

# <sup>18</sup>F-FDG PET/CT 不同重建算法对肺结节 SUV 的影响

赵斌 郭彬威 黄斌 梁梦 秦志星 郝新忠 李思进 武志芳

山西医科大学第一医院核医学科、分子影像精准诊疗省部共建协同创新中心,太原 030001

通信作者:武志芳, Email: wuzhifang01@163.com

**【摘要】目的** 比较<sup>18</sup>F-脱氧葡萄糖(FDG) PET/CT 的 4 种重建算法对肺结节标准摄取值(SUV)的影响。**方法** 回顾性收集 2018 年 2 月至 2019 年 7 月在山西医科大学第一医院行<sup>18</sup>F-FDG PET/CT 检查的 46 例实性肺结节患者[男 27 例,女 19 例,中位年龄 66(44~82)岁]的 PET/CT 图像,采用有序子集最大期望值迭代法(OSEM)、OSEM+飞行时间(TOF)、OSEM+TOF+点扩散函数(PSF)及正则化算法(BSREM)进行图像重建(方法依次以 G1~G4 表示),通过视觉和半定量方法分析肺结节及背景参数。根据肺窗所测结节直径,分为小结节(直径≤10 mm)和大结节(10 mm<直径≤30 mm)。行 Kruskal-Wallis 秩和检验及 Bonferroni 法分析不同算法间 SUV 的差异,行 Spearman 相关分析探讨 SUV 变化率(%ΔSUV)与结节直径的相关性,行受试者工作特征(ROC)曲线分析探讨 SUV 对肺结节良恶性的诊断效能。**结果** 共 114 个结节,大结节 55 个,小结节 59 个。在视觉分析中,G4 较 G1~G3 的小结节视觉检出率分别提高了 55.93%(33/59)、44.07%(26/59)和 20.34%(12/59)。在 114 个肺结节中,最大 SUV(SUV<sub>max</sub>)、平均 SUV(SUV<sub>mean</sub>)在不同算法间比较差异有统计学意义(中位 SUV<sub>max</sub>:2.65~5.29,中位 SUV<sub>mean</sub>:2.05~2.99;H 值:20.628 和 17.749,均 P<0.001),G4 对 G1 的 SUV<sub>max</sub>(中位数分别为 5.29 和 2.65)和 SUV<sub>mean</sub>(中位数分别为 2.99 和 2.05)有明显提升(均 P<0.001)。%ΔSUV<sub>max</sub>(中位数:4.45%~52.96%)、%ΔSUV<sub>mean</sub>(中位数:1.69%~47.56%)与结节直径呈负相关[9.75(6.20,16.58) mm;r<sub>s</sub> 值:-0.371~-0.354,-0.371~-0.320,均 P<0.001]。在 59 个小结节中,G4 对 G1 的 SUV<sub>max</sub>(中位数分别为 4.05 和 2.14)有明显提升(H=18.327, P<0.001),G4 对 G1 和 G3 的 SUV<sub>mean</sub>(中位数分别为 2.31、1.26 和 1.53)有提升作用(H=16.808,均 P<0.05)。在 55 个大结节中,SUV 在不同算法间的差异无统计学意义(H 值:0.812~7.290,均 P>0.05)。G1~G4 的 SUV<sub>max</sub> 诊断良恶性的最佳截断值分别为 4.335、5.185、5.410、5.745,曲线下面积(AUC)分别为 0.747、0.699、0.756 和 0.778,四者的 SUV<sub>mean</sub> 及 SUV<sub>peak</sub> 最佳截断值对应的 AUC 也显示出类似趋势。**结论** 在 4 种重建算法中,BSREM 可明显提高图像质量和直径 10 mm 以下肺结节的 SUV<sub>max</sub> 及 SUV<sub>mean</sub>,其 SUV 良恶性诊断阈值应适当上调。

**【关键词】** 肺肿瘤;正电子发射断层显像术;体层摄影术,X 线计算机;图像处理,计算机辅助;脱氧葡萄糖

**基金项目:**国家自然科学基金(81571716, 81971655);山西省回国留学人员重点科研资助项目(2017-重点 5);山西省重点研发计划项目(201703D421020)

DOI:10.3760/cma.j.cn321828-20191122-00266

## Effects of different reconstruction algorithms on SUV of pulmonary nodules in <sup>18</sup>F-FDG PET/CT

Zhao Bin, Guo Binwei, Huang Bin, Liang Meng, Qin Zhixing, Hao Xinzong, Li Sijin, Wu Zhifang

Department of Nuclear Medicine, the First Hospital of Shanxi Medical University; Collaborative Innovation Center of Molecular Imaging Precision Diagnosis and Treatment, Taiyuan 030001, China

Corresponding author: Wu Zhifang, Email: wuzhifang01@163.com

**【Abstract】** **Objective** To compare four reconstruction algorithms of <sup>18</sup>F-fluorodeoxyglucose (FDG) PET/CT on standardized uptake value (SUV) of pulmonary nodules. **Methods** A total of 46 patients (27 males, 19 females; median age: 66 (range: 44~82) years) with solid pulmonary nodules from February 2018 to July 2019 in the First Hospital of Shanxi Medical University who performed <sup>18</sup>F-FDG PET/CT imaging were enrolled. All PET/CT images were retrospectively reconstructed by using four algorithms reconstructions including ordered subset expectation maximization (OSEM), OSEM+ time of flight (TOF), OSEM+ TOF+ point spread function (PSF) and block sequential regularized expectation maximization (BSREM) (G1~G4). Nodule and background parameters were analyzed semi-quantitatively and visually. The maxi-

mum of SUV ( $SUV_{\max}$ ) , mean of SUV ( $SUV_{\text{mean}}$ ) and peak of SUV ( $SUV_{\text{peak}}$ ) were collected by the region of interest (ROI). Nodules were divided into small nodule group (diameter  $\leq 10$  mm) and large nodule group ( $10 \text{ mm} < \text{diameter} \leq 30$  mm). Kruskal-Wallis rank sum test and Bonferroni method were performed to compare the differences of SUVs between G1-G4, and Spearman correlation analysis was used to analyze the correlation between the change rate of SUV (% $\Delta$ SUV) and the diameter of nodules. The receiver operating characteristic (ROC) curve analysis was used to analyze the diagnostic efficacy of SUV for the differential diagnosis of pulmonary nodules and to get the optimal threshold. **Results** There were 114 pulmonary nodules (large nodules,  $n=55$ ; small nodules,  $n=59$ ). In visual analysis, the visual detection rates of small nodules in G4 were 55.93% (33/59), 44.07% (26/59), 20.34% (12/59) higher than those in G1-G3. Of 114 pulmonary nodules in 46 patients, there were differences in  $SUV_{\max}$  and  $SUV_{\text{mean}}$  between G1-G4 (median  $SUV_{\max}$ : 2.65-5.29, median  $SUV_{\text{mean}}$ : 2.05-2.99;  $H$  values: 20.628 and 17.749, respectively, both  $P < 0.001$ ), G4 had significant increases compared to G1 in  $SUV_{\max}$  (median 5.29 and 2.65,  $P < 0.001$ ) and  $SUV_{\text{mean}}$  (median 2.99 and 2.05,  $P < 0.001$ ). The % $\Delta$ SUV<sub>max</sub> (median: 4.45%-52.96%) and % $\Delta$ SUV<sub>mean</sub> (median: 1.69%-47.56%) were negatively correlated with the diameter of nodules (9.75(6.20, 16.58) mm;  $r_s$  values: -0.371 to -0.354, -0.371 to -0.320, all  $P < 0.001$ ). In 59 small nodules, G1 significantly increased the  $SUV_{\max}$  of G4 (median 4.05 and 2.14,  $H = 18.327$ ,  $P < 0.001$ ), while G4 significantly increased the  $SUV_{\text{mean}}$  of G1 and G3 (median 2.31, 1.26 and 1.53,  $H = 16.808$ ,  $P < 0.05$ ). There was no significant difference in SUVs between G1-G4 in 55 large nodules ( $H$  values: 0.812-7.290, all  $P > 0.05$ ). The optimal threshold values of  $SUV_{\max}$  in G1-G4 were 4.335, 5.185, 5.410, 5.745 and the area of under curves (AUCs) were 0.747, 0.699, 0.756, 0.778 respectively. The AUC of  $SUV_{\text{mean}}$  and  $SUV_{\text{peak}}$  also showed a similar trend. **Conclusion** Among the four reconstruction algorithms, BRERM can not only enhance the image quality, but also significantly improve the  $SUV_{\max}$  and  $SUV_{\text{mean}}$  of lung nodules diameter below 10 mm, and thus its diagnostic threshold of SUV should be appropriately increased.

**[Key words]** Lung neoplasms; Positron-emission tomography; Tomography, X-ray computed; Image processing, computer-assisted; Deoxyglucose

**Fund program:** National Natural Science Foundation of China (81571716, 81971655); Shanxi Key Scientific Research Projects for Returned Overseas Staff (2017-Key Point 5); Shanxi Key Research and Development Program (201703D421020)

DOI:10.3760/cma.j.cn321828-20191122-00266

<sup>18</sup>F-脱氧葡萄糖(fluorodeoxyglucose, FDG) PET/CT 显像已常规应用于肿瘤的诊断、分期及疗效评估, 在评估肺结节方面也有一定价值。2018 年第 5 版美国国家综合癌症网络(National Comprehensive Cancer Network, NCCN)指南推荐直径 8~14 mm 实性肺结节术前评价行 PET/CT<sup>[1]</sup>, 我国专家同样建议直径>8 mm 无法定性的实性肺结节行 PET/CT<sup>[2]</sup>。标准摄取值 (standardized uptake value, SUV) 是 PET 显像中最常用的评估代谢活性的半定量指标, 包括最大 SUV (maximum SUV,  $SUV_{\max}$ )、平均 SUV (mean SUV,  $SUV_{\text{mean}}$ ) 和标准摄取峰值 (peak of SUV,  $SUV_{\text{peak}}$ )。SUV 受诸多因素的影响:生物因素(患者体质量、病灶特征、呼吸运动和受检者的血糖水平等)可影响 FDG 在体内的分布;技术和采集因素(注射后显像时长、显像设备、信号采集等)会影响所测图像与 FDG 在体内真实分布情况的差距<sup>[3]</sup>;此外, 其量化结果还取决于 PET 不同的图像重建算法<sup>[4]</sup>。

目前 PET 主要有 2 种重建算法:分析法和迭代法<sup>[5]</sup>, 后者较前者明显改善了信号噪声比 (signal to noise ratio, SNR)<sup>[6-7]</sup>。随着迭代技术的不断改进, 目前临床常用算法有有序子集最大期望值迭代法 (ordered subset expectation maximization, OSEM) 、

OSEM+飞行时间 (time of flight, TOF)、OSEM+TOF+点扩散函数 (point spread function, PSF) 以及正则化算法 (block sequential regularized expectation maximization, BSREM)。Teoh 等<sup>[8]</sup>在<sup>18</sup>F-FDG PET/CT 肺小结节显像研究中发现, BSREM 较 OSEM+TOF 可增加病灶  $SUV_{\max}$ , 提升图像的 SNR 以及信号背景比 (signal to background ratio, SBR), 另外还发现 BSREM 的视觉分析比其半定量分析方法对结节的评估和良恶性诊断效能更佳; Howard 等<sup>[9]</sup>研究指出在 BSREM 算法中随着决定惩罚函数强度<sup>[10]</sup>, 即  $\beta$  值的减小, 平均直径为 8 mm 的肺小结节的平均  $SUV_{\max}$  及视觉评分递增, 但无论  $\beta$  值为多少, BSREM 视觉评估效果均高于 OSEM。本研究旨在探讨<sup>18</sup>F-FDG PET/CT 的 4 种重建算法 OSEM、OSEM+TOF、OSEM+TOF+PSF 及 BSREM 对肺结节 SUV 的影响。

## 资料与方法

1. 研究对象。回顾性收集 2018 年 2 月至 2019 年 7 月于本科行<sup>18</sup>F-FDG PET/CT 检查的 46 例患者 [男 27 例, 女 19 例, 中位年龄 66 (44~82) 岁] 的 PET/CT 图像, 采用 OSEM、OSEM+TOF、OSEM+TOF+PSF 及 BSREM 方法进行图像重建。入选标准: 直径≤

30 mm 的实性肺结节<sup>[2]</sup>。排除标准:(1)直径≤30 mm 的纯磨玻璃密度或混合性磨玻璃肺结节;(2)采集条件不一致者。本研究符合《赫尔辛基宣言》的原则。

2. PET/CT 检查及重建方法。采用美国 GE 公司 Discovery MI PET/CT 仪,<sup>18</sup>F-FDG 由日本住友公司 HM-10 型回旋加速器及 FDG 合成模块生产(放化纯>99%)。患者检查前禁食 4~6 h, 血糖水平处于正常范围, 于肘静脉按体质量注射<sup>18</sup>F-FDG 2.96~4.44 MBq/kg, 注射后安静休息 50~70 min 行全身显像。扫描范围自颅顶至股骨中部, 采集方式为列表模式, PET 显像在正常潮气呼吸下采集, 3 min/床位, 采集 5~7 个床位。CT 扫描参数: 管电压 120 kV, 管电流 60~150 mA(xy 方向/z 方向自动调制), 噪声指数为 18, 螺距 0.984, 层厚 3.75 mm, 旋转时间 0.5 s, 探测器覆盖范围为 40 mm。采用 4 种不同的迭代算法进行重建, 即 OSEM(2 次迭代, 17 个子集, 6.4 mm 高斯滤波器)、OSEM+TOF(飞行时间+3 次迭代, 17 个子集, 6.4 mm 高斯滤波器)、OSEM+TOF+PSF(点扩散函数+飞行时间+3 次迭代, 17 个子集, 6.4 mm 高斯滤波器)及 BSREM(惩罚因子  $\beta=350$ )。利用 CT 投射扫描数据对 PET 图像进行衰减校正, 把校正后的 PET 图像与 CT 图像进行融合, 分别得到横断位、矢状位及冠状位的 CT、PET 及 PET/CT 融合图像。

3. PET 图像视觉评估。由 2 位经验丰富的 PET/CT 医师进行 PET 图像盲法判读, 对使用不同重建方法的图像结合其清晰度、平滑度及结节检出性能, 进行 0~4 分视觉评估(0 为差, 1 为一般, 2 为良好, 3 为很好, 4 为优秀)。

4. PET/CT 图像半定量分析。(1)背景参数的获取。分别将 OSEM、OSEM+TOF、OSEM+TOF+PSF 及 BSREM(以下分别简称为 G1、G2、G3 及 G4)重建的 PET 图像与原始采集的 CT 融合。在肝右叶测量直径为 3 cm 球形感兴趣区(region of interest, ROI)的 SUV 标准差(standard deviation of SUV, SUV<sub>sd</sub>), 在降主动脉测量直径为 1 cm 球形 ROI 的 SUV<sub>mean</sub>。每个结节的 SBR 定义为结节 SUV<sub>max</sub>/降主动脉 SUV<sub>mean</sub>; SNR 定义为结节 SUV<sub>max</sub>/肝脏 SUV<sub>sd</sub>。

(2)结节代谢参数的获取。将 CT 和经 4 种算法重建的 PET 图像共同加载到 PETVCAR 软件(美国 GE Health)中, 在横断面、矢状面和冠状面观察病灶的 PET 灰度和 PET/CT 融合图像。将采样框放置在图像上, 软件自动辨识肺结节轮廓, 在三维方向上检查和调整取样框大小。重复该过程直至每个 PET 阳性区域被选择和优化。使用 PETVCAR 中的

自适应迭代算法, 得到 ROI 的 SUV<sub>max</sub>、SUV<sub>mean</sub> 和 SUV<sub>peak</sub>。同时在肺窗横断面测量结节直径, 以 10 mm 为界分为 2 组: ≤10 mm 为小结节,>10 mm 且≤30 mm 为大结节。不同重建算法下所测得的 SUV 记为 SUV<sub>算法</sub>, 将不同算法间的 SUV 差值记为 ΔSUV, 其下降百分比称为变化率, 记为%ΔSUV, 如 G1、G2 间 SUV<sub>max</sub> 变化率%ΔSUV<sub>max</sub> = (SUV<sub>G2max</sub> - SUV<sub>G1max</sub>) / SUV<sub>G1max</sub> × 100%。

5. 随访。所有患者均行病理随访, 随访截止时间为 2019 年 10 月 7 日, 中位随访时间为 12.5(3~20)个月。

6. 统计学处理。采用 IBM SPSS 21.0 软件进行数据分析。符合正态分布的定量资料用  $\bar{x} \pm s$  表示, 不符合正态分布的定量资料以  $M(P_{25}, P_{75})$  表示。采用 Kappa 检验分析 2 名医师视觉评分间的一致性; 采用  $\chi^2$  检验分析不同重建算法间视觉评分的差异; 采用 Kruskal-Wallis 秩和检验分析不同重建算法间 SBR、SNR 和 SUV 的差异; 两两比较行 Bonferroni 法; 采用 Spearman 相关分析整个队列%ΔSUV 与结节直径的相关性。SUV<sub>max</sub>、SUV<sub>mean</sub> 及 SUV<sub>peak</sub> 对肺结节良恶性的诊断效能采用受试者工作特征(receiver operating characteristic, ROC)曲线分析, 计算曲线下面积(area under curve, AUC), 寻找最佳截断值, 比较不同截断值的诊断效能。P<0.05 为差异或相关性有统计学意义。

## 结 果

1. 肺结节 FDG 摄取的视觉评估。2 名医师在 G1~G4 重建算法间视觉评分具有高度一致性( $\kappa$  值: 0.986、0.926、0.940、0.944, 均 P<0.001)。对于 114 个肺结节(大结节 55 个, 小结节 59 个), G1 视觉评分 0~4 分者分别有 60、28、12、12 和 2 个, G2 分别有 48、30、18、16 和 2 个, G3 分别有 23、32、41、16 和 2 个, G4 分别有 11、29、25、27 和 22 个, 不同方案的视觉评分构成差异有统计学意义( $\chi^2=111.72$ , P<0.001)。小结节中 G1~G4 视觉评分为 0 者分别有 41、34、20 和 8 个, 大结节中对应有 19、14、3 和 3 个; G4 较 G1~G3 的小结节视觉检出率分别提高了 55.93%(33/59)、44.07%(26/59) 和 20.34%(12/59)。有 12 个结节(均为小结节)在 G4 中目测为阳性(评分不为 0 分), 但在其他 3 组中目测为阴性(评分为 0 分)。

2. 背景参数分析。在整个队列中, 结节在 G1~G4 的中位 SBR 分别为 2.09(1.45, 4.03)、2.90(1.75, 4.90)、3.28(1.77, 5.64)、4.05(2.60, 6.51), 中位 SNR

分别为 5.83(3.93, 13.15)、6.54(3.78, 13.45)、19.90(12.86, 36.89)、24.90(18.32, 42.22), 四者 SBR 和 SNR 的差异有统计学意义( $H$  值: 24.855 和 168.934, 均  $P < 0.001$ )。两两比较发现, SBR 在 G1 和 G4 间、G2 和 G4 间的差异有统计学意义(调整后均  $P < 0.001$ ), 余组间差异均无统计学意义(均  $P > 0.05$ ); SNR 在 G1 和 G3 间、G1 和 G4 间、G2 和 G3 间、G2 和 G4 间差异有统计学意义(调整后  $P < 0.001$ ), 余组间差异均无统计学意义(均  $P > 0.05$ )。

3. 所有肺结节同一 SUV 指标在不同重建算法间的比较。114 个肺结节中位直径为 9.75(6.20, 16.58) mm。在 G1~G4 算法重建后的图像中, 病灶中位  $SUV_{max}$  分别为 2.65(1.98, 6.78)、4.20(2.33, 8.36)、4.51(2.29, 8.89) 和 5.29(3.32, 9.54), 中位  $SUV_{mean}$  分别为 2.05(1.13, 3.91)、2.39(1.35, 5.16)、2.44(1.27, 4.85) 和 2.99(1.88, 5.37), 中位  $SUV_{peak}$  分别为 1.88(1.03, 4.21)、2.06(1.24, 4.44)、1.94(1.18, 4.16) 和 2.08(1.30, 4.69)。同一患者同一结节在 G1~G4 算法重建下的<sup>18</sup>F-FDG PET/CT 图像见图 1。

4 组间  $SUV_{max}$  的差异有统计学意义( $H=20.628$ ,  $P < 0.001$ )。两两比较发现,  $SUV_{max}$  在 G1 和 G4 间的差异有统计学意义(调整后  $P < 0.001$ ), 余组间差异均无统计学意义(均  $P > 0.05$ )。组间  $\Delta\%SUV_{max}$  最大、最小的 2 组分别为 G4 和 G1[52.96%(31.86%, 83.09%)]、G2 和 G1[16.42%(8.18%, 31.61%)];  $\%ΔSUV_{max}$ (中位数范围为: 4.45%~52.96%)和结节直径间存在负相关( $r_s$  值: -0.371~-0.354, 均  $P < 0.05$ )。

4 组间  $SUV_{mean}$  差异亦有统计学意义( $H=17.749$ ,  $P < 0.001$ )。两两比较发现,  $SUV_{mean}$  在 G1 和 G4 间的差异有统计学意义(调整后  $P < 0.001$ ), 余组间差异均无统计学意义(均  $P > 0.05$ )。组间  $\Delta\%SUV_{mean}$  最大、最小的 2 组分别为 G4 和 G1[47.56%(27.29%, 79.71%)]、G2 和 G1[17.13%(8.79%, 29.83%)];  $\%ΔSUV_{mean}$ (中位数范围为: 1.69%~47.56%)和结节直径间存在负相关( $r_s$  值: -0.371~-0.320, 均  $P < 0.05$ )。

4 组间  $SUV_{peak}$  差异无统计学意义( $H=1.677$ ,  $P=0.642$ )。组间  $\Delta\%SUV_{peak}$  和结节直径不存在相关性( $r_s$  值: -0.198~-0.152, 均  $P > 0.05$ )。

4. 不同大小肺结节 SUV 在不同重建算法间的比较(表 1)。114 个肺结节按直径进行分组, 小结节(直径≤10 mm)共 59 个, 其中<8 mm 43 个, 中位直径 6.2(4.8, 8.4) mm; 大结节(10 mm<直径≤30 mm)共 55 个, 中位直径 16.70(13.10, 21.55) mm。小结节组 G1~G4 的中位  $SUV_{max}$  和  $SUV_{mean}$  分别为 2.14(1.46, 4.34)、2.65(1.92, 5.44)、2.59(1.92, 5.54)、4.05(2.86, 7.44) 和 1.26(0.79, 2.43)、1.58(1.08, 3.08)、1.53(1.03, 2.98)、2.31(1.62, 4.30), 四者  $SUV_{max}$  和  $SUV_{mean}$  间差异有统计意义( $H$  值: 18.327 和 16.808, 均  $P < 0.05$ ),  $SUV_{max}$  在 G1 和 G4 间、 $SUV_{mean}$  在 G1 和 G4 间、G3 和 G4 间的差异有统计学意义(调整后均  $P < 0.05$ ), 而  $SUV_{peak}$  在 4 种重建算法间差异无统计学意义( $H=1.564$ ,  $P>0.05$ )。大结节组  $SUV_{max}$ 、 $SUV_{mean}$ 、 $SUV_{peak}$  在 4 种算法间比较差异均无

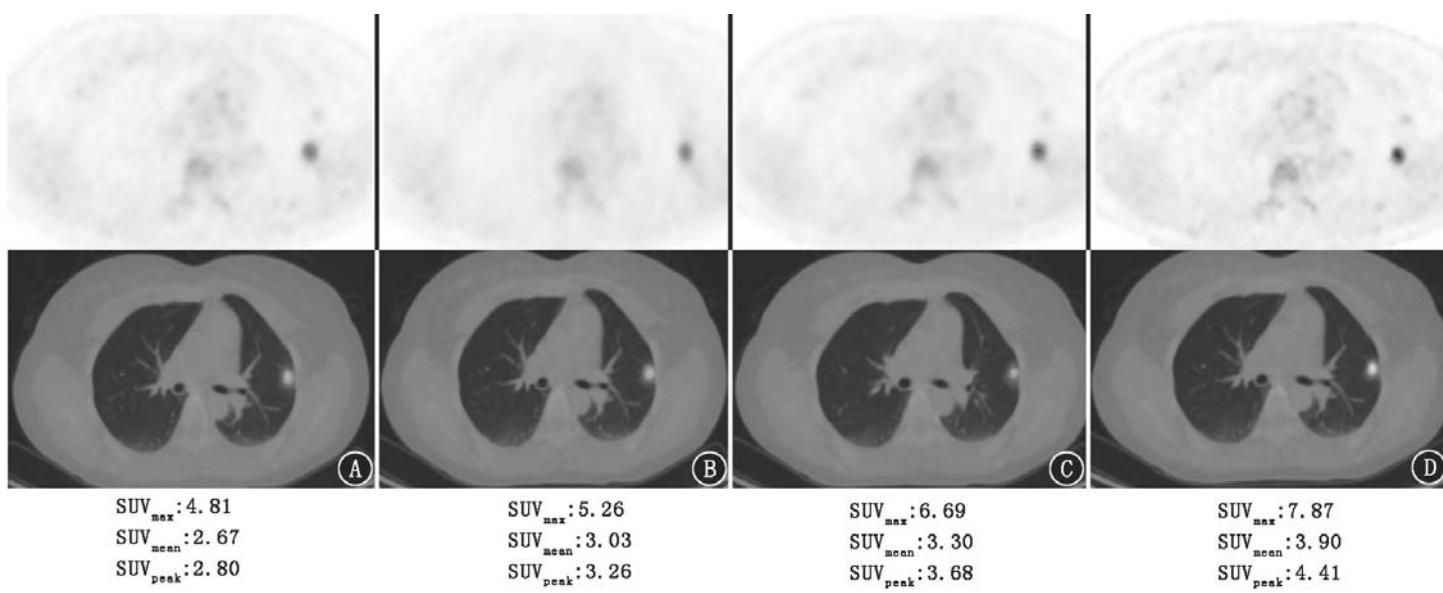


图 1 淋巴瘤化疗后复查患者(女, 28岁)的同一肺结节在 4 种重建算法下的<sup>18</sup>F-脱氧葡萄糖(FDG)PET/CT 显像图。A~D 分别表示有序子集最大期望值迭代法(OSEM)算法、OSEM+飞行时间(TOF)算法、OSEM+TOF+点扩散函数(PSF)算法、正则化算法(BSREM)重建后的 PET 及 PET/CT 融合图, 显示左肺上叶尖后段可见一直径约 13 mm 的结节, 边界模糊, 代谢增高, 其中在 BSREM 中左肺上叶尖后段结节视觉效果更佳, 且标准摄取值(SUV)逐渐递增; SUV<sub>max</sub> 为最大 SUV, SUV<sub>mean</sub> 为平均 SUV, SUV<sub>peak</sub> 为标准摄取峰值

统计学意义( $H$  值:0.812~7.290, 均  $P>0.05$ )。

5. 诊断肺结节良恶性的 ROC 曲线分析(表 2)。根据随访及病理结果,共 63 个肺结节纳入 ROC 曲线分析,其中恶性结节 41 个,良性结节 22 个。41 个恶性结节中位直径 9.6(3.0~27.1) mm,其中非小细胞肺癌 24 个(经病理证实 14 个:腺癌 13 个,鳞状细胞癌 1 个;转移瘤 17 个,均通过随访或影像资料明确)。22 个良性结节中位直径 10.15(4.5~25.6) mm,其中肺结核 12 个,感染性病变 6 个,结节病 3 个,错构瘤 1 个,均通过随访或影像资料明确。63 个结节进行了良恶性诊断效能的 ROC 曲线分析,其中 G4 的  $SUV_{max}$ 、 $SUV_{mean}$  及 G3 的  $SUV_{peak}$  诊断肺小结节良恶性的 AUC 在其组间比较中各自最大,但 SUV 在 4 种算法间的 AUC 差异均无统计学意义( $z$  值:0.157~1.738, 均  $P>0.05$ )。

## 讨 论

肺结节患者早期无明显临床症状和体征,仅在临床体格检查时偶然发现。良恶性结节形态学特征表现有一定重叠,因此其诊断较困难。结节代谢改变的发生早于形态学变化,故 PET 提供的分子信息非常有利<sup>[11]</sup>,其中最常用的半定量指标是 SUV。随着 PET 图像重建算法日益涌现,分析法逐渐被迭代法所取代,其中 OSEM 临床应用广泛,其通过逐步迭代从原始数据中找到最有可能接近真值的图像,每次迭代后会得到 1 幅更加接近真实数据的图像,但随着迭代次数的增加,噪声在图像中变得更加突

出<sup>[12]</sup>。因此临幊上使用相对较低的迭代次数来获得可接受的图像,但这将导致图像欠收敛,从而低估 SUV,尤其对 FDG 显像的小病灶有较大影响<sup>[8]</sup>。随着迭代技术的不断革新,在 OSEM 基础上 TOF 及 PSF 相继与其联合应用来提高 SUV 的准确性。TOF 图像重建技术是利用具有快速闪烁晶体和电子线路的设备测出光子到达探测器的时间差,从而提高定位及 PET 统计数据质量<sup>[13]</sup>;PSF 则通过点扩散功能响应模型对 PET 仪中的不同空间位置上空间分辨率损失进行补偿<sup>[3,14]</sup>。上述方法都是在图像噪声和完全收敛的定量精度之间进行权衡。

近期,BSREM 作为一种新型的优化 PET 迭代重建算法被提出。与其他惩罚函数相比,BSREM 是相邻体素间差异的函数,充当噪声抑制项,从而抑制噪声,允许迭代次数增加,其强度由  $\beta$  值控制<sup>[8]</sup>。BSREM 能更好地对小型结构进行显像,并可提供更精确的 SUV<sup>[9,15]</sup>,但不同重建算法对 SUV 的影响尚未明确。Teoh 等<sup>[8]</sup>应用 BSREM 和 OSEM+TOF 处理 121 例肺结节的图像,指出 BSREM 能够增加 OSEM+TOF 的  $SUV_{max}$ ;Brendle 等<sup>[4]</sup>分析了 27 例恶性肿瘤患者的<sup>18</sup>F-FDG PET/CT 图像发现,与 OSEM+TOF 相比,PSF 的加入会使得  $SUV_{max}$ 、 $SUV_{mean}$  及  $SUV_{peak}$  明显增加;Akamatsu 等<sup>[16]</sup>使用 OSEM, OSEM+PSF、OSEM+TOF、OSEM+PSF+TOF 回顾性研究了 41 例转移淋巴结的  $SUV_{max}$  和  $SUV_{mean}$ ,PSF 和 TOF 的各自加入均增加了 SUV,但 PSF 和 TOF 的结合使 SUV 增幅最大;吴天棋等<sup>[17]</sup>进行了 OSEM 及 OSEM+TOF

表 1 不同重建算法下不同大小肺结节  $SUV_{max}$ 、 $SUV_{mean}$  及  $SUV_{peak}$  的比较 [ $M(P_{25}, P_{75})$ ]

算法	$SUV_{max}$		$SUV_{mean}$		$SUV_{peak}$	
	小结节	大结节	小结节	大结节	小结节	大结节
G1	2.14(1.46, 4.18) <sup>a</sup>	5.94(2.34, 9.58)	1.26(0.79, 2.43) <sup>a</sup>	3.57(1.37, 5.58)	1.12(0.85, 2.53)	3.88(1.58, 7.36)
G2	2.65(1.92, 5.40)	5.96(2.59, 11.25)	1.58(1.08, 3.08)	3.62(1.43, 6.74)	1.29(0.99, 2.81)	4.23(1.71, 7.52)
G3	2.59(1.92, 5.54)	6.51(3.21, 11.49)	1.53(1.03, 2.99) <sup>a</sup>	3.87(1.91, 6.45)	1.30(0.91, 2.67)	4.07(1.70, 7.10)
G4	4.05(2.89, 7.45)	7.03(3.43, 12.28)	2.31(1.62, 4.29)	4.45(2.33, 7.11)	1.35(0.87, 2.97)	4.60(1.72, 7.95)

注:大结节(10 mm< 直径≤30 mm)55 个,小结节(直径≤10 mm)59 个;G1 为有序子集最大期望值迭代法(OSEM)算法,G2 为 OSEM+飞行时间(TOF)算法,G3 为 OSEM+TOF+点扩散函数(PSF)算法,G4 为正则化算法(BSREM); $SUV_{max}$  为最大标准摄取值, $SUV_{mean}$  为平均标准摄取值, $SUV_{peak}$  为标准摄取峰值;<sup>a</sup> 表示与 G4 比较差异有统计学意义

表 2 不同重建算法对应的  $SUV_{max}$ 、 $SUV_{mean}$ 、 $SUV_{peak}$  对 63 个肺结节的 ROC 曲线分析结果

算法	$SUV_{max}$				$SUV_{mean}$				$SUV_{peak}$			
	灵敏度	特异性	截断值	AUC(95% CI)	灵敏度	特异性	截断值	AUC(95% CI)	灵敏度	特异性	截断值	AUC(95% CI)
G1	50.0%	97.4%	4.335	0.747(0.653~0.840)	53.2%	92.1%	2.335	0.752(0.660~0.845)	66.1%	78.9%	1.615	0.765(0.674~0.856)
G2	50.0%	94.7%	5.185	0.699(0.598~0.800)	45.2%	97.4%	3.365	0.712(0.614~0.811)	51.6%	94.7%	2.795	0.758(0.666~0.850)
G3	56.5%	94.7%	5.410	0.756(0.663~0.848)	54.8%	97.4%	2.890	0.766(0.675~0.857)	50.0%	94.7%	2.670	0.776(0.686~0.866)
G4	59.7%	94.7%	5.745	0.778(0.689~0.867)	59.7%	94.7%	3.200	0.782(0.694~0.870)	50.0%	94.7%	2.915	0.760(0.667~0.852)

注:AUC 为曲线下面积,ROC 为受试者工作特征

的模型实验,发现 TOF 技术能有效提升图像对比度,但未行相关临床试验。

与另外 3 种重建算法相比,BSREM 所得图像视觉质量更好,整个队列的结节 SBR 和 SNR 随着重建算法的改进而递增,与之前的报道一致<sup>[8,10]</sup>。在 2 名医师视觉评分高度一致性的情况下,发现有 PSF 模型的重建算法比无 PSF 模型的重建算法所得图像质量评分更高,尤其在小结节中,与 Teoh 等<sup>[8]</sup>的结果一致,因为在直径≤10 mm 的病灶中,通过 PSF 模型对所得 SUV<sub>max</sub> 更接近于模型真值<sup>[12]</sup>。BSREM 的 SNR 几乎为 OSEM 的 4 倍,这可能是因为相较于 PSF 的有限收敛,BSREM 可达到完全收敛<sup>[18]</sup>。

Teoh 等<sup>[8]</sup>比较了 OSEM+TOF 和 BSREM 处理的 121 例肺结节<sup>18</sup>F-FDG PET/CT 图像,发现无论在所有肺结节中,还是在小结节组或大结节组中,BSREM 所得 SUV<sub>max</sub> 均明显高于 OSEM+TOF,结节大小与%ΔSUV<sub>max</sub> 呈负相关,当结节直径≤10 mm 时,平均%ΔSUV<sub>max</sub> 为 92.1%,当结节直径>10 mm,平均%ΔSUV<sub>max</sub> 为 45.3%;Howard 等<sup>[9]</sup>发现当 β = 350 时,BSREM 所测肺小结节的 SUV<sub>max</sub> 与 OSEM 无明显差异,随着 β 值的降低,2 种重建方法所测的 SUV<sub>max</sub> 差异越大。在本研究中,所有结节随着重建算法的改进,所得 SUV<sub>max</sub>、SUV<sub>mean</sub> 均明显高于经 OSEM 重建后的数据;与另外 3 种重建算法相比,BSREM 对 SUV<sub>max</sub>、SUV<sub>mean</sub> 有明显提升,且结节大小与%ΔSUV 呈负相关。但与 Teoh 等<sup>[8]</sup>的结论不同的是,本研究仅在小结节中发现 SUV<sub>max</sub> 和 SUV<sub>mean</sub> 在不同重建方法中差异具有统计学意义,这是由于部分容积效应所致<sup>[19]</sup>,当病灶越小,SUV 被低估的程度越高<sup>[3]</sup>;另一部分原因可能是 Teoh 等<sup>[8]</sup>的研究中恶性大结节占比(至少 82%)高于本研究的 65.79%,且病理类型分布不同。Messerli 等<sup>[20]</sup>认为在肺腺癌中,BSREM、OSEM 所得 SUV<sub>max</sub> 变化率比肺鳞状细胞癌更高,在本研究中因样本量受限,未进行 SUV 与病理类型的相关分析。另外,与 Howard 等<sup>[9]</sup>的结论不同是因为纳入标准的不同,其选择了平均直径为 8 mm 的处于肺癌高风险的 32 例肺小结节,还融入了视觉评估标准进行综合评价。本研究中发现 OSEM+TOF 与 OSEM+TOF+PSF 重建的 PET 数据无明显差异,但 Akamatsu 等<sup>[16]</sup>的研究发现转移淋巴结 SUV 在 OSEM+TOF 及 OSEM+TOF+PSF 存在明显差异,后者对前者有明显提升作用,结果不同的原因可能是相同重建算法间的重建参数(即高斯滤波器的半峰全

宽)不同(Akamatsu 等<sup>[16]</sup>使用 2 mm,本研究所用为 6.4 mm),该数值不仅是 BSREM 的最适值,同时也体现了实验的可对照性。

Teoh 等<sup>[8]</sup>应用 OSEM + TOF 和 BSREM 的 SUV<sub>max</sub> 进行 ROC 曲线分析,发现使用 2 种算法的 AUC 差异无统计学意义( $P > 0.05$ )。在本研究中,SUV 在 4 种重建算法间 AUC 差异也不具有统计学意义,当应用 SUV<sub>max</sub> 和 SUV<sub>mean</sub> 做 ROC 曲线分析时,BSREM 的 SUV 不仅高于其他重建算法,伴随着预期的特异性降低还提高了诊断灵敏度。BSREM 的特异性通常较低,若结合其他诊断参数,如临床病史或先前的影像学表现,可提高特异性。但本研究的 4 种重建算法的灵敏度较既往文献偏低<sup>[8]</sup>,这可能是因为本研究中病灶整体偏小且占比较高,恶性和良性结节中小结节占比均高于 50%,而结节偏小时,PET 能提供的细节更少。

综上所述,4 种不同重建算法中,BSREM 不仅在视觉方面可改善图像质量,还可明显提高 10 mm 以下肺结节的 SUV<sub>max</sub> 及 SUV<sub>mean</sub>,其 SUV 良恶性诊断阈值应适当上调,但对于良恶性结节的诊断能力无统计学意义,可能与本研究样本量有限有关,今后需加大样本量进一步随访研究。

**利益冲突** 所有作者均声明不存在利益冲突

## 参 考 文 献

- [1] Ettinger DS, Aisner DL, Wood DE, et al. NCCN Guidelines Insights: Non-Small Cell Lung Cancer, Version 5.2018 [J]. J Natl Compr Cane Netw, 2018, 16(7): 807-821. DOI: 10.6004/jnccn.2018.0062.
- [2] 中华医学会呼吸病学分会肺癌学组,中国肺癌防治联盟专家组. 肺结节诊治中国专家共识(2018 年版)[J]. 中华结核和呼吸杂志, 2018, 41(10): 763-771. DOI: 10.3760/cma.j.issn.1001-0939.2018.10.004.
- [3] Adams MC, Turkington TG, Wilson JM, et al. A systematic review of the factors affecting accuracy of SUV measurements[J]. AJR Am J Roentgenol, 2010, 195(2): 310-320. DOI: 10.2214/AJR.10.4923.
- [4] Brendle C, Kupferschläger J, Nikolaou K, et al. Is the standard uptake value (SUV) appropriate for quantification in clinical PET imaging? —Variability induced by different SUV measurements and varying reconstruction methods[J]. Eur J Radiol, 2015, 84(1): 158-162. DOI: 10.1016/j.ejrad.2014.10.018.
- [5] Khalaf M, Abdel-Nabi H, Baker J, et al. Relation between nodule size and <sup>18</sup>F-FDG-PET SUV for malignant and benign pulmonary

- nodules[J]. J Hematol Oncol, 2008, 1: 13. DOI:10.1186/1756-8722-1-13.
- [6] Nomori H, Watanabe K, Ohtsuka T, et al. Visual and semiquantitative analyses for F-18 fluorodeoxyglucose PET scanning in pulmonary nodules 1 cm to 3 cm in size[J]. Ann Thorac Surg, 2005, 79 (3): 984-988; discussion 989. DOI:10.1016/j.athoracsur.2004.07.072.
- [7] Kim SK, Allen-Auerbach M, Goldin J, et al. Accuracy of PET/CT in characterization of solitary pulmonary lesions[J]. J Nucl Med, 2007, 48(2): 214-220.
- [8] Teoh EJ, McGowan DR, Bradley KM, et al. Novel penalised likelihood reconstruction of PET in the assessment of histologically verified small pulmonary nodules[J]. Eur Radiol, 2016, 26(2): 576-584. DOI:10.1007/s00330-015-3832-y.
- [9] Howard BA, Morgan R, Thorpe MP, et al. Comparison of Bayesian penalized likelihood reconstruction versus OS-EM for characterization of small pulmonary nodules in oncologic PET/CT[J]. Ann Nucl Med, 2017, 31(8): 623-628. DOI:10.1007/s12149-017-1192-1.
- [10] Teoh EJ, McGowan DR, Macpherson RE, et al. Phantom and clinical evaluation of the Bayesian penalized likelihood reconstruction algorithm Q.Clear on an LYSO PET/CT System[J]. J Nucl Med, 2015, 56(9): 1447-1452. DOI:10.2967/jnumed.115.159301.
- [11] Mosmann MP, Borba MA, de Macedo FP, et al. Solitary pulmonary nodule and <sup>18</sup>F-FDG PET/CT. Part 1: epidemiology, morphological evaluation and cancer probability[J]. Radiol Bras, 2016, 49(1): 35-42. DOI:10.1590/0100-3984.2014.0012.
- [12] Bettinardi V, Presotto L, Rapisarda E, et al. Physical performance of the new hybrid PET / CT Discovery-690[J]. Med Phys, 2011, 38(10): 5394-5411. DOI:10.1118/1.3635220.
- [13] Jaskowiak CJ, Bianco JA, Perlman SB, et al. Influence of reconstruction iterations on <sup>18</sup>F-FDG PET/CT standardized uptake values [J]. J Nucl Med, 2005, 46(3): 424-428.
- [14] Tong S, Alessio AM, Kinahan PE. Noise and signal properties in PSF-based fully 3D PET image reconstruction: an experimental evaluation [J]. Phys Med Biol, 2010, 55(5): 1453-1473. DOI:10.1088/0031-9155/55/5/013.
- [15] Park R, Miller BW, Jha AK, et al. A prototype detector for a novel high-resolution PET system: BazookaPET [J]. IEEE Nucl Sci Symp Conf Rec (1997), 2012, 2012: 2123-2127. DOI:10.1109/NSSMIC.2012.6551486.
- [16] Akamatsu G, Mitsumoto K, Taniguchi T, et al. Influences of point-spread function and time-of-flight reconstructions on standardized uptake value of lymph node metastases in FDG-PET[J]. Eur J Radiol, 2014, 83(1): 226-230. DOI:10.1016/j.ejrad.2013.09.030.
- [17] 吴天棋,庄静文,谢峰,等.飞行时间技术和高清技术对 PET/CT 图像质量的影响[J].中国医学装备,2016,13(1): 2-5. DOI:10.3969/J.ISSN.1672-8270.2016.01.001.  
Wu TQ, Zhuang JW, Xie F, et al. Research on impact of time of flight acquisition and high definition reconstruction on PET/CT image quality[J]. China Medical Equipment, 2016, 13 (1): 2-5. DOI:10.3969/J.ISSN.1672-8270.2016.01.001.
- [18] Maffione AM, Grassetto G, Rampin L, et al. Molecular imaging of pulmonary nodules[J]. AJR Am J Roentgenol, 2014, 202 (3): W217-223. DOI:10.2214/AJR.13.11733.
- [19] Shankar LK, Hoffman JM, Bacharach S, et al. Consensus recommendations for the use of <sup>18</sup>F-FDG PET as an indicator of therapeutic response in patients in National Cancer Institute Trials[J]. J Nucl Med, 2006, 47 (6): 1059-1066.
- [20] Messerli M, Kotasidis F, Burger IA, et al. Impact of different image reconstructions on PET quantification in non-small cell lung cancer: a comparison of adenocarcinoma and squamous cell carcinoma[J]. Br J Radiol, 2019, 92 (1096): 20180792. DOI:10.1259/bjr.20180792.

(收稿日期:2019-11-22)

## · 读者 · 作者 · 编者 ·

### 关于投稿提供伦理委员会批准文件及受试对象知情同意书的通告

根据中华医学会杂志社的相关规定,当论文的主体是以人为研究对象的试验时,作者应该说明其遵循的程序是否符合负责人体试验的委员会(单位性的、地区性的或国家性的)所制定的伦理学标准,并提供该委员会的批准文件(注明批准文件号)及受试对象的知情同意书。

本刊编辑部