

# 心率变异性基本尺度熵的多尺度化研究\*

黄晓林<sup>1)</sup> 崔胜忠<sup>2)</sup> 宁新宝<sup>1)†</sup> 卞春华<sup>1)</sup>

1) 南京大学电子科学与工程系近代声学教育部重点实验室, 生物医学电子工程研究所, 南京 210093)

2) 南京医科大学生理学系, 南京 210029)

(2009 年 3 月 17 日收到, 2009 年 4 月 7 日收到修改稿)

将基本尺度熵的方法在时间上做多尺度化的扩展, 并将其应用到心跳间隔序列的分析研究中. 研究发现, 健康人的心率变异性是小时间尺度下的模式特定性与大时间尺度下的模式丰富性相结合的, 而充盈性心衰患者则正好相反. 这说明充盈性心衰患者在小时间尺度下心脏动力系统的控制不力, 导致随机性增加, 而在大时间尺度下对外界环境变化反应又不够丰富, 从而导致生命更容易受威胁. 据此提出了以小时间尺度下的基本尺度熵值相对于大时间尺度下平台区基本尺度熵值的变异参数  $\delta$  作为区分健康人和充盈性心衰患者的诊断依据. 通过对 72 例健康人和 44 例充盈性心衰患者的计算, 发现两组样本差异显著, 证明了参数  $\delta$  的有效性.

关键词: 多尺度化的基本尺度熵, 心跳间隔序列, 心率变异性, 充盈性心衰

PACC: 0547, 8770

## 1. 引 言

健康人的心跳是一个典型的多输入的非线性系统, 它在自发性节律的基础上同时受自主神经系统, 即交感神经和副交感神经等多种因素的协调控制, 因而呈现复杂的变异性, 即心率变异性(heart rate variability, HRV). HRV 研究既可以采用心跳间隔序列, 也可以采用心率波动序列. 本文中采用的是前者. 作为一种评价自主神经系统活动水平的无创性手段, HRV 在预测心源性猝死以及其他与自主神经控制功能密切相关的疾病(如糖尿病、癫痫、昏厥等)的诊断方面都有广泛的临床应用<sup>[1-3]</sup>, 另外在非治疗医学方面, 例如醉酒检测、运动医学等, HRV 也显示出其独特的优越性<sup>[4-6]</sup>. 因此, HRV 的研究一直是国内外众多学者研究的热点.

20 世纪 90 年代以来, 非线性的概念和研究方法被广泛应用到 HRV 的分析中, 并取得了一定的进展<sup>[7,8]</sup>. 其中比较有代表性的一类方法是熵分析法. 熵是热力学中一种度量混乱与随机程度的测度, 后来作为信息量或复杂性的一种测度, 熵也被广泛地应用到时间序列分析中. 心跳间隔序列, 作为确定性非线性系统产生的非随机信号, 仅用概率分布函数

描述是不够的. 间隔排列的先后顺序蕴含了大量的内在动力学特性, 熵测度则能够从一定程度上衡量这种时间前后的关联性. 常用的熵分析有近似熵(approximate entropy, AE)、样本熵(sample entropy, SE)、基本尺度熵(base-scale entropy, BE)等. 它们都因为对数据量要求不高且具备较好的抗噪能力而在心率波动或其他生理信号分析方面有比较成功的应用<sup>[6,9-15]</sup>. 其中, 李锦等提出的 BE 方法<sup>[12]</sup>因为算法简单快速, 适用于 500 点的短时数据, 能很好地排除数据的非平稳干扰, 而在 HRV 分析上获得了较成功的应用. 据报道, 该方法能够成功区分健康年轻人、健康老年人和心脏病患者, 而且冠心病患者、心肌梗死(myocardial infarction, MI)患者和充盈性心衰(congestive heart failure, CHF)患者的 BE 值都显著地高于健康人. 健康老年人相对于健康年轻人也呈现 BE 值升高的趋势<sup>[12-15]</sup>. 然而, 在对更多数据计算时, 作为一个临床诊断依据, BE 的区分能力并不最佳. 另外, 一般认为健康状态下的心脏动力系统应该比疾病状态下具有更高的复杂性. 因此熵作为复杂性的测度, 在健康人中应该表现得熵值更高. 这一点也与李锦的结论不一致. 其原因是 BE 算法本身只提取了数据的高频成分进行熵分析. 然而, 被 BE 算法滤除的低频波动是否真的完全来源于非平稳干扰

\* 国家自然科学基金(批准号:60701002)资助的课题.

† 通讯联系人. E-mail: xbning@nju.edu.cn

而与心脏固有的内在动力学特性无关呢? 已有研究表明, 作为生命信号的一种表现形式, 心跳间隔序列是一种典型的长时相关信号且具有多尺度性<sup>[16, 17]</sup>. 其低频成分尽管到目前为止还没有发现明确的生理对应, 但是可以肯定的是至少部分来源于心脏动力系统的长时相关成分. 因此, 要全面了解心脏动力系统的内在动力学本质, 短时的、单一的尺度分析或单一的节律表示都是不够的. Costa 等<sup>[18, 19]</sup>提出的多尺度熵 (multiscale entropy, MSE) 是应用得最广泛的多尺度方法之一, 它考察原始时间序列在不同时间尺度下的 SE, 以此来获得对原始序列复杂性的全面了解. 该方法在心率波动信号分析、基因代码分析以及人造心跳间隔序列和实际生理心跳间隔序列的区分、脑磁、动脉血压以及胎儿心电图信号分析等方面都获得了较成功的应用<sup>[19-23]</sup>. 但是, MSE 的方法计算长时数据时速度较慢, 计算效率不高, 另一方面, 对心跳间隔序列的 MSE 分析更多反映的是年龄上的差别, 而对 CHF 无法形成临床上的诊断依据<sup>[24, 25]</sup>.

为全面考察心跳间隔序列的复杂性, 我们将 BE 的方法在长时数据上扩展为多尺度化的基本尺度熵 (multiscale base-scale entropy, MBE), 发现尽管在小的时间尺度下 CHF 患者的 BE 值高于健康人, 但在大时间尺度下, CHF 患者的 BE 值却低于健康人. 这说明健康人的心率波动在小的时间尺度下集中在特定波动模式, 而在大的时间尺度下却呈现更丰富的波动模式, 由此才能对外界环境的各种变化作出尽可能精确的反应. 我们还发现, 在考察的时间尺度下, 尽管任何一个单一尺度下的 BE 都不能完全区分健康人和 CHF 患者, 但是 MBE 曲线的变异参数  $\delta$  却能对两类人作出很好的区分, 从而可以作为一个临床诊断的依据.

## 2. 基本尺度熵

基本尺度熵方法基于符号动力学理论, 将心跳间隔序列在幅度上符号化后求熵测度, 下面给出计算方法<sup>[12]</sup>.

对于数据长度为  $N$  的时间序列  $\{u(i) : 1 \leq i \leq N\}$  将其嵌入  $m$  维相空间,

$$X(i) = [u(i), u(i+L), \dots, u(i+(m-1)L)], \quad (1)$$

其中  $L$  是延迟时间. 取  $L=1$  时, 则共有  $N-m+1$  个  $m$  维矢量. 对于每一个  $m$  维矢量, 计算出基本尺

度  $\lambda$ ,  $\lambda$  定义为  $m$  维矢量中所有相邻点的差值均方根值, 即

$$\lambda(i) = \sqrt{\frac{\sum_{j=1}^{m-1} ((u(i+j) - u(i+j-1)))^2}{m-1}}. \quad (2)$$

根据基本尺度, 选择符号划分标准为  $\alpha\lambda$ , 然后把每一个  $m$  维矢量转换成  $m$  维矢量符号序列  $S_i(X(i)) = \{s(i), s(i+1), \dots, s(i+m-1)\}$ ,  $s \in \{0, 1, 2, 3\}$ . 具体的转换方法如下:

$$s(i+k) = \begin{cases} 0 & (\bar{u} < u_{i+k} \leq \bar{u} + \alpha\lambda) \\ 1 & (u_{i+k} > \bar{u} + \alpha\lambda) \\ 2 & (\bar{u} - \alpha\lambda < u_{i+k} \leq \bar{u}) \\ 3 & (u_{i+k} \leq \bar{u} - \alpha\lambda). \end{cases} \quad (3)$$

这里  $i=1, 2, 3, \dots, N-m+1$ ;  $k=0, 1, 2, \dots, m-1$ ;  $\bar{u}$  代表第  $i$  个  $m$  维矢量中各分量的平均值;  $\lambda$  则代表第  $i$  个  $m$  维矢量的基本尺度; 符号  $0, 1, 2, 3$  仅仅作为每一个划分区域的记号, 具体数值的大小并没有实际的意义;  $\alpha$  是符号化的尺度因子,  $\alpha$  取值过大会导致细节信息的丢失, 不能很好捕捉到信号中的动态信息,  $\alpha$  取值过小会导致对噪声过于敏感.

为了计算 BE 值, 先统计  $m$  维矢量符号序列  $S_i$  的分布概率  $P(S_i)$ . 包含  $0, 1, 2, 3$  四种符号的  $m$  维矢量符号序列  $S_i$  共有  $4^m$  种不同的组合形式. 统计每一种不同的组合形式在所有  $N-m+1$  个  $m$  维矢量中所占的概率为

$$p(l) = \frac{\alpha(l)}{N-m+1} \quad (4)$$

其中  $\alpha(l)$  表示序列第  $l$  种组合形式的个数,  $1 \leq l \leq 4^m$ . 则  $m$  维矢量的 BE 值定义为

$$E_b = - \sum p(l) \ln p(l). \quad (5)$$

为了计算方便,  $m$  的取值一般为  $m=3, 4, 5, 6, 7, N$  的取值一般只要大于  $4^m$  即可. (5) 式中对所有可能组合形式的分布概率求信息熵.  $E_b$  描述了时间序列中  $m$  个连续值所包含的信息. 熵值越高, 说明序列蕴含的信息量越大, 随机性也越大, 而可预测性越小. 熵值越低, 则说明序列蕴含的信息量越小, 序列越规则并接近确定性信号, 从而也更容易预测.

从 BE 的计算可知, 符号化的依据与时间序列的实时均值和相邻数据增量的均方根相关, 使得相似波形有相同的编码模式, 如图 1 所示. 因此该方法相当于对数据做了一次高通滤波, 仅提取原始信号的高频波动特征而忽略掉幅度特性和低频基底信号 (图 1 中的细线) 的影响, 故具有很好的抗非平稳性

能力.

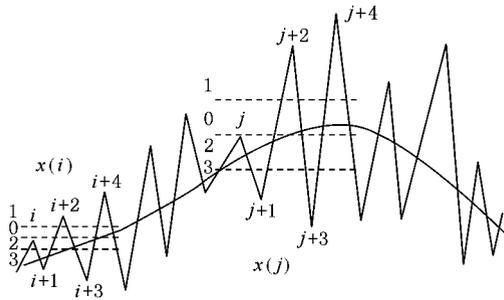


图 1 相似波动模式示意图

李锦等<sup>[12-15]</sup>将上述的 BE 方法应用到 HRV 的分析,发现冠心病患者、MI 患者和 CHF 患者的  $E_b$  都显著地高于健康人,健康老年人相对于健康年轻人也呈现  $E_b$  的升高趋势.这说明在适当的符号化以后,健康人的心率波动信号更集中在特定的波动模式.一般认为,健康状态下的心脏动力系统应该比疾病状态下具有更高的复杂性,因此作为复杂性的测度,熵值在健康人中应该表现得更高.我们认为,上述这种不一致性的产生是由于 BE 的方法在符号化过程中将低频波动成分当作非平稳干扰完全滤除,因而其结果主要是反映的是信号高频成分的复杂性测度.以往研究表明,健康人的心率波动中存在  $1/f$  成分<sup>[26]</sup>是一种典型的长时相关信号<sup>[16]</sup>,其中的低频成分不完全是非平稳干扰,而至少部分地来源于心脏动力系统自身的低频节律.只在一个时间尺度下考察短时 HRV 的 BE 不能获得全面的认识.由此,我们将 BE 的方法在时间域上也做了多尺度化的展开,并用来对长时 HRV 数据进行分析,以期获得对心脏动力系统更全面的认识.

### 3. 多尺度粗粒化

仿照 Costa 等<sup>[18]</sup>在 MSE 中的做法,我们也对原始的心跳间隔序列做了时域上的粗粒化处理,下面给出具体的方法.

给定一维离散序列  $\{x_1, \dots, x_i, \dots, x_N\}$ ,对于给定的时间尺度因子  $\tau$ ,原始序列按下式粗粒化<sup>[17]</sup>:

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{(j-1)\tau+1}^j x_i \quad (1 \leq j \leq N/\tau). \quad (6)$$

粗粒化的本质是对长时数据做低通滤波.不同的时间尺度,则代表了不同的低通截止频率.然后,对于不同的  $\tau$ ,计算序列  $\{y_j^{(\tau)}\}$  的  $E_b$ ,得到  $E_b$  随尺度  $\tau$

的变化曲线,即实现了 MBE 分析.

### 4. 心跳间隔序列的 MBE 分析

我们研究的数据来自 MIT-BIH (Boston's Beth Israel Hospital) 数据库中的 chf 和 chf2 数据库,以及 nsr 和 nsr2.其中包括 46 例健康老年人(平均年龄为  $65.87 \pm 3.97$  岁),26 例健康年轻人(平均年龄为  $34.67 \pm 7.37$  岁),44 例 CHF 患者(平均年龄为  $55.59 \pm 11.31$  岁).计算中根据他人的经验选取的参数为  $m=4$ ,符号化尺度因子  $\alpha=0.1$ ,原始序列总长  $N=20000$  时间尺度  $\tau$  从 1 到 20 变化.计算结果如图 2 所示.为使图示更清晰,我们在图 2 中只画出了上述三类数据的平均值结果.

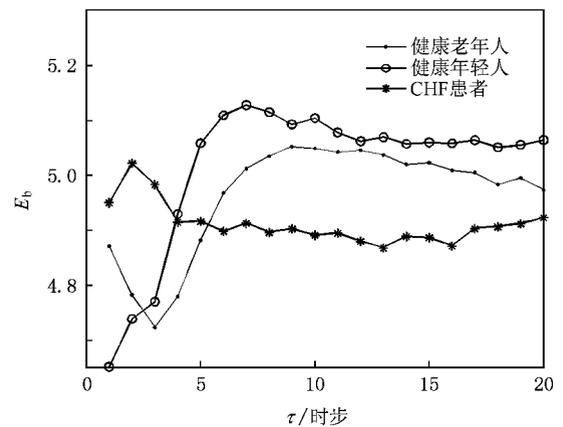


图 2 心跳间隔序列的 MBE 分析

从图 2 可以看到,在尺度  $\tau=1$  时,三组数据的  $E_b$  与文献 12—15 所得到的结果一致,以健康年轻人为最低,健康老年人有所升高,而 CHF 病人则最高.但是,随着时间尺度的增加,健康年轻人的  $E_b$  会逐渐增加,直到  $\tau \geq 7$  时进入平台区.而健康老年人则是先降低,当  $\tau=3$  时降到最低值,然后开始回升,也在  $\tau \geq 7$  以后基本进入平台区.至于 CHF 患者则以  $\tau=2$  时的  $E_b$  为最高,之后减小,在  $\tau \geq 4$  时进入平台区.而三组样本的平台区对应的  $E_b$ ,以健康年轻人为最高,健康老年人次之,CHF 患者为最低,与尺度  $\tau=1$  下的情况刚好相反.这说明健康的心跳间隔序列尽管在小的时间尺度下更集中于特定的波动模式,但是在大的时间尺度( $\tau \geq 7$ )下却具有更丰富的波动模式.因此,健康年轻人的心率变异性是小尺度时的规则性加上大尺度时的不确定性,从而能适应外界环境的不同时间尺度的复杂变化.

在我们考察的样本中,单独用某个特定尺度下的  $E_b$  并不能将三类数据很好的区分,例如图 3(a), (b)(c)(d)分别给出了在  $t$  检验下  $p$  值最小的四

个尺度 1, 2, 3, 15 下的  $E_b$  在三类样本中的分布情况.由此可以看到,对于健康人和 CHF 患者的区分,效果都不是特别理想.

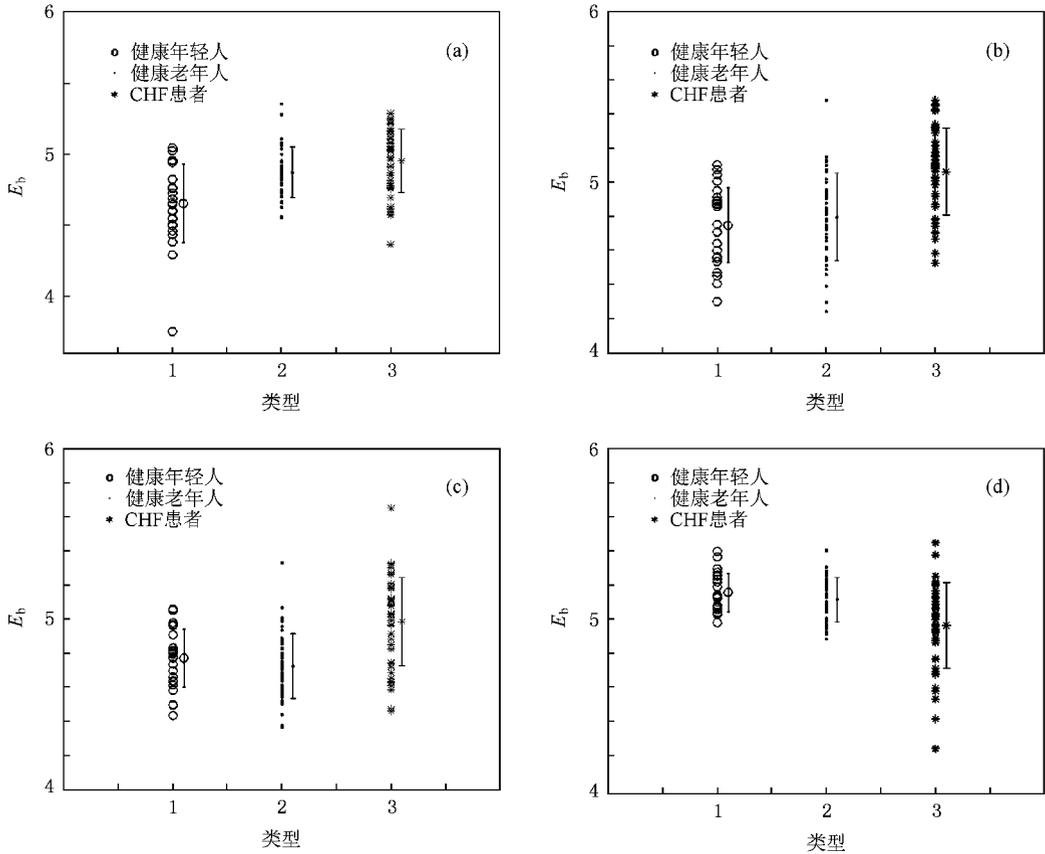


图 3 不同尺度下的 BE 在三类样本的取值分布 (a)  $\tau=1$ , (b)  $\tau=2$ , (c)  $\tau=3$ , (d)  $\tau=15$

从图 2 可以看到,相对于平台区,健康年轻人的 BE 变化范围最大,健康老年人次之,而 CHF 患者则最小,而且从小时间尺度变化到大时间尺度时,健康人的 BE 动态范围在平台区的下方,而 CHF 患者 BE

的动态范围则在平台区的上方.我们定义了 MBE 的变异性  $\delta$  来度量平台区 ( $\tau \geq 10$  的 BE 平均值)与小尺度 ( $\tau < 7$ ) 段的最大差异值.若平台区高,则  $\delta$  为正,若平台区低,则  $\delta$  为负.该参数实际上对 MBE 以

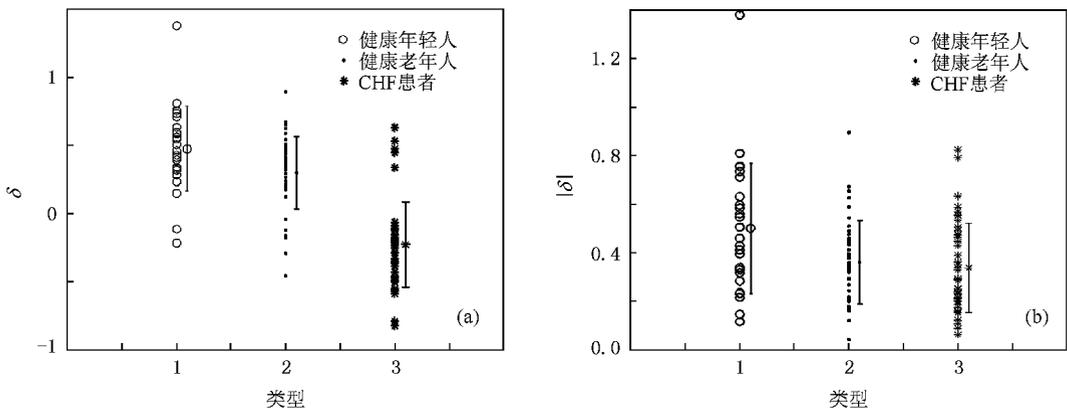


图 4 参数  $\delta$  和  $|\delta|$  在三组样本上的分布 (a)  $\delta$ , (b)  $|\delta|$

平台区为参照的动态范围及方向给出了定量的描述. 针对三类样本数据, 我们分别计算了 MBE 的  $\delta$ , 结果在图 4(a) 中给出.

从图 4(a) 可以看到, 总体上, 健康人的  $\delta$  为正, 而 CHF 患者则以负值为主导, CHF 患者组能明显地区别于健康年轻人和健康老年人(  $t$  检验的  $p$  值为  $10^{-13}$  ). 因此该参数用于区分健康人和 CHF 患者的效果比较理想. 我们还在图 4(b) 中给出了  $|\delta|$  在三组样本上的结果. 从图 4(b) 可以看到, CHF 患者的  $|\delta|$  较健康年轻人有较显著的降低(  $p = 0.003$  ), 说明了健康年轻人 MBE 的动态范围大于 CHF 患者, 但是健康老年人与 CHF 患者之间差异不显著(  $p = 0.5$  ). 故作为区分健康与 CHF 的诊断参数, 考虑了变异方向的  $\delta$  更合适, 即 MBE 中小尺度相对于平台区的变异方向更能理想地区分两类样本, 因此作为诊断 CHF 疾病的参考阈值,  $\delta$  可以取为零.

相比另一个常用的多尺度分析方法 MSE, MBE 的方法在计算速度上大大超过 MSE 方法, 同样对于 20000 点的时间序列, 做时间尺度从 1 到 20 的 MSE 分析时, 在 CPU 主频 2.5 GHz 的计算机上, 平均要消耗 70 min, 而 MBE 则大约 30 s 即可完成.

## 5. 结 论

通过对心跳间隔序列进行 MBE 研究发现, 当时间尺度变化时, 心跳间隔序列的 BE 值在各类人群中的取值情况都会有所变化. 在尺度 1 下, 以 CHF

患者的 BE 值最高, 健康老年人居中, 健康年轻人最低. 随着尺度的增加, 健康年轻人的 BE 值会呈现上升的趋势, 直到进入平台区, 健康老年人则出现先下降后上升的趋势, 最后也进入平台区, CHF 患者则是先上升后下降, 并迅速进入平台区. 而大尺度下的平台区 BE 值在三组样本上的分布, 则刚好与尺度 1 下的情况相反, 以健康年轻人为最高, 健康老年人次之, CHF 患者最低. 该结果说明健康年轻人的心率变异性是小时间尺度下的特定性与大时间尺度下的丰富性相结合的, 而 CHF 患者的心率变异性则是在小尺度下显现出更多的不确定性, 这可能是内在心脏动力系统的短程控制不力造成( 致使系统随机性增加, 熵值增加 ), 大尺度下又缺乏足够的应变性, 因此, 总体上显现出对外界环境的适应性差, 因而生命安全更容易受到威胁.

由于单一尺度下的 BE 不能很好区分考察的三类样本, 我们提出了 MBE 的变异性参数  $\delta$  及  $|\delta|$  来衡量心脏的健康状态. 结果发现, 尽管健康年轻人的  $|\delta|$  较显著地大于 CHF 患者, 即健康年轻人的 MBE 变异范围大于 CHF 患者, 但健康老年人与 CHF 患者之间差异不显著, 而  $\delta$  则能很好地区分健康人和 CHF 患者. 这说明在心率波动信号的 MBE 分析中, 小尺度区相对平台区的变异方向是健康人与 CHF 患者最大的差别.

由于数据有限, 本文只分析了 CHF 一种疾病, 在样本库中增加疾病的种类, 考察 MBE 在不同疾病之间的共性与差别, 是我们下一步要深入的方向.

- [ 1 ] Kleiger R E, Miller J P, Bigger J T, Moss A J 1987 *Am. J. Cardiol.* **59** 256
- [ 2 ] Huikuri H V, Makikallio T H, Airaksinen K E J, Seppanen T, Puukka P, Raiha I J, Sourander L B 1998 *Circulation* **97** 2031
- [ 3 ] Stein P K, Kleiger R E 1999 *Annu. Rev. Med.* **50** 249
- [ 4 ] Evrengül H, Tanriverdi H, Dursunoglu D, Kaftan A, Kuru O, Unlu U, Kilic M 2005 *Epilepsy Res.* **63** 131
- [ 5 ] Weise F, Krell D, Brinkhoff N 1986 *Drug Alcohol Depend.* **17** 89
- [ 6 ] Zhuang J J, Ning X B, Zou M, Sun B, Yang X 2008 *Acta Phys. Sin.* **57** 2805 ( in Chinese ) [ 庄建军、宁新宝、邹 鸣、孙 飙、杨 希 2008 物理学报 **57** 2805 ]
- [ 7 ] Ning X B, Bian C H, Wang J, Chen Y 2006 *Chin. Sci. Bull.* **51** 385
- [ 8 ] Bian C H, Ning X B 2004 *Chin. Phys.* **13** 633
- [ 9 ] Pincus S M 1991 *Proc. Natl. Acad. Sci. ( USA )* **88** 2297
- [ 10 ] Cao B, Xiang Y P, Lu X Q, Zeng M, Huang S S 2008 *Chin. Phys. B* **17** 865
- [ 11 ] Richman J S, Moorman J R 2000 *Am. J. Physiol. Heart Circ. Physiol.* **278** H2039
- [ 12 ] Li J, Ning X B, Wu W, Ma X F 2005 *Chin. Phys.* **14** 2428
- [ 13 ] Li J, Ning X B 2006 *Phys. Rev. E* **73** 052902
- [ 14 ] Li J, Ning X B 2005 *Chin. Sci. Bull.* **50** 1269
- [ 15 ] Li J, Ning X B 2007 *Physica A* **384** 423
- [ 16 ] Ivanov P C, Chen Z, Hu K, Stanley H E 2004 *Physica A* **344** 685
- [ 17 ] Yang X D, Ning X B, He A J, Du S D 2008 *Acta Phys. Sin.* **57** 1514 ( in Chinese ) [ 杨小冬、宁新宝、何爱军、都思丹 2008 物理学报 **57** 1514 ]
- [ 18 ] Costa M, Goldberger A L, Peng C K 2002 *Phys. Rev. Lett.* **89** 068102
- [ 19 ] Costa M, Goldberger A L, Peng C K 2005 *Phys. Rev. E* **71** 021906

- [ 20 ] Costa M , Goldberger A L , Peng C K 2002 *Comp. Card.* **29** 137 *Trans. Biomed. Eng.* **53** 119
- [ 21 ] Wang J , Ma Q L 2008 *Chin. Phys. B* **17** 4424 [ 24 ] Wessel N , Schirdewan A , Kurths J 2003 *Phys. Rev. Lett.* **91** 119801
- [ 22 ] Angelini L , Maestri R , Marinazzo D , Nitti L , Pellicoro M , Pinna G D , Stramaglia S , Tupputi S A 2007 *Artif. Intell. Med.* **41** 237 [ 25 ] Thuraisingham R A , Gottwald G A 2006 *Physica A* **366** 323
- [ 23 ] Ferrario M , Signorini M G , Magenes G , Cerutti S 2006 *IEEE* [ 26 ] Kobayashi M , Musha T 1982 *IEEE Trans. Biomed. Eng.* **29** 456

## Multiscale base-scale entropy analysis of heart rate variability<sup>\*</sup>

Huang Xiao-Lin<sup>1)</sup> Cui Sheng-Zhong<sup>2)</sup> Ning Xin-Bao<sup>1)†</sup> Bian Chun-Hua<sup>1)</sup>

<sup>1</sup> *Key Laboratory of Modern Acoustics of Ministry of Education , Institute of Biomedical Electronic Engineering , Department of Electronic Science and Engineering , Nanjing University , Nanjing 210093 , China*

<sup>2</sup> *Department of Physiology , Nanjing Medical University , Nanjing 210093 , China*

( Received 17 March 2009 ; revised manuscript received 7 April 2009 )

### Abstract

Multiscale base-scale entropy ( MBE ) is used to analyze heart beat interval series of both congestive heart failure ( CHF ) sufferers and the healthy people . The results indicate that , in contrast to the higher base-scale entropy ( BE ) value in short time scale , the data of CHF sufferers have actually lower BE value compared with those of the healthy people in long time scale . This supports the view that CHF patients ' heart beats are more random in short time scale due to impaired cardiac dynamic system , but are not versatile enough in long time scale to adapt to external impacts . More importantly , we found that  $\delta$  , a parameter in MBE analysis , can be used to distinguish effectively the CHF patients from the healthy , which is proved in analysis of a database containing 72 healthy people and 44 CHF sufferers .

**Keywords** : multiscale base-scale entropy , heart beat interval series , heart rate variability , congestive heart failure

**PACC** : 0547 , 8770

<sup>\*</sup> Project supported by the National Natural Science Foundation of China ( Grant No. 60701002 ) .

<sup>†</sup> Corresponding author . E-mail : xbning@nju.edu.cn