

用定量形态学测量法检测再发性脊柱骨折

吴春营 李 皎 樊 波 李景学 刘忠厚

摘要 应用定量形态学测量法(quantitative morphometric assessment, QM)和半定量分析法(semi-quantitative assessment)诊断 335 名绝经期后妇女的再发性脊椎骨折,以比较这两种方法在诊断上的差异。由三名受过训练的放射科医师分别用 SQ 法对比每位受检者的脊椎平片的基础片和随访片作出初诊,再由该三名医师共同复阅所有平片,结合各自诊断和 QM 的结果得出会诊阅片结果(Consensus Reading, CR)。并使用 QM 法的六点数字式技术测量平片,采用四种形态学定义作为诊断标准。结合每一标准所选用的四个常用划分骨折的阈值作出再发性骨折的诊断,即(1)随访片内椎高减少百分比(PHr);(2)椎高减少绝对值(AHr);(3)PHr 加 AHr 以及(4)椎高比率减少绝对值(ARr)。比较定量法与半定量法的诊断结果显示,CR 诊断再发性骨折率为 0.955%,三名医师之间以及各自与 CR 之间的诊断均有很高的符合率,Kappa 值(K 值)达 0.86~0.96;敏感性为 84%~93%;特异性为 99.9%。QM 与 CR 的 K 值为 0.27~0.63;敏感性为 30%~80%,特异性为 95%~99.9%。其中最高的 K 值为 AHr 的 0.63,相应的敏感性为 51%,特异性为 99.98%。结果显示,定量形态测量法与半定量分析法对诊断再发性脊椎骨折的符合率仅为轻至中度。对于流行病学和临床药效的研究,单独使用定量形态测量法并非是一可靠的手段。因此该法应结合受训过的放射学医师或有相当经验的临床医师的阅片结果一并使用。

关键词: 骨质疏松 骨折,闭合性 形态学测量

The application of quantitative morphometry for assessment of incident vertebral fractures, comparison with semiquantitative technique

Wu Chunying, Li Jiao, Fan Bo, et al

Department of radiology, University of California, San Francisco, CA 94143-0628, USA

Abstract: This study compared semiquantitative reading and quantitative morphometry for the assessment of incident vertebral fractures. Radiographs of the thoracolumbar spine taken one year apart for 335 postmenopausal women with suspected vertebral fractures were evaluated. Quantitative morphometry (QM) and visual semiquantitative grading by 3 trained radiologists independently and in a joint consensus reading (CR) session were used to determine incident vertebral fractures. QM was performed using a six-point-digitization technique, and incident deformities using QM were defined using a range of cutoffs with four different morphometric criteria. By the CR 9.55% of the women and 1.33% of the vertebrae had incident fractures after 1 year. Excellent agreement was found among the radiologists' readings as well as between the individual readings and the CR for the assessment of incident fractures, with kappa scores between 0.86 and 0.96. Sensitivities ranged from 84% to 93% with specificities of 99.9% relative to the CR. Using selected cutoff thresholds with the four QM criteria, kappa scores between QM and the CR ranged from 0.27 to 0.63, with corresponding sensitivities from 30% to 80% and specificities from 95% to 99.9% relative to the CR. The assessment of incident vertebral fractures using QM showed only fair to moderate agreement with the consensus reading or with the individual read -

ings. The assessment of incident fractures using QM alone may not be sufficiently reliable for detection of incident vertebral fractures in epidemiological studies and in clinical trials, and it should be performed principally in conjunction with the reading of a trained radiologist or a highly experienced clinician.

Key words: Osteoporosis Fractures, closed Morphometry

确定椎骨骨折初发率(prevalent vertebral fracture)和再发率(incident vertebral fracture)对于骨质疏松流行病学研究和药效判定是一个重要指标。X线平片评定脊椎骨折的再发率则是常用的最终手段^[1]。临床上通常是通过观察病人的脊椎平片诊断脊椎骨折,以便制定出相应的治疗方案。但是许多学者认为此法相当主观,这种单纯通过目测诊断的方法,使得检查的结果变异性甚大。这种变异尤其见于经验不足的阅片者。定量形态学测量法(quantitative morphometry assessment, QM)则具有客观性强、重复性高的优点,但它的使用却可能因诊断阈值的选择以及与正常值的比较,或所使用的数字式技术的不统一或误差而带来一些问题。许多研究人员报道了使用各种半定量性和定量性技术诊断脊椎骨折初发率^[2,3]。但运用这项技术诊断再发性骨折却未引起足够的关注^[4,5]。作者比较上述两种诊断再发性脊椎骨折的方法,即由有经验的放射科医师采用半定量法阅片,以及采用定量性形态测量技术的六点数字式方法,结合不同的形态学诊断标准确定再发性脊椎骨折,旨在探讨运用这两种方法在诊断上的差异,以达到进一步完善对脊椎骨折的诊断步骤。

1. 材料和方法

受检者选自 335 名绝经期后妇女。选择标准包括:双能 X 线吸收法(DXA)测得正位腰椎骨密度低于 T 数的两个标准差;每位病人的基础片内显示至少有一个以上,但少于七个椎体骨折。这些平片检查包括:首次正侧位胸腰椎基础片,一年后的侧位胸腰椎随访片。两套平片均采用标准的摄片技术获得。

1.1 半定量检测椎体变形:由三名放射科医师分别采用半定量法目测椎体变形的程度。其中一名(R1)为经验丰富者,另两名(R2, R3)为受

过训练者。采用 Genant 等的半定量分级标准将两套平片内的 13 个椎体(T4-L4)分为正常(0 度),轻度骨折(1 度),中度骨折(2 度),或重度骨折(3 度)(附图)。如随访片内的某椎体得分高于基础片内的对应椎体则视为再发性骨折^[3]。全部平片均由其中一名医师(R2)标定椎体水平。所有医师均采用该标定椎体水平进行阅片。在三名医师独立阅片后,再由该三名医师一同复阅所有平片。此次阅片则参考各位医师的诊断结果,以及定量形态学测量法的结果(QM 只采用椎体绝对高度减少 15%作为诊断标准)。将此共同阅片结果做为“会诊结果”(Consensus Reading, CR)。

1.2 形态测量法检测椎体变形:所有平片内 T4~L4 均经定量形态学测量法测量。于每一椎体轮廓四角以及上下终板中点(即椎体前后缘之间。如因投照因素使终板出现上下二缘,则取二缘上下中点)处共取六点用笔标好。上述标点所测得的位置均坐落于数字测量仪的数字式台板的 X-Y 坐标上,再由计算机程序根据坐标计算出所有椎体的前、中和后的高度。在所有病人平片内的 4 355 个椎体中,54 个椎体因显示不清而未行测量。再选出 40 例经 CR 诊断无再发性骨折的基础和随访平片重复测量,以分析 QM 法的重复精度,用方差均值表示(Root Mean Square)。根据基础和随访片的测量值,作者使用四种形态学标准定义再发性脊椎骨折:(1)椎体任一高度减少百分比(PHr);(2)椎体任一高度减少绝对值(AHr);(3)椎体高度减少百分比加减少绝对值(PAHr);(4)椎体高度比率减少的绝对值(ARr)。每一个标准内又选用四个常用参数作为诊断阈值(表 1)。

表1 四种诊断标准及其各法选用的阈值

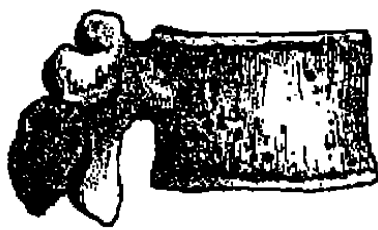
方法	诊 断 标 准			
	PHr	AHr	PAHr	ARr
1	$\geq 15\%$	$\geq 3\text{mm}$	$\geq 15\% + 3\text{mm}$	≥ 0.12
2	$\geq 17.5\%$	$\geq 4\text{mm}$	$\geq 17.5\% + 3\text{mm}$	≥ 0.14
3	$\geq 20\%$	$\geq 5\text{mm}$	$\geq 20\% + 4\text{mm}$	≥ 0.16
4	$\geq 25\%$	$\geq 6\text{mm}$	$\geq 25\% + 4\text{mm}$	≥ 0.18

注:PHr,任一椎体高度减少百分比;AHr,任一椎体高度减少绝对值;PAHr,PHr加AHr;ARr,任一椎体高度比减少值,即: H_a/H_p , H_m/H_p , H_p/H_p+1 或 H_p/H_p-1 (H_a 为椎体前高, H_m 为椎体中高, H_p 为椎体后高, H_p+1 为上一椎体后高, H_p-1 为下一椎体后高)。

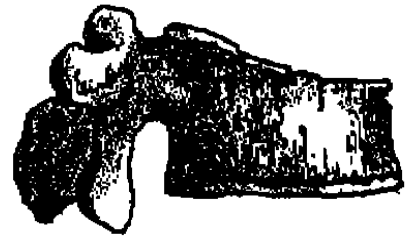
1.3 统计学处理:所有阅片者之间、阅片者与CR之间、QM与CR之间的敏感性、特异性和Kappa值(K值)经统计学处理得出。Kappa统计学分析是用于衡量定性诊断结果之间(或诊断结论只包括阳或阴性的数据)的一种相关检验法。该值由-1至1。-1和1分别表示完全负相关和完全正相关。K值大于0.8以上为极具相关性,K值等于0.6~0.8为高度相关性,K值为0.4~0.6为中度相关性,小于0.4则为弱或无相关性。

2. 结果

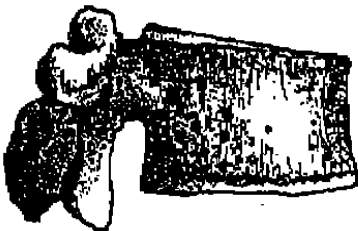
我们使用的QM法测量重复精度分别为4.73%(前高)、4.35%(中高)和3.71%(后高)。表2分别显示应用半定量法、定量法和CR所得出的再发性椎体骨折率的绝对值和发生率。采用CR的骨折再发率为1.33%(57个)。三名放射科医师诊断的发生率1.23%~1.26%。根据表1的阈值用定量形态学方法测出的骨折再发率PHr为0.5%~2.6%、AHr为0.8%~6.1%、PAHr为0.5%~2.2%、ARr为1.1%~5.1%。表3显示三名放射科医师及QM法与CR的比较结果。由表可看出三名医师之间有很高的符合率,K值达0.8~0.84。各位医师与CR的诊断结果比较则显示K值为0.86~0.96。但QM法与CR之间的K值仅为弱至中度相关。所有的形态测量诊断结果,K值最高者为0.63,是采用ARr的任意一个椎体高度减少6mm以上的阈值而得出。其相应的敏感性为50.9%、特异性为99.9%。由各位医师诊断再发性骨折出现部位与CR之间



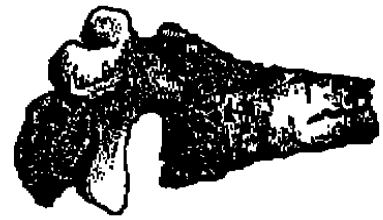
0 级



2 级



1 级



3 级

注:0级:正常椎体,1级:轻度楔形骨折,椎体前高减少20%~25%,2级:中度楔形骨折,椎体前高减少25%~40%,3级:重度楔形骨折,椎体前高减少>40%(椎体中高和后高减少使用相同分眼法,随访片内的椎体分级大于基础片内的相应椎体分级相应则定义为再发性骨折)

附图 半定量法判定骨折的椎体前高减少分级示意图

表 2 半定量法和定量法诊断再发骨折结果

诊 断 指 标	再发骨折率(%)	
	椎体数	病人数
CR	57(1.33)	32(9.55)
R1	53(1.23)	29(8.66)
R2	53(1.23)	28(8.36)
R3	54(1.26)	31(9.25)
PHr		
≥15%	110(2.56)	66(19.70)
≥17.5%	62(1.44)	42(12.54)
≥20%	30(0.91)	27(8.06)
≥25	24(0.56)	20(5.97)
AHr		
≥3mm	260(6.05)	135(40.30)
≥4mm	110(2.56)	64(19.10)
≥5mm	58(1.35)	39(11.64)
≥6mm	35(0.81)	26(7.76)
PAHr		
≥15%+3mm	96(2.23)	59(17.61)
≥17.5%+3mm	55(1.27)	38(11.34)
≥20%+4mm	34(0.79)	24(7.16)
≥25%+4mm	22(0.51)	18(5.37)
ARr		
≥0.12	220(5.12)	127(37.91)
≥0.14	130(3.02)	88(26.26)
≥0.16	80(1.86)	62(18.51)
≥0.18	49(1.14)	41(12.24)

注:标准诊断法(CR),各位放射医师(R₁、R₂、R₃),定量法使用四种诊断标准。每一种标准选用四个阈值

亦极相似。当采用较宽松的阈值作为定量形态学方法诊断再发性骨折时,其发生率往往高于CR的发生率。反之,当采用较严格的阈值时,其发生率则比CR要低得多。

3. 讨论

对如何运用更客观和更恒定的手段来诊断椎体骨质疏松性骨折已进行了大量的研究。一些作者曾报道了使用定量形态测量法的数字式技术诊断初发性椎体骨折^[3,6,7]。但是,由于这些研究中没有引用公认的诊断标准作为比较,因此其结果是否可靠使人置疑。此外,运用当前的定量性方法取得的结果常常大相径庭^[4,8~10]。大多数用于诊断初发性椎体骨折的运算法多不适用再发性椎体骨折的诊断。QM诊断初发性椎体骨折多依赖于将每一椎节结果与正常人群的相应椎节进行比较。而对于诊断再发性

骨折则需直接将随访片的每一椎节的测量结果与基础片的相应椎节进行比较。本文中,作者采用数种形态测量定义脊椎再发骨折。每一标准又选用不同的诊断阈值来观察基础和随访片内椎体形态上的差别。

QM的六点数字式测量技术已很成熟,报道显示其测量重复性令人满意。椎高测量精度为7%~9%、椎面积测量精度为6%~11%^[8,11]。我们测量的重复精度为2%~4%^[12]。理论上QM法似乎是一理想的、客观的诊断方法。如选用低于重复精度误差的诊断阈值,应能得到可靠的、并且具有高敏感性和特异性的诊断结果。但是,作者比较了QM法与CR诊断结果,四个形态测量标准所选用的阈值与后者只呈轻度至中度的符合率。两者间的这种差异有几个因素可以解释。(1)测量时选择六点上的误差。(2)尽管基础片和随访片使用所谓标准的投照技术和质量监控程序,两套平片的影象仍存在不一致性。基本上说,误差来源于病人投照体位变动造成的影响和放大率的不同。诸如体重的增减,脊柱侧弯和旋转、X线束中心线的变动、管片距离不一致等。这些投照因素影响平片质量上的不稳定性均可对数字式测量的取点产生实质性的影响。因此,可以理解对同一病人的基础片与随访片的测量者自身的重复精度要比重复测量同一平片的重复精度为低。我们对随访片采用校正法计算,以消除因放大率不同造成的误差。这种校正法理论上是可行的,然而由于必需将靶片距离一并考虑,故增加了计算的复杂性还可能对测量结果产生新的误差。Leidig-Bruckner等^[13]将本文使用的半定量标准与定量法所测得的椎体变形及骨折变形指数(SDI)的比较结果显示,QM法诊断结果特异性为86%。QM与SDI两法间存在差异主要见于相对较小的椎体,如上胸段。但总而言之,两法间全部椎节的对比有较好的相关性。具有争议的问题是使用QM法时单独采用某一诊断阈值来确定再发性骨折是否合适。因为采用某一诊断标准为阳性时,另一标准则可能为阴性。尤其有的标准受影像放大率的影响,有的则不

表 3 诊断符合率的统计学结果

诊断指标	符合率(%)	Kappa 统计(K 值)	95(%)可信性	敏感性(%)	特异性(%)
R1	99.86	0.96	0.90~0.99	92.98	100
R2	99.68	0.87	0.80~0.94	84.21	99.88
R3	99.65	0.86	0.79~0.93	84.21	99.86
PHr					
≥15%	98.12	0.51	0.40~0.61	75.44	98.42
≥17.5%	98.81	0.57	0.45~0.68	59.65	99.27
≥20%	99.02	0.56	0.42~0.69	47.37	99.72
≥25%	99.00	0.46	0.31~0.62	33.33	99.88
AHr					
≥3mm	94.72	0.27	0.18~0.36	78.95	94.93
≥4mm	97.98	0.47	0.36~0.58	70.18	98.35
≥5mm	98.95	0.60	0.49~0.72	54.39	99.45
≥6mm	99.20	0.63	0.50~0.75	50.88	99.86
PAHr					
≥15%+3mm	99.44	0.55	0.45~0.66	75.44	98.75
≥17.5%+3mm	98.98	0.60	0.48~0.72	59.65	99.51
≥20%+4mm	99.14	0.59	0.46~0.72	47.37	99.84
≥25%+4mm	99.05	0.48	0.32~0.64	33.33	99.93
ARr					
≥0.12	95.42	0.27	0.17~0.37	70.18	95.76
≥0.14	97.23	0.36	0.25~0.48	61.40	97.76
≥0.16	98.16	0.41	0.29~0.54	50.88	98.80
≥0.18	98.51	0.39	0.24~0.54	36.84	99.34

受影响。因此,如果合并应用不同诊断标准可能提高诊断的准确性。正如本文中所显示的,当合并使用绝对和相对椎体高度减少的特定阈值(PAHr)时,显示诊断特异性较高而并不丧失其敏感性。

从临床的角度看,由于再发性骨折的发生率很低,故诊断方法应具有较高的敏感性和特异性。因为即使敏感性高也可能得出再发性骨折的绝对值高于实际状况,造成对临床药效的误判^[14]。此外单纯使用形态测量法的标准,不可避免的会漏诊许多显而易见的骨折。McCloskey 等采用新的运算法判定一组患转移癌妇女的初发和再发椎体骨折率。使用两个较严格的标准使诊断的假阳性率降低,但敏感性也相应降低。Ross 等^[15]运用椎高低于相应正常人群的未骨折椎体的三个标准差作为阈值。发现只有 70%的再发骨折的椎体被检出。也有人用椎高减少 20%作为诊断阈值^[16]。但本组结果表明,肉眼所见的骨折超过 50%未被 QM 法

检出。

在观察再发性骨折的半定量阅片过程中,阅片者可同时对比随访片与基础片的同一椎节。此法可肉眼纠正椎体因投照因素不同而呈现在两片中的形态上的差异。正如本文的结果显示的 SQ 法与各阅片者之间存在很高的相关性。但当诊断初发性骨折时,由于无前片作为参照则 SQ 法诊断结果的差异可能较大。总之,作者认为 SQ 法诊断脊椎骨折具有较高的可靠性,尤其适用于检测再发性骨折。尽管如此,由于肉眼观测可受到某些投射因素的影响,因而即使多人集体阅片所得的诊断也欠理想。虽然此法并非完善,本文中使用的统计比较方法有利于 SQ 法,而使 QM 法和 SQ 法的相关性显得偏低。尽管如此,作者仍认为 SQ 可作为诊断脊椎再发性骨折的重要手段。可以想象,将 QM 和 SQ 合并使用可能是一适当的途径。QM 法较松的阈值可确定可疑的骨折造成的椎体变形,而 SQ 法则可基于此筛选诊断再发骨折。对

两法有分歧的病例则由一经验丰富的阅片者复审来确定真实的再发骨折。采用这种步骤可提高诊断准确性,使统计结果更具可靠性。

新进展的技术,如形态测量 X 线分析(MXA)可克服某些定量测量技术对传统平片检测的不足。尤其它能减少投射技术不恒定而产生的放大率和投射角度的变化。但此项新技术尚存在图像清晰度欠高及噪声较高的问题。在确认它在诊断再发性骨折的作用之前,尚需通过实验进一步了解其使用价值。

作者的结论是,单独使用 QM 法诊断再发性脊椎骨折不应被视为一可靠的诊断方法。选择某一适用的诊断阈值虽可提高定量性技术和 CR 之间的符合率,但仍有相当的差异。在进行流行病学和临床药效的研究时,该技术的不足应引为重视。有必要对受检者摄片过程进行质量控制(Quality Assurance)以减少其技术上可能造成的误差。此外尚需进一步探讨是否通过应用新的定量性诊断阈值,或合并使用不同阈值或将摄片过程规范化以提高此项技术的诊断能力。但是,目前应将定量性形态测量技术与有经验的放射学专家或临床医师的阅片结合使用以诊断脊椎的再发性骨折。SQ 法具有很高的重复性,能可靠的确定再发性脊椎骨折。故应将其作为主要的诊断手段。

致谢:本研究曾在美国加州大学旧金山医学分校的 Genant 教授的指导和 Jergas 医师的帮助下进行。

参 考 文 献

- 1 Kleerekoper M, Nelson DA, Peterson EL, et al. Outcome variables in osteoporosis trials. *Bone*, 1992, 13: S29.
- 2 Hedlund LR, Gallagher JC. Vertebral morphometry in diagnosis of spinal fractures. *Bone Miner*, 1988, 5: 59.
- 3 Genant HK, Wu CY, vanKuijk C, et al. Vertebral fracture assessment using a semi-quantitative technique. *J Bone Miner Res*, 1993, 8: 1137.
- 4 McCloskey E, Spector TD, Eyres KS, et al. The assessment of vertebral deformity: A method for use in population studies and clinical trials. *Osteoporosis Int*, 1993, 3: 138.
- 5 Riggs BL, Hodgson SF, O'Fallon WM, et al. Effect of fluoride treatment on the fracture rate in postmenopausal women with osteoporosis. *N Engl J Med*, 1990, 322: 802.
- 6 Gallagher JC, Hedlund LR, Stoner S, et al. Vertebral morphometry; normative data. *Bone Mineral*, 1988, 4: 189.
- 7 Davies KM, Recker RR, and Heaney RP. Normal vertebral dimensions and normal variation in serial measurements of vertebrae. *J Bone Miner Res*, 1989, 4: 341.
- 8 Smith-Bindman R, Cummings SR, Steiger P, et al. A comparison of morphometric definitions of vertebral fracture. *J Bone Miner Res*, 1991, 6: 25.
- 9 McCloskey EV, Spector TD, Khan S, et al. Prevalence of vertebral deformity and concordance between definitions of fracture. In: *Proceedings of Fourth International Symposium on Osteoporosis and Consensus Development Conference*, Hong Kong, 1993, C. Christiansen and B. Riis (ed.). *ApS Osteopress*, 62. from a single radiograph to identify fractures. *Calcif Tissue Int*, 1992, 51: 95.
- 10 Leidig G, Storm T, Genant HK, et al. Comparison of two methods to assess vertebral fractures. In: Christiansen C. and Overgaard K. *Proceedings of Third International Symposium on Osteoporosis*. (ed.). Denmark: Copenhagen, 626-628.
- 11 Nelson D, Peterson E, Tilley B, et al. Measurement of vertebral area on spine X-rays in osteoporosis; reliability of digitizing techniques. *J Bone Res*, 1990, 5: 707.
- 12 Li J, Wu CY, Jergas M, et al. Comparison of Semi-quantitative and Quantitative Methods for the Assessment of Prevalent Vertebral Fractures. *Osteoporosis Int* submitted for publication.
- 13 Leidig-Bruckner G, Genant HK, Minne HW, et al. Comparison of a semiquantitative and a quantitative method for assessing vertebral fractures in osteoporosis. *Osteoporosis Int*, 1994, 4: 154.
- 14 Kanis JA, McCloskey EV, Khan S, et al. What is a vertebral fracture? In: Christiansen C. and Riis (ed.). *Proceedings of Fourth International Symposium on Osteoporosis and Consensus Develop-*

- ment Conference, Hong Kong, 1993, ApS Osteopress, 26-28.
- 15 Ross PD, Davis JW, Epstein RS, et al Ability of vertebral dimensions from a single radiograph to identify fractures. *Calcif Tissue Int*, 1992, 51:95.
- 16 Cummings SR, Melton III LJ, Felsenberg D. Osteoporosis Foundation Working Group on Vertebral Fracture. Report: Assessing vertebral fractures. *J Bone Miner Res*, (in press).

• 会讯 •

中华老年医学会代谢性骨病学组 关于召开第三届全国学术交流会的通知

自一九九三年十月第二次全国老年代谢性骨病和骨密度测量学术交流会以来,各地在这一领域中的基础与临床研究方面取得了不少新的进展和经验。为交流和研讨这一方面的学术成果,经中华老年医学会批准,代谢性骨病学组决定:1996年9月初在上海举行第三次全国老年代谢性骨病和骨密度测量学术交流会。兹将有关事项通知如下:

一、征文内容:

1. 原发性骨质疏松症(老年性及绝经后)的病理生理与防治的基础和临床研究;
2. 原发性骨质疏松症(老年及绝经后)的病因、病理、诊断和防治研究;
3. 双能X线骨密度仪在代谢性骨病诊治中的临床应用;
4. 老年代谢性骨病的动物模型研究;
5. 各种代谢骨病的诊治及保健的研究;

二、征文时间和注意事项:

1. 时间:征文截稿日期为1996年5月20日。
2. 论文要求:800字的摘要两份,A4纸打印。
3. 摘要按题目、作者、单位(地址、邮编)、前言、材料和方法、结果、结论顺序写。
4. 论文须在全国性杂志和学术会议上未录用过的。
5. 论文经单位审核,另附推荐介绍信,并加盖公章。
6. 论文寄上海市斜土路2094号,上海医科大学老年医学(基础部)研究中心,王洪复教授收,邮政编码200032。一欢迎您参加本次学术交流会议,为老年医学和保健事业作贡献!

中华医学会老年医学学会

代谢性骨病学组

一九九五年四月。