



(内部刊物)

足球科技信息

Football Science and Technology Information

中国足球协会技术部 主办

2019年 第4期 总第5期

足球科技信息

Football Science and Technology Information

中国足球协会技术部 主办

2019 年 7 月

《足球科技信息》编委会

主 编：蔡 勇

副主编：克瑞斯、郭炳颜、王景波

责任编委：王景波

编 委：费 建、孙 雯、肇俊哲、郑超勇、张 昭

尹煜华、赵英杰、陈庆云、井伟祯、张宇宁、席子杰

足球科技信息

Football Science and Technology Information

第4期(总第5期)2019年7月

目 录

资讯

如何更好地量化足球?	1
大数据是足球的未来	3
足球科技与创新现状——佩德罗·马奎斯(本菲卡).....	5

理论探讨

自由度还是奴役度?	7
巴塞罗那足球俱乐部:我们的新研究如何帮助解开“巴萨之路”	11
利物浦的压迫系统	13

技术报告

2018年国际足联U-17女足世界杯技术报告	23
------------------------------	----

大数据

足球运动员在精神压力下运动表现能力的量化分析	30
训练计划混淆了急慢性工作负荷比与受伤的关系	54

学术论文

精英青少年足球人才选拔中多维度测评与教练评估的有益协作	58
在足球比赛中有赢得比赛的最佳阵容年龄?	74

如何更好地量化足球？

国际足联

与以往全面量化足球比赛相比，现在有大量资源可以提供更多的数据来描述和分析足球比赛。为了消除新的争议，国际足联希望帮助利益相关者更好地了解其中一些关键统计数据。



尤其像《点球成金》这类电影，已经尝试将包括足球在内的每项体育运动变成数字游戏。随着视频质量和自动图像处理技术的提高，还有大量资源用于赛后分析，足球世界获得的统计数据比我们所能注意到的更多。由于媒体、广播公司、球队和广大观众能够获取这些数据，数据科学开始关注足球“大数据”，国际足联对这些信息有效性的现状进行了评估。

量化足球比赛的一种最常用的方法就是“事件数据”。简单来说，就是将比赛分

成一系列“事件”，例如传球、射门、进球或铲球，从而在技术上对比赛每一时刻所发生的事情进行描述。通常，根据所提供信息的深度，会将比赛分成几千个事件。然后，这些事件可以用于对球员、比赛或整个联赛进行统计分析，以及对球队、赛季或球员之间进行具体比较。

最近在2018年国际足联世俱杯比赛期间开展的一项内部研究对大量主要提供商进行了比较，深入分析了各种数据集的有效性、可靠性和可比性。因此，国际足联

通过提醒强调某些趋势：

1. 即使是最基本的事实决策，提供商量化匹配的方式仍然存在差异。虽然进球数量趋于一致，但是任意球、角球甚至黄牌的数量并不总是准确的。

2. “成功铲球”或“完成传中”等更多主观指标依赖于一定程度的动作定义，显示出提供商之间的差异超过 50%。

3. 由于业务的性质，几乎所有数据提供商仍然严重依赖于人工处理。虽然良好的提供商具有非常高的实时准确度，但大

部分数据可验证的速度不够快。

4. 数据的质量控制仍处于起步阶段，例如通过新技术进行质控。

以上问题表明，对这些数据产生一定程度的怀疑是合理的，特别是在比较不同的赛事或一段时间的数据时。虽然尽可能快速地获得尽可能多的数据的愿望是可以理解的，但是确保这些数据准确反映比赛仍然是应尽的责任。这一点在制定表现指标或尝试判断不同数据集之间的球员表现时变得尤为重要。

大数据是足球的未来

西蒙·奥斯汀



塞维利亚体育总监蒙奇表示：“大数据是足球的未来”，因而工程师和物理学家的需求比球探更迫切。

在罗马并不愉快的两年经历之后，2019年4月1日，蒙奇这位西班牙人重新回到了他曾效力17个赛季的俱乐部。他说俱乐部以及比赛整体与他在2000年开始工作时的情况截然不同。

他告诉 EPC Tracker：“大数据和人工智能是任何想要成长的俱乐部都会提到的两个名词”。最近，我们开始找的不再是球探，而是工程师、数学家、物理学家和统计学或算法方面的专家。

“大数据是足球的未来。不是因为你打算根据数据购买球员，而是因为它可以降

低风险。”

大数据是指可用于揭示模式和趋势的非常大的数据集。利物浦有一支四强研究团队，由伊恩·格雷厄姆 (Ian Graham) 领导，专门从事研究和开发、建模、机器学习和分析工作。

团队成员拥有哈佛大学和剑桥大学等机构的数学和物理学博士学位。去年，阿森纳聘请了俄罗斯数据科学家米哈伊尔·希尔金 (Mikhail Zhilkin)，他拥有物理和数学硕士学位，之前曾参与开发过“糖果传奇”游戏。

这与 21 世纪初的情形相去甚远，当时还使用锥桶来计算距离和速度。

蒙奇回忆道：“回想起 2000 年，我们通过放置一些锥桶来计算必须覆盖的距离”，我不得不打电话给提供者告诉他对巴塞罗那对阵马德里竞技的比赛进行录像，因为在五周之后我们将与他们对阵。他将视频发给了我。

或者，当我给我在拉科鲁尼亚工作的表弟打电话，让他帮忙了解拉科鲁尼亚队

时，他会将资料邮寄给我。你将收到 15,000 张纸质文件，非常占用空间。

现在，在训练中，你在球员后背放一个 GPS 就能测量一切。在训练课程结束两分钟后就能获得数据。以同样的方式，在虚拟平台中，瞬间就能获得球员的图像和数据。

在与同事共享的服务器上，我们可以同时查看文档和报告。我可以派一名教练到秘鲁，并在比赛开始 30 分钟后获得报告。”

足球技术和创新现状——佩德罗·马奎斯（本菲卡）

ST Staff



佩德罗·马奎斯是本菲卡俱乐部的技术总监。他于2018年7月加入俱乐部，负责青少年足球的技术协调，致力于将年轻球员培养成未来的明星。在加盟本菲卡之前，马奎斯在曼城度过了八年，他帮助曼城队建立了执教方法论和比赛风格。在那段时间里，这支英超球队赢得了所有四项国内冠军，包括三个联赛冠军。在城市足球集团于2014年成立后，他帮助协调将俱乐部的比赛模式扩展到其他五个国家的姊妹队。

本菲卡是三支从未降级过的葡超联赛球队之一。这家位于里斯本的俱乐部赢得了创纪录的 79 个国内奖杯，其中包括 36 个葡超联赛冠军和 26 个杯赛冠军。本菲卡最近赢得了 2016/2017 赛季的两项冠军。俱乐部还在 20 世纪 60 年代赢得了两次欧洲联盟杯（联盟杯是欧洲冠军联赛的前身）。

根据马奎斯的观点，过去十年间足球技术的最大进步是“技术资源和相关服务变得更容易获得，且成为主流”。计算机和笔记本电脑的价格已经下降，相机变得更加强大，现在甚至可以自动跟踪球员的移动，并且有更多方式可以在线访问和查看视频。这些变化使得几乎每个人都可以使用工具来深入分析球员表现，并改善球员的成长和训练。“总的来说，这使得我们能够更好地了解比赛、战术以及大多数情况下的准备工作。”

虽然像视频这样的技术现在被广泛使用，但代沟仍然是采用技术所面临的主要挑战。“大多数资深主教练对技术都不太了解，”马奎斯解释道。“我们可以看到年轻一代的教练以更自然的方式对技术进行应用整合，他们更多的是将它看成一种附加值，而不会担心人们可能因此而对他的足球知识产生负面认知。年轻的教练将使用技术来促进信息的流动，而且技术的使用也可能会对一代又一代教练产生共鸣。

马奎斯认为，使用人工智能和机器学习来提取视频片段进行查看，以及对这些片段中的事件进行统计，将为教练工作提供支持。“距离技术成熟仍然尚需时日，但

市场上已经有各种公司在努力解决这个问题，”他指出。“有了正确的教练 / 分析师的大脑来支持其相应的算法，这种助力将是巨大的。它将极大地影响分析师的日常工作流程，使他们最终能够真正将更多时间集中在分析上而不是视频剪辑上。”

资金最雄厚的俱乐部和机构可能会继续探索足球技术的边界，但是马奎斯预计大多数球队将专注于使用大数据来确定真正影响球队表现的因素。他相信，一旦他们将技术的使用集中在这些领域，就会产生有意义的见解。

并非所有技术都有益，马奎斯对视频比赛游戏的不断普及及其对真正比赛的直接影响表示担忧。“街头足球正在消失，”他说。“俱乐部已经注意到对当前球员的创造力和适应性的影响 虽然不仅仅是对抗这种日益增长的电子竞技趋势，但如何应对非正式练习日渐衰落的解决方案可能会依赖于虚拟博彩业，以及我们如何将沙发游戏与推广户外比赛的新方式进行结合。”

关于虚拟和增强现实教练辅助工具，马奎斯不确定这些是否能为未来的足球运动发挥重要作用，或者它们是否仅适用于消费者和粉丝参与。“我认为足球首先需要有经验，使球在脚上具有生命力，尽情比赛，感受队友和对手之间的互动，然后寻求进球。这是比赛本质的一部分，我们不应该破坏它。”他解释说，“如果我们抹除任何元素，就有可能危及学习和训练环境，而球员需要培养和适应这些环境才能踢出我们喜爱的比赛。”

自由度还是奴役度？

莫里西·洛佩斯

如何才能恰当而科学地理解足球？在这里，建议开发对比赛环境更敏感的动力学模型的第一步是解决选择“正确的自由度”（the right degrees of freedom 缩写 DoF）的问题（Krieger）。根据 Martin Krieger 的说法，DoF 是分解一个系统以获得对其“控制”的各个部分，即人们可以改变和控制的变量，预期整体行为的诱导变化是系统性的，并对系统的工作方式有所启发。有趣的是，通过使用动态系统理论（DST）作为建模框架，原则上可以使用我们选择的任何 DoF 来描述系统，但只有所有可能的 DoF 集合的子集才能正确处理变量，从而使我们能够预测和修改行为（Krieger, 1992）。这不是小问题。在某种程度上，发现正确的 DoF 的任务可能与设计模型的任务一样困难，而在足球中，似乎正是如此。在过去，当研究有意识运动时，比如伯恩斯坦著名的锤击任务，对肌肉力量甚至关节角度的测量，已被明确地确定为只能提供“较差”测量的纯解剖单位（Turvey, 1986）。事实上，结构层面上的众多单位使得解释这种行为极具挑战性，因此，运动控制理论提出了一种更为功能性的运动方法（Shaw 和 Turvey, 1981）。当时马达理论家的观点之一正如 Ghiselin (1981) 所述：“支持行动的结构不应与行动相混淆”。

如今，我在足球中遇到了同样的问题。例如，当测量球员相对于球队重心的距离，而不是测量获取系统合法关系的功能单元时，这些指标是系统解剖（即结构）级别的示例。也就是说，借用 Ghiselin (1981) 的概念，足球运动员的结构分布在距统计点（即重心）一定距离处，不应该与以合法结构的方式出现在比赛特定情境中发挥作用的高阶行为相混淆。可以肯定的是，足球确实会产生紧急过程，个人的行为在某些方面就像群体一样，一些统计数字可以揭示比赛的当前和未来状态等等。然而，足球和球队运动需要什么，一般而言，首先是足球理论必须回答基本问题的原则性和结构性指导，这些问题又会揭示出从中可以推导出适当模型的相关 DoF。（见图 1）

当一个人踢足球或训练其他球员时，他面临同样的问题，即找到合适的 DoF 来改变和改善他和球队的表现。从这种相当注重实践的方法中，可以得出一些初步的观察结果，以引导更多的理论驱动的科学方法来解决适当的问题。踢足球是一种在一生中不断发展的技能。随着时间的推移，技能的这种变化是人们可以检测足球运动员变得越来越适应的重要参数。我认为，其中一个这样的参数显而易见，实际上也是最重要的，即球的位置。在孩子们一起

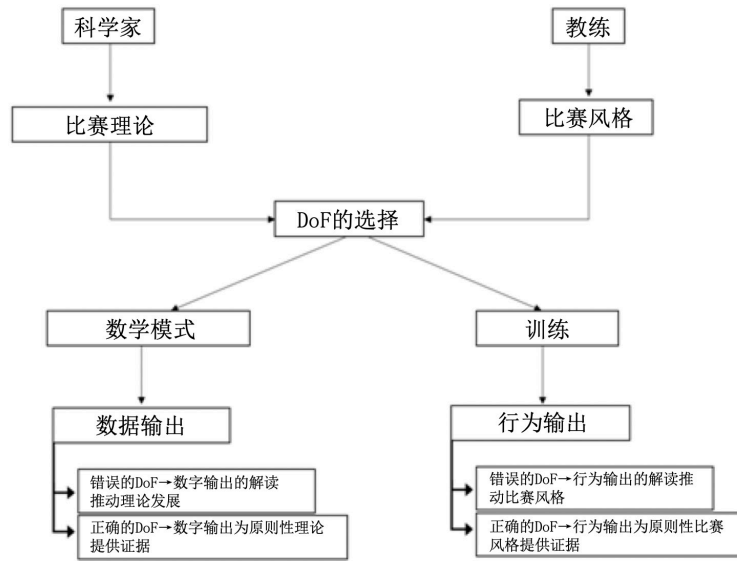


图 1 它代表了科学家和教练在比赛系统方面的工作所不得不面临的类似挑战。在这两种情况下，不正确的自由度选择将从根本上改变输出（即数字或行为），迫使科学家或教练根据他们对 DoF 的选择理解输出，而不是作为原则性理论 / 比赛风格的可测试预测的证据。

玩耍的第一阶段，球的位置似乎直接并且几乎完全驱动他们的行为。他们对团队合作、战术动作、三角站位甚至盘带都知之甚少。无论其他球员在做什么，初学的孩子都会同时跟随球的运动。他们对团队结构没有敏感性，也不会与其他队友和对方团队协调。当拥有球时，一个孩子通常只是试图与球一起跑向球门，而当没有球时，他只会试图通过追逐直接将球抢回。球的位置和控球球员因此构成了球员行为的基本预测因素。也就是说，需要提出的第一个问题就是“球在哪里？”以及“谁控制它？”。

现在考虑更有经验的足球运动员的情况。当然，他们的行为模式要复杂得多，而且似乎取决于比赛的很多参数。球的“位置”和“谁”仍然是推动他们行为的主要参数吗？

(见图 2)

虽然巴萨的足球大师会证实这一点，

但是他们的科学家同行会建议这些参数是否有助于模拟足球运动员的行为，这正是该模型应该凭经验证明的。因此，任何足球动力学模型，无论是比赛的哪个方面或其所针对的分析水平，只要它是足球模型，就必须包含描述球的位置和控球球员的参数。足球是一种集体现象，但可以肯定的是，与鸟类的群集行为不同，它是一种以位置为中心的运动，而球员拥有一种非常特殊的集体元素，即球。（见图 3）

总结而言，本文认为，当试图理解科学家和训练（教练）所研究的系统（即足球）时，球和控球球员似乎是要考虑的相关 DoF。因此，教练和科学家一定要注意，如果他们选择的 DoF 不合适的话，可能会影响他们的比赛理论（科学家）或他们的比赛风格（教练）。也就是说，当科学家对 DoF 的选择不正确时，他们的理论就成

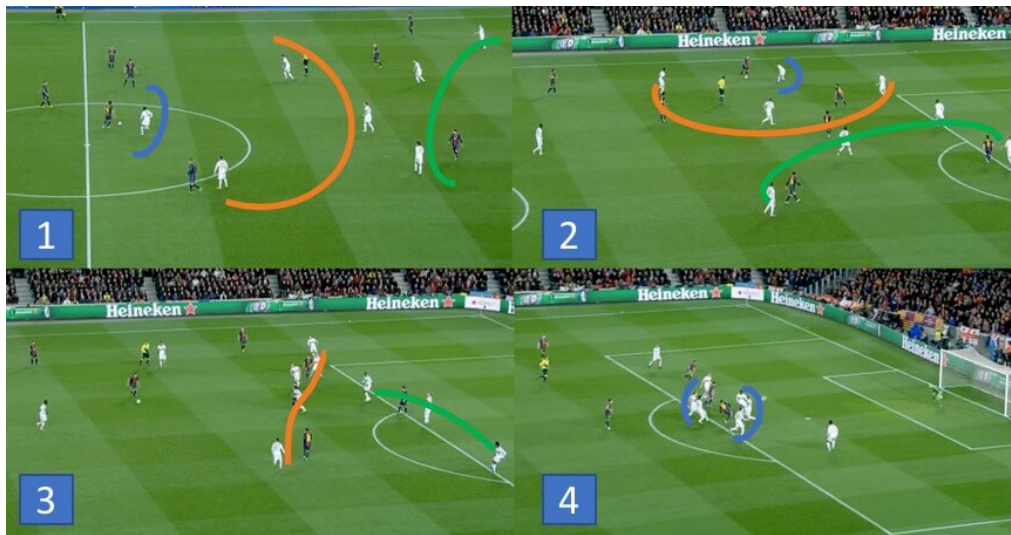


图 2 显示防守球员的位置。正如 López-Felip & Turvey (2017) 所提出的，有三个功能区表示球员相对于球的位置的关系。蓝色曲线表示由于相对于球的距离较近而处于直接干预区；橙色线表示与干预区内球员相互配合的区域；而绿线表示与其他区域的合作区域，因为它们与球的距离较远。

了数学建模解释的奴隶。同样地，如果教练没有在每次训练中使用正确的变量来约束球队行为到所需的比赛风格，那么，他们的比赛风格就会因教练错误解读实现所期望比赛风格的比赛变量 (DoF) 而受到影

响。

以后的文章将讨论足球的本体问题，以回答以下问题：DoF 之间的哪些关系在足球中具有系统性？系统第一位是什么？以及类似性质的其他问题。

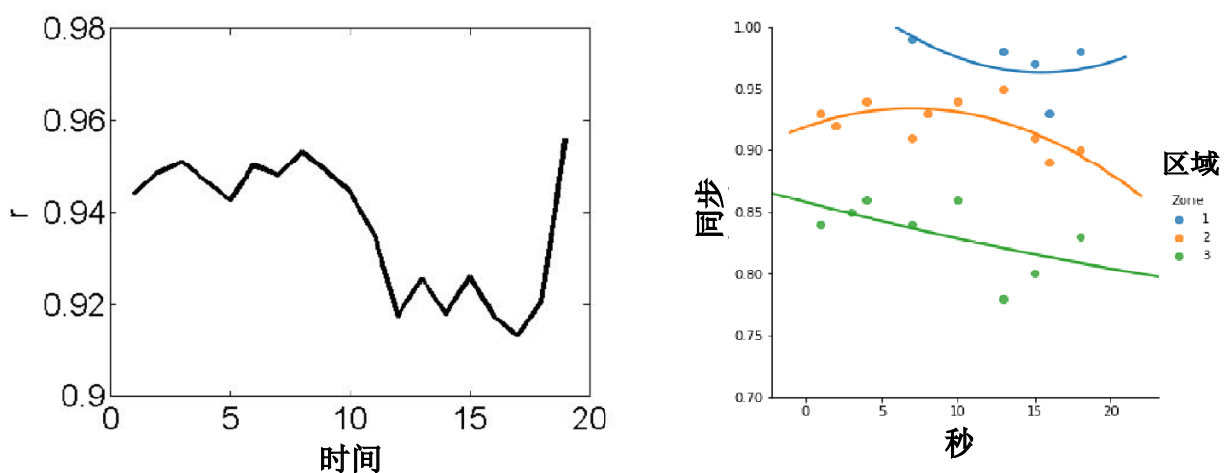


图 3 a) 表示球队一场比赛恢复阶段 (约 20 秒) 的整体同步情况。 b) 显示在图 1 所示比赛期间球队内部协调的行为模式作为球员与球的距离的关系。分析结果表明，球员离球越远时，球员间的协作就会越差。

巴塞罗那足球俱乐部：我们的新研究如何帮助解开“巴萨之路”

保罗·布拉德利

巴塞罗那足球俱乐部是有史以来最成功的球队之一。有人认为他们在俱乐部各个层面取得成功的关键因素之一就是教练们从学院到一线队所采用的独特理念和训练方法。

现在，我们将第一次通过研究提供一些有关训练方法的见解，并尝试揭示俱乐部成功的秘诀。这些“秘密”可能与帕科·塞路尔·洛教练提出的理论以及它们如何与最前沿的体育科学紧密关联。

该研究由巴塞罗那足球俱乐部运动表现部的三名工作人员安德烈·马丁·加西亚（Andres Martin-Garcia）、安东尼奥·戈麦斯·迪亚兹（Antonio Gomez Diaz）和弗朗西斯科·科斯·莫雷拉（Francesc Cos Morera）进行。我受邀担任顾问，与欧洲大西洋大学的大卫·卡萨米卡纳（David Casamichana）一起工作。

我们一起在2015/2016赛季使用最先进的跟踪技术来监控球员在训练和比赛期间的体能消耗情况。每个训练日的方法都经过精心详细的说明，展示球员如何为即将到来的比赛做好准备，并在可能的情况下深入了解这种方法有效的原因。

这些发现令人大开眼界，并回答了为什么球员每场比赛都能始终在身体、技术和战术上做好准备，以及他们如何在漫长

的赛季中保持这种状态。

模仿比赛的训练

比赛开始前三四天，“巴萨模式”要求球员通过健身房和户外训练课程（后者包括小场比赛和使用不同大小场地的位置训练）相结合的方式达到一周之内体能的巅峰。

这意味着所有训练课程包括关注综合能力（身体、技术和战术）的练习。据传，一些俱乐部仍然会进行大量基于跑步的身体准备，但巴萨特别关注模拟比赛的关键元素和某些比赛情况的实际战术训练。

保持新鲜感

由于巴萨模式结合了训练的各个方面，这些训练的时间比其他顶级俱乐部的时间少了十分钟（虽然很小但是会在整个赛季不断累积）。这能够让球员保持新鲜感。在比赛开始前一两天，该模式主要侧重于通过连续进行控球和传球练习进行技战术准备，每队人数较少的位置练习，以及定位球练习。

训练负荷指标，例如球员在训练中所覆盖的高强度距离和加速度数值，随着比赛接近而减少（在重要比赛之前减少练习）。这表明教练正在减缓身体的负荷，而重点

磨练球员的战术和技术准备状态，以确保他们在大赛中保持健康和敏捷。

做好“比赛准备”

我们的研究揭示的另一个关键差异是巴萨如何让那些没有参与足够比赛的球员保持比赛状态。在巴萨模式中，球员需要在比赛后的几天内进行训练。例如，这些课程会将球队分成两个训练小组。第一组由参加比赛超过 60 分钟的球员组成，主要进行低强度活动并结合恢复性训练。

参加比赛不到 60 分钟且需要“补充”身体和战术强度的球员参加技战术循环练习（一个接一个练习），随后进行高强度的位置训练和小场地比赛。这种额外的训练提供了适当的刺激，以保持球员的体能，并且是教练用来确保上场时间有限的球员

被派上场时能够在身体、技术和战术上做好准备的重要工具。

看上去巴萨模式的目的是在每一周和整个赛季中改变球员的身体 / 战术负荷，以使所有球员的表现达到顶峰并保持高水准——包括那些没有获得比赛时间的球员。

那么，什么是“巴萨之路”？看上去并不全都是 tiki-taka 一脚触球练习，而是一种独特的理念，融合了优秀的教练训练和前沿的运动科学。它涉及到对是什么让球员提高的更为复杂的理解。

战术很关键，休息和恢复也是如此。许多顶级俱乐部当然都会有自己的理解。但是我们的研究提供了一个独特的见解，让人们了解欧洲最伟大的足球俱乐部之一的成功秘诀。

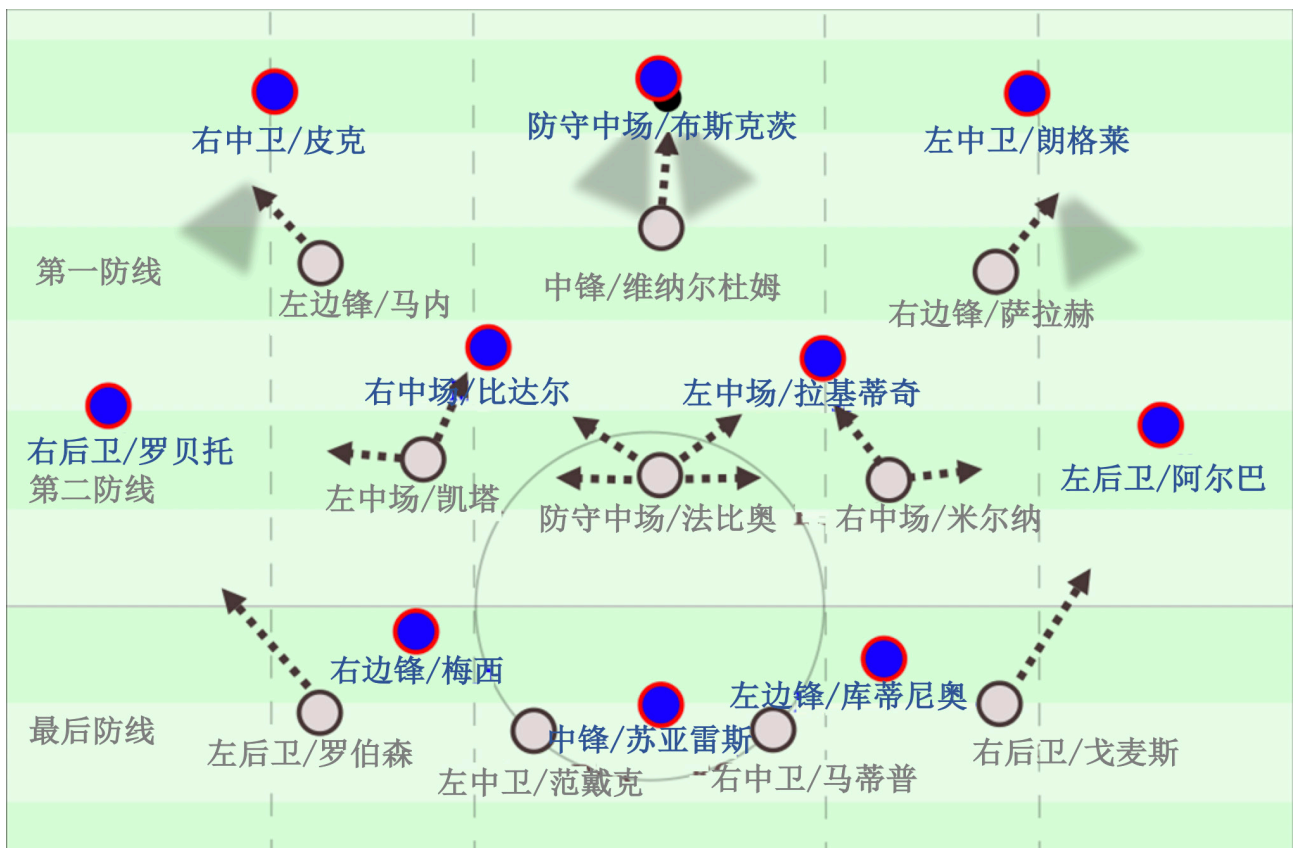


利物浦的压迫系统

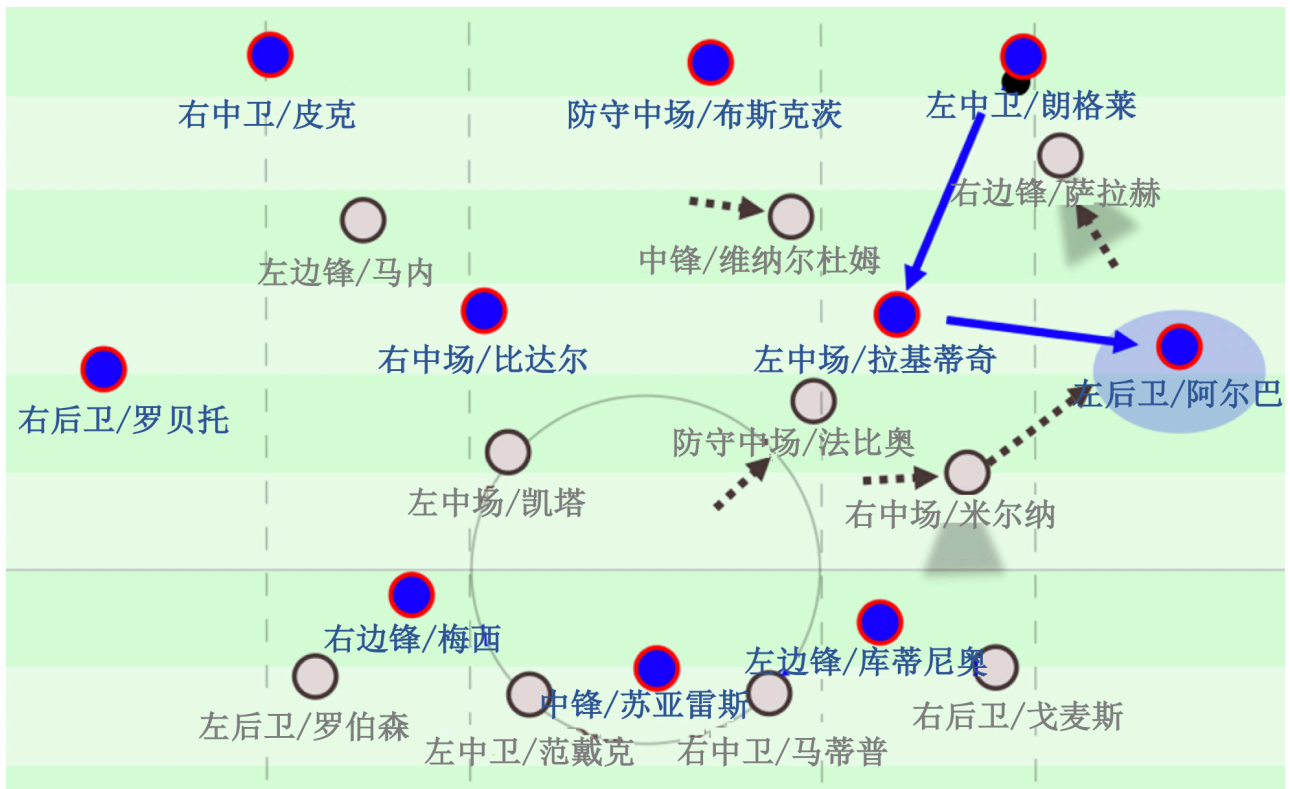
VON IB

利物浦 4-3-3 压迫系统的基础和主要原则是将球保持在中路区域，防止将球横向移动，如果做得好的话，这通常能在纵向和横向上延展防守结构。

要实现这一目的，他们的阵型需要兼顾传球通道和选位：第一条防线切断传球路线，防止对方向边后卫 / 边锋大范围传球，而第二条防线更倾向于选位，因此他们的原则是选择可以接近 2 名进攻球员的站位（防守需要根据对手的阵型、距离、球速进行移动）。与此同时，最后一条防线，主要是后卫也有选择的方向：如果第二防线不能创造最佳的机会接近对方的后卫，那么他们会压上，在后场留下相同人数。理想情况下，最后一条防线需要保持在适当位置，以便在中路区域创造更好的长传球线路。



欧冠第一回合对阵巴塞罗那时的基本压迫系统：第一条防线以封锁传球路线为主，第二条防线快速移动并选位，罗伯森和戈麦斯在需要时会前压，接近罗贝托-阿尔巴的位置。



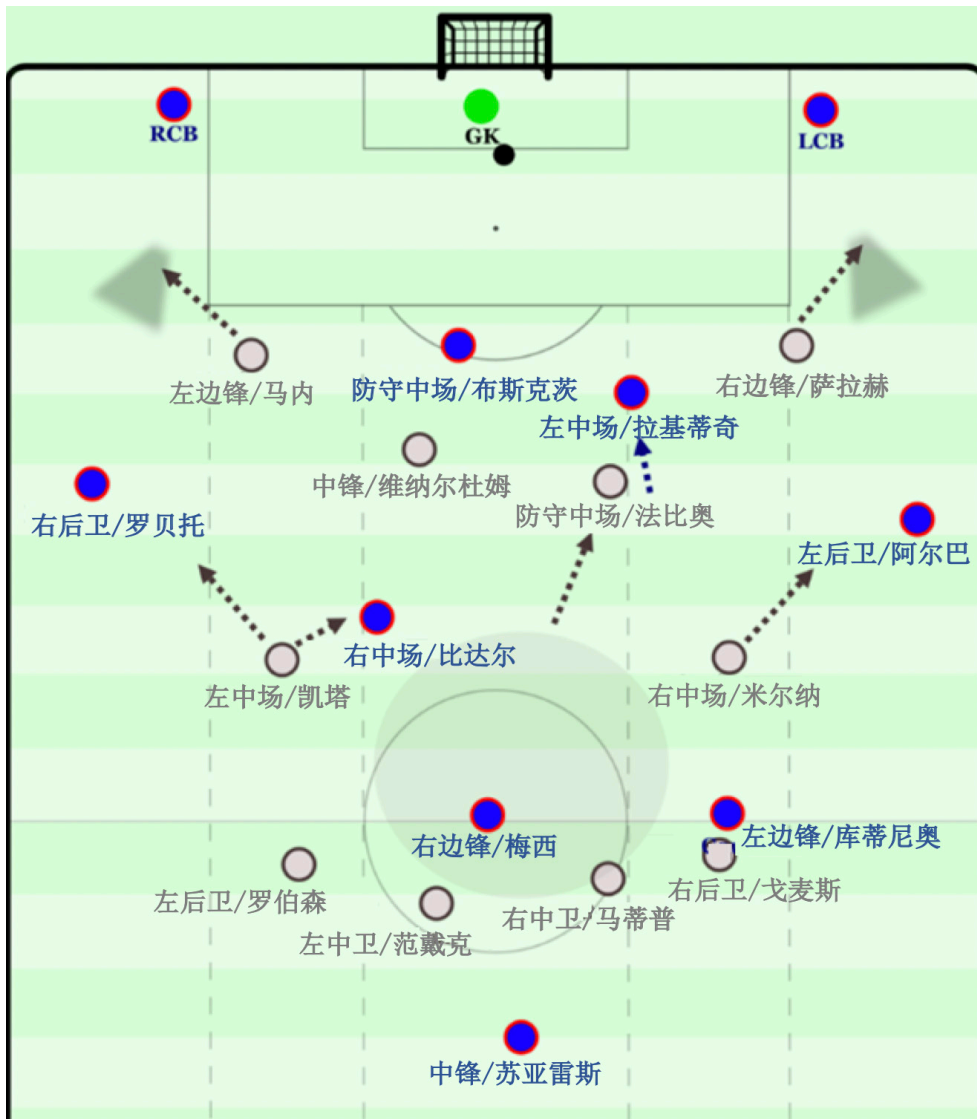
即使对手可以找到有价值的中路选项或广泛的区域，他们也有不同的解决方案/制造陷阱来进行防守，尤其是在侧翼。

迫使对手的进攻组织进入陷阱，模型来自第一回合：对手左中卫被右边锋/萨拉赫贴身紧逼并封挡向外侧传球，这打开了向中场的纵向通道，然后中场可以传球给左后卫拉开宽度。如果右后卫/戈麦斯无法压过最后一条防线，那么第二条防线将水平移动以覆盖横向的区域。

动态移动

边锋正在压迫外角，而中锋会根据对手的阵型总是设法压迫离球较近的中场或防守中场。这种移动使得所示的离球较近

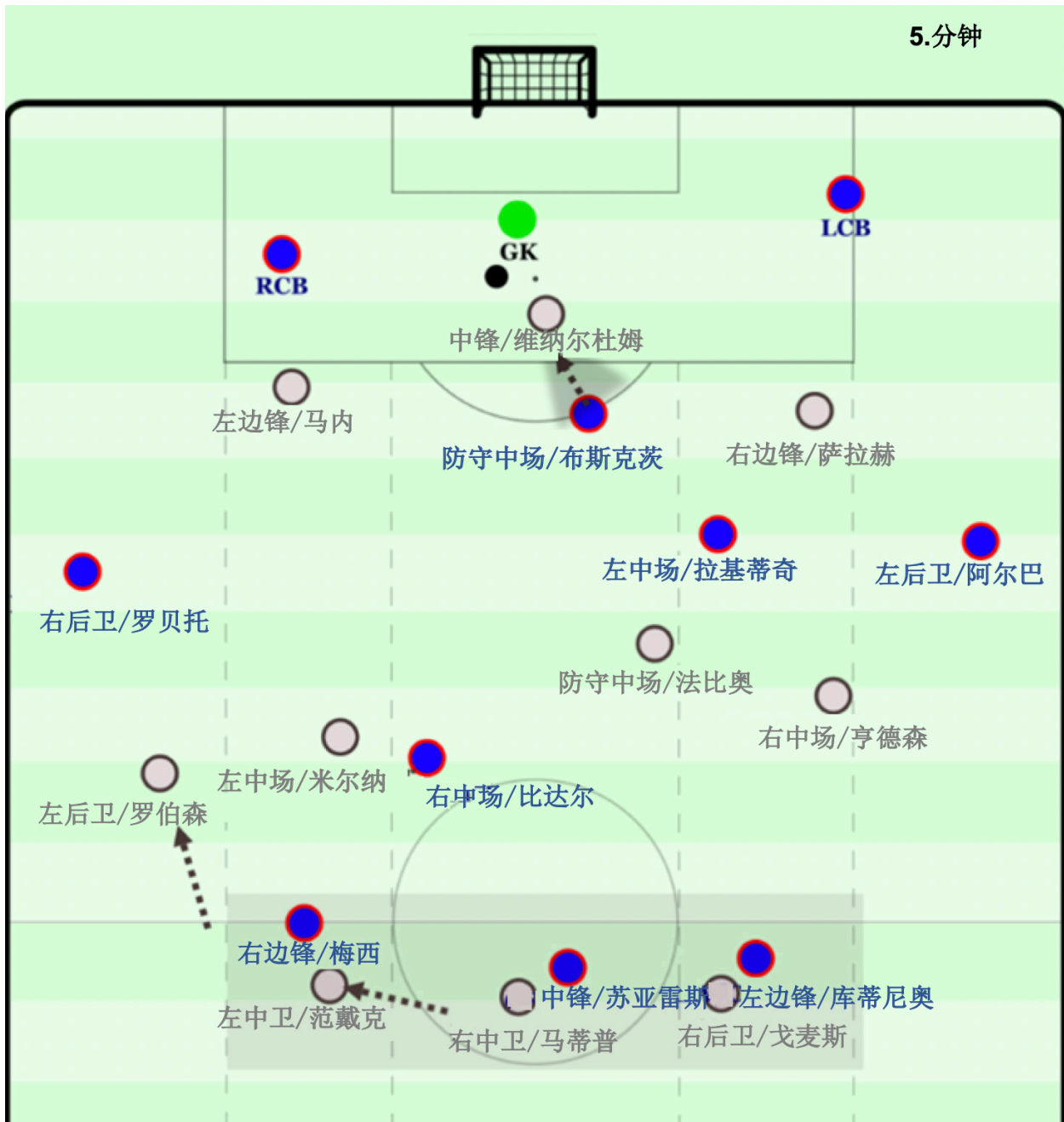
的中场/米尔纳移动得更宽一些并且控制传球路线的纵深，防止他过于专注于对手。由于他的位置更宽，如果他必须向外移动以压迫对手的边后卫，则他可以更容易的进入广阔空间。由于第二条防线大多在一条平线上交错排列，因此防守中场可以灵活地移动到更高的位置。对于一个单独的防守中场，通常没有必要这样做，因为中锋/菲尔米诺做得很好，不断用他的跑动覆盖这种选项，使得第二防线可以专注于覆盖背后的深度区域。对于2个位置更深的中场球员（这里是3-2结构），模式是防守中场/法比奥前压逼抢离球近的中场，阻止他在接球后转身。当球出边线时，防守中场有时间重新调整他的位置。



压迫巴萨的组织，第一回合：如果中后卫接到球，边锋 / 萨拉赫和马内跑动上前压迫，同时封住向边后卫传球的路线。如果纵深站位的防守中场 - 左中场接球，更容易前压，还可以以更好的压迫角度压缩纵向选项，在回传时向守门员施压。如果边后卫 / 阿尔巴 - 罗贝托直接接球，那么中场 / 凯塔和米尔纳会横向压迫，他们的站位相对开阔以减少移动距离。

尽管利物浦的基本阵型在中路留出了更大的空间，但巴萨并不能顺利地从中后场开始组织，无法有效地与第二条线进行联

系，这导致他们未能更好地利用中路的空间。通过这种方式拉扯拉基蒂奇是一个聪明的解决方案，因为这拉出了法比奥，在中场身后开辟了更多的空间。拉出 4 名球员可以成为为球队打开宝贵空间的有效方式，他们可以在球门球发出后进行更好的组织并争抢第二落点的球。这也让利物浦的最后一条防线保持良好的站位，所以尽管巴萨在长传球之后重新获得控球权，但是 4 名后卫能够切断身后任何可能的跑动，从而减慢控球速度。



另外一名关键的球员是比达尔，更靠近中路站位的他需要面对左中场 / 凯塔和米尔纳，通过把他向内拉扯可以为右后卫制造更多接球的空间。由于这个选项只使用了一次，巴萨基本上都在对抗高压。可是利物浦也有相应的解决方案。

解决方案：防线的移动 -> 在这个场景中，罗贝托直接接到特尔施特根传球，

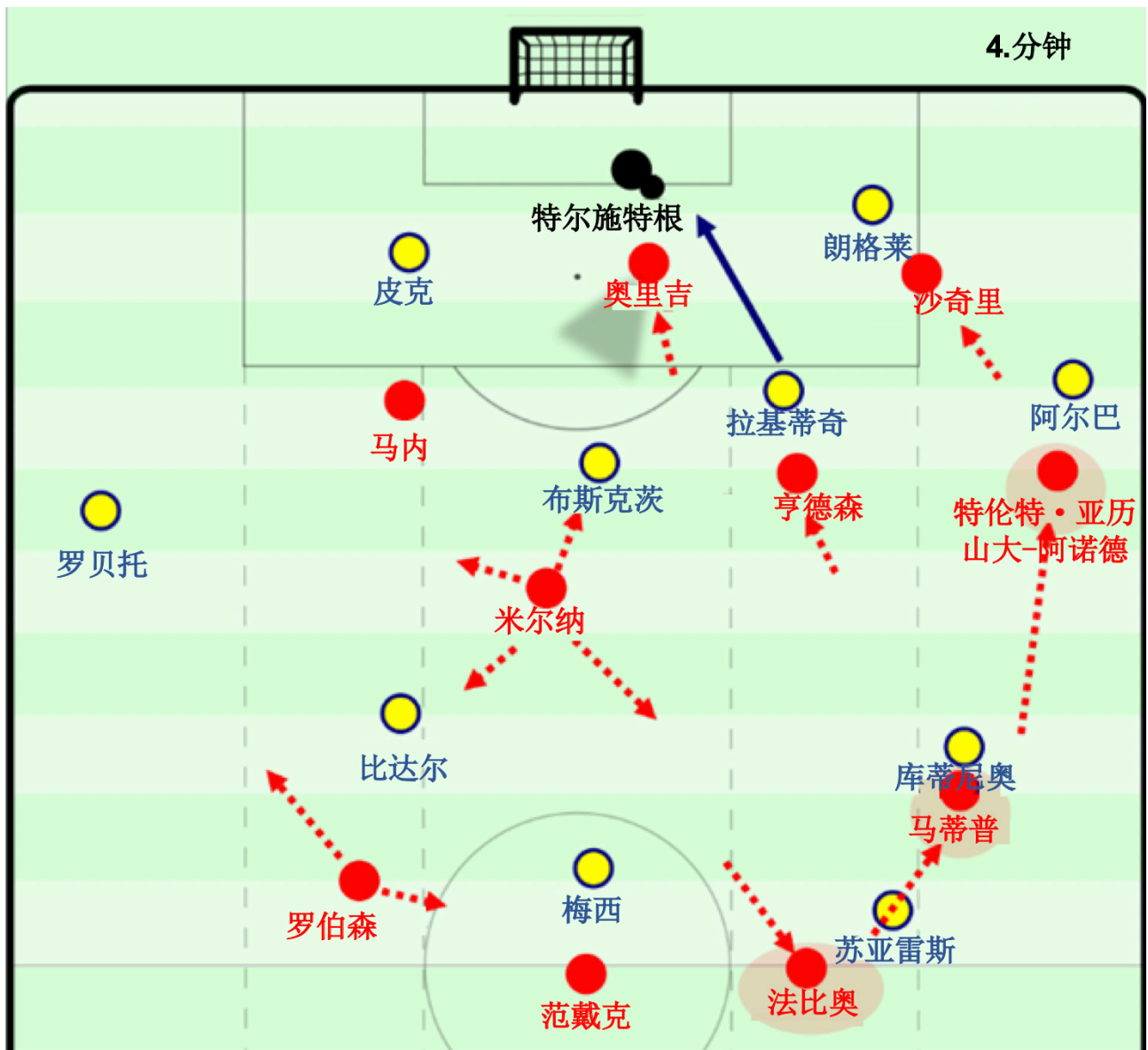
马内深度逼抢迫使其向后传球，同时让罗伯森有机会压上。利物浦的最佳做法就是保持移动，在后防制造 3 对 3 的局面，但是没有最佳的中场串联，也无法利用梅西 - 库蒂尼奥之间的宽度距离（例如直传球，将防线拉宽），因此范戴克 - 戈麦斯可以在长传球之后很容易地利用他们的身体优势。

第二回合

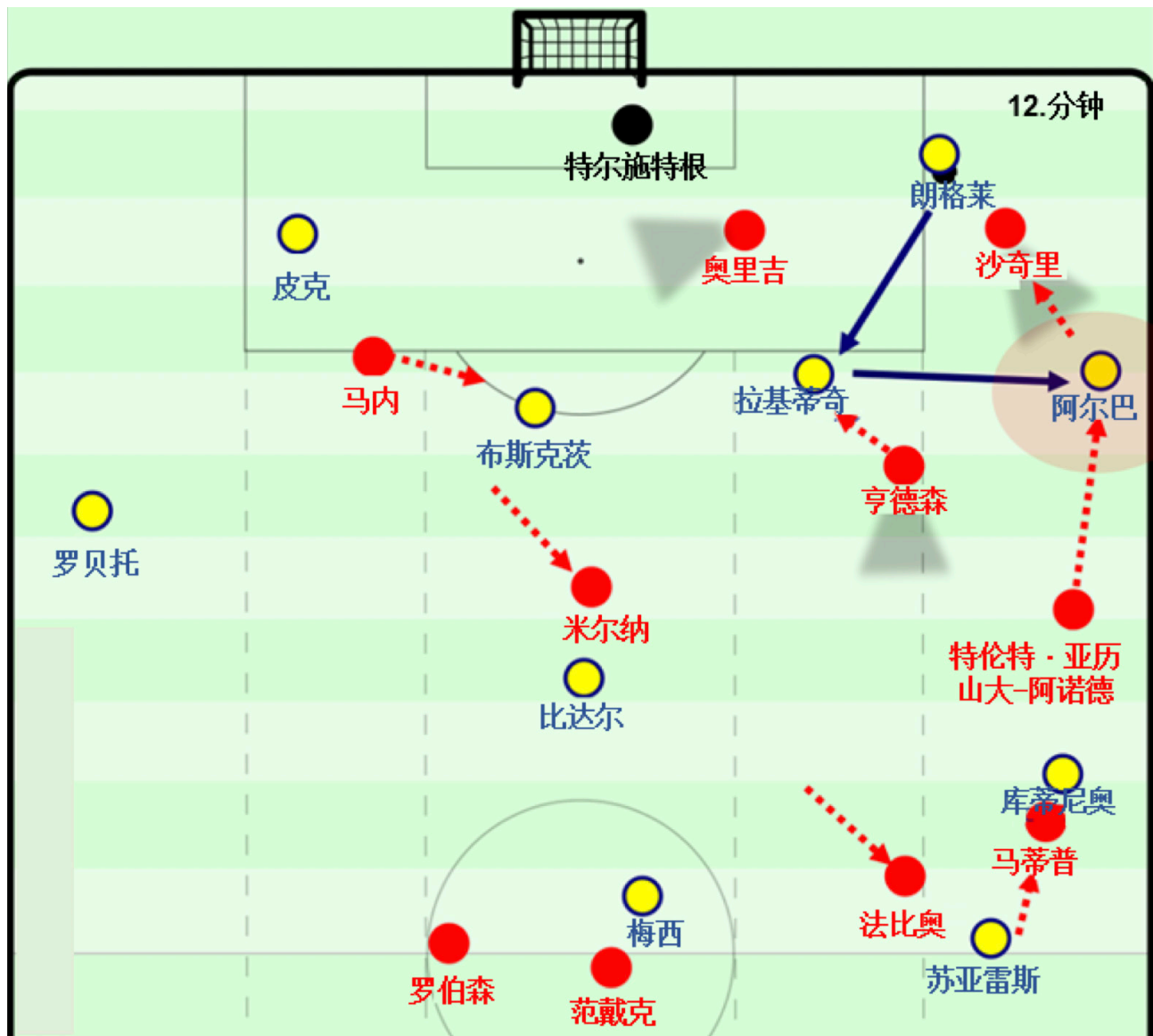
在第二回合，利物浦对他们压迫的方式做了一些关键调整。首先，中锋 / 奥里吉的主要作用是引导巴萨向左路组织进攻，以防止他们在更强大的右路将球传给梅西。为了实现这一目标，奥里吉的站位非常靠近布斯克茨，他的逼抢封堵迫使这一侧只能将球回传给特尔施特根，然后通过朗格莱进行组织。在向后传球后，奥里吉立即

通过弧线跑动追击球的移动，封锁右路的传球——主要原则：防止球的换边，使其保持在初始区域。

第二回合的压迫方案：迫使向后传球给特尔施特根，然后奥里吉立即跑动压迫，迫使特尔施特根使用他的左脚。防线保持移动换位，法比奥在马蒂普和范戴克之间深度回撤。



从第一回合吸取教训，当时巴萨可以在某些情况下有效地利用广阔的区域，他们也改变了阵型的移动方式。为了更好地接近对手的边后卫（阿尔巴和罗伯托），最后一条防线通常站位更宽以减少压上的距离，这使得亨德森和亚历山大·阿诺德在对手接到传球后可以立即向边后卫施压，同时整个后防线保持移动。但是促成这一变化的关键在于对中场线的改变：与第一回合的战术相反，防守中场 / 法比奥在任何情况下都保持更深的站位，作为防线身前的自由人，并且在马蒂普前提压迫对方边锋 / 库蒂尼奥时，由他回撤来占据中卫的位置。这也改变了离球较远的中场 / 米尔纳的角色，他必须在防守中场 / 布斯克茨和右中场 / 比达尔之间选择更高的位置：如果需要防止布斯克茨斜传，则立即压上，防止球的换边——或者如果阿尔巴接球，则深度回撤。



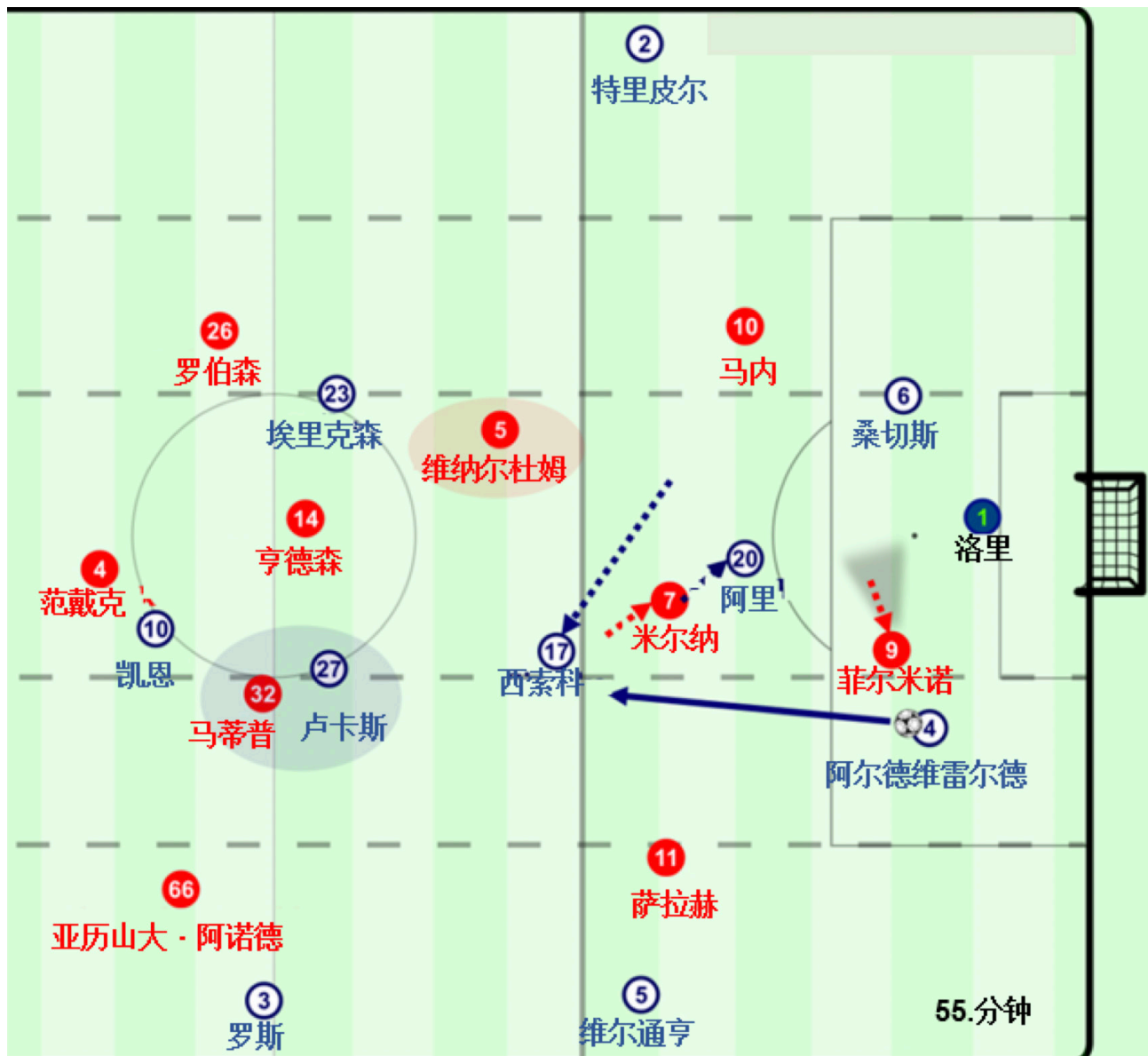
给阿尔巴设置的压迫陷阱：奥里吉再次利用他的防守面积切断右路的路线，而沙奇里则向上压迫朗格莱，切断其给阿尔巴的传球路线，打开中路。亨德森的主要移动：不要过早压上，首先关注的是切断纵深传球道，然后在拉基蒂奇接球时进行上抢压迫。故意留出给阿尔巴的传球路线，随后由亚历山大-阿诺德进行逼抢，后者从随球移动的防线位置前提——加上法比奥的回撤。

压迫中场回传球的模型：中锋 / 奥里吉切断传给远端中后卫 / 皮克的路线，迫使其向后传球给中后卫 / 朗格莱，然后从对角线对其进行逼抢——但离球较远的边锋 / 马内也会在可能给中后卫 / 皮克的横向传球路线上进行压迫。关键 -> 防守布斯克茨：理想情况下他每次都被离球较远的中场 / 米尔纳逼抢，以防止其进行斜传。这种压迫战术也能使离球较近的中场 + 防守中场切断传球通道。



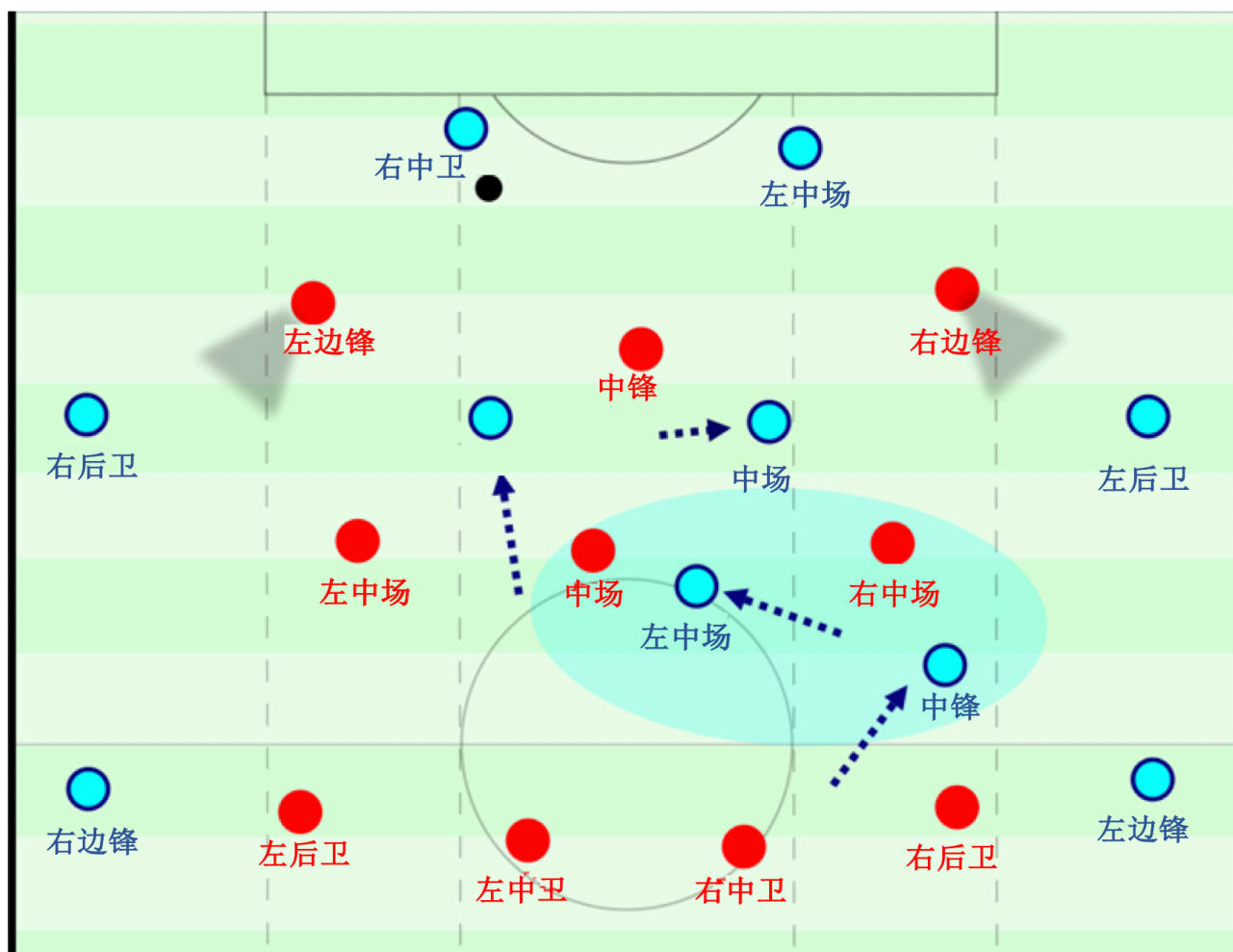
切断横向传球路线意味着对手基本上有三种方式来进行渗透 / 利用：

- 聚集 -> 会减少前锋线的覆盖范围，可能会增加中场中路的球员以利用空间
- 例如建立菱形中场（2-2 斜向站位）



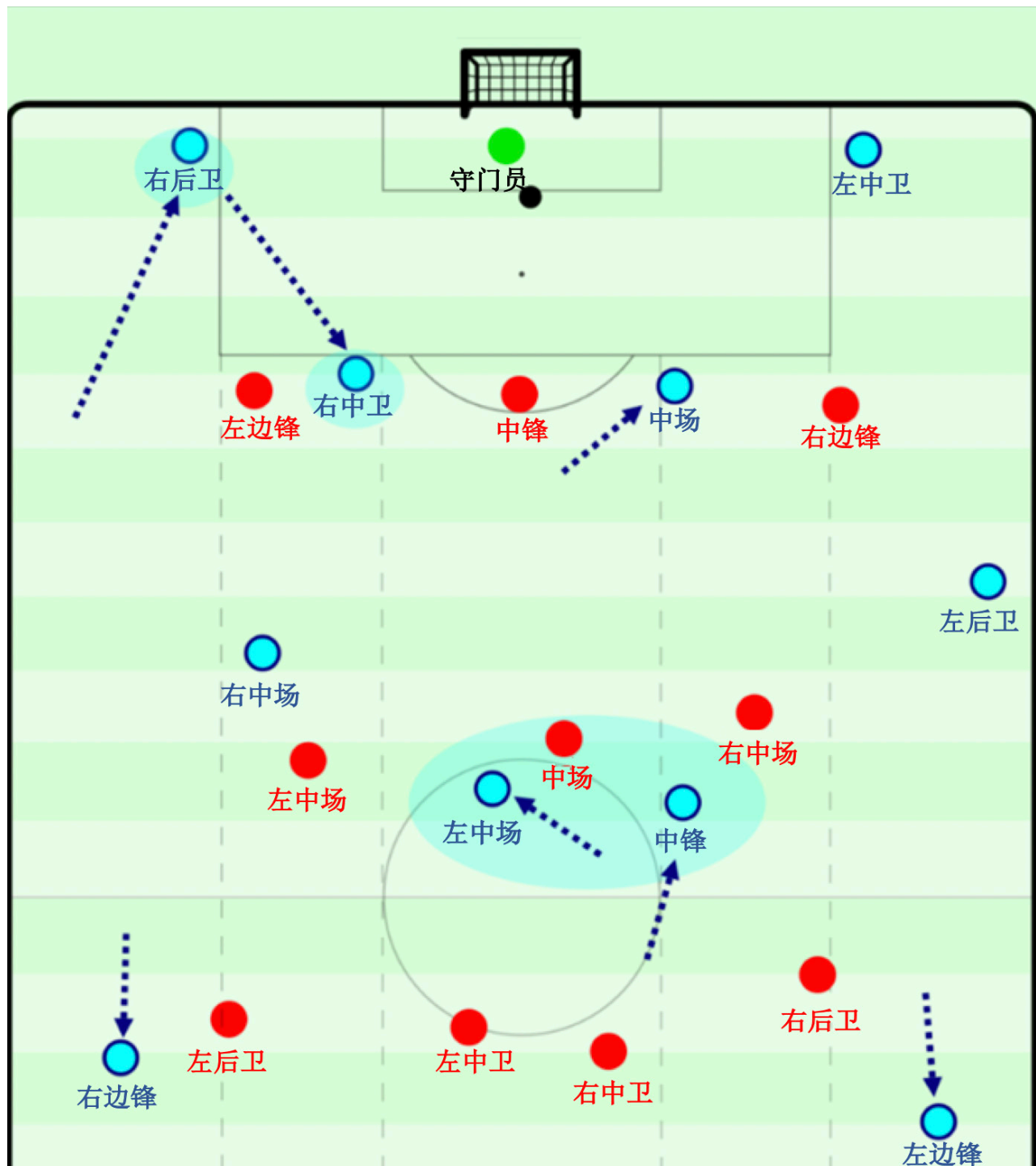
热刺的解决方案——锋线身后采用菱形中场站位，西索科 - 阿里两人盯防米尔纳，并给菲尔米诺制造方向问题，减少锋线的传球路线覆盖范围，打开直传给西索科的路线。

- 双倍宽度 → 拉伸其阵型，减少宽度的覆盖，这样可以更有利于中后卫的处理球（这里有两个关键原则：
 1. 中后卫在球的对角线方向，利用中路连接并拉回锋线，然后利用宽阔的空间。
 2. 中后卫的距离 → 彼此更紧密的站位，将对手边锋向内拉，为边后卫直接接球或通过配合开辟更多空间，如中后卫 - 防守中场 - 左后卫路线）。
- 例如拉宽站位的 4-3-3/4-1-4-1——具有不对称性 + 使用假 9 号位置，因为中后卫不会跟随移动而且边后卫的注意力集中在边锋上。



曼城在1月对阵利物浦时的基本战术：边后卫和边锋都拉开宽度以拉扯利物浦的结构，加上更加不对称的站位，右中场回撤，左中场更靠近中路，中锋向边路移动。这种中场的不对称处理，可以让他们占据在广阔的空间，以便右中场 - 中锋在中场接球。

- 双方 → 4-2-2-2-0——利用斜向站位加守门员在后场建立5对3的局面，有利于球更流畅的运转，并利用球场宽度同时使利物浦的防线随球移动）。



曼城的目标是创造 4-2-2-2-0：右中卫向内移动到 6 号位置（也可能与深度回撤的中场一样），右边卫深度回撤，增加对方锋线的轮转难度，如果需要的话，右中场可以选择向外移动——拉伸中场线。左中场围绕着防守中场移动，中锋也回撤作为伪 9 号以牵制右中场，延缓其在左后卫拉伸宽度时所施加的压迫。

2018 年国际足联 U-17 女足世界杯技术报告

国际足联技术研究小组



▶ 进球

本届赛事是历史上进球最少的一届，在 32 场比赛中仅有 86 个进球（平均每场比赛 2.69 个）。实际上，随着通往球门的道路变得越来越困难，通过定位球得分数量显著增加，24 个进球，占进球总数 28%，或超过四分之一进球来自各种各样的定位球，凸显了在定位球方面的准备和执行工作非常优秀。其中包括日本队在手抛球后取得的 2 个进球，彰显出如果经过正确的准备、计划和组织，任何死球都可能造成威胁。日本队显然很注重这方面，她们的

8 个进球中有 4 个来自定位球。与此同时，墨西哥队一路杀入决赛，尽管她们只取得了 6 个进球，其中 1 个来自角球——在乌拉圭世界杯通过角球取得的进球不少于 8 个。韩国队证明了她们是这方面的大师，她们的 7 个进球中有 5 个来自角球。她们出色的空中力量——这一主题将在后续进行更深入的讨论——以及攻门的时机，都比对手更为出色。另一个很好的例子就是新西兰队与日本队在四分之一决赛中所取得的进球，她们的跑动时机近乎完美。

然而，虽然有的国家——例如前面提到的墨西哥和新西兰——似乎专注于将定

位球作为其取得进球的重要方式，而其他国家则可能将注意力转移到其他地方，也许是因为在她们的军械库中拥有其他武器。例如，西班牙队是唯一一支未能在比赛中通过定位球得分的队伍，但是她们却最终捧得了奖杯。她们的15个进球中有8个来自传中或倒三角传球，她们对场地的宽度的充分利用（后续讨论的另一个方面）是取得成功的重要因素。最后一脚传球的质量使她们前锋的工作变得更容易，而冷静的射门——她们每2.3次射门就能取得一个进球，而赛事平均值为4.8——这令她们一骑绝尘。西班牙队出现在罚球区内的次数也更多，这是她们充分利用宽度并将球传递到罚球区的直接结果：当对手的防守未能充分解围时，她们在罚球区里的球员就会将球拦下并组织进攻。另外两名奖牌获得者墨西哥队和新西兰队出现在罚球区内的次数也相当高，收获是12个进球中有10个是在罚球区内完成的。

技术

虽然在U-17级别的比赛中战术的成熟度仍然处于发展阶段，但有趣的是我们看到了个别球员天赋的展现以及这些年轻人勇于表现自己的无畏精神。从这个角度看，具有出色控球能力的技术娴熟且充满自信的球员在未来几年可能会非常普遍。事实上，那些在1对1面对防守球员时敢于突破并能娴熟控球的球员，是最引人注目的，并且能制造最大的威胁。例如，冠军西班牙

队中有两个这样的人物，其中一位是伊娃·纳瓦罗（Eva Navarro），在右路活动的她送出了并列赛事最多的三次助攻，而另一位则是综合实力最强的萨尔玛·帕伦罗（Salma Paralluelo）——刚刚15岁零八天的她也是赛事最年轻进球球员——在左路活动。二人都是西班牙成功的关键因素，她们的节奏变化和假动作为克劳迪娅·皮娜（Clàudia Pina）创造了机会，而后者通过出色的射门能力一击致命。这些只是众多能够单独创造得分机会的球员的一些例子，但最重要的是保持球权并充满信心地传球。冠军西班牙队以65%的控球率和79%的传球成功率表现出色，但她们并不是唯一拥有出色控球能力的队伍。芬兰队——她们的控球率排名第二，为64%，传球成功率则最高——带来了一种新的打法，她们利用守门员有效地进行后场传球。他们的两名门将安娜·库文娜（Anna Koivunen）和艾玛·以曼尼（Emma Immonen）平均每场传球次数最多，分别为63次和59次。这凸显了她们在他们球队组织中的作用，并在这个层面上代表了一种新的现象，随后将在守门员部分进行讨论。

除了短传和建设性的组织外，出色的控球技术也很普遍。德国队的夏奇拉·马丁内斯（Shekiera Martinez）和吉娅·科莉（Gia Corley）展现了优秀的第一脚触球和近距离控球在保持控球权或躲避对手方面的重要性，而这也有助于加快她们的动作。尽管战术意识仍在发展中，她们的控球能力具备了明显的优势，并为她们提供了更

多的时间和空间，通过控球或加速来展示她们的个人技巧。

广阔的区域就是许多个人加速的选择区域，其中包含大量的1对1进攻，加纳对新西兰的第二个进球就是教科书般的例子，说明了这是多么的致命。节奏、带球技巧和对球的自信，这些都使得球员能够向防守的空隙进攻并进入危险区域完成传中或倒三角传球，给插上的中场或前锋制造机会。

这种进攻很难防守，但是通过良好的比赛阅读和正确的身体位置和方向，防守球员至少可以获得阻截或铲抢的机会，甚至出现在球和进攻球员之间的路线上。在1对1防守中，在压力之下有效利用身体并且保持硬度和弹跳对于防守球员做出关键解围或防止前锋接球至关重要。在这方面，许多后卫看上去都训练有素。

我们已经提到过乌拉圭球员的个人自信，这带来了大量的远射——每队平均5.9次，16%的进球来自罚球区外，其中许多射门越过守门员。与此同时，乌拉圭的另一项突出技术特点则来自空中。8个进球来自头球，还有几个来自定位球，这证明了在训练中——至少在这个级别——对于定位球的演练更多从进攻角度出发。此外，在这个年龄段的头球能力正在提高，这在进攻和防守中都能看到，跳跃的时机、选位、对比赛的阅读和力量都有助于头球的快速发展。

从进攻组织到进攻三区

随着守门员越来越多地参与到球队的组织当中，球队有时在控球时的站位会比无球时更加靠后，这听起来似乎有点违反直觉。通过这种方式，他们能够从后场通过二过一和短传配合进行组织，通常利用短传将球从一侧转移到另一侧，而不是急于将球权交出。

短传在这个年龄段的比赛中更为常见，虽然涉及三名或更多球员的精细配合非常罕见，但简单而有效的换位却能很好地执行并达到摆脱对手逼抢的目的。相当数量的双脚球员在接球时表现出娴熟和镇定，同时还展现出了良好的视野和比赛阅读能力，完成转身并找到队友。

墨西哥和加拿大的比赛就是一个明显的例子，即使在后场也努力保持控球和传递：发挥她们的优势，通过短传和二过一有效地从后场组织进攻。日本和韩国的比赛风格也类似。即使在压力之下，她们也很擅长通过短传找到解决方案，并且当所有传球路线都被封堵时，她们并不惧怕将球传给守门员。

带球推进并吸引防守球员之后将球传给空位的队友是我们观察到的另一个有效策略。这是一种创造人数或位置优势的方式，以及在对手被成功吸引的情况下及时出球，往往能够创造进球或得分的机会。



▶ 高压

除了表现出信心之外，勇敢——无论是个人还是球队——当然也必不可少。

几乎所有的球队都采用了高压战术，不论对手控球还是不控球，防线都处于高位，并且没有人会坐等对手向其发动进攻。这使得比赛非常有趣，并且也展示了教练如何鼓励他们的球员在场上勇于主动表现，而不仅仅是被动反应。新西兰在这一方面体现得尤为突出，她们在对阵加拿大时仅仅 20 秒就取得进球，创造本项赛事有史以来的最快进球绝非巧合，当时她们有两名球员逼抢守门员并迫使其匆忙做出决定，随后被她们加以利用。虽然这种做法很大胆，但这是来自大洋洲国家的一种特别聪明的方法，她们希望利用自己的优势，包括她们的内在竞争力和永不放弃的精神来

对球施压。

除了勇敢和冒险的战术，许多教练还展示了他们的独创性，利用惊奇元素来击败他们的对手。这在新西兰对阵日本的四分之一决赛中有明显体现。承认他的球队在技术上处于劣势，新西兰教练莱昂·伯尼（Leon Birnie）和他的教练组制定了一个获胜的计划，这也要归功于在比赛的期间所进行的心理工作，比赛最终取得了胜利。比尼说，如果你问球员，他们是否有信心击败日本队，90% 球员会说有。

他和他的教练团队对日本队进行了研究，虽然她们的强大给他们留下了深刻印象，但他们并不会束手就擒。相反，他们大胆地选择了其他人未曾尝试过的事情——在场上采取高位逼抢并迫使日本队进行防守，几乎让她们在自己的半场中窒息。

只有在体能水平可以维持的情况下，这种战术才有效，但这仍然足以帮助他们将比赛拖入点球大战，并最终取得胜利。

守门员也为高压战术做出了贡献，提前占据有利位置，从而对对手的进攻做出预判。许多守门员经常勇敢地冲出罚球区，这让她们的队友更有信心前压，因为她们知道守门员会覆盖她们的身后和后防线之后的空间。

与此同时，墨西哥采取了一项明确的计划，保持中等强度逼抢并改变她们的压

迫节奏，提高更广泛区域的强度，并作为一个整体行动。“我坚信细节的作用，我认为细节决定成败”，他们的教练莫妮卡·范盖拉（Mónica Vergara）说道，她的球队在对乌拉圭时的整体表现证明了这一点。

球员在压迫时的跑动也很重要。有时候，他们会从盲区绕到对手身边，令对手猝不及防，而当对手向后传球时（比如传给守门员），强度会发生变化——这时会施加更多的压力。因此，压迫过程往往是动态而复杂的。



▶ 利用球场宽度

西班牙超过一半的进球来自传中或倒三角传球，她们对场地宽度的利用是成功的关键之一。她们平均每场比赛传中 26 次，而赛事的平均值为 18 次，而这些传球的质

量往往也非常高。

由于几乎所有球队都采取高压战术，反击是非常宝贵的，但这并没有否定在控球时利用横向通道的必要性，经常将比赛从一侧转移到另一侧，以寻找边锋 1 对 1 的机会。在这方面，西班牙并不是唯一一个



边锋站在对方半场拉开宽度的球队，通常处于开阔区域，来完成接球。墨西哥、日本和加纳同样采用这种策略。

与此同时，边后卫拥有更多的支持，而不只是交叉跑动。他们的任务是将球交给边锋之前带球推进并让边锋利用他们的控球发起进攻。在边后卫的支持下，边锋还可以选择将球回传，由边后卫从不同角度完成斜传到罚球区；这些战术非常有效，带来了许多进球。



► 守门员

由于女子 U-17 守门员的身高仍在发育，处理高空球射门仍然是一项挑战，本次大赛中很大一部分进球都来自远距离的高空球射门。在处理这些球时，守门员必须表现出良好的站位、预判和移动步法，以便他们能够很好地起跳，并最终用强有力的手和指尖触球，完成扑救。在本届世界杯期间有一些很好的例子，守门员展现出这些特质并完成重要的扑救。

考虑到这一点，应该鼓励守门员重点训练跳跃高度以及完成高空球射门扑救所需的其他素质。这也可以通过复制女子足球 U-17 级别比赛中的远射轨迹并且提高对这种射门的预判能力来实现。

除了扑救之外，守门员的整体角色已经多样化，芬兰队特别展示了一种全新的方法。她们的第一门将安娜·库文娜实际上是球队的第 11 名场上球员，她与球门的平均距离为 17.3 米——这意味着她大部分时间都在罚球区外而不是罚球区内，并且有意愿且有能力通过传球帮助球队从后场组织进攻。“这是全新的发现，我很惊讶。很高兴看到这样的变化”，国际足联守门员专家帕斯卡尔·祖贝布勒 (Pascal Zuberbühler) 说。“她就像一个额外的场上球员；她使用双脚触球并且技术娴熟，从而给她的队友们提供了信心，她们也愿意让她参与进来——这是她们在训练中就在做的。”

两名在乌拉圭比赛出场的芬兰队守门员的传球次数是赛事守门员平均传球次数的两倍。大部分都是短传或中距离传球，而不是长传，而提升比赛参与度的战术是芬兰队关注的焦点。“在每支青年国家队中，我们都会努力让守门员成为一名额外的场上球员”，芬兰队的守门员教练埃梅里·雷普内（Eemeli Reponen）解释道。“这是我们足球理念的一部分。”

虽然芬兰的守门员表现出足够的脚下技术，但手抛球是一个可以在未来投入更多时间的研究领域，因为手抛球通常可以比长传更准确地将球传给队友。守门员的快速思考，以及准确的手抛球和队友聪明的预判和跑位，可以让球队在对手失去控球后立即从后场发起进攻。后卫和中场球

员在识别机会并向空挡跑动所表现出的敏锐的意识，证明了球队在训练场上所付出的努力，这是另一个可以通过一起训练获益的方面，守门员作为额外的球队成员越来越多地参与其中。

这不仅是在进攻阶段，在防守时同样如此——正如前面关于高位压迫部分的讨论——守门员的贡献对球队的整体表现至关重要。尤其是球队采用高位压迫战术时，守门员的位置意味着他们可以产生相当大的影响，并且许多守门员都表现出了冲出罚球区冒险的勇气。

“当守门员发起进攻位置很高时，他们能够对进攻做出预判。”祖贝布勒说道，“守门员冲出罚球区并表现勇敢，这是非常积极的变化。”



足球运动员在精神压力下运动表现能力的量化分析

洛特·布兰森，彼得·罗伯雷希茨，扬·范哈伦，杰西·戴维斯

1 简介

英超联赛 2011/2012 赛季的最后一个比赛日产生了一场经典的比赛，赛前曼城需要一场胜利才能获得赛季冠军，然而在这场比赛中却以一球落后女王公园巡游者，这种局面一直持续到伤停补时阶段。此时此刻，球员们会有怎样的发挥成了关键的问题。在精神压力下，他们会超常发挥还是失常？接下来的比赛书写了曼城队史上经典的一页：埃丁·哲科打进扳平一球，随后塞尔吉奥·阿奎罗攻入了致胜一球，两人奇迹般地让曼城夺得了 44 年以来第一个联赛冠军。尽管并非所有足球比赛都像这一场一样压力巨大，但球员的确需要面对各种情况，而这些情况会带来很大的精神压力。

绝大多数现有的足球运动员表现指标关注的都是球员的技能 and 体能，往往忽略了球员的技能 and 体能会受精神压力的影响。然而，在分析球员或球队的发挥时，精神压力是一个反复出现的概念。因此，精神压力对球员的影响的量化参数能为足球俱乐部提供四个重要的应用案例。

(1) 球员转会：俱乐部显然希望自己的球员在精神压力下能有良好的发挥。如果能更准确地评估转会目标球员的这一素

质，就能对目标球员多一份了解，从而对球员转会操作提供帮助。例如：我们确认侯瑟姆·奥亚尔和谢尔丹·沙奇里能替代莱斯特城队球星里亚德·马赫雷斯。

(2) 训练：如果俱乐部知道某名球员在某些压力情况下会做出欠佳的选择，那么可想办法解决。例如，教练可指导这名球员在具体情况下该怎样处理。此外，也可以设计定制的训练课，帮助球员克服这些困难。例如：我们发现奥兰多城（北美职业足球大联盟）的前锋多姆·德怀尔在压力下经常无谓犯规，而这个缺点是能改正的。

(3) 战术决策：当足球比赛中出现压力很大的情况时，某些战术会更有价值或更容易带来成功。了解这些战术能帮助教练对比赛做出针对性的战术布置。例如：我们发现曼联在比赛中面临压力很大的情况时，没有充分意识到传中的战术价值。

(4) 首发和替补：了解球队的哪些球员在压力下能发挥良好，这一点能帮助教练进行人员布置，决定让哪些球员进入重要比赛的首发阵容，在场上情势紧张时替换哪些球员上场或下场。例如：尤文图斯的四名中后卫贝纳蒂亚，博努奇，基耶利尼和鲁加尼在常规情况下实力相当，但在压力很大的时候，贝纳蒂亚和博努奇的发

挥更好。

对以下问题进行过客观分析的论文还没有，本文将开创先河：当球员在足球赛场上面临很大的精神压力时，他的发挥会受到怎样的影响？为了研究这一问题，我们采取了以下方法：

(1) 针对赛场上的每一种情况，我们开发了机器学习模型，运用多种赛场情景（例如，比赛是否竞争激烈，两队的差距和联赛排名）和目前的比赛状态（例如比分和比赛剩余时间）来评估控球球员承受了多少精神压力。

(2) 由于精神压力会以多种方式影响球员的发挥，例如他的决策（即他怎样从几个选项中选择下一个动作）或在做出决策后他的发挥怎么样，所以我们开发了机器学习模型来评估每个行为的三个方面：行为的选择，行为的执行，以及行为对比赛得分的预期贡献。

(3) 为了评估一名球员怎样应对精神压力，我们比较了不同压力级别下他的表现指标。

运用这一方法，我们分析了7个联赛6858场比赛的事件数据。与追踪数据不同，全球很多联赛都提供了事件数据。因此，球队不仅能用我们的算式（metric）更好地评估自己的球员在比赛关键时刻的贡献，也能据此交易球员，并比较自己的球员与其它球队和联赛的球员。

2 测量精神压力

我们假设巨大的精神压力可能对球队实现自身目标产生重大影响。具体来说，有两个因素会影响压力水平：

(1) 比赛本身的纷纷扰扰会影响球员的赛前压力。例如，竞争激烈的比赛或直接影响升降级的比赛会比一场赛季末无关大局的普通比赛产生更大的压力。

(2) 比赛中发生的事件会影响比赛中的压力水平。比赛时双方比分接近时会让压力上升，尤其当比赛时间越来越少时，一个进球会让球队更有可能得到理想的结果。相反，如果双方比分悬殊，那么一个进球对结果影响很小，这时球员的压力会减小。

因此，我们设计了两个全新的算式：一个用来检测赛前压力，另一个检测比赛中压力。具体而言，比赛 g ，目前处于比赛 xt 状态，总的比赛压力是这两个算式之和：

$$\text{Pressure}(g,xt)=\text{Pressurepre-game}(g)$$

$$\text{Pressurein-game}(xt)$$

接下来，我们将分析怎样计算每一个算式。

2.1 赛前精神压力

理想情况下，每场比赛都有一个对应的压力水平级别（例如，零压力、压力较低、压力正常、压力较高）。然而，这些压力水平级别并不存在，而且要精确地定义各个级别，并以人工标记这些数据可能有难度。不同的压力级别之间存在顺序关系，我们的研究是在这个基础上进行的。此外，



英超联赛2017/2018赛季
埃弗顿俱乐部

赛前压力

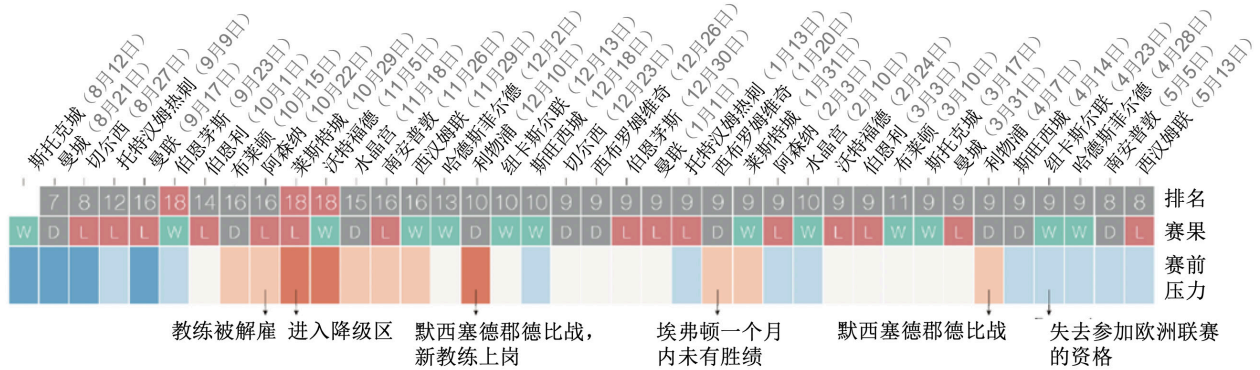
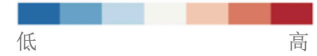


图 1 英超 2017/2018 赛季埃弗顿赛前压力水平。埃弗顿赛季初的糟糕战绩让压力值开始上升，当他们进入降级区时压力值到达顶峰。新教练阿勒代斯上岗后带领埃弗顿回到了积分榜中游，压力下降。

判断两场比赛中哪一场的重要性更高比为这两场比赛设定压力级别要简单。因此，我们收集了足球专家们给出的判断，并用这些数据来训练基于机器学习的排列程序。排列程序学习了为两场比赛设定分数，分数更高的比赛说明赛前压力更高。这些分数定义了我们的压力算式。

图 1 显示了埃弗顿在 2017/2018 赛季每场比赛的赛前压力级别。在第六场比赛结束后压力开始上升，原因是球队表现欠佳，这导致教练在第九场比赛结束后被解雇。埃弗顿一直在降级区徘徊，压力值一直居高不下。在取得几场胜利后，埃弗顿的排名有所上升，压力也减小了。赛季收官阶段，压力一直保持在适中水平，直到埃弗顿失去参加下赛季欧洲联赛的资格为止。

模型

比赛的风险程度和赢球的重要性是导致压力的重要因素。因此，我们需要建立描述这种关系的特征。我们在研究中考虑了以下四大类特征：

球队目标：每支球队都有赛季目标，例如夺冠或避免降级，目标会影响球队的赛前压力级别。我们按以下方式计算球队的赛季目标，根据上一赛季的成绩，球队最好的 20 名球员的转会价值，租借球员费用和教练的名气值（这反映了俱乐部的名气有多大）将每个联赛中的球队分成四组。

比赛重要性：要统计一场比赛对球队赛季目标有怎样的影响，这需要评估这场比赛的结果会对球队实现赛季既定目标（例如避免降级）的概率有怎样的影响。我们使用了基于埃洛模型（Elo-based model）来解决这个问题。根据球队在赛前的联赛

排名，我们模拟了这个赛季的结果。接下来，我们使用了肯德尔等级相关系数（Kendall-Stuart tau-c）来计算每场比赛结果（胜 - 平 - 负）的关系和本赛季预计球队成绩（例如，降级，夺冠）。

近期表现：足球俱乐部也会因近期的战绩而感到压力。尤其是大的俱乐部，几场比赛连续表现欠佳将累积起压力。我们使用近期 5 场比赛的得分，以及使用埃洛模型来估算实际与预期成绩的差距，来计算这一个压力源。

比赛环境：一场比赛的特定特征会影响压力，包括比赛场地（即主场或客场），由足球教练数据决定的对手的竞争性，观看比赛的现场观众数量和教练执教的时间。

我们请了一些足球专家给出成对的比赛排名，并用它们去拟合梯度增强排名树算法模型（Gradient Boosted Ranking Trees model）数据和模型详情请参阅附录 A.1。

2.2 比赛中的精神压力

在比赛进行中，根据比赛的进程，球队的压力会随时间而变化。我们提出，当打进一球能增加球队获得预期结果的机会时，压力会增加，而当打进一球对预期结果影响有限时，压力会减小。分析进球或丢球对预期结果的影响时，有一个办法是分析两种情况的差异，前一种情况是在目前比赛状态下球队获胜的可能，后一种情况是假设主队或客队额外打进一球时球队获胜的可能。

如果球队在比赛中的某个时间点 t 打入一球，那么球队获胜增加的的概率与打平增加的的概率之和，就是球队在比赛中这个时间点 t 的压力级别。英超 2017/2018 赛季埃弗顿和沃特福德的比赛过程十分激烈，图 2 显示了这场比赛球队的压力级别和获胜可能的变化。我们的压力模型会对影响球队获胜可能的事件（例如进球）作出反应。例如，在沃特福德开场先进一球后，埃弗顿的压力上升了，而在沃特福德攻入两球后，埃弗顿的压力下降了。然后，当埃弗顿打入第一球后，他们的压力有了显著增长，在他们打入扳平比分的一球时，沃特福德的压力级别也提高了。

模型

足球比赛进球不多的特点让建立赢球概率模型变得更为困难。因此，我们没有建立预测胜 - 平 - 负概率的模型，而是建模预测了在比赛剩余时间内可能的进球数量。通过预估影响胜 - 平 - 负赛果的每一种比赛发展机率，我们的模型能评估出在比赛尾声球队获胜，打平和告负的概率。具体来说，我们会根据比赛在 t 时刻当前的局面，建模评估两支球队从 $t+1$ 时刻到比赛结束的进球数量概率分布。由此我们可推导出两支球队的进球数量概率分布，球队在比赛中的进球数即为在 t 时刻已打进的进球数和 t 时刻以后预期的进球数之和。

由于补时的存在，比赛的时间长度存在不确定性，因此我们将比赛时间划分为 100 个时段，每个时段对应 1% 的比赛时间。

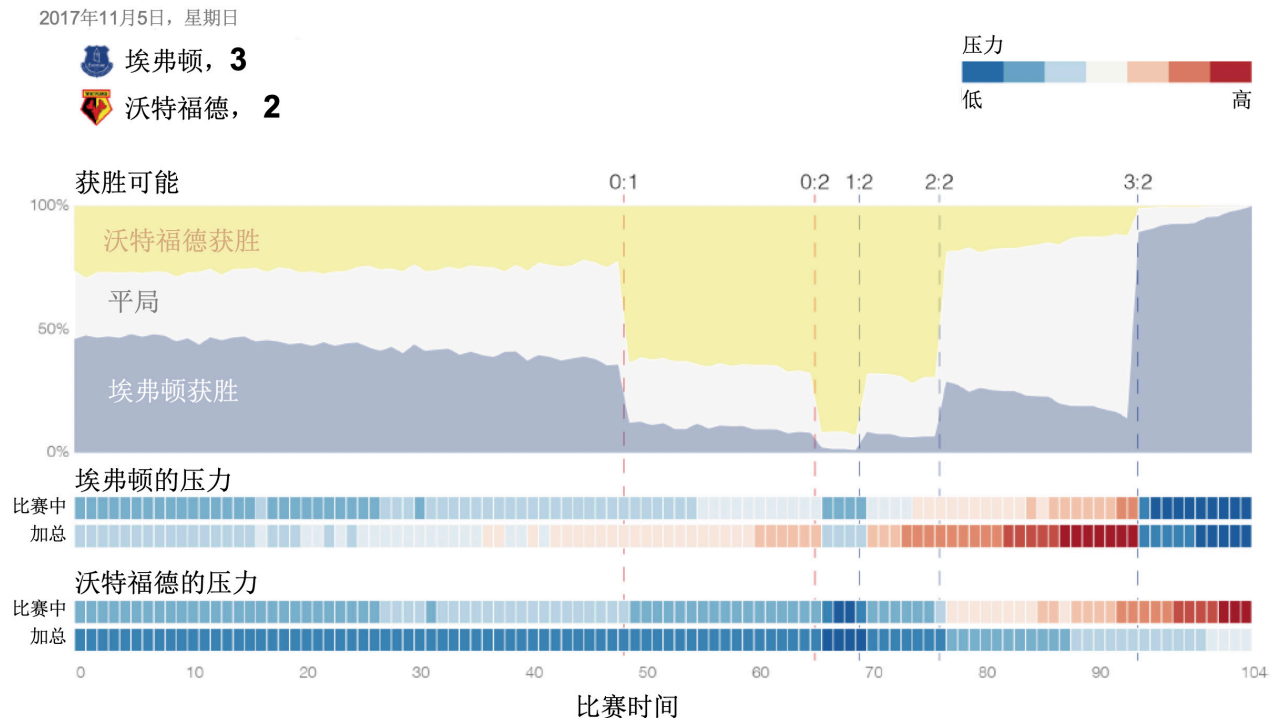


图 2 2017/2018 英超赛季埃弗顿 3 比 2 战胜沃特福德的比赛中压力级别变化（底部）和获胜可能变化（顶部）。当进球能增加球队获得预期结果的机会时，压力会增加，而当打进一球对预期结果影响有限时，压力会减小。在比赛开始时，由于仍有足够的时间追平进球，并赢得比赛，所以两队的压力级别都比较低。

根据独立二项分布原理，我们建模预测了主队 ($y > t, \text{home}$) 和客队 ($y > t, \text{away}$) 在 t 时刻后进球的数量：

$$y > t, \text{home} | \theta t, \text{home} B(T - t, \theta t, \text{home}),$$

$$y > t, \text{away} | \theta t, \text{away} B(T - t, \theta t, \text{away}),$$

算式中的参数 θ 表示第 t 时刻每支球队预计的进球强度。我们根据目前的比赛局面来预测这些进球强度，预测时会考虑以下特征：已打进的进球数，两队的比分差距，红黄牌数量，两队埃洛评级 (Elo-ratings) 差距，在 t 时刻之前 10 个时段内进攻传球的平均数，在 t 时刻之前 10 个时段内争抢获胜的平均概率。由于比赛情形

的特征会随时间发生非线性变化，因此我们使用临时随机过程来对进球强度建模。该方法可以传递信息，还能在时段之间执行连贯推理。

我们使用了基于 PyMC3 's 的自动微分变分推理 (ADVI) 算法，用从 2014/2015 到 2017/2018 赛季英超、西甲和德甲的数据来验证我们的比赛模型。本模型的详细变分推理过程请参阅附录 A.2。

3 衡量球员的表现

压力可能以不同的方式影响球员的表

现。因此，当球员在比赛中做一个动作时，理论上可以从三个重要的方面来评估这个行为：

(1) 对比赛的贡献：这个行为对增加球队得分或阻止对手得分而言有多大的帮助？例如，一个高质量的直塞传球，让球友能形成面对守门员的单刀机会，这个行为就很有价值。

(2) 决策的质量：球员在当下是否做出了最优决策？即使这个行为有助于球队，但仍有可能有更好的行为存在。亦即球员可以做出更好的决策。例如，在 2018 年 10 月 29 日托特汉姆热刺对阵曼城的比赛中，大卫·席尔瓦离球门只有几米远，而且有射门机会，他这时选择了传球，没有射门（见

图 3）。这显然并非最优决策。

(3) 执行的质量：球员做出决定后，对动作的执行质量如何？球员也许决策正确，例如在有射门机会时选择射门，但射门的质量却很差（例如，射门后足球从球门上方越过）。

近年来，研究人员提出了一些球员表现指标来评估球员行为的贡献。不过，据我们所知，尚未有球员表现指标能衡量球员决策或行为执行的质量。因此，我们使用了现有的指标来计算行为的贡献，并提出两个新算法来计算球员决策和行为执行的质量。下表给出了一些图文并茂的例子，列举了由我们算法得出的影响低、中、高得分的因素。

案例	贡献	决策	执行
中场球员在离球门 35 米处远射得分，而直接传球可以创造队友面对守门员时单刀的机会。	高	低	高
左后卫发起 60 米的长传球，球传给了右边锋。	中	中	高
中卫想传直塞球，想让前锋在对方禁区内得球。然而，直塞球并不精确，被对方守门员没收了。	低	高	低



图 3 2018 年 10 月 29 日托特汉姆热刺对阵曼城的比赛中，大卫·席尔瓦在离球门只有几米远，有射门机会时没有射门，选择把球传给斯特林。这是截止 2018 年 11 月 11 日，2018/2019 赛季英超比赛中得分最低的球员决策。

3.1 贡献率

我们在计算贡献率时用到了德克鲁斯 (Decroos) 等人提出的算式。该算式考虑了 22 种动作，例如射门、传球、过人等，每种动作都有成功或失败的可能。无论是成功或失败，每个动作的结果都会改变比赛的局面。计算贡献率的目的是衡量这些改变比赛局面的动作的价值。比赛局面的价值在于产生进球的概率，而该算式会算出动作执行前后比赛局面的变化，从而衡量动作的价值。比赛局面为 s_i 时，会造成后果 o_i 的动作 a_i 的贡献价值按下式计算：

$$CR(s_i, a_i, o_i) = V(s_j) - V(s_i),$$

当比赛局面为 s_i 时，会造成后果为 o_i 的动作 a_i 在执行后将比赛局面改变为 s_j ，而 $V(s)$ 是比赛局面为 s 时的价值。

动作的贡献取决于几方面因素，包括它的类别，在场上的位置，产生的结果。动作执行后，如果将比赛局面改变为有利球队得分（例如通过成功的过人和传球，足球转移到了更有威胁的区域），或减少对手得分的机会（例如守门员防住了对手的射门），那么该动作的贡献率为正。反之，动作执行后，如果将比赛局面改变为不利球队得分（例如失败的传球），或增加对手得分的机会（例如对手拦截了传球），那么该动作的贡献率为负。

3.2 决策得分

面对各种比赛局面，球员在决定动作

时考虑几种可能。为了了解球员是否做出了好的决策来执行动作，我们必须（1）对动作造成的结果进行抽象化处理（即我们要考虑动作成功与失败会造成的结果），（2）考虑球员选择的动作与比赛局面导致的其它动作的关联性。为了得出这一结论，我们比较了球员选择的动作的预期贡献率和比赛局面下其它所有可能的动作的预期贡献率的差异，并据此判断球员决策的质量。比赛局面为 s_i 时，动作 a_i 的决策得分按下式计算：

$$DR(s_i, a_i) = ECR|s_i, a_i - ECR|s_i.$$

第一项是选择的动作的预期贡献率。第二项要求判断出比赛局面下其它可能动作，由于我们使用的是事件数据，并不知道球员在场上的具体位置，因此难度很大。因此，我们通过评估当前比赛局面下一个动作的贡献率，以此来预估所有可能动作的预期贡献得分。

计算选择的动作的预期贡献率

由于动作可能成功或失败，所以我们按两种结果的加权和来计算选择的动作的预期贡献率：

$$ECR|s_i, a_i = P(o_i+) \cdot CR(s_i, a_i, o_i+) + P(o_i-) \cdot CR(s_i, a_i, o_i-)$$

上式中 $P(o_i+)$ 和 $P(o_i-)$ 分别代表了动作成功和失败的概率。

为了预估动作成功的概率，我们使用每一类动作（例如传球、过人或拦截）的历史比赛数据来调试二元分类器。我们使用

了梯度提升决策树算法模型，原因在于它预计的概率准确性较高（详见附录 A3），这对于本次任务很重要。正值代表了成功的动作（例如成功传球到队友脚下或射门打进），而负值代表了失败的动作（例如射门不中）。相关的特征包括动作开始和结束的位置，球员做动作时使用的身体部位和按照顺序前面的动作开始和结束的位置。由于有些类别的动作的结束位置可能含有动作成功的信息（例如射门），所以我们在计算中排除了这一特征。

预估所有动作的预期贡献率

我们统计了与当前比赛局面高度相似的比赛局面下球员执行的动作的历史数据，据此预估出了当前比赛局面下所有可能动作的预期贡献率。为实现这一目标，我们调试了梯度提升决策树算法模型，使用它根据当前比赛局面推算出下一个动作的贡献率。我们使用了以下特征来描述当前的比赛局面：足球当前位置，前两个动作开始和结束的位置，前两个动作的类别，根据足球移动的距离和耗时计算的动作序列（由前两个动作构成）速度。对该模型的正式定义和评估详见附录 A.4。

3.3 执行得分

动作执行得分的目的是评估球员执行动作的好坏，不会考虑动作的选择是否恰当。对于高难度动作，例如成功传了直塞

球或接到队友的长传，我们会在得分上奖励球员。同样，如果球员在做简单动作时失败，例如在没有压力的前提下横向传球给无人盯防的队友失败，那么我们会在得分上惩罚球员。为实现这一目标，我们比较了动作的实际结果（例如传球是否传到了队友脚下，或射门是否打进）和动作的预期成功概率的差异，据此计算出了执行得分。举例来说，洛杉矶银河与洛杉矶 FC 的比赛中兹拉坦·伊布拉西莫维奇的一脚高难度远射进球得分，让守门员无计可施，我们的模型给了这个动作执行很高的分数。假设动作 a_i 产生了结果 o_i ，其执行得分按下式计算：

$$ER(a_i, o_i) = [o_i+] - P(o_i+)$$

如果结果 o_i 成功了，那么 $[o_i+]$ 的值为 1，否则为 0，而 $P(o_i+)$ 由上一个动作成功预测算式得出。

4 衡量球员在精神压力下的表现

为了分析球员在不同精神压力级别下的发挥，我们需要结合我们的压力模型和三个表现指标。首先，我们使用了我们的压力模型来计算所有已分析的比赛中的所有局面下精神压力级别的排名。我们将各种局面下的压力级别归类如下：

- (1) 高压局面排在前 20%；
- (2) 中压局面排在前 60%；
- (3) 低压局面排在前 20%。

第二，根据我们的三个表现指标，我们分别使用它们来计算所有统计的比赛中

的球员动作得分。第三，对于已知球员或球队，我们聚合这三个表现指标，来计算执行动作时球员的精神压力级别。在我们的分析中，在球员表现算式中考虑了以下数据的聚合：

- 比赛 90 分钟内贡献率均值；
- 每种动作类别在 90 分钟比赛内的贡献率均值；
- 90 分钟比赛内贡献率的百分等级；
- 决策得分均值；
- 每种动作类别的执行得分均值；
- 每种动作类别的执行得分的百分等级。

由于球员的上场时间严重制约了他的贡献，所以我们通常报告的是比赛 90 分钟内归一化的贡献率。由于不同类别的动作的执行得分分布不同（例如，射门通常的成功率比传球的成功率要低），所以我们没有对动作类别的执行得分进行数据聚合。

5 应用案例

我们分析的事件数据来自于英超、西甲、德甲、意甲、法甲、荷甲和北美职业足球大联盟（MLS）联赛。针对每个联赛，我们考虑了 2016/2017、2017/2018 和正在进行的 2018/2019 赛季的数据，收集到的数据的截止日期是 2018 年 11 月 11 日。另外，我们的分析也纳入了 2014/2015 和 2015/2016 两个赛季的英超联赛，用于分析球员的长期发展。

在我们分析的所有比赛局面中，我们

针对球员执行的动作计算了压力级别和我们的三个表现指标。我们比较了球员 90 分钟比赛内的贡献均值、决策均值、低压力和高压力情况下的动作执行得分。我们只计算了比赛时间总计在 900 分钟以上的球员，而且高压力和低压力情况下的比赛时间至少要分别达到 180 分钟。

在本节的剩余部分中，我们将展示我们的算式怎样帮助俱乐部和教练来解决以下关键问题：

(1) 球员收购：球员在压力下是否能有好的发挥？

(2) 训练：球员在特定压力情况下有哪些反复出现的决策失常问题，可在训练中加以解决？

(3) 战术指导：哪些战术计划在高压情况下更容易成功？

(4) 首发和替补：在关键比赛时教练应安排哪些球员打首发，或在比赛局面紧张时做出怎样的人员轮替安排？

5.1 球员收购：替代莱斯特城队的里亚德·马赫雷斯

在这个使用案例中，我们将假设自己是 2018 年夏季莱斯特城队技术总监。在满足阿尔及利亚籍前锋里亚德·马赫雷斯的愿望，希望前往联赛排名前六名的球队踢球后，技术总监必须找到一位合适的球员来替代他。因此，本次分析仅考虑截止 2017/2018 赛季的数据。

由图 4a 可见，我们计算了里亚德·马



寻找马赫雷斯的替代球员

推荐的候选人:

- 加扎尔
- 奥亚尔
- 沙里奇
- 普通中卫

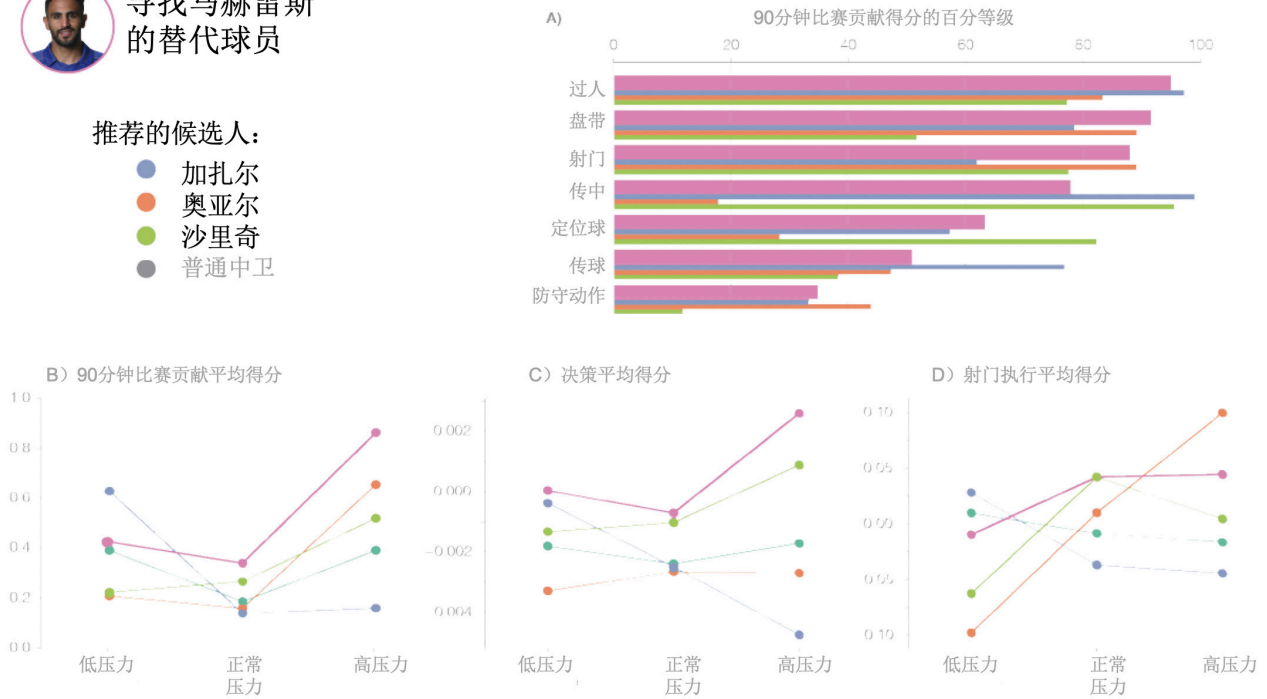


图 4 里亚德·马赫雷斯和替代他在莱斯特城位置的候选球员比较：奥亚尔、加扎尔和沙里奇。表 A 比较了这些球员每个动作类别的贡献得分。表 B、C 和 D 比较了他们在压力下的发挥。

赫雷斯每个动作的贡献率，其显示这位前锋对莱斯特城队的贡献主要在于过人、盘带、射门和传球的贡献。另外，我们也对他的 986 分钟高压情形下的发挥做了分析，显示他在高压情形下发挥尤其良好。马赫雷斯在高压情形下的决策和射门都比低压力情形下更好，他对莱斯特城的贡献在高压情形下比低压力情形下要大得多。

纵观欧洲五大联赛，我们找到了 28 岁以下的三位候选球员，他们和里亚德·马赫雷斯一样，擅长过人、盘带、射门和传球：谢尔丹·沙奇里（斯托克城），侯瑟姆·奥亚尔（里昂）和拉希德·加扎尔，而加扎尔最终于 2018 年 8 月 5 日从摩纳哥转会到莱斯特城。图 4b 显示了比赛 90 分钟内这三

名球员在低压力和高压情形下的累计贡献。虽然贡献幅度小于马赫雷斯，但奥亚尔（547 分钟高压数据）和沙奇里（1477 分钟高压数据）在压力下的贡献都有所增长。与之相反，莱斯特城队选择的加扎尔（358 分钟高压数据）在压力下似乎发挥失常。图 4c 和 4d 显示决策和射门执行的得分均值与压力级别的关系。马赫雷斯和沙奇里在高压时决策能力有所提高，而只有马赫雷斯和奥亚尔能在高压时提高他们的动作执行得分。在压力之下，加扎尔的决策能力下滑，在动作执行上也遇到了问题。

我们得出的结论是，加扎尔也许并非莱斯特城购入并替代马赫雷斯的最佳选择。我们注意到马赫雷斯在高压状态下能有

好的发挥，然而加扎尔在高压状态下却发挥失常。因此，我们的分析认为奥亚尔和沙里奇技能与马赫雷斯相当，包括在高压状态下发挥出的水平，所以这两人应该是更好的选择。

5.2 训练：多姆·德怀尔的犯规

如果俱乐部能知道某位球员在特定的紧张局面下会出现决策失误，那就能解决这一问题。教练能让球员知道他的决策会对球队带来怎样的负面影响，并教导他做出不同的决策。此外，也可以设计特定的训练课，帮助球员克服这些困难。为了演示这个应用案例，我们比较了奥兰多城中锋多姆·德怀尔在低压力状态（916分钟）和高压状态（2135分钟）下的犯规情形。如图5所示，德怀尔在高压状态下，在对方罚球区内的犯规明显增多。此外，我

们的算式计算出他在高压状态下的大部分犯规在决策得分上得到了负分，说明这些犯规直接导致对方获得球权（而不是，比如说破坏了对方的反击）。他的教练可以使用这些信息，在训练时在这一点帮助德怀尔提高。

5.3 战术计划

曼联：过人还是传中？

曼联在高压比赛状态下（1818分钟高压数据）依靠过人，其90分钟比赛的过人次数从19.41次翻倍至43.88次。然而，如图6所示，对于进攻球员而言，除了拉什福德以外，在压力下其他人比赛90分钟过人的贡献平均得分在降低，所以这并不是一个很成功的战术。我们的算式显示传中会是一个更成功的战术。尤其是像瓦伦

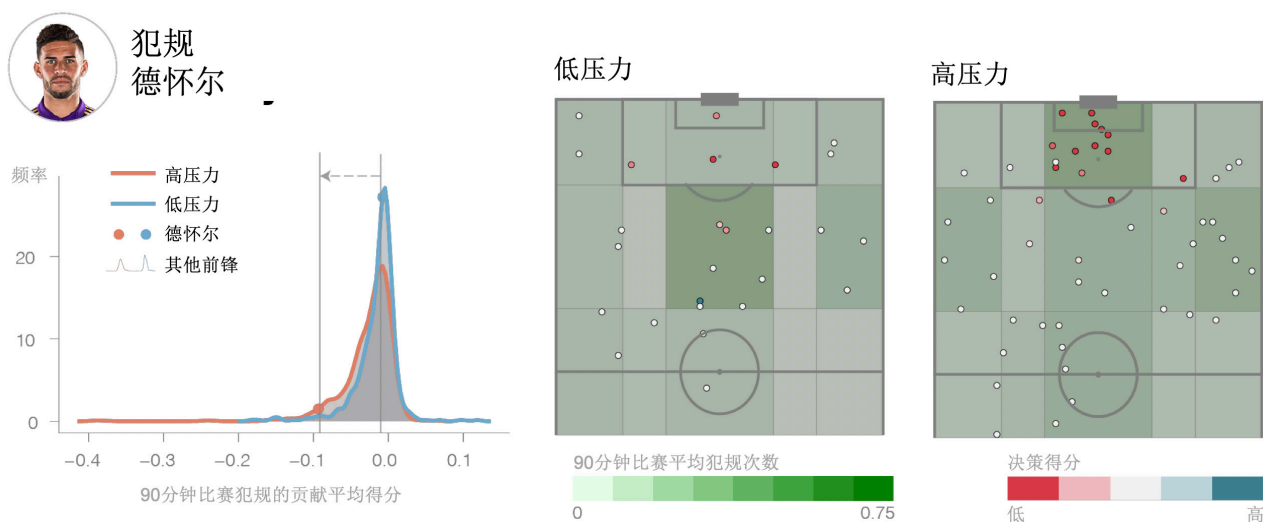


图5 多姆·德怀尔在低压力和高压状态下的犯规比较。在压力下，德怀尔的无谓犯规明显增多，直接将球权交给了对手。

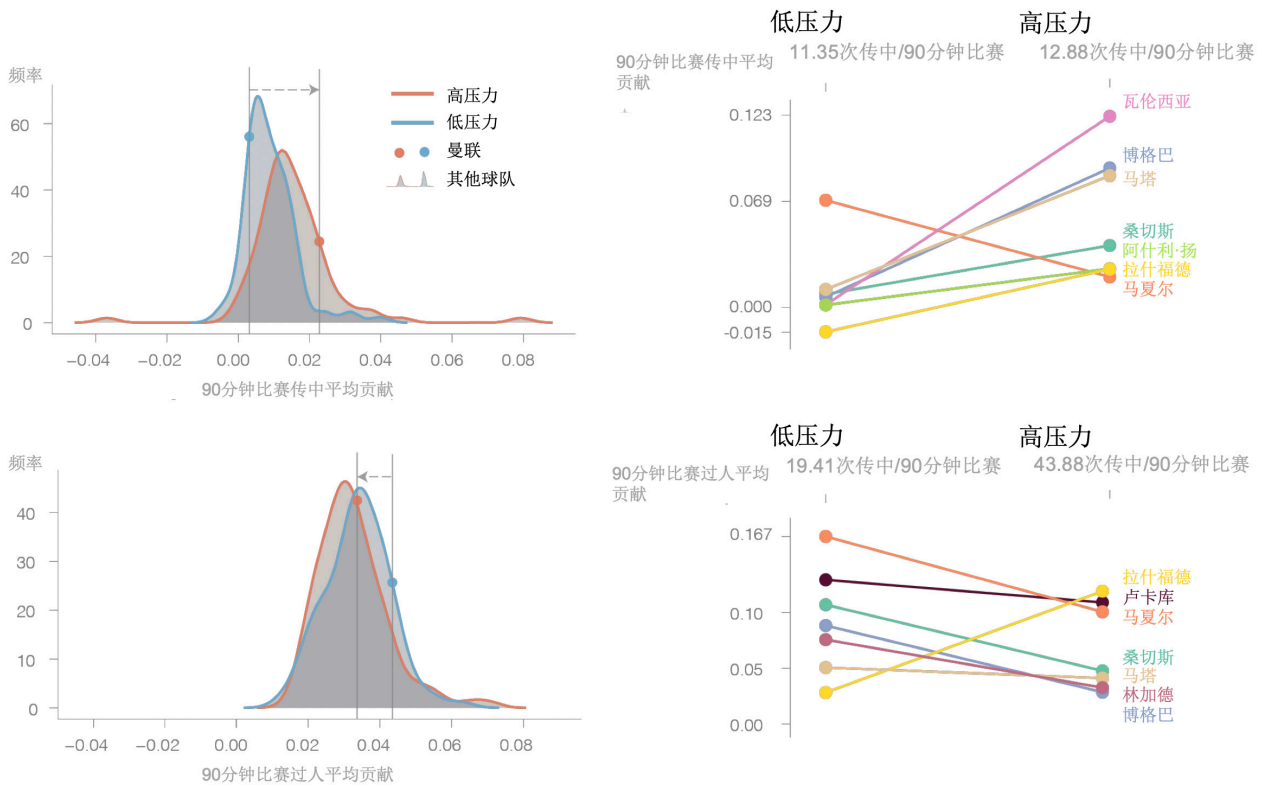


图 6 在高压力情景下，曼联通常会使用过人形成突破。然而，球员在压力下过人的贡献会减少，而传中的贡献会增加。这一数据显示在激烈的比赛中，传中可能比过人在战术上价值更高。

西亚、博格巴和马塔等球员在压力下的传中能极大地增加得分机会。此外，考虑到曼联拥有卢卡库和费莱尼这样的球员，他们接到传中后有破门的能力。

掷界外球：高压力状态下的更佳决策

掷界外球是我们观察到在高压力和低压力状态下差异最明显的动作类别。几乎所有球队掷界外球时的贡献和决策得分都会增加，而他们执行的得分保持不变。这个现象很有趣，说明球队在比赛早期掷界外球时错过了创造更多威胁的机会。另外，利物浦的教练尤尔根·克洛普也意识到掷

界外球的重要性增加了，他在 2018 年的夏天聘用了一位专职掷界外球教练。作为范例，图 7 显示了莱斯特城在比赛 90 分钟内掷界外球的决策得分，以及在低压力和高压力状态下掷界外球的差异。也许有人认为在低压力状态下掷界外球的决策得分更低是因为向后场掷球（例如确保控球），但实情并非如此。恰恰相反，在低压力状态下球员更倾向于向前场掷球。在靠近对手球门线附近掷界外球是个例外，在高压力状态下球员会更多地把球投掷到罚球区内。莱斯特城队教练根据这一信息训练他的队员，调整他们在比赛中高压状态下掷界外球的表现。同样，莱斯特城队的对手

掷界外球 莱斯特城

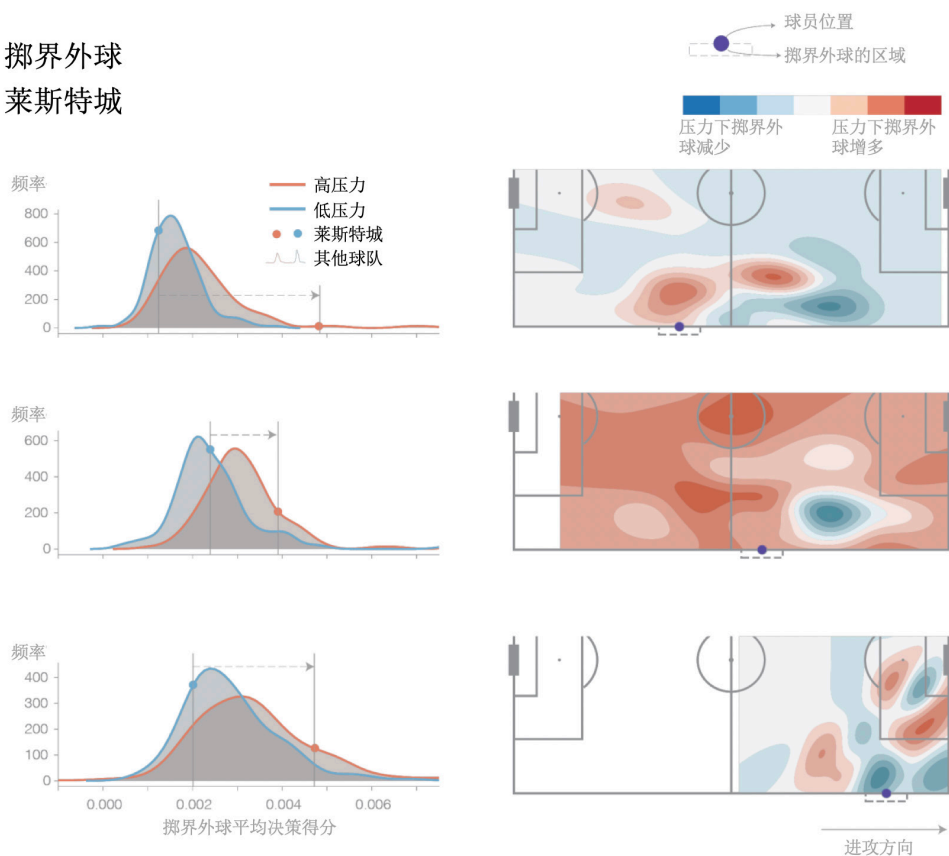


图7 莱斯特城90分钟比赛掷界外球决策平均得分与英超联赛平均水平比较（左），和这些界外球掷入场内后落地区域的概率图（右）。数据建模后形成了二维高斯分布，红色区域显示莱斯特城队在压力下更有可能将球掷入场内的区域（比较压迫较低时的区域），而蓝色区域是他们不太可能将球掷入场内的区域。

球队教练在制定战术时可讨论这些信息，决定在和莱斯特城比赛时，对手在不同比赛局面下掷界外球时他的球员应怎样调整防守站位。

5.4 首发和替补

让合适的球员上场比赛是教练的重要任务之一。归根到底，这个任务要求教练预测哪些球员在既定条件下能有最好的发挥。球员在压力下的预期表现因此成为了教练决策时的重要变数。我们的算式能帮

助教练决定哪些球员能在高压力比赛中打首发，或哪些球员在高压力比赛中应该轮换。例如，图8对比了尤文图斯中卫马德希·贝纳蒂亚（457分钟高压力比赛数据），莱昂纳多·博努奇（201分钟高压力比赛数据），吉奥吉奥·基耶利尼（470分钟高压力比赛数据），达尼埃莱·鲁加尼（278分钟高压力比赛数据），巴黎圣日耳曼中后卫安赫尔·迪马利亚（451分钟高压力比赛数据），朱利安·德拉克斯勒（359分钟高压力比赛数据），切尔西的前锋奥利维尔·吉鲁（617分钟在阿森纳队的高压力比赛数据

和 127 分钟切尔西的数据），阿尔瓦罗·莫拉塔（477 分钟高压比赛数据）90 分钟比赛的平均贡献。针对每一种情况，我们的算式得出的结果说明应该派不同的球员参加高压力的比赛。尤文图斯的各位中锋在普通压力下比赛中的平均贡献相似，贝纳蒂亚和博努奇在高压下的发挥比基耶利尼和鲁加尼更好。迪马利亚在高压下也发

挥得比德拉克斯勒更好。虽然莫拉塔的整体贡献比吉鲁好，但他在压力下时似乎会发挥失常，而吉鲁却证明自己在切尔西是善于进球，杀死比赛的前锋。但要注意，我们统计的数据中包含了吉鲁在阿森纳的数据，因为其在切尔西的高压比赛中的数据不足。

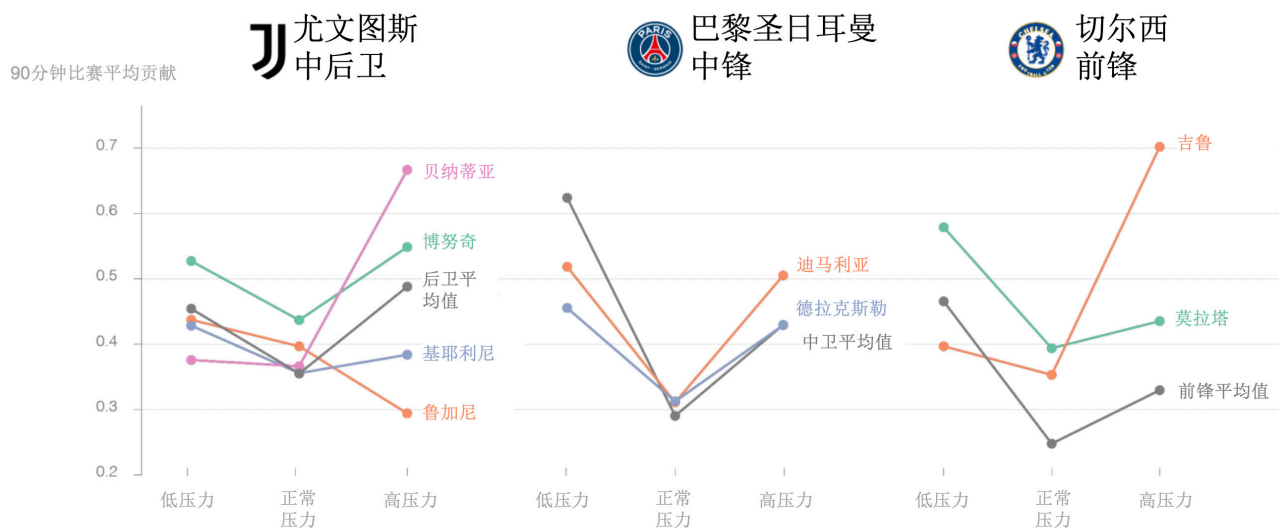


图 8 尤文图斯后卫，巴黎圣日耳曼进攻型中卫和切尔西前锋 90 分钟比赛内平均贡献比较。这些图展示了在高压比赛中选派不同的球员上场是有意义的。因为吉鲁在切尔西的高压比赛中的数据不足，所以我们的计算中包括了他在阿森纳的数据。

6 观察

本节提供一些基于我们的球员发挥和压力算式分析而观察到的有趣数据。首先，我们将介绍内马尔在压力下发挥失常。其次，我们分析三位北美职业足球大联盟明日之星在压力下的发挥。再次，我们会指出利物浦倾向于购买在高压下发挥得更好的球员。最后，我们展示了我们的研究结

果并非是因为球员因疲劳而产生的状态下滑。

6.1 内马尔在压力下发挥失常

图 9 显示内马尔（先后效力于巴塞罗那和巴黎圣日耳曼）在高压力和低压力状态下发挥的变化。在低压力状态下（基于 997 分钟的低压力数据）内马尔是一位 90

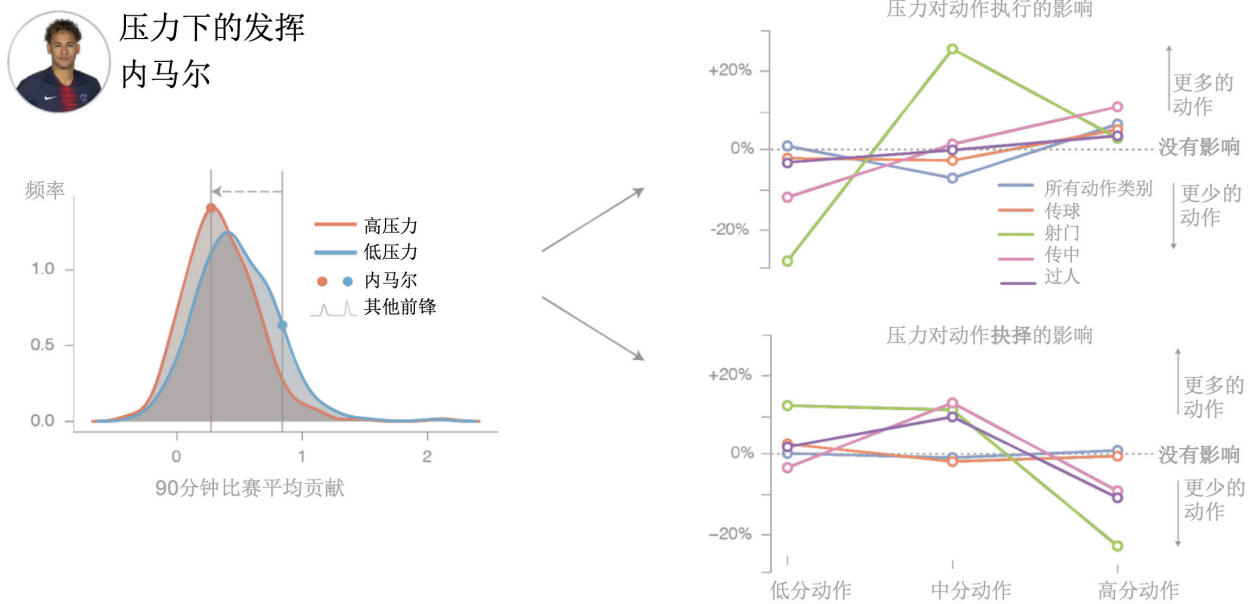


图9 在压力下,内马尔从90分钟比赛平均贡献得分高的球员变成了平均贡献得分普通的球员。他在发挥上的失常主要原因在于高压状态下决策不当。

分钟比赛时间内平均贡献很高的球员，然而在高压力状态下（基于472分钟的高压力数据），却成为了一名平庸的球员。进一步的分析显示他的发挥失常原因在于他的动作决策很糟糕。内马尔在高压力状态下做的决定更多的得到了低分或中分，只有少数决定能得高分。相反，内马尔在压力下的动作执行不会失常，各种动作执行反而更好。

6.2 北美职业足球大联盟的明日之星

近来，欧洲俱乐部越来越多地从北美职业足球大联盟中寻找好的苗子，拜仁慕尼黑签下了阿尔方索·戴维斯，莱比锡签下了泰勒·亚当斯就是其中两例。亚特兰大的中锋埃塞基耶尔·巴尔科是另一位很快能到欧洲踢球的北美职业足球大联盟球

员。他们三位代表了北美职业足球大联盟2018赛季90分钟比赛内U20球员平均贡献得分最高的球员。

图10显示了三位球员在2016、2017、2018赛季中处于不同压力级别时防守动作的贡献得分的百分等级（例如，拦截、抢断、解围和犯规），进攻动作的贡献得分的百分等级（例如传中，射门，过人和盘带）和传球的贡献得分的百分等级。由于每类动作每90分钟贡献得分的百分等级分布不一样，所以我们展示了每一类动作的得分的百分等级。戴维斯（1271分钟高压数据）在压力下似乎表现不佳，他是联盟中低压力状态下发挥最好的球员之一，然而在高压力状态下他的发挥却低于平均值。他在90分钟比赛时间内的进攻动作、防守动作、传球的贡献得分会随着压力的增加而减少。与之相反，亚当斯（1804分钟高压数据）

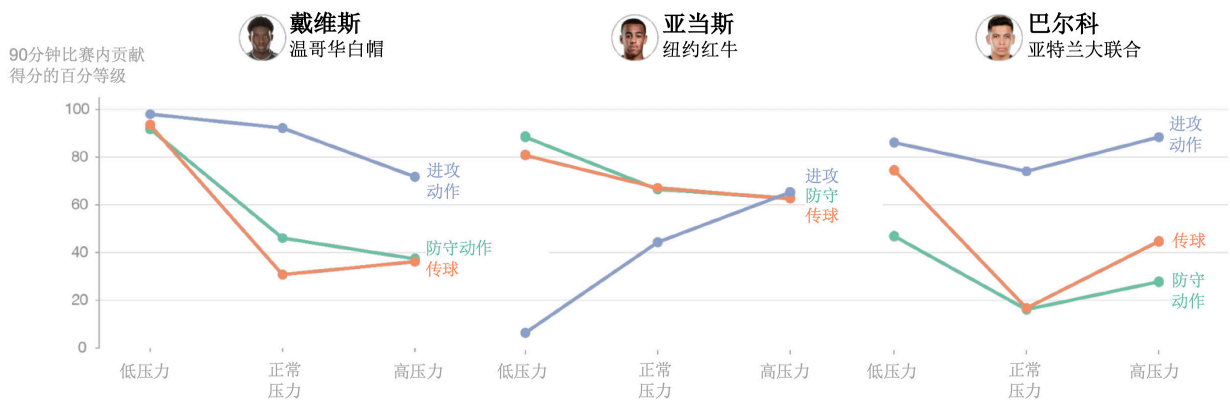


图 10 北美职业足球大联盟的明日之星阿尔方索·戴维斯、泰勒·亚当斯和埃塞基耶尔·巴尔科在不同压力级别下的贡献得分的百分等级。

和巴尔科（760 分钟高压数据）在压力下的发挥有好有坏，有些动作执行得更好，有的执行得更差。

加盟利物浦后高压状态下的比赛时长比他在老东家罗马时要短。

6.3 利物浦的球员交易

我们分析了利物浦自 2017/2018 赛季开始的球员交易，数据显示红军倾向于购买在压力下能有较好发挥的球员。对于阿利松（罗马），亚历克斯·奥克斯拉德-张伯伦（阿森纳），安德鲁·罗伯逊（赫尔城），法比奥（摩纳哥），纳比·凯塔（莱比锡），维吉尔·范迪克（南安普敦）和谢尔丹·沙奇里（斯托克城）这些球员，图 11 显示了每名球员在各自的俱乐部踢球时 90 分钟比赛内分别在低压力状态和高压状态下的平均贡献。除了凯塔外，其余这些新加盟利物浦的球员在高压状态下的贡献得分都比低压力状态下得分更高。我们的分析中没有包括穆罕默德·萨拉赫，原因是他

6.4 疲劳

疲劳感和精神压力一样会在比赛中不断增加。因此，我们进行了一次分析来检验我们的研究结果并非是因为球员疲劳而导致的状态下滑。图 12 比较了每名球员在精神压力下的表现（统计了球员在 90 分钟高压比赛环境下的平均贡献）和疲劳时的发挥（统计了球员每场比赛最后 15 分钟的平均贡献，排除了高压情形）。根据得出的皮尔森相关系数值仅为 0.136，我们的结论是研究结果和球员在比赛尾声因疲劳而出现的状态下滑不一样。

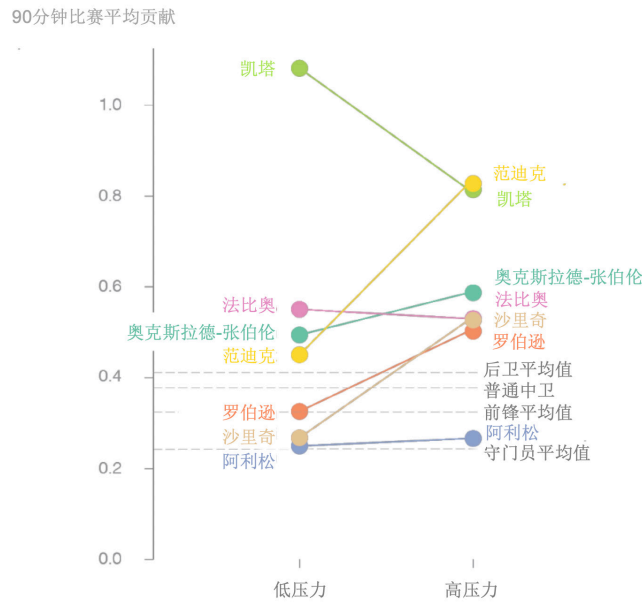


图 11 利物浦新近签约的球员在老东家 90 分钟比赛中低压力和高压力状态下的平均贡献。为了进行比较，我们的分析中纳入了不考虑压力状态下各位置上普通球员的贡献平均值。从 2017/2018 赛季开始，利物浦签下了多名在高压状态下有良好发挥的球员。纳比·凯塔是唯一的例外。

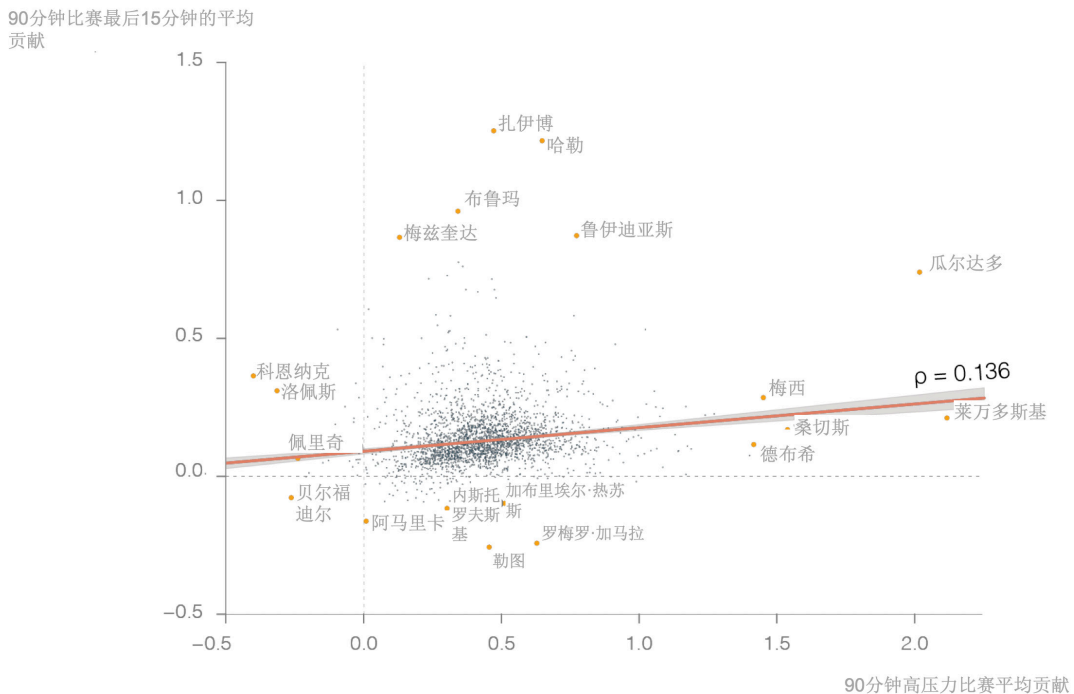


图 12 球员在高压比赛环境和疲劳状态下（比赛最后 15 分钟，不包括高压情形）的发挥比较。球员的发挥以 90 分钟比赛中的平均贡献来统计。皮尔森相关系数的值仅为 0.136，说明我们的压力算法计算的不是疲劳值。

7 相关工作

压力对运动员发挥的影响在其他背景和体育领域也有过研究。其中一种方法是人为定义高压状态，美国职业篮球联盟（NBA）的做法就是这样（例如比赛到了最后五分钟，双方比分差距只有五分），棒球比赛也有相关的做法（例如跑垒员处于得分位置），然后用表现算式（performance metrics）来对高压状态下的球员进行测算。有些篮球和棒球领域的方法会考虑比赛中获胜概率模型来定义高压（或“激烈”）情形。不过，这些方法只考虑了有限数量的动作（例如忽略了传球、射门未果、犯规等）。我们的方法与上述方法有着重要的区别。首先，我们的压力模型考虑了比赛前和比赛中的因素。第二，我们的模型考虑的动作类型更广泛，包括进攻和防守动作。第三，我们的模型对每种动作的三个方面进行评分。此外，认识科学领域就运动员在压力下发挥的机制和调节也有广泛的研究。我们的研究主要以实验为主，通过人为数据控制、抽象和自我聚集来诱导压力情形。

研究者们考虑了用不同的组成元素来构建压力模型。赛前压力方面，我们考虑了NBA季后赛的淘汰赛制，以此来研究输球后果很严重的风险会对球员发挥有什么样的影响。这一研究定义了一场高压比赛，在这场比赛中球队一旦输球将面临淘汰。他们并没有像我们一样考虑很多的特征。比赛中获胜概率模型的研究对象集中

在篮球、美式橄榄球和棒球方面。比较之下，他们对足球的关注不多，也许因为足球比赛得分不多的特点导致对它的分析更有难度。另外，在足球领域和其它体育领域有大量的研究工作关注设计高级算式来评估行为和决策的价值。

8 结论

本文的目的是设法以客观方式了解高精神压力状态对足球运动员表现的影响。由于足球运动员经常面临这种情况，对压力和表现的关系加深了解能在很多方面有所帮助。我们以图文并茂的详实应用案例介绍了我们的研究成果能怎样帮助球员转会决策，教导个别球员，做战术决策和制定比赛先发和替补球员名单。我们的部分研究内容有着直接的实用价值。例如，我们的分析指出几乎所有的足球队都能从调整掷界外球战术中受益。另外，我们给出了一些有趣的观察论述，例如利物浦近来在购买球员时，无论是有意识还是无意识的，都倾向于购买在压力下有良好的发挥的球员。内马尔在压力下的发挥会有所下滑。

附录

A.1 赛前压力评估模型

实际数据的缺乏是这一任务的主要难题。因此，我们利用 19 位专家组成的足球专家组来学习和评价我们的等级分类器。我们从专家组获得 330 对排列数据，而基础数据是随机选择的 170 场 2016/2017，2017/2018 赛季的英超、德甲和西甲联赛比赛。我们使用了这些成对排列数据来拟合梯度提升决策树算法模型（max_depth:6, n_estimators:50）。另外，为让我们的模型更有说服力，我们获取了 20 场比赛的 483 对排列数据，包括重要的升降级相关比赛，老对手之间的比赛，表现不佳的球队比赛和赛季临近结束时无关利害的比赛。一个准确的模型给出的数据应满足以下标准：接近专家们对这些比赛的压力级别的聚合（部分）排序。然而，专家对每一对数据的排序方式可能不一样（即没有一致性的排序），所以准确的模型给出的排序一致性应该专家们共同认可的排序一致性。我们的模型在经过机器学习后，注释器排序的一致性为 73.91%，接近专家共同认可的 79.79% 的排序一致性。

A.2 赛中获胜概率模型设定和评估

如同本文的主要章节所述，我们对主队 ($y>t,home$) 和客队 ($y>t,away$) 在未来进球的数量建的模型为独立二项分布：

$y>t,home | \theta_{t,home} B(T - t, \theta_{t,home})$ 和
 $y>t,away | \theta_{t,away} B(T - t, \theta_{t,away})$,

算式中的参数 θ 表示第 t 时刻预计的主队和客队的进球概率。这些进球概率是根据当前的比赛状态特征： $x_{t, home}$ 和 $x_{t, away}$ 预计的，本文的主要章节已讨论了相关内容。由于这些比赛状态的特征重要性将随时间出现非线性变化，所以我们使用了临时随机过程来对进球概率建模。具体来说，我们使用了以下定义来对进球概率建模：

$\text{logit } \theta_{t, home} = \alpha_t x_{t, home} + \beta + H_a$
 with $\alpha_t \sim N(\alpha_t - 1, \sigma^2)$

$\text{logit } \theta_{t, away} = \alpha_t x_{t, away} + \beta$
 $\beta \sim N(0, 10)$

$H_a \sim N(0, 10)$

H_a 表示主场优势。

我们根据英超、西甲、德甲从 2014/2015 到 2017/2018 赛季（4232 场比赛）的事件数据来预测并验证了对应模型。因此，我们用了 70% 的数据来试验模型，30% 的数据来测试，建立了比赛的临时排序。为了检测我们模型的质量，我们用模型预估了每场比赛的胜、平、负概率，用 $x\%$ 来表示预测数据与真实结果的偏差率。图 A.1 中的概率校正曲线显示了我们的预测概率与实际概率的差异。我们的模型的差异贴近 0% 的线，0% 的线表示最理想（即最佳）的结果。我们的预测偏差通常在 2.5% 这条线以内，最大差异值是 9.61%。

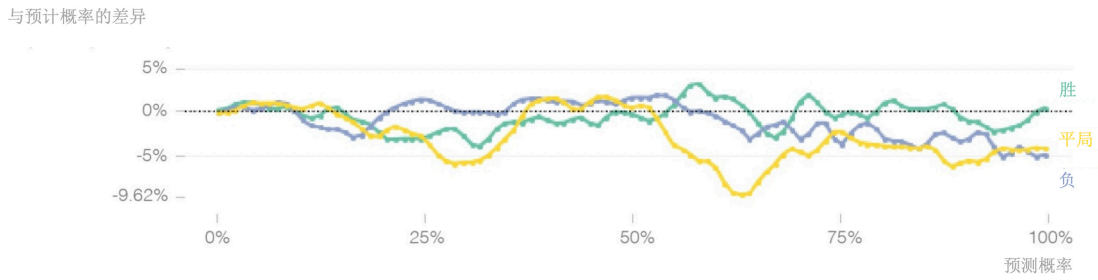


图 A.1 赛中获胜概率模型的概率校正曲线。我们的模型预测的概率与实际结果很接近，说明了我们的模型经过了很好的校正。

A.3 动作成功模型的评估

为了预测动作能否成功的概率，我们试验了一个二元梯度提升决策树算法分类器，试验的数据是 2014/2015 和 2015/2016 赛季的英超、西甲、德甲、意甲、法甲、荷甲和比利时甲级联赛数据。在删除没有相关数据的比赛（例如动作类别）后，我们得到了由 2430 场比赛和 296 万个动作的训练集。我们用 2016/2017 赛季相同联赛的

其它比赛（共 2404 场比赛，292 万个动作）来验证我们的模型。

由于每种动作类别的成功定义不同，我们为每个动作都调试了一个单独的模型。表 A.1 列举了这些模型的精度召回曲线下方面积（AUC-PR），比较了它和相应动作类别的基本成功率。

表 A.1 预测每一类动作是否成功的二元梯度提升决策树算法分类器的精度召回曲线下方面积。基本成功率当作基线。

动作类别	验证集的精度召回曲线下方面积的基本概率（等于基本概率）	我们模型的验证集的精度召回曲线下方面积
传球	0.8265	0.9561
传中	0.3144	0.5397
掷界外球	0.8821	0.9677
角球（传中）	0.4582	0.6550
角球（短角球）	0.8514	0.9829
定位球（长传）	0.5612	0.7429
定位球（短传）	0.8761	0.9783
定位球（射门）	0.0668	0.1011
盘带	0.6991	0.8357
抢断	0.4575	0.7267
射门	0.1094	0.3689
点球	0.7480	0.7480
过人	0.9916	0.9916
球门球	0.5665	0.8310

除了保证准确性很高以外，我们需要校准预测的概率，以便将这些预测的成功概率用到我们的动作决策和执行算式中。因此，我们分析了所有动作类别累计的校正曲线（图 A.2）。由于预测概率与事实成功概率的偏差保持在 3% 以内，我们得出的结论是我们的模型经过了良好的校正。

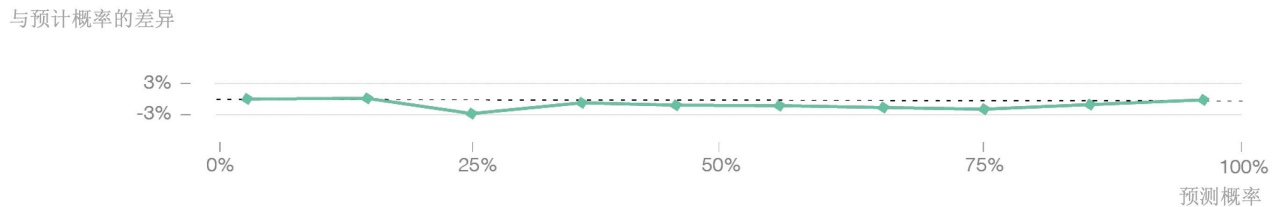


图 A.2 动作成功模型的概率校正曲线。预测的成功概率与实际动作结果很接近，证明我们的模型经过了良好的校正。

A.4 预计下个动作的贡献得分的评估

两个梯度提升决策树模型是我们预计下个动作的贡献得分的基础。假设现在的比赛状态为 s_i ，第一个模型预测的是下一个比赛状态 s_{i+1} 的进球概率 $P(g_{i+1} + | s_i)$ ；第二个模型预测的是下一个比赛状态 s_{i+1} 丢球概率 $P(g_{i+1} - | s_i)$ 。这两个概率的和如下：

$$ECR | s_i = [P(g_{i+1} + | s_i) - P(g_i + | s_{i-1})] - [P(g_{i+1} - | s_i) - P(g_i - | s_{i-1})].$$

我们使用了和附录 A.3 相同的数据集来调试和评估这两个模型。表 A.2 显示了两个模型在验证集上的均方差 (MSE)。作为基线，我们使用了在相同位置开始的所有动作的下一个比赛状态的平均值。

表 A.2 预测下一比赛状态进球或丢球的模型的均方差 (MSE)。在相同位置开始的所有动作的下一个比赛状态的平均值当作基线。

模型	基线的均方差	我们模型的均方差
$P(g_{i+1} + s_i)$	0.0023	0.0019
$P(g_{i+1} - s_i)$	4.9139e-06	4.0413e-06

训练计划混淆了急性和慢性工作负荷比与受伤的关系

卢克·博恩，帕特里克·沃德，达西·诺曼

1 简介

在过去十年中，人们在使用主观（即感知运动评级）和客观（即惯性传感器输出）球员监控策略来了解运动受伤风险上进行了大量工作（Halson, 2014; Bourdon 等, 2017）。特别值得注意的一个指标是急慢性工作负荷比（ACWR），其在时间 t 上的定义为：

$$ACWR_t = L_{7t} / L_{28t}$$

其中 L_{7t} 和 L_{28t} 分别是最近 7 天和 28 天的球员累积工作量（Hulin 等, 2014; Hulin 等, 2015）。在过去的 5 年中，多个体育项目的大量学术论文得出结论：ACWR 是可预测受伤风险的，因此 ACWR 已经成为职业体育管理运动员工作量的专业标准做法（Malone 等, 2017a; Malone 等, 2017b; Colby 等, 2017 年）。

在本论文中，我们使用来自意甲（意大利足球）和 NFL（美式橄榄球）的两支职业球队的赛程和训练数据，表明训练计划混淆了 ACWR 受伤关系，导致高估了其关系的重要性，并对以前忽略了这种混淆的研究提出质疑。

2 数据

我们的数据包括意大利足球和美式橄榄球一个赛季内的每日训练负荷。在这两种情况下，数据都是从惯性传感器收集，并与人体测量结果相结合，以获得每名球员每次训练所经历的总身体负荷的概念。虽然我们专注于单一指标（球员负荷）来简化演示，但请注意我们在论文中观察到的混淆适用于大多数球队收集的指标，例如加速和冲刺数，以及主观测量，如主观疲劳感知评估。由于这些位置的特殊体能需求，我们去除了守门员和四分卫，并将每场比赛的球员负荷进行平均以获得每日全队平均训练负荷。我们将急性负荷计入训练课程的 7 天平均值，并将慢性负荷计算为相应的 28 天平均值。比如说，计算 10 月 29 日的 ACWR 是“10 月 22 日至 28 日的球员负荷总和除以 7”，“10 月 1 日至 28 日的球员负荷总和除以 28”。在非训练日假设负荷为零。在图 1 中，我们显示了整个赛季的原始数据图像，以及急性（红色）和慢性（蓝色）负荷。

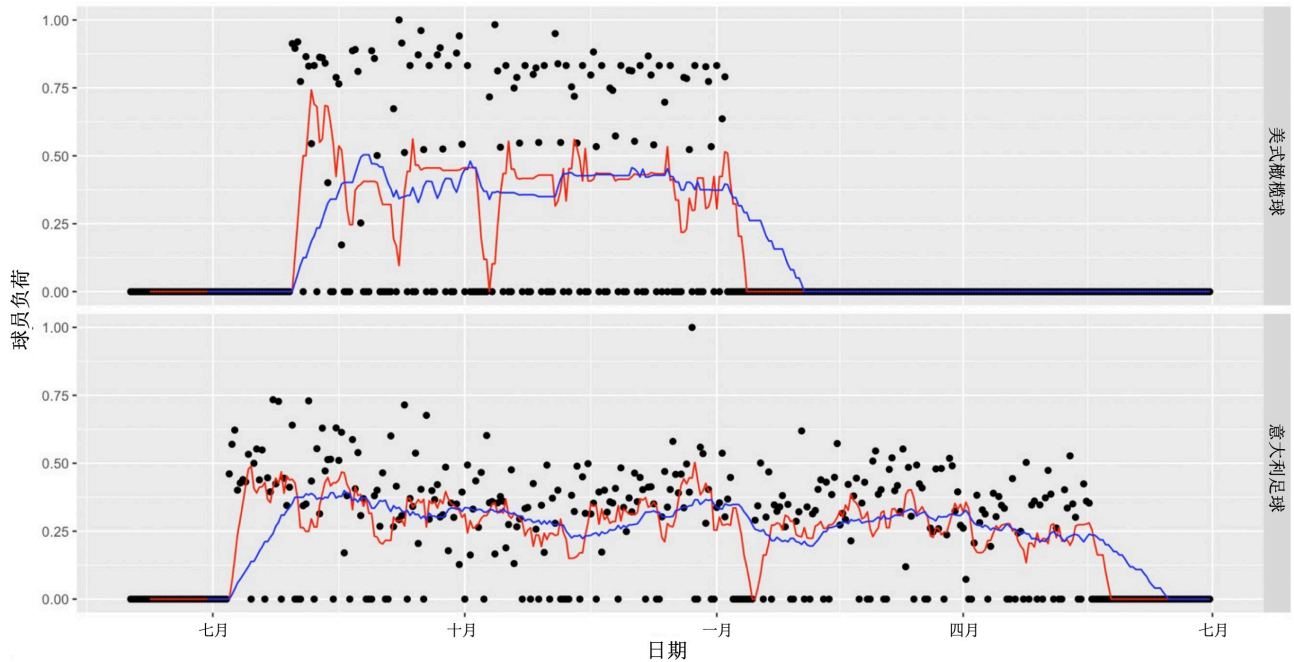


图1 原始数据显示每项运动中的标准化负荷,以及整个赛季的急性(红色)和慢性(蓝色)负荷。

3 一种无效的损伤模型

本文的基本观察是,即使我们模拟了长达一个赛季的训练时间表(和伤病),使得伤病仅取决于当前训练中的球员负荷(因此 ACWR 没有影响),当我们在赛季

结束研究这些模拟值时 ACWR 与受伤显著相关。如图所示,理解 ACWR 受伤关系的标准方法(图2,左)忽视了混淆的训练时间表的影响 Φ (图2,右),因此错误捕获了这种关系。

为了研究这种混淆,我们通过从上述

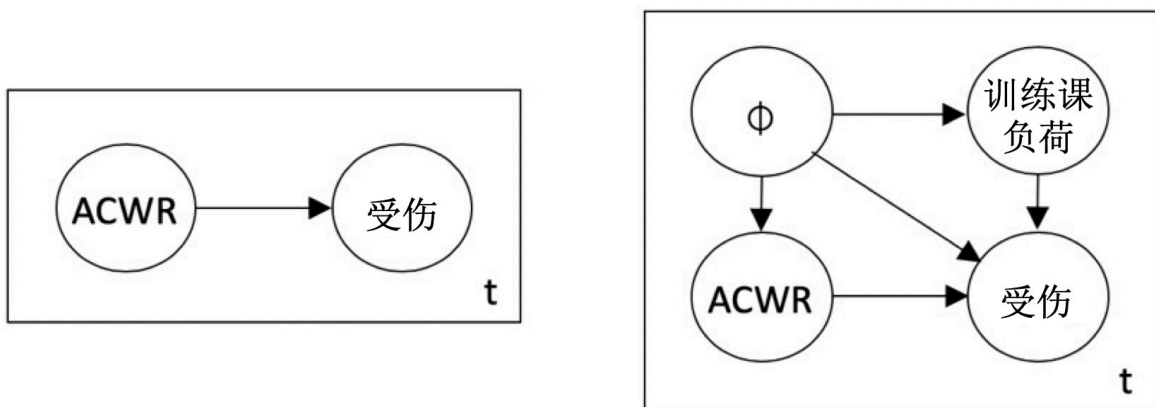


图2 ACWR- 受伤关系图忽略了时间表(左)和时间表影响 Φ (右)的混淆

球队收集的基础训练负荷进行了大量模拟。具体来说，我们使用观察到的球队赛程和训练负荷来模拟 1000 个球员赛季情形，其中给定训练课中受伤的概率与该赛段中的球员负荷成正比。我们对此进行了校准，以便在大约 5% 的训练课程中引发受伤，尽管本论文的结果对于更高或更低的受伤率是可靠的。通过这种方式模拟，我们指定训练课之前的训练要素与球员受伤的可能性无关。

接下来，我们将急慢性标志（“AC 标志”）定义为 ACWR 超出 (0.8,1.3) 范围的任何训练课 (Gabbett, 2016)。在模拟了 1000 个赛季之后，我们发现在受伤和非受伤期间 AC 标志流行率存在统计学显著性差异（足球: $p=2.6 \times 10^{-12}$; 橄榄球: $p=1.4 \times 10^{-7}$ ）。因此，即使我们以 ACWR 对受伤风险没有影响的方式模拟赛季，一项事后研究发现 ACWR 与受伤之间存在有意义的关系！事实上，预计训练课前存在 AC 标志会导致意甲和 NFL 的伤病风险分别增加 6.8% 和 10.5%，即使按结构计算两者都应为 0%。

迄今为止，以类似于上述的方式研究 ACWR 与受伤关系已经进行了大量的工作，大多数人发现了相似的统计学显著性水平。这导致 ACWR 被广泛用作各级运动受伤的诊断工具。虽然我们的工作并没有篡改这些研究，但确实有可能这些研究中发现的显著性可能是出现计划 (schedule) 混淆的结果。

4 修正计划和受伤训练课负荷

本论文的基本论点是，由于体育训练负荷（训练营、国际休息日等）的独特性，ACWR- 受伤关系与计划相混淆。更具体地说，ACWR 是因当前训练课负荷而与受伤相关联的。图 4 说明了美式橄榄球和意大利足球的这种效应，其中我们看到对引发当前训练课的 AC 标志，带有该标志的训练课比不带该标志的训练课具有更高的负荷。

我们现在探讨对当前训练课负荷的调节是否减轻了计划的混淆。具体来说，我们拟定两个逻辑回归模型，其中受伤作为结果，第一个仅使用 AC 标志作为协变量 (covariate)，第二个也包括训练课负荷。正如预期的那样，初始模型产生与上面列出的基线测试相同的 p 值。然而，在对训练课负荷进行调节后，这种关系消失了（足球: $p = 0.33$; 橄榄球: $p = 0.70$ ）。

虽然对当前训练课负荷的调节消除了混淆效应，但这是在模拟环境中（通过设计）才形成的独特的可能性。在实践中，来自受伤球员的数据经常被完全审查或丢失，因为球员的治疗优先于数据的收集。这种审查提出了一种调整混淆的替代方法，即残存模型 (survival models)。当受伤的球员的数据在受伤之前可用时，可以直接应用残存模型。具体来说，一个人将包括直到受伤时刻作为受伤时间的球员负荷，并且每一个非受伤的球员都将在他们观察到的球员负荷下接受审查。也就是说，在观察到受伤之前，健康的球员被视为审查状

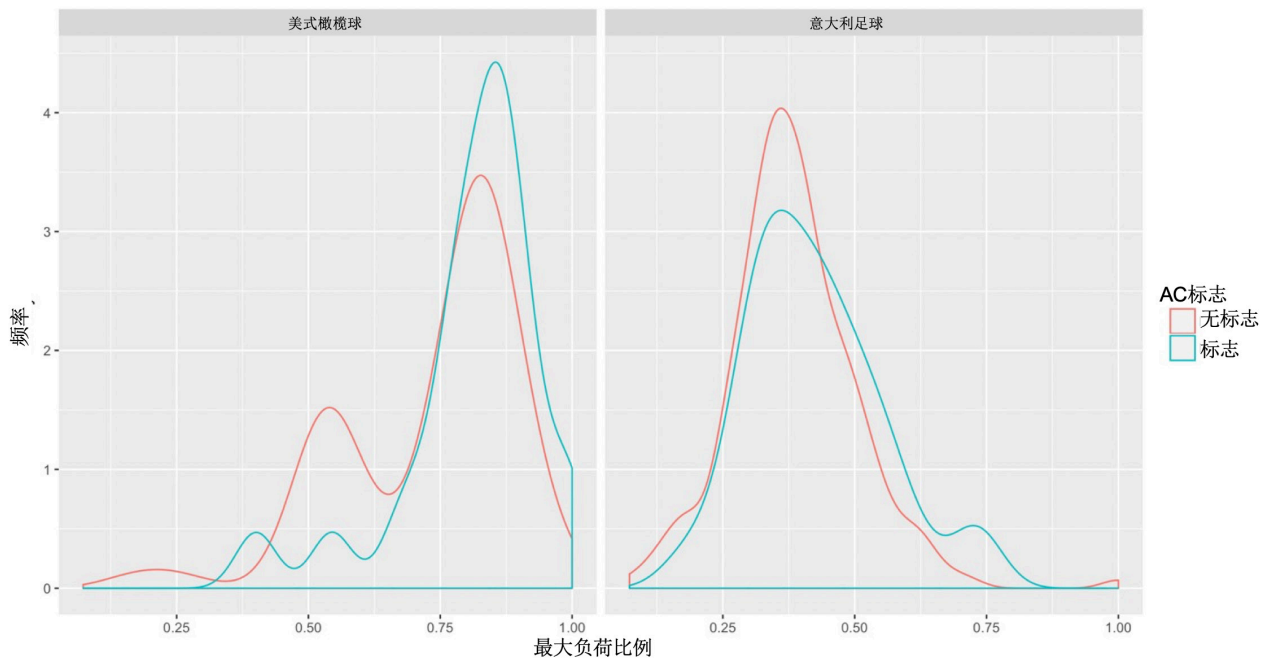


图 4 标记训练课 (flagged sessions) 和非标记训练课之间的负荷分布不同，表明计划混淆了 ACWR- 负荷的关系，因此也混淆了 ACWR- 受伤的关系。

态。事实上，通过随机模拟训练课中的受伤时间将这种技术应用于我们的数据，AC 标志和受伤之间的关系确实变得不重要（足球： $p = 0.22$ ；橄榄球： $p = 0.41$ ）。

当没有受伤球员的数据时，我们可以使用队友的信息来了解球队或位置级别 (position level) 的伤病情况。例如，Ward 等人 (2018) 使用位置平均值来显示在位置级别的美式橄榄球的训练课负荷和受伤之间的关系。在这种情况下，潜在的因素会从个人的受伤风险变为特定位置小组内受伤的可能性。或者，如果聚焦在个人层面，可以使用残存模型和潜在变量模型，输入或推断受伤球员的缺失负荷数据。然而，这些方法还需要对贝叶斯分层模型和因果推断相当了解。有兴趣的读者可以参考 Gelman 等人 (2013 年) 和 Imbens 与 Rubin

(2015 年) 对此类方法的广泛处理。

5 结论

总之，我们证明了关于 ACWR- 受伤关系的因果结论很容易与计划混淆。我们使用蒙特卡罗方法结合来自两项运动的训练负荷数据来说明年度训练日历对 ACWR 受伤关系的影响。然后，我们提出了减轻这种混淆的方案。我们的研究不仅影响围绕 ACWR 的学术论述，而且还使从业者对其预测受伤的价值有了更实际的期望。

精英青少年足球人才选拔中 多维度测评与教练评估的有益协作

罗兰·西哈特斯雷涅，克劳迪娅·祖贝尔，马克·齐邦

摘要：由于青少年足球的广泛开展，该领域从业者面临着一个难题，如何找到早期人才选拔最有效的解决方案。尽管科学界已经提出多维模型有一段时间，但教练评估和运动能力测试的方法仍然很普遍。先前的研究明确了这些不同方法的优缺点。当前的研究直接对各种人才选拔方法的有效性进行了比较（教练评估、运动能力测试与多维数据之间的对比）。由青少年足球运动员、他们的父母和教练组成的 117 名人员参加了 U14 年龄组的多维度测评（教练评估、运动能力测试、心理特征、家庭支持、训练史和生物成熟度）。受试者工作特征 (ROC) 曲线下面积 (AUC [95% CI]) 表明了每种方法在评估预测五年后 U19 球员的球员身份的有效性（职业与非职业）。运动能力测试 (0.71[0.58; 0.84]) 的 AUC 低于多维数据 (0.85[0.76; 0.94], $p = 0.02$)，而教练评估与其他两个没有区别 (0.82[.74; .90])。此外，组合人才选拔的方法 (combined talent selection approaches)，尤其同时使用教练评估和多维数据，在预测 U19 球员身份方面明显更好 (0.93 [0.87; 0.98], $p = 0.02$ 与仅使用多维数据相比)。尽管某些局限可能会阻碍产生更深入的见解（数据汇总、忽略使用非线性统计），但事实证明，在人才选拔中使用多维度的科学主张确实富有成效。特别是主观的教练眼光与科学数据的结合，可以减轻这些不同方法共同弱点的影响。未来的研究应该集中于优化有前景的多维模型的输出。了解与这些模型中特定维度相关的详细参数，并采取增强的非线性统计，可以进一步改进人才选拔工作。

关键词：人才选拔；人才诊断；足球；多维测试；教练评估；预测有效性

引言

在过去几十年间，足球运动大受人们追捧，这说明它已成为世界上最具竞争力的运动之一。与此同时，培养优秀的足球运动员也已成为俱乐部和各国家协会有利可图的工作，并可因此获得较高的声望（Relvas 等，2010 年）。在这一人才培养过程中，人才识别与选拔起着关键作用。人才选拔指的是将确定的人才纳入培养计划（Williams 和 Reilly, 2000）。下文虽只提及人才选拔，但就此概念而言，我们也指识别有潜力成为精英球员的参与者的过程（人才识别）。人才选拔的作用是识别和选择最具前途的青少年球员，以使其在足球组织发展体系中获得优越的学习环境（如专业教练）（Williams 和 Reilly, 2000）。总体而言，似乎很明显，对任何年轻足球运动员进行优化和持续的晋级，将是利用整个人口潜力最有前景的足球人才培养模式（Côté 和 Hancock, 2015）。然而，足球组织内部的资源仍然有限。因此，必须将人才选拔和淘汰（意味着失去潜力）作为一个不可避免但必要的过程加以考虑，以便把资源集中在最具潜力成为未来精英球员的身上（Suppiah 等，2015）。

由于人才选拔的必要性，这一领域的从业者持续受到一个问题的困扰，即人才选拔过程中最有效的方法是什么。尽管多维人才选拔方法已经提出一段时间（Abbott 等，2005；Vaeyens 等，2008；Williams 和 Reilly, 2000），但大多数俱乐

部和协会仍然仅依赖于教练评估的主观数据（Christensen, 2009；Larkin 和 Reeves, 2018）。除了教练的眼光，在几个培养项目的人才选拔中，常见的只有几项特定的客观数据（例如，从运动能力测试中获得）（Höner 等，2017）。因此，科学界的建议和该领域目前实行的程序之间仍存在差距（Larkin 和 Reeves, 2018）。除了经常讨论的与教练培训需求相关的问题（Figueiredo 等，2014）之外，造成这一差距的一个原因，可能是缺乏科学证据证明多维度的人才选拔方法优于常用的教练评估或运动能力测试。迄今为止，这些不同的人才选拔方法（教练评估、运动能力测试与多维数据）之间尚没有直接的对比。因此，这三种方法在预测年轻足球运动员未来成功的潜力方面可能存在的差异尚不明确（Schorer 等，2017）。

关于教练评估在足球人才选拔中的效用，科学意见也不统一。一方面，从整体而言，教练可以从几个维度整合信息，并从整体上评判球员（Buckers 等，2015）。Jokuschies 等（2017 年）通过将瑞士足球协会五名青年国家队教练的人才标准系统化，认可教练眼光这一积极观点。他们能够证明教练依据一定的人才标准给出的评分是可靠的，对球员总体潜力的评估也有效。此外，教练对球员总体表现和总体潜力的评分显示，评分者之间有较高的可靠性（Fenner 等，2016；Güllich 等，2017；Zuber 和 Conzelmann, 2014）。然而，教练在人才选拔中的决定似乎由主观感受所

引导 (Johansson 和 Fahlén, 2017; Lund 和 derströ, 2017) , 而该领域的从业人员也没有形成普遍接受的人才模型 (Jokuschies 等, 2017 年) 。此外, 比赛中的整体表现评分还受到球员比赛时的动作数量影响 (Tromp 等, 2013) 。尽管经验丰富的教练可能会对生物成熟度作出有效评估 (Romann 等, 2017 年) , 但生物成熟度仍会影响其对球员比赛表现的主观评分 (Cripps 等, 2016 年) 。

有几个案例通过测量生理数据和总体运动能力 (Dodd 和 Newans, 2018; Gonaus 和 Müller, 2012; Le Gall 等, 2010; Murr 等, 2018) 以及球员技能 (Forsman 等, 2016; Höner 和 Votteler, 2016; Sarmiento 等, 2018) , 证明运动能力测试对足球人才选拔的价值。然而, 由于受到发育相关因素的影响, 如生物成熟度和相对年龄, 生理数据 (如有氧能力) 和总体运动能力测试 (如冲刺成绩) 对年轻球员长期天赋的预测有效性受到了强烈质疑 (Johnson 等, 2017; Malina 等, 2017 年; 米勒等人, 2017; Romann 等, 2018) 。鉴于此, 有研究认为, 尽管针对特定领域的测试项目 (如球员技能) 其可靠性通常较低 (Lidor 等, 2009) , 但却能达到比总体运动能力测试更高的预测有效性。因此, 运动能力测试对足球人才选拔的总体价值仍在讨论当中 (Leyhr 等人, 2018) 。

除了教练评估和运动能力测试之外, 为人才选拔多维模型提出的常见科学建议, 还将心理特征、家庭支持和训练史当作预测未来成功的潜在因子 (Figueiredo 等,

2009; Huijgen 等, 2014 年; Williams 和 Reilly, 2000) 。尤其是心理特征, 越来越受到该领域的关注。例如, 动机、意志和自我调节的技巧尤其重要 (Gledhill 等, 2017; Zuber 等, 2015) 。值得注意的是, 教练对精英青少年足球运动员天赋的看法, 主要受心理特征的影响 (Jokuschies 等, 2017) 。然而, 综合影响 (如对随时间推移产生的个性变化的了解有限, 心理因素操作上的困难, 研究中使用的设计种类多样) 抑制了对青少年足球人才选拔中心理特征价值的清晰看法的形成 (Gledhill 等, 2017; Sarmiento 等, 2018) 。家庭支持的影响可以通过情感、资金或组织方式来表达, 这种影响一般都是在人才培养的背景下进行讨论 (Côté, 1999; Knight 等, 2017) , 而其对人才选拔的预测能力几乎尚未有人调查 (Zibung, 2014) 。因此, 仍有必要进一步理解家庭支持对人才选拔的可能影响 (Sarmiento 等人, 2018 年) 。最后, 对于训练史的预测价值, 尤其是 12 岁以下球员的训练史, 人们展开了激烈的辩论。尽管有证据表明, 在早期使用包含各种足球活动的专业抽样模式 (Ford 和 Williams, 2017) , 似乎对后来的成功有一定的成效, 但数据仍然存在矛盾 (Hornig 等, 2016) 。

总体而言, 常用和推荐的人才选拔方法各有利弊。教练评估具有内在的主观性, 在考虑心理测量特性时, 这一直是为人们争论的焦点 (Johansson 和 Fahlén, 2017; Jokuschies 等, 2017; Lund 和 Söderström,

2017年)。然而，教练评估的全面性反映了教练可以从几个不同方面整合信息以及更多地从整体上评判球员的潜力。这展示出明显优于其他评估方法的优势，因此在选拔或取消选拔一名球员时依据整体评估做决策更容易 (Buekers 等, 2015)。相比之下，对于运动能力测试和多维度测评，就客观性和可靠性而言，心理测量的特性通常被科学界和该足球领域从业者所接受 (Höner 等, 2017)。然而，只有有限的证据能够支持每个维度在预测球员未来成功方面的有效性 (Sarmiento 等, 2018)。此外，运动能力测试和多维测试组合所提供的结果，包含来自不同项目和维度的数个变量，而将这些变量整合到整体评估中并确定具体变量负载这一问题至关重要 (Bergman 和 Trost, 2006; Till 等, 2016; Till 等, 2018)。考虑到对单个球员作出全面决策的重要性，这一点尤其是个问题。

鉴于目前不同人才选拔工具的使用具有不确定性，人们越来越有兴趣并要求直接比较它们在预测有效性方面的相对值 (Buekers 等, 2015; Schorer 等, 2017)。如果所有方法都有各自的优势和劣势，那么，哪种方法对预测后来的成功最有用：教练评估、运动能力测试还是多维数据？Schorer 等在女子手球队运动员的样本中部分考虑了此问题。他们发现，运动能力测试的逻辑回归模型在十年内所预测的正确选出女子手球人才的比例 (85.2%) 高于国家队教练评估 (79.3%)。然而，由于研究的探索性，他们避免使用推论统计，并且

没有测试人才选拔工具之间在统计学上的显著差异。

因此，当前研究通过直接比较教练评估、运动能力测试和多维数据这些工具的预测有效性，以此检验这些工具是否能在精英青少年足球人才选拔方面表现出更高的成功率。第二个目的是解释清楚以下假设后面的支撑证据，即组合评估方法可能带来更好的预测，因为使用组合在足球领域十分常见 (教练评估和运动性能测试；Höner 等, 2017)，或者能够提供关于每个球员最全面的视角 (教练评估和多维数据)。后者对于预测体育人才未来的成功似乎特别有效 (Reilly, 2006; Zuber 等, 2016)。

方法

研究设计和参与者

当前研究是《瑞士足球人才选拔和培养》这一纵向研究项目的一部分。它整合了数个维度，全面评估人才培养 (例如运动能力、教练评估、心理特征、家庭支持和训练史)。该项目跟踪调查了瑞士足球协会人才选拔系统中大量1999年出生的球员。瑞士足球协会的人才选拔系统遵循了人才培养的金字塔标准模式，这在其他联邦人才培养计划中也很常见 (Bailey 和 Collins, 2013)。将球员早早选入人才基地和地区球队是从U12年龄组开始的 (大约占注册球员的6%；Romann 和 Fuchslocher, 2013)。精英青少年培养计划 (瑞士青年国

家队，约占注册球员的 1%) 从 U15 年龄组开始。自 2012 年以来，瑞士一直跻身于国际足联成年国家队排名前 12 名的国家中，不得不将瑞士足球协会的人才选拔系统视为一套高效的系统。

当前研究使用了 117 名运动员的样本。在 2017/2018 赛季，这些球员中有 20 名 (17.1%) 参加了瑞士的甲级到丙级联赛，或者被提名到瑞士 U19 青年国家队 (职业球员)。其余 97 名球员参加丁级或更低级别的联赛，并被归类为非职业球员。五年前 (2012/2013 赛季，U14 年龄组)，所有

球员、家长和俱乐部教练都自愿参加教练评估和多维度测评 (见表 1)。

测评

教练评估。球员所在的俱乐部教练执行视觉量表评估程序，以评估球员当前的比赛表现。对于教练的评分，他们使用了 0 到 100 之间的视觉量表。当肯德尔和谐系数 $W = 0.89$ 时，该教练评估工具的评分者信度可视为令人满意 (Zuber 和 Conzelmann, 2014)。由于球员是不同地区球队的一部分，

表1 测量变量和项目一览表

维度	变量	参考	可靠性
年龄, 生物成熟度	实际年龄 (年)	Mirwald等, 2002	$r_{tt} = .96$
	相对年龄 (月)		
	巅峰速度时的年龄 (年)		
人体测量学	身高 (厘米)		$r_{tt} = .99$
	体重 (公斤)		$r_{tt} = .99$
教练评估	比赛表现 (分数)	Zuber和Conzelmann, 2014	$W = .89$
总体运动能力	YoYo测试 (米)	Bangsbo等, 2008	$r_{tt} = .93$
	下蹲跳 (厘米)	Casartelli等, 2010	$ICC = .96$
	40米冲刺 (秒)	Zuber等, 2016	$r_{tt} = .96$
	敏捷性测试 (秒)	Höner等, 2015	$r_{tt} = .83$
技术	运球 (秒)	Höner等, 2015	$r_{tt} = .56$
	传球 (秒)	Zuber等, 2016	$r_{tt} = .68$
	颠球 (分数)	Höner等, 2015; Zuber等, 2016	$r_{tt} = .79$
心理特征	成就动机 (净希望)	Wenhold等, 2009	$\alpha = .76/.73$
	成就目标定向 (得分)	Elbe, 2004	$\alpha = .80/.72/.81$
	自我决定 (指标)	Pelletier等, 1995; Demetriou, 2012	$\alpha = .86$
家庭支持	足球在家庭中的重要程度 (得分)		$r_{tt} = .63$
	父母和学校对运动的优先考虑 (得分)		$r_{tt} = .52$
	经济投入 (瑞士法郎/年)		$r_{tt} = .78$
	时间投入 (小时/周)		$r_{tt} = .59$
训练史	12岁之前的训练和比赛 (小时)	Hopwood, 2015	$0.59 \leq r_{tt} \leq 0.97$

因此共有 15 名俱乐部教练参与了对球员比赛表现的评分。为了应对这一状况，设计了一个标准化程序将教练引入测试工具（例如假想的青年国家队队员得分应该在 90 到 100 之间，而非常差的队员得分应该在 0 到 10 之间；Zuber 和 Conzelmann, 2014）。

出于几种原因，我们限制了对教练评估的进一步标准化（例如，单独一个教练，他应该给所有的球员评分）。首先，所提到的令人满意的评分者信度表明研究具有适当的客观性。第二，应用运动科学应该考虑生态效度（ecological validity）。就足球联盟和协会内全国人才培养系统的人才选拔过程而言，这可能意味着有必要将不同地区的不同教练纳入进来。

运动能力测试。在整个赛季中，运动员参加了两次（2012 年秋季和 2013 年春季）由七个项目组成的组合测试，以确定运动能力。赛季表现由两组测试的平均值计算而来。由于几种原因（例如受伤、生病和学校活动），一些运动员错过了某个测量点（29.9%），在这种情况下，则以其他测量作为测试分数。由于在为期六个月的赛季内，在相似的年龄组中没有发现运动能力发生重大变化（Francioni 等人，2016 年），并且在足球的长期发展分析中，常用到相似的程序（Gonaus 和 Müller, 2012; Höner 等, 2015），因此，可以说这一程序是恰当的。

1 级 Yo-Yo 间歇恢复测试（YY）测量的是间歇耐力的表现（ $rtt = 0.93$; Bangsbo 等, 2008）。在垂直下蹲跳测试（CMJ，手臂不摆动）中，五次尝试的最高值通过加速测

量系统得出（Myotest, 瑞士锡永; $ICC = 0.96$; Casartelli 等, 2010）。40 米冲刺测试（40 米）是在起跑线和终点线各用一个双光电传感器（Microgate, 意大利博尔扎诺）进行的（ $rtt = 0.96$; Zuber 等, 2016）。至于敏捷性测试（AG），运动员进行短距离冲刺，改变方向绕着三个杆跑，并在测试完成前反向重复这些动作（Höner 等, 2015）。至于冲刺测试，则使用双光电传感器测量时间（ $rtt = 0.83$ ）。运球测试（DR）与敏捷性测试的方式相同，唯一的区别在于它是在有球而不是无球的情况下进行的（ $rtt = 0.56$; Höner 等, 2015）。传球测试（PA）是对霍纳等（2015）所用测试的改编。球员在限定区域内尽可能快地传球，然后依次将球从四面墙上弹开，每个方向一次。在第四次传球之后，以相反的顺序重复此前同样的传球（总共传球九次）。用秒表手动测量时间（ $rtt = 0.68$; Zuber 等, 2016）。在颠球测试（JU）中，球员轮流沿着一个如 8 字形状的路线交替用左脚和右脚颠球。他们每完成四分之一一个圆圈，就得 1 分。一旦出现错误（例如，一只脚连续颠球两次，或者球触地），或者最迟在 45 秒后，停止测试。测试成绩等于得分总数（Höner 等, 2015），达到 $rtt = 0.79$ （Zuber 等, 2016）。

整个测试过程遵循一个标准计划（热身、测试顺序、训练有素的测试团队），并且只在干燥的人工草皮上进行。如遇下雨天气，另择期进行。对于 40 米冲刺、AG、DR、PA 和 JU，取两次尝试中最好的一次成绩用于数据分析。对于尽全力的 YY

测试，只有一次尝试机会。

心理学问卷。与早期在《瑞士足球人才选拔和培养》项目（Zuber 等，2015）中对心理评估的研究一致，通过问卷调查确定了以下动机特征。

成就动机是通过简化的《成就动机量表—体育》德国版，用净希望（net hope）（对成功的希望减去对失败的恐惧）来测量的（Wenhold 等，2009）。每个量表由五项组成，采用四分制反应量表（从 0 = 完全不适用于我到 3 = 完全适用于我）。内部一致性结果达到可以接受的水平，其中 α 对成功的希望 = 0.76， α 对失败的恐惧 = 0.73。

成就目标定向使用德国版的《运动定向问卷》进行测量（赛事定向、获胜定向和目标定向的平均得分；Elbe, 2004）。每个量表由六项组成，采用五分制反应量表（从 1 = 强烈不同意到 5 = 强烈同意）。内部一致性达到可接受水平（ α 竞争 = 0.80， α 获胜 = 0.72， α 目标 = 0.81）。

自我决定是用德国版的《运动动机量表》测量的（Demetriou, 2012; Pelletier 等，1995）。这一量表包括七个子量表：内在动机（三个子量表：了解、完成、体验）、外在的、内投的和确定的规则以及动机。每个量表由四项组成，采用的是七分制的反应量表（从 1 = 完全不对应到 7 = 完全对应）。这七个分量表合起来形成一个自我决定指数（Vallerand, 2001）。当 $\alpha = 0.86$ 时，内部一致性被视为可以接受。

环境问卷。与早期在《瑞士足球人才选拔和培养》项目（Sieghartsleitner 等人，

2018 年及 Zibung 和 Conzelmann 2014）中对人才培养环境的研究一致，通过问卷调查确定了以下几个方面。

对家庭支持的测量试图涵盖情感、资金和组织几个方面。这些方面用足球在家庭中的重要程度（从 1 = 低重要性到 5 = 高重要性）、父母和学校对运动的优先考虑（从 1 = 上学更重要到 5 = 足球更重要）、经济投入（父母整个赛季花费的绝对金额，如为设备或训练的花费）和时间投入（父母平均一周花费的绝对时间量，如往返交通或进行鼓励所花费的时间）四个维度进行测量。就这些项目的心理测量特性而言，在连续两个赛季的测量之间得到了 $0.52 \leq r \leq 0.78$ 的差异稳定性记录。

训练史是通过俱乐部内部组织的足球训练量和 12 岁以下的全部足球自由活动量来测量的。运动量是通过追溯性调查问卷收集起来的，方法是将自参加体育运动以来的每一个年龄组的周运动小时数相加。这些数值之和就是 12 岁以下的运动总小时数（Sieghartsleitner 等，2018）。实践证明这种方法的可靠性可以接受（Helsen 等，1998; Hopwood, 2015）。

生物成熟度。为控制生物成熟度，以出生月份（1 月 = 1，12 月 = 12）为单位，对以巅峰速度时的年龄（Mirwald 等，2002）和相对年龄（RA）表示的成熟时间（Cumming 等，2017）进行了评估。前一种方法的心理测量特性已被证明是可接受的（Müller 等，2015）。

数据分析

通过五年前在 U14 年龄组的运动表现收集的数据，数据分析对五个分类模型进行计算，来预测 U19 球员身份（职业或非职业）。表 2 对这些分类模型和综合变量进行了概述。第一个模型仅使用教练评估的比赛表现（教练眼光模型，1 个变量）。第二个模型由来自运动能力测试的数据（运动能力模型，2 个变量）计算而来，并使用总体运动能力项目（YY 测试，CMJ，40 米冲刺，AG）的平均 z 值和技能项目（DR，PA，JU）的平均 z 值。第三个模型（多维模型，6 个变量）使用总体运动能力项目的平均 z 值、技能项目的平均 z 值、成熟时间（APHV）和相对年龄（RA，出生月）的平均 z 值、所有心理特征的平均 z 值、所有家庭支持项目的平均 z 值以及训练史的 z 值。第四个模型结合了来自教练评估的比赛表现和运动能力测试的数据（教练眼光和运动能力模型，3 个变量），因为这种组合在该领域较为常见。最后，第五个模型使用了所有可用的信息，结合了教练评估中的比赛表现和多维模型中的 6 个 z 值（整体模型，7 个变量），这反映了对每个球员的整体看法（Reilly, 2006; Zuber 等, 2016）。

每个模型都使用来自 R（R 核心小组，2017 年）的二元逻辑回归（BLR）作为鲁棒分类器（Antonogeorgos 等，2009），并使用来自 R 包 pROC（Robin 等，2011）的受试者工作特征（ROC）来确定分类的辨别能力。在这一过程中，BLR 首先计算

了每个人被归类为职业或非职业球员的可能性。用似然比检验（模型系数的综合检验；Zeileis 和 Hothorn, 2002）和 Hosmer-Lemeshow 检验对系数的显著性和 BLR 模型的适当标定进行测试（Hosmer 等，2013），其中系数的显著性表明比基线模型有所改进，适当标定则可以证明模型是否与数据相符。两项测试的显著性 α 水平均设置为 $p < 0.05$ 。根据相应的零假设，系数的显著性用 $p < 0.05$ 表示，适当标定的显著性用 $p > 0.05$ 表示。在表示出适当的标定后，来自 BLR 的似然值被用于创建 ROC。通过 DeLong 非参数检验，使用曲线下面积（AUC）——一个衡量分类质量的指标——及其标准误差，对 ROC 模型之间的相互差异进行比较（DeLong 等，1988; Robin 等，2011）。同样，显著性 α 水平最初也设置为 $p < 0.05$ 。由于要对五种分类模型进行多次比较，因此用错误发现率来适当地调整 p 值（Benjamini 和 Hochberg, 1995）。

除了能够确定分类质量和各模型的即时统计可比性之外，ROC 还可以通过描述敏感性、特异性和准确性的变化以及甄别阈值的变化，改进 BLR 的统计数据（Robin 等，2011）。这些值通常被称为正确选拔的人才比例（敏感性）、正确撤销选拔的非人才比例（特异性）和所有选拔决策的正确百分比（准确性）。与 BLR 设定的固定甄别阈值相比，ROC 创造了能够确保最有力的甄别阈值之一的可能性，即约登指数（Youden, 1950）。该指数使敏感性和特异性的总和最大化，并提供了超过 BLR 的

额外好处。在人才培养系统同等重视正确选拔的人才和正确撤销的非人才的情况下，YI（约登指数）可以代表系统最有效的人才选拔阈值。

结果

表 3 对职业和非职业运动员被测变量和项目的描述性特征进行了概述。根据

BLR 分析的结果（见表 4），经过适当标定（ $0.50 < p < 0.95$ ），全部五种模型都是具有显著性（ $p < 0.01$ ），并且显示从 0.14 至 0.55（Nagelkerkes R 方）模型与数据都拟合。表 5 给出了 ROC 的描述值。AUC[95% CI] 指示值区间为 0.71[0.58; 0.84] 至 0.93[0.87; 0.98]。分类模型的敏感性指示值在 70% 和 95% 之间，这意味着这些模型能够正确识别 20 名职业球员中的 14 至 19 名。特异性

表2 五种计算出的分类模型和所包含的变量概览

分类模型	变量数	变量
1 教练眼光模型	1	来自教练评估的比赛表现的z值
2 运动能力模型	2	总体运动能力项目的平均z值 技能项目的平均z值
3 多维模型	6	总体运动能力项目的平均z值 技能项目的平均z值 成熟时间和相对年龄的平均z值 全部三种心理特征的平均z值 所有家庭支持项目的平均z值 训练史的z值
4 教练眼光和运动能力模型	3	模型1和模型2组合
5 整体模型	7	模型1和模型3组合

表3 职业和非职业运动员被测变量和项目的平均值 (±标准差)

维度和变量	职业 运动员 (n = 20)		非职业运动员 (n = 97)		合计 (n = 117)	
年龄, 生物成熟度						
实际年龄 (年)	13.53	(0.38)	13.54	(0.30)	13.54	(0.31)
相对年龄 (月)	4.80	(3.89)	5.07	(3.39)	5.03	(3.47)
巅峰速度时的年龄 (年)	13.81	(0.54)	14.05	(0.60)	14.01	(0.95)
人体测量学						
身高 (厘米)	161.1	(7.1)	160.0	(8.4)	160.2	(8.1)
体重 (公斤)	48.3	(6.0)	47.5	(8.0)	47.6	(7.7)
教练评估						
比赛表现 (分数)	79	(12)	60	(16)	63	(17)
总体运动能力						
YoYo测试 (米)	1032	(406)	1101	(352)	1089	(361)
下蹲跳 (厘米)	29.0	(2.8)	31.0	(3.8)	30.6	(3.7)
40米冲刺 (秒)	6.41	(0.34)	6.44	(0.33)	6.43	(0.33)
敏捷性 (秒)	8.12	(0.25)	8.10	(0.29)	8.11	(0.28)
技术						
运球 (秒)	10.12	(0.39)	10.35	(0.59)	10.31	(0.57)
传球 (秒)	16.2	(1.6)	16.8	(1.6)	16.7	(1.6)
颠球 (分数)	9.5	(6.1)	6.2	(5.2)	6.7	(5.5)
心理特征						
成就动机 (净希望)	1.91	(0.88)	1.79	(0.88)	1.81	(0.87)
成就目标定向 (得分)	4.70	(0.23)	4.49	(0.39)	4.53	(0.38)
自我决定 (指标)	10.25	(2.36)	8.91	(2.85)	9.14	(2.81)
家庭支持						
足球在家庭中的重要程度 (得分)	4.70	(0.47)	4.34	(0.68)	4.40	(0.66)
父母和学校对运动的优先考虑 (得分)	3.45	(1.15)	2.67	(0.97)	2.80	(1.04)
经济投入 (瑞士法郎/年)	2625	(2218)	1723	(1940)	1877	(2009)
时间投入 (小时/周)	12.3	(7.0)	9.1	(8.7)	9.7	(8.4)
训练史						
12岁之前的训练和比赛 (小时)	3311	(1039)	3160	(1017)	3187	(1017)

(Specificity) 值从 66% 到 88% 不等，由此分类模型正确识别了 97 名非专业运动员中的 64 至 85 名。

敏感性和特异性共同导致 YI 为 0.61 或教练眼光模型中总体的正确人才选拔决策达到 71%。换句话说，在 U14 年龄组的 117 名精英青少年足球运动员的样本中，该模型将断定 52 名运动员会成为职业球员（19 个有效预测，1 名职业球员被忽略）。在运动能力模型中，最终的准确率是 73% 的正确预测 (YI = 0.43)，这意味着对 40 名球员的断定更为严格，结果只有 14 个有效预测（6 名职业球员被忽略）。多维模型显示的准确率为 82% (YI = 0.67)，因此对 34 名球员的判定中有 17 个有效预测（3 名职业球员被忽略）。教练眼光和运动能

表4 五种二元逻辑回归分类模型的显著性、标定和模型拟合值

维度和变量	模型系数的综合检验			Hosmer-Lemeshow检验			模型拟合
	χ^2	df	p	χ^2	df	p	Nagelkerkes R ²
教练眼光模型	23.63	1	<.01	3.42	7	0.84	0.31
运动能力模型	10.19	2	<.01	2.81	8	0.95	0.14
多维模型	25.47	6	<.01	5.09	8	0.75	0.33
教练眼光和运动能力组合模型	37.02	3	<.01	7.34	8	0.50	0.45
整体模型	47.46	7	<.01	5.00	8	0.76	0.55

表5 五种分类模型的受试者工作特征曲线的描述值

分类模型	AUC* [95% CI]	敏感性[95% CI]	特异性 [95% CI]	准确性[95% CI]	YI†
教练眼光模型	0.82 [0.74; 0.90]	0.95 [0.75; 1.00]	0.66 [0.47; 0.81]	0.71 [0.56; 0.82]	0.61
运动能力模型	0.71 [0.58; 0.84]	0.70 [0.35; 0.95]	0.73 [0.40; 0.96]	0.73 [0.53; 0.89]	0.43
多维模型	0.85 [0.76; 0.94]	0.85 [0.60; 1.00]	0.82 [0.57; 0.98]	0.82 [0.63; 0.93]	0.67
教练眼光和运动能力组合模型	0.88 [0.81; 0.95]	0.95 [0.70; 1.00]	0.71 [0.56; 0.96]	0.75 [0.63; 0.93]	0.66
整体模型	0.93 [0.87; 0.98]	0.90 [0.75; 1.00]	0.87 [0.70; 0.99]	0.88 [0.72; 0.96]	0.77

*曲线下面积，†约登指数

表6 分类模型的AUC与DeLong非参数检验 (DeLong等, 1988) 的比较

用于AUC比较的模型‡		AUC [95% CI]	AUC [95% CI]	Z	p
教练眼光模型	与 运动能力模型	0.82 [0.74; 0.90]	0.71 [0.58; 0.84]	1.50	0.13
教练眼光模型	与 多维模型	0.82 [0.74; 0.90]	0.85 [0.76; 0.94]	-0.60	0.55
运动能力模型	与 多维模型	0.71 [0.58; 0.84]	0.85 [0.76; 0.94]	-2.40	0.02 [§]
教练眼光模型	与 教练眼光和运动能力组合模型	0.82 [0.74; 0.90]	0.88 [0.81; 0.95]	-1.86	0.06
运动能力模型	与 教练眼光和运动能力组合模型	0.71 [0.58; 0.84]	0.88 [0.81; 0.95]	-3.23	< .01 [§]
教练眼光模型	与 整体模型	0.82 [0.74; 0.90]	0.93 [0.87; 0.98]	-2.97	< .01 [§]
多维模型	与 整体模型	0.85 [0.76; 0.94]	0.93 [0.87; 0.98]	-2.35	0.02 [§]
教练眼光和运动能力组合模型	与 整体模型	0.88 [0.81; 0.95]	0.93 [0.87; 0.98]	-1.69	0.09

‡ 由于缺乏实践或理论相关性, 出于经济原因, 两种比较 (运动能力模型与整体模型比较、多维模型与整体模型比较) 被省略。§ 分类模型间具有显著差异 ($p < 0.05$; 错误发现率调整后的p阈值: 0.031; Benjamini和Hochberg, 1995)。

表7 二元逻辑回归整体模型系数

变量#	β	SE	Wald	df	p	机会比率 [95% CI]
比赛表现	2.53	0.73	11.86	1	< 0.01	12.53 [2.97; 52.83]
总体运动能力	-2.20	0.71	9.66	1	< 0.01	0.11 [0.03; 0.44]
家庭支持	1.65	0.58	8.21	1	< 0.01	5.20 [1.68; 16.07]
技术	0.35	0.56	0.38	1	0.54	1.41 [0.47; 4.24]
心理特征	-0.15	0.58	0.06	1	0.80	0.86 [0.28; 2.71]
训练史	0.08	0.35	0.05	1	0.83	1.08 [0.55; 2.12]
生物成熟度	-0.05	0.57	0.01	1	0.93	0.95 [0.31; 2.92]
常数	-3.31	0.71	21.63	1		0.04

变量按照 β 系数绝对值排序。

力组合模型正确预测了 75% 的球员 (YI = 0.69)，这意味着对 47 名球员的预测限制较小，预测到 19 名职业球员 (1 名职业球员被忽略)。最后，整体模型的准确率为 88% (YI = 0.77)，做出了最有效的人才选拔决策。就对职业球员的判定而言，该模型将只选拔 30 名球员，其中有 18 个有效预测 (2 名职业球员被忽略)。

表 6 显示了比较 AUC 的非参数方法的结果。前三行是三种人才选拔方法之间的成对比较。接下来，将教练评估和运动能力测试组合与单独使用教练眼光或运动能力测试的模型进行比较。最后三项比较的目标是整体模型的价值 (它是通过与教练眼光进行对比来检验的)、多维数据以及教练眼光和运动能力测试的组合模型。

为了确定分类模型中单个维度的值，表 7 显示了整体模型的 BLR 回归系数，该模型包含全部 7 个用到的变量。到目前为止，比赛表现具有最高的正面影响，机会比率 (OR) 为 12.53 ($p < 0.01$)，而总体运动能力与职业运动员身份呈负相关 (机会比率 =

0.11, $p < 0.01$)。后来发现家庭支持对回归模型有进一步的显著影响 (OR = 5.20, $p < 0.01$)。技能 ($p = 0.54$)、心理特征 ($p = 0.80$)、训练史 ($p = 0.83$) 和成熟度 ($p = 0.93$) 对回归模型没有显著影响。由于比赛表现在这一模型中占主导地位，表 8 中给出了多维模型的 BLR 回归系数，可以进一步观察不同被测变量的值。家庭支持 (OR = 4.24, $p < 0.01$) 和技能 (OR = 3.00, $p = 0.02$) 对获得职业运动员有显著的积极影响，而总体运动能力再次呈负相关 (OR = 0.37, $p = 0.04$)。心理特征 ($p = 0.27$)、成熟度 ($p = 0.32$) 和训练史 ($p = 0.82$) 对多维模型没有显著影响。

讨论

这些最新发现表明，通过使用 U14 数据，五种精英青少年足球人才选拔分类模型中的每一种都对 U19 球员身份 (职业与非职业) 的预测做出了显著贡献。即使是运动能力模型中最低的 YI (0.43) 也能达到

表8 二元逻辑回归多维模型系数

变量#	β	SE	Wald	df	p	机会比率 [95% CI]
家庭支持	1.45	0.49	8.75	1	< .01	4.24 [1.63; 11.05]
技术	1.10	0.47	5.54	1	0.02	3.00 [1.20; 7.48]
总体运动能力	-0.99	0.48	4.22	1	0.04	0.37 [0.15; 0.96]
心理特征	0.54	0.49	1.20	1	0.27	1.71 [0.66; 4.43]
生物成熟度	0.48	0.48	0.99	1	0.32	1.61 [0.63; 4.14]
训练史	0.07	0.29	0.05	1	0.82	1.07 [0.60; 1.90]
常数	-2.21	0.38	33.33	1		0.11

变量按照 β 系数绝对值排序。

73% 的正确选拔决策。这表明，所有被调查的人才选拔方法（教练评估、运动能力测试和多维数据）以及它们的特定组合都是强有力的工具。因此，它们有理由被普遍运用并被推荐使用（Christensen, 2009; Höner 等, 2017; Williams 和 Reilly, 2000）。

不同人才选拔方法比较

通过对不同方法预测有效性的比较发现，教练评估与运动能力测试或多维数据没有区别，而运动能力测试数据的预测通过多维度测评得到了提升。因此，如果人才培养系统仅依赖于人才选拔的测量数据，那么，应当优选多维数据。这符合以前的科学建议（Abbott 等, 2005; Vaeyens 等, 2008）。

关于工具组合的潜在益处，运动能力测试并不能提高教练眼光的预测有效性，但是在教练眼光和运动能力组合模型中增加教练评估，运动能力测试就会从中受益。此外，教练评估和多维数据组合而成的整体模型优于各个单独的部分，但与教练眼光和运动能力测试的组合没有区别。这为在精英青少年足球中普遍使用教练评估和测量数据来选拔人才的做法提供了有力证据。对于测量数据，在与教练评估相结合的情况下，只存在有必要使用多维数据的趋势，因为整体模型并不明显优于教练眼光和运动能力组合的模型。综上所述，对人才选拔进行整体判断的思路似乎富有成

效（Reilly, 2006; Zuber 等, 2016）。然而，如果某些原因禁止多维数据测量（如经济原因），只在教练评估中添加运动能力测试的数据可能会产生类似的结果。

教练评估

根据不同选拔模型之间的比较，总体而言，教练评估对球员比赛表现的评分是对球员将来足球事业成功的有力预测（即只有整体模型明显优于教练眼光模型）。此外，整体模型回归系数中的 OR 也突出了这一点：OR 为 12.53[2.97; 52.83] 时，获得一个教练评分更高的标准差可以提高成为职业球员的概率。比赛表现这一变量所具有的高度影响与教练评估在该领域的普遍使用（Christensen, 2009）及其在科学界日益增长的正面支持（Fenner 等, 2016; Jokuschies 等, 2017; Romann 等, 2017）是一致的。教练评估的预测有效性之所以较高，一个可能的原因就是它具有全面性（Buckers 等, 2015）。在当前调查中，这还可能对比赛表现这一变量产生影响。俱乐部教练对每个球员都有广泛的了解（例如，关于家庭支持和训练史），因此在他们进行评分时，比赛表现可能不仅包括基于比赛的方面（例如技术和战术技能，总体运动能力），还可能进一步受到不同方面知识的影响。就与短期教练评估的可比性而言，这似乎是对当前调查的一种相关限制（即教练和球员互不了解）。教练评估和能力标准之间的某种依赖性（职业与

非职业)可能会对当前结果产生进一步限制。如前所述,在该领域使用教练评估非常普遍(Christensen, 2009),因此成为一名职业球员大部分时间取决于通过教练评估而被选入职业球队。

运动能力测试和多维数据

与当前研究中的其他方法相比,虽然运动能力测试的预测有效性最低,但有两个方面值得注意。首先,总体运动能力与成为职业运动员呈负相关。尽管Leyhr等(2018)报告称U12至U15年龄组的冲刺表现和成人表现水平之间没有显著关系,但这仍让人感到出乎意料(Gonaus和Müller, 2012; Murr等, 2018)。然而,成熟度(Malina等, 2017)样本的同质性以及对U14至U19的长期预测的混合影响,可能在一定程度上解释了这一点。此外,技能对精英青少年足球人才选拔的假设预测值(Höner和Votteler, 2016; Sarmiento等, 2018)在多维模式而不是整体模式中具有较高的正面影响。因此,在整体模型中,教练评估中的主要比赛表现可以解释技能上的一定差异,而后者在多维模型中仅用测量数据就达到了3.00 [1.20; 7.48]的显著且相关的机会比率。由于教练认为技能在这个年龄段非常重要(Larkin和O'Connor, 2017),因此这非常合理。

尽管多维模型的表现不如整体模型,但前者通过证明其优于运动能力模型,证实了其在科学界得到的认可(Williams和

Reilly, 2000)。这种优势主要是基于家庭支持的贡献,这一因素显示了有意义的OR值4.24[1.63; 11.05],并支撑了这一领域迄今为止被低估的对人才选拔的实用性(Zibung和Conzelmann, 2014)。尽管心理特征和训练史在与当前项目不同的调查中显示出了预测有效性(Sieghartsleitner等, 2018; Zuber等, 2015),而成熟度显示出与成人水平的成功有一定关系(奥斯托吉奇等, 2014),但它们目前对多维模型预测值的贡献并不显著。

本研究的一个局限影响了运动能力测试以及多维数据的结果:所用的统计方法。如前所述,在整体评估中整合几种变量并确定某些变量的负载至关重要(伯格曼和特罗特, 2006年; Till等人, 2016年; Till等人, 2018年)。主要问题是如何将多个变量或项目放入总体运动能力或多维模型中。由于当前研究的目的是获得不同人才选拔方法的直接比较,因此我们通过曲线模型(BLR)和代表最独立变量的平均 z 值来建立各种模型。因此,在一个发展中的体育人才的整体性中,假设的几个变量或项目之间的相互作用和补偿现象不能被适当地复现(Meylan等, 2010)。超越这种方法,从一般线性模型到非线性方法(例如以人为本的方法,ergman等, 2003),将可能实现这种相互作用和补偿现象(Conzelmann等, 2018)。此外,这些以人为本的方法将会适应人才培养问题的良好理论背景,而这些问题来自发展科学(即年轻人才是正在发展的人)及其动态互动方法(Zuber等,

2016)。然而，非线性的和以人为本的分析无法在不同的人才选拔方法之间进行直接比较，因此，使用目前的方法是当前研究问题的早期想法的必然结果。因此，我们必须考虑到，将几个项目整合到平均 z 值会阻碍进一步的观察，运动能力测试和多维数据的价值可能会被低估——参见本项目早期研究中心理特征和训练史的相关预测有效性 (Sieghartsleitner 等, 2018; Zuber 等, 2015)。由于其整体性 (Buckers 等, 2015) 并在当前调查中只在单一项目中有所体现，因此，教练评估中的比赛表现不受这一方法问题的影响，这也解释了其高预测有效性以及教练评估和测量数据组合模型的高预测有效性。

结论

本研究检验了不同人才选拔方法的预测有效性。总体而言，教练评估和测量数据之间的结合似乎对精英青少年足球人才选拔是有益的，因为这可以缓冲它们的共同优势和劣势。最好这种组合策略中的测量数据是多维的，而不是仅仅基于运动能力，这通过整体模型实现了最高的预测有效性。

这些结果表明，经常被推荐的多维方法 (Williams 和 Reilly, 2000) 显示出了相对于不太复杂的选拔策略的优势。然而，这种多维度的优势需进一步加以研究，以优化这些方法在该领域的人才选拔价值。尽管有几项调查已经完成，但仍需要进一步

了解多维数据中某些维度和项目的相关性 (Figueiredo 等, 2009)。更重要但几乎未被研究的领域，是在人才选拔中最大限度地利用多维信息的方法论方面 (Conzelmann 等, 2018; Till 等, 2016)。出于几种原因，本研究所应用的统计分析对不同人才选拔方法进行直接比较是必需的，但可能不适合分析多维数据。各项目的总和以及曲线模型的使用抑制了某些维度之间可能的相互作用和补偿现象 (Meylan 等, 2010)，并暗示了相同的统计模型对样本中的每个个体都有效 (Bergman 等, 2003)。一些非线性调查 (Pfeiffer 和 Hohmann, 2012; Pion 等, 2017) 和以人为本的方法 (Sieghartsleitner 等, 2018) 试图解决这些限制问题。然而，还需要进一步实施和优化这些方法，以便更广泛地了解它们在人才选拔方面的可能价值。因此，我们希望当前对多维方法的倡导能够鼓励更多的体育科学家超出线性统计，利用多维数据的假设力量，提高俱乐部和协会人才培养系统中人才选拔的质量。

在足球比赛中有赢得比赛的最佳阵容年龄？

拉法莱·波利，卢瓦克·拉夫内尔，罗杰·贝松

1 简介

研究报告对 31 个欧洲顶级联赛的球队年龄结构进行了分析。据此，我们可以从国际层面说明差异，同时还可以对球队年龄结构与成绩之间的联系进行调查。

各联赛和俱乐部的球员年龄差别很大。虽然最具竞争力的球队和冠军不是最年轻的，但通常也不是最有经验的球队。观察到的差距也反映出存在不同的文化和战略。

这项研究从 2009 年到 2017 年历时 9 年。为了纳入分析，球员必须在 10 月 1 日出现在被调查俱乐部一队阵容中。此外，他还应该在参考赛季期间参加过国内联赛，或者如果不是这种情况，则应该在前两个赛季均参加过成年锦标赛。2 号和 3 号守门员也是考察对象。

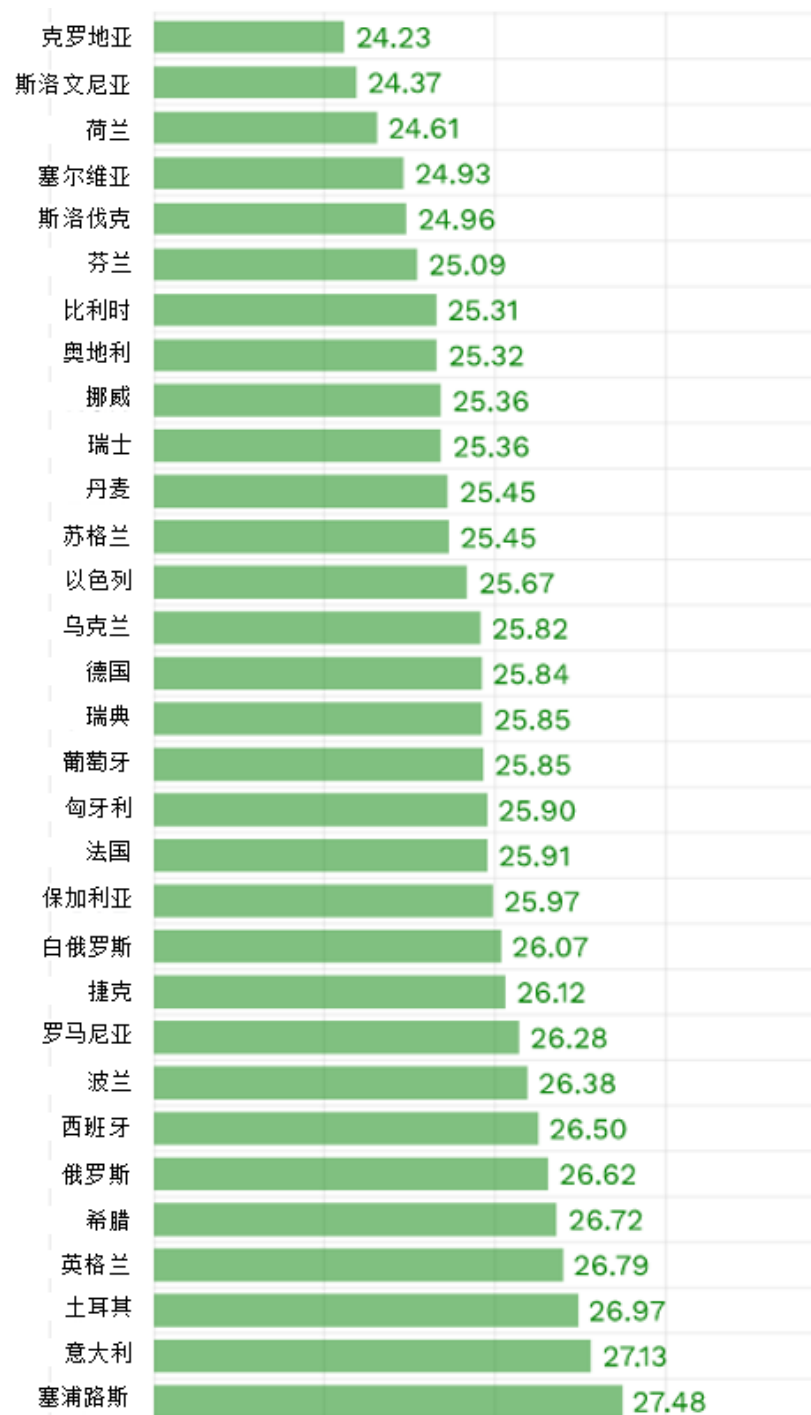
表 1 调查的联赛和球员 (2009-2017)

奥地利	奥超联赛	848
比利时	比甲联赛	1510
白俄罗斯	白俄罗斯超级联赛	1025
保加利亚	保加利亚甲级联赛	1346
克罗地亚	1.克罗地亚甲级联赛	1190
塞浦路斯	1.塞浦路斯甲级联赛	1541
捷克	捷克联赛	1251
丹麦	丹麦超级联赛	1005
英格兰	英超联赛	1753
西班牙	西甲联赛	1721
芬兰	芬超联赛	1080
法国	法甲联赛	1778
德国	奥超联赛	1468
希腊	希超联赛	1677
匈牙利	匈牙利甲级联赛	1314
以色列	以色列超级联赛	1155
意大利	意甲联赛	1742
荷兰	荷甲联赛	1550
挪威	挪超联赛	1333
波兰	波兰甲级联赛	1427
葡萄牙	葡超联赛	1771
罗马尼亚	罗马尼亚甲级联赛	1648
俄罗斯	英超联赛	1429
苏格兰	苏超联赛	1165
塞尔维亚	塞尔维亚超级联赛	1576
瑞士	希超联赛	909
斯洛文尼亚	塞尔维亚超级联赛	1139
斯洛文尼亚	1.斯洛文尼亚甲级联赛	1016
瑞典	瑞典超级联赛	1324
土耳其	土超联赛	1636
乌克兰	英超联赛	1223

2 年龄差距

根据联赛的不同，欧洲各球队的年龄差异很大。塞浦路斯顶级联赛是在 2009 年至 2017 年期间对最有经验球员进行重组的赛事：平均 27.5 岁。意大利甲级联赛的球队年龄也同样很大（27.1 岁），比克罗地亚和斯洛文尼亚的联赛平均年龄大三岁左右。所有 31 个顶级联赛 9 年分析的球员平均年龄为 25.9 岁。

表 2 球队平均年龄（按联赛）（2009-2017）



球队之间的球员平均年龄的差异也很大。表 3 列出自 2009 年以来所调查的 31 个顶级联赛中球员年龄最大的俱乐部。2015 年切沃队创下历史新高：30.6 岁。另外，还有两支球队的球员平均年龄也超过 30 岁：塞浦路斯联赛的 AEK 拉纳卡队（2016 年）和希腊人竞技队（2012 年）。在年龄最大的 10 支球队中有 8 家俱乐部位于地中海地区（意大利、塞浦路斯、土耳其和希腊）。这是表现区域特征的第一个指标。

表 3 平均年龄最大的球队，31 个欧洲顶级联赛（2009-2017）

切沃（意大利）	2015	30.58
拉纳卡AEK（塞浦路斯）	2016	30.44
切沃（意大利）	2016	30.40
希腊人竞技（塞浦路斯）	2012	30.38
拉纳卡AEK（塞浦路斯）	2017	30.29
AC米兰（意大利）	2011	29.98
图拉兵工厂（俄罗斯）	2016	29.76
拉齐奥（意大利）	2011	29.70
AC米兰（意大利）	2009	29.69
伊斯坦布尔巴萨克赛尔（土耳其）	2017	29.68

另一方面，自 2009 年以来，没有一支球队的球员年龄低于 2013 年的贝尔格莱德拉德队：平均 21.0 岁。10 支最年轻的球队中没有一支球队属于最具竞争力的欧洲联赛。其中 6 支球队来自塞尔维亚、克罗地亚和芬兰。在排名前 10 的俱乐部中，一些俱乐部的阵容非常年轻主要反映出俱乐部的财务问题。2013 年阿尔基拉纳卡俱乐部和哈茨俱乐部情况尤其如此。

表 4 平均年龄最小的球队，31 个欧洲顶级联赛（2009-2017）

拉德（塞尔维亚）	2013	21.01
阿尔基拉纳卡（塞浦路斯）	2013	21.18
斯塔卡卡姆斯科耶（乌克兰）	2017	21.24
哈茨（苏格兰）	2013	21.33
万塔PK35（芬兰）	2016	21.62
萨格勒布火车头（克罗地亚）	2012	21.63
洪卡（芬兰）	2014	21.65
AS维塔（斯洛伐克）	2014	21.74
斯普利特海杜克（克罗地亚）	2013	21.80
OFK贝尔格莱德（塞尔维亚）	2014	21.80

不同位置的球员平均年龄也有很大差异。一般规律是守门员比后卫年龄大，后卫比中场球员年龄大，中场球员则比前锋年龄大。在所调查的 31 个联赛中，守门员的年龄比前锋差不多大 1.5 岁。主要原因是场上位置对身体素质要求不同，以及守门员经验很重要。后卫相对于中场球员和前锋也同样如此。



图 1 各位置平均年龄，31 个欧洲顶级联赛（2009-2017）

另一个规律是，最具竞争力的联赛通常汇集的有经验球员多于最不具竞争力的联赛。球队年龄与欧足联排名之间存在的统计学上的显著负相关性证实了这一观察结果。但是，也有许多例外情况。例如，荷兰和克罗地亚俱乐部在欧洲杯中的表现明显超过他们球员的平均年龄。相反，塞浦路斯队则没有达到人们对其球员经验水平的预期结果。

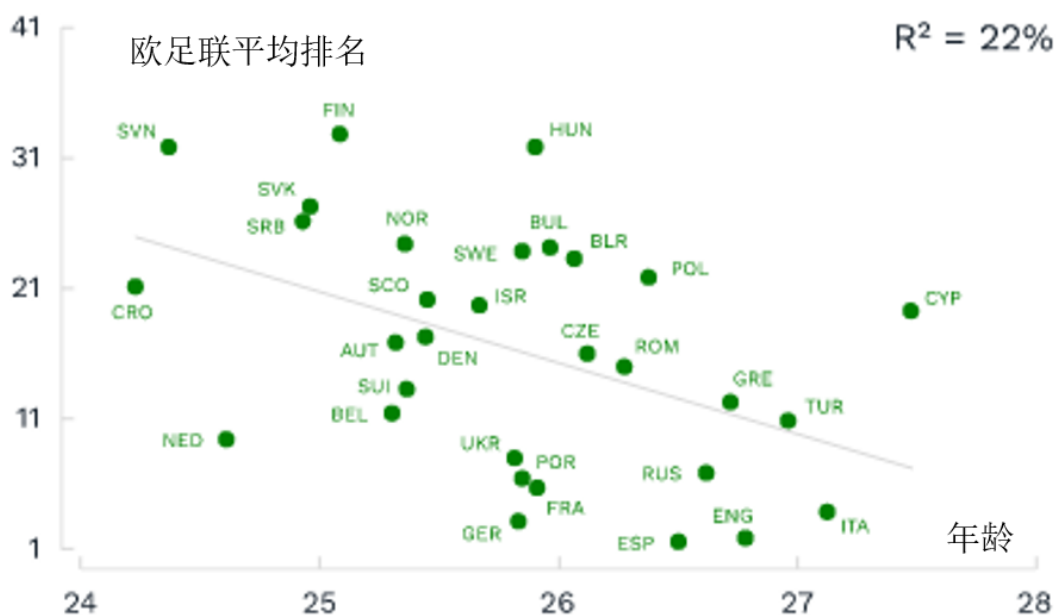


图 2 各联赛平均年龄与欧足联排名之间的相关性（2009-2017）

俱乐部的运动水平与球队的平均年龄之间也存在相关性。运动水平分类方法所考虑的因素有足协代表队在欧洲俱乐部赛事中的表现，球员所属俱乐部在国内联赛中的级别以及所取得的成绩。

一般来说，俱乐部越具有竞争力，阵容年龄更大。此规律特别适用于图 3 中所

示的四个最低水平俱乐部类别。然而，在水平最高的三个俱乐部类别中没有观察到显著差异。这一发现表明，在运动和经济实力达到一定的门槛后，俱乐部的年龄结构趋于一致。我们的分析表明，球员的最佳平均年龄在 26 岁至 27 岁之间。

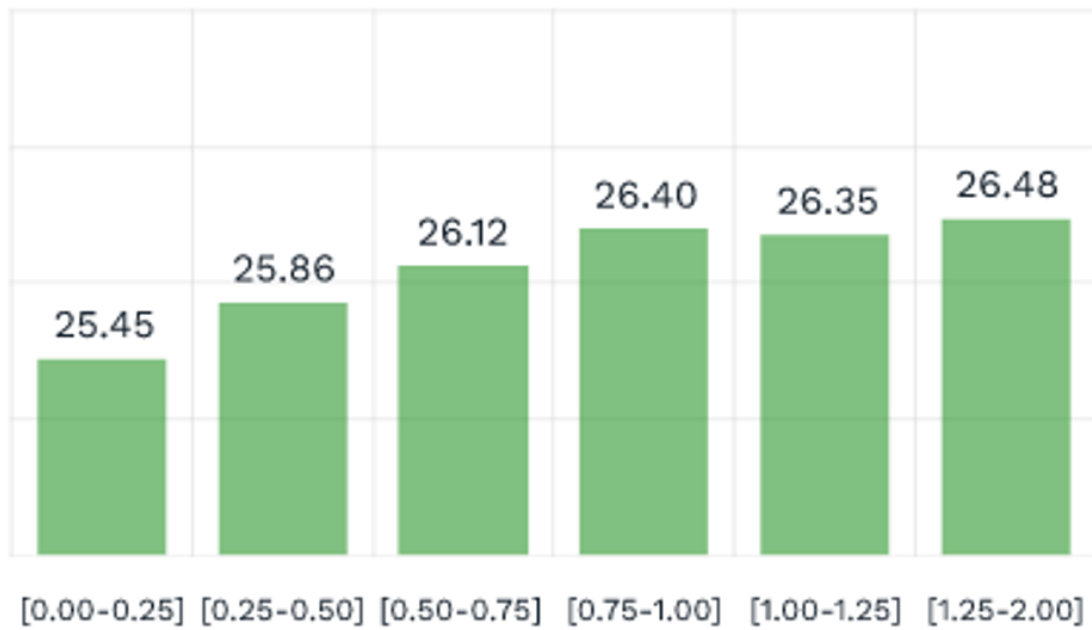


图 3 球员平均年龄和运动水平

3 冠军球队的年龄

在分析期间，10 支最年轻的冠军球队中没有一支属于最具竞争力的欧洲联赛。斯洛伐克的特伦钦足球俱乐部在排行榜上领先两家荷兰俱乐部：阿贾克斯和埃因霍温足球俱乐部。后面的例子表明，即使在欧足联排名前十的国家，年轻阵容也有可能赢得冠军。然而，在欧洲足球金字塔的最顶层，情况通常并非如此。在调查期间，年龄最小的五大联赛冠军是 2011/2012 赛季的多特蒙德队（24.7 岁）。

表 5 平均年龄最小的冠军，31 个欧洲顶级联赛（2009-2017）

AS维塔（斯洛伐克）	2014	21.74
AFC阿贾克斯（荷兰）	2012	22.66
PSV埃因霍温（荷兰）	2014	22.84
康斯坦察未来（罗马尼亚）	2016	23.01
NK马里博尔（斯洛文尼亚）	2010	23.02
马尔默（瑞典）	2013	23.20
萨尔茨堡红牛（奥地利）	2015	23.23
北西兰（丹麦）	2011	23.56
鲍里索夫BATE（白俄罗斯）	2009	23.68
凯尔特人（苏格兰）	2012	24.02

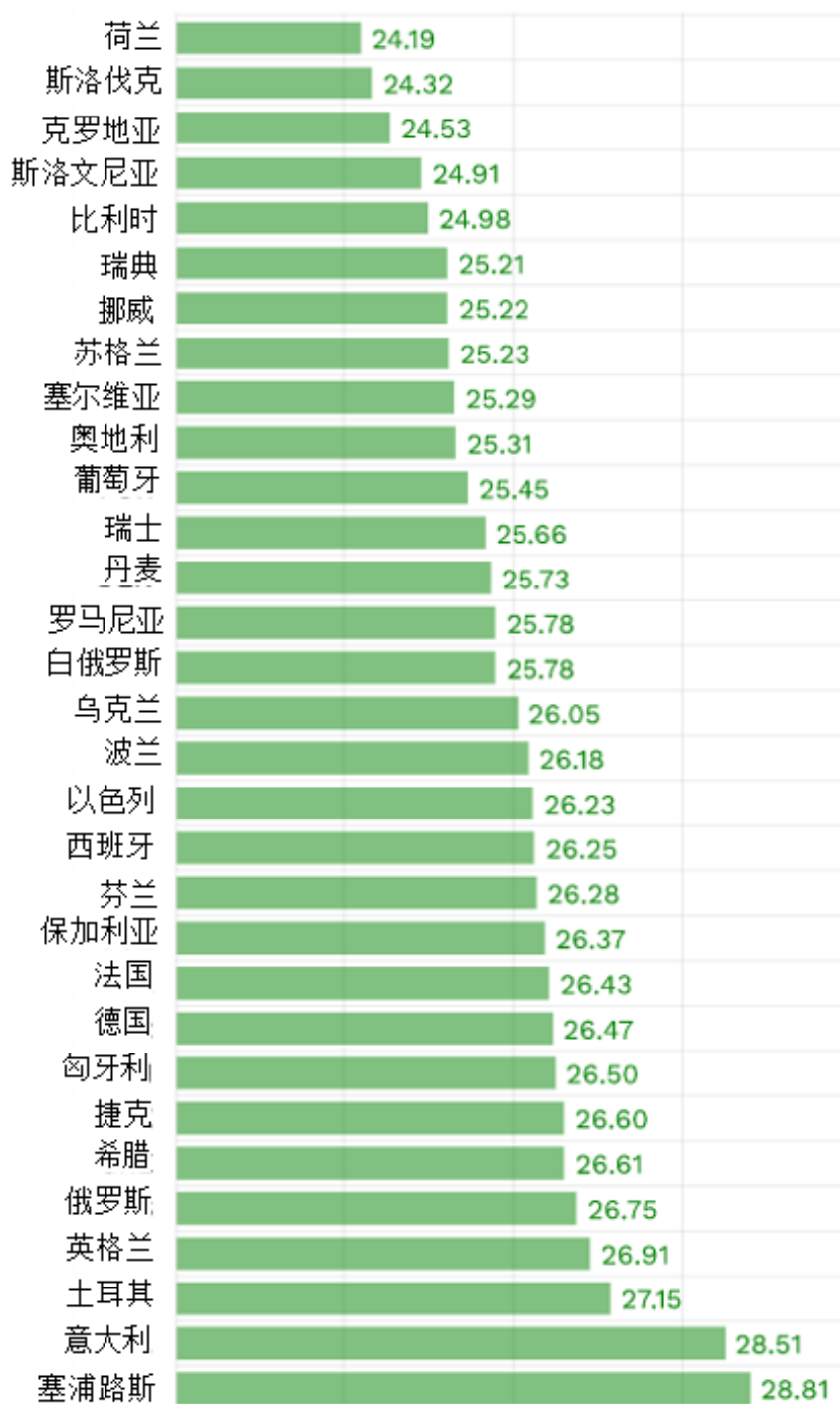
2009 年至 2017 年期间平均年龄最大的冠军俱乐部前 10 名名单显示出意甲联赛具有特殊性。在球队最有经验的 4 支冠军球队中有 3 支来自意甲球队。来自特别偏爱有经验球员的国家另一支球队是：塞浦路斯的希腊人竞技队。这支球队也是唯一一支赢得联赛冠军时其球员平均年龄超过 30 岁的冠军球队。

表 6 平均年龄最大的冠军，31 支欧洲顶级联赛（2009-2017）

希腊人竞技（塞浦路斯）	2012	30.38
AC米兰（意大利）	2010	29.20
国际米兰（意大利）	2009	28.92
尤文图斯（意大利）	2013	28.86
萨尔茨堡红牛（奥地利）	2009	28.35
皮尔森（捷克）	2014	28.30
圣彼得堡泽尼特（俄罗斯）	2011	28.28
费内巴切SK（土耳其）	2013	28.28
比尔舒华夏普尔（以色列）	2016	28.17
拜仁慕尼黑（德国）	2016	27.98

对不同联赛冠军平均年龄的比较分析还显示出对一个球队最有利的年龄结构的认知存在文化差异。一个极端是，最年轻的冠军来自于荷兰（24.2岁），那里的球队非常重视年轻球员的培养和发展。而另一个极端是塞浦路斯冠军队，调查期间球队年龄最大（28.8岁）。

表 7 冠军球队平均年龄（按联赛）（2009-2017）



4 结论

阵容组合是一门艺术，只有细心组织才能获得成功。有证据表明，要持久达到最高水平，必须保持良好的平衡。平衡的年龄结构允许年轻球员和有经验的球员一起成长，并逐步取代他们成为球队的支柱。这也是长期保持令人满意的稳定性和竞技水平的必要先决条件。

虽然年龄结构与成功之间的关系并不唯一，但 2009 年至 2017 年期间欧洲五大联赛冠军的年龄平均数是一个能够作为基准的很好指标。这一数字是 26.5 岁。根据之前的调查结果，我们可以说，为了取得可持续的成功，俱乐部未满 27 岁的球员数量要尽可能多于超出这一年龄门槛的球员数量。

足球科技信息

