

Doi: 10.3969/j.issn.1005-5185.2011.12.

论著 Original Research

采用欧洲腰椎体模对多中心腰椎定量CT的精密度和准确度的评估

Assessment of Precision and Accuracy of Spinal QCT in a Multi-centre Clinical Research with a European Spine Phantom

陈祥述¹ Chen Xiangshu程晓光¹ Cheng xiaoguang彭俊红² Peng Junhong李葆青³ Li Baoqing苏晋生⁴ Su Jinsheng戴春来⁵ Dai Chunlai

作者单位

1. 北京积水潭医院放射科 北京
2. 武汉市普爱医院放射科 北京
3. 北京市石景山医院放射科 北京
4. 太原市中心医院放射科 北京
5. 吉林大学第一医院放射科 北京

通讯作者 程晓光

Department of Radiology, Beijing Jishuitan Hospital, Beijing 100035

Author Correspondence to: Cheng xiao guang
E-mail: xiao65@263.net

基金支持或者利益申明

北京市卫生系统高层次卫生技术人才培养项目资助(2009-2-03)

中国图书资料分类法分类号:
R

收稿日期: 2011-09-14

修回日期: 2011-10-05

中国医学影像学杂志
2011年 第19卷 第12期:Chinese Journal of Medical Imaging
2011 Volume 19(12): 页码

【摘要】目的 采用欧洲腰椎体模(ESP)作为标准评估多中心腰椎定量CT(QCT)骨密度测量的重复性和准确性误差。**材料和方法** 2个中心(3台CT分别是东芝64排CT、东芝16排CT及GE64排CT)按照常规腰椎定量CT骨密度测量条件和方法对欧洲腰椎体模(ESP)进行重复测量各10次;另外4个中心(CT分别是东芝64排、东芝16排、西门子64排、GE双排)分别用欧洲体模(ESP)扫描测量一次。**结果** 3台QCT(CT顺序依次是东芝64排,东芝16排,GE64排)对欧洲体模测得的相对精密性CV分别是0.512%、0.523%、0.514%。7台QCT机器间的精密性误差CV分别是L1 4.213%、L2 2.752%、L3 4.786%。7台CT的准确度变化范围和平均值括号内数值分别为L₁1.20%~8.79%(5.1);L₂-0.25%~9.15%(3.8);L₃-7.03%~5.69%(-0.5)。**结论** 参与研究的多中心QCT系统的机器内精密性、机器间精密性及准确度较好,结果准确可靠。

【关键词】 骨密度测量;精密性;准确度;定量CT(QCT);欧洲腰椎体模(ESP)

【Abstract】 Objective To evaluate the reproducibility and accuracy of spinal QCT in a multi-centre clinical research using a European Spine Phantom. **Material and Methods** Three CT scanners (64-row Toshiba, 16-row Toshiba and 64-row GE CT scanner) in two centers the ESP was scanned 10 times each in routine technique without repositioning the phantom. The ESP was scanned one time at the same protocols on each CT scanner in another four centers, whose CT scanners are 64-row Toshiba, 16-row Toshiba, 64-row Siemens and 2-row GE CT scanner respectively. **Results** The coefficient of variation of the first three QCT systems were respectively 0.512%, 0.523% and 0.514% for integral phantom. The inter-instrument precision errors of seven QCT systems were 4.213%, 2.275%, and 4.786% in L1, L2 and L3 respectively. The range (and mean) of accuracy error of seven QCT systems were respectively 1.20% to 8.79% (5.1), -0.25% to 9.15% (3.8) and -7.03% to 5.69% (-0.5) in L1, L2 and L3. **Conclusions** The precision, inter-instrument precision and accuracy of spinal QCT in the multi-centre clinical research are great. The results are accurate and reliable.

【Key words】 Bone densitometry; precision; accuracy; QCT; ESP

骨质疏松症可以通过无创性的骨密度(bone mineral density, BMD)检查进行诊断,骨密度可反应骨强度的75%以上^[1]。测量骨密度的方法主要是双能X线吸收测定法(dual X-ray absorptiometry, DXA)和定量CT(quantitative CT, QCT),QCT比DXA有一些明显的优势^[2]。新近研究显示,中国50岁以上的男性和女性罹患骨质疏松症比例分布是10.4%和31.2%^[3],对骨密度测量的需求将会日

益增高。建立大样本量的数据库对于骨质疏松及其相关疾病的诊断、治疗及预防起着非常大的作用^[3-6]。在多中心研究中对参与的各中心的机器的稳定性、一致性及差异进行评估是非常重要的^[7]。精密性按照评价指标可分为绝对精密性和相对精密性^[8],精密性误差通常是在一定时间内对体模或志愿者重复测量取得^[9-10],精密性误差还可有机器内精密性误差和机器间精密性误差之分^[7];而实际

上国际标准协会 (the International Organization of Standardization , ISO) 定义了精密度两个亚分类: 重复性 (repeatability) 和再现性 (reproducibility)。骨密度测量人员所说的精密度一般是指 ISO 定义的重复性方面的内容^[11~12]。欧洲腰椎体模 (ESP) 是国际公认的骨密度测量评估标准, 主要是对骨密度测量仪器进行标准化和相互校准的^[13~14]。本研究的目的是用一个内部“真实密度”已知的欧洲腰椎体模分别对多中心的腰椎 QCT 骨密度测量系统的精密度和准确度进行评估。

1 材料与方法

6 家中心 (7 台 CT 扫描仪) 参与了这项研究, 作为比较, 又选择了两台 DXA 扫描仪进行测量计算。7 台 CT 扫描仪为中国现阶段常用的 CT 机器。两台 DXA 扫描仪分别是 GE 公司的 Lunar Prodigy 和 Lunar iDXA。7 台 CT 进行定量 CT 检查时分别选用的校准体模是美国 Mindways 公司的 5 样本固体体模 (3 台)、美国 Image Analysis 公司的 3 样本 (3 台) 及 4 样本固体体模 (1 台)。具体各中心编号、位置、医院、CT 机型及校准体模配置见表 1。

表 1 各中心编号、位置、医院、CT 机型及配置校准体模类型

编号	城市	医院	公司机型编号	校准体模类型
1	北京	积水潭医院	东芝 16 排 (QCT1)	Mindways 公司 5 样本固体体模
			东芝 64 排 (QCT2)	Mindways 公司 5 样本固体体模
2				
3	北京	石景山医院	GE 64 排 (QCT3)	Image Analysis 公司 3 样本固体体模
	北京	隆福医院	GE 双排 (QCT4)	Image Analysis 公司 4 样本固体体模
4	武汉	普爱医院	东芝 16 排 (QCT5)	Image Analysis 公司 3 样本固体体模
5	太原		东芝 64 排 (QCT6)	Mindways 公司 5 样本固体体模
6	长春		西门子 64 排 (QCT7)	Image Analysis 公司 3 样本固体体模

笔者采用对同一个欧洲腰椎体模 (编号 NO145, 德国 ORM 公司生产) 进行了测量分析。欧洲腰椎体模 (ESP) 是一种模拟人体腰椎的体模, 是用来对骨密度测量仪器进行标准化和相互校准的, 对 DXA 和 QCT 都适用, 也可用来进行日常的质量控制^[14]。欧洲腰椎体模是由环氧树脂制成的塑料加上其他各种成分达到水和骨等效的固体材料组成, 包含有三个松质骨密度不等的椎体, 三个椎体的松质骨的羟基磷灰石密度分别是 50mg/cm³、100 mg/cm³、200 mg/cm³, 在用 DXA 检查时后前位 (PA) 投照下, 三个椎体的面积骨密度 (BMDa) 分别是 0.50 g/cm²、1.00g/cm²、1.50g/cm²^[7, 9, 13-17]。所有参加研究的 QCT 系统 (除 GE 双排外) 的 CT 按照设定的常规扫描条件进行扫描, 扫描条件及重建方式均为 kV120、mAs125、螺旋扫描采集、标准体部重建方式、层厚为 1mm 的薄层 Volume 数据、DFOV400mm, 床高以各系统初始化时所定的床高为准。GE 双排的扫描条件是 kV120、mA120、椎体中心层面断层扫描。两台 DXA 仪均按标准腰椎后前位扫描条件进行扫描。QCT1、QCT2、QCT3 及 2 台 DXA 按照上述扫描条件及常规腰椎骨密度检查步骤对欧洲腰椎体模 (ESP) 进行扫描分析, 由同一检查者重复

进行 10 次, 无需重新摆位, 2 次重复检查间时间不超过 1 分钟, 以分别计算每台机器的精密度和准确度误差。QCT4-7 分别按照上述条件对欧洲腰椎体模进行 1 次扫描, 以评估各机器的准确度误差情况。以上 QCT 系统 (除 QCT4 外, QCT4 的扫描数据传至 Image analysis 公司的 QCT-5000 工作站进行分析处理) 扫描数据统一在中心 1 的一骨密度工作站 (Mindways 公司的 QCT PRO 工作站) 上进行分析处理, 使用软件的 3D 测量模式进行测量, 分别测出欧洲腰椎体模各椎体松质骨骨密度 (单位 mg/cm³)。DXA 数据在其各自工作站上进行分析处理, 测出各椎体及 L₁-L₃ 的骨密度 (单位 g/cm²)。以上测量时, 除非出现明显错误操作时, 均使用工作站软件的自动功能, 如自动探测边界、自动生成感兴趣区 (ROI) 等。精密度误差用标准差 (s, 绝对精密度) 和变异系数 CV (相对精密度) 表示。

每台机器测量各椎体的变异系数 CV 公式^[8]为:

$$CV_{SD} = \frac{SD}{\bar{X}} \cdot 100\%$$

(1)

其中 $\bar{X}$ 、SD 分别是每个椎体测量 10 次所得的平均值和标准差。

当各检查椎体间方差齐性时,可用下列标准差及变异系数公式^[8,13]计算各机器对欧洲腰椎体模整体的精密度误差:

$$SD = \sqrt{\sum_{j=1}^m SD_j^2 / m} \quad (2)$$

$$CV_{SD} = (SD / \sum_{j=1}^m \bar{x}_j / m) \cdot 100\% \quad (3)$$

$$CV_{SD} = \sqrt{\sum_{j=1}^m CV_j^2 / m} \quad (4)$$

m 为参与的研究对象的数目,在此为欧洲腰椎体模椎体的数量,因此按 $m=3$ 代入计算。 $\bar{X}_j$ 、 SD_j 、 VC_j 分别是各椎体十次测量的平均值、方差和变异系数。

准确度误差用测量值与欧洲腰椎体模各椎体“真实值”比较求得,公式如下:

准确度误差=(测量值-真实值)/真实值 (5)

统计分析采用SPSS16.0统计软件。比较各机器的结果、各机器的精密度差异时使用 t 检验、非参数检验和方差分析等。使用相关分析对QCT与DXA测量结果进行相关分析^[3,7]。

2 结果

QCT1、QCT2、QCT3及2台DXA测得的各椎体BMD的平均值、标准差 s 、变异系数 CV 、对体模整体标准差 SD 及整体变异系数 CV 见表2。由方差分析和线性回归分析显示三台QCT测得的各椎体精密度误差(CV)与椎体的密度呈一定的负相关(Pearson相关系数为 -0.781 , $P=0.013$),即椎体的密度越低,精密度误差越高,反之即椎体的密度越高,精密度误差就越低;而2台DXA测得结果没有此规律。3台QCT对欧洲腰椎体模各椎体测量(L_1 、 L_2 、 L_3)的 CV 平均值分别是 0.818% 、 0.479% 、 0.416% 。

表2 QCT1~3及2台DXA检测结果

仪器型号	L1		L2		L3		b 整体精密度	
	$\bar{X} \pm SD$	CV (%)	$\bar{X} \pm SD$	CV (%)	$\bar{X} \pm SD$	CV (%)	s	CV (%)
QCT1	51.24±0.41	0.807	99.75±0.46	0.464	185.95±0.78	0.419	0.575	0.512
QCT2	53.75±0.36	0.671	103.36±0.57	0.553	195.95±0.82	0.420	0.613	0.523
QCT3	54.40±0.53	0.976	109.15±0.46	0.421	210.21±0.86	0.408	0.640	0.514
							L1-L3	
							$\bar{X} \pm S$	CV (%)
Prodigy	a0.611±0.004	0.692	1.094±0.003	0.316	1.583±0.014	0.889	1.134±0.007	0.588
iDXA	a0.615±0.005	0.867	1.111±0.004	0.398	1.631±0.010	0.586	1.158±0.003	0.260

注:QCT测量的结果单位为 mg/cm^3 ,DXA测量结果的单位为 g/cm^2 ;a处两结果间差异无统计学意义 $P=0.081$;b处按照公式(3)计算得出。

对这5台设备的测量结果进行相关分析显示,每两个设备的测量结果间存在明显的正相关($P<0.01$),Spearman等级相关系数变化范围为 $0.862 \sim 0.934$,其中QCT1(东芝64排CT)与

QCT3(GE64排CT)的相关性最高。

7台QCT和2台DXA测得结果与欧洲腰椎体模相应椎体的真实值比较计算所得的准确度误差及机器间精密度误差见表3和图1、2。

表3 7台QCT和2台DXA测得的准确度误差及机器间精密度误差 $[\text{mg}/\text{cm}^3, \text{百分率}(\%)]$

仪器	型号	L ₁		L ₂		L ₃	
		BMD	准确度	BMD	准确度	BMD	准确度
QCT	CT机型						
QCT1	东芝64排	51.24	(2.5)	99.75	(-0.3)	185.95	(-7.0)
QCT2	东芝16排	53.75	(7.5)	103.36	(3.4)	195.05	(-2.5)
QCT3	GE64排	54.40	(8.8)	109.15	(9.1)	210.21	(5.1)
QCT4	GE双排	50.60	(1.2)	103.50	(3.5)	199.50	(-0.3)
QCT5	东芝16排	50.77	(1.5)	102.03	(2.0)	200.17	(0.1)

(续 表)

仪器	型号	L ₁		L ₂		L ₃	
		BMD	准确度	BMD	准确度	BMD	准确度
QCT7	西门子 64 排	50.95	(1.9)	104.44	(4.4)	211.38	(5.7)
平均值	52.56	(5.1)	103.76	(3.8)	198.92	(-0.5)	
	标准差	2.2143		2.8554		9.5212	
	变异系数 CV	(4.213)		(2.752)		(4.786)	
DXA	机型						
1	Prodigy	0.6109	(22.2)	1.0942	(9.4)	1.5831	(5.5)
2	iDXA	0.6150	(23.0)	1.1113	(11.1)	1.6312	(8.7)
	平均值	0.6130	(22.6)	1.1028	(10.3)	1.6072	(7.1)
	标准差	0.00290		0.01209		0.034011	
	变异系数	(0.472)		(1.096)		(2.116)	

QCT1-3、两台DXA使用的是十次测量结果的平均值；欧洲腰椎体模各椎体L₁、L₂、L₃体积骨密度为50、100、200mg/cm³,面积骨密度分别为0.5、1.0、1.5g/cm²

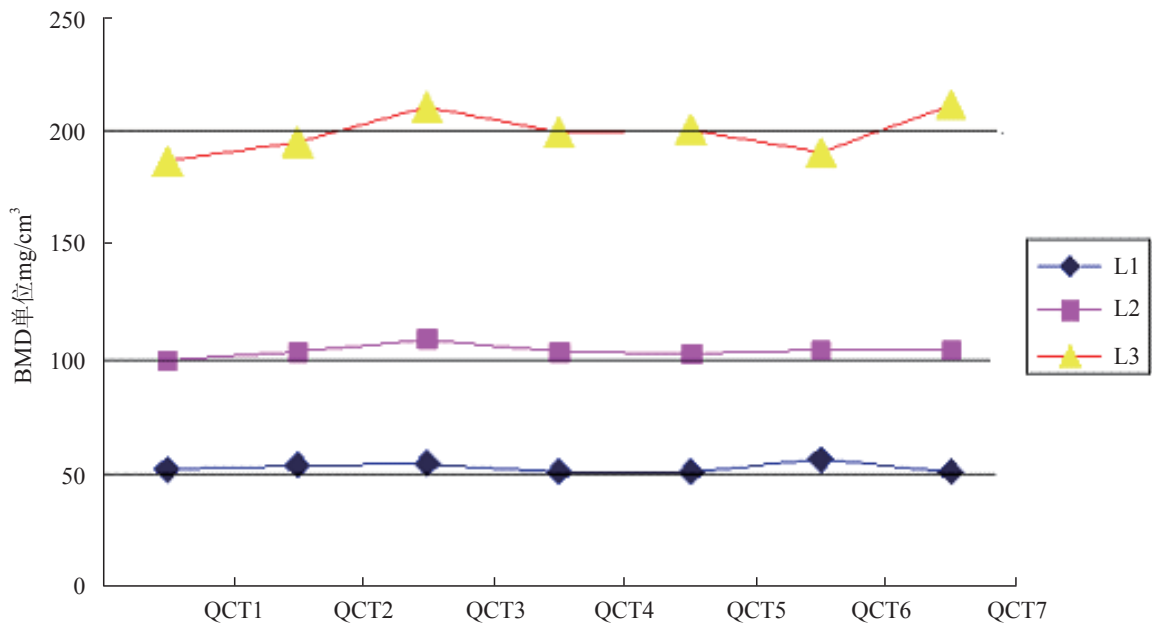


图1 7台QCT测得结果与欧洲腰椎体模相应椎体的真实值比较（欧洲腰椎体模L₁、L₂、L₃体积骨密度分别是50、100、200mg/cm³）

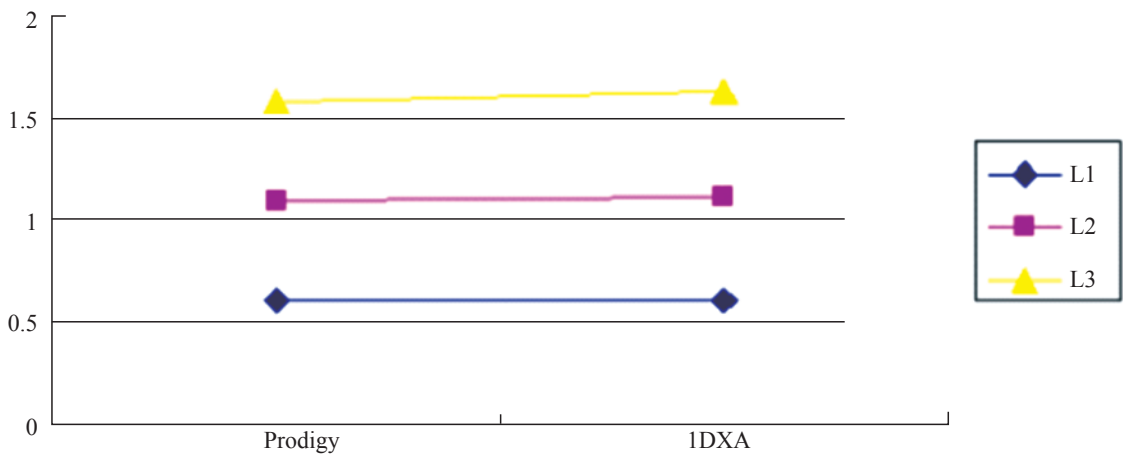


图2 2台DXA测得结果与欧洲腰椎体模各椎体真实值比较（欧洲腰椎体模L₁、L₂、L₃面积骨密度分别是0.5、1.0、1.5g/cm²）

7 台 QCT 对欧洲腰椎体模各椎体 (L_1 、 L_2 、 L_3) 机器间的相对精密度误差 CV 分别是 4.213%、2.752%、4.786%。可见机器间精密度误差远大于各机器的精密度误差。7 台 QCT 测得值与欧洲腰椎体模的“真值”的 Pearson 相关系数均大于 0.9999, $P < 0.05$ 。

3 讨论

本研究显示, 3 台 QCT 的机器内精密度误差、2 台 DXA 的机器内精密度误差在机器之间差异无统计学意义, 这与 Genant 等^[18]、Lees 等^[7]用欧洲腰椎体模测得 DXA 机器内精密度误差结果相一致。同时我们发现 3 台 QCT 测得的机器内精密度误差与所测的椎体密度值呈负相关 (Pearson 相关系数为 -0.781 , $P = 0.013$), 即密度越低的椎体 QCT 对其测得精密度误差越高, 重复性越差。而 DXA 检查却无此规律。这与 Gliier 等^[8]认为骨质疏松的患者人群比正常人群所测的精密度误差要高相一致。由于正常情况下, 松质骨的代谢活性是皮质骨的 8 倍, Adams^[2]认为松质骨为主的部位比皮质骨为主的部位骨密度的变化要大, 这也与笔者的结果相一致。

在本研究中, 7 台 QCT 对欧洲腰椎体模各椎体 (L_1 、 L_2 、 L_3) 机器间的精密度误差 CV 分别是 4.213%、2.752%、4.786%, 机器间精密度误差远大于各机器的精密度误差。结果较 Pearson 等^[13]用欧洲腰椎体模测得机器间精密度误差要高, 这可能是由于笔者测得的结果是 7 台 QCT 机器总的精密度误差, 由于参与的机器较少, 没有按厂家分别进行计算, 而 Pearson 等的结果是分厂家机器分别计算的。如果采用一些新技术和方法, QCT 间的精密度误差会降低, 如 Li 等^[19]介绍的自动注册系统和 3D 测量、结构分析三种最新技术方法结合在一起, 进行腰椎和髋部骨密度测量, 既能减少用户主观因素方面的影响, 又能减少运算时间, 重要之处是还能提高精密度, 这将会进一步推动 QCT 在临床及试验方面的应用。

7 台 QCT 对欧洲腰椎体模的测量的准确度误差比 2 台 DXA 测量的准确度误差低, 7 台 QCT 的准确度误差均在文献报导的 4% ~ 15% 范围内^[2, 18, 20-22]。QCT 可以去除一些对 DXA 影响较明显的因素, 如周围组织结构的影响等。对于 2 台 Lunar DXA 的准确度误差较高, 特别是对 L_1 椎体的过高评估的原因, Lees 等^[7]认为是由于 Lunar 机器使用的边界探测程序去除更多低密度的面积, 从而造成骨密度增高,

使结果偏离真实值。

虽然 7 台 QCT 的准确度误差在可接受范围内, 但是机器之间存在一定的系统误差, 如 QCT1 测的欧洲腰椎体模的各椎体普遍比 QCT2、QCT3 低, 因此比较不同机器间结果时必须进行校正, 对每台机器建立校正方程后, 数据进行校正后才可以进行比较^[7, 13, 23]。欧洲腰椎体模作为标准的校准体模在国外广泛应用于 DXA 的临床与研究工作中^[7, 9, 11, 13, 23], 而国内应用的较少^[3], 应用于 QCT 的研究就更少。但是随着骨密度测量学界重新对 QCT 的逐渐重视, 对于 QCT 一些新的认识逐渐达成共识^[2, 24-25], 对于欧洲腰椎体模的也将会逐渐重视起来。

本研究显示, 参加研究的多中心 QCT 系统的机器内精密度、机器间精密度及准确度较好, ESP 作为标准评估 QCT 机器精密度和准确度是可靠的, 可为临床上开展 QCT 多中心研究提供基础保证。

参考文献

- [1] Osteoporosis: review of the evidence for prevention, diagnosis and treatment and cost-effective analysis. *Osteoporos Int*, 1998, 8(Suppl 4): S3-S6.
- [2] Adams JE. Quantitative computed tomography. *Eur J Radiol*, 2009, 71(3): 415-424.
- [3] Zhu H, Fang J, Luo X, et al. A survey of bone mineral density of healthy Han adults in China. *Osteoporos Int*, 2010, 21(5): 765-772.
- [4] Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. Report of a WHO Study Group. *World Health Organ Tech Rep Ser*, 1994, 843: 1-129.
- [5] Bonjour JP, Ammann P, Rizzoli R. Importance of preclinical studies in the development of drugs for treatment of osteoporosis: a review related to the 1998 WHO guidelines. *Osteoporos Int*, 1999, 9(5): 379-393.
- [6] Prevention and management of osteoporosis. *World Health Organ Tech Rep Ser*, 2003, 921: 1-164.
- [7] Lees B, Garland SW, Walton C, et al. Evaluation of the european spine phantom in a multi-centre clinical trial. *Osteoporos Int*, 1997, 7(6): 570-574.
- [8] Gliier CC, Blake G, Lu Y, et al. Accurate assessment of precision errors: how to measure the reproducibility of bone densitometry techniques. *Osteoporosis Int*, 1995, 5(4): 262-270.
- [9] Young M. Cross-calibration of axial DEXA systems. *Radiography*, 2001, 7(1): 79-83.
- [10] Tothill P, Hannan WJ. Precision and accuracy of measuring changes in bone mineral density by dual-energy X-ray absorptiometry. *Osteoporos Int*, 2007, 18(11): 1515-1523.

- [11] Sydney Lou Bonnick. The Third Edition of Bone Densitometry in Clinical Practice. New York: Humana Press, 2010: 257-278.
- [12] Engelke K, Glüer CC. Quality and performance measures in bone densitometry: part 1: errors and diagnosis. *Osteoporos Int*, 2006, 17(9): 1283-1292.
- [13] Pearson J, Dequeker J, Henley M, et al. European semi-anthropomorphic spine phantom for the calibration of bone densitometers: assessment of precision, stability and accuracy. The European Quantitation of Osteoporosis Study Group. *Osteoporosis Int*, 1995, 5(3): 174-184.
- [14] Kalender WA, Felsenberg D, Genant HK, et al. The European Spine Phantom a tool for standardization and quality control in spinal bone mineral measurements by DXA and QCT. *Eur J Radiol*, 1995, 20(2): 83-92.
- [15] Pearson D, Cawte SA, Green DJ. A comparison of phantoms for cross-calibration of lumbar spine DXA. *Osteoporos Int*, 2002, 13(12): 948-954.
- [16] Kolta S, Ravaud P, Fechtenbaum J, et al. Accuracy and precision of 62 bone densitometers using a european spine phantom. *Osteoporos Int*, 1999, 10: 14-19.
- [17] Ellis KJ, Shypailo RS, Steinberg FM, et al. Reproducibility of fan-beam DXA measurements in adults and phantoms. *J Clin Densitom*, 2004, 7(4): 413-418.
- [18] 林强, 姚金朋, 余卫. 定量 CT 在骨密度测量中的应用. *中华放射学杂志*, 2009, 43(2): 219-221.
- [19] Li W, Sode M, Saeed I, et al. Automated registration of hip and spine for longitudinal QCT studies: integration with 3D densitometric and structural analysis. *Bone*, 2006, 38(2): 273-279.
- [20] Glüer CC, Engelke K, Lang TF, et al. Quantitative computed tomography (QCT) of the lumbar spine and appendicular skeleton. *Eur J Radiol*, 1995, 20(3): 173-178.
- [21] Guglielmi G. Quantitative computed tomography (QCT) and dual X-ray absorptiometry (DXA) in the diagnosis of osteoporosis. *Eur J Radiol*, 1995, 20(3): 185-187.
- [22] Guglielmi G, Gluer CC, Majumdar S, et al. Current methods and advances in bone densitometry. *Eur Radiol*, 1995, 5(2): 129-139.
- [23] Genant HK, Grampp S, Glüer CC, et al. Universal standardization for dual X-ray absorptiometry: patient and phantom cross-calibration results. *J Bone Miner Res*, 1994, 9(10): 1503-1514.
- [24] Reid DM, Mackay I, Wilkinson S, et al. Cross-calibration of dual-energy X-ray densitometers for a large, multi-center genetic study of osteoporosis. *Osteoporos Int*, 2006, 17(1): 125-132.
- [25] Engelke K, Adams JE, Armbrrecht G, et al. Clinical use of quantitative computed tomography and peripheral quantitative computed tomography in the management of osteoporosis in adults: the 2007 ISCD Official Positions. *J Clin Densitom*, 2008, 11(1): 123-162.

(责任编辑 纳琨)