with CTPA, including a significant number of technically suboptimal studies (6% in PIOPED 2), high breast radiation exposure (a particular concern in premenopausal women), and contrast-related side effects, such as allergy and renal impairment (*1*). None of these limitations apply to V/Q SPECT scintigraphy. It should also be noted that in the PIOPED 2 study, over 40% of patients did not undergo CTPA because of renal impairment, contrast allergy, or too poor a state of health (*6*). This hardly endorses the notion that CTPA should be regarded as the primary screening test for the imaging of PE. Although CTPA has the advantage of being able to detect other lung diseases, it should be noted that V/Q scintigraphy can detect conditions other than PE (*8*). It is also important to determine whether the other findings detected by CTPA are actually the cause of the patient's clinical symptoms, rather than being ancillary and unrelated. In such cases, the ability of CTPA to detect other diseases may be at the expense of a lower overall sensitivity in the detection of PE. Although there is no doubt that referring clinicians would like a diagnostic imaging test for PE that is highly sensitive and specific for disease detection, as well as being perfectly safe, it is apparent from the PIOPED 2 study that CTPA does not meet these criteria. It is important that referring clinicians understand the limitations of the technology (particularly in relation to sensitivity), despite the appeal that the binary reporting of CTPA may have.

It is unfortunate that V/Q scintigraphy in the United States continues to be done almost universally with planar imaging, and using ^{133}Xe as the ventilation agent in many cases. As long as V/Q scintigraphy is performed in the same way that it was in the 1970s and 1980s, it is evident that the test will struggle to compete with the rapidly evolving CT technology. V/Q SPECT can be adequately performed with diethylenetriaminepentaacetic acid aerosols, and in many European countries as well as Canada and Australia, superior ventilation agents (such as Technegas) are available. V/Q SPECT is clearly superior to planar imaging and, combined with recent developments in computing and camera hardware, has the potential to be further enhanced with quantification and fusion imaging (*9*). In addition to its diagnostic role, V/Q SPECT has the ability to quantify the extent of PE, and this may be valuable in guiding treatment decisions (*10*). For all these reasons, it is disappointing that the Invited Perspective downplays the role of V/Q scintigraphy and completely ignores the advances that have occurred with the technique in regions where SPECT has long replaced planar V/Q scintigraphy.

REFERENCES

1. Strashun AM. A reduced role of V/Q scintigraphy in the diagnosis of acute pulmonary embolism. *J Nucl Med.* 2007;48:1405–1407.
2. Reinartz P, Wildberger JE, Schaefer W, Nowak B, Mahnken AH, Buell U. Tomographic imaging in the diagnosis of pulmonary embolism: a comparison between V/Q lung scintigraphy in SPECT technique and multislice spiral CT. *J Nucl Med.* 2004;45:1501–1508.
3. Bajc M, Olsson CG, Olsson B, Palmer J, Jonson B. Diagnostic evaluation of planar and tomographic ventilation/perfusion lung images in patients with suspected pulmonary emboli. *Clin Physiol Funct Imaging.* 2004;24:249–256.