

DOI:10.3969/j.issn.1671-024x.2025.04.008

基于心率变异性的水平运动员睡眠阶段分期分析

常宁^{1,2},肖云瀚²,万帅²,胡鑫²,张少帆²,崔玉婷³,耿家先⁴

(1. 西安交通大学 生命科学与技术学院, 西安 710049; 2. 西安体育学院, 西安 710068; 3. 宝鸡文理学院, 陕西宝鸡 721013; 4. 湖州师范学院 体育学院, 浙江 湖州 313000)

摘要: 为了评估基于心率变异性(HRV)的可穿戴设备在高水平运动员睡眠监测中的可行性,设计 Firstbeat 监控系统与多导睡眠监测(PSG)的对比实验。选取 30 名高水平运动员连续 2 晚同步佩戴 Firstbeat Bodyguard 3 和 PSG 记录仪,收集 60 晚睡眠数据,比较清醒期、浅睡眠、慢波睡眠(SWS)和快速眼动睡眠(REM)的分期结果,比较 2 种方法在睡眠阶段分期中的差异。结果表明:Firstbeat 与 PSG 在所有睡眠阶段时长监测中均无显著差异($p > 0.05$);在清醒期特异性达 0.87,敏感性 0.93,准确性 0.90;SWS 阶段特异性 0.91,敏感性 0.78,准确性 0.83;REM 阶段特异性 0.8,敏感性 0.6,准确性 0.84;混淆矩阵显示各阶段准确率分别为清醒 76.6%、浅睡眠 67.2%、SWS 70.6%、REM 58.4%;说明该系统可有效监测清醒期和 SWS,为运动员睡眠评估提供了可行的替代方案。

关键词: 心率变异性;睡眠阶段分期;Firstbeat 监控系统;高水平运动员

中图分类号: TP391;G804.49 文献标志码: A 文章编号: 1671-024X(2025)04-0052-08

Analysis of sleep stage classification in elite athletes based on HRV

CHANG Ning^{1,2}, XIAO Yunhan², WAN Shuai², HU Xin², ZHANG Shaofan², CUI Yuting³,
GENG Jiaxian⁴

(1. School of Life Science and Technology, Xi'an Jiaotong University, Xi'an 710049, China; 2. Xi'an Physical Education University, Xi'an 710068, China; 3. Baoji University of Arts and Sciences, Baoji 721013, Shaanxi Province, China; 4. College of Physical Education, Huzhou University, Huzhou 313000, Zhejiang Province, China)

Abstract: To evaluate the feasibility of heart rate variability (HRV)-based wearable devices for sleep monitoring in elite athletes, a comparative study was conducted between the Firstbeat monitoring system and polysomnography (PSG). Thirty elite athletes underwent simultaneous monitoring using Firstbeat Bodyguard 3 and PSG recorders during two consecutive nights, yielding 60 nights of synchronized sleep data. Sleep stages including wakefulness, light sleep, slow wave sleep (SWS), and rapid eye movement (REM) sleep were systematically compared. The results showed no significant differences in the duration of all sleep stages measured by Firstbeat and PSG ($p > 0.05$). For wakefulness, specificity was 0.87, sensitivity was 0.93, and accuracy was 0.90. For SWS, specificity was 0.91, sensitivity was 0.78, and accuracy was 0.83. For REM sleep, specificity was 0.8, sensitivity was 0.6, and accuracy was 0.84. The confusion matrix revealed accuracy rates of 76.6% for wakefulness, 67.2% for light sleep, 70.6% for SWS, and 58.4% for REM sleep. These findings indicate that the Firstbeat system can effectively monitor wakefulness and SWS, providing a feasible alternative for sleep assessment in elite athletes.

Key words: heart rate variability (HRV); sleep stage classification; Firstbeat monitoring system; elite athletes

在竞技体育领域,充足且高质量的睡眠不仅是运动员恢复体能、预防伤病的重要保障,更是维持最佳

竞技状态的关键因素。睡眠过程中,机体通过复杂的生理调节实现能量恢复、组织修复及神经调控,对运

收稿日期: 2025-02-13

基金项目: 陕西省体育局科技攻关项目(203072202);浙江省哲学社会科学规划课题(22NDJC157YB);西安体育学院教育教学改革研究重点项目资助(XTJY2401)

第一作者: 常宁(1985—),男,博士研究生,讲师,主要研究方向为柔性可穿戴系统在竞技运动中的应用。E-mail:111028@tea.xaipe.edu.cn

通信作者: 耿家先(1984—),男,博士,副教授,主要研究方向为竞技运动员运动表现提升。E-mail:wylingzhi@126.com

运动员的竞技表现与心理健康具有深远影响^[1]。睡眠周期在整个夜间睡眠期间交替进行,健康成年人的睡眠周期在一个晚上大约交替4到5次,互相连接,周而复始^[2]。根据美国睡眠医学学会(American Academy of Sleep Medicine, AASM)的睡眠及其相关判断手册的分期标准^[3],睡眠周期可分为清醒期(wake)、非快速眼动期(non-rapid eye movement, NREM)和快速眼动期(rapid eye movement, REM),其中NREM期又分为N1、N2和N3期。这种不同阶段之间的变化与生理指标的变化相关,如肌肉张力^[4]、血压^[5]、温度调节^[6]、心率和心率变异性(heart rate variability, HRV)^[7]。在NREM阶段,深度睡眠(N3)(slow wave sleep, SWS)和较轻睡眠阶段(N1和N2)(lighter sleep stages, LSS)之间也有一些生理差异,但这些差异不如REM和NREM阶段^[8]之间的差异明显。具体来说,睡眠越深,副交感神经调节就越强^[9]。

尽管多导睡眠监测(PSG)提供了可靠的睡眠数据,被认为是测量睡眠阶段的金标准^[9],是区分不同睡眠阶段的唯一方法,但随着睡眠障碍诊断的日益增多,由于其成本高、操作复杂、受试者舒适度差且耗时耗力,很难满足诊断需求^[10]。因此,开发一种经济高效、易操作且可靠的睡眠研究工具有重要的临床意义和实际价值^[11]。

近年来,随着可穿戴技术的迅速发展,HRV作为一种反映自主神经系统活动的非侵入性生物标志物,不仅在评估压力反应、应激恢复和运动能力方面得到了广泛应用^[12-14],而且其变化与睡眠阶段转换密切相关,能够反映睡眠过程中不同阶段的生理特征^[7]。基于这一认识,Firstbeat科技公司开发了基于HRV与加速度计数据相结合的睡眠分析设备——Firstbeat Bodyguard 3,其通过单导联心电信号精准测量HRV,同时采集运动数据,为用户提供个性化的睡眠数量与质量反馈,并在一定程度上检测心律异常^[15]。这一方法不仅有望突破传统PSG在实际应用中的限制,也为高水平运动员的常规睡眠监测提供了一种经济高效、便捷实用的替代方案。

本研究聚焦于评估基于HRV和加速度计的Firstbeat睡眠分析方法与PSG之间的相关性。通过对比2种方法在敏感性、特异性和准确性等方面的表现,探讨Firstbeat方法在细化睡眠阶段分期中的可行性与局限性,进而为其在竞技体育场景以及临床睡眠监测中的应用提供理论和实证依据,推动可穿戴设备在运动员睡眠健康管理中的应用,并为未来整合更多生理信号提供新的思路和方法。

1 睡眠监测试验

1.1 试验参与者选择

本课题组于2024年4月至6月在西安体育学院招募30名从事专业体育训练的高水平运动员作为参与者,参与者均系国家二级以上运动员。这项研究得到了西安体育学院伦理委员会的批准,所有程序都遵循《赫尔辛基宣言》及其后来的修正案,符合《涉及人的生命科学和医学研究伦理审查办法》要求,所有参与者均签署了书面知情同意书。

入选标准包含:①年龄18~37周岁;②现役注册运动员且每周专项训练 ≥ 20 h;③身体健康,无睡眠障碍或神经方面病史,没有服用任何药物;④近3个月内未参与其他干预性研究;⑤睡眠地点和时长相对稳定(通常在宿舍固定床位23:00至次日07:00之间);⑥自愿参加本试验,并签署知情同意书。

排除情形包括:①不同意参加者;②任何诊断出的睡眠障碍,存在神经系统器质性疾病(如癫痫、脑外伤后遗症等);③使用任何可能影响睡眠的药物(抗抑郁药、苯二氮卓类药物等);④合并急慢性疾病(甲状腺功能异常、未控制的高血压等);⑤难以随访者;⑥妊娠或哺乳期女性。

出现以下情况予以剔除:①依从性不足(干预措施执行率 $< 80\%$);②主动要求退出研究;③失访(连续3次未按计划完成评估);④其他原因未完成整个试验。

中止试验的情形包括:①出现不良反应(事件);②受试者出现意外伤害病需住院治疗;③研究期间被选拔至国家队集训或外训;④受试者个人原因申请退出;⑤其他需要中止试验的情况。

最终,本研究共招募到30名高水平运动员,参与者的一般资料如表1所示。

表1 参与者一般资料

Tab.1 General characteristics of participants

参与者参数	参数	数值
人数	男	15
	女	15
年龄/岁	最大	37
	最小	18
	$X \pm s$	22.5 \pm 4.5
BMI/(kg·m ⁻²)	最大	27
	最小	12
	$X \pm s$	21.6 \pm 2.8
PSQI评分 ^[16]	最大	14
	最小	4
	$X \pm s$	7.4 \pm 2.2

注: BMI为身体质量指数; PSQI为匹兹堡睡眠质量指数量表。

参与者中男、女各 15 名,性别分布均衡;年龄范围为 18~37 岁;BMI 指数为 12~27(21.6 ± 2.8);匹兹堡睡眠质量指数(Pittsburgh sleep quality index,PSQI)量表^[16]评分为 4~14(7.4 ± 2.2)分,样本整体的睡眠质量差异适中,无脱落和剔除。

1.2 试验过程

参与者被要求连续 2 个晚上分别进行 PSG 和基于 HRV 的 Firstbeat 监测,并且在整个测试期间,严格控制潜在干扰因素:要求参与者在测试当日下午 16 点后避免摄入含酒精或咖啡因的饮品。同时,参与者被要求像往常一样正常睡眠,遵循其日常作息规律,睡前不进行剧烈的活动。此外,参与者被要求睡前将手机及其他带有信号发射功能的电子设备置于距离床铺至少 2 m 的位置,以确保监测设备不受外部信号干扰。

根据参与者的个人睡眠时间安排,研究人员在参与者睡眠之前上门,为参与者佩戴测量电极并启动监测设备。每位参与者均被鼓励按照其平常的睡眠习惯进行自然睡眠。次日清晨,研究人员在参与者醒来约 30 min 内抵达现场结束监测并回收设备。

所有研究操作均由接受过专业培训的团队完成,团队成员均为具备硕士及以上学历且拥有中级及以上职称的实验师,以确保数据采集的规范性和可靠性。2 种监测设备的实物和测试示意图 1 所示。

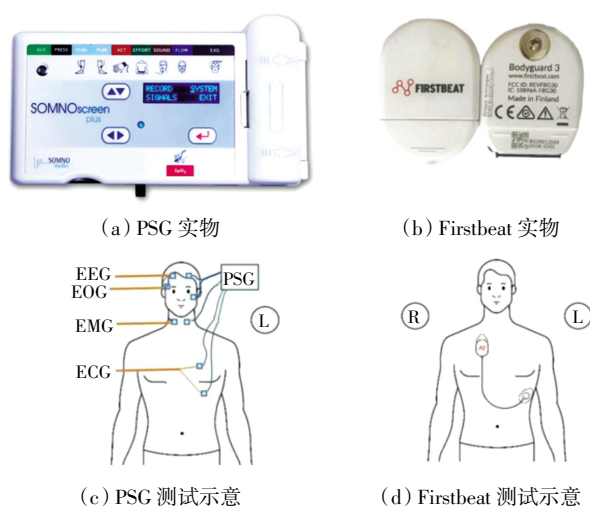


图 1 2 种监测设备的实物和测试示意

Fig.1 Physical photographs and test schematic diagrams of two monitoring devices

1.2.1 PSG 监测

本研究使用德国索姆诺医疗技术有限公司(SOMNOmedics GmbH)研制的便携式多导睡眠监测仪(SOMNOscreen plus PSG+)进行睡眠监测。记录参数包括:左、右脑电图(electroencephalogram,EEG);左、

右眼电图(electro-oculogram,EOG);左、右肌电图(electromyogram,EMG);心电图(electrocardiogram,ECG)。PSG 的设置与移除由经验丰富的实验师进行,以确保监测过程的规范性与数据的准确性。

EEG 信号的采集遵循标准化的 10-20 系统电极放置方案,共记录 8 通道 EEG 测量数据,采样率为 256 Hz。其中,F3、F4、C3、C4、O1、O2 为 6 个脑电图位置;A1、A2 为 2 个通道。使用一次性粘接电极测量 ECG、EOG 和 EMG,其中 2 个位置用于 ECG 和 EOG,3 个位置为 EMG。所有信号均参考 A1 和 A2 电极的平均信号,并且在 0.5~40 Hz 的通频带范围内进行滤波,以减少噪声并保留关键生理信息。该监测方案通过严谨的参数设置与标准化操作,确保了 PSG 数据的高质量采集与后续分析的可靠性。

PSG 数据的睡眠阶段评分由 2 名经验丰富的实验师使用 DOMINO program 2.7(SOMNOmedics GmbH,德国)进行手动标注。评分依据为 AASM 发布的规则^[3],以 30 s 为单位时间序列进行逐段分析。2 名实验师独立检查数据,并在每个时间段达成一致后完成评分。如果存在分歧,则暂停评分流程,双方共同仔细审阅数据,直至达成共识后继续评分,以确保评分的准确性与一致性。

1.2.2 Firstbeat 监测

本研究采用 Firstbeat Bodyguard 3 进行 HRV 的监测。该设备包括 2 个胸部电极,用于采集心电信号。Firstbeat 睡眠分析方法使用一种基于神经网络的算法,结合 HRV 数据、HRV 推导的呼吸频率以及加速度计数据,评估受试者的清醒状态或睡眠阶段。算法将睡眠分为清醒(wake)、轻度睡眠(light,涵盖 N1 和 N2)、深度睡眠(N3)和快速眼动睡眠(REM)。

为统一数据处理方式,本文将 PSG 测量中的睡眠分期(wake-N1-N2-N3-REM)进行相应整合,将 PSG 的 N1 和 N2 阶段合并为 Firstbeat 定义的轻度睡眠(light)。这种分期方式确保了 2 种方法之间的可比性,为后续数据分析提供了科学依据。

Firstbeat 监测采用简便的佩戴方式和算法分析,为睡眠分期提供了非侵入性的高效解决方案,其结合多种生理参数的设计有助于提高睡眠状态评估的准确性。

1.3 统计分析方法

为评估 Firstbeat 监测系统监测实际睡眠开始时间的能力,本文通过计算 PSG 与 Firstbeat 间的时间差异并使用配对样本 *t* 检验进行统计分析。PSG 与 Firstbeat 的睡眠分期比较均以 PSG 确定的实际睡眠开

始时间为基准,确保比较的起始点一致。

(1) 通过配对样本 t 检验, 比较了 2 种方法在睡眠开始、清醒、浅睡眠、深度睡眠(SWS)和快速眼动(REM)睡眠阶段时间上的差异。这种分析方法能够量化 2 种方法之间的系统性偏差, 并验证其统计显著性。

(2) 逐段计算 Firstbeat 与 PSG 在监测睡眠和清醒阶段的敏感性(监测实际睡眠的能力)、特异性(监测实际清醒的能力)以及准确性(同时监测睡眠与清醒的综合能力)。以 PSG 为“金标准”, 对比 Firstbeat 在各阶段(清醒、浅睡眠、SWS、REM)中的真阳性(TP)、真阴性(TN)、假阳性(FP)和假阴性(FN)数值。将 PSG 清醒阶段时长的均值作为“阈值”, 当 PSG 和 Firstbeat 时长均 \geq 阈值, 统计其个数为 TP; 当 PSG \geq 阈值, 而 Firstbeat $<$ 阈值, 统计其个数为 FN; 当 PSG $<$ 阈值, 而 Firstbeat \geq 阈值, 统计其个数为 FP; 当 PSG 和 Firstbeat 均 $<$ 阈值, 统计其个数为 TN。这些指标通过标准公式计算, 敏感性 = $TP/(TP + FN)$, 特异性 = $TN/(TN + FP)$, 准确性 = $(TP + TN)/(TP + TN + FP + FN)$, 通过 Bootstrap 法检验, 并结合混淆矩阵进行验证。

(3) 构建混淆矩阵, 对比 PSG 和 Firstbeat 在各阶段(清醒、浅睡眠、SWS、REM)中的真阳性、真阴性、假阳性和假阴性数值。当 PSG 和 Firstbeat 均将某 30 s 时段判定为睡眠时, 记为真阳性; 均判定为清醒时, 记为真阴性; 若 PSG 判定为睡眠而 Firstbeat 判定为清醒, 记为假阳性; 反之, 记为假阴性。混淆矩阵提供了清晰的阶段性分类结果, 并量化了 2 种方法在睡眠分期上的一致性。

(4) 为直观展示 PSG 与 Firstbeat 在各睡眠阶段的评分差异, 采用 Bland-Altman 分析法绘制散点图, 分别针对清醒、浅睡眠、SWS 和 REM 睡眠阶段进行评估。通过散点图展示 PSG 和 Firstbeat 测量值之间的差异分布, 并标明有多少观测值落在 30 min 窗口期的范围内, 从而评估 2 种方法的一致性及其偏差分布。

上述多层次的统计分析方法, 结合差异检验、敏感性评估、混淆矩阵和可视化图表, 为全面评价 Firstbeat 与 PSG 在睡眠阶段分期中的一致性和准确性提供了科学依据。

2 PSG 和 Firstbeat 睡眠分期结果的比较

2.1 各阶段时长及其占比比较

本研究基于 30 名参与者的 60 个夜间睡眠数据, 分析了各睡眠阶段的时长及其在总睡眠中的占比, 结果如表 2 所示。

表 2 所有参与者睡眠各阶段时长及占比

Tab.2 Sleep stage duration and proportions among all participants

指标	时长/min	在总睡眠中占比/%
睡眠总时长	425±46	100
清醒阶段	20.49±2.22	4.82±1.79
浅睡眠阶段	199.67±21.61	46.98±8.46
SWS 阶段	96.39±10.43	22.68±5.37
REM 阶段	108.45±12.24	25.52±6.13

注:入睡时间(hh:mm)为 23:05±1:24;清醒时间(hh:mm)为 7:15±0:16

由表 2 可知:参与者的睡眠时间安排整体较为规律;浅睡眠阶段占主导地位,与正常生理睡眠结构一致;SWS 和 REM 的时长和占比符合健康成年人的预期分布。这些特征为后续分析 Firstbeat 与 PSG 在不同睡眠阶段监测中的表现提供了重要参考依据。

2.2 配对 t 检验

表 3 显示了 Firstbeat 和 PSG 在各睡眠阶段时长之间的配对 t 检验结果及平均差异。

表 3 Firstbeat 和 PSG 睡眠各阶段时长之间的配对 t 检验比较和平均差异

Tab.3 Paired t -test comparison and mean differences in sleep stage durations between Firstbeat and PSG

各阶段	SD	SE	95%置信区间/min	t	p
清醒	5.351	0.98	-2.050~1.946	-0.053	0.958
浅睡眠	44.647	8.15	-20.577~12.767	-0.479	0.635
SWS	15.454	2.82	-8.256~3.285	-0.881	0.386
REM	24.158	4.41	-6.295~11.747	0.618	0.541

由表 3 可知:

(1) 清醒阶段, 2 种方法的平均时长差异无统计学显著性($t = -0.053$, $p = 0.958$), 95%置信区间为 $-2.050 \sim 1.946$ min, 这表明 Firstbeat 在清醒阶段监测方面表现出与 PSG 高度一致的能力。

(2) 浅睡眠阶段, Firstbeat 和 PSG 的平均差异较大, 变异性较高($SD = 44.647$), 但未达到统计显著性($t = -0.479$, $p = 0.635$), 95%置信区间为 $-20.577 \sim 12.767$ min。这反映出 2 种方法在监测浅睡眠阶段的结果具有一定一致性, 但仍存在数据波动。

(3) SWS 阶段, 2 种方法的监测结果同样未表现出显著性差异($t = -0.881$, $p = 0.386$), 95%置信区间为 $-8.256 \sim 3.285$ min。Firstbeat 在这一阶段的监测与 PSG 较为一致, 且数据变异性相对较低($SD = 15.454$)。

(4) REM 阶段, 2 种方法的监测时长差异无统计学显著性($t = 0.618$, $p = 0.541$), 95%置信区间为 $-6.295 \sim 11.747$ min, 进一步表明 Firstbeat 在 REM 阶段

监测中的表现与 PSG 一致性较高。

总体而言,配对 t 检验结果显示,Firstbeat 与 PSG 在所有睡眠阶段的时长监测中均未表现出显著差异 ($p > 0.05$)。尽管浅睡眠阶段的变异性相对较大,但在清醒、SWS 和 REM 阶段,Firstbeat 的监测结果与 PSG 具有较高的一致性。这表明 Firstbeat 作为一种便携的监测工具,能够可靠地评估夜间睡眠结构。

2.3 特异性、敏感性和精确性比较

与 PSG 相比,Firstbeat 在监测不同睡眠阶段的能力包括特异性、敏感性和准确性等方面存在一些差异,如表 4 所示。

表 4 Firstbeat 和 PSG 监测睡眠各阶段的特异性、敏感性和准确性

Tab.4 Specificity, sensitivity, and accuracy of Firstbeat vs. PSG in monitoring each sleep stage

睡眠阶段	特异性	敏感性	准确性
清醒	0.87±0.11	0.93±0.03	0.90±0.05
浅睡眠	0.63±0.21	0.65±0.18	0.66±0.16
SWS	0.91±0.05	0.78±0.13	0.83±0.12
REM	0.80±0.13	0.60±0.24	0.84±0.12

由表 4 可以看出:

(1) 清醒阶段,Firstbeat 表现出较高的特异性、敏感性和准确性,显示出其在监测清醒状态时与 PSG 具有高度一致的能力。这表明 Firstbeat 能够可靠地区分清醒与睡眠状态。

(2) 浅睡眠阶段,Firstbeat 的监测性能有所下降。这表明浅睡眠阶段的生理特征可能较难通过 HRV 和加速度计的数据进行明确区分,从而导致监测一致性较低。

(3) SWS 阶段,Firstbeat 在特异性和准确性方面表现良好,但敏感性相对较低。由此可见,Firstbeat 在监测 SWS 阶段的性能接近 PSG 标准,尤其在排除非 SWS 阶段方面表现较为出色。

(4) REM 睡眠阶段,Firstbeat 的监测性能表现出一定的局限性。虽然 REM 阶段的特异性和准确性尚属可接受范围,但敏感性较低,可能与 REM 阶段的生理特征较难通过 HRV 和加速度计数据准确捕捉有关。

总体来看,Firstbeat 在清醒和深度睡眠阶段的监测性能接近 PSG 的“金标准”,具有较高的特异性和准确性。然而,其在浅睡眠和 REM 阶段的监测性能相对较弱,尤其是敏感性偏低。这一结果表明,Firstbeat 在监测生理特征较为明显的睡眠阶段(如清醒和 SWS)时效果较好,但在生理标记相对不明确的阶段(如浅

睡眠和 REM)则存在改进空间。

2.4 混淆矩阵比较

通过混淆矩阵^[17]来比较 Firstbeat 和 PSG 在测量夜间所有的睡眠阶段中清醒以及浅睡眠、SWS 和 REM 睡眠阶段的真阳性、真阴性、假阳性和假阴性,以及各阶段的准确率,结果如表 5 所示。

表 5 Firstbeat 对比 PSG 在睡眠各阶段的混淆矩阵及准确率
Tab.5 Confusion matrix and accuracy of Firstbeat compared to PSG across sleep stages

睡眠阶段	清醒	浅睡眠	SWS	REM	共计
清醒	2 923 ^a	1 650	141	3 876	8 590
浅睡眠	554	14 329 ^a	2 463	554	17 900
SWS	108	3 018	6 600 ^a	45	9 771
REM	231	2 331	147	6 273 ^a	8 982
共计	3 816	21 328	9 351	10 748	45 243
准确率/%	76.6	67.2	70.6	58.4	

注:a 对角线表示真阳性的数量。

由表 5 可知:

(1) 清醒阶段,Firstbeat 正确识别了 2 923 个清醒时段(真阳性),准确率为 76.6%。尽管表现出较高的一致性,但仍有 1 650 个浅睡眠时段和 141 个 SWS 时段被误分类为清醒。这些误分类可能与浅睡眠和清醒阶段在生理特征上的部分重叠有关。

(2) 浅睡眠阶段,真阳性为 14 329 个,对应的准确率为 67.2%。Firstbeat 将 554 个清醒时段和 554 个 REM 时段误分类为浅睡眠,同时将 2 463 个 SWS 时段误分类为浅睡眠。这种较高的误分类率可能与浅睡眠阶段的生理特征较为模糊、难以明确区分有关。

(3) SWS 阶段,Firstbeat 正确分类了 6 600 个时段,准确率为 70.6%。尽管大多数 SWS 时段被正确分类,但仍有 3 018 个时段被误分类为浅睡眠,108 个时段误分类为清醒。这表明 Firstbeat 在区分深度睡眠与浅睡眠时存在一定困难,但在总体上能够有效识别 SWS 阶段。

(4) REM 阶段,Firstbeat 正确识别了 6 273 个时段,准确率为 58.4%。这一阶段的准确率最低,主要原因可能是有 2 331 个 REM 时段被误分类为浅睡眠,同时 231 个清醒时段和 147 个 SWS 时段也被错误分类为 REM。这反映出 REM 阶段特征在 HRV 和加速度计数据中的区分度较低。

总体而言,Firstbeat 在清醒、浅睡眠和深度睡眠阶段的监测性能相对较高,但在浅睡眠和 REM 阶段的表现较弱。这一结果强调了 Firstbeat 在特定阶段(如浅睡眠和 REM)中进一步优化算法的必要性。混淆矩

阵提供了细致的分类错误模式分析,为改进睡眠分期的准确性提供了重要参考。

2.5 Bland-Altman 散点图比较

为直观展示 PSG 与 Firstbeat 在各睡眠阶段的评分差异,本文基于 30 min 测量期的 Bland-Altman 散点图比较 Firstbeat 和 PSG 所监测到的睡眠不同阶段时长的差异,如图 2 所示。

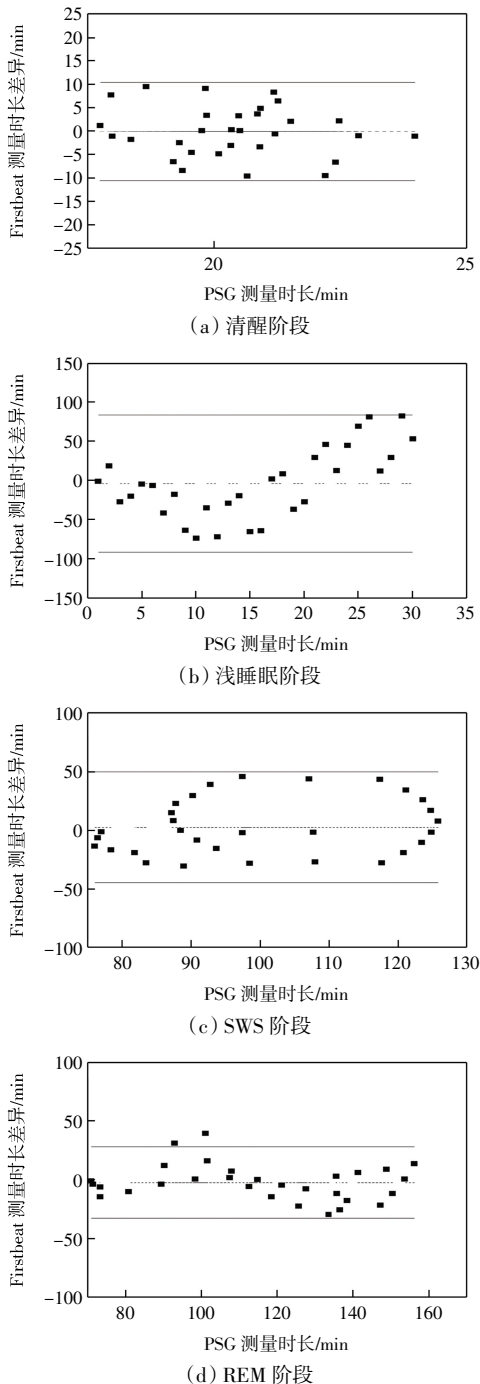


图 2 Firstbeat 和 PSG 在睡眠不同阶段的 Bland-Altman 散点图

Fig.2 Bland-Altman plots across sleep stages between Firstbeat and PSG

由图 2 可知:

(1) 清醒阶段,Firstbeat 与 PSG 的大多数观测值聚集在零差异附近,且偏离中位线的范围较小,表明 Firstbeat 和 PSG 在清醒阶段具有较高的一致性和可靠性。数据的均值差异接近零,表明 2 种方法在清醒阶段的时长测量基本一致。

(2) 浅睡眠阶段,Firstbeat 与 PSG 在浅睡眠阶段的时长差异可以观察到较大范围的偏离值,尤其在较长时长的测量中,Firstbeat 倾向于低估浅睡眠时长。这种差异反映出浅睡眠阶段的生理特征较为模糊,导致 Firstbeat 的监测表现相较于 PSG 略显不足。

(3) SWS 阶段,Firstbeat 与 PSG 在 SWS 阶段的差异呈现集中分布,大部分点位于允许范围内,说明 2 种方法在该阶段的监测结果具有良好一致性。然而,部分观测值显示出轻微的系统性误差,Firstbeat 在较短时长的深度睡眠中略有低估趋势。

(4) REM 阶段,2 种方法在 REM 阶段的大多数观测值位于允许范围内,但 Firstbeat 在 REM 阶段的误差分布相较于其他阶段更为显著,尤其在较长时长的 REM 阶段,表现为一定的高估。这表明快速眼动阶段的监测可能因其复杂的生理特征而受到一定影响。

总体而言,Bland-Altman 散点图表明,Firstbeat 与 PSG 在清醒和 SWS 阶段具有较高的一致性,而在浅睡眠和 REM 阶段存在一定程度的误差。这种误差可能与各阶段生理特征的信号区分度相关,为进一步改进 Firstbeat 的算法提供了依据。

2.6 比较结果分析

可穿戴设备已在健康和福祉消费市场上占据了重要地位,并随着技术进步不断涌现新设备和算法。目前,大部分研究都是基于加速度传感器的数据来监测睡眠质量和持续时间^[18],但也有研究探索了呼吸信号等其他测量方法^[19]。多篇综述评估了基于加速度计的睡眠可穿戴设备在监测睡眠与觉醒方面的准确性通常在 0.81~0.91 之间,敏感性在 0.87~0.99 之间,而特异性相对较低,介于 0.10~0.52 之间^[20-22]。例如,监测轻度睡眠的准确率介于 69%~81% 之间,监测 SWS 的准确率介于 36%~89% 之间,而监测 REM 的准确率介于 62%~89% 之间。这些结果显示,单独加速度信号本身可能不足以准确识别复杂的睡眠阶段,尤其在更复杂的睡眠阶段分类(如清醒、浅睡眠、深度睡眠及快速眼动期)时面临挑战^[23]。

相比之下,HRV 作为一种反映自主神经活动的生物标志物,显示出监测睡眠阶段的潜力^[7]。文献^[23]研究表明,结合 HRV 与呼吸信号的监测方法对于 SWS

的识别准确率可高达 89%;文献[24]研究表明,HRV 与加速度计联合应用可成功监测 75%的 SWS 和 70%以上的 REM 睡眠,但浅睡眠阶段的识别准确率较低,仅为 42%~52%。本文研究结果也具有类似的准确性,这进一步支持了将加速度计和基于 HRV 的测量相结合,以实现可重复的睡眠分期的概念,为 HRV 和加速度计结合应用于睡眠阶段监测的可行性提供依据。

本研究评估了基于 HRV 和加速度计的 Firstbeat 方法在监测睡眠阶段分期的能力,结果表明,Firstbeat 在监测浅睡眠和 SWS 时,与金标准 PSG 相比无统计学显著差异。然而,在 REM 和清醒时间的监测中,Firstbeat 表现出一定偏差。尽管存在这些差异,但考虑到夜间睡眠时长,这些误差在实际应用中具有可接受性,特别是在连续监测中。Firstbeat 对睡眠开始的监测非常准确,这与他人的研究结果^[25]一致。

需要注意的是,根据定义,睡眠阶段的划分依赖于 EEG、EOG、EMG 和 ECG 模式的变化。因此,缺少 EEG、EOG 和 EMG 信号的直接测量使得 REM 的监测特别困难。虽然 HRV 在本研究和其他研究^[7]中均表现出较强的阶段反映能力,但依靠其他生理指标来区分睡眠阶段始终是基于二手信息,其无法替代对 EEG、EOG 和 EMG 信号即时变化的准确监测。然而,研究表明,尽管心率变异辅助睡眠阶段监测在监测特定睡眠阶段方面的准确性较低,但仍能很好地评估睡眠结构。

尽管如此,本研究结果表明,与 PSG 相比,Firstbeat 方法在监测觉醒阶段具有良好的特异性和出色的灵敏度及准确性。对于浅睡眠阶段,特异性和灵敏度的表现相对较弱。这说明 Firstbeat 方法在监测具有显著副交感神经活动特征的睡眠阶段表现最佳,而浅睡眠由于其生理特征的不明显区分导致监测难度增加^[26]。

3 结 论

将 HRV 与加速度计数据相结合,用于夜间睡眠阶段的分期监测,可充分有效地测量夜间睡眠阶段的变化。研究结果显示:

(1) Firstbeat 在监测清醒、SWS 和 REM 方面表现出较高的准确性,而在浅睡眠阶段的特异性和灵敏度相对较弱。尽管存在一定的局限性,Firstbeat 系统因其便携性、经济性和易用性,仍是消费者级睡眠监测的潜在选择,能够为用户提供睡眠结构的概况和相关健康指导。

(2) 本研究在设计上具有较高的生态效度,参与者的性别分布均衡,监测场景设置于真实的生活环境

(即参与者宿舍),并在 2 晚对 Firstbeat 方法进行了验证。然而,本研究也存在一定的局限性。首先,尽管样本在 PSQI 评分上的睡眠质量差异一定程度上反映了群体多样性,但本研究未纳入被诊断患有睡眠障碍或其他相关健康问题(如心血管疾病)的参与者;其次,结果仅适用于健康个体,在这些特殊群体中的适用性仍需进一步验证。

(3) 未来研究将重点探索 Firstbeat 方法在临床环境中的应用潜力,尤其是在患有睡眠障碍或慢性病的受试者群体中。同时,改进算法以提高浅睡眠阶段监测的特异性和灵敏度,进一步提升其适用性和可靠性。结合多种传感器数据(如 HRV、加速度、呼吸频率等)开发个性化算法,可能是提高睡眠监测准确性的关键方向。

参考文献:

- [1] SILVA A C, SILVA A, EDWARDS B J, et al. Sleep extension in athletes: What we know so far—A systematic review[J]. *Sleep Medicine*, 2021, 77: 128–135.
- [2] CARSKADON M, DEMENT W, KRYGER M, et al. Normal human sleep: An overview[C]// DEMENT W C, KRYGER M H, ROTH T. *Principles and practice of sleep medicine*. Philadelphia: Elsevier Saunders, 2005: 13–23.
- [3] IBER C, ANCOLIISVAEL S, CHEESON A, et al. *The AASM Manual for the Scoring of sleep and Associated Events: Rules, Terminology and Technical Specifications*[M]. Westchester: American Academy of Sleep Medicine, 2007.
- [4] BRUNNER D P, DIJK D J, BORBÉLY A A. A quantitative analysis of phasic and tonic submental EMG activity in human sleep[J]. *Physiology & Behavior*, 1990, 48(5): 741–748.
- [5] PAIMEGGIANI P L. Physiologic regulation in sleep[C]// DEMENT W C, KRYGER M H, ROTH T. *Principles and Practice of Sleep Medicine*. Philadelphia: Elsevier Saunders, 2005: 185.
- [6] SZYMUSIAK R. Body temperature and sleep[C]// ROMANOVSKY A A. *Handbook of Clinical Neurology*. Philadelphia: Elsevier, 2018: 341–351.
- [7] HERZIG D, ESER P, OMLIN X, et al. Reproducibility of heart rate variability is parameter and sleep stage dependent[J]. *Frontiers in Physiology*, 2018, 8: 1100.
- [8] BERLAD I, SHLITNER A, BEN-HAIM S, et al. Power spectrum analysis and heart rate variability in Stage 4 and REM sleep: Evidence for state-specific changes in autonomic dominance[J]. *Journal of Sleep Research*, 1993, 2(2): 88–90.
- [9] BOUDREAU P, YEH W H, DUMONT G A, et al. Circadian variation of heart rate variability across sleep stages[J]. *Sleep*, 2013, 36(12): 1919–1928.

- [10] ACQUAVELLA J, MEHRA R, BRON M, et al. Prevalence of narcolepsy and other sleep disorders and frequency of diagnostic tests from 2013–2016 in insured patients actively seeking care[J]. *Journal of Clinical Sleep Medicine*, 2020, 16(8): 1255–1263.
- [11] BOE A J, MCGEE KOCH L L, O'BRIEN M K, et al. Automating sleep stage classification using wireless, wearable sensors[J]. *NPJ Digital Medicine*, 2019, 2: 131.
- [12] CRESPO-RUIZ B, RIVAS-GALAN S, FERNANDEZ-VEGA C, et al. Executive stress management: Physiological load of stress and recovery in executives on workdays[J]. *International Journal of Environmental Research and Public Health*, 2018, 15(12): 2847.
- [13] MUTIKAINEN S, HELANDER E, PIETILÄ J, et al. Objectively measured physical activity in Finnish employees: A cross-sectional study[J]. *BMJ Open*, 2014, 4(12): e005927.
- [14] SMOLANDER J, AJOVIITA M, JUUTI T, et al. Estimating oxygen consumption from heart rate and heart rate variability without individual calibration[J]. *Clinical Physiology and Functional Imaging*, 2011, 31(4): 266–271.
- [15] LOWN M, YUE A M, SHAH B N, et al. Screening for atrial fibrillation using economical and accurate technology (from the SAFETY study)[J]. *The American Journal of Cardiology*, 2018, 122(8): 1339–1344.
- [16] BACKHAUS J, JUNGHANNS K, BROOCKS A, et al. Test-retest reliability and validity of the Pittsburgh Sleep Quality Index in primary insomnia[J]. *Journal of Psychosomatic Research*, 2002, 53(3): 737–740.
- [17] KAHAWAGE P, JUMABHOY R, HAMILL K, et al. Validity, potential clinical utility, and comparison of consumer and research-grade activity trackers in Insomnia Disorder I: In-lab validation against polysomnography[J]. *Journal of Sleep Research*, 2020, 29(1): e12931.
- [18] BARON K G, DUFFECY J, BERENDSEN M A, et al. Feeling validated yet: A scoping review of the use of consumer-targeted wearable and mobile technology to measure and improve sleep[J]. *Sleep Medicine Reviews*, 2018, 40: 151–159.
- [19] DUNN J, RUNGE R, SNYDER M. Wearables and the medical revolution[J]. *Personalized Medicine*, 2018, 15(5): 429–448.
- [20] DANZIG R, WANG M X, SHAH A, et al. The wrist is not the brain: Estimation of sleep by clinical and consumer wearable actigraphy devices is impacted by multiple patient- and device-specific factors[J]. *Journal of Sleep Research*, 2020, 29(1): e12926.
- [21] HAGHAYEGH S, KHOSHNEVIS S, SMOLENSKY M H, et al. Accuracy of wristband fitbit models in assessing sleep: Systematic review and meta-analysis[J]. *Journal of Medical Internet Research*, 2019, 21(11): e16273.
- [22] DE ZAMBOTTI M, CELLINI N, GOLDSTONE A, et al. Wearable sleep technology in clinical and research settings [J]. *Medicine and Science in Sports and Exercise*, 2019, 51(7): 1538–1557.
- [23] LONG X, FONSECA P, AARTS R M, et al. Detection of nocturnal slow wave sleep based on cardiorespiratory activity in healthy adults[J]. *IEEE Journal of Biomedical and Health Informatics*, 2017, 21(1): 123–133.
- [24] MUZET A, WERNER S, FUCHS G, et al. Assessing sleep architecture and continuity measures through the analysis of heart rate and wrist movement recordings in healthy subjects: Comparison with results based on polysomnography[J]. *Sleep Medicine*, 2016, 21: 47–56.
- [25] SCOTT H, LACK L, LOVATO N. A systematic review of the accuracy of sleep wearable devices for estimating sleep onset[J]. *Sleep Medicine Reviews*, 2020, 49: 101227.
- [26] PENZEL T, KANTELHARDT J W, LO C C, et al. Dynamics of heart rate and sleep stages in normals and patients with sleep apnea[J]. *Neuropsychopharmacology*, 2003, 28(Suppl 1): S48–S53.

本文引文格式:

常宁,肖云瀚,万帅,等.基于心率变异性的高水平运动员睡眠阶段分期分析[J].天津工业大学学报,2025,44(4): 52–59.

CHANG N, XIAO Y H, WAN S, et al. Analysis of sleep stage classification in elite athletes based on HRV [J]. *Journal of Tiangong University*, 2025, 44(4): 52–59(in Chinese).