New Application of Dual Point $^{18}$F-FDG PET/CT in the Evaluation of Neoadjuvant Chemoradiation Response of Locally Advanced Rectal Cancer

Hai-jeon Yoon, MD,* Seok-ki Kim, MD,* Tae-Sung Kim, MD,* Hyung-Jun Im, MD,* Eun Seong Lee, MD,* Hyun Chul Kim, MD,† Ji Won Park, MD,† Hee Jin Chang, MD,† Hyo Seong Choi, MD,† Dae Yong Kim, MD,† and Jae Hwan Oh, MD†

Purpose: FDG PET/CT has been suggested as the most reliable modality to predict pathological tumor responses after neoadjuvant chemoradiotherapy (CRT) in locally advanced rectal cancer (LARC). However, several confounding factors including radiation-induced inflammation could not be easily avoided with the commonly used single-point FDG PET/CT. Our aim was to evaluate the accuracy of a dual-point PET/CT protocol in LARC response prediction to CRT.

Patients and Methods: Sixty-one LARC patients were enrolled and treated with neoadjuvant CRT. PET/CT was performed before and after CRT. Dual-point acquisition was applied to post-CRT PET/CT. Post-CRT SUVmax (postSUV), pre/post-CRT SUVmax change (RI), and dual-point index (DI) of post-CRT PET/CT were compared with the Dworak tumor regression grade (TRG) as a gold standard. Univariate and multivariate analyses, as well as receiver operating characteristic curve analysis, were used to evaluate the predictive ability of demographic, clinical, and metabolic PET parameters.

Results: Fifteen patients of TRG3-4 were defined as pathological responders, and 46 patients of TRG1-2 were nonresponders. The resulting response index (RI) ranged from $-13$ to 94.8% (59.1 $\pm$ 22.0%), and delay index (DI) ranged from $-45.2$ to 25.0% ($-9.1 \pm 12.1$%). Univariate analysis resulted in PET parameters (postSUV, RI, and DI) as significant predictors ($P = 0.004$, $P < 0.001$, $P < 0.0001$). According to multivariate analysis, RI and DI remained as significant predictors ($P = 0.04$ and $P = 0.0004$). Receiver operating characteristic analysis showed that DI had significantly higher area under the curve compared with RI (0.906 vs 0.696, $P = 0.018$). Delay index had 86.7% sensitivity, 87.0% specificity, 68.4% positive predictive value, 95.2% negative predictive value, and 86.9% accuracy.

Conclusions: Dual-point post-CRT PET/CT can predict pathological tumor response better than conventional single time point pre- and post-CRT PET/CT.

Key Words: rectal cancer, neoadjuvant chemoradiotherapy, dual point PET/CT, Dworak tumor regression grade, DI

Received for publication February 17, 2012; and accepted for publication June 1, 2012.

From the *Department of Nuclear Medicine, and †Center for Colorectal Cancer, Research Institute and Hospital, National Cancer Center, Goyang, Republic of Korea.

Conflicts of interest and sources of funding: This research was supported by the Pioneer Research Center Program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (2011-002110) and the Bio & Medical Technology Development Program of the National Research Foundation (NRF) funded by the Korean government (MEST) (2011-0019794). Reprints: Seok-ki Kim, MD, Department of Nuclear Medicine, National Cancer Center, 111 Jungbalsan-ro, Ilsandong-gu, Goyang-si, Gyeonggi-do 410-769 Republic of Korea. E-mail: sskim@ncc.re.kr.

Copyright © 2012 by Lippincott Williams & Wilkins ISSN: 0363-9762/13/3801-0007

Reprints: Seok-ki Kim, MD, Department of Nuclear Medicine, National Cancer Center, 111 Jungbalsan-ro, Ilsandong-gu, Goyang-si, Gyeonggi-do 410-769 Republic of Korea. E-mail: sskim@ncc.re.kr.

Copyright © 2012 by Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.
acquisition. Early phase imaging was obtained from the skull base to the proximal thigh, and then, delayed phase imaging was obtained covering the lower abdomen and pelvic cavity. All data were acquired in 3-dimensional mode on a single PET/CT system (Discovery STE, General Electric Medical Systems, Waukesha, Wis).

PET/CT Response Assessment

Our experienced nuclear medicine physicians reviewed and analyzed PET/CT scans at the work station with AW 4.3 software (General Electric Medical Systems). For response evaluation, we measured maximum standardized uptake values (SUVmax) by placing standardized volumes of interest (VOIs) on areas of pathologically increased metabolic activity.

Three SUVmax values were measured including SUVmax before CRT (preSUV) and dual-point SUVmax values after CRT (postSUVearly = postSUV, postSUVdelayed). Then, using these 3 values, percentage SUVmax reduction (response index [RI]) and dual-point index (delay index [DI]) were calculated and described as follows:

\[ \text{Response Index} = \frac{(\text{PreSUV}) - (\text{PostSUVearly})}{\text{(PreSUV)}} \times 100 \]

\[ \text{Delay Index} = \frac{(\text{PostSUVearly}) - (\text{PostSUVdelayed})}{\text{(PostSUVdelayed)}} \times 100 \]

Surgical Approach and Pathological Response Assessment

All enrolled patients underwent surgery within 6 to 8 weeks after CRT completion. After operation, experienced pathologists at our institution reviewed surgical specimens and classified pathological tumor regression grade (TRG). TRG was divided into 5 categories according to the Dworak grading system. According to TRG system, TRGs 3 and 4 were considered as the responding group and TRGs 0 to 2 were considered as the nonresponding group.

Statistical Analysis

All statistical analyses were performed using MedCalc software (MedCalc Software version 11.4.4, Belgium). Demographic and clinical parameters including age, sex, clinical stage, histological grade, treatment option, level of CEA, tumor distance from anal verge (AV), and PET/CT parameters, including postSUV, RI, and DI, were compared with pathological TRG as a gold standard.

To evaluate the ability of the parameters to predict pathological response, the Mann-Whitney test or Fisher exact test were used for univariate analysis. Multivariate stepwise logistic regression analysis was used to identify independent predictors among the parameters, which were statistically significant on univariate analysis. The optimal cutoff value of each PET/CT parameter was calculated for response prediction by receiver operating characteristic (ROC) analysis. The area under curve (AUC) of single parameter was compared with each other.

RESULTS

Patients

Sixty-one LARC patients were successfully assessed with PET/CT. The main clinical features of the patients are described in Table 1. After curative resection, 46 cases of TRG1-2 (9 TRG1, 37 TRG2) were defined as pathological responders, and 15 cases of TRG3-4 (8 TRG3, 7 TRG4) were defined as responders according to TRG system.

PET/CT Parameters for Pathological Response Prediction

PostSUV, RI, and DI values of pathological responders and nonresponders were plotted on a scattergraph (Fig. 1A, B, and C). We divided the range of them into 3 zones according to the scatter plot patterns: responder zone, mixed zone, and nonresponder zone. The responder zone is the positive predictive value–maximized range, whereas the nonresponder zone is the negative predictive value–maximized range, and the intervening range is the mixed zone.

The postSUV value ranged from 1.0 to 9.5 (4.4 ± 1.9). The nonresponder zone was more than 1.5 (Fig. 1A). As we can see from the scatterplot, there was considerable overlap between the 2 groups, and 86.7% (13/15) of...
pathological responders were included in the mixed zone. Only 2 pathological responders were included in the responder zone, in which FDG uptakes regressed to the almost background level. The RI ranged from 13% to 94.8% (59.1%). The non-responder zone for RI was less than 20.0%, and that for the responder zone was more than 85.0% (Fig. 1B). More pathological responders distributed toward the responder zone, although most of them still could not escape the mixed zone. The distribution of the responders in the mixed zone slightly decreased from 86.7% (13/15) to 80.0% (12/15).

The DI ranged from 45.2% to 25.0% (9.1%). The responder zone for DI was more than 0%, and that for the nonresponder zone was less than 25.0% (Fig. 1C). Delay index showed less distribution of pathological responders in the mixed zone (26.7%, 4/15) but more distribution in the responder zone than postSUV and RI. Excluding 4 pathological responders who were residing in the mixed zone, all negative DI cases (<0%) were pathological non-responders (Fig. 2).

We analyzed 11 pathological responders who were correctly discriminated by DI. In 2 patients, postSUV decreased to the almost background level of 1.0, and RI was 79.2% and 94.8%. They were discriminated correctly also by postSUV or RI (Fig. 3).

However, in the other 9 patients, residual postSUV ranged from 2.4 to 5.2, and RI ranged from 22.4% to 79.8%; they resided in the mixed zone and could not be correctly discriminated by postSUV and RI. Four patient showed 0.0% of DI (no change between postSUVearly and postSUVdelayed), and other 5 patients showed positive value of DI (decrease of postSUVdelayed compared with postSUVearly) (Fig. 4).

**Univariate and Multivariate Analyses**

Demographic and clinical parameters including age, sex, clinical stage, histological grade, treatment option (CRT, surgery), level of CEA, and tumor distance from AV were not significant pathological response predictors on univariate analysis, whereas all of PET/CT...
parameters including postSUV, RI, and DI were significant (Table 2). Multivariate logistic regression analysis selected RI and DI as significant predictors by excluding postSUV (P = 0.04 and P = 0.0004).

Comparison of Predictive Performances Between PET/CT Parameters

Predictive performances of postSUV, RI, and DI were compared using ROC analyses. Sensitivity, specificity, positive and negative predictive values (PPV and NPV), and accuracy were calculated and compared (Fig. 5).

PostSUV resulted in 66.7% sensitivity, 60.9% specificity, 35.7% PPV, 84.8% NPV, and 63.9% accuracy with 3.7 cutoff value (AUC, 0.703). RI showed 60.0%, 71.7%, 40.9%, 84.6%, and 68.9% with 67.9% cutoff (AUC, 0.696). Delay index showed 86.7%, 87.0%, 68.4%, 95.2%, and 86.9% with $\Delta t_{5.7\%}$ cutoff (AUC 0.906). Delay index parameter resulted in a significantly higher AUC compared with RI (P = 0.02) as well as postSUV (P = 0.03).

Discussion

We introduced dual-time protocol to reduce the influence of previously mentioned confounding factors in neoadjuvant CRT response prediction. In vivo studies demonstrated FDG uptake in tumor cells continuously increase for 3~4 hours, whereas FDG uptake reached their peak after 30~60 minutes in the case of benign lesion. In the case of a remnant tumor, FDG uptake would increase further on the delayed phase, and false positives incurred as a result of partial volume effect would be improved. In the case of radiation-induced inflammation, FDG uptake would be monotonic or decrease on the delayed phase, and false negatives from inflammation would be avoided. Our results were satisfactory by showing that DI resulted in a significantly higher predictive performance with a significantly higher AUC than RI parameter (0.696 vs 0.906, P = 0.02).

Most previous studies with PET or PET/CT showed a sensitivity ranging from 80% to 100% and a specificity ranging from 60% to 90% for predicting pathological responder with conventional RI parameter. However, SUV quantification method, time interval for post-CRT PET or PET/CT, pathological response criteria,
and clinical rationale for ROC cutoff point determination differ among those studies. Considering previously mentioned confounding factors and inconsistent study protocols, our study’s low sensitivity and specificity with RI parameter (60.0% and 71.7%) are not unexpected. Relatively low sensitivity of our study is closely related with a number of pathological responders remained in the mixed zone who showed residual FDG uptake and moderate RI values. These residual FDG uptake and/or moderate RI values seem to be attributed to the inflammatory reaction induced by radiotherapy, or nonspecific physiological bowel uptake. With DI parameter, sensitivity, specificity, and accuracy were increased to 86.7%, 87.0%, and 86.9%, respectively. This improvement indicates that potential false-positive or false-negative cases can be correctly discriminated with DI parameter.

It has been an important issue to determine the optimal time interval for post-CRT PET/CT. Interim PET/CT around 1 or 2 weeks from the start of CRT could be advantageous to predicting a response earlier and modifying CRT protocol before surgery. Although we applied dual-time protocol to presurgical PET/CT scan (5–7 weeks after CRT), this protocol can be applied to any time interval for response monitoring. We speculate that dual-time protocol is independently useful irrespective of time interval after CRT. Even interim protocol cannot avoid previously mentioned confounding factors. Unlike single-time acquisition protocol, dual-time protocol can also distinguish physiological bowel uptake also from residual tumor uptake by comparing uptake pattern of early and delayed phase scan. Another strong point of DI is that it does not need pre-CRT PET/CT acquisition theoretically, although it is necessary for staging. Delay index can be convenient and have advantages with regard to costs and radiation exposure reduction. Prospective study with a combined protocol of dual-point acquisition and interim PET/CT scan will be promising in the future. This study could raise several debates and limitations. First, our institution’s protocol to perform presurgical scan at 5 to 7 weeks after CRT (late interval protocol) could not provide early information about tumor response. Therefore, presurgical modification of CRT protocol was not possible. However, using our late interval protocol, modification of surgical planning according to the predicted tumor response still remains possible. Moreover, the clinical meaning of early or interim protocol has not been fully established yet.

Acute radiation-induced inflammation is known to occur within 6 weeks, whereas chronic radiation inflammation can occur months to years after radiation exposure. Transiently reduced FDG uptake, so-called metabolic stunning, can occur shortly after chemotherapy and/or radiotherapy. In terms of postradiation inflammation, the optimal period is not known to avoid both acute and chronic inflammation. Interim PET/CT around 1 or 2 weeks from the start of CRT could be advantageous to avoid chronic radiation-induced inflammation and to modify CRT protocol before surgery. However, it is even controversial whether interim PET/CT could predict the response with better accuracy. Recently, there were 2 studies that performed multiple interim PET/CT scans during CRT. Rosenberg et al reported that the accuracy of the presurgical scan (76%) was slightly superior to the interim scan on day 14 (72%). However, Janssen et al reported that percentage reduction of SUVmax on day 15 interim PET/CT was the most optimal predictor.

Unlike interim protocol, late-time protocol could be helpful to decrease the influence of acute postradiation inflammation and potential overestimation caused by temporarily stunned tumor cells. The exact time interval is controversial, but at least 6 weeks after the end of therapy has been suggested. Second, more than 50% of our enrolled population was assessed as TRG2, which were defined as the pathological nonresponder group, whereas the portion of TRG3 or 4 was too small (24.5%). TRG2 means moderate regression and somewhat equivocal results. Therefore, we could not exclude the possibility that there may be some underestimated cases among the TRG 2 cases.

Third, FDG uptake could increase with time in pathological responders in contrast to expectations. In reality, 4 pathological responders showed further FDG uptake on delayed phase images and were misclassified with DI on scattergraph (Fig. 1). This phenomenon is somewhat confusing because our study started on the assumption that delayed FDG uptake can discriminate tumor cells from inflammatory cells. However, a previous study with in vivo and clinical data demonstrated that a radiation reaction can cause delayed uptake as well.

CONCLUSIONS

This consecutive study is about the new application of dual-point PET/CT protocol for pathological response prediction in neoadjuvant treatment setting. In our experience, dual-point FDG PET/CT showed better accuracy than single-point PET/CT in predicting response to CRC in LARC patients.

REFERENCES


