Maximum Standardized Uptake Values on Positron Emission Tomography of Esophageal Cancer Predicts Stage, Tumor Biology, and Survival

Robert J. Cerfolio, MD, and Ayesha S. Bryant, MSPH, MD

Division of Cardio-Thoracic Surgery, Department of Surgery, University of Alabama at Birmingham, and Department of Epidemiology, University of Alabama at Birmingham School of Public Health, Birmingham, Alabama

Background. The stage of esophageal cancer is currently determined by the anatomic TNM classification system as opposed to information about tumor biology.

Methods. A retrospective review was made of a prospective electronic database. Patients had esophageal cancer, dedicated positron emission tomography (PET) using F-18-fluorodeoxyglucose (FDG-PET) and maximum standardized uptake value (maxSUV) measured. Biopsies were obtained from suspicious nodal and systemic locations, and when indicated, resection with complete lymphadenectomy was performed.

Results. There were 89 patients (53 men). The median maxSUV for patients with high grade dysplasia, stage I, IIa, IIb, III, and IVa esophageal cancer was 1.7, 2.9, 8.9, 7.7, 9.5, and 12, respectively. Multivariate analysis showed patients with a high maxSUV were more likely to have poorly differentiated tumors (risk ratio 1.89, $p = 0.032$) and advanced stage (risk ratio 2.6, $p < 0.001$). The maxSUV correlated better ($r^2 = 0.85$) than the current TNM staging system for survival ($r^2 = 0.68$). Receiving operator characteristics curve demonstrated a maxSUV of 6.6 to be the optimal cut-off point. The 4-year survival of patients with a maxSUV of 6.6 or less was 89%, whereas it was only 31% for those patients with values greater than 6.6 ($p < 0.001$).

Conclusions. The maxSUV of an esophageal cancer on dedicated FDG-PET scan is an independent predictor of stage, tumor characteristics, and survival. It predicts survival better than the current TNM staging system. This information may help guide treatment strategies.

(Ann Thorac Surg 2006;82:391–5)

The treatment of most solid organ tumors, including esophageal cancer, depends on the stage. The stage is currently assessed using the TNM classification system [1], which focuses on the presence or absence of cancer in specific anatomic locations. Despite efforts to accurately stage and treat patients with esophageal cancer, the 5-year survival for those with pathologic stage I, II, and III is only 50%, 23%, and 8%, respectively [2]. Integrated positron emission tomography (PET) with computed tomography (CT) using integrated F-18-fluorodeoxyglucose (FDG) PET/CT or dedicated FDG-PET are increasingly available noninvasive tests that are Medicare approved for the clinical staging of patients with esophageal cancer. The FDG-PET provides a quantitative value of the biological aggressiveness of a malignancy by reporting the maximum standardized uptake value (maxSUV). The maxSUV, which is less variable than the mean SUV [3], represents the amount of metabolic activity (radioactive glucose uptake by the cancer cells) at a pixel. It is calculated by the software contained in the PET machine by a formula that uses variables such as the amount of FDG injected and the patient's weight. Because this value is related to the clinical behavior of a specific tumor in a specific patient, the maxSUV may provide quantitative information that can be used as a tool for guiding therapy as well as predicting prognosis in patients with malignancies [4, 5]. Positron emission tomography centers across the country are moving toward standardizing their techniques so the maxSUV reported by one facility is more translatable to the maxSUV reported by others. Studies have shown that the maxSUV in patients with esophageal cancer may predict resectability [6]. The objective of this study was to assess whether the maxSUV of esophageal cancer predicts survival for patients with this malignancy.

Patients and Methods

Patients

This is a retrospective analysis of an electronic prospective database. Patients who presented to one surgeon between May 2000 and June 2005 with biopsy-proven, apparently resectable (no evidence of T4 or M1 disease) esophageal cancer were eligible. Patients with high-grade dysplasia were also candidates for this study. All patients underwent clinical staging with a CT scan, endoscopic ultrasound fine-needle aspiration (EUS-FNA), and FDG-PET scanning. Patients were excluded if they were less than 19 years old, had a history of type I

Accepted for publication March 6, 2006.

Presented at the Forty-second Annual Meeting of The Society of Thoracic Surgeons, Chicago, IL, Jan 30–Feb 1, 2006.

Address correspondence to Dr Cerfolio, Division of Cardiothoracic Surgery, University of Alabama at Birmingham, 1900 University Blvd, THT 712, Birmingham, AL 35294; e-mail: rcerfolio@uab.edu.

Published by Elsevier Inc

0003-4975/06/$32.00
doi:10.1016/j.athoracsur.2006.03.045

diabetes mellitus, had any chemotherapy or radiotherapy before the maxSUV calculation on PET, or refused entry into the study. The University of Alabama at Birmingham's Institutional Review Board approved both the prospective database used for this study and this trial. Consent was obtained to include patient data in our prospective database, which patients were informed may be used for future studies. The Institutional Review Board waived patient consent for this specific study.

Imaging

The FDG-PET scans were performed on a dedicated ECAT EXACT PET scanner (CTI, Knoxville, Tennessee) or on an integrated PET-CT scanner (Discovery LS PET/CT Scanner; General Electric, Milwaukee, Wisconsin). Patients were asked to fast for 4 hours and then subsequently received 555 MBq (15 mCi) of FDG intravenously followed by PET after 1 hour. The scans were performed from the skull base to midthigh level. Attenuation correction of PET images for the ECAT system was performed with standard transmission scanning using 68 Germanium sources (three rods). The scanning time for emission PET was 6 minutes, and transmission using 68 Germanium rods was 4 minutes per bed position. For the Discovery system, a CT examination was used for attenuation correction of PET images. The scanning time for emission PET was 5 minutes per bed position. Iterative reconstruction with CT attenuation correction was performed. Maximum SUV was determined by drawing regions of interest on the attenuation corrected FDG-PET images around the primary tumor. It was then calculated using the formula [7]:

$$\text{maxSUV} = \frac{C(\mu Ci/mL}{\frac{ID(\mu Ci)}{w(kg)}}$$

In the formula shown above, C is activity at a pixel within the tissue defined by an region of interest, ID is injected dose in μCi, and w is the patient's body weight in kg. The maxSUV within the selected regions of interest was used exclusively.

Procedures, Staging, and Surgery

All patients were clinically staged by CT scan, EUS-FNA, and PET scan. Suspicious sites (on PET defined as a maxSUV $>$ 2.5) were further investigated and pathologically staged. Nodal disease was assessed by EUS-FNA as previously described [8]. Patients with suspected M1 disease in the liver, adrenal, or contralateral lung underwent definitive biopsy to prove or disprove M1 cancer. If the bone or brain was suspected to harbor metastases, magnetic resonance imaging was considered the standard reference.

Patients who were T1N0M0 after staging underwent resection through an Ivor Lewis esophagogastrectomy with complete thoracic lymphadenectomy and removal of celiac and left gastric lymph nodes, as previously described [8]. The final postresection stage was used in this study for these patients. Patients who had metastatic cancer in lymph nodes or had T2 or greater lesions underwent neoadjuvant chemoradiotherapy. For these patients, who were not resected until after the completion of their neoadjuvant therapy, their clinical stage was used, not their postresection pathologic stage. However, nodal disease (N1 and M1a disease) as well as other suspected M1 sites underwent definitive investigation or biopsy before the initial stage assessment and before the start of their neoadjuvant therapy. Pathologic review was performed by standard techniques, and immunohistochemically staining was employed when appropriate. The pathologic stage was assessed using the international staging system [1]. Survival data was obtained through clinic letters, hospital computer information systems, treatment updates, letters from oncologists, the Social Security death index, and telephone calls.

Statistics

The primary endpoint was survival, which was from the date of surgery to the date of the last follow-up or death. If patients received neoadjuvant therapy, the start time was the first date of the initiation of treatment. Patients still alive at the end of our study were censored. A χ^2 analysis or Fisher exact test was used to evaluate discrete dichotomous variables. Analysis of variance was used for discreet nondichotomous variables. For continuous variables, the Student *t* test or the Mann-Whitney *U* test was used to compare means. All comparisons were two-sided with a *p* value of less than 0.05 used to indicate statistical significance. A receiver operating characteristics (ROC) curve was generated to identify the optimal maxSUV value that maximized the specificity and sensitivity of survival. Kaplan-Meier analysis was performed initially to assess for differences among the maxSUV and survival. Univariate analyses were performed with a two-sided log-rank test [9]. Variables with a significant difference between groups based on results of the univariate analyses were entered as candidate variables in a multivariate analysis with a Cox proportional-hazard model with both forward and backward stepwise inclusion of factors, with an inclusion criterion of *p* 0.05 or less . Patients who died within 30 days of surgery or before discharge (operative mortalities) were excluded from the survival analysis. All statistical analysis was performed using SAS v. 9.0 (SAS Institute, Cary, North Carolina).

Results

Patient Characteristics

Of the total of 92 patients, 3 operative deaths were excluded from further analysis, thus leaving 89 patients (53 men) with a median age of 64 years (range, 29 to 81). Patient characteristics, pathology, the median maxSUV, and outcomes are summarized in Table 1. Complications included atrial fibrillation in 8 patients, pneumonia in 7 (caused by aspiration in 4), chylothorax in 2, and liver failure, an ischemic cecum, acute renal failure, superior mesenteric embolus, and deep venous thrombosis in 1 each. The 3 operative deaths were from liver failure,

Table 1. Patient Characteristics, Median Maximum (MaxSUV) Values, and Outcomes for Subgroups

	Number of patients	Median maxSUV	*p* Value[a]
Sex			0.653
Male	53	7.2	
Female	36	6.8	
Histology			0.318
High-grade dysplasia	6	1.7	
Adenocarcinoma	47	8.9	
Squamous cell	32	9.1	
Others	4	4.9	
Stage			0.024
High-grade dysplasia	6	1.7	
I	11	2.9	
IIA	32	8.9	
IIB	14	7.7	
III	18	9.5	
IV	8	10.0	
T status			0.008
High-grade dysplasia	6	1.7	
T1–T2	29	6.0	
T3–T4	24	9.8	
N status			0.009
N0	67	4.9	
N1	22	6.8	
M status			0.755
M0	81	6.9	
M1a	6	7.0	
M1b	2	4.8	
maxSUV			NA
< 6.6	39	NA	
≥ 6.6	50		

[a] Compares the maxSUV values.

maxSUV = maximum standardized uptake value; NA = not applicable.

superior mesenteric embolus with infracted small bowel, and aspiration pneumonia in 1 patient each. There were no anastomotic leaks. There was no relationship between survival and the development of a complication. The median maxSUV of the primary tumor increased as the stage increased ($p = 0.024$). Figure 1 illustrates the ROC curve. It identified 6.6 as the maxSUV value that optimized the sensitivity and specificity for predicting survival (area under curve = 0.85). The Kaplan-Meier univariate 5-year survival was significant for numerical stage ($p = 0.032$), T status (0.010), N status (0.034), and maxSUV equal to or greater than 6.6 or less than 6.6 ($p < 0.001$) (Fig 2). Variables that were found to be independent predictors of survival by Cox hazards regression analysis were TNM staging ($p = 0.032$) and maxSUV ($p = 0.014$). Patients with a maxSUV of greater than 6.6 had a significantly worse survival (31% versus 89%, $p < 0.001$). Linear regression showed a better correlation between maxSUV and survival ($r^2 = 0.85$) as compared with TNM staging and survival ($r^2 = 0.68$).

Comment

The treatment of esophageal cancer, like most solid organ tumors, is dependent on the stage. The current TNM staging system for esophageal cancer is based only on anatomic as opposed to biological factors. However, there is increasing evidence that biological factors influence prognosis just as much, if not more than, anatomical factors [10–12]; FDG-PET may be a noninvasive modality that aids in the detection of some of these genetic, oncologic, and biological factors. Most importantly, it quantifies all these factors after taking into account the patient's own immunosurveillance system. That is probably even more powerful than just biopsying a tumor and performing multiple pathologic tests on it alone. Takita and colleagues [12] in 2003 reported that patients with squamous cell esophageal cancer have a high expression of Glut-1, a glucose transporter that is a marker of poor survival. In this study, we found that squamous cell cancers had a higher median maxSUV than adenocarcinoma.

The findings of our study show that the maxSUV as determined by an FDG-PET scan of a patient with a primary esophageal tumor is an independent predictor of survival. In fact, similar to the report of Sasaki and associates [13] in 2005, our results show that it is a better predictor of survival than the current TNM staging. Interestingly, our findings in this study are quite similar to the ones we reported in 2005 on a study of 315 patients with nonsmall-cell lung cancer. We found the maxSUV value in patients with nonsmall-cell lung cancer to be better correlated to survival, prognosis and recurrence than the TNM staging system [14].

There are strengths and limitations to every study. Strengths of this study include the prospective database

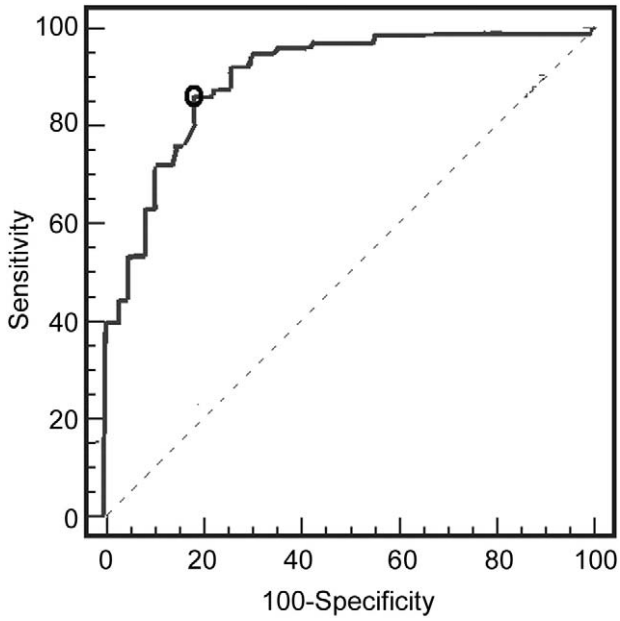

Fig 1. Receiver operating characteristics curve (circle indicates optimal value: sensitivity 87%, specificity 82%).

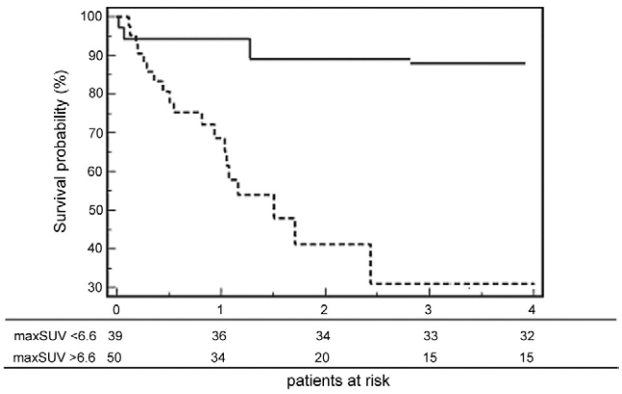

Fig 2. Survival probability of patients with a maximum standardized uptake values (maxSUV) less than 6.6 (solid line) compared with patients whose maxSUV was 6.6 or greater (dashed line; $p < 0.001$).

used, the use of one surgeon, which limits confounders, the use of pathologic instead of clinical staging, the requirement of nodal or metastatic site biopsies, and the careful follow-up. Limitations to this study include the select group of patients chosen (patients with stage IVb disease were not included, as few came to our surgical clinic, and the several who did had chemotherapy or radiotherapy, or both, already started before their initial PET scan). Other limitations include the use of several different PET centers, the selected use of immunohistochemically staining, and the use of several different pathologists.

There are many possible clinical imports of these data. Perhaps a patient with a T1N0M0 esophageal cancer that has a high maxSUV (≥ 6.6) may benefit from neoadjuvant therapy. Perhaps a patient with a high maxSUV tumor, but an early pathologic staged and resected tumor, is more likely to recur systemically and deserves more careful follow-up or even adjuvant chemotherapy. The maxSUV from a FDG-PET scan may provide clues of undiscovered oncogenic, molecular, or biological factors that affect survival. Further studies are needed to answer these provocative questions.

In conclusion, the maxSUV of an esophageal cancer as calculated by dedicated FDG-PET scan is an independent predictor of stage, tumor characteristics and survival. It predicts survival better than the current TNM staging system.

References

1. Sobin LH, Wittekind C. TNM Classification of malignant tumours. 6th ed. New York: Wiley & Sons, 2003.
2. National Cancer Institute, physicians inquiry 2003. Available at: www.nci.gov.
3. Lee JR, Madsen MT, Bushnel D, et al. A threshold method to improve standardized uptake value reproducibility. Nucl Med Comm 2000;21:685–90.
4. Cerfolio RJ, Bryant AS, Ojha B, et al. Improving the inaccuracies of clinical staging of patients with NSCLC: a prospective trial. Ann Thorac Surg 2005;80:1207–13.
5. Mountain CF. Revisions in the international systems for staging lung cancer. Chest 1997;111:1710–7.
6. Van Westreenen HL, Plukker JT, Cobben DC, et al. Prognostic value of the standardized uptake value in esophageal cancer. AJR Am J Roentgenol 2005;185:436–40.
7. Nabi HA, Zubeldia JM. Clinical applications of F18-FDG in oncology. J Nucl Med Technol 2002;30:3–9.
8. Cerfolio RJ, Bryant AS, Ojha B, et al. The accuracy of endoscopic ultrasonography with fine-needle aspiration, integrated positron emission tomography with computed tomography, and computed tomography in restaging patients with esophageal cancer after neoadjuvant chemoradiotherapy. J Thorac Cardiovasc Surg 2005;129:1232–41.
9. Cox DR. Regression models and life-tables. J R Stat Soc 1972;34:187–220.
10. Wijnhoven BP, Pignatelli M, Dinjens WN, et al. Reduced p120ctn expression correlates with poor survival in patients with adenocarcinoma of the gastroesophageal junction. J Surg Oncol 2005;92:116–23.
11. Yamazaki K, Hasegawa M, Ohoka I, et al. Increased E2F-1 expression via tumour cell proliferation and decreased apoptosis are correlated with adverse prognosis in patients with squamous cell carcinoma of the oesophagus. J Clin Pathol 2005;58:904–10.
12. Takita KH, Myazaki T, Nakajima M, et al. Correlation of 18-F-flourodeoxyglucose (FDG) accumulation with glucose transporter (Glut-1) expression in esophageal squamous cell carcinoma. Anticancer Res 2003;23:3263–72.
13. Sasaki R, Komaki R, Macapinlac H, et al. [18F] Fluorodeoxyglucose uptake by positron emission tomography predicts outcome of non-small-cell lung cancer. J Clin Oncol 2005;23:6799–800.
14. Cerfolio RJ, Bryant AS, Ojha B, et al. The maximum standardized uptake values on positron emission tomography of a non-small cell lung cancer predict stage, recurrence and survival. J Thorac Cardiovasc Surg 2005;130:151–9.

DISCUSSION

DR STEVEN R. DEMEESTER (Los Angeles, CA): Rob, very nice and nicely presented. It's difficult to really believe that the size of the tumor is not really what you're looking at, because we know that if you have a bigger tumor, you're going to have more uptake on PET. We also know that if you have a bigger tumor, a longer tumor, it's going to be deeper and more likely to have lymph nodes. By the time you have a 6 cm or 7 cm circumferential tumor, almost 100% of patients will have positive lymph nodes. So it's difficult from your paper here to be sure you're not just looking at the size of the tumor. What would really be convincing is if you showed me similar T-staged tumors, both 2 cm or 3 cm, invasive just into the musculus propria, both N0 or both N1, and then showed me a survival difference based on the maxSUV. Can you give us any of that information, stage-comparable, size-comparable tumors that have different survival based purely on the maxSUV?

DR CERFOLIO: Steve, that's a great question. As you and I spoke before the conference started, that really is the golden ring or brass ring, that we'd like to be able to show that. We showed that in lung cancer because we had the N. We were not really able to show that in this series, but I think that's coming. I can tell you parenthetically that we have had patients with small tumors with hot maxSUVs that have recurred, T1N0, small, but a maxSUV of 8, and I've talked to them about, you know, do we do neoadjuvant. And they say, "Well, Doc, do I

need to?" We say, "No, there're no data. You're T1N0 by EUS." We resect them, and I know of 2 of them who have come back with recurrent systemic cancer, 1 in the liver and 1 in the brain. These are just case reports, bedtime stories, fairy tales. You're right. Hopefully we'll be able to present that one day with a larger N.

DR WAYNE L. HOFSTETTER (Houston, TX): Doctor Cerfolio, I enjoyed your talk very much. As you know, we have an interest in the predictive value of PET scanning at M.D. Anderson as well. Specifically, I wanted to ask a question about one of your conclusions in terms of looking at patients with apparently earlier stage disease and using PET SUV as a potential predictor for more advanced disease. This could potentially categorize a patient as someone you would perhaps want to send on to neoadjuvant therapy or take straight to surgery. I recently queried our database, and I was trying to decipher retrospectively whether the SUV of an FDG-avid esophageal lesion in earlier stage patients had predictive value for stage and outcome. We have a previously published paper showing this correlation in the patients with more advanced disease, and you're showing this as well; you can really see the ones that are very hot, with SUVs in the teens, 20s, et cetera. Those are the ones that you're sending off to get neoadjuvant therapy. The patients who are in the high-grade dysplasia, early cancer, T1, T2 range are the ones you're trying to ferret out here with the SUVs, and frequently you're getting values that are 1, 2, less than 4, in that range. We're seeing the same kind of values come up with Barrett's esophagus with ulceration, and with esophagitis. Did you go back and look at those to see if there was any incumbent esophagitis or Barrett's ulcers that were involved as a confounding factor? Secondly, were you using the SUVs alone as an indicator to take the patient straight to surgery or to refer for neoadjuvant therapy? Thank you and congratulations on an excellent presentation.

DR CERFOLIO: Thank you very much for your comments. As for your two questions, one, we only had 6 patients with high-grade dysplasia, so to really go back and look at which ones are ulcerated or not, probably we wouldn't be able to make much of a determination with that. As for your second question, no, right now we're not using the maxSUV to make determinations for the use of neoadjuvant therapy. We're using EUS-FNA or the presence of biopsy-proven nodal disease or T3 lesions, so if the patients have nodal disease or if they're T3N1. And then the T2N0 is controversial. That's another lecture. But in some of those patients, we'll use neoadjuvant as well. I am not currently using the maxSUV, although it's entering into my conversation with the patients, because when I see that it's high, my bias is that they should go on to neoadjuvant therapy, but just like the patient I said before who was T1N0 who had a maxSUV of 9, we resected that patient without the use of neoadjuvant. These provocative data suggest that maybe we should be considering it in that preoperative treatment, but I have no data to recommend that yet.

DR JOHN R. ROBERTS (Nashville, TN): Robert, I accept that we probably should do more thinking in terms of biological indicators as far as making our treatment decisions, that we don't do that very well yet. I didn't follow how you determined that maxSUV was a better predictor than the TNM staging mechanism. Can you give some further details about that?

DR CERFOLIO: Sure. Well, it's a linear regression analysis. So what you do is you take the one factor, compare it to the other, and see which one is a better predictor of overall survival, and you do that with linear regression, and you use that area under the curve, the R-squared values that I showed you. I didn't want to get too much into a statistical lecture, but that's how that is done. It showed that the maxSUV was actually superior to the TNM. Now, our current TNM for esophageal cancer probably isn't that good, and so you're comparing maybe a guy who has been in the major leagues a few years to a rookie, but on the other hand, the guy who has been in the major leagues a few years looks like he has got to be changed and has got to be tweaked a little bit because we know that that system is not very good.

DR THOMAS FABIAN (New Haven, CT): I was wondering what percentage of your patients underwent neoadjuvant therapy, and that percentage of patients who underwent neoadjuvant therapy, were you restaging them with PET scan? Furthermore, did you analyze the reduction in maxSUV (delta SUV) between the patients before undergoing neoadjuvant and after, and did that correlate in any way to survival? Could you draw any conclusions on what is most important prognostic sign? Is it the initial maxSUV, deltaSUV, or the postinduction SUV?

DR CERFOLIO: That's a great question. We have actually published on that, and that's a separate group of patients, and I won't bore you with more data and more slides on another study we've done, but we've presented this, and I'm not sure if it has been published yet, but it's coming. But the answer is yes, that the delta or the change in the maxSUV absolutely predicts who is a complete responder. When you have greater than an 80% decrease in the maxSUV of the primary, they have a 94% or 95% chance, somewhere in there—I forget the number—a very high percent chance that they are a complete responder, and there have been data before that show those complete responders are most likely to be alive at 5 years. So we have done that. We have presented a study at the Western that re-EUS'd everybody and re-PET'd everybody and showed that their repeat maxSUV was actually a little more accurate than the repeat EUS, which it's sometimes hard to tell a T1 from a T2 because of scar.