The Performance Science Index: relationships with estimated market value and relative overall sporting performance of a selection of elite football (soccer) teams

M. Buchheit, [1][2][3]**Sam Gormley,** [1] **K. Hader,** [1] **D. McHugh** [1]

[1] Kitman Labs, Dublin, Ireland
[2] Lille OSC, Lille, France
[3] HIITscience, Revelstoke, Canada

Sport science | Performance science | Winning | Economics | Key performance indicators | Benchmarking

Headline

Alex Ferguson said about 10 years ago that "sport science was the biggest and most important change in (his) life-time". Since then sport science hasn't stopped growing in elite sports, with most elite clubs employing multiple sport scientists, data analysts and external scientific consultants looking at their data in search of making a difference and facilitating success (Training Ground Guru).

Monitoring training load, response to training (i.e., wellness and fatigue), sleep and many other performance- and health-related variables has become common (daily) practises, mostly thanks to advances in technology - in which teams invest exponentially (2, 4, 6). There is also a growing need to register and classify injury data, which helps clubs to benchmark their own incidence and burden in relation to both performance on the pitch and clubs' financial status (5). While there is a persistent feeling that all these human and technological investments serve team performance, very little is known about the actual sport performance science practises in real-life scenarios, in terms of process, variables monitored and frequency of collection, storage, treatment and use to make decisions. Overall, there is a lack of consensus in terms of best performance science practices.

To provide initial direction to practitioners in clubs and allow some types of benchmarking practices, the Performance Science Index (PSI) (1) was developed at Kitman Labs. The PSI is designed to assess a club's current performance science practises in the main 13 sport and medical science categories. It allows practitioners to understand where gaps may exist in their typical data workflow, collection, storage and data analysis. By better understanding what is being captured, and what may be missing, people can get a clear view on which performance science category to prioritize, where to move forward and what to implement. Nevertheless, the actual drivers for sport performance practices implementation are still unclear (e.g., cultural vs economics), and it is still unknown whether performance practices are actually team sporting performance.

Aim

The aim of this paper is to examine the associations between a selection of elite football (soccer) clubs' PSI overall scores, and both their estimated market value and relative overall sporting performance (i.e., Elo ranking).

Performance Science Index

The input to the Performance Science Index (1) is a survey of a team's data sources. The data sources are segmented into 13 distinct Performance Science categories (one for each Performance Science category, see below), with a total of 135 questions overall. A key driver behind the Index is to be able to benchmark teams against their peers based on their data collection practices. That requires a score of some sort to be developed which aggregates the survey responses into a single measure. There are clearly many ways to do this, and decisions such as whether all questions are weighted equally or certain questions or categories are weighted more heavily than others require input from domain experts. What matters for comparison purpose is not only consistency but how the score is constructed - the weightings etc - should be transparent if a PSI is being quoted outside the organisation. After testing multiple formats we settled on a score that combines all questions' responses to arrive at a numeric value on a scale of 0-100. A perfect score of 100 would correspond with an organisation that collects each type of data in the survey at the maximum possible frequency, and stores all of that data within a single centralised platform. Centralising all data allows for optimal data manipulation and overall analytic process automation.

Performance Science categories

1. Programming information (e.g., session content and objectives, team and individual)
2. Testing and screening for performance (e.g., anthropometry, body composition, fitness, strength, speed, agility, lab vs field)
3. Testing and screening for injury risk mitigation (e.g., injury history, all sorts of screening, strength, mobility, flexibility, movement quality and technique)
4. Psychological profiles (e.g., interview, questionnaires, profiles)
5. Monitoring load during training (e.g., internal and external load, quantitative and qualitative)
6. Monitoring load during competition (e.g., internal and external load, quantitative and qualitative)
7. Match technical performance (e.g., events, videos)
8. Monitoring player's responses to load (e.g., subjective such as typical wellness scores, objective such as fitness, neuromuscular status and overall status using HRV, blood markers, hydration for example)

9. Capturing the context - everything that may affect load (stimuli), and in turn adaptation and responses (e.g., use of recovery strategies, nutritional interventions, away matches travel duration and time the team came back from the away match, temperature, heat index, surface type, etc.)
10. Injury epidemiology (e.g., mechanisms, diagnostics, exposure, delays)
11. Return to play (session contents, testing, qualitative vs quantitative assessments)
12. Dealing with missing data (e.g., international breaks)
13. Collecting player's information from their previous clubs (e.g., injuries, performance testing, load and match data)

Example of Performance Science Index Questions

The questions revolve around the types of variables collected, the frequency of their collection and their storage location e.g., centralised or not. For example, people at a given club collect Rate of Perceived exertion (RPE) daily as measure training load, and they store it in the Kitman Labs platform.

They also measure heart rate response to a submaximal run every 6 weeks as a measure of response to load (fitness), and those data are in a database in MS Excel, stored in one of the staff's computer.

Market value and overall sporting performance (Elo rankings)

- We used Transfermarkt for the market values,
- We used https://footballdatabase.com/ for the ELO Ratings. The Elo rating system is a method for calculating the relative skill levels of players in zero-sum games such as chess. It is named after its creator Arpad Elo, a Hungarian-American physics professor.

Results

The associations between a selection (based on completion date and data health checks) of representative clubs' PSI overall scores and both their estimated market value and Elo ranking are shown in Figure 1 and 2.
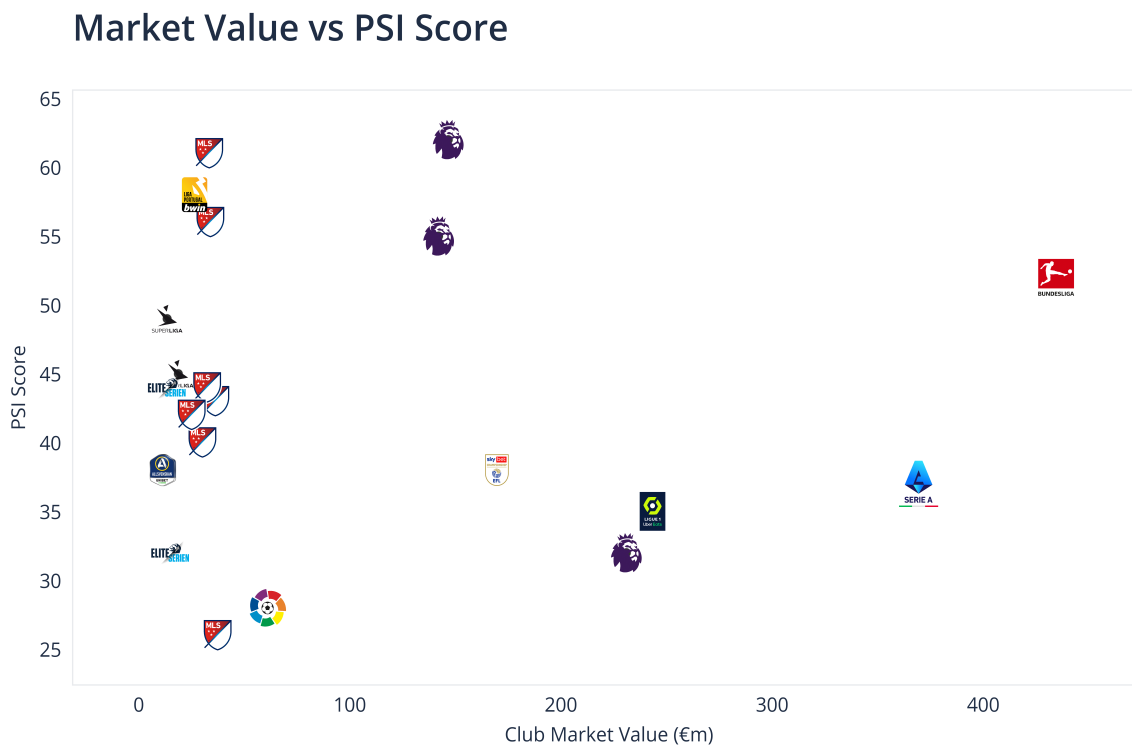

Fig. 1. Relationship between overall PSI scores and market value for a selection of representative clubs that completed the Performance Science Index survey at the same time (i.e., start of the 2021-2022 season) and with similar data health checks.
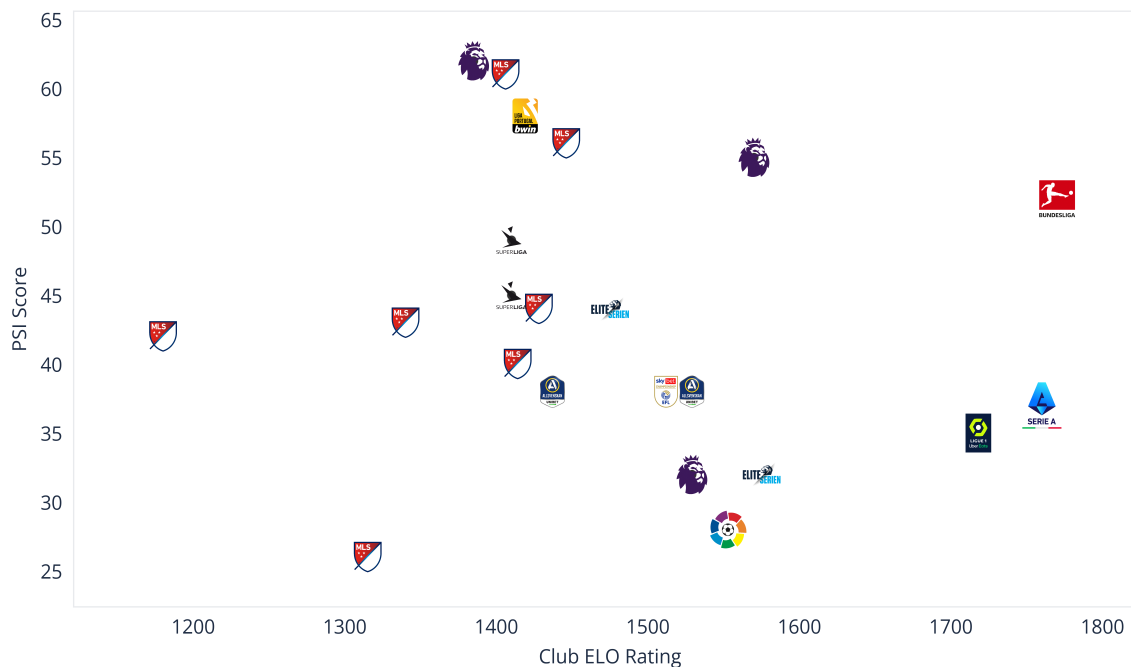
Fig. 2. Relationship between overall PSI scores and Elo ranking for a selection of representative clubs that completed the Performance Science Index survey at the same time (i.e., start of the 2021-2022 season) and with similar data health checks.

Discussion

The main results show a large variability of overall PSI scores. More importantly, overall PSI scores didn't seem to correlate with clubs' wealth (inferred from their estimated market value) (Figure 1). This highlights once more that there are definitely many ways to skin a cat, and that cultural biases and beliefs are likely more important than economics when it comes to implementing and using performance science.

Football is the most popular sport in the world, and the fact that it is so unpredictable is likely a great explanation for that. And if performance science was to have a direct, immediate impact on teams' overall performance, every single club would have embraced Sir Ferguson's vision. But the reality, again, in the authors' experience, is that there may be more important factors behind the ability of teams to win matches and trophies, such as player roster, competition squad composition and tactical models. While this belief was confirmed by the lack of correlation between PSI scores and Elo value (Figure 2), PSI scores should however also be examined in relation to other important aspects of the game, such as players availability and health, or even injury burden. This would provide another dimension to the importance of the PSI, and should be the matter of future investigations.

Limitations

The data analyzed and presented are as good as what the practitioners were happy to provide; there are therefore possible mismatches between what they reported doing, and what they actually do in their own context. Also, the PSI Score (as outlined in the calculation section), contains an element which awards higher points for teams that store data in a centralized platform (here, it's obviously referring to the Kitman platform rather than another vendor's); this means that they don't provide a fully objective metric for data collection practices. Finally, there's very little documentation for the calculation of the ELO ratings on `https://footballdatabase.com/`, so it's hard to verify them. Market Values on Transfermarkt have been documented to have flaws in their aggregation.

Conclusion and practical applications

- The Performance Science Index (PSI) is a tool to assess a club's current performance science practices, and allows practitioners to understand where gaps may exist in their typical data workflow, collection, storage and data analysis.
- There are no good or bad PSI scores; these scores should only be viewed used to direct practitioners toward the most pressing and impactful performance sciences areas to be developed in their structure.
- The lack of association between PSI scores and economical, performance and cultural contexts reflects practitioners' vision on the value of sport science - which may (or may not) be associated with success on the pitch - this is likely related to how practitioners use the information provided per se.
- Examining the association between PSI scores and other important aspects of the game, such as players availability and health, or even injury burden would provide another dimension to the importance of the PSI, and should be the matter of future investigations.

Twitter: M. Buchheit (@mart1buch), K. Hader (@Karad70), D. McHugh (@DerekMcHugh3)

References

1. M. Buchheit S. Gormley K. Hader & D. McHugh. The Performance Science Index, Oct 22.

2. Akenhead R, Nassis GP. Training Load and Player Monitoring in High-Level Football: Current Practice and Perceptions. Int J Sports Physiol Perform. 2016 Jul;11(5):587-93. doi: 10.1123/ijspp.2015-0331. Epub 2015 Oct 9.

3. Berndsen J, McHugh D. Comment on "Black Box Prediction Methods in Sports Medicine Deserve a Red Card for Reckless Practice: A Change of Tactics is Needed to Advance Athlete Care". arXiv preprint arXiv:2204.02402. 2022.

4. Drust B. Applied science and soccer: a personal perspective on the past, present and future of a discipline. Sport Performance & Science Reports, March 2019, #56, v1.

5. Eliakim E, Morgulev E, Lidor R, Mecke Y. Estimation of injury costs: financial damage of English Premier League teams' underachievement due to injuries. BMJ Open Sport Exerc Med. 2020 May 20;6(1):e000675. doi: 10.1136/bmjsem-2019-000675. eCollection 2020.

6. Impellizzeri FM, Marcora SM, Coutts AJ. Internal and External Training Load: 15 Years On. Int J Sports Physiol Perform. 2019 Feb 1;14(2):270-273. doi: 10.1123/ijspp.2018-0935. Epub 2019 Jan 6.