



# 数据驱动精准训练:理论内涵、实现框架与推进路径

钟亚平<sup>1</sup>, 吴彰忠<sup>1</sup>, 陈小平<sup>2,3\*</sup>

(1. 武汉体育学院 体育大数据研究中心, 湖北 武汉 430079; 2. 国家体育总局体育科学研究所, 北京 100061;  
3. 宁波大学 体育学院, 浙江 宁波 315211)

**摘要:**“数据驱动”和“精准训练”已成为运动训练领域的主旋律和高频词,但相关研究系统性和深入程度有待提升。通过文献梳理、逻辑推理、辩证分析等方法,对数据驱动精准训练的理论内涵、实现框架和推进路径进行了深入探究。研究认为:1)数据驱动精准训练是运动训练主体通过大数据技术与运动训练实践的深度融合推动运动训练客体最优化发展的一种运动训练理论与方法新范式,具有突出负荷精准量化、注重损伤精准预警和追求效能精准评估的鲜明特征;2)数据驱动精准训练的实现框架是在建设训练数据中台和开发智能训练干预引擎的基础上,借助“精准训练智能大脑”(数据驱动精准训练服务平台),通过“智控”+“人控”相结合的方式,对竞技训练过程中的八大核心关键环节进行全程性、立体化、剂量式的微观和滴定控制;3)数据驱动精准训练在推广应用过程中应铺好“平台建设”奠基石,清除“数据素养”拦路障,迈实“落地应用”试探步,为科技助力奥运备战、促进竞技训练高质量发展、形成“数据把脉、全面互动、精准反馈、轻负高效”的运动训练新发展格局开拓新思路。

**关键词:**数据驱动;精准训练;训练数据;科技助力;奥运备战;竞技体育

**中图分类号:**G808.12 **文献标识码:**A

“精准”一词释义为精益、准确。“精准”一词应用于现代运动训练领域始见于20世纪末,Burke(1998)提出了“精准训练”(precise training)一词,并重点论述了可通过监测训练过程中运动员的心率水平来调控训练负荷,进而达到提升训练效果的目的。然而,囿于当时数据采集与处理的技术及条件,精准训练的理念未能普遍推广。伴随信息时代新观念、新技术的迅猛发展,特别是泛在网络、云计算、智能可穿戴、人工智能、大数据等新一代数字技术的社会化进程逐渐加快,数据驱动精准训练在当今体育国际大赛和职业联赛中发挥的作用越趋重要,成为支撑竞技训练生态圈可持续、高质量发展的重要引擎。

党的十八大以来,在数字中国建设背景下,我国积极实施科技助力奥运备战和国家重点研发计划“科技冬奥”重点专项,关于数据驱动精准训练的理论研究和实践应用逐渐“升温”。在理论上,部分学者开始关注奥运备战训练中的海量数据问题,提出在整合分期的训练框架下,通过数据科学与大数据技术助力竞技训练的精准化、个性化、智能化发展(陈小平,2018a;胡海旭等,2021;袁守龙,2018)。在实践应用领域,我国跳水、田径、赛艇国家队利用大数据技术对运动员体能、技术、恢复、营养等训练细节进行全方位把控,最终在东京奥运会上实现了突

破。但现有研究多停留在理念普及、场景解析和案例介绍层面,不够系统和深入。鉴于此,本研究拟对数据驱动精准训练的基本概念、内涵特征、实现框架等关键性基础问题进行探讨,以期紧随国际前沿,助力我国运动训练数字化、智能化、个性化发展,促进我国运动训练科学化水平的全面提升、“科技助力奥运备战”的高质量开展,以及“数据把脉、全面互动、精准反馈、轻负高效”的竞技训练新发展格局的形成。

## 1 数据驱动精准训练的理论内涵

### 1.1 基本概念

数据驱动是指主体运用大数据技术,通过挖掘和分析数据资源及相关数据产品来提供决策和支持方案,从而

收稿日期:2021-07-14; 修订日期:2021-11-08

基金项目:国家社会科学基金项目(19BTY099);国家体育总局决策咨询研究重点项目(2021-B-09-02);湖北省中央引导地方科技发展专项(2019ZYD054)。

第一作者简介:钟亚平(1968-),男,教授,博士,博士研究生导师,主要研究方向为运动训练、数字体育,E-mail:zhongyaping@whsu.edu.cn。

\*通信作者简介:陈小平(1956-),男,教授,博士,博士研究生导师,主要研究方向为运动训练,E-mail:chenxiaoping@ciss.cn。

推动客体的运行(蔡玉卿, 2019), 其内涵一般包括以下3个方面: 1) 以数据总量大、类型多、流转速度快为前提条件; 2) 以通过数据挖掘技术分析和处理数据为基本路径; 3) 以能够为下一步的行为提供有价值的指导信息为核心目的。当前, 随着智能可穿戴、物联网、云计算等数字技术在训练实践中的广泛应用, 各种基于生物学、心理学、营养学、环境学等多维训练指标的数据能够被连续动态采集, 并逐步形成训练大数据。当这些训练数据流在训练的各个环节不断产生与闭环运行时, 就会形成具有自我修正功能的正向反馈机制, 进而驱动运动训练过程, 实现精细化控制。

从驱动范式看, 运动训练可分为经验驱动型和数据驱动型, 前者是指运动训练主体以直觉经验驱动训练决策为前提, 基于单一的主观本体感受所形成的单中心训练模式; 后者是指运动训练主体以大数据驱动训练决策为前提, 基于多元的客观训练指标监控所形成的多中心训练模式(杜长亮等, 2019)。数据驱动精准训练遵循“数据采集—数据处理—模型训练—测试评估—服务应用”技术路线, 从数据驱动的“神经中枢”切入, 汇聚多源异构的训练数据类型, 通过对训练大数据的深度挖掘和多源分析, 能够快速发现运动员训练短期适应效应量与长期适应预期目标值, 辅助训练团队提出更精准的个性化“训”方案, 指导运动员进行更精准的高质量“练”, 进而推动数字时代竞技训练生态圈健康可持续发展(李爱霞等, 2021)。

基于此, 本文将数据驱动精准训练定义为运动训练主体(运动员和教练员团队等)将数据科学(大数据技术)与运动训练实践相结合, 在数据化运动员竞技能力系统发展过程的基础上, 通过对训练过程每一环节相关指标的微观和滴定控制来推动运动训练客体(竞技能力和运动成绩)最优化发展的一种运动训练理论与方法新范式。其中, “提升训练效能”是根本目的, “训练大数据”是实现基础, “精准训练”是实现手段。

## 1.2 支撑理论

### 1.2.1 “数据驱动决策”理论及模型

Wladawsky-Berger(1999)将数据驱动决策定义为“基于对数据的整合分析, 而不是仅仅依靠直觉进行决策的实践”。Ackoff(1989)构建的“DIKW体系”认为当数据和信息积累到一定数量后就会形成完备的知识体系和生成智慧, 进而驱动科学决策。当前, 由数据集成、数据认知、决策优化、搜索挖掘、预测干预等过程构成的大数据驱动范式被广泛应用于过程控制、信息管理、决策支持、趋势预测等诸多方面。数据驱动科学训练决策(图1)借助智能数据分析理论及模型对海量运动训练数据进行分析、挖掘, 辨析训练负荷及训练方式对运动员竞技能力及运动表现的相关性及因果性, 发现运动训练复杂体系的训

练规律, 以促进不同主体更好地开展竞训活动: 1) 针对教练员, 数据驱动能够帮助其更好地了解运动员的训练状态和训练需求, 为完善个性化训练方案、靶向干预训练决策、优化运动训练过程、全面评估训练效果等提供科学依据; 2) 针对运动员, 数据分析结果有利于其进行自我导向训练、自我训练评价和规划自我训练目标的实现, 自我导向训练有利于使抽象的训练过程更具体、更有目标且更加安全; 3) 针对科研人员, 数据驱动强调理解和关注训练结果(无效训练、训练适应、超量恢复、过度训练等)产生的内在机理和多因素机制, 为发现训练规律、创新训练理论、复制训练成功经验提供重要保障。

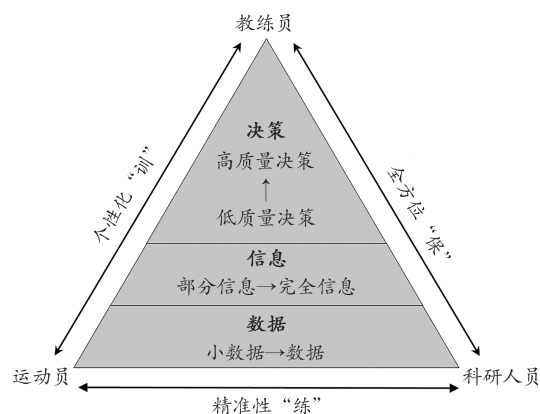


图1 数据驱动科学训练决策模型

Figure 1. Model of the Data-Driven Scientific Training Decision

### 1.2.2 “应激-适应”理论及模型

加拿大病理学家Selye(1950)的“应激-适应”理论将机体对刺激的应答过程分为报警、抵抗、疲惫3个阶段, 并认为不同的应激源会使机体产生与之相适应的应答反应, 进而破坏有机体的内环境稳态, 随后有机体必须通过应激方式寻回平衡或稳定。该理论在训练实践中不断演化发展, 为运动训练构建了一个宏观的生物学基础。McGuigan(2017)基于“应激-适应”理论构建了竞技训练的“一般适应综合征”模型, 将运动员机体对训练负荷刺激的适应过程分为4个时相(图2)。“应激-适应”理论及其模型提示, 对应激源(训练负荷)在量度、强度和时机上进行精准控制是确保运动员产生最佳训练适应的先决条件, 负荷过小不能累积训练效果, 负荷过大会徒增过度训练综合征风险, 负荷时机把控不准则无法利用训练痕迹效应实现超量恢复。数据驱动精准训练的优势在于可以根据运动员的个体差异在最佳时机给予机体最宜刺激, 使其产生最优适应, 防止无效训练和过度训练的发生, 进而提升训练的质量效益。

此外, 国际训练学界在“应激-适应”理论的基础上, 融合了数学、生理学、物理学、数据科学、计算机科学提出的临界功率(critical power)模型、训练负荷-效应(impulse-

response)模型、竞技潜能元(performance potential meta, PerPot)模型等,已经实现了对运动训练复杂生物系统的方程式、计量化表达,这些理论模型从关注“刺激”到关注“反应”,再到关注“过程与结果”,为数据驱动精准训练的实现提供了重要基础(胡海旭,2016;Morton,2006;Perl,2004;Thierry,2003)。如Perl(2003)提出了二维水平的

PerPot模型,在简化复杂训练要素的基础上,通过计算机软件来模拟训练负荷刺激对竞技能力产生短期与长期反应的动力过程,设置应答流与应变流延迟时间(时滞)的不同参数,调控运动训练中机体适应的正负效果,是一种用于优化即刻训练计划和预防过度训练的计算模型,实现了对运动训练定性现象的定量描述。

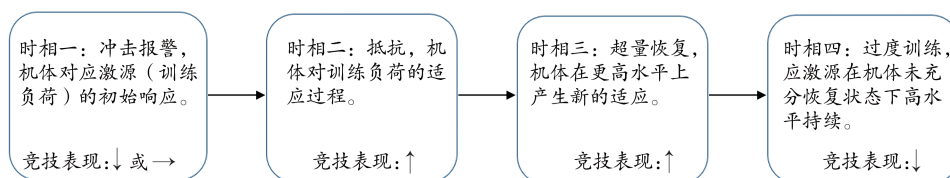


图2 “一般适应综合征”模型(McGuigan,2017)

Figure 2. Model of the General Adaptation Syndrome (McGuigan,2017)

### 1.2.3 整合分期理论及模型

Bompa等(2018)将整合分期定义为运动员不仅需要技术、身体和战术训练,同时还需要在周期中整合心理训练和营养计划,这是将所有训练要素纳入一个综合性年度训练计划的分期模式。杨国庆(2020)认为,整合分期理论试图解决经典分期与板块分期中忽视训练空间维度的核心要素问题,其精髓是去单个要素最优化,重视将动作模式、运动技术、训练负荷、运动心理、运动营养等诸多训练要素的参数加以数字化整合和量化权衡,以取得训练效益的新突破。可见,整合分期模式是对原有分期理论时间维度的空间立体化拓展,侧重当前分期阶段上的空间结构与功能维度的要素调节。在训练实践中,建立整合分期的应用模型必须要解决:1)在分期阶段维度上回答每个训练分期阶段的时间长短、内容调和、负荷配置等问题;2)在分期内容维度上回答各项分期因素在不同分期阶段的负荷节奏问题;3)在分期负荷维度上解决不同分期内容的训练负荷能否让各分期阶段(或训练时刻)的运动员机体产生最佳适应的问题。显然,传统的经验驱动训练模式难以满足整合分期多维立体化管理各训练要素的要求,而数据驱动精准训练可以基于整合分期中瞬态生理、心理等多模态数据的集成与挖掘,并借助深度学习等算法在时空维度上对运动训练各要素进行有机整合与平衡配置,使制定高效的个性化整合分期训练方案得以实现。

## 1.3 应用特征

### 1.3.1 突出“负荷精准量化”

运动训练监控的内容涉及运动员的选材、训练、参赛、管理等各个方面,而数据驱动精准训练的出发点和落脚点是“训练”,因此对训练负荷的精准量化是数据驱动精准训练的基础和重心。训练负荷的精准量化分为外部负荷量化和内部负荷量化,外部负荷量化用于评价运动员动作稳定性、协调性及做功情况,内部负荷量化用于评

价机体的生理和心理系统对训练负荷刺激的适应程度,相同的外部负荷会导致不同运动员(或同一运动员的不同训练阶段)产生不同内部反应(Impellizzeri et al.,2019),如举重和马拉松运动员在进行100 kg深蹲时会有不同的感知和生理反应,中长跑运动员在运动损伤恢复前后对1500 m跑会有不同的内部生理反应(Johann,2017)。

外部负荷的精准量化具体表现为通过GPS、陀螺仪、加速度计、功率计等设备对练习的组数、时间、质量、距离、频率、间歇时间等进行量化。如通过GPS监控1名足球运动员在1次训练课中的跑动距离为9725 m,并在此基础上对该运动员的跑动速度、变向次数、负荷累计等相关指标进行分析,以评价运动员的运动能力和进行损伤预警(Malone et al.,2021;Mauro et al.,2021)。内部负荷的精准量化具体表现为通过生化测试等方法对运动员的心率及心率变异性、呼吸与能量代谢、摄氧量等心肺指标,血液、汗液、唾液、尿液等体液指标,肌电图、脑电图、近红外光谱等神经肌肉指标进行量化(Rago et al.,2020)。此外,还可以通过问卷方式,如主观疲劳量表(rating of perceived exertion, RPE)能够对运动员训练前后的主观体验进行测评,以监控机体承受负荷刺激的水平。

### 1.3.2 注重“损伤精准预警”

在运动训练实践中,负荷安排过度会使机体无法及时从训练中得到恢复,运动员不仅无法产生良好运动适应和提高运动表现,反而会增加损伤和患病风险,甚至发生死亡。Veugelers等(2016)研究发现,在持续1周的高负荷训练累积或周间训练负荷变化幅度较大的情况下,橄榄球运动员受伤的风险会明显增加。通过对其他运动项目(特别是同场对抗类运动)的研究也陆续证明了运动员的受伤风险与训练负荷的累加、骤变和恢复不足有很大关系(Anderson,2018;Weiss et Al.,2017)。Nibali等(2015)认为,长时间的有氧耐力高强度训练会削弱机体的免疫功能,增加运动员上呼吸道感染的患病风险。



数据驱动精准训练注重优化运动损伤的风险预警机制,通过对运动员健康生活指标数据和竞技训练指标数据的多维采集和整合分析,借助人工智能算法得出运动员的“训练窗”(即在最佳训练时间进行最适合机体现实状态的最优训练内容和最宜训练强度),以最大限度降低训练过度 and 运动损伤的发生风险。如 omegawave 竞技状态综合诊断系统作为一款能够实时同步监测运动员中枢神经系统和心肺系统训练准备状态的无创式、便携化评估系统,可以通过测量当前运动员的心电图、脑电波、静心率、训练时安全心率阈值、心率变异性、肌肉功能状态、疲劳指数、能量代谢指数等数据,分析出运动员的现实状态和疲劳程度,并匹配算法引擎给出“训练窗”建议。教练员可以在此基础上调控运动员训练的“开窗时间”(根据当前机体状态,建议的最佳训练负荷和训练内容)和“闭窗时间”(当机体疲劳时,建议降低训练负荷或进行恢复训练),以精准有效预防运动损伤。此外,一款由芬兰开发的 firstbeat 运动员专用系统被广泛应用于竞技训练中,该系统仅 3 min 即可掌握运动员的神经疲劳恢复情况,从中发现疲劳恢复较差的运动员,据此个性化调整训练节奏、制定赛前减量方案,避免运动员因过度训练导致运动损伤的发生(Jack et al., 2019)。

### 1.3.3 实现“效能精准评估”

人体运动的内在本质是能量代谢(黎涌明等, 2014),运动训练科学化的目的之一是通过适宜的训练刺激(应激源)使机体发生预期的良性应激反应,提升机体能量代谢、肌肉做功以及适应训练刺激的效能。陈小平(2020)指出,在训练实践中(特别是体能类、周期性运动项目),运动员训练时的现实能量代谢和肌肉做功效能与该运动项目在比赛中实际需要的供能方式和最佳做功难以始终保持良好匹配,这是制约训练适应效能提升的关键。任何时候,人体的三大供能系统都会同时参与供能,三大系统供能百分比是项目特征的重要组成部分,教练员训练理念的形成以及训练计划的制定和实施必须建立在对这个特征认识的基础之上。

数据驱动精准训练可以在对不同运动单项能量代谢和肌肉做功特征进行精准评估的基础上,通过定量施加训练负荷使运动员的生理、心理适应水平与该运动项目的能量代谢特征和最佳做功方式实现高度拟合,让机体对训练负荷刺激产生最佳适应效能(精准良性应激),避免出现过度训练和训练不足,进而对提高运动员成才率、缩短培养过程、延长运动寿命、提升训练质量效益发挥积极作用。黎涌明等(2018)通过量化的方法对男子赛艇运动员 240 次卧拉练习的有效性进行的研究发现,40% RM 强度的 240 次卧拉练习属于高强度训练,最高血乳酸浓度为  $(8.60 \pm 1.28)$  mmol/L。因此,240 次卧拉练习非但不利于提高赛艇运动员实际所需的力量耐力,还有可能破坏

通过低强度有氧训练提高起来的力量耐力。

## 1.4 实践价值

数据驱动精准训练吻合 Jim Gray 提出的数据密集型科研“第四范式”,其将带来一场运动训练理论与实践的“范式”变革。数据驱动精准训练高度契合了当前运动训练立体化智控、精准化施控和个性化调控的时代需求和未来趋势,助力运动训练工作从“要素驱动”向“数据驱动”和“创新驱动”转变,推动运动训练科学化水平再上新台阶。

### 1.4.1 助力“运动训练模式”立体化智控创新

传统的运动训练监控局限于运动员训练过程中生理和心理负荷指标的监控,忽视了对运动表现(技术、战术)、负荷变量(负荷量、负荷强度、负荷总量、训练环境、运动损伤、疾病)、生活方式(营养、睡眠、生活压力、社会关系)等因素的全方位立体化监控。随着信息时代数字技术的发展,数据驱动精准训练模式下的数字化监控逐渐趋于立体化,监控的维度越来越广、准确度越来越高、颗粒度越来越细,数据采集的方式也将更加便捷,“数据石油”和“算法引擎”的匹配将驱动运动训练模式数字化立体化监控的“智慧化”,把教练员从诸多常规化、机械化的工作中解脱出来,使其将工作重心聚焦到创新性的训练设计及个性化的训练指导上来,进而促进训练效益质的提升。

### 1.4.2 增强“运动训练过程”精准化施控能力

数据驱动精准训练贵在“精准”,训练质量效益的提升需要科研人员全方位“保”、教练员个性化“训”、运动员精准性“练”。陈小平(2018b)认为,当代的运动训练过程控制是一个系统工程,须因人而异、因时而进、因势而新,是点对点 and 数字化的定量调控,而不再是凭主观经验的定性调控。仅凭主观经验的定性判断已无法满足运动训练过程控制的精准高效要求,运动训练中无时无刻不在生成的海量数据和日新月异的信息技术更迭为数据驱动精准训练提供了先决条件,基于海量训练数据的科学训练决策将使运动训练过程控制实现从抽象到具体、从定性到定量、从共性到个性的根本性转变。

### 1.4.3 满足“运动训练适应”个性化调控需求

适应理论指出,机体对应激源(训练负荷)的适应具有鲜明的个体化特征。不同运动员的应激能力受性别、年龄、生理、心理、遗传、个性和成长环境等内外部因素影响而千差万别,同一名运动员在不同的训练阶段、训练条件和训练环境下对训练安排均具有个性化要求。因此,对不同运动员在不同时空条件下的运动训练适应进行个性化调控是运动训练科学化发展的必然要求和时代趋势。数据驱动精准训练是基于数据科学的训练信息整合,不仅可以通过数据挖掘技术发现运动员竞技能力演化的群体特征和运动项目对竞技能力需求的一般规律,

还可以结合个体的适应水平和应激能力等特征识别不同运动员竞技能力提升的个性化需求,并通过算法自动生成与运动员适应水平相匹配的个性化训练方案,真正实现竞技训练的“轻负高效”。

## 2 数据驱动精准训练的实践观察

### 2.1 数据驱动精准训练的应用场景呈现

随着竞技体坛国际竞争的日益激烈,以及5G、物联网、无线传感、人工智能、云计算、大数据等数字技术的社会化进程不断加快,数据驱动科学决策理念在竞技训练领域备受推崇。Rajsp等(2020)调研发现,采用数据智能分析理论提升训练效能的运动项目涉及篮球、足球等31个项目。在此基础上,本文摘取部分在国内外较有影响力的数据驱动精准训练应用案例,并依照Fister等(2019)提出的运动训练过程的4大环节(计划、实现、控制、评估)进行分类呈现,以便对数据驱动精准训练的全貌形成具体认知和整体把握。

第一,在计划环节,数据驱动精准训练主要体现在:1)训练目标设定。如Velloso等(2013)通过对田径运动员前6周的营养和体能数据进行采集和建模,可以准确预测第7周800 m和5 000 m的运动成绩,这可以为教练员设定短期训练目标提供重要参考。2)分期阶段划分。如我国冬残奥国家队使用的“冬残奥数据服务平台”,参考Schelling(2021)等提出的调度支持系统的运行思路,把影响训练计划的因素分为稳定因素(比赛日期、比赛地等)和非稳定因素(运动员现实状态、训练的气候环境等),并借助计算机演算关联出这些影响因素之间的关系,建立一个训练分期安排的决策系统,实现了训练日程安排的自动规划和调整。3)训练方案制定。如Eriksson等(2021)通过搜集、整理、数据化处理大量优秀游泳运动员的年度训练计划集成训练计划数据库,然后借助机器学习构建了训练计划自动生成器,能够为教练员智能推荐实施训练计划时的负荷强度。

第二,在实现环节,数据驱动精准训练主要体现在:1)热身阶段的运动员状态(生理、心理)诊断。如在英格兰超级联赛球队中广泛使用的Viper Pod训练背心可以监测运动员热身时的步频、步幅、双脚的压力负荷等指标数据,并且会在热身开始15 min后生成反馈报告,以评估运动员的训前准备状态,帮助团队科学合理地安排后续训练任务。俱乐部在使用这款训练背心后,单赛季球员训练伤病的发生风险降低了约50%(Smpokos et al., 2018)。2)训练实施中的负荷(内部负荷、外部负荷)监控。如美国奥林匹克委员会为拳击等格斗对抗项目的运动员提供了Hykso腕带,该腕带是一个具有传感功能的便携穿戴设备,可以实时记录和传输运动员出拳的速度、次数和轨迹等指标数据(Stanislav, 2020),并在训练中使用了一种特

制的具有测量打击力度功能的拳袋,教练员可以在便携式的显示终端上即时掌握运动员出拳力量的峰值,以及达到力量峰值和维持力量峰值的时间,并在此基础上精准施加和调控训练负荷(Kilduff et al., 2013)。3)负荷刺激后机体疲劳程度的评估。如训练后通过可穿戴设备采集运动员的睡眠数据(就寝时间、醒来时间、睡眠潜伏期等),使用基于计算机的反应时间和快速视觉信息处理任务来测量运动员精神运动速度,并利用这类数据量化评估运动员的疲劳水平,以防训练过度,促进机体实现超量恢复(Halson, 2014)。

第三,在控制环节,数据驱动精准训练主要体现在:1)技术细节改进,如美国职业篮球联赛(National Basketball Association, NBA)利用Noah Basketball公司推出的数据服务系统测量运动员投篮的出手点、球的弧度、球与篮圈的偏度等数据,追踪运动员的投篮技术,以发现影响运动员投篮命中率的技术细节,并为运动员制定极具个性化的投篮训练方案。2)运动战术优化。如McIntyre等(2016)运用Sport VU系统对NBA联赛中27万个挡拆数据进行分析,并通过机器学习将防守挡拆的策略分为挤过、绕过、包夹、换防4种类型,以帮助球队教练员制定防守挡拆进攻战术的针对性策略。以勒布朗·詹姆斯为例,数据分析发现其在打挡拆时面对这4种防守方式时每回合的平均得分分别约为0.9、1.0、1.1和0.8,这意味着普遍来讲防守勒布朗·詹姆斯打挡拆的最好策略应为换防。3)动作模式改善。如我国田径队使用的FASTMOVE智能捕捉和动作诊断系统,弥补了传统的红外和惯性动作捕捉系统必须配套专用灯光和佩戴反射标记点或传感器硬件的局限,可以实时捕捉摄像机前运动员的身体动作,并以人体模拟棍图、模拟骨架图等形式呈现,实现对关节位置、关节角度、落地姿势等动作功效指标的可视化分析,帮助运动员优化动作模式,有效预防关节损伤。

第四,在评估环节,数据驱动精准训练主要体现在:1)运动损伤预警。如加拿大的Kinduct运动员智能管理系统利用多种数据集成技术,记录了1万多名运动员150多万次训练课的数据,并通过算法分析出了不同运动项目、不同类型运动员发生运动损伤的规律和特征,以帮助教练员有效预防运动损伤特别是重大伤病的发生。以橄榄球项目为例,在采集分析大量损伤数据的基础上,发现对男子运动员来说损伤最严重的部位是膝关节前交叉韧带(缺训近180天),频率最高的病因是脑震荡。这些数据分析的结果能够指导教练团队在组织训练时实施针对性的预防措施。2)运动表现分析。如中国跳水梦之队引入的“3D+AI”跳水训练系统,利用高速相机、计算机视觉、3D建模等设备和360°全方位采集运动员在1~2 s内完成的跳水动作指标,并完成身体姿态的三维重建及量化评估,将先前“起跳不够”的模糊定性评估转变为“比标准



起跳高度低了 10.8 cm”的精准定量评估,还可以通过机器学习对动作的完成质量和难度系数进行打分。3) 竞技人才选拔。如全球范围内超过 750 家球队俱乐部使用的一款名为“Catapult Sport”的穿戴设备,这款设备集成了 GPS、心率带、多轴加速度仪、陀螺仪、电子罗盘等组件,能够有效采集和记录运动员的跑跳、加速、变向、力量、单次负荷、负荷累积等指标,同时还配备分析设备以帮助俱乐部评价和预估运动员的竞技能力水平和未来成就(Kuzmits et al., 2008)。

## 2.2 数据驱动精准训练的典型案例剖析

前文按照“计划—实现—控制—评估”的逻辑进路对近年来国内外数据驱动精准训练的应用场景进行了分类呈现,为了进一步揭示数据科学和大数据技术给竞训领域带来的新变化和新效益,本研究选取在东京奥运会上取得优异成绩的中国国家赛艇队作为我国高水平竞技训练依托数据驱动逐渐实现数字化转型、迈向高质量发展的典型案例并加以剖析,以展现训练大数据在现代高水平训练场景下对体能、技术、机能、疲劳、营养、睡眠、伤病预防、运动表现等方面的深度介入。

第一,数据驱动划桨技术细节精准优化。赛艇的船体在水中划行受流体力学的作用影响很大,其技术训练相对复杂,需要兼顾运动员与桨之间,桨与船体之间的连接关系。研究表明,同等条件下船体匀速前进时所受阻力最小(茅洁,2011),因此,如何让运动员在完成“提桨—拉桨—按桨—推桨”这一周期性技术流程过程中保持艇速稳定以减少动力损失成为赛艇技术训练的关键环节。显然,经验丰富的教练员也无法通过肉眼和经验去实时识别和修正这些重要信息的细微变化,一旦某些欠佳的技术细节不能得到及时优化,其负面影响会在动作周期性重复中被无限放大,进而使技术优化陷入瓶颈。为此,中国赛艇队量身定制了一套综合智能训练辅助系统。首先是数据精准采集。通过艇(桨)上安装的测力、测速、测稳等无线传感装置和无人机视频追踪技术实时采集船体前进过程中人、船、桨的视频数据,并将这些数据即时传输到流数据处理平台(streaming data platform, SDP)。SDP 作为一个可弹性扩展平台可用于实时获取、存储和分析边缘场景下连续的训练流数据(果青,2021)。其次是数据精准分析。综合智能训练辅助系统在 SDP 提供的海量数据源和算力基础上,调用基于 25 个关键点的运动生物力学算法分析模型,对视频中的每一帧画面进行姿态检测,对关键部位在划桨周期动作变化过程中的动作模式进行实时综合评估,发现训练中动作速度、角度、力度以及起桨的幅度、稳定性等数值偏差,并利用可视化技术动态呈现每个周期的动作的划动功率、桨叶抓水效果、每桨重复率等关键指标。同时,考虑到运动员个体生理差异使得技术动作很难存在人人都适合的“黄金标准”,

系统在预定义的类似“黄金标准”的模型基础上,采用深度学习 CNN 算法将运动员的训练实时数据与历史数据进行比较,从而迭代出灵活性极高且适合不同类型运动员实际情况的模型标准。其三是结果精准反馈。为运动员设计的内置摄像头和传感器等装置能够使其实时看到自己的每一个动作,并根据教练员基于动作和船体数据给予的科学提示进行纠正或微调,进而优化自己的动作技术细节,实现竞技技术表现的突破。

第二,数据驱动赛艇专项体能精准提升。赛艇是典型的以有氧为主的力量耐力型运动项目,比如,在 2 km 赛程中,运动员需要持续高强度运动 5.5~8 min,划桨 210~230 次,划桨的峰值功率接近 400 W,其无氧供能占比达到 80% 以上。因此,专项体能的强弱特别是力量耐力水平是影响赛艇竞技成绩的基础因素,而对训练负荷的精准控制在强化赛艇运动员的专项体能过程中起着重要的杠杆作用。一方面,中国赛艇队依据“世界冠军模型”对运动员体能训练内、外部负荷目标进行精准设定,如在外负荷方面,训练过程中把 2 000 m 约 8 min 的陆上测功仪训练任务分为开始、赛中和结束阶段,其中开始阶段的 1 min 桨频为 42~48 桨/min,每桨需达到 1 000~1 500 N 的力;赛中阶段的 6 min 桨频为 32~38 桨/min,每桨需达到 500~700 N 的力;结束阶段的 1 min 桨频为 38~40 桨/min,每桨需达到 500~700 N 的力;针对内部负荷,可以通过测试运动员的血乳酸浓度、RPE 等指标来实现对训练强度的精准把控,如把血乳酸浓度小于 2 mmol/L、RPE 得分小于 17 的中低强度训练作为赛艇运动员的主要训练内容(Nugent et al., 2020)。另一方面,中国赛艇队借助陆上测功仪、可穿戴设备等数字化体能训练器材实现了对训练负荷实施过程的精准调控,如具有阻力系数调节功能,以及力量、做功、成绩(模拟划行 2 000 m 所需的时间)、桨频等信息显示功能的陆上测功仪能够让赛艇运动员在进行陆上体能训练时通过电子屏幕上的运动划桨频率、肌肉力量、峰值划桨功率、速度心率等内外部负荷指标数据,及时调整自己的训练状态和负荷强度,使体能训练过程更加科学可控。更重要的是,综合智能训练辅助系统自动收集和预处理训练过程中实时产生的内外部负荷量化数据,并采用聚类算法进行深度分析,建立适用于国家赛艇队运动员个体特征的体能训练负荷节奏调控模型,这种超适宜的负荷刺激再配上超适应的恢复带来的是专项体能训练负荷极限阈值的瓶颈突破。正是通过这种数据驱动专项体能精准提升的训练模式,中国赛艇队最终实现了“强化体能、恶补短板”的目标。以测功仪 30 min 20 桨训练为例,按照国际标准,运动员需要以 20 桨/min 的频率持续做功 30 min,根据体质量不同,平均 500 m 的配速达到 1 min 48 s~1 min 52 s,女队此前仅 1 人能达到国际标准,现在达标率为 100%,这为东京奥运赛场上女子四

人双桨项目取得优异成绩奠定了基础。

第三,数据驱动运动员膳食营养精准摄入。通过科学化膳食摄入满足不同专项运动和训练负荷特点的营养素,是支撑运动员保质保量完成高强度训练和比赛任务的物质基础。为了满足不同运动员或同一运动员在完成不同训练负荷后所需营养类型及营养量的个性化需求,中国赛艇队引入了智慧餐盘系统,以保证运动员膳食结构的科学化和营养补充的精准化。饮食前,每位运动员的数据会提前输入到营养师的运动员科学膳食系统,系统会自动生成运动员当日所需的食物热量,以及蛋白质、脂肪、碳水化合物、维生素等营养素的摄入量。饮食中,运动员通过人脸识别系统进入餐厅后,会拿到一个与其信息相匹配的内置芯片系统的餐盘,随后打进餐盘的食物热量、质量、糖分都会通过餐盘的热量显示系统、称重系统和辐射系统呈现给运动员并进入大数据分析系统。饮食后,运动员当天摄入食物的热量、质量、糖分等指标数据会传输到营养师和保障人员的移动终端设备上,使其能够结合数据分析运动员的饮食情况,并给出第二天用餐的合理化建议。除了自由取餐区,还有针对重点或伤病队员的“定餐区”。营养师会根据运动员当天的体能状况、身体状况和营养短板提前完成配餐,以最大限度保证科学化膳食。基于大数据技术的精准营养系统既保证了运动员膳食结构的科学性,又满足了轻量级运动员控制体重的需要,从源头上解决了运动训练所需的能源储备问题(陈思彤,2021)。

第四,数据驱动最佳竞技表现精准释放。数据驱动精准训练的最终目的是提升训练的综合效能,让运动员在预定时间段(如重大赛事备战和比赛期间)发挥出最佳竞技表现。整合分期理论认为,运动员不仅需要技术、体能的训练,还要在周期中整合心理建设、疲劳管理、损伤康复等训练要素。中国赛艇队使用的综合智能训练辅助系统正是基于整合分期训练理论,围绕运动员5年东京奥运会备战周期中体能、技能、营养、睡眠、生理、心理等训练要素的海量数据展开分析,以此作为运动员参赛选拔、训练周期划分、训练状态诊断、训练计划制定、训练负荷施加、训练过程调控、运动损伤预防、技战术安排的重要参考。随着赛艇国家队联合科研攻关团队将先进数字技术在赛艇竞技训练中的逐渐深度融合应用,赛艇运动员24 h内产生的所有与运动表现相关的要素信息,如体能(身体机能、内外部负荷等)、技术(峰值浆力点、抓水效果、浆频、浆幅因子等)、营养(摄入的碳水化合物、脂肪、蛋白质等)、睡眠(睡眠时长、睡眠阶段分布、睡眠中断频率、睡眠效率等)等都被作为数据源进行全面采集、汇集储存,并通过边缘计算及深度卷积神经网络算法对海量训练数据进行集成、分析及语义化处理,以阶段性、动态化的图表形式呈现运动员竞技状态的变化,发现影响人体适应的

复杂因素与最佳竞技表现形成之间关系,进而识别运动员现有短板,为整体性理解、解释和实现竞技表现提升提供决策依据,最终使运动员的最佳竞技表现在东京奥运会比赛期间精准释放(徐思佳,2021)。

未来,随着大数据、人工智能等新一代数字技术的迅速发展,数据驱动精准训练覆盖的运动项目日趋增多,并且将从竞技训练延伸到全民健身领域;多源数据采集将更加丰富和便捷,能够实现训练学、动力学、生理学、心理学等结构化、半结构化及非结构化数据的动态采集、无线传感、同步显示、即时分析;从机器学习到深度学习,从一般数据解释到可视化分析,智能分析理论与模型的应用将日益完善和科学;与运动训练实践需求的联系将日趋紧密,从仅关注运动训练中动作模式改善、动作技术优化、训练负荷监控、疲劳损伤预警等单一问题向关注“运动中的人”的多维综合分析发展。需要注意的是,当前数据驱动在运动训练中的应用还处在对“小数据”和“局部信息”的挖掘,虽然这类“小数据”驱动能对精英运动员的运动表现提升起到积极作用,但从全局和长远来看,由训练大数据生成智能科学训练决策的数据驱动精准训练范式无疑对竞技体育生态圈的健康可持续发展具有重要战略意义。

### 3 数据驱动精准训练的实现框架

基于对上述数据驱动精准训练支撑理论及实践示例综合分析,认为数据驱动精准训练的实现框架是借助由海量训练数据模型、深度学习算法、高度计算力等智能化技术与算法融合构建的“精准训练智能大脑”,在时空结构上对复杂动态的运动训练过程进行规律化认识和表述,并对竞技训练过程加以优化决策和预测干预。“精准训练智能大脑”是物联网大脑架构的产物,基于“应激-适应”“整合分期”等训练理念,以信息科学为指引,利用云反射弧链接训练主体、训练客体与训练环境的类脑神经网络,将多源异构的海量训练数据汇聚起来,并采用智能算法和前沿技术建构的智能训练中枢系统。具体而言,“精准训练智能大脑”在本研究中特指数据驱动精准训练服务平台,主要包括训练数据中台、智能训练干预引擎、服务运行模式等核心环节(图3)。该平台顺应数据集成处理、系统分析、智能应用的理论逻辑对基础训练数据进行实时追踪、智能处理、结构化组织与意义生成,并发挥数据挖掘对竞训过程的描述、解释、预测和干预功能,形成竞技训练服务新样态。

#### 3.1 训练数据中台

数据中台是一个面向混合式环境的数据汇聚中心,实现资源和业务互通,并提供高效的数据服务。训练数据中台是数据中台在运动训练领域中的应用,是为了满足竞技训练的特定化需求,对全域训练数据进行采集处理、



优化集成,使之成为可理解、可使用、可管理的数据资产,进而服务于竞技训练生态体系,特别是在运动员全息画像、运动训练与专家诊断模型构建、运动表现预测、运动损伤预警、训练负荷干预等方面具有基础保障性作用,使构建“精准训练智能大脑”成为可能。其价值体现在:1)实现数据管理标准化、规范化和模式化,数据中台可以

将多源、异构的数据进行标准化存储,消除异构数据源,实现数据资产的统一管理,便于数据的公开、共享;2)提升数据决策精准化、个性化和智能化,数据中台可以实现数据与业务的集成管理和相互赋能,实现训练数据从无形资产到有形价值的转变;3)关注数据治理,强调训练大数据服务化和安全性。

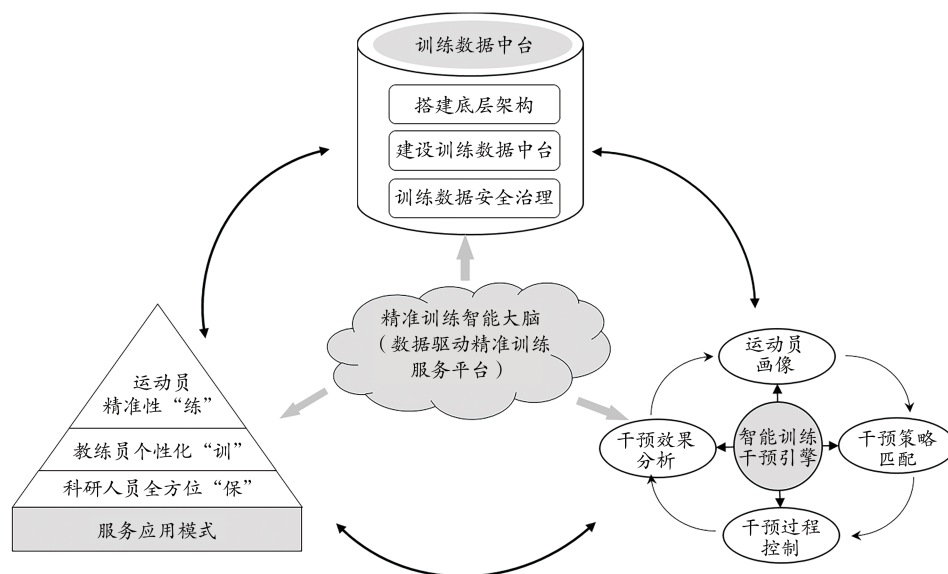


图3 数据驱动精准训练的实现框架

Figure 3. The Implementation Framework of Data-Driven Precise Training

训练数据中台包括搭建底层架构、建设训练数据中台、训练数据安全治理等关键技术,具体架构如图4所示:1)搭建底层架构是结构化基础,主要借助数字技术和生化技术建成训练环境,通过自动采集与手动采集相结合的方式完成对竞技训练全过程中运动员机体的内隐数据和外显数据的多元化、连续性采集。2)训练数据中台建设是核心环节,是对数据进行集成、分析以及语义化的过程。首先,需基于不同来源、不同方式集成多元化数据,构建全量数据仓库;其次,要结合实际训练中的应用需求,通过智能算法、云计算等技术对数据仓库中的集成数据进行深度处理和分析;最后,可通过数据可视化、虚拟现实等技术将大数据分析的结果以直观的、便于理解的形式展示给运动训练主体,以赋能竞技训练个性化、高效式发展。3)训练数据安全治理是重要保障,可以通过制定执行严格的数据开放共享标准、流量管控等技术方法保障数据的安全流通,加强对运动员数据隐私安全的保护,以构建安全可信、融汇互通、良性循环的竞技训练生态圈。

### 3.2 智能干预引擎

智能训练干预引擎是一个以“运动员画像—干预策略匹配—干预过程控制—干预效能分析”为核心的循环结构干预模型,能够为数据驱动精准化、个性化、智能化训练提供有力支持。一方面,智能训练干预引擎可以汇聚

海量优质训练资源,并借助人工智能、云平台等先进数字技术构建一个科学化、高质量的运动训练专家知识库;另一方面,智能训练干预引擎通过对运动员外显行为和内隐特征数据的测试评估、动态监测和分析挖掘,实时把控运动员的真实训练情况,给予运动员及时、动态、精准、个性化指导。具体而言,智能训练干预引擎的循环结构主要包括4个环节:1)运动员画像是对运动员现实状态的准确识别和判断,是智能训练干预的起点。其通过对运动员机体多层次(分子、细胞、器官、组织、系统)、多因素(体能、技战术、心理、营养等)、多阶段数据的系统分析和综合诊断,识别出关键特征,构建标签化运动员模型,概括出运动员和运动队的信息全貌,便于找准训练干预的立足点和发力方向。2)干预策略匹配使训练决策过程从经验走向数据智能驱动,始终指向更加优化、更具效能的训练方案,是保证智能训练干预效果的关键。其依据运动员实时动态画像的关键表征,采用多重匹配算法,从运动训练专家知识库中提取出与运动员现实状态匹配度最高的干预方案或优化策略。因此,智能训练干预引擎的关键在于不断丰富训练专家知识库中的干预方案,以满足不同运动员以及同一运动员不同阶段的差异化、个性化训练需求。3)干预过程控制强调训练应激源(负荷刺激)的适宜性和可量化,使训练过程控制从定性走向定量,从宏观走向微观,始终指向竞技能力的提高、运动损伤和过



度训练(或训练不足)的避免、最佳运动表现的实时出现。其在干预策略最佳匹配的基础上,进一步将训练策略细分为可在训练过程控制中量化控制、持续跟踪的具体训练指标,并在训练中对跟踪指标进行即时比对,一旦发现偏差变化便会通过实时反馈功能提醒用户调节优化、纠正偏差,避免无效训练,提高训练的质量效益。4)干预效果分析是对已实施的智能训练干预过程的阶段性效果进

行测试评估的过程,这意味着前一阶段智能训练干预过程的结束和新一轮智能训练干预过程的开始,依次循环往复。干预效果分析使训练评价的主体和指标转向“去中心化”的多元协同机制,避免“一言堂”和“仅凭运动成绩说话”的单一式评价,可以真正总结和发现训练过程中的有效经验和实际问题,进而促进智能训练干预引擎的不断修正和持续优化。

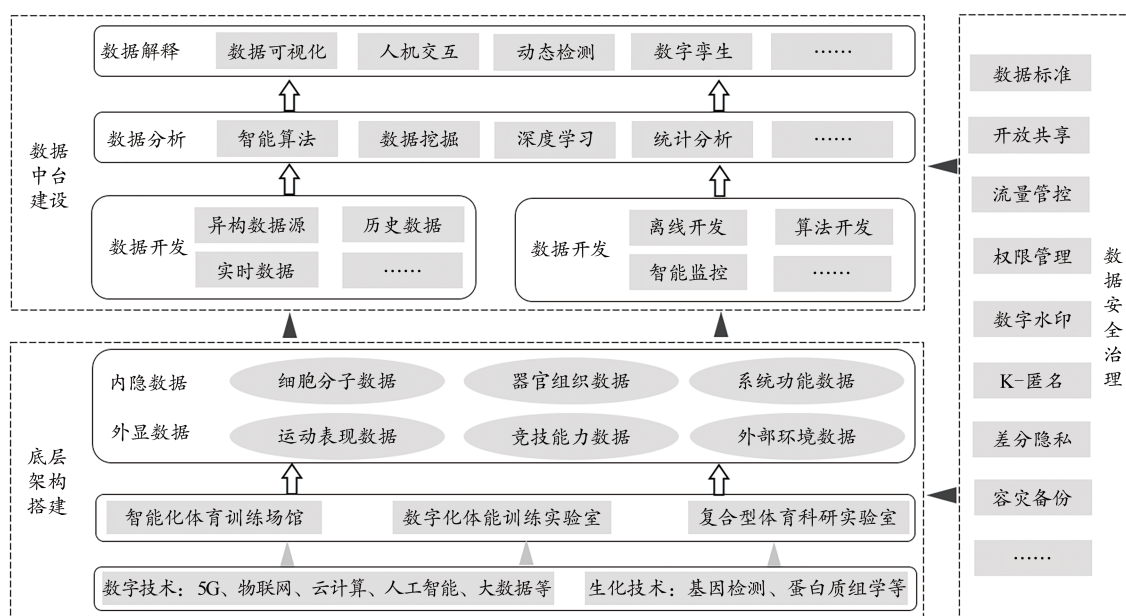


图4 训练数据中台架构体系

Figure 4. The Architecture System of Training Data Center

### 3.3 服务运行模式

建设训练数据中台和智能训练干预引擎的最终目的是让训练大数据真正地为竞训实践服务。本研究基于数据驱动决策理论、应激-适应理论、整合分期理论、运动训练过程控制理论以及由这些理论衍生出来的计算模型,提出了数据驱动精准训练的服务运行模式。整个运行模式需依托训练数据中台和智能训练干预引擎,通过“智控”+“人控”相结合的方式,实现对竞技训练过程中8大核心关键环节的全程性、立体化、剂量式的微观和滴定控制(图5)。

1)精准诊断起始状态是数据驱动精准训练的逻辑起点,利用大数据技术对上述影响运动员训练水平的内部和外部起始状态进行数据化处理,并概括和提炼出与运动员竞技能力和参训水平相关的信息全貌,进而为后续训练环节的精准控制提供客观依据。诊断的目的是生成给“运动员画像”;诊断的意义是为了制定符合运动员现状的训练方案,降低损伤风险;诊断的内容包括运动员机体的内部起始状态(机体的生理、心理准备状态)和外部起始状态(影响运动员训练质量的训练环境、自然和社会环境状况)。

2)精准设定训练目标是数据驱动精准训练的方向指引,旨在让运动员经过严格的体能、技能、战术、心理、营养、恢复等方面的训练后,使其竞技能力的整体功能按运动项目的制胜要求定向演变。利用大数据技术制定训练目标的优势在于可以结合精准起始状态诊断下的“运动员画像”,运用最近发展区理论,为运动员设定高度契合自身能力的可量化结果预期,避免训练目标设定的盲目性、模糊性和同一性,促进训练目标从抽象到具体、从定性到定量、从共性到个性的转变,以更好地发挥训练目标的导向、激励和评价作用。

3)精准定制训练内容可将训练目标任务化,并赋予其操作性。运动员竞技能力结构的非平衡性特征以及不同项目对各项竞技子能力的非平衡性需求,导致了训练内容安排的复杂性和多变性,使得某项训练内容可能是有效的,但不一定是最佳的。而精准定制训练内容回答的就是“有效训练”与“最佳训练”的矛盾问题,其利用信息处理和人工智能分析技术,帮助运动训练主体在纷繁复杂的训练内容中定制一套提升运动员竞技能力整体功能的最佳方案。

4)精准选用训练方法的目的是最大程度释放运动训

练内容的功效, 强调训练方法在选择和使用两个方面的精准性。从选择的层面来看, 精准选用训练方法就是结合运动员个体的起始状态和目标需求, 通过分析海量的训练数据来深度剖析不同训练方法之间的本质区别和联系, 进而帮助运动训练主体选择最适宜的运动训练方法。

从使用的层面来看, 精准应用训练方法可以将训练过程中机体的状态变化通过生理生化、动作速率与轨迹等反映训练质量的参数指标进行量化和可视化, 进而实现对运动员在训练中的即刻生理、心理适应水平进行动态监测和细分控制。

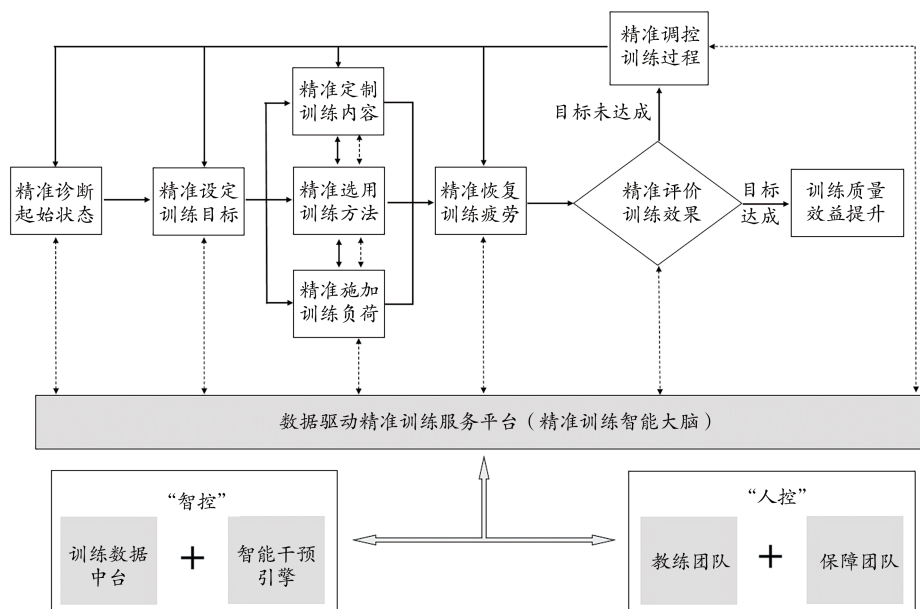


图5 数据驱动精准训练的服务运行模式

Figure 5. Service Operation Mode of Data-Driven Precise Training

5) 精准施加训练负荷是利用竞技能力系统中的临界效应和阈值现象, 充分发挥负荷刺激对提升训练质量效益的杠杆作用。在大量的训练实践中, 运动员的实际承载负荷很难始终与其当时的需求负荷实现完美匹配。而精准施加训练负荷可以利用大数据技术、运动表现分析技术以及生理生化监测技术手段对项目的负荷特点、机体的耐受能力、训练的外部环境等因素进行量化分析和综合考虑, 将负荷刺激下机体发生的形态和功能变化清晰而直观地呈现给运动训练主体, 帮助教练员根据竞技能力系统内部相互作用的阈值设定达到阈强度的训练负荷, 实现对机体深度有效的阈刺激, 从而充分发挥训练负荷的杠杆作用。

6) 精准恢复训练疲劳的关键在于如何最优化解决消耗与补充、训练与间歇之间的动态平衡问题。精准恢复训练疲劳能够将大数据技术与先进的疲劳监测、耗能评估、能量摄入、疲劳缓解等技术有机结合起来, 以确保训练疲劳被适度滴定, 并在系统分析机体疲劳产生的原因(神经系统、能量消耗、物质代谢等)、类型(外周疲劳、中枢疲劳)及部位(不同的肌肉群)的基础上, 对营养摄入、间歇时长、恢复手段进行定量控制, 进而提升运动员的竞技表现和对训练负荷的适应效能。

7) 精准评价训练效果是对竞技实践结果满足运动训

练主体需要的程度进行准确判断的活动, 其最重要的目的不在证明, 而在改进。通过数据挖掘、深度学习等技术可以对训练过程现实的或潜在的价值进行实时分析、诊断、评价与反馈, 帮助运动训练主体发现训练实践中的成效和问题。根据冰山理论, 训练效果有显性和隐性之分, 显性的训练效果可以根据运动训练主体的直觉来判断, 但其所占比重很小, 而精准评价训练效果的优势在于可以充分发挥数据的描述、解释和预测功能, 对占更大比重的隐性训练效果进行挖掘和呈现, 并加以综合评价, 为后续精准调控训练过程提供客观依据。

8) 精准调控训练过程是指在精准评价前一分期阶段训练效果的基础上, 以精准设定的训练预期目标为参照标准, 对数据驱动精准训练过程中的前6个环节进行针对性干预, 直到精准设定的训练目标达成, 训练的质量效益满足了运动训练主体的要求并呈现良性向好的发展趋势。竞技训练的多因素(机体内部的生理心理适应水平、运动员的生物动作能力、外部训练环境)、多周期(单峰、双峰、多峰)、多阶段(训练分期的基础准备、专项准备、赛前减量、转化实现、调整过度; 机体适应的报警、抵抗、疲劳、恢复)、多结果(训练不足、最佳适应、过度训练、运动损伤)特征决定了精准调控训练过程的必要性, 且调控时机和调控内容把握越精准, 越有利于训练质量效益的提升。



#### 4 数据驱动精准训练的推进路径

“无数据、不训练”“无监控、不训练”已成为高水平精英运动员竞技能力提升的基本准则(闫琪等,2018),必须加快推进运动训练模式由经验驱动向数据驱动转型,从构建数据驱动精准训练平台、提升训练主体数据素养、对接运动项目应用等方面着手,让数据驱动精准训练赋能竞技训练高质量发展的愿景更快实现。

##### 4.1 铺好“奠基石”:构建数据驱动的精准训练服务平台

构建基于数据驱动精准训练服务平台,是开展数据驱动精准训练的“奠基石”,该平台需在具备泛在数据采集与融合功能的基础上,提供多源数据挖掘与决策服务,并能够实现动态数据的实时监测与可视化。

###### 4.1.1 具备泛在数据采集与融合功能

“泛在”(Ubiquitous)即广泛存在,泛在数据采集是指对任何人或任何物在任何时间和任何地点所产生的数据进行收集和储存的过程。泛在数据收集大大增强了无线传感器网络部署的灵活性,不需要依靠固定式接收器收集整个网络的传感器数据,而是利用移动设备(如掌上电脑或智能手机)从周围传感器中进行泛在数据采集。数据驱动精准训练服务平台的泛在数据采集与融合功能可以突破时空的限制,实现对训练中的全员、全程、全方位数据采集,为精准训练提供海量的数据支持。“全员”是指数据采集的对象涵盖个人项目的单个运动员和集体项目的团队运动员以及包括教练员在内的其他辅助人员;“全程”是指数据采集的时间覆盖训练过程的准备阶段、实施阶段、调整阶段;“全方位”是指数据采集的范围不仅局限于运动中的人,还覆盖运动所处的整个外部环境。

###### 4.1.2 提供多源数据挖掘与决策服务

数据驱动精准训练服务平台的泛在数据采集与融合功能为精准训练提供了海量的数据集,但若平台不能为其提供数据挖掘与决策服务,数据资源将会沦为“数据垃圾”。大数据的本质是通过数据分析历史、解释现在、预测未来,数据驱动精准训练服务平台的重要职能之一就是利用大数据技术对海量多源的训练数据资源进行深度研读,挖掘数据之间的内在关联,归纳其本质、探究其规律,揭示出超越经验和传统的认知,发现新的训练规律和运动员竞技能力的增长特征及演化趋势,进而为训练实践活动提供更具针对性和更加高效的方案决策,优化教练员的时间分配和资源分配,从而推动训练资源供给体制改革,解决优质训练资源的短缺和失衡问题。

###### 4.1.3 支持动态数据实时监测与可视

运动训练过程和运动员的竞技能力系统是一个非线性的、复杂的动态系统,这就决定了运动训练数据具有动态性,静态只是训练数据的一种短时的阶段性特征。因此,数据驱动精准训练服务平台需支持对动态训练数据的实时监测,这将大大提升训练数据采集的价值,增强主

体对训练过程控制的精准程度。此外,数据可视化是关于数据视觉表现形式的科学技术,其不仅可以直观表达训练数据的属性和变量,还能够对海量、多元、复杂的训练数据进行有意义的比较,提炼出通过其他方法不易发现的训练信息,弥补表格呈现方式直观性、易懂性不足的问题。数据驱动精准训练服务平台对训练动态数据的实时监测与可视化处理将大幅提升数据的真实性、直观性和可用性,实现了对抽象、宏观的运动训练过程和运动员竞技能力增长模式的具体和微观化处理,进而让运动员的训练更具信心、更有目标。

##### 4.2 清除“拦路障”:聚焦提升运动训练主体的数据素养

在对训练数据价值充分认可的情况下产生的“数据饥渴”与海量训练数据资源暴露之间,横亘着运动训练主体数据素养缺乏的巨大鸿沟,这是当下制约数据驱动精准训练深入发展的一大“拦路障”。因此,提升运动训练主体的数据意识素养、数据处理素养和数据应用素养是确保数据驱动精准训练顺利推进的重要支柱。

###### 4.2.1 提升数据意识素养

运动训练主体包括运动员个人和运动训练团队,他们的数据意识素养主要体现在对数据的敏感性和批判性两个方面。敏感性指的是运动训练主体特别是教练员要积极主动地了解 and 关注大数据相关的基础知识和理论,并高度认同大数据技术在提升运动训练质量效益中的重要地位。同时,还要结合运动训练的规律和需要,在训练实践中主动寻求用数据决策、用数据管理、用数据创新的训练思维。批判性指的是要辩证地认识到对数据的过度依赖可能会带来误导和隐私泄露等不容忽视的问题,要以理性、客观、规范的态度评估数据价值,既要看到数据给运动训练科学化发展带来的机遇,也要看到随之而来的挑战和潜在风险(江小涓等,2020)。

###### 4.2.2 提升数据处理素养

运动训练主体的数据处理素养主要体现在对训练数据的广泛收集和整合分析过程中。就广泛收集而言,对运动员训练数据的广泛采集是实施数据驱动精准训练的重要引擎,而训练数据的广泛采集需要运动员和教练员的积极配合才能完成。因此,他们的数据素养直接影响训练数据能否被采集以及采集的数据是否及时、准确和具有可用性。就整合分析而言,随着训练大数据平台的建立和完善,训练大数据的来源将更加广泛、形式更具多样、内容更显复杂,这就要求运动训练主体特别是教练员必须具备重新解释数据的能力,以及数据筛选、数据转化能力。

###### 4.2.3 提升数据应用素养

运动训练主体的数据应用素养主要体现在利用大数据驱动科学训练决策、精准训练施策和有效训练评策的能力上。数据驱动科学训练决策的能力是指在宏观层面

通过训练大数据来决定“练什么”“怎么练”“恢复什么”“怎么恢复”的问题;数据驱动科学精准施策的能力是指在微观层面通过训练大数据来指导“练多少”“练多久”“恢复多少”“恢复多久”的问题;数据驱动有效训练评策的能力是指从客观角度通过训练大数据来解释“为什么这样练”“为什么这样恢复”的问题。总之,数据应用素养就是将数据意识素养和数据处理素养与训练实践活动进行深度融合,利用大数据驱动来弥补经验驱动在解决训练问题中的缺陷与不足,从而加强对运动训练过程的控制,提升运动训练的科学化水平。

#### 4.3 迈实“试探步”:推进数据驱动精准训练的落地应用

数据驱动精准训练已成为世界各国增强竞技体育国际竞争力的新引擎,是点燃运动训练科学化创新动能的导火索,将会成为一种统领运动训练模式重大变革的全新思维方法。我国要加紧推进数据驱动精准训练落地应用,在加快训练场馆数字化升级、优化运动训练复合型团队、支持重点潜优势运动项目优先试行等方面精准发力,以数字化信息整合助力竞技体育实力全面提升。

##### 4.3.1 加快运动训练场馆数字化改造升级

当今国际竞技体坛激烈竞争的背后实则是科学技术的竞争。近年来,我国部分基础大项在现代科技助力下取得了全面发展和进步(如短跑运动员苏炳添在百米项目多次突破10 s大关并创造了亚洲纪录),加剧了体育工作者对训练大数据、智能化训练场馆、数字化体能实验室、复合型体育科研实验室的渴望和需求。加快运动训练场馆数字化改造升级,打造一批具有动作捕捉反馈和多重处理、力量和速度等体能指标定量评估、运动机能无创监测、运动营养和微量元素补充、训练后系统恢复、运动伤病预防和干预、运动员学习和教练员智能管理、大数据挖掘和云端共享等功能的智能化科学训练基地是科技助力奥运备战发展的必然要求,也是数据驱动精准训练最直接的落地形式。

##### 4.3.2 优化运动训练复合型团队工作模式

数据驱动精准训练是一项以团队形式综合专门人才,协同配合、精准发力的复杂系统工程,单纯依靠专项技战术教练员个体“包打天下”的传统训练模式无法适应数据驱动精准训练的运作要求。数据驱动精准训练落地应用需要运动训练方法专家(运动生物学和训练相关专家)、运动心理学专家、运动营养学专家与数据分析专家之间的协同配合。我国当前的运动训练复合型团队存在运行机制不完善、工作效能不高、专业人才缺乏、成员职权不清晰、团队建设缺乏延续性等问题,传统意义上的领队或教练无论在知识储备还是能力结构方面,均已不能胜任当前对复合型团队负责人的新要求。目前,一些国家创新了复合型团队的管理模式,团队负责人由原来的主教练变为“竞技表现主管”,专门负责运动队训练和管理的

协调工作,大大提升了数据驱动精准训练的应用效能,我国可以结合具体实际加以学习和借鉴(杨国庆等,2020)。

##### 4.3.3 支持重点潜优势运动项目优先试行

《体育强国建设纲要》强调,要在保持传统优势项目领先地位的基础上,做大做强基础项目;恶补短板,不断提高冰雪竞技水平。这是数据驱动精准训练助力体育强国建设的出发点和着力点。为实现上述目标,数据驱动精准训练在推进过程中应该在项目布局上做好整体规划,转变以优势项目为中心的奥运战略布局,综合评估不同竞技体育项目发展潜力和价值,支持数据驱动精准训练模式在“三大球”“体能主导类项目”“冰雪项目”等重点潜优势项目上优先试行推广,助力奥运会等国际大赛中影响力较大的项目在成绩上实现突破,推动潜优势项目和弱势项目不断向优势项目转化。以篮球项目为例,可利用政策杠杆构建政府主导、部门协同、高校保障、社会力量积极参与的“数据驱动篮球项目精准训练示范工程”,运用政策激励、资金补贴、人才保障等措施调动中国男子篮球职业联赛各俱乐部试行数据驱动精准训练的积极性。

## 5 结语与展望

### 5.1 结语

数据驱动精准训练作为一种运动训练理论与方法的新范式,是运动训练实践经验从量变积累到质变飞跃的必然转型,也是提升运动训练科学化水平、增强竞技体育综合实力、实现体育强国战略目标的必然选择。未来有必要联合行政管理层、运动训练界、体育科学界、数据科学界等各方力量进行集中论证和研讨,并充分发挥中国特色的运动训练体制机制的优势,尽快搭建全样本、全过程、全景式的训练大数据共享平台,加强训练大数据采集便携化、分析准确化、传输可视化的研究,创新数字化、智能化训练的应用场景,使数据驱动精准训练的重大价值得以更快显现和更大释放。

### 5.2 展望

运动训练融合创新和精准个性化调控的时代趋势不可阻挡,运动训练过程控制亟需也必将会在数据驱动的支持下变得越趋精准和科学。但是,随着数据科学和大数据技术与运动训练实践的深度融合,运动训练活动对主体的参与能力提出了新的更高要求,相关主体必须要有战略眼光和超前意识,主动迎接和科学应对大数据时代背景下运动训练科学范式变革带来的机遇和挑战。

1)数据驱动精准训练有利于推动“主教练包办”的传统训练方式向“分工统筹、流水作业”等一系列产业化工业模式转变。但其在实践中应树立人本理念,明确对运动员的训练不是“精密数控机床”,不存在完全统一的规程标准。因此,对个体训练负荷控制和竞技状态调整的



精度如何把握和界定是影响数据驱动精准训练成效的关键问题。

2)数据驱动精准训练无疑将全面提升运动训练的科学化水平,但也容易导致数据沉溺或数据依赖,滋生唯数据主义。唯数据主义推崇数据流量最大化与数据自由至上,致力于将运动员的一切生命活动数据化,这虽然有益于运动训练过程控制的精准性,但若不能充分保护个人数据权力,就可能出现侵害运动员人格权、侵犯隐私和歧视行为的发生。

3)以提高竞技能力、降低损伤风险、适时出现最佳运动表现为根本目的的竞技训练是一个长期、动态和复杂的非线性系统,训练负荷作为竞技训练质量效益提升的关键,会引起运动员生理和心理上的应激。当前,人们对生物体的了解无论在认知层面还是技术层面均远不足以达到对机体运动进行全面和精准控制的要求,相关的理论和实践研究有待引起广泛重视。

4)大数据给运动训练的科学研究带来了新的洞察力,有助于训练质量效益的提升。然而,我国运动训练的领域的的数据资源虽十分丰富,但数据之间的交叉融合不够,“数据孤岛”问题普遍存在。加之运动训练工作者普遍缺乏数据素养,既有运动训练科学知识又有专业数据研判和分析能力的复合型人才更为稀缺,这已成为制约数据驱动精准训练模式高质量发展的一个亟待解决的现实问题。

#### 参考文献:

蔡玉卿,2019.大数据驱动式社会监督:内涵、机制与路径[J].河南社会科学,27(8):52-58.

陈思彤,2021.科技为竞技体育插上翅膀[N].中国体育报,2021-07-28(05).

陈小平,2018a.从师徒传技到奥运攻关:对运动训练集成化、科学化的思考[J].体育科学,38(7):14-15.

陈小平,2018b.科技助力奥运训练:形势、进展与对策[J].体育学研究,1(1):76-82.

陈小平,2020.当代运动训练经典理论与方法[M].北京:人民体育出版社:32-59.

杜长亮,胡海旭,仇乃民,等,2019.大数据背景下运动训练研究范式探索[M].北京:中国海洋大学出版社:34-60.

果青,2021.戴尔科技如何通过AI助力国家赛艇队训练?技术人员揭秘幕后故事[EB/OL].[2021-10-30].<http://www.techweb.com.cn/internet/2021-10-12/2860177.shtml>.

胡海旭,2016.竞技能力增长理论模型及其演进[J].体育科学,36(2):14-24,40.

胡海旭,金成平,2021.智能化时代的个性化训练:机器学习应用研究进展与数字化未来[J].体育学研究,35(4):9-19.

江小涓,李姝,2020.数字化、全球化与职业体育的未来[J].上海体育学院学报,44(3):1-16.

李爱霞,舒杭,顾小清,2021.打造教育人工智能大脑:教育数据中台技术实现路径[J].开放教育研究,27(3):96-103.

黎涌明,纪晓楠,资薇,2014.人体运动的本质[J].体育科学,34(2):

11-17.

黎涌明,李博,王欣欣,2018.男子赛艇240次卧拉练习的能量代谢特征[J].上海体育学院学报,42(6):104-108,118.

茅洁,2011.我国赛艇项目特点及制胜规律的研究[D].北京体育大学.

徐思佳,2021.科技改变运动,中国赛艇夺金背后的科技秘密[EB/OL].[2021-10-30].<https://new.qq.com/omn/20210803/20210803A00HIV00.html>.

闫琪,廖婷,张雨佳,2018.数字化体能训练的理念、进展与实践[J].体育科学,38(11):3-16.

杨国庆,2020a.整合分期:当代运动训练模式变革的新思维[J].体育科学,40(4):3-14.

杨国庆,彭国强,戴剑松,等,2020b.中国竞技体育复合型训练团队的发展问题与创新路径[J].北京体育大学学报,43(06):10-19,34.

袁守龙,2018.体能训练发展趋势和数字化智能化转型[J].体育学研究,1(2):77-85.

ACKOFF R L, 1989. The circular organization: An update[J]. Acad Manage Exec (1987-1989),3(1):11-16.

ANDERSON J L, 2018. Quantification of physical loading, energy intake and expenditure In English Premier League Soccer Players [D]. Liverpool: Liverpool John Moores University.

ARNE G, STEPHEN S, EIKE E, 2009. Training methods and intensity distribution of young world-class rowers[J]. Int J Sports Physiol Perform, 4(4):448-460.

BURKE E R, 1998. Precision Heart Rate Training[M]. Champaign, IL: Human Kinetics:7-20.

ERIKSSON R, NICANDER J, 2021. Automated generation of training programs for swimmers generating weekly training plans in the style of a professional swimming coach using genetic algorithms and random trees[D]. Goteborg, Sweden: Chalmers University of Technology.

FISTER I, FISTER J I, FISTER D, 2019. Computational Intelligence in Sports[M]. Berlin: Springer.

HALSON S L, 2014. Monitoring training load to understand fatigue in athletes[J]. Sports Med, 44(Suppl 2): S139-S147.

IMPELLIZZERI M F, MARCORA S M, COUTTS A J, 2019. Internal and external training load: 15 years on[J]. Int J Sports Physiol Perform,14(2):270-273.

JACK R A, LINTNER D M, HARRIS J D, et al., 2019. Omega-wave: An emerging technology and application in professional baseball pitchers[J]. J Sports Med Phys Fitness, 59(7): 1144-1149.

JOHANN W, 2017. How do training and competition workloads relate to injury?The workload-injury aetiology model[J]. Brit J Sports Med, 51(5):428-435.

KILDUFF L P, FINN C V, BAKER J S, et al., 2013. Preconditioning strategies to enhance physical performance on the day of competition[J]. Int J Sports Physiol Perform, 8(6):677-681.

KUZMITS F E, ADAMS A J, 2008. The NFL combine: Does it predict performance in the National Football League?[J]. J Strength Cond Res, 22(6): 1721-1727.

MALONE S, COLLINS K, MCROBERTS A, 2021. Understanding the association between external training load measures and injury risk in Elite Gaelic football[J]. J Sports Med Phys Fitness, 61(2): 233-243.

- MAURO M, RAFAEL O, NUNO L, 2021. Load measures in training/match monitoring in soccer: A systematic review[J]. *Int J Environ Res Public Health*, 18(5):2721.
- MCGUIGAN M, 2017. *Monitoring Training and Performance in Athletes*[M]. Champaign, IL: Human Kinetics Inc.
- MCINTYRE A, BROOKS J, GUTTAG J, et al., 2016. Recognizing and analyzing ball screen defense in the nba[C]//22<sup>nd</sup> ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD). Boston, MA, USA: Association for Computing Machinery: 11-12.
- MORTON R H, 2006. The critical power and related whole-body bioenergetic models[J]. *Eur J Appl Physiol*, 96(4):339-354.
- NIBALI M L, TOMBLESON T, BRADY P H, 2015. Influence of familiarization and competitive level on the reliability of counter-movement vertical jump kinetic and kinematic variables[J]. *J Strength Cond Res*, 29(10):2827-2835.
- NUGENT F J, FLANAGAN E P, WILSON F, et al., 2020. Strength and conditioning for competitive rowers[J]. *Strength Cond J*, 42(3):6-21.
- PERL J, 2003. On the long-term behaviour of the performance-potential-metamodel PerPot: New results and approaches[J]. *Int J Comput Sci Sport*, 2(1):80-92.
- PERL J, 2004. PerPot-a meta-model and software tool for analysis and optimisation of load-performance-interaction[J]. *Int J Perform Anal Sport*, 4(2):61-73.
- RAGO V, BRITO J, FIGUEIREDO P, et al., 2020. Internal training load monitoring in professional football: A systematic review of methods using rating of perceived exertion[J]. *J Sports Med Phys Fitness*, 60(1): 160-171.
- RAJSP A, FISTER I F, 2020. A systematic literature review of intelligent data analysis methods for smart sport training[J]. *Appl Sci*, doi:10.3390/app10093013.
- SCHELLING X, FERNANDEZ J, WARD P, et al., 2021. Decision Support System Applications for Scheduling in Professional Team Sport. The Team's Perspective[J]. *Front Sports Act Living*, 3:678489.
- SELYE H, 1950. Stress and the general adaptation syndrome[J]. *Brit Med J*, 1(4667):1383-1392.
- SMPOKOS E, MOURIKIS C, LINARDAKIS M, 2018. Seasonal physical performance of a professional team's football players in a national league and European matches[J]. *J Hum Sport Exerc*, 13(4):720-730.
- STANISLAV E P, 2020. Performance of striking techniques among qualified Muay Thai athletes of different weight classes[J]. *Int J Perform Anal Sport*, 20(2):294-304.
- THIERRY B, 2003. Variable dose-response relationship between exercise training and performance[J]. *Med Sci Sports Exerc*, 35(7): 1188-1195.
- BOMPA T O, BUZZICHELLI C, 2018. *Periodization-6th Edition: Theory and Methodology of Training*[M]. Champaign, IL: Human Kinetics.
- VEUGELERS K R, YOUNG W B, FAHRNER B, et al., 2016. Different methods of training load quantification and their relationship to injury and illness in elite Australian football[J]. *J Sci Med Sport*, 19(1):24-28.
- VELLOSO E, BULLING A, GELLERSEN H, et al., 2013. Qualitative activity recognition of weight lifting exercises[C]//The 4<sup>th</sup> Augmented Human International Conference. Stuttgart, Germany: ACM: 116-123.
- WEISS K J, ALLEN S V, MCGUIGAN M R, et al., 2017. The relationship between training load and injury in men's professional basketball players[J]. *Int J Sport Physiol*, 12(9):1-20.
- WLADAWSKY-BERGER I, 1999. Turning points in information technology[J]. *IBM Syst J*, 38(2/3):449-452.

## The Theoretical Connotation, Implementation Framework and Promotion Path of Data-Driven Precise Training

ZHONG Yaping<sup>1</sup>, WU Zhangzhong<sup>1</sup>, CHEN Xiaoping<sup>2, 3\*</sup>

1. *Sports Big-Data Research Center, Wuhan Sports University, Wuhan 430079, China;*

2. *China Institute of Sport Science, Beijing 100061, China;*

3. *College of Physical Education, Ningbo University, Ningbo 315211, China*

**Abstract:** “Data-driven” and “precision training” are the main theme and high-frequency words in the field of sports training, but systematic and in-depth research is needed in this topic. This study reviewed and explored the theoretical connotation, implementation framework and promotion path of data-driven precise training, it is suggested that: 1) data-driven precision training is a new paradigm of sports training theory and method which has the distinctive characteristics of precise load, precise alert of injury and precise evaluation; 2) the implementation framework of data-driven precise training is based on the construction of a training data platform and the development of an intelligent training intervention engine, with the help of “precise training of the intelligent brain” and through “intelligent control” + “human control”, it can control the eight key links in sports training; 3) for the advancement and application of data-driven precise training, three aspects of platform construction, data literacy and application should be highlighted, and then to form a new development pattern of sports training.

**Keywords:** *data-driven; precise training; training data; support of sci-tech; preparation for Olympic Games; competitive sports*