Am J Nucl Med Mol Imaging 2021;11(5):428-442
www.ajnmmi.us /ISSN:2160-8407/ajnmmi0134369

Original Article

Reproducibility of FDG PET/CT image-based cancer staging and standardized uptake values with simulated reduction of injected FDG dose or acquisition time

Ryan D Niederkohr[1], Stephen P Hayden[1], James J Hamill[2], Judson P Jones[2], Joshua D Schaefferkoetter[2], Edison Chiu[1]

[1]*Department of Nuclear Medicine, Kaiser Permanente, Santa Clara, CA, USA;* [2]*Siemens Medical Solutions USA, Inc., Knoxville, TN, USA*

Received April 16, 2021; Accepted August 17, 2021; Epub October 15, 2021; Published October 30, 2021

Abstract: ^{18}F-fluorodeoxyglucose (FDG) PET/CT is widely used for oncologic imaging. This study aimed to evaluate, using data simulation, if reduction of injected FDG dose or PET acquisition time could be technically feasible when utilizing a sensitive commercial PET/CT imaging system, without sacrificing image quality, image-based staging accuracy, or standardized uptake value (SUV) accuracy. De-identified, standard of care oncologic FDG PET/CT datasets from 83 adults with lymphoma, lung carcinoma or breast carcinoma were retrospectively analyzed. All images had been acquired using clinical standard dose and acquisition time on a single PET/CT system. The list mode datasets were retrospectively software reprocessed to achieve undersampling of counts, thus simulating the effect of shorter PET acquisition time or lower injected FDG dose. The simulated reduced-count images were reviewed and compared with full-count images to assess and compare qualitative (subjective image quality, stage stability) and semi-quantitative (image noise, SUVmax stability, signal-to-noise and contrast-to-noise ratios within index lesions driving cancer stage) parameters. While simulated reduced-count images had measurably greater noise, there appeared to be no significant loss of image-based staging accuracy nor SUVmax reproducibility down to simulated FDG dose of 0.05 mCi/kg at continuous bed motion rate of 1.1 mm/sec. This retrospective simulation study suggests that a modest reduction of either injected FDG dose or emission scan time might be feasible in this limited oncologic population scanned on a single PET/CT system. Verification of these results with prospectively acquired images using actual low injected FDG activity and/or short imaging time is recommended.

Keywords: Fluorodeoxyglucose (FDG), PET/CT, oncology, staging, radiation dose

Introduction

^{18}F-fluorodeoxyglucose (FDG) positron emission tomography/computed tomography (PET/CT) has proven to be a powerful and frequently used tool for oncologic imaging, with well-established applications for staging, assessment of treatment response and surveillance of multiple cancer types. Historically, FDG PET/CT has been considered a relatively higher-radiation exposure imaging exam resulting from the injected FDG dose (effective dose estimated to be 0.019 mSv/MBq or 0.703 mSv/mCi injected dose in adult patients) plus the concurrent CT dose [1]. Modern PET equipment, with technological improvements in PET detector sensitivity, electronics and reconstruction algorithms, offers a real opportunity to reduce radiation exposure to patients via lower injected FDG dose. Alternatively, this sensitivity could be leveraged to reduce PET scan acquisition time with resultant gains in patient comfort and scanner throughput.

The selection of injected FDG dose often varies considerably among PET imaging centers, and such variation may derive from a variety of factors. US FDA labeling for FDG (various manufacturers) cites a recommended adult dose of 5-10 mCi (185-370 MBq) in the oncologic setting, but does not provide further guidance; specifically, the labeling does not address scan acquisition time, options of dosing adjusted for body weight/size, or accommodations for differences in PET hardware. The 2015 (version 2.0) European Association of Nuclear Medicine

(EANM) procedure guideline for tumor imaging with FDG PET/CT considers this topic in greater detail, offering a mathematic equation to guide selection of FDG dose which accounts for patient body weight and PET bed position overlap [2]. The EANM guideline, however, acknowledges that other relevant parameters such as PET detectors with enhanced sensitivity, increased volume of detectors resulting in extended axial field of view, and continuous bed motion are not considered in these equations and have potential to reduce administered FDG activity. Despite this EANM procedure guideline being jointly approved by the Society of Nuclear Medicine and Molecular Imaging (SNMMI) (https://www.snmmi.org/ClinicalPractice/content.aspx?ItemNumber=6414#Onc, accessed 12/27/20), injected FDG dose in routine clinical practice in the United States often exceeds levels predicted by the EANM guideline. For example, published survey results have reported that use of fixed FDG doses on the order of 10-15 mCi remain relatively commonly at many centers across the United States [3, 4].

The objective of this study was to evaluate, via retrospective simulation using software techniques, if injected FDG dose or PET emission scan time could be feasibly reduced in the oncologic imaging setting, without sacrificing subjective image quality, image-based staging accuracy, or accuracy of standardized uptake value measurements, when utilizing a higher-sensitivity commercial PET/CT imaging system.

Materials and methods

This single-center, data-only study involving retrospective analysis of existing clinical PET/CT images (which were de-identified to protect patient confidentiality) was approved by our institutional review board. All scans included in this study had been performed as part of standard-of-care clinical management and were acquired using the clinical standard technique (i.e., standard FDG dose and PET acquisition time) in place at the time of this cohort's studies.

Patient characteristics and clinical PET/CT protocol

FDG PET/CT scans from 83 adult (age ≥18 years) patients (21 with Hodgkin or high-grade non-Hodgkin lymphoma, 38 with non-small-cell lung carcinoma and 24 with ductal breast carcinoma) were selected for retrospective analysis. Patients were prepared for FDG PET/CT scanning following standard institutional procedures in place at the time of this cohort's studies, which included a pre-scan fasting interval of at least 6 hours, suspension of exogenous insulin administration for at least 6 hours prior to the scan and confirmation of blood glucose less than 200 mg/dl prior to FDG injection. Standard clinical protocol at the time was to inject a fixed FDG dose which averaged 8.0 mCi ± 6% in this cohort. Mean patient body weight of this cohort was 77 kg ± 27% (range, 39.9-157.9 kg), and mean injected FDG dose per kilogram actual body weight was 0.110 mCi/kg ± 25.8%; see **Figure 1**. This injected FDG dose was generally concordant with 2015 EANM/SNMMI guidelines (calculated at 0.095 mCi/kg; assuming bed overlap >30% and assuming that 1.1 mm/sec continuous table motion corresponds with an equivalent acquisition time of 2 minutes per bed position if using conventional "step and shoot" technique). Imaging acquisition began following a radiopharmaceutical uptake time of 60 ± 10 minutes. All scans were acquired using a single Biograph mCT 20 PET/CT scanner (Siemens Medical Solutions; Knoxville, TN, USA) with TrueV (4 rings of LSO crystals with axial field of view coverage of 21.6 cm) and FlowMotion (continuous bed motion) technology. A low-energy, noncontrast CT scan (used for attenuation correction and anatomic localization) was acquired, followed by list mode PET emission scan acquisition with continuous bed motion rate of approximately 1.1 mm/sec.

Data processing to simulate reduction of injected FDG dose or PET acquisition time

All PET/CT datasets were de-identified prior to any study-related tasks. The list mode PET datasets were retrospectively reprocessed using software techniques to achieve undersampling of counts, thus simulating the effect of a shorter PET acquisition time or a lower injected FDG dose. Following techniques validated in prior studies [5], decimated reduced count images were simulated by randomly discarding events in the PET list mode dataset according to 4 preset fractions (80%, 60%, 40% and 20% of full count images), with no attempt made to adjust the ratio of prompt and random events

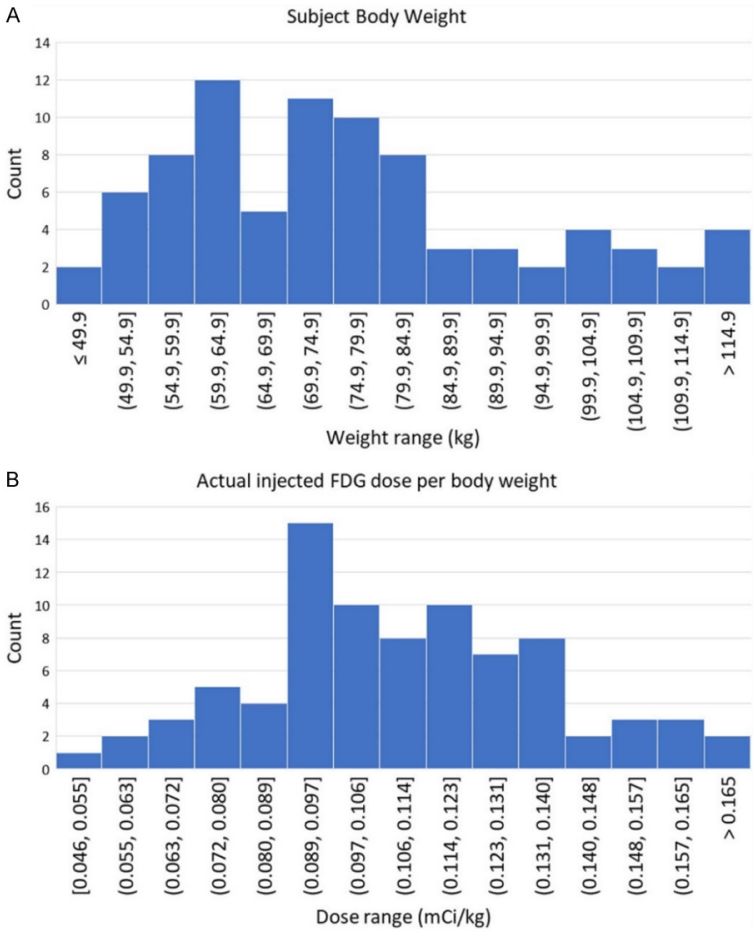

Figure 1. Histograms summarizing patient body weight (A) and actual injected FDG dose per kilogram body weight (B) for scans (N=83) included in this study. Mean patient body weight in this cohort was 77 kg ± 27% (range, 39.9-157.9 kg) and mean injected FDG dose was 0.110 mCi/kg ± 25.8%.

as a function of activity. All datasets were reconstructed using ordinary Poisson iterative reconstruction with attenuation correction and use of point spread function and time-of-flight corrections. Data were reconstructed using 2 iterations, 21 subsets, into a 200×200 (4.07283 mm×4.07283 mm) transaxial volume, postfiltered with a 4 mm isotropic Gaussian, and resampled with the same axial spacing as the corresponding CT series. The number of 5 mm axial slices varied from patient to patient, depending on the continuous bed motion scan length. Decimated images were scaled by the inverse of the decimation fraction to achieve the same quantitative scaling as the 100% image, then converted to DICOM, using volume scaling.

Administered FDG activity per body weight (mCi/kg) was known for all full-count images, and was calculated for the simulated reduction of injected dose at all decimated lower-count images by multiplying full-count level by the decimation fraction (e.g., if full-count/100% image was 0.1 mCi/kg, then 80% image was assigned an simulated dose level of 0.08 mCi/kg). Because the number of true coincidences scales linearly with the in-field activity under usual clinical PET scanning conditions (i.e., situations in which saturation, dead-time, etc. can reasonably be ignored), this calculation based on the assumption of a linear relationship between counts and injected dose is felt to be valid.

Qualitative image analyses

Image quality assessment task: all reconstructed PET images were reviewed by two board-certified nuclear medicine physicians (RN, EC) and an experienced nuclear medicine technologist (SH) who were tasked to subjectively assess overall PET image quality. These reviewers, blinded to all clinical data, visually reviewed and scored every scan using a 5-point scale (1=nondiagnostic, 2=poor, 3=moderate, 4=good, 5=excellent) to reflect subjective quality of the PET images at every decimation level.

Staging task: the two physicians, blinded to all clinical data except scan indication (lymphoma, lung cancer or breast cancer), reviewed PET/CT images (replicating standard clinical interpretative conditions as closely as possible) to arrive at an image-based staging assessment for every patient, following AJCC 8th Edition guidelines for breast carcinoma and lung carcinoma and the Lugano Classification System for Hodgkin and high-grade non-Hodgkin lymphomas [6, 7]. Specifically, nodal (N) stage, distant metastasis (M) stage and overall stage group (0-IV) were recorded for all breast and lung cancer cases, and overall stage group (0-IV) was recorded for all lymphoma cases. Each set of images were reviewed by each observer in a

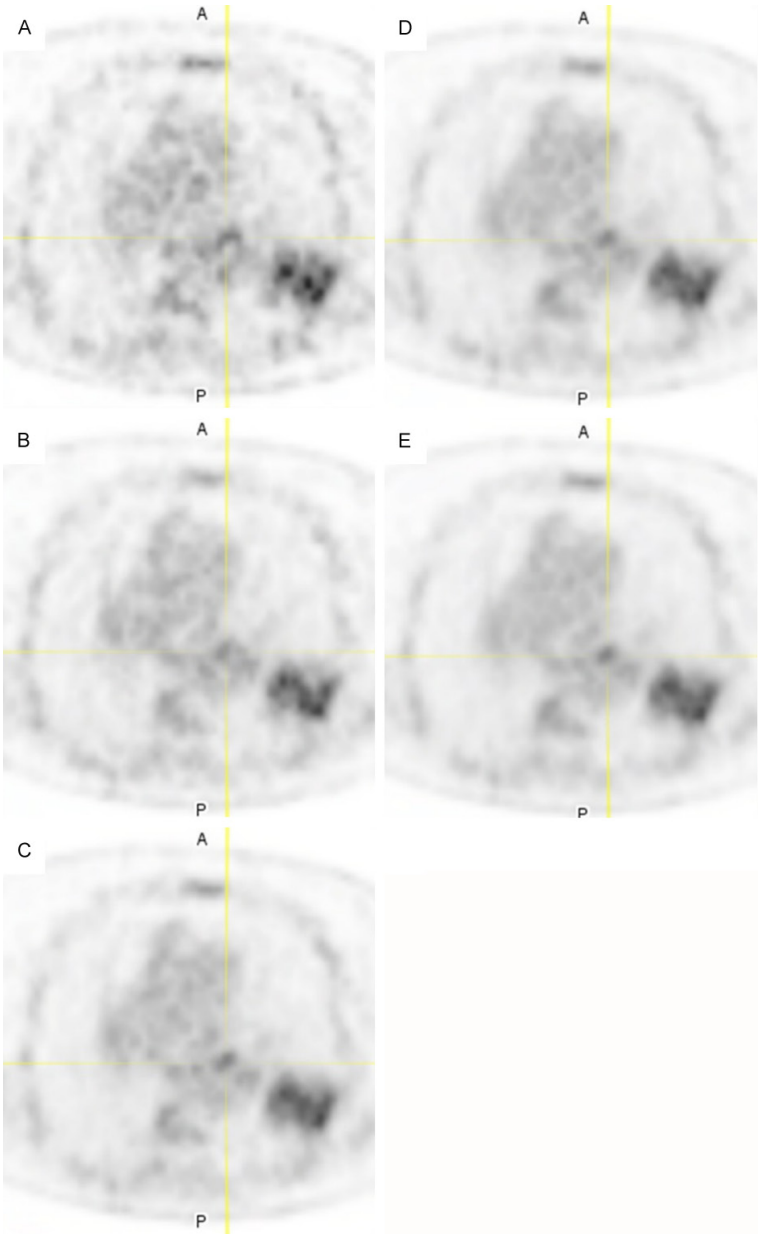

Figure 2. Representative aggregate simulated images at sampled decimations (A: 20%=0.019 mCi/kg; B: 40%=0.038 mCi/kg; C: 60%=0.057 mCi/kg; D: 80%=0.076 mCi/kg; E: full count/0.095 mCi/kg reconstructions), demonstrating a left hilar node index lesion in setting of left lung carcinoma. These demonstrate high noise and poor index lesion contrast on the simulated lower count images (A/B; observer conspicuity and staging confidence scores 1-2), with improving lesion contrast on images with 60% or greater simulated counts (C-E; observer conspicuity and staging confidence scores 3-4).

consistent order, starting with the 20% decimated PET images, followed by the 60% decimated PET images, and then the full-count (100%) PET images. 40% and 80% decimated images were reviewed if there was a staging discordance between 20/60% and 60/100% decimated images, respectively. Subjective observer confidence of overall stage group was assigned using a 4-point scale (1=no, 2=low, 3=moderate, 4=high confidence). Concordance versus discordance of staging determination among the various PET imaging decimations was assessed.

Using the full-count (100%) images, the 1-3 FDG-avid lesions driving the highest overall stage group (called *index lesions*) for each scan were tabulated (total of 154) for subsequent quantitative analytic purposes as detailed below. In addition, each index lesion was subjectively scored by each observer with regard to its visual/qualitative conspicuity using a 4-point scale (1=not detectable or questionably detectable; 2=low contrast, 3=medium contrast, 4=high contrast). A composite set of representative images (to demonstrate these scores) is presented in **Figure 2**.

Quantitative image analyses

Overall image noise: The liver has been widely used to quantify overall FDG PET image noise due to its relatively homogeneous FDG uptake [4]. As a result of Poisson statistics of positron emission, the signal-to-noise ratio of PET images depends on the injected activity and acquisition time, among other parameters. To assess overall image noise, a 3 cm diameter spherical volume of interest (VOI) was placed in normal right hepatic lobe (mid portion, taking care to exclude central ducts and vessels and any FDG-avid lesions, following PERCIST guidelines) on the 100% images and propagated identically onto all lower-decimated images [9]. Mean standardized uptake value (SUVmean) and standard deviation within this

VOI were measured for every (full-count and all simulated lower count) image. Coefficient of variation of liver noise, represented as (standard deviation)/(SUVmean) within the liver VOI, was used as a metric to represent overall image noise for purposes of analysis.

Index lesion analyses: Maximum SUV (SUVmax) was measured for every index lesion on all images. SUVmax as measured on simulated lower-count images was compared with full-count (100%) images to assess for possible bias resulting from count reduction image noise. Signal-to-noise ratio (SNR, defined as [SUVmean or SUVmax in lesion]/[standard deviation of SUV in an adjacent background ROI]) and contrast-to-noise ratio (CNR, defined as [SUVmean in lesion-SUVmean in adjacent background ROI]/[standard deviation of SUV in an adjacent background ROI]) were measured for every index lesion on full count and all decimated images. Use of these metrics has been validated in similar prior studies [8, 10, 11]. The background ROI was one voxel wide and was separated from the lesion by at least one voxel. Although the 100% image was used to define the background region, the same set of voxels was used to define SNR and CNR in the 80, 60, 40 and 20% images.

Results

Image quality and noise assessment

Results of subjective image quality score (scoring system: 1=nondiagnostic; 2=poor; 3=moderate; 4=good; 5=excellent), pooled for all 3 observers, are plotted in terms of relationship with simulated injected FDG dose (mCi/kg) and quantitative liver noise (coefficient of variation; CoV) in **Figure 3**. As expected, subjective image quality was noted to improve with higher simulated injected FDG dose and lower image noise (using liver noise CoV as a marker). Observations were grouped into scores of 1-2 (nondiagnostic or poor quality) and 4-5 (good to excellent quality) or 3-5 (moderate to excellent quality) for analytic comparison. The differences in mean simulated FDG dose for scans with scores 1-2 (mean simulated FDG dose=0.022 mCi/kg; mean liver noise CoV=0.235) versus scores 4-5 (mean simulated FDG dose=0.087 mCi/kg; mean liver noise CoV=0.113) or scores 3-5 (mean simulated FDG dose=0.075 mCi/kg; mean liver noise CoV=0.127) were all statistically significant (T-test, two-sample assuming unequal variances, $P<0.001$). No scores of 1 or 2 (nondiagnostic or poor quality) were observed at a simulated injected FDG dose greater than 0.041 mCi/kg.

The relationship between liver noise (CoV) and simulated injected FDG dose was explored in greater detail, as presented in a scatterplot in **Figure 4**. These data clearly fit the expected physical model [noise=1/sqrt (counts)]. Note is made of a more rapid increase in image noise when simulated injected FDG dose decreased to a level below 0.05 mCi/kg.

Staging task

Data for concordance vs. discordance of PET/CT image-based nodal (N) staging of breast and lung cancers, distant metastasis (M) staging for breast and lung cancers, and overall stage group (0-IV) for breast cancer, lung cancer and lymphoma are presented in **Figure 5**. Results were considered concordant or discordant for each simulated reduced-count image via comparison with its corresponding full-count image (which was considered the reference standard), and are presented for observer 1 individually, observer 2 individually, and pooled results from both observers. Differences in injected dose per body weight between concordant and discordant observations were found to be statistically significant for all comparisons (Mann-Whitney-Wilcoxon test, 2-tailed, $P<0.05$ for all comparisons).

Breast cancer: a total of 24 patients were included, of which 20 had detectable FDG-avid lesions and 4 had no detectable FDG-avid lesion on full-count (100%) scans. Of the total of 96 pooled staging observations by both observers, discordant results were identified in 2 (=2.1%) for nodal (N) stage, 2 (=2.1%) for distant metastasis (M) stage, and 2 (=2.1%) for overall stage group. Neither of the discordant N stage observations resulted in any change in overall stage group. Both of the discordant M stage observations (incorrectly assigned as M0 rather than M1) resulted in a change in overall stage group. The discordant staging observations occurred in a range of simulated injected FDG doses between 0.012 and 0.025 mCi/kg; no discordant staging was observed above 0.025 mCi/kg simulated FDG dose. Of the patients with no detectable FDG-avid disease

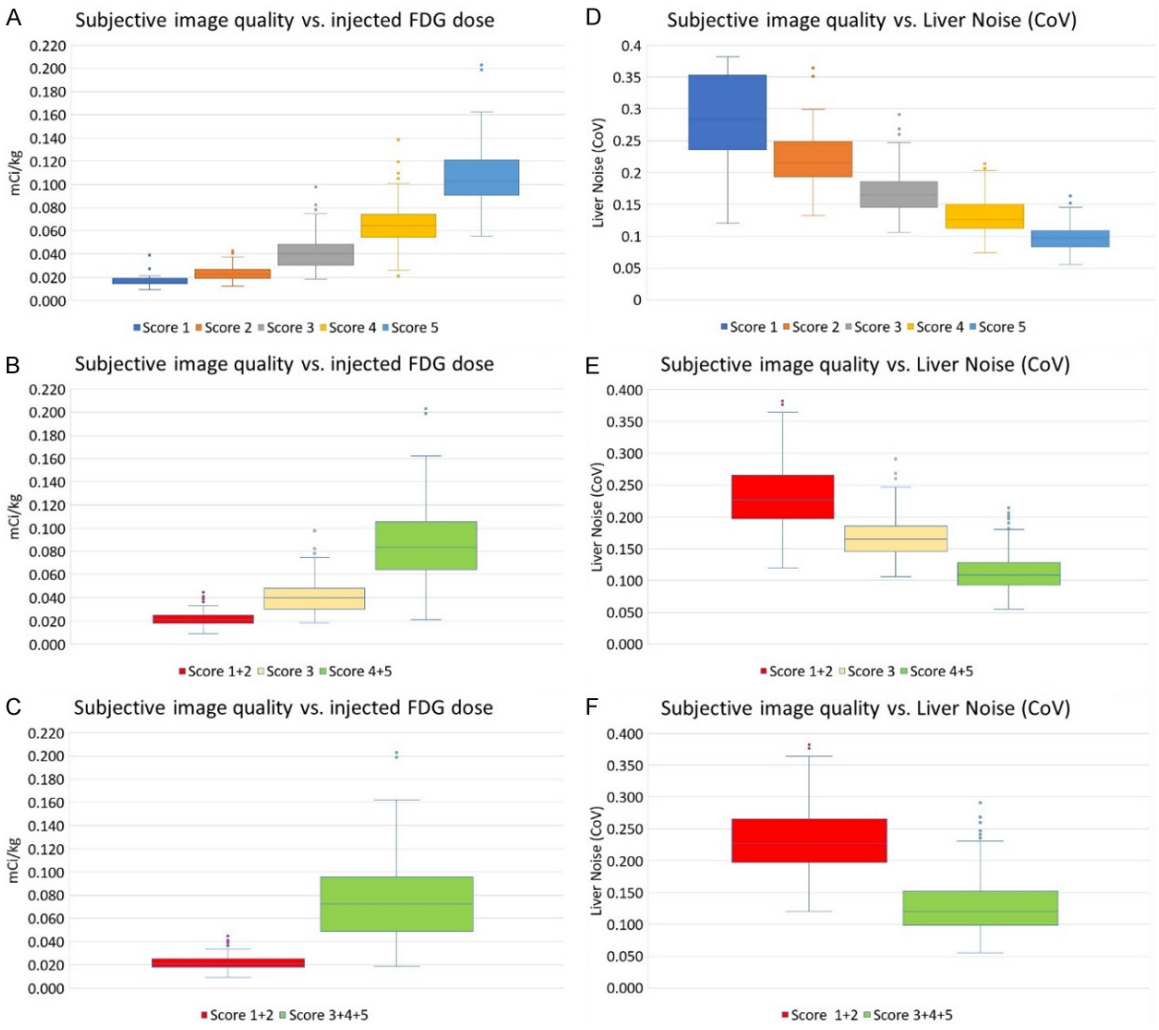

Figure 3. Data for subjective image quality score (scoring system: 1=nondiagnostic; 2=poor; 3=moderate; 4=good; 5=excellent), pooled for all observers and all observations, as related to simulated injected FDG dose level (A-C) and a quantitative measure of overall image noise (coefficient of variation, CoV, of liver SUVmean; D-F) are presented using box plots. As expected, subjective image quality was noted to improve with higher simulated FDG dose and lower image noise. Observations were grouped into scores of 1-2 (nondiagnostic or poor quality) and 4-5 (good to excellent quality) or 3-5 (moderate to excellent quality) for analytic comparison. Differences in simulated FDG dose and liver noise CoV between observations scored 1-2 (mean simulated FDG dose=0.022 mCi/kg; mean liver noise CoV=0.235) and scores 4-5 (mean simulated FDG dose=0.087 mCi/kg; mean liver noise CoV=0.113) or scores 3-5 (mean simulated FDG dose=0.075 mCi/kg; mean liver noise CoV=0.127) were all statistically significant (Kruskal-Wallis test for 3 groups or Mann-Whitney-Wilcoxon test for 2 groups, two-tailed, $P<0.001$). No scores of 1 or 2 (non-diagnostic or poor quality) were observed at a simulated injected FDG dose greater than 0.041 mCi/kg.

on full-count (100%) scans, no false-positive lesions were identified on the simulated lower-dose images.

Lung cancer: a total of 38 patients were included, of which 37 had detectable FDG-avid lesions and 1 had no detectable FDG-avid lesion on full-count (100%) scans. Of the total of 152 pooled observations by both observers, discordant results were identified in 17 (=11.1%) for nodal (N) stage, 3 (=1.9%) for distant metastasis (M) stage, and 15 (=9.9%) for overall stage group. The discordant N stage observations resulted in a change in overall stage group in 10 of 17 patients (in both directions, i.e., some erroneously understaged while others were erroneously overstaged). All discordant M stage observations (all from M0 to M1) resulted in a change in overall stage group. The discordant staging observations occurred in a

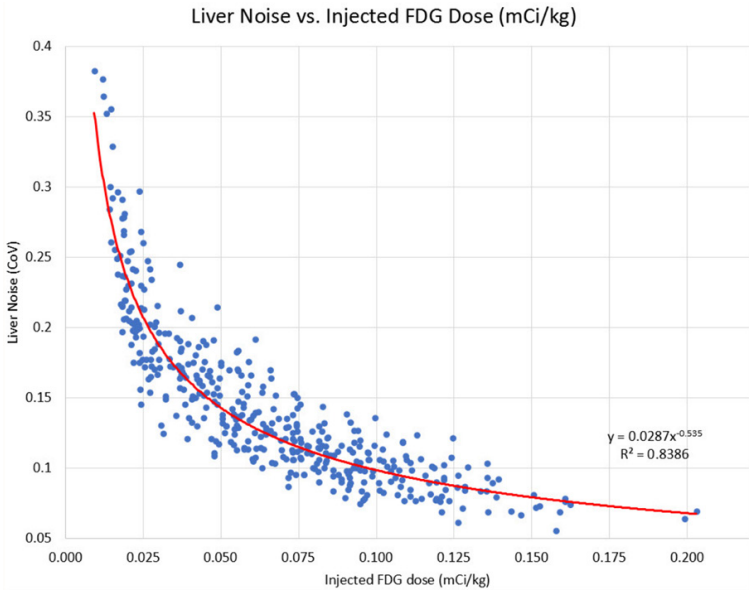

Figure 4. Scatterplot demonstrating the relationship between liver noise (represented as coefficient of variation, CoV, of SUVmean in a 3 cm region of interest in normal liver) and simulated injected FDG dose for all included scans (N=83, each with multiple iterations). Fit with the expected physical model [noise=1/sqrt (counts)] is excellent. The knee of this curve appeared to occur at a simulated injected FDG dose level just below 0.05 mCi/kg, below which noise was noted to increase rapidly.

range of simulated injected FDG doses between 0.014 and 0.05 mCi/kg; no discordant staging was observed above 0.05 mCi/kg. Of the patients with no detectable FDG-avid disease on full-count (100%) scans, no false-positive lesions were identified on the simulated lower-dose images.

Lymphoma: a total of 21 patients were included, of which 15 had detectable FDG-avid lesions and 6 had no detectable FDG-avid lesion on full-count (100%) scans. Of a total of 84 pooled observations by both observers, discordant results were identified in 3 (=3.6%) for overall stage group. Of these discordant results, 2 observations represented erroneous overstaging (i.e., false-positive result) and 1 observation represented erroneous understaging (i.e., false-negative result) on simulated lower-dose images. The discordant staging observations occurred in a range of simulated injected FDG doses between 0.009 and 0.023 mCi/kg; no discordant staging was observed above 0.023 mCi/kg.

Subjective staging confidence scores are presented in **Figure 6** (scoring system: 1=no confidence; 2=low confidence; 3=medium confidence; 4=high confidence), including both full-count (100%) and simulated reduced count images. There was a clear trend toward higher staging confidence as simulated injected FDG dose increased. When rebinned into discrete groups representing "unacceptable" (score 1 or 2) vs. "acceptable" (score 3 or 4) observer confidence, the difference in simulated FDG dose between observations scored 1 or 2 (no or low confidence; mean= 0.023 mCi/kg) and observations scored 3 or 4 (medium or high confidence; mean=0.069 mCi/kg) was found to be statistically significant (Mann-Whitney-Wilcoxon test, two-tailed, P<0.001) for each observer individually and for pooled observations from both observers. No score of 1 or 2 (no or low confidence) was observed above a simulated injected dose of 0.044 mCi/kg.

SUV reproducibility assessment

A scatterplot of SUVmax bias (comparing measured SUVmax on 20/40/60/80% decimated images to 100% images) as a function of simulated injected FDG dose is presented in **Figure 7**. This shows a trend toward higher SUVmax bias as image noise increases (i.e., with reduction of simulated injected FDG dose), and appears to increase exponentially as injected dose levels decrease below 0.04 mCi/kg.

The relationship between SUV bias and simulated injected FDG dose was explored in greater detail to quantify bias within specific FDG dose ranges and at specific simulated FDG dose cutoffs. These analyses are presented in **Figure 8**. For a cutoff of 0.04 mCi/kg: using actual (not absolute) percent bias values, scans with simulated FDG dose of less than 0.04 mCi/kg had mean SUVmax bias of 7.6% (95% CI 5.7-9.5%), while scans with simulated FDG dose of 0.04 mCi/kg or greater had mean SUVmax bias of 1.2% (95% CI 0.7-1.8%); this difference is statistically significant (Mann-Whitney-Wilcoxon test, two-tailed, P<0.01). Using absolute values of percent bias, which represents a more conservative measure since it magnifies

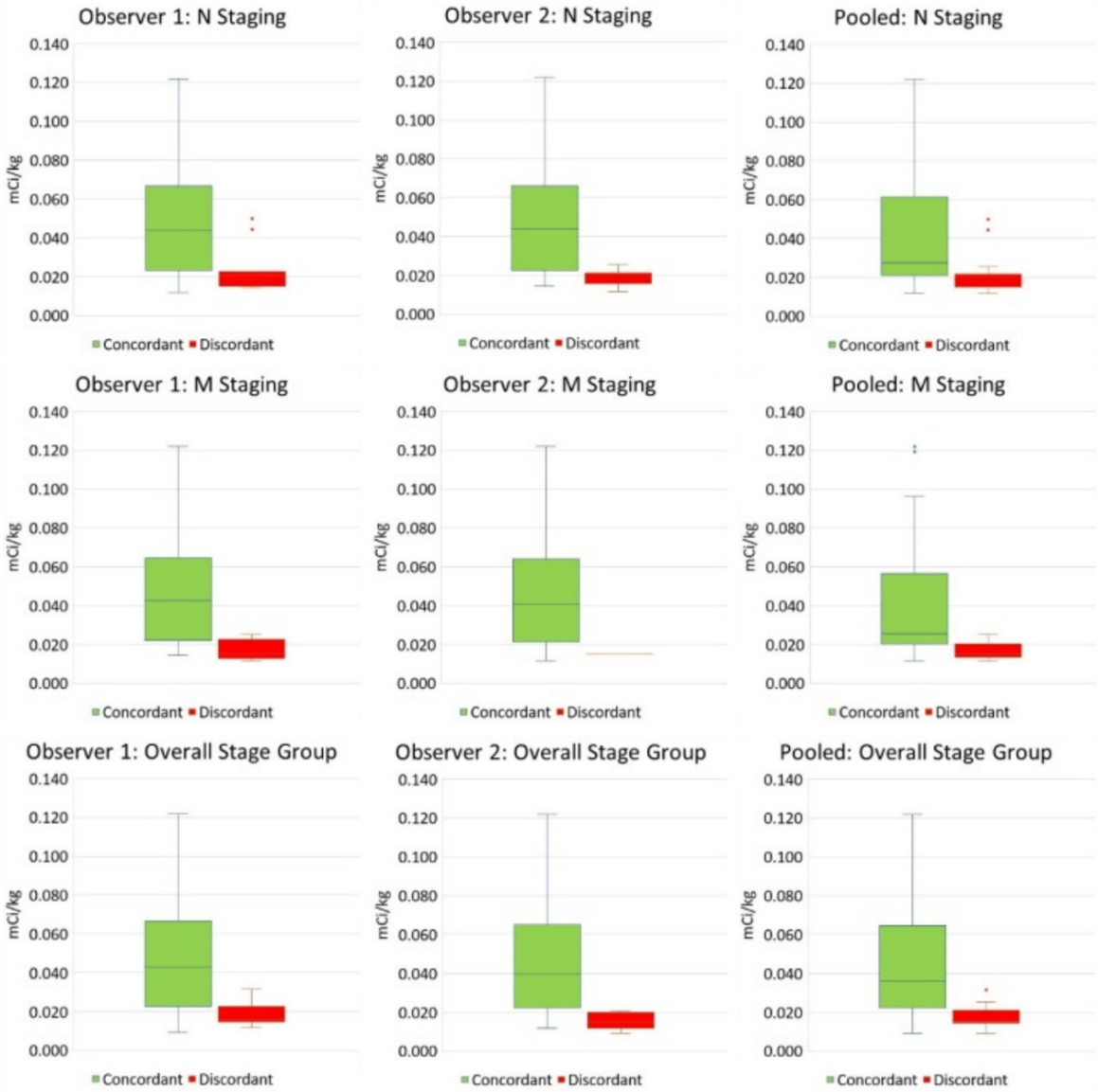

Figure 5. Staging task: data for concordance vs. discordance of PET/CT image-based nodal (N) staging of breast and lung cancer, distant metastasis (M) staging for breast and lung cancer, and overall stage group (0-IV) for breast cancer (N=24), lung cancer (N=38) and lymphoma (N=21) cases, as related to injected FDG dose level, are presented using box plots. Results were considered concordant or discordant for each simulated reduced-count image via comparison with its corresponding full-count image (which was considered the reference standard), and are presented for observer 1 individually, observer 2 individually, and pooled results from both observers. Differences in injected dose (actual for 100% images, simulated for lower-decimated images) between concordant and discordant observations were found to be statistically significant for all comparisons using Mann-Whitney-Wilcoxon test (2-tailed, $P<0.05$).

the effect of differences in the mean, scans with simulated FDG dose of less than 0.04 mCi/kg had mean SUVmax bias of 12.0% (95% CI 10.6-13.4%), while scans with simulated FDG dose of 0.04 mCi/kg or greater had mean SUVmax bias of 4.0% (95% CI 3.6-4.4%); this difference is also statistically significant ($P<0.01$). 93.8% of scans with simulated FDG dose of 0.04 mCi/kg or higher had bias less than 10% and 98.0% of scans had bias less than 15%.

For a cutoff of 0.05 mCi/kg: using actual (not absolute) percent bias values, scans with simulated FDG dose of less than 0.05 mCi/kg had mean SUVmax bias of 3.4% (95% CI 2.0-4.9%),

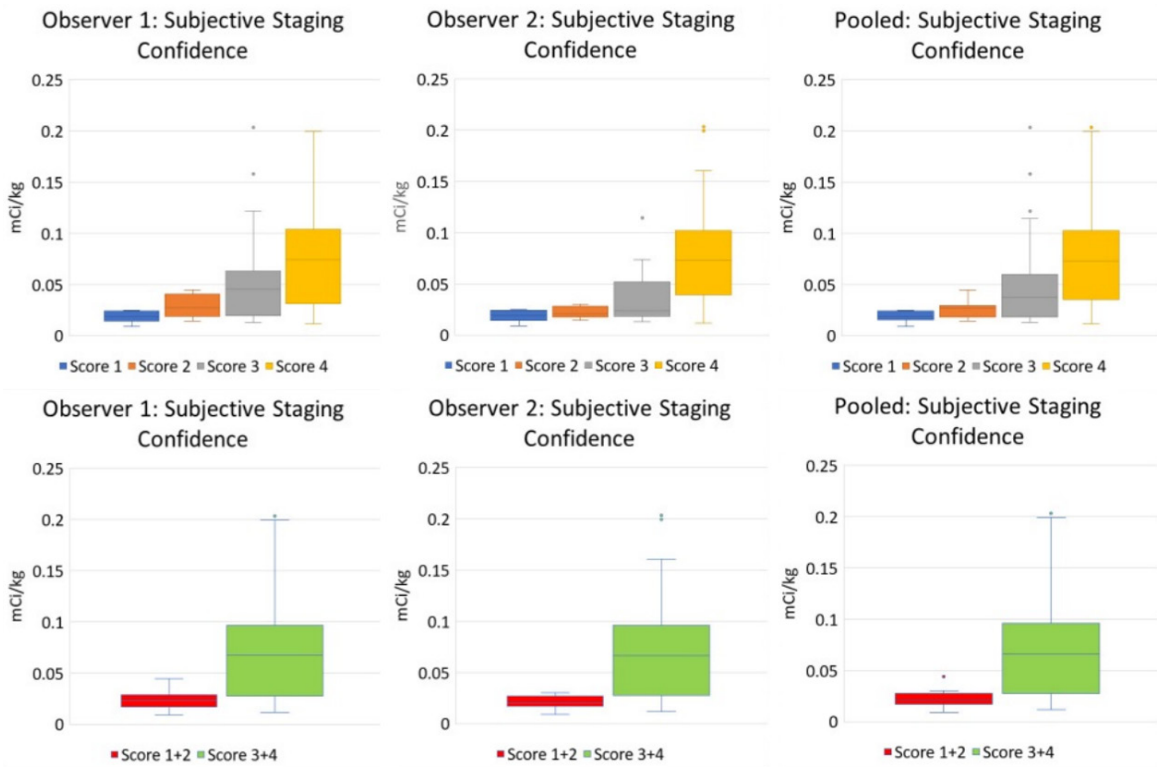

Figure 6. Staging task: subjective observer confidence scores (scoring system: 1=no confidence; 2=low confidence; 3=medium confidence; 4=high confidence), as related to simulated injected FDG dose level are presented using box plots. There was a clear trend toward higher observer staging confidence as simulated injected FDG dose increased. When re-binned into discrete groups representing "unacceptable" (score 1 or 2) vs. "acceptable" (score 3 or 4) observer confidence, the difference in simulated FDG dose between observations scored 1 or 2 (no or low confidence; mean=0.023 mCi/kg) and observations scored 3 or 4 (medium or high confidence; mean=0.069 mCi/kg) was found to be statistically significant (Mann-Whitney-Wilcoxon test, two-tailed, P<0.001) for each observer individually and for pooled observations from both observers.

while scans with simulated FDG dose of 0.05 mCi/kg or greater had mean SUVmax bias of 0.8% (95% CI 0.3-1.3%); this difference is statistically significant (Mann-Whitney-Wilcoxon test, two-tailed, P<0.01). Using absolute values of percent bias, scans with simulated FDG dose of less than 0.05 mCi/kg had mean SUVmax bias of 9.3% (95% CI 8.1-10.4%), while scans with simulated FDG dose of 0.05 mCi/kg or greater had mean SUVmax bias of 3.5% (95% CI 3.1-3.8%); this difference is also statistically significant (P<0.01). 95.2% of scans with simulated FDG dose of 0.05 mCi/kg or higher had bias less than 10% and 99.1% of scans had bias less than 15%.

For simulated FDG doses between 0.04 and 0.06 mCi/kg, using absolute values, there was no statistically significance difference in SUVmax bias among the subgroups (0.04-0.045, 0.045-0.5, 0.05-0.055, 0.055-0.06) as tested (Kruskal-Wallis test, P=0.483).

Index lesion conspicuity assessment

Results of subjective index lesion (N=154) conspicuity, pooled for both observers, are plotted in terms of relationship with contrast-to-noise ratio (CNR, using lesion SUVmean) and signal-to-noise ratio (SNR, using lesion SUVmax and SUVmean) in **Figure 9** (scoring scale: 1=not detectable or questionably detectable; 2=low contrast, 3=medium contrast, 4=high contrast). As expected, lesions with higher visual/qualitative observer scores (3 or 4) had overall higher (mean) CNR and SNR. The differences in CNR and SNR between these groups were found to be statistically significant (Mann-Whitney-Wilcoxon test, 2-tailed, P<0.001). However, there was substantial overlap in CNR and SNR between the groups, with some lesions that were scored as having high visual/qualitative contrast having rather low CNR and SNR, in the same range of other lesions with low visual scores (1 or 2).

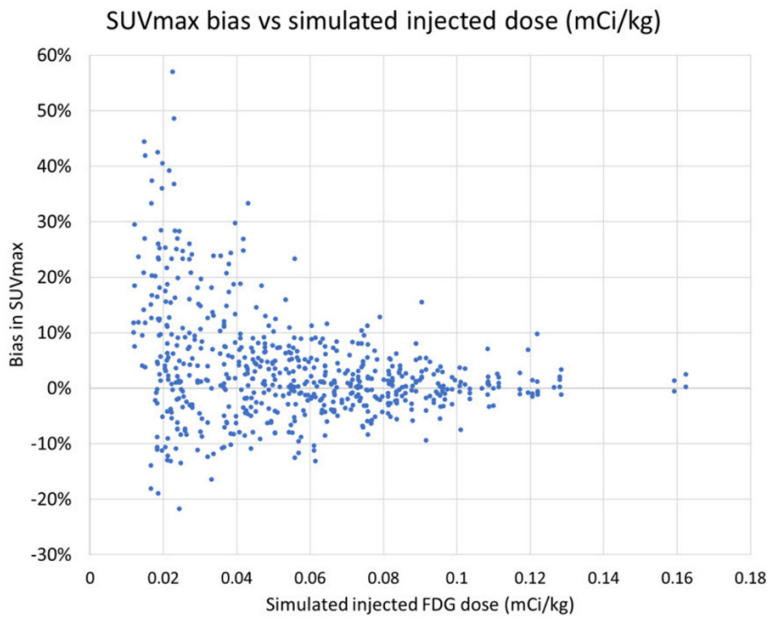

Figure 7. Scatterplot of SUVmax bias (comparing measured SUVmax on 20/40/60/80% decimated images to 100% images; N=154 index lesions) as a function of simulated injected FDG dose, showing a trend toward higher SUVmax bias as image noise increases (i.e., with reduction of simulated injected FDG dose). Bias was noted to be fairly low (± 10%) at simulated FDG dose levels down to around 0.04-0.05 mCi/kg, and appeared to increase exponentially as simulated injected dose levels decreased below 0.04 mCi/kg.

To explore potential causes of this observation, visual/qualitative index lesion conspicuity scores were compared with simulated injected FDG dose, as the latter was demonstrated to correlate well with overall image noise as described above and shown in **Figure 10**. The difference in simulated FDG dose (mCi/kg) between observations scored 1 or 2 (mean=0.039 mCi/kg) and observations scored 3 or 4 (medium or high confidence; mean=0.072 mCi/kg) was statistically significant (Mann-Whitney-Wilcoxon test, two-tailed, P<0.001). No index lesion visual/qualitative score or 1 (not detectable or questionably detectable) was observed above a simulated injected dose of 0.044 mCi/kg.

Discussion

Despite widespread use of FDG PET/CT for oncologic imaging, there remains considerable variation in selection of injected FDG dose and PET image acquisition time among imaging centers worldwide. Guidance as specified in the 2015 EANM procedure guideline (taking into account patient body weight, emission scan acquisition time and extent of overlap between PET bed positions) is useful as a starting point. However, the EANM guideline does acknowledge that several other parameters are not considered, including but not limited to enhanced sensitivity of modern PET detection equipment, increased volume of detectors resulting in extended axial field of view, and continuous bed motion [2]. All of these parameters, which are increasingly common features on modern PET/CT imaging systems, have potential to reduce required administered FDG activity and/or acquisition time. Having these options offers practical benefit, including lower overall radiation exposure to the patient (which may be of particular value in patients who undergo serial follow-up scans) and/or shorter scan time (which can be helpful from the perspective of scanner throughput and patient comfort, including reduction or elimination of patient motion during longer acquisitions which can lead to artifacts and degradation of overall image quality).

The mean actual injected FDG dose in this cohort (0.11 mCi/kg) was very similar to the recommended dose as calculated by use of the 2015 EANM guideline (0.095 mCi/kg), accounting for the limitations described above. However, it is important to note that the particular scanner used in this study had 4 rings of PET detectors (higher sensitivity secondary to increased axial field of view) and continuous bed motion technology (increasing effective bed overlap by comparison with “step and shoot” bed motion), neither of which are accounted for in the 2015 EANM guideline, such that the dose as recommended by the EANM equation likely represents an overestimation of actual dose required with use of this specific higher-sensitivity system. While 30% of the patients in this study population had an injected FDG dose lower than EANM recommended dose, 84% of these were within 25% of the recommended EANM dose, thus representing a relatively small difference in the context of

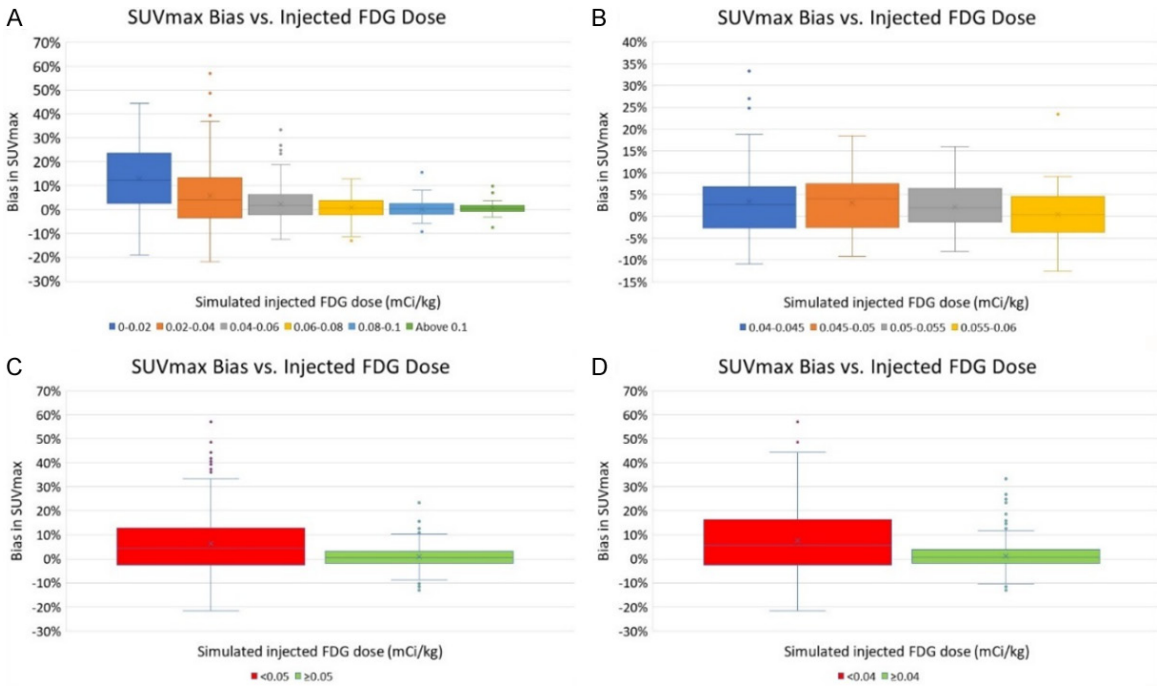

Figure 8. SUVmax bias subanalyses. A: Relationship between SUVmax bias and simulated injected FDG dose ranges are presented in box plot format (N=154 index lesions). These findings are essentially an alternative representation of data in the **Figure 7** scatterplot and confirm that SUVmax bias generally remains low (± 10%) down to a simulated injected dose level of around 0.04 mCi/kg. B: Subanalyses looking at narrower simulated injected FDG dose ranges between 0.04 and 0.06 mCi/kg, presented in box plot format. There was no statistically significant difference in SUVmax bias among these subgroups (Kruskal-Wallis test, P=0.483). C and D: Relationship between SUVmax bias and simulated injected FDG dose ranges using 0.05 and 0.04 mCi/kg as cutoffs, respectively. Differences in SUVmax bias between these groups was found to be statistically significant (Mann-Whitney-Wilcoxon test, two-tailed, P<0.05).

typical clinical practice and felt to be mitigated by the higher-sensitivity nature of the particular scanner used.

Our results indicate that subjective image quality was strongly associated with overall image noise, using normal liver SUV coefficient of variation as a surrogate marker of image noise. This association is well-known and expected. However, our results also suggest that, despite degradation of overall image quality secondary to noise, the accuracy of image-based oncologic staging remained stable at simulated FDG dose levels down to approximately 0.05 mCi/kg at constant PET acquisition time. In addition, SUVmax measurements also appeared stable (mean bias of less than 5%; 95% of scans within 10% bias and 99% of scans within 15% bias) for scans down to 0.05 mCi/kg simulated dose at constant PET acquisition time. This dose level, or an equivalent reduction in image acquisition time while maintaining standard FDG dose, would represent a reduction of nearly 50% from values recommended in the 2015 EANM guideline.

Assessment of index lesion conspicuity as a function of simulated dose reduction yielded interesting results. As would be expected, lesions with higher visual conspicuity as assessed by experienced PET readers had significantly higher mean contrast-to-noise (CNR) and signal-to-noise (SNR) ratios. However, there was substantial overlap between the groups. Of note, some lesions that were readily detected visually had low CNR/SNR. Furthermore, several of the more subtle index lesions in this study were noted to have higher subjective/visual conspicuity scores with increased simulated FDG dose, while CNR and SNR remained essentially constant. While of uncertain etiology, a few possibilities are speculated for this phenomenon. First, lesions with very low CNR/SNR may be sensitive to measurement variance, especially when the number of noise iterations is low. Secondly, overall image

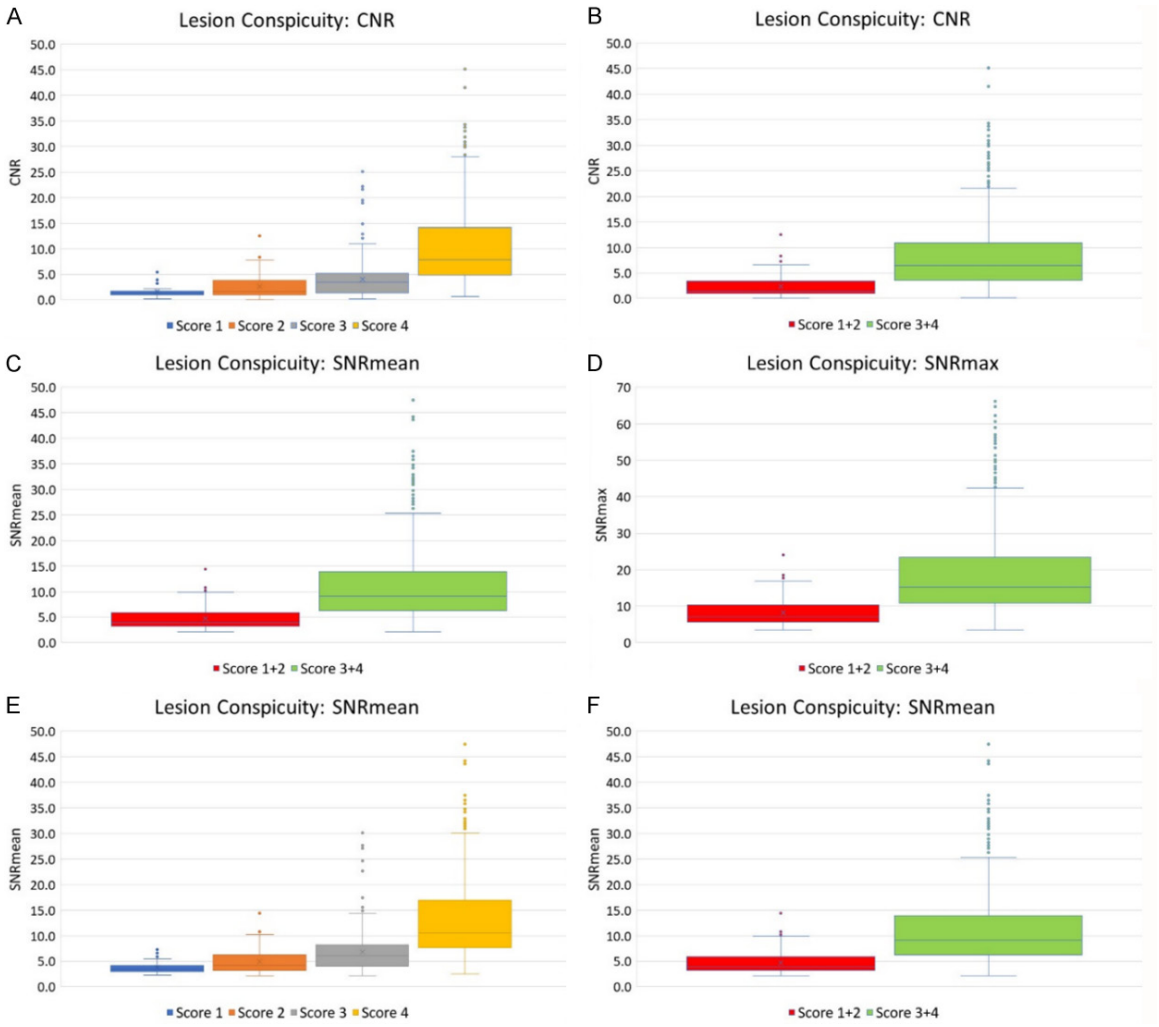

Figure 9. Subjective conspicuity data for index lesions (N=154; scoring scale: 1=not detectable or questionably detectable; 2=low contrast, 3=medium contrast, 4=high contrast), pooled for both observers, as related to contrast-to-noise ratio (CNR, using lesion SUVmean; A, B) and signal-to-noise ratio (SNR, using lesion SUVmax (C, D) and SUVmean (E, F)), are presented in box plot format (N=72). As expected, lesions with higher visual/qualitative observer scores (3 or 4) had overall higher CNR and SNR values. When rebinned into discrete groups representing "unacceptable" (score 1 or 2) vs. "acceptable" (score 3 or 4) subjective contrast, the differences in CNR and SNR between observations scored 1 or 2 vs. 3 or 4 were all statistically significant (Mann-Whitney-Wilcoxon test, two-tailed, $P<0.001$). However, there was substantial overlap in CNR and SNR between the groups, with some lesions that were scored as having high visual/qualitative contrast having rather low CNR and SNR, in the same range of other lesions with low visual scores (1 or 2).

noise may play an important role: as overall image noise outside the lesion increases, the appearance of artifactual "hot spots" throughout the image may draw the observer's eye away from a subtle lesion of interest, and the reduction of overall noise may allow for the lesion to be more readily detected. The lesion itself may also take on a more jagged shape/appearance which may mimic noise and further contribute to loss of visual conspicuity and/or diagnostic confidence. Finally, it is possible that the background standard deviation (noise) might be a combination of actual noise and image variations due to actual variability in the uptake as a result of OSEM reconstruction.

This study has several limitations. One potential limitation of this study is its retrospective nature in which reduced dose/count images were simulated using software techniques. As such, it could be argued that the observed results might not necessarily translate to imag-

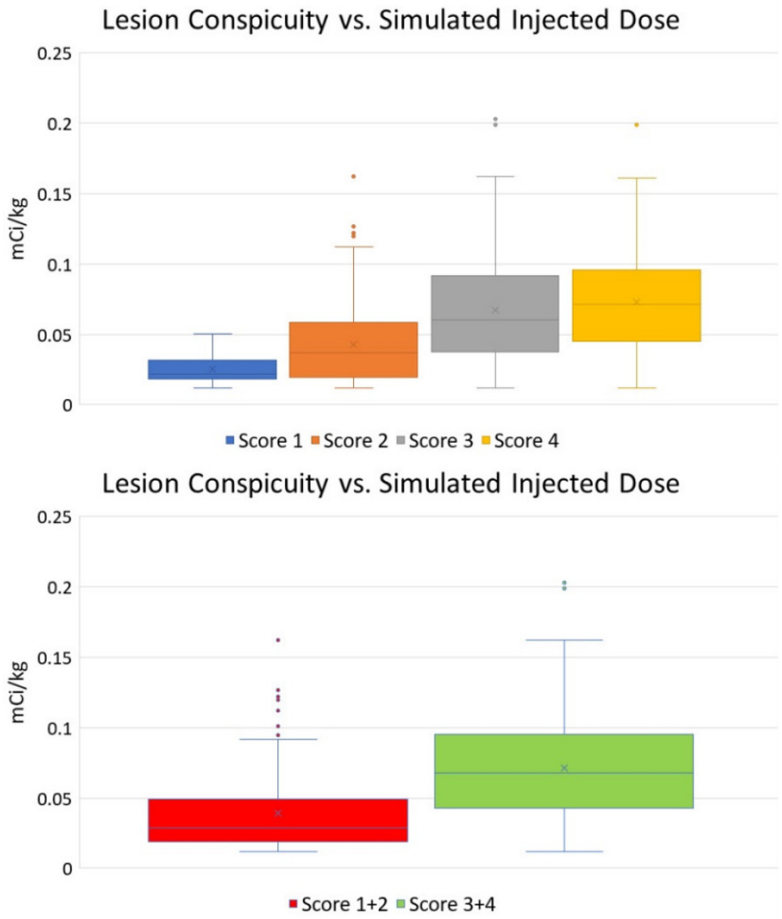

Figure 10. Subjective conspicuity data for index lesions (N=154; scoring scale: 1=not detectable or questionably detectable; 2=low contrast, 3=medium contrast, 4=high contrast), pooled for both observers, as related to simulated injected FDG dose, presented in box plot format. The difference in simulated FDG dose (mCi/kg) between observations scored 1 or 2 (mean=0.039 mCi/kg) and observations scored 3 or 4 (medium or high confidence; mean=0.072 mCi/kg) was statistically significant (Mann-Whitney-Wilcoxon test, two-tailed, P<0.001).

es acquired using actual low injected FDG activity and/or short image acquisition time. Of note, prompt and random coincidences do not scale in the same manner relative to in-field activity: as in-field activity decreases, the random fraction becomes smaller, not larger. Thus, our simulated reduced-count data contained more random fraction and therefore more noise than would be expected to occur when actually reducing either injected FDG dose or acquisition time. This effect should therefore bias our results toward more conservative conclusions [5]. However, verification with prospectively acquired images using actual low injected FDG activity and/or short imaging time is recommended.

Another potential limitation is that histopathology was not used as a standard of reference to confirm accuracy of image-based staging in this study. Nevertheless, this scenario (i.e., absence of histopathologic confirmation for every lesion) is common in routine clinical practice. In addition, the primary intent of this study was for intra-patient data comparison among different simulated FDG doses, rather than a full analysis of the performance of FDG PET as validated by histopathology. Also, only index lesions driving highest stage were assessed in this study; no attempt was made to evaluate all identifiable lesions, such that some lesions (which did not affect overall stage group) could possibly have been missed on lower-count images.

Finally, it is important to note that all images in this study were acquired using a single PET/CT scanner model, and thus the conclusions of this study would be limited to this system or systems with equivalent performance. Injected dose and/or acquisition time reduction on less sensitive scanners would result in more significant image quality degradation; conversely, more sensitive scanners could potentially result in a larger reduction in dose and/or acquisition time. Therefore, the results of this study are not cross-applicable to all scanner models or types. In addition, body habitus has a significant effect on image quality, such that injected dose or count reduction may not be feasible (due to unreasonable degradation in image quality) in obese patients.

Conclusion

The results of this retrospective data simulation study suggest that a modest reduction of

either injected FDG dose or PET emission scan time to levels below the range specified in the EANM/SNMMI procedure guideline might possibly be feasible using the specific equipment studied. While simulated reduction in injected FDG dose or acquisition time resulted in images with subjectively and objectively greater noise (particularly in obese patients), image-based staging accuracy and semi-quantitative (SUVmax) measurements appeared to be reproducible down to a simulated injected FDG dose level of approximately 0.05 mCi/kg at a continuous bed motion rate of 1.1 mm/sec. Verification with prospectively acquired images using actual low injected FDG activity and/or short imaging time is recommended. We acknowledge that ongoing evolutions in PET hardware (e.g., digital detector systems, further increases in detector volume/transaxial field of view, etc.) and reconstruction algorithms (e.g., advanced de-noising algorithms) may rapidly reduce the practical relevance of this study's findings.

Acknowledgements

The authors would like to thank Mark Westley and Victoria Peckham (Kaiser Permanente Division of Research), and Bijan Farboud and Vijay Shah, PhD (Siemens Medical Systems USA, Inc.) for administrative support and technical insight. This study was supported by Siemens Medical Solutions USA, Inc. as part of an academic collaboration.

Disclosure of conflict of interest

None.

Address correspondence to: Ryan D Niederkohr, Department of Nuclear Medicine, Kaiser Permanente, 700 Lawrence Expy, Santa Clara, CA, USA. Tel: 408-851-5615; E-mail: ryan.d.niederkohr@kp.org

References

[1] International Commission on Radiation Protection (ICRP). Radiation dose to patients from radiopharmaceuticals-addendum 3 to ICRP Publication 53, ICRP Publication 106. Approved by the Commission in October 2007. Ann ICRP 2008; 38: 1-197.

[2] Boellaard R, Delgado-Bolton R, Oyen WJ, Giammarile F, Tatsch K, Eschner W, Verzijlbergen FJ, Barrington SF, Pike LC, Weber WA, Stroobants S, Delbeke D, Donohoe KJ, Holbrook S, Graham MM, Testanera G, Hoekstra OS, Zijlstra J, Visser E, Hoekstra CJ, Pruim J, Willemsen A, Arends B, Kotzerke J, Bockisch A, Beyer T, Chiti A and Krause BJ; European Association of Nuclear Medicine (EANM). FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. Eur J Nucl Med Mol Imaging 2015; 42: 328-354.

[3] Alessio AM, Farrell MB and Fahey FH. Role of reference levels in nuclear medicine: a report of the SNMMI dose optimization task force. J Nucl Med 2015; 56: 1960-1964.

[4] Becker MD, Butler PF, Siam M, Gress DA, Ghesani M, Harkness BA, Yoo DC and Oates ME. U.S. PET/CT and gamma camera diagnostic reference levels and achievable administered activities for noncardiac nuclear medicine studies. Radiology 2019; 293: 203-211.

[5] Schaefferkoetter J, Nai YH, Reilhac A, Townsend DW, Eriksson L and Conti M. Low dose positron emission tomography emulation from decimated high statistics: a clinical validation study. Med Phys 2019; 46: 2638-2645.

[6] Cheson BD, Fisher RI, Barrington SF, Cavalli F, Schwartz LH, Zucca E and Lister TA; Alliance, Australasian Leukaemia and Lymphoma Group; Eastern Cooperative Oncology Group; European Mantle Cell Lymphoma Consortium; Italian Lymphoma Foundation; European Organisation for Research; Treatment of Cancer/Dutch Hemato-Oncology Group; Grupo Español de Médula Ósea; German High-Grade Lymphoma Study Group; German Hodgkin's Study Group; Japanese Lymphorra Study Group; Lymphoma Study Association; NCIC Clinical Trials Group; Nordic Lymphoma Study Group; Southwest Oncology Group; United Kingdom National Cancer Research Institute. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. J Clin Oncol 2014; 32: 3059-3068.

[7] American Joint Committee on Cancer: AJCC Cancer Staging Manual, Eighth Edition. In: Amin AB, editor. New York: Springer; 2017.

[8] Yan J, Schaefferkoetter J, Conti M and Townsend D. A method to assess image quality for low-dose PET: analysis of SNR, CNR, bias and image noise. Cancer Imaging 2016; 16: 26.

[9] Wahl RL, Jacene H, Kasamon Y and Lodge MA. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. J Nucl Med 2009; 50 Suppl 1: 122S-150S.

[10] Schaefferkoetter JD, Yan J, Sjohom T, Townsend DW, Conti M, Chung Tam JK, Soo RA and Than I. Quantitative accuracy and lesion detectability of low-dose 18F-FDG PET for lung cancer screening. J Nucl Med 2017; 58: 399-405.

[11] Schaefferkoetter JD, Yan J, Townsend DW and Conti M. Initial assessment of image quality for low-dose PET: evaluation of lesion detectability. Phys Med Biol 2015; 60: 5543-5556.