Review

International Journal of Sports Science & Coaching

The use of player tracking data to analyze defensive play in professional soccer - A scoping review

International Journal of Sports Science & Coaching
2022, Vol. 17(6) 1567–1592

DOI: 10.1177/17479541221075734
journals.sagepub.com/home/spo
SAGE

Leander Forcher[1], Stefan Altmann[1,3], Leon Forcher[1,2], Darko Jekauc[1] and Matthias Kempe[4]

Abstract

The analysis of tracking data in tactical game analysis is a topic of rising interest, as more detailed insights into performance structure in soccer can be obtained compared to traditional (e.g. notational) analyzes. Compared to the variety and detailed analyzes of offensive play, the number of studies analyzing the defensive play is low. However, in recent years, an increasing number of studies investigating defensive play have been published, so it seems useful to provide an overview of the current state of research in this area. Therefore, this study aims to identify the approaches that have been used to analyze the defensive play in professional soccer using player tracking data and to reveal the findings on successful defensive play. A systematic literature search of electronic databases (PubMed ($n = 604$), Web of Science ($n = 593$), and SPORTDiscuss ($n = 872$)) was conducted according to the PRISMA extension for Scoping Reviews (PRISMA-ScR). Studies that were included used tracking data of professional adult male soccer and analyzed defensive play. The result is a total of 23 studies that were analyzed in detail using the standardized quality assessment checklist for systematic reviews in sports science. The synthesis of results was carried out descriptively by organizing the results into different levels of tactical play (individual level, group level, team level). All included studies were of good methodological quality. The approaches to investigate defensive play using tracking data are highly heterogeneous (e.g. analysis of defensive pressure, analysis of synchronization, behavioral analyzes, ball recoveries). Successful defensive play is characterized by high pressure at the individual level, by high inter-team and intra-team synchronization and balanced defense at the group level, and by a compact coordinated organization at the team level. By summarizing the state of research on defensive play in soccer using sophisticated analysis approaches that showcase the possibilities of tracking data, this study provides an important foundation for future research in this area.

Keywords

Association football, performance indicators, tactical game analysis

Highlights

- The approaches to analyze defensive play in soccer using tracking data are highly heterogenous.
- Most promising approaches to analyze defensive play are the analysis of defensive pressure (at the individual level) and the quantification of compact organization (at the group level).
- Successful defensive play is characterized by high defensive pressure (at the individual level), inter-team and intra-team synchronization as well as a balanced defense (at the group level), and the contraction of organization (at the team level).

Introduction

Match analysis in soccer has been in the focus of interest for the scientific community and coaches for more than seventy years.[1] The majority of publications in science and practice are based on notational analysis, which typically identifies the who, what, where, and when of key game events.[2]

Reviewers: Filipe Clemente (Polytechnic Institute of Viana do Castelo, Portugal)
Bill Gerrard (University of Leeds, UK)
Carlos Lago-Peñas (University of Vigo, Spain)

[1]Institute of Sports and Sports Science, Karlsruhe Institute of Technology, Karlsruhe, Germany
[2]TSG 1899 Hoffenheim, Zuzenhausen, Germany
[3]TSG ResearchLab gGmbH, Zuzenhausen, Germany
[4]Center for Human Movement Sciences, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

Corresponding author:
Leander Forcher, Institute of Sports and Sports Science, Karlsruhe Institute of Technology, Engler-Bunte-Ring 15, 76131 Karlsruhe, Germany.
Email: leander.forcher@kit.edu

While this approach provides a basis for quantitative match analysis, it provides little information about the interactions within and between teams.[3] In contrast, recent reviews highlight the benefits of the use of tracking data[4,5] to achieve this. The tracking data include the positions of all players on the pitch and the ball using multiple measurement techniques like multiple camera tracking systems, global positioning systems (GPS), or radio-frequency identification (RFID).[6] The use of tracking data helps to produce more complex and reliable performance metrics compared to traditional analyses[7,8] and take the opponent interaction more precisely into account.[9] This is crucial to evaluate the performance in soccer and improves the validity of new performance indicators[10] that can help to understand the dynamics of the game and the underlying factors of success. This illustrates the importance of tracking data for match analysis in soccer.

Earlier research using tracking data focused on the physical performance of soccer players.[11,12] However, in recent years, the analysis of tactical performance with player tracking data received more attention, predominantly investigating the offensive play[13] focusing on actions from the attacking team's perspective.[14] These offensive and defensive soccer tactics can be divided into different levels: Individual tactics, group tactics, and team tactics.[1] This classification is used to structure the findings in this paper in order to provide a structured overview of the contributions to the different levels of individual, group and team tactics.

Subjects of interest in attacking play in soccer on an individual level were the evaluation of passes,[15–19] shots on goal (e.g. expected goals),[20] the danger of every attacking action,[8] or the space control of offensive players.[21] On a team level the organization on the pitch was investigated.[22]

In contrast, there is a lack of a comparable multitude of defensive analyzes.[23] Some studies already draw attention to the importance of defensive actions (e.g. ball recovery) as a starting point of the offensive phase.[24,25] However, various authors support the idea that soccer has a primarily defensive focus as it is played on a larger pitch with the difficulty of controlling the ball with the feet resulting in low numbers of scored goals compared to other team sports.[2,24] This idea can be supported by the results of Lepschy et al.[26] and Georgievski et al.[27] both showing that most of the critical success factors in soccer are defensive actions. Furthermore, Winter et al.[25] mentioned the behavior after the loss of possession to be important for success and Bosca et al.[28] reported that it is more important to improve defensive rather than offensive efficiency to be successful. Accordingly, Vilar et al.[2] strengthened the idea of the defensive characteristic of soccer by detecting the numerical superiority of defensive players to offensive players in sub-areas closer to the own goal. This illustrates the importance of defense in soccer and that it is essential to focus more on defensive game analysis.

However, the majority of studies evaluating defensive play use simple match statistics (e.g. tacklings, fouls, interceptions, yellow & red cards) or inverted offensive statistics (e.g. goals conceded, shots against) to analyze defensive play.[29–31] With the lack of quantity and quality (e.g. percentage of ball possession reveals little information about the actual game) of the used defensive statistics[23] those studies are not able to portray the complexity of defensive play in a highly dynamic team sport like soccer.

However, in recent years, there is a rising number of studies that investigated defensive play using tracking data in a sophisticated way.[32–34] The novelty of this topic leads to a high diversity in study designs missing a common basis as starting point of the research. Those heterogeneous studies investigated different aspects of defensive play (e.g. defensive pressure, synchronization, ball recoveries) and use different analyzing approaches (e.g. use of performance indicators, use of computational models) to extract important information out of tracking data. Following the PRISMA guidelines' extension for scoping reviews (PRISMA-ScR),[35] where a scoping review is characterized as a summary of a body of knowledge that is heterogeneous in methods, it appears meaningful to summarize the variety and nature of the evidence in the analysis of defensive play using tracking data in a scoping review to build a basis for future research in this area.[35]

Therefore, the aim of this study is twofold. First, this study aims to discover the approaches used for the analysis of defensive play at different tactical levels (individual level, group level, team level) in professional soccer using tracking data. Secondly, this study aims to reveal the findings about successful defensive play at different tactical levels (individual level, group level, team level) that have been obtained by analyzing defensive play in professional soccer using player tracking data.

To the best of the authors' knowledge, this study is the first to summarize the contributions to defensive analysis in professional soccer with up-to-date analysis methods (using tracking data).

Methods

Study design / search strategy

The scoping review was conducted according to the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyzes extension for Scoping Reviews (PRISMA-ScR)[35] and the guidelines for performing systematic reviews in sports science.[36] It was preregistered on open science foundation (osf.io) (DOI 10.17605/OSF.IO/29MWZ).

The literature search was conducted on 06/02/2021 using the following electronic databases: PubMed, Web of Science, and SPORTDiscuss. Every databases was

Table 1. Eligibility criteria.

Criteria	Inclusion criteria	Exclusion criteria
Population	- Studies in professional adult male soccer	- Studies in other sports than soccer (e.g. futsal, beach soccer, handball, basketball, american football) - Studies in other than professional level (e.g. amateur soccer) - Studies in other than adult male soccer (e.g. woman soccer, youth soccer)
Analysis	- Analysis of defensive play in 11 vs 11 matches	- Analysis of small sided games (SSGs)
	- Analysis of open play situations	- Analysis of dead ball situations (set pieces)
Data	- Studies using spatiotemporal tracking data	- Studies using other data than tracking data (e.g. notational data) - Studies analyzing simple match statistics (e.g. yellow/red cards, conceded goals)
Study design	- No restrictions with regard to study design	
Other	- Studies written in English	- Not available in English
	- Studies published in a peer-reviewed journal	- Not published in a peer-reviewed journal

examined with the following combination of keywords in the same structure of code lines:

("defensive" OR "defense" OR "defence" OR "defending")

AND ("soccer" OR "football")

AND ("performance analysis" OR "game analysis" OR "match analysis" OR "match analytics" OR "game performance" OR "match performance" OR "observation" OR "sports analytics" OR "team sports evaluation")

In addition, a secondary literature scan was executed by examining the reference lists of retrieved full-text articles to identify additional articles not identified by the initial search.

The exclusion and inclusion criteria for the decision on the inclusion of a study in this review are depicted in Table 1.

Data extraction

After the initial literature search, the study selection consisted of four phases as shown in Figure 1. The first phase consisted of the removal of duplicates. In the second phase, the abstract screening was conducted. Thereafter, the full texts of the remaining studies were screened. The aforementioned inclusion and exclusion criteria were applied to the inclusion and exclusion of studies in the second and third phases. The selection process from phase one to three was conducted by the first author of this review (LF). The final decision of inclusion of all remaining studies was discussed with two additional experienced researchers (MK, SA, LF). In the fourth phase, the full texts included in this review were assessed and data were extracted using a Microsoft Excel 2016 spreadsheet (Microsoft Corporation, Redmond, Washington, USA) containing the following information: Sample size (league, season), study design (performance indicator), defensive performance analysis, main results, quality assessment (see Table 2). The items *study design* and *defensive performance analysis* account for main aim of this review about the data analyzing approaches to analyze defensive play and represent the primary outcome. The item *main findings* shows the outcomes for the secondary aim of this study about the findings on effective defensive play.

In addition, the connection to success in every study was detected and shown in the main results as this connection to success is highly important to validate and examine the actual significance of the new processing approaches evaluating position data.[4] To provide a structure to this heterogeneous body of knowledge, all included studies are divided into the different levels of tactical defensive play (individual level, group level, team level) according to the work of Rein and Memmert.[1] Furthermore, all studies that did not fit the above classification because they used measures at different levels of tactical play were assigned to the mixed approaches section. This subdivision was determined by the first author (LF) and two experienced researchers (MK & SA).

Quality assessment

The quality of the included studies was assessed using the standardized quality assessment checklist provided by the authors of the Guidelines for performing systematic reviews in sports science.[36] Of the 23 items of this checklist examining the overall quality of sport science articles, 20 items were identified as applicable for this review and were used for the assessment. Two items of *technology used* were excluded because no technology guidelines were used in the considered studies and one item of *method* regarding dropouts was excluded because dropouts

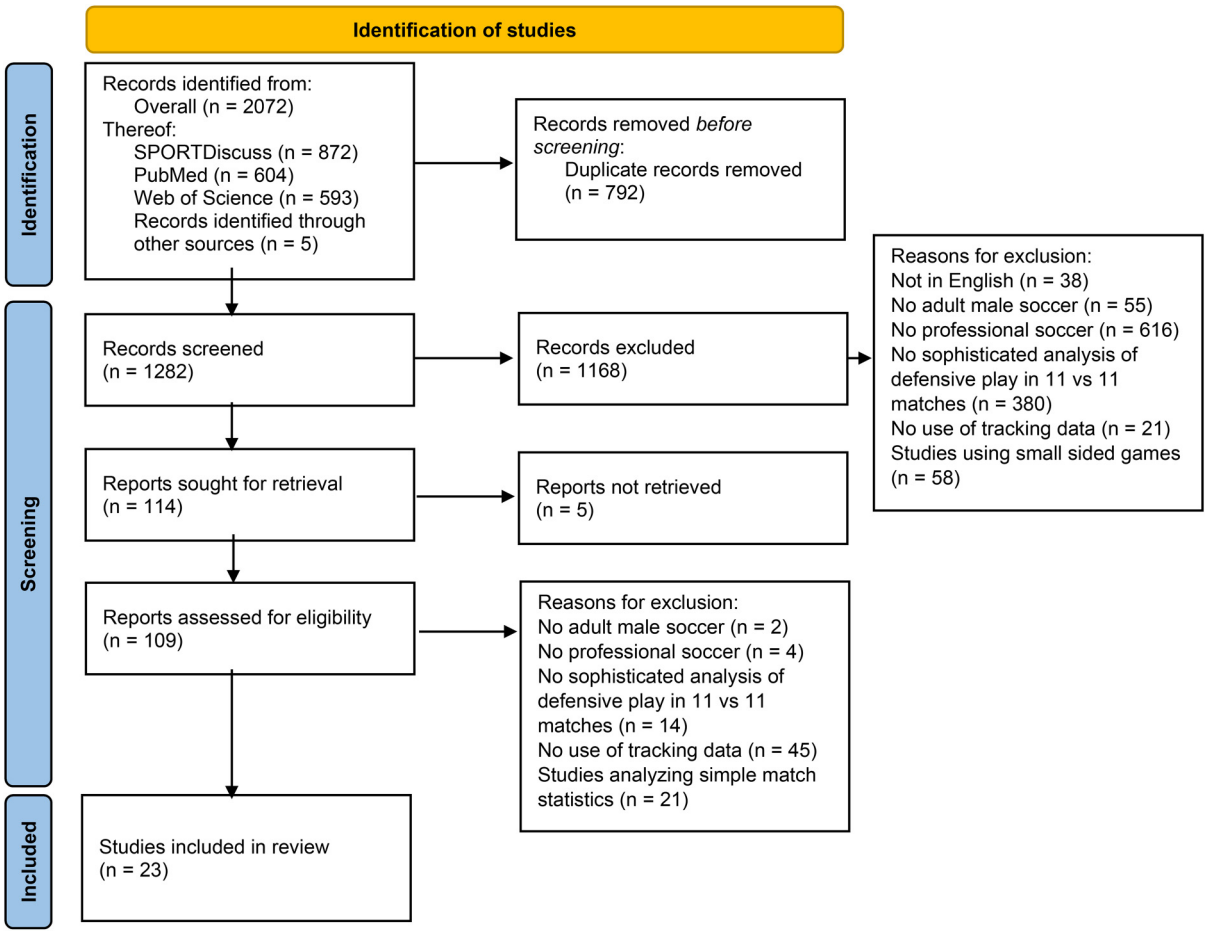

Figure 1. Literature search results.

are not expected in studies using tracking data of professional soccer matches. Each item could be answered either yes or no, with a score of 1 for yes and 0 for no. The total score was obtained by summing these scores across all items. Following the procedure of similar quality assessments,[50] the final scores were rated as: low methodological quality for final scores below 50% (final score: 1–9 of 20), good methodological quality for final scores between 50–75% (final score: 10–15 of 20), excellent methodological quality for final scores over 76% (final score: 16–20 of 20).

A risk of bias assessment across the included studies was not applied because it is not applicable to the expected heterogeneity of studies in scoping reviews.[35]

Results

Search results

The initial literature search resulted in a total of 2072 records (PubMed $n = 604$, Web of Science $n = 593$, & SPORTDiscuss $n = 872$) of which 4 records were added through the scan of secondary literature. In phase one, 792 duplicates were removed. The following title and abstract screening (phase two), using the inclusion and exclusion criteria, excluded 1168 records for different reasons (see Figure 1). The full texts of the remaining 109 records (full texts of 5 records could not be retrieved) were screened also using the defined inclusion and exclusion criteria (phase three). Finally, a total of 23 studies were included in this review.

Quality assessment

All studies reached at least good methodological quality (50–75%; final score: 10–15 of 20) (See Table 2). Beyond that, 14 studies reached excellent methodological quality (76–100%; final score: 16–20 of 20).

Sample characteristics

The study characteristics are depicted in Table 2. All included studies were published after 2011 with the most studies being published in 2017 ($n = 5$) and 2021 ($n = 4$). The number of published studies increased from 8 between 2012–2016 to 15 between 2017–2021.

Table 2. Study characteristics of included studies.

Study	Sample size (league, season)	Study design (performance indicator)	Defensive performance analysis	Main results	QA
Individual level					
Defensive pressure					
Andrienko et al.[37]	4 matches (of Borussia Dortmund of German Bundesliga, season 2015/16)	Validation of the developed pressure model (comparing mean pressure values with pressure values in successful defensive pressure situations). Analysis of pressure behaviors in case studies. (Defensive pressure or Pressing)	1. Pressure on the ball 2. Pressure on attacking players	Validity: - Association of high pressure values (on the ball & on the attacking players) with successful defensive actions due to high pressure (prove of validity) Pressure analysis: - Pressure on opposing players is higher in the defending team's half and in close distance from the ball (0–15 m) - Pressure analysis can help to investigate and provide additional insights into defensive tactics Relation to success: - Yes: Success of defensive pressing is defined as ball possession regain, ball going out of play, or ball turning away from the defended goal	17 (20)
Analysis of offense with opponent interaction (defensive pressure)					
Goes et al.[38]	299 matches (of Dutch Eredivisie, season 2018/19)	Computation of pass risk and pass reward. Analysis of the effect of positional role on the risk and reward of passing decisions. Pass risk: Classifier of the probability of a pass being successful. Identification of alternative pass options. Predicting the probability of successful passes to all pass options. Pass reward: Longitudinal outplayed opponents, potential of scoring opportunity, & Euclidean distance. (pass risk & pass reward with variables: Pass length, pass angle, path density, passer direction, receiver direction, forward displacement, goal distance pass, goal angle pass, goal distance reception, goal angle reception, pressure on pass, & pressure on receiver)	Defensive pressure on passer & defensive pressure on receiver	- Successful passes are characterized by higher predicted probability of being successful - Higher pass risk is not correlated with higher pass rewards - Defenders took the least amount of pass risk while forwards took the most pass risk - The model can accurately model pass risk and reward (validity) Relation to success: - No (but risk and reward of successful and unsuccessful passes are compared)	18 (20)

(continued)

Table 2. (continued)

Study	Sample size (league, season)	Study design (performance indicator)	Defensive performance analysis	Main results	QA
Lago-Ballesteros et al.[39]	12 matches (of one professional Spanish soccer team, season 2009/10)	Analysis of effects of different (offensive, defensive and situational) variables on score box possessions of one soccer team. (Offensive: duration, starting zone, team possession type, pass number, players in possession, and passing options, defensive: opponent number, and defensive pressure, situational: match location, quality of opposition, and match status)	Opponent number (Number of defending players located between the ball and their goal), defensive pressure (distance between the player with the ball and a direct pressing opponent player(s))	- Of all considered possessions, 33.4% produced score-box possessions, 52.5% achieved progression, and 14.1% did not reach any progression - Direct attacks and counterattacks are three times more effective than elaborate attacks (in achieving a score-box possession) - Attacks are more successful when starting from middle pitch zones and playing against less than six defending players compared to attacks starting in defensive pitch zones and against a balanced defense - No differences in the possession outcome were found for defensive pressure Relation to success: - Yes: Score box possession is defined as success criterion	16 (20)
Link et al.[8]	64 matches (of German Bundesliga, season 2014/15)	Validation of *Dangerousity* by comparing the *Dangerousity* values to the evaluation by semi-professional football coaches in 100 match scenarios. Exemplary game analysis based on the *Dangerousity* calculations. (*Dangerousity*: 1. Position of the ball, 2. Ball control, 3. Pressure that is put on the player by the opponent, 4. Density of opponent players in front of the goal)	Two defensive variables in *Dangerousity*: 3. Pressure that is put on the player by the opponent, 4. Density of opponent players in front of the goal	Validity: - *Dangerousity* quantification is in the same range as human observers *Dangerousity:* - *Dangerousity* shows higher correlation with betting odds compared to other key performance indicators - The derived metrics of *Dangerousity* (Action Value, Performance, Dominance) portray the match performance more accurate than traditional performance indicators (e.g. ball possession, shots, tackles, pass rates) because they better rule out effects of chance Relation to success: - Yes: Success is defined with betting odds	15 (20)
Szczepański & McHale[7]	760 matches (of English Premier League, seasons 2006/07 & 2007/08)	Development of passing model, calculating various predictions and estimating the passing ability of players. (Passing ability model: Based on the probability of a pass being successful Variables: time since the previous pass & the	Defensive pressure on passer (proxied with variables: origin of the pass, time since the previous pass & pass number, game time, action followed a duel, player's average position), defensive pressure on pass receiver (proxied with variables: intended destination	Validity: - The passing model shows a better fit to the logistic regression model of the end result compared to simple completion rate of passes Game analysis:	16 (20)

(continued)

Table 2. (continued)

Study	Sample size (league, season)	Study design (performance indicator)	Defensive performance analysis	Main results	QA
		pass number in the current sequence of passes for that team, game time, type of pass, action followed a duel, player is at his home ground, players' average position)	of the pass, time since the previous pass & pass number, game time)	- Headed passes are less accurate compared to passes played by foot and further have negative effects on the following pass - Passes are more likely to be accurate when played immediately after a ball recovery - The completion probability of a pass is about 76% for half of the passes and 90% for a quarter of the passes - Only about a quarter of the passes are more likely to be unsuccessful than to be successful Relation to success: - Yes: Success is defined with game outcome (logistic regression)	
Group level					
Behavioral analyzes					
Castellano & Casamichana[40]	655 matches (of Spanish BBVA (first-division) & Adelante League (second-division), season 2013/14)	Division of all teams into four groups using the end table of the season: top 10 teams in BBVA, bottom 10 teams in BBVA, top 10 teams in Adelante Leagues, bottom 12 teams in Adelante Leagues. Analysis of differences between the four groups. (Width, depth, height of defense, total running distance covered by team, center kicks, shots at goal, corner kicks, total passes made, & percentage of successful passes)	Height of defense (distance between the furthest-back defender and the defended goal)	- Higher ranked teams of BBVA performed better than the other three groups (for almost all variables) - In variables width and depth of play and height of defense lower ranked teams of BBVA performed significantly better than higher ranked teams of Adelante League - Teams in top half of BBVA league positioned their players closer to the opponents' goal compared to teams in the bottom half - Defense line of teams in top half of BBVA was closer to their own goal than that of both groups of Adelante league Relation to success: - Yes: Success is defined with end result of the season (top teams, bottom teams) and different leagues (1st and 2nd division)	19 (20)
Clemente et al.[41]	3 matches (of Portuguese Primeira Liga, season 2012/13)	Analysis of the influences of final score and half of the match factors on positions, defensive area of play and the number of triangulations in a given region of a defensive area of play.	Tactical position while defending & defensive play area	Tactical position: - Game ended in a draw: Highest number of defensive players (3.51) - Game ended in a loss: Smallest number	19 (20)

(continued)

Table 2. (continued)

Study	Sample size (league, season)	Study design (performance indicator)	Defensive performance analysis	Main results	QA
		(Tactical position (estimation of the number of defensive, midfield and attacking players), defensive play area (partial surface area between tactical positions))		of defensive players (3.05), highest number of midfielders (3.39) and attacking players (3.56) - Game ended in a win: Smallest number of attacking players (3.30) - No statistical differences between the first and second halves. Defensive play area: - Greatest defensive pressing occurred in the midfield region (compared to the backward and attacking region) - Greatest amount of defensive coverage was identified in the first half comparing both halves Relation to success: - Yes: Influence of end result (success criterion) on the variables is considered (but no effectiveness of results of performance indicators is discussed)	
Synchronization					
Folgado et al.[42]	4 matches (of English Premier League, season 2010/11)	Measurement of intra-team movement synchronization tendencies in 2 levels: 1. Between opposing teams, by comparing the winning and losing team in each match, 2. Within the analyzed team, by comparing different matches of one team according to the final outcome. (Intra-team synchronization on dyade level (2 players)	Intra-team synchronization between 2 players (dyads) with defensive role (all dyads formed by 2 defenders or by a defender and a midfielder)	- Teams losing tend to exhibit a lower value of movement synchronization - More synchronized behavior in defensive dyads compared to offensive dyads (indicating varying cooperating strategies across the pitch) - Synchronization comparisons are more useful when comparing the same team when the formation used is more likely the same (compared to between team analyzes) Relation to success: - Yes: Effect of synchronization on end result is discussed	19 (20)
Goes et al.[43]	118 matches (of Dutch Eredivisie, 4 seasons (years unknown))	Identification of subgroups using an automated formation descriptor. Calculation of centroids of subgroups (intra-team (between subgroups of one team) & inter-team (between subgroups of opposing teams, e.g. defense of one team and offense of other team)) and comparison between unsuccessful and successful attacks (using a spatiotemporal tool	Subgroup interaction of defending teams and their relation to attacking subgroups	- Defenders on the attacking team and attackers on the defending team show a decreased inter-team synchronization - Inter-team subgroup interactions were characterized by in-phase synchronous behavior - During successful attacks a slightly	17 (20)

(continued)

Table 2. (continued)

Study	Sample size (league, season)	Study design (performance indicator)	Defensive performance analysis	Main results	QA
		determining the probability of an attack resulting in a scoring opportunity). (Dynamic sub-group interaction)		a-synchronous behavior on a team-level especially in the longitudinal direction was found - Analysis of subgroup interactions provides more in-depth information compared to team-level variables (e.g. centroids of full team) Relation to success: - Yes: Success of an attack is defined with a spatiotemporal tool determining the probability of an attack resulting in a scoring opportunity	
Numerical superiority					
Vilar et al.[2]	1 match (of English Premier League, season 2010/11)	Characterizing the patterns emerging from player interactions in different sub-areas of the field. Investigating the stability and instability originating primarily in local numerical superiority by examining how teams place their players on the field during the game (collective offensive and defensive performance). Identification of key sub-areas of play that are important to stability and instability of a team. (Numerical superiority in sub-areas of the field)	Numerical superiority of defense in defensive sub-areas	- Teams place more players than their opponents in sub-areas closer to their own goal - Teams rarely place more players than their opponents in sub-areas more distant from their own goal - Sub-areas in the center of the field had a higher importance compared to the wings - Most frequent patterns of coordination were + 1-player numerical superiorities measured in the center-back sub-areas (importance of stabilizing and securing these defensive regions) - The winning team maintains higher defensive stability while concurrently risks more players to offensive sub-areas of field in comparison to the losing team Relation to success: - No: Only end result of one game is considered and analyzed (no calculations)	14 (20)
Analysis of offense with opponent interaction (density & opponent number)					
Lago-Ballesteros et al.[39]	12 matches (of one professional Spanish soccer team, season 2009/10)	Analysis of effects of different (offensive, defensive and situational) variables on score box possessions of one soccer team. (Offensive: duration, starting zone, team possession type, pass number, players in	Opponent number (Number of defending players located between the ball and their goal), defensive pressure (distance between the	- Of all considered possessions, 33.4% produced score-box possessions, 52.5% achieved progression, and 14.1% did not reach any progression - Direct attacks and counterattacks are	16 (20)

(continued)

Table 2. (continued)

Study	Sample size (league, season)	Study design (performance indicator)	Defensive performance analysis	Main results	QA
		possession, and passing options, defensive: opponent number, and defensive pressure, situational: match location, quality of opposition, and match status)	player with the ball and a direct pressing opponent player(s))	three times more effective than elaborate attacks (in achieving a score-box possession) - Attacks are more successful when starting from middle pitch zones and playing against less than six defending players compared to attacks starting in defensive pitch zones and against a balanced defense - No differences in the possession outcome were found for defensive pressure Relation to success: - Yes: Score box possession is defined as success criterion	
Link et al.[8]	64 matches (of German Bundesliga, season 2014/15)	Validation of *Dangerousity* by comparing the *Dangerousity* values to the evaluation by semi-professional football coaches in 100 match scenarios. Exemplary game analysis based on the *Dangerousity* calculations. (*Dangerousity*: 1. Position of the ball, 2. Ball control, 3. Pressure that is put on the player by the opponent, 4. Density of opponent players in front of the goal)	Two defensive variables in *Dangerousity*: 3. Pressure that is put on the player by the opponent, 4. Density of opponent players in front of the goal	Validity: - *Dangerousity* quantification is in the same range as human observers *Dangerousity*: - *Dangerousity* shows higher correlation with betting odds compared to other key performance indicators - The derived metrics of *Dangerousity* (Action Value, Performance, Dominance) portray the match performance more accurate than traditional performance indicators (e.g. ball possession, shots, tackles, pass rates) because they better rule out effects of chance Relation to success: - Yes: Success is defined with betting odds	15 (20)
Team level					
Behavioral analyzes					
Bartlett et al.[44]	10 matches (of elite European soccer (league unknown), season unknown)	Four categories of open attacking plays. The various measures of team centroid and dispersion between the two teams were correlated and compared between successful (goals & shots and headers on goal) and unsuccessful attacks (tackles & other losses of possession).	Measures of centroid and dispersion while defending	Centroid: - Strong positive correlations between attacking and defending team centroids for all four types of attacks - Little evidence for the decrease of inter-team distances between team centroids before critical match events (almost no crossing of the centroids)	15 (20)

(continued)

Table 2. (continued)

Study	Sample size (league, season)	Study design (performance indicator)	Defensive performance analysis	Main results	QA
		(2 measures of team centroid, 9 measures of dispersion)		Dispersion: - No contraction of defenders in successful defensive play compared to unsuccessful defensive play - Expansion contraction relationship (contracting in defense and expanding in offense) is not supported Relation to success: - Yes: Success of an attack is defined (successful: goal, shot on goal / unsuccessful: defensive tackle, turn-over of ball possession)	
Clemente et al.[41]	3 matches (of one professional football team (league unknown), season unknown)	Analysis of the influence of halves of a match and possession status on five collective metrics was conducted. (no statistic is mentioned) (Weighted centroid, weighted stretch index, surface area, & effective area of play)	Collective metrics while defending	- Teams' dispersion and average position on the field decrease during the 2nd half of the match - Tendency of expansion of players' placement on the field in ball possession (stretch index, surface area and effective area of play) - Tendency of contraction of players without ball possession - Support of the expansion contraction relationship Relation to success: - No	14 (20)
Moura et al.[12]	8 matches (of Brazilian First Division Championship, season unknown)	Analysis of team's coverage area and spread on the pitch when teams were attacking or defending. Analysis of difference of both measures between possession outcomes shot on goal and successful tackle. (Teams' coverage area (convex hull area) & teams' spread (Frobenius norm))	Teams' coverage area and spread while defending	- Decrease of coverage area and spread during defensive play (after loss of ball possession) - Increase of coverage area and spread during offensive play (after win of ball possession) - Support of the expansion contraction relationship - Teams present a greater surface area and spread in unsuccessful defending play sequences (shots on goal) compared to successful ones (successful tackles) - Teams present a greater surface area and spread in unsuccessful attacking sequences (successful tackle) compared to successful ones (shots on goal).	13 (20)

(continued)

Table 2. (continued)

Study	Sample size (league, season)	Study design (performance indicator)	Defensive performance analysis	Main results	QA
				Relation to success: - Yes: The success of an attack is defined with attack outcomes (successful: shot on goal, unsuccessful: successful tackle)	
Welch et al.[45]	2 matches (of Norway Eliteserien, in 2013 (season unknown))	Determination of the durations spend in each observed form of collective state during each phase of play (attacking, defending, or out-of-play). Investigation of global analysis, collective movement analysis by game phase, and collective state transitions. (Collective behavior/state (order parameter: polarization & angular momentum, mean group speed, team surface area))	Collective behavior in defending phase (order parameter: polarization & angular momentum, mean group speed, team surface area)	- Teams form polar and swarm collective states and switch between those collective states through direct and relatively fast transitions - With well aligned motion with polar nature teams achieve higher average group speeds - In the defensive phases of play the collective movement is more ordered, compact (i.e. lower surface area, high level of collective order) and faster moving in contrast to the attacking and out-of-play phases Relation to success: - No	17 (20)
Synchronization					
Frencken et al.[46]	1 match (of Champions League, season 2008/09)	Calculation of longitudinal and lateral inter-team distances for both halves. Identification of critical match periods (periods of high variability of inter-team distances over a moving 3-s window). (Inter-team distances (distance between longitudinal and lateral team centroids)	Inter-team distances (distance between longitudinal and lateral team centroids) (changes in inter-team distance might reflect changes in team pressure or switches between attacking and defensive pressure)	- High inter-team interactions of teams are discovered in both halves (strongest in longitudinal compared to lateral direction) - Inter-team distances are minimally related to critical match events (variability threshold was exceeded before three of sixteen critical match events (goals, goal scoring opportunities)) - The identified match periods with high variability of inter-team distances were predominantly (93%) connected with collective defensive pressure on the opponents (defending players moving forward to pressurize players of the attacking team) Relation to success: - Yes/No: Relation to critical match	16 (20)

(continued)

Table 2. (continued)

Study	Sample size (league, season)	Study design (performance indicator)	Defensive performance analysis	Main results	QA
				events (e.g. goals and goal-scoring opportunities), but no analysis or consideration of performance/success	
Ball recoveries					
Fernandez-Navarro et al.[47]	10 matches (of Spanish La Liga, season 2010/11)	Clustering of general defensive behaviors used by teams based on different quantitative defensive variables. Examination of the influence of categorial variables on this defensive behavior and the defensive outcome. (Categorical variables: Team analyzed, match status, venue, quality of opposition, match period Quantitative variables: Distance from the least advanced outfield defender to his goal line when defending teams gained the ball, distance between the player in possession of the ball to the nearest defender, length of the last pass that attacking teams made when conceded possession of the ball to defending teams, number of passes made before defending teams gained the ball, duration of defensive pieces of play, outcome of defensive pieces of play, outcome of offensive pieces of play following ball gains)	Defensive quantitative variables: Distance from the least advanced outfield defender to his goal line, duration of defensive pieces of play initiated in the opposing half, distance between the player in possession of the ball to the nearest defender, length of the last pass that attacking teams made when conceded possession of the ball to the analyzed team, number of passes made by attacking teams before analyzed teams gained the ball	Defensive playing styles: - Defense close to the own goal, mid-positioned defense with less intense pressure to attacking players, mid-positioned defense with more intense pressure, high-pressure defense Influence of categorial variables: - Defensive play was effected by match status and quality of opposition - Teams winning recovered more balls in pitch zones closer to their own goal - Teams losing recovered more balls in pitch zones closer to the opponents' goal - The weaker the quality of the opponent the more balls are recovered in advanced zones of the pitch Relation to success: - Yes/No: Influence of context variable match status is considered (winning, losing, drawing), but solely descriptive analysis of defending behavior according to match status	20 (20)
Santos & Lago-Peñas[32]	13 matches (of one elite Spanish soccer team, season unknown)	Investigation of the effect of three situational variables on five positional variables throughout ball recovery situations. Use of linear regression model to investigate how the situational variables affected each positional variable. (Ball recovery situations with positional variables: team compactness, compactness behind ball location, compactness forward ball location, team width, game center index situational variables: match location, quality of opposition, match status)	Ball recovery situations with positional variables: team compactness, compactness behind ball location, compactness forward ball location, team width, game center index	- Match location has no influence on the analyzed positional performance indicators in ball recovery situations - Match status and strength of opposition has an influence - Teams winning, had a greater approach between players and the ball compared to even scores - Teams winning or losing reduce their game center index and advance their team width compared to even scores - Teams playing against strong opponents advance their compactness forward (distance between the ball recovery	18 (20)

(continued)

Table 2. (continued)

Study	Sample size (league, season)	Study design (performance indicator)	Defensive performance analysis	Main results	QA
				and the player closest to the opponents goal line) Relation to success: - Yes/No: The influence of match status on ball recovery variables is considered, but no effectiveness of different ball recovery variables on performance/success is discussed	
Santos et al.[48]	13 matches (of one elite Spanish soccer team, season unknown)	Investigation of the influence of the situational variables on the team performance variables during ball recovery. (Team performance variables: ball recovery location in the field, position of the defensive line, position of the offensive line situational variables: match location, quality of opposition, & match status)	Defensive team performance variables: ball recovery location in the field, position of the defensive line	- Match status, match location, and quality of opposition have an influence on team variables during ball recovery - Teams losing (compared to winning or even scores), playing at home (compared to playing away), and against a weaker opponent (compared to stronger opponents) advance their ball recovery location as well as the position of the defensive and offensive lines further to the opponents' goal Relation to success: - Yes/No: The influence of match status on ball recovery variables is considered, but no effectiveness of different ball recovery variables on performance/success is discussed	18 (20)
Analysis of offense with opponent interaction (teams' organization)					
Goes et al.[15]	18 matches (of Dutch Eredivisie, season 2017/18)	Validation of *D-Def* (measure of passing effectiveness) with 3 steps: 1. Examination of connection between *D-Def* and the overall movement of the players of the defending team, 2.Investigation whether *D-Def* score can differentiate the performance of different passers and passes, 3. Assessment of predictive value of pass attributes (pass velocity, pass length, pass angle) on *D-Def*. (*D-Def*: Passing effectiveness computed with the defensive disruption following a pass)	Defensive disruptiveness (changes in defensive organization) following a pass	Validity: - Overall individual movement of the defending team is highly correlated with the D-Def score - D-Def is able to clearly distinguish between different passes and passers - Passing angle is not a determining factor for the D-Def score - Passing length and passing velocity are prominent factors to achieve a decrease in defensive organization Relation to success: - No	19 (20)
Mixed approaches (individual, group, & team level)					
Computational models					
Matsuoka et al.[33]	2 matches (of Japanese J1-League, in 2016 (season unknown))	Construction of defensive items (variables) via qualitative analysis (causal structure was	24 variables of defensive play	Defensive playing styles: - Set defense (maintain defensive	14 (20)

(continued)

Table 2. (continued)

Study	Sample size (league, season)	Study design (performance indicator)	Defensive performance analysis	Main results	QA
		extracted by soccer specialists using the Delphi method and the causal-effect analysis). Successful defensive play was defined as ball gain and unsuccessful defensive play was defined as a play that failed to gain the ball. Set of split value to differentiate between successful and unsuccessful defensive play by classification binary tree to reveal success criteria of defensive tactical play. (24 variables of defensive play)		organization), control defense (restrict and guide the opponent's offence move), and concentration defense (increase player density around the ball) Model: - Fourteen valid items (of all 24 items) are selected by evaluating item characteristics of soccer defensive tactical play items (unidimensionality of items, reliability, validity) - Ability score of soccer defensive tactical play is significantly higher in successful defensive play (with ball gain) compared to unsuccessful defensive plays (without ball gain) - The introduced ability score is a valid measurement of defensive play Relation to success: - Yes: Success of a defensive play was defined (successful: ball gain, unsuccessful: other than ball gain)	
Toda et al.[23]	45 matches (of Japanese J1-League, in 2019 (season unknown))	Validation of the classifier (VDEP) by comparing the prediction of true events compared to existing classifiers. Examination of the relationship between VDEP and the team performance in actual matches and during a season. Demonstrating of examples of game evaluation or complete season evaluation of a team. (Valuating Defense by Estimating Probabilities (VDEP) evaluates the potential increase in the number of ball recoveries and the potential decrease in the number of effective attacks with 73 variables (not all included variables are clearly stated))	VDEP (but no specific performance indicators)	- The VDEP method predicts true positives of events correctly - The VDEP value shows moderate correlation to the season outcome (winning points) - The introduced model can be a consistent indicator to evaluate both attacks (after the ball recovery) and defense itself (prevention of being attacked and the ball recovery) Relation to success: - Yes: Success is defined with season outcome (VDEP method correlated with season outcome)	15 (20)
Prediction models					
Le et al.[49]	100 matches (of professional soccer league (league unknown), season unknown)	Automatic identification of the tactical formation (player role alignment). Use of recurrent neural networks and a popular deep-learning tool to learn the fine-grained behavior model for each role in the tactical formation for each time step. Prediction of	Defensive ghosting (prediction of defending behavior)	- The ghosting model is able to maintain solid defensive formation and structure - The average level of deviation from the actual defending behavior across all players and teams is comparatively	10 (20)

(continued)

Table 2. (continued)

Study	Sample size (league, season)	Study design (performance indicator)	Defensive performance analysis	Main results	QA
		motion patterns of the league average team or of a particular team. (Defensive ghosting (prediction of defending behavior))		low (~4 meters) - The total expected goal value coming from different ghosting styles is highly correlated with the overall number of goals conceded in reality Relation to success: - Yes/No: Expected goals is defined as success criterion. Relationships between ghosting models and resulting expected goal values are discussed.	
Stöckl et al.[34]	1200 matches (of the top 5 European football leagues (leagues unknown), season unknown)	Split of dataset into 90% train and 10% test set. Execution of the predictions xReciever, xPass, and xThreat. Presentation of approach in case studies. (1. xReceiver: likelihood of a player receiving the next pass 2. xThreat: likelihood that there will be a shot on goal in the next 10 s 3. xPass: likelihood of the success of the pass 4. Player Availability: availability of every attacker off the ball at each frame using outputs of xReceiver and xPass 5. Defensive Impact: Detection of high level defensive concepts (e.g. ball and man orientated defending), defensive position play and off ball runs 6. Disruption Maps: Global visual representations of defending teams' ability to disrupt the oppositions attacking strategy)	5. Defensive Impact: Detection of high level defensive concepts (e.g. ball and man orientated defending), defensive position play and off ball runs 6. Disruption Maps: Global visual representations of defending teams' ability to disrupt the oppositions attacking strategy	- The accuracy of all three models (xPass, xReciever, xThreat) is satisfiying - Different defensive styles (man and ball orientated defending) and off ball runs are identified - Exemplary single game analysis with the calculated model values can reveal how a forward of one team was nullified by the other team's defense (e.g. forwards probability of receiving a pass was below average, only 3 runs were attacking forward runs) Relation to success: - No: Just analysis of one game where the end result is considered with no calculations	10 (20)

Table 3. Approaches used to analyze defensive play using tracking data.

Level of tactical play	Analyzing approach	Explanation	Metrics used	Aim of analysis
Individual level	Defensive pressure (see Figure 2, d)	- The magnitude of pressure exerted from all pressers on the pressure target - Andrienko et al. [35] used the distance and the direction of the closest defender(s) to the pressure target	- Pressure on ball [35] - Pressure on opponents [35]	- Quantification of the space a attacker has to perform actions - Quantification of the influence of defenders on the attackers
Group level	Numerical superiority	- Difference in the number of players between both teams in the same subarea of the pitch	- Numerical superiority in subareas of the pitch [2]	- Measurement of pitch areas where a team is more dominant - Measurement of pitch areas where a team has a higher probability for gaining the ball or maintain ball possession
	Behavioral analysis (see Figure 2, c)	- Analysis of the defensive behavior of subgroups	- Positioning of defensive line [36] - Tactical position while defending & defensive play area [37]	- Analysis of subgroups (e.g. defensive line)
	Synchronization (see Figure 2, a)	- Synchronous behavior is characterized by players moving in the same direction with the same speed and the same acceleration	- Intra-team synchronization on dyad level (pairings of two players) [38] - Inter-team subgroup interactions [39]	- Analysis of the synchronization of movement behavior of subgroups - Identification of coordination patterns of subgroups
Team level	Behavioral analysis	- Analysis of the defensive behavior of the whole team	- Centroids, surface area, spread of full team [12,40,41] - Polarization and angular momentum [42]	- Analysis of the whole team
	Synchronization (see Figure 2, b)	- Synchronous behavior is characterized by players moving in the same direction with the same speed and the same acceleration	- Inter-team distances of centroids of both teams [43]	- Analysis of the synchronization of movement behavior of the whole team - Identification of coordination patterns of the whole team
	Ball recovery	- The goal of defending play is to regain the ball from the opponent, also called ball recovery	- Ball recoveries (Fernandez-Navarro et al. 2019; Santos et al. 2017; Santos & Lago-Penas, 2019)	- Analysis of the goal of defensive play (to regain the ball)
Mixed approaches	Computational models	- Use of computational methods to develop a model about defensive play	- Ability score of soccer defensive tactical play [47] - Valuating Defense by Estimating Probabilities (VDEP) [23]	- Identification of hidden patterns in unstructured data
	Prediction models	- Use of computational methods to predict defensive play	- Defensive ghosting [48] - xReceiver, xThreat, xPass [49]	- Predictions of defensive behavior

The studies had an average sample size of 145 games. The majority of 15 studies used less than 30 games as a sample.

Most analyzes were carried out using data from the big European soccer leagues. There were five investigations in Spanish La Liga, three in English Premier League, three in Dutch Eredivisie, and two in German Bundesliga. Furthermore, there were nine studies using data from other Leagues (e.g. 3× Eredivisie, 2× Japanese J1 league) and four studies where no specific league was mentioned and therefore remained unknown.

Tracking data measurement systems were predominantly provided by tracking companies. 14 studies used tracking systems of data providers such as AMISCO (Amisco Pro, Nice, France), TRACAB by CHYRONHEGO (Chyron, New York, United States), OPTA by STATS and ProZone by STATS (Stats Perform, London, United Kingdom). The tracking systems of six studies remained unknown.

Main findings

A summary of all included studies is provided in Table 2. A summary of the approaches used to analyze defensive play using tracking data is depicted in Table 3 and a summary of findings on successful defensive play is presented in Table 4.

Every level of tactical play (individual level, group level, team level, mixed approaches) is structured in the following way. First, the details of the approaches to analyze defensive play are presented to account for the first research question to give an overview of the heterogeneous study

Table 4. Main findings of successful defensive play.

Level of tactical play	Analyzing approach	Finding
Individual level	Defensive pressure	- Successful defensive plays are characterized by high defensive pressure (on the ball & on the attackers) [35]
Group level	Numerical superiority	- Teams place more players than their opponents in sub-areas closer to their own goal - Teams rarely place more players than their opponents in sub-areas more distant from their own goal - Winning teams maintain higher defensive stability while concurrently risk more players to offensive sub-areas of field compared to losing teams [2]
	Synchronization	- Winning teams show more intra-team synchronization on dyade level (2 players) compared to losing teams [38] - Defenders on the attacking team and attackers on the defending team show a decreased inter-team synchronization in successful attacks [39]
	Positioning of defensive line	- Better teams position their defensive line further up the pitch [50]
	Number of defenders	- Attacks are less successful when starting in defensive pitch zones and playing against more than five defenders (balanced defense) [51]
Team level	Synchronization	- Slightly a-synchronous behavior on a team-level especially in the longitudinal direction during successful attacks [39]
	Contraction expansion relationship	- Contraction of players during defensive play (out of ball possession) and expansion of players during offensive play (in ball possession) → Support of contraction expansion relationship [12,42,52] - Teams present a stronger contraction (smaller surface area and spread) in successful defending play sequences compared to unsuccessful ones [12]
	Influence of categorial variables	- Teams recover more balls closer to their own goal when winning (compared to even scores or losing), playing away (compared to playing at home), and playing against a stronger opponent (compared to weaker opponents) [44–46]

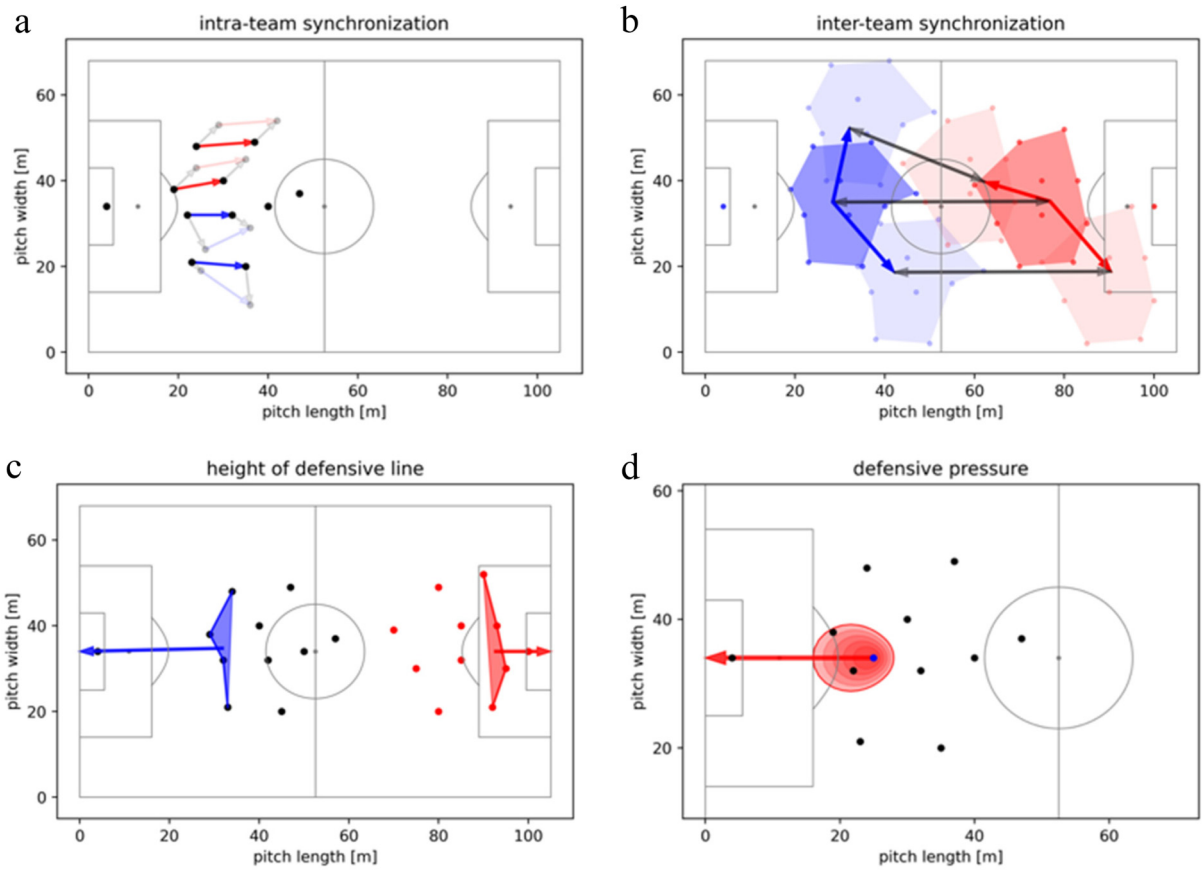

Figure 2. Visualization of exemplary key performance indicators:
a: visualization of intra-team synchronization on dyad level, with perfect synchronous movement behavior at the top (red) and asynchronous movement behavior at the bottom (blue).
b: Visualization of inter-team movement synchronization of opposing teams' centroids.
c: Visualization of the measurement of the height of the defensive line (blue).
d: Visualization of the measurement of defensive pressure: Two defending players (black) exerting defensive pressure on an attacking player (blue) with the threat direction towards the goal (red arrow).

approaches. Second, it is shown whether the approaches were presented in connection to success, which is important to show the actual significance of the new processing approaches and substantial for the interpretation of the results which are presented third in every section. Both the second and third paragraphs account for the second research question of this paper about the findings on successful defensive play.

Due to the heterogeneous study designs of the included papers, the results are only comparable with each other to a limited extend and are therefore only presented if applicable.

Individual level. Five studies investigated defensive play at an individual level. All studies used defensive pressure (on the ball and on the opponents) to investigate defensive play. Andrienko et al.[37] developed a method to quantify pressure with two different pressure measures (1. pressure on the ball, 2. pressure on the attacking players) based on the idea that pressing is the magnitude of pressure exerted from all pressers on a pressure target (visualized in Figure 2, d). This study was the only one to focus solely on the analysis of defensive pressure. In contrast, the other four studies analyzed offensive play considering defensive pressure as a variable of opponent interaction. Goes et al.[38] used the pressure model of Andrienko et al.[37] among other variables to evaluate passing ability, Szczepański and McHale[7] only proxied the pressure on the passer and receiver (without giving exact calculations) also evaluating passing ability. Both Lago-Ballesteros et al.[39] and Link et al.[8] used similar approaches to Andrienko et al.[37] by measuring the distance to attackers and the pressure on the ball leading player, respectively.

Regarding the investigation in connection to success Szczepański & McHale[7] and Link et al.[8] used the end result as success criterion at a game level. At a possession level, Lago-Ballesteros[39] defined the success of an attack as score box possession and Andrienko et al.[37] defined

the success of defensive play as ball possession regain, ball going out of play, or ball turning away from the defended goal. In contrast, Goes et al.[38] did not consider success in their investigation.

With regards to the findings to defensive play, Andrienko et al.[37] showed that the mean pressure values during a whole game were notably lower than the pressure values during successful defensive play.[37] This difference was evident for pressure on the ball as well as for pressure on the attacking players. The authors also found that the pressure is higher in areas close to the ball (0–15 m from the ball) compared to areas further away from the ball. The other four studies taking the opponent interaction of defensive pressure into account did not aim to reveal outcomes for defensive play and therefore did not present relevant results.

Group level. At a group level, seven studies analyzed defensive play. The used approaches include behavioral analyzes,[40,41] synchronization,[42,43] analysis of numerical superiority,[2] and analysis of offense including the density of players behind the ball as opponent interaction. In behavioral analyzes, Castellano and Casamichana[40] investigated the height of the defensive line (visualized in Figure 2, c) and Clemente et al.[41] used the tactical formation to define the defensive area of play. To evaluate synchronization (interaction) between or within teams, Goes et al.[43] assessed the intra-team and inter-team synchronization of subgroups (defensive, midfield, & attacking line) and Folgado et al.[42] investigated the intra-team movement synchronization at dyad level (pairings of 2 players) (visualized in Figure 2, a). Vilar et al.[2] investigated numerical superiority in dynamical sub-areas of the pitch. The other two studies explored offensive play considering the opponent interaction as the number of defensive players still able to defend an approaching attack. While one study defined this as density of defensive players[8] another referred to the term as opponent number.[39]

Six of the seven studies analyzing defensive play at a group level investigated their approaches in connection to success. Castellano and Casamichana[40] used the end result of a season. At the game level, the end result[8,41,42] was used as success criteria. At the possession level, the success of an attack was defined by the score box possession[39] and with a spatiotemporal tool determining the probability of an attack resulting in a scoring opportunity.[43] Solely Vilar et al.[2] did not consider success in their investigation.

Regarding the results of the studies, group level behavioral analysis of defensive play revealed that stronger teams positioned their defensive line further up the field compared to weaker teams.[40] Additionally, greater defensive pressure occurred in the midfield region, compared to attacking and defensive regions.[41] Looking at the synchronization at a group level, successful attacks are characterized by a decreased inter-team synchronization between the defenders of the attacking team and attackers of the defending team.[43] Furthermore, Folgado et al.[42] revealed that losing teams tend to have lower inter-team synchronization compared to winning teams and that defensive dyads (pairings of two players) are more synchronized than offensive dyads. Vilar et al.[2] found that soccer teams place more players than their opponents in areas closer to their own goal, especially in the center-back sub-areas of play. Furthermore, their analysis revealed that winning teams were able to risk more players in sub-areas closer to the opponents' goal (forward areas) while concurrently maintaining higher stability in sub-areas closer to their own goal (defensive areas) compared to the losing team. Taking the opponent interaction into account, solely one study showed results regarding defensive play. Lago-Ballesteros et al.[39] indicated that attacks are more successful when playing against less than six defending players compared to attacks against a balanced defense (six or more defenders).

Team level. Nine studies analyzed defensive play at a team level. The approaches used to analyze defensive play in this tactical level were the investigation of ball recoveries (Fernandez-Navarro et al. 2019; Santos et al. 2017; Santos & Lago-Penas, 2019), behavioral analyzes,[12,44,45,51] analyzes of synchronization,[46] and analyzes of offense with the defensive organization as a measure of opponent interaction.[15] All three studies exploring ball recoveries focused on the influence of situational variables (match status, match location, or quality of opponent) on the ball recoveries.[32,47,48] Investigations focusing on defending behavior at a team level used collective metrics of the full team to describe the movement of players on the field, e.g. team centroids, coverage area, spread.[12,44,51] Solely Welch et al.[45] used other order measures to investigate team behavior (polarization and angular momentum). Frencken et al.[46] assessed the inter-team distances between the team centroids to measure synchronization of opposing teams (visualized in Figure 2, b). Analyzing passing effectiveness Goes et al.[15] calculated the defensive disruptiveness following a pass using measures to describe the defensive organization (centroids, surface area, and spread).

Regarding the connection to success, solely Bartlett et al.[44] and Moura et al.[12] defined success at a possession level with attack outcomes (e.g. shot on goal). Additionally, four studies partially considered the relation to success, for example by investigating the influence of match status on measurements without the discussing effectiveness.[32,46–48] The remaining three studies did not consider success in their investigations (Clemente et al. 2013; Goes et al. 2019; Welch et al. 2021).

Regarding the results on defensive play, the studies on ball recoveries show that teams gain more balls closer to their own goal when playing away (vs. playing at home),

winning (vs. losing or drawing), and playing against a stronger opponent (vs. weaker opponent), and vice-versa (Fernandez-Navarro et al. 2019; Santos et al. 2017; Santos & Lago-Penas, 2019). Solely Santos and Lago-Peñas[32] could not find an effect of match location on ball recovery situations. Investigating defensive behavior, Moura et al.[12] showed that in ball possession, the coverage area and the spread increase compared to sequences of play without ball possession (concentration-expansion relationship). They also identified that in successful defensive plays, the teams showed smaller coverage areas and spreads compared to unsuccessful defensive plays.[12] Clemente et al.[51] also found a decreasing trend in dispersion and coverage measures of teams without ball possession (concentration-expansion relationship). Welch et al.[45] showed that the movement of players in the defensive phase is more ordered and compact (supporting the concentration-expansion relationship) and faster-moving compared to the other phases of play. In contrast, Bartlett et al.[44] could not find evidence supporting the mentioned concentration-expansion relationship. Furthermore, they could identify only a few critical events (e.g. shots, goals) by analyzing the interactions of team centroids. Similarly, Frencken et al.[46] reported that inter-team distances of the opposing teams are rarely related to critical match events (e.g. scoring opportunities, goals). However, those critical match periods identified by inter-team distances were associated with collective defensive actions (e.g. pressurize players of the attacking team).[46] Goes et al.[15] did not discuss results regarding defensive play.

Mixed approaches. Four studies using a combination of individual, group, and team tactical variables and were assigned to this section. The used approaches were computational models[23,33] and prediction models.[34,49] Regarding the computational models, Matsuoka et al.[33] introduced a model of defensive tactical play and Toda et al.[23] presented a model called *Valuating Defense by Estimating Probabilities* estimating the probability of a ball gain. Predicting defensive play, Le et al.[49] demonstrated a real-time prediction of defending behavior using team or league average behavior predictions (ghosting) and Stöckl et al.[34] used different predictions (xPass, xReceiver, xThreat) to measure the defending teams' ability to disrupt the oppositions attacking strategy and to detect high-level defensive concepts (e.g. ball or man orientated defense).

To measure their models in connection to success, Toda et al.[23] used the season outcome and Matsuoka et al.[33] defined success of defensive play as gain of ball possession. Stöckl et al.[34] only partially took success into account and Le et al.[49] did not consider the effects of success at all.

Concerning the findings of the included studies, Matsuoka et al.[33] identified three sub-defensive tactical types of play, which were set defense (maintaining defensive organization), concentration defense (increasing player density around the ball), and control defense (restricting or guiding the attacking play). Furthermore, they demonstrated that their model shows higher values for successful defensive play (ball gain) compared to unsuccessful (no ball gain) defensive play.[33] Toda et al.[23] found that the distance of the closest defender to the ball had the highest impact on their model and the values calculated from this had a moderate positive correlation with the season outcome, defined as winning points. Le et al.[49] focused on the characteristics of their predictions and showed that their movement predictions for the players can maintain solid defensive formation and therefore can quantify, analyze and compare defending behavior on a fine-grained level. By analyzing a single-match situation they showed that the ghosting model of a good defense reduced the expected goal value from ~70% (of the actual sequence where the analyzed team conceded a goal) to ~40%.[49] Similarly, Stöckl et al.[34] demonstrated that their predictions were able to create a compact visual representation of the teams' defensive performance. Moreover, the single-game analysis using the predictions could show how one team's defense was able to negate the offensive actions of the other team's forwards (e.g. striker's probability of receiving a pass was below average).

Discussion

The aim of this scoping review was twofold. First, we aimed to explore the approaches used to analyze defensive play in professional soccer using player tracking data. Secondly, we wanted to reveal the findings about successful defensive play using those approaches.

Following the aims, the discussion is structured as follows. First, the approaches analyzing defensive play are discussed in the different levels of tactical play (individual, group, team level) to account for the first research question. Second, the findings on successful defensive play are presented in the same structure to answer the second research question. Then, overlapping limitations of included studies are discussed and future directions are derived accordingly. Finally, practical implications of the results of this review are exemplified.

Approaches

The approaches used to analyze defensive play are widespread. The range of approaches is depicted in Table 3. The analysis of defensive pressure appears to be an informative indicator of defensive play. While the analysis of synchronization and defending behavior provide information about the defending process, the investigations on ball recoveries focused on the moment of a ball gain and did not consider the course of the defensive play. Both computational and prediction models include promising data processing approaches that can reveal hidden information

about unstructured data but do not provide practice-relevant information.

Individual level. The only approach used to analyze defensive play at the individual level was defensive pressure (visualized in Figure 2, d). However, the results of Andrienko et al.[37] indicate that this defensive pressure is a substantial part of successful defending. Also, the widespread use of defensive pressure as an opponent interaction in the analysis of offensive play indicates the awareness of the importance of defensive pressure in soccer.[7,8] Therefore, the[38,39] analysis of defensive pressure is a promising performance indicator for defensive play that should be further investigated.

Group level. The analysis of defending behavior at the group level is more informative compared to team interactions,[43,46] as the measurement of team interactions might be too simplistic for the portray of complex game situations. This is especially true to the analysis of defending behavior and synchronization. Both approaches can reveal important information about the defending process on a fine-grained level. However the behavioral approaches on a group level (height of the defensive line & defensive playing area) in particular are very simplistic and reveal very little information about defensive tactical play.

The analysis of numerical superiority, which describes the placement of players in defined areas of the pitch, is related to the approach of measuring the density of players behind the ball who are still able to defend an approaching attack. Both measures are very simple measurement techniques that help to analyze the tactics of defensive play. However, due to their simplicity, they cannot reveal detailed information about the spatial relation of players to each other, so the information gain using these approaches remains low.

Team level. The approaches to analyze defensive play at the team level were diverse. Three studies focused on the influence of situational variables on ball recoveries. As ball recoveries are one of the main goals of defensive play (next to prevent goals of opponents), they can provide important information about defensive playing tactics. However, they merely consider the moment of the ball recovery (outcome of defensive play) and not consider the whole defending process, although Fernandez-Navarro[47] used some variables related to the whole defensive play. Furthermore, all three studies focused on the influence of situational variables and not on the tactics of successful defensive play. Therefore, little information can be obtained about the tactics of the defending process.

The approaches to analyze defending behavior and synchronization at the team level, as mentioned before, are expected to provide fewer insights compared to the analysis at the group level. However, both approaches (analysis of synchronization and behavior) help to analyze the team's spatial organization on the pitch. Accordingly, Goes et al.[15] introduced the quantification of defensive organization in a sophisticated way analyzing passing effectiveness. This quantification of defensive organization, which is crucial for successful defensive play,[12] holds great potential for future research in defensive play at the team level.

Mixed approaches. Both the computational models and the prediction models used a vast amount of variables at different levels of tactical play (e.g. individual parameters, group parameters, team parameters). In both approaches, the computational data processing methods help to extract the most important information from the great amount of unstructured data. This procedure can help to identify hidden patterns in the data and provide new insights in the constraints of successful defensive play from a computational perspective. To reveal this information the studies need to report all included parameters and discuss the resulting weights of the individual parameters. This is crucial for other researchers and practitioners to retrace the data processing and to gain practice-relevant information. However, it is the weighting of individual parameters that is largely lacking, severely limiting the usefulness of these approaches. Therefore, computational models and prediction models can help to deal with the huge amount of unstructured tracking data and to reveal hidden patterns but fail to reveal practice-relevant results, which needs to be addressed in the future.

Successful defensive play

The results on successful defensive play are depicted in Table 4. Successful defensive play is characterized by high defensive pressure on the ball and on the opponents. With more players behind the ball and numerical superiority in areas closer to the own goal, the amount of spatial defensive pressure increases which is also related to successful defensive plays. Furthermore, success in defensive play is indicated by the contraction of the team's organization. Ultimately, a higher inter-team and intra-team synchronization is a good indicator for successful defensive play.

Individual level. The main findings on successful defensive play at the individual level are that successful defensive plays are characterized by high defensive pressure (on the ball & on the attackers).[37]

There is a big research potential in this area, as the mentioned studies using the opponent interaction of defensive pressure did not discuss results regarding defensive play and Andrienko et al.[37] only exemplarily showed the measurement of pressure using a small sample size (4 games). The conditions of pressure, team differences depending

on different tactics, and the how/when/where of efficient defensive pressure need to be addressed in the future.

Group level. The main findings of successful group tactical defensive play are as follows. Winning teams maintain higher defensive stability (numerical superiority) while concurrently risking more individual players to offensive sub-areas of the field compared to losing teams.[2] Furthermore, attacks are less successful when played against a balanced defense (more than five defenders)[39] and better teams position their defensive line further up the pitch.[52] Furthermore, intra-team as well as inter-team synchronization of defensive players is an important factor of successful defensive play.[42,43]

As teams place more players than their opponents in sub-areas closer to their own goal,[2] the spatial and temporal pressure on the opponents in those areas rises due to this numerical superiority. Therefore, those findings describe the tendency of teams applying higher defensive pressure on opponents when they are closer to their goal and indicate the importance of defensive pressure.

The findings of Lago-Ballesteros et al.[39] that attacks are less successful when played against more than five defenders are supported by the work of Castellano et al.,[52] who identified the fewest offensive players when a team won compared to a draw or loss. This higher number of defending players is also considered in the study of Link et al.[8] taking the density of defenders in front of the ball into account. However, Link et al.[8] revealed no results regarding successful defensive play.

The finding of higher placement of the defensive line in better teams compared to weaker teams indicates a high pressing playing style applied by better teams. This defensive tactical playing style, where teams try to quickly regain the ball in areas closer to the opponents' goal to control the game or to achieve scoring opportunities is therefore associated with success.

Regarding the synchronization during defensive play winning teams show more intra-team synchronization on dyad level (2 players) compared to losing teams[42] and defenders on the attacking team and attackers on the defending team show a decreased inter-team synchronization in successful attacks (Goes et al. 2021).

Overall, the positioning of defensive players and their synchronization is important for the success of defensive play.

Team level. At a team tactical level, the main finding is that in successful defensive plays teams present a stronger contraction compared to unsuccessful defensive plays.[12]

Three studies showed this contraction of players in defensive playing phases (players moving closer together) compared to an expansion in offensive playing phases (players diverge).[12,45,51] This indicates the defensive principle to close gaps between defenders, thus narrowing the possible space for attackers in order to prevent the opponent from creating scoring opportunities. This contraction-expansion relationship depending on ball possession was not found in one study.[44] However, in this context, solely one study focused on the connection to success and revealed that the contraction is stronger in successful defensive play.[12] The measurement of defensive organization on the pitch in the context of success needs to be addressed in the future as the contraction of the organization during defensive play is indicated to be a highly important performance indicator for defensive success. The presented approaches show possibilities for how team tactical variables can be used to quantify this defensive organization.

Team level analysis of ball recoveries revealed that defensive play is affected by situational variables but provides little information about successful defensive play. Teams recover more balls closer to their own goal when winning (vs. losing or drawing), playing away (vs. playing at home), and playing against a stronger opponent (vs. weaker opponents).[32,47,48] This highlights that teams adapt their defending playing style to contextual factors. For example, winning teams tend to defend in areas closer to their own goal and focus on quick counter-attacking after ball gains in order to save the score.

Another finding is that the inter-team synchronization (short distances between team centroids) is not clearly related to critical match events (e.g. shots, goals).[44,46] Rather, it is evident that those moments of exceptional interactions indicate high defensive pressure.[46]

Mixed approaches. The findings on successful defensive play of the mixed approaches were sparse. Both the computational models and the prediction models focused on the model characteristics and the results regarding their model. For example, Matsuoka et al.[33] showed that their model indicates higher values in successful defensive play (ball gain) compared to unsuccessful defensive plays (without ball gain). Solely Toda et al.[23] reported a resulting weighting of parameters and found that the distance of the closest defender to the ball had the highest impact on their model. This distance can be interpreted as defensive pressure on the ball and once again shows the importance of this defensive performance indicator. This example points out that the weighting of individual parameters can contribute to the understanding of the models and can provide insights into the measurement of defensive play to increase their practical value.

Limitations and future directions

The majority of the studies included in this review used insufficiently small sample sizes (fewer than 30 games). The highly dynamic team sport soccer leads to a great variability of performance indicators and in this context smaller sample sizes lead to a diminished overall generality of results.[53] Furthermore, almost half of the studies did not investigate their approaches in connection to success. This is important to show the actual significance

and importance of the developed performance indicators or processing approaches. Both limitations lead to a diminished general validity of the results. Those studies taking the connection to success into account defined success at different levels (season, game, possession level). Thereby, the analysis of success at a possession level can reveal the most fine-grained insights into tactical play because it is subject to the slightest effects of chance compared to match or season level. Depending on the study design the definition of success at a game level can be informative as well. The analysis of success at a season level can only provide rough insights into tactical play as the conditions of an attacking or defending sequence cannot be viewed. Therefore, the advantages of the different levels of success analyzed in the different studies should be considered in future investigations.

Accordingly, it is important to investigate the whole defending process and not just a snapshot of a playing situation, e.g. a ball recovery situation. To increase the practical relevance, it is important to find performance-determining factors of defensive play. Yet, most studies lack the practice-oriented discussion of results, so that the practical relevance of most publications remains low.

However, with the possibilities of the evaluation of tracking data, these issues can be counteracted. The availability of large amounts of data from competitive matches at the professional level (e.g. German Bundesliga, British Premier League) and the possibility of time-saving and fine-grained analysis compared to traditional (e.g. notational) analysis methods indicates an interesting potential for future game analysis. In this context, it should be noted that most of the studies included in this review used tracking data from tracking companies. This shows that researchers in soccer using tracking data rely on tracking companies to measure and to provide the tracking data.

Overall, the approaches and study designs shown in this review are very heterogeneous and build various starting points for future research. These future investigations should use sophisticated data processing approaches and focus on practice-orientated discussion and usability of their results. Future research in defensive play using tracking data should focus on the following:

1. The conditions of defensive pressure. Team differences depending on different tactics and the how/when/where of efficient defensive pressure.
2. Quantification of the compact organization during defensive play in connection to success.
3. The practice-oriented evaluation and discussion of results to increase the practical impact of introduced approaches evaluating tracking data.

4.4 Practical implications

The usage of new analyzing approaches in soccer, such as the evaluation of tracking data, has the potential to provide new insights into the performance structure of defensive play. Yet, there is a widespread lack of implementation in practice of this new analytic approach about performance in soccer. Therefore, the next section shows examples of how the gained information of this review can impact practice.

The results of this review on successful defensive play referring to the second research question are summarized in Table 4. Those findings can help practitioners (e.g. coaches, analysts, managers) in designing match tactics, planning training content, or analyzing opponents. In this context, the connection of the used approaches of the different studies to success (e.g. end result of a game) is important. Solely studies that discuss their results in connection to success can reveal findings of successful tactics in soccer and draw conclusions of the success factors in soccer. This is highly important for the practical impact of the results in this research area. Subsequently, some examples how to use the results of this review are presented in more detail.

The purpose of defensive play is to protect the own goal and recover the ball.[33] This ball recovery is impacted by situational variables (match status, match location, quality of opponent). Therefore, practitioners should bear these situational variables in mind when analyzing match performance.[47]

Successful defensive play is characterized by applying high pressure on attackers[37] and attaining a coordinated compact team organization. This contraction of organization in defensive playing phases[12,32,42,45,48] is important to minimize the space for the attacking team to advance their attack and thereby enable the defenders to quickly put pressure on the attackers after a ball move. Furthermore, the intra-team synchronization is important for defensive success.[42,43] Therefore, these tactics should be considered by coaches, experts, and analysts.

Findings like these can impact the game of soccer if discussed in a practical manner and implemented by practitioners.

Conclusion

Overall, the approaches used to analyze defensive play were highly heterogeneous ranging from the analysis of ball recoveries to computational models. Most promising approaches to analyze defensive play were the analysis of defensive pressure and the quantification of defensive compact organization. The findings on successful defensive play using those approaches were that successful defensive play at the individual level is characterized by high defensive pressure. At the group level, inter-team and intra-team synchronization as well as a balanced defense are important

for successful defensive play. At the team level, the contraction of the spatial organization is important for success.

In general, defensive play is highly important for success in soccer.[26,27] However, so far it has received less attention compared to offensive play. This review highlights the importance of defensive play and builds a basis for future research in this area. Furthermore, it presents important aspects that should be taken into account when conducting research in this area (e.g. sample size, connection to success, practical discussion of results). In conclusion, there are various starting points for future research on the analysis of defensive play in soccer using the possibilities of tracking data to improve the understanding of defensive play.

Authors' contribution

- Conceptualization: LeaF, SA, MK
- Investigation: LeaF, LeoF
- Methodology: LeaF, SA, MK
- Supervision: SA, DJ, MK
- Validation: LeaF
- Writing: LeaF, SA, LeoF, DJ, MK

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

ORCID iD

Leander Forcher https:orcid.org/0000-0002-6428-8643

Supplemental material

Supplemental material for this article is available online.

References

1. Rein R and Memmert D. Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science. *SpringerPlus* Epub ahead of print 1 December 2016; 1: 1–16. DOI: 10.1186/s40064-016-3108-2
2. Vilar L, Araujo D, Davids K, et al. Science of winning soccer: emergent pattern-forming dynamics in association football. *Journal of Systems Science & Complexity* 2013; 26: 73–84.
3. Herold M, Goes F, Nopp S, et al. Machine learning in men's professional football: current applications and future directions for improving attacking play. *Int J Sports Sci Coach* 2019; 14: 798–817.
4. Memmert D and Rein R. Match analysis, Big data and tactics: current trends in elite soccer. *Deutsche Zeitschrift für Sportmedizin* 2018; 69: 65–72.
5. Goes FR, Meerhoff LA, Bueno MJO, et al. Unlocking the potential of big data to support tactical performance analysis in professional soccer: a systematic review. *Eur J Sport Sci* 2021; 21: 481–496.
6. Gudmundsson J and Horton M. Spatio-Temporal analysis of team sports. *ACM Comput Surv* 2018; 50: 1–34.
7. Szczepanski L and McHale I. Beyond completion rate: evaluating the passing ability of footballers. *Journal of the Royal Statistical Society Series A-Statistics in Society* 2016; 179: 513–533.
8. Link D, Lang S and Seidenschwarz P. Real time quantification of dangerousity in football using spatiotemporal tracking data. *PLoS One* 2016; 11: 1–16.
9. Grehaigne J-F, Bouthier D and David B. Dynamic-system analysis of opponent relationships in collective actions in soccer. *J Sports Sci* 1997; 15: 137–149.
10. Tenga A, Holme I, Ronglan LT, et al. Effect of playing tactics on goal scoring in Norwegian professional soccer. *J Sports Sci* 2010; 28: 237–244.
11. Carling C, Bloomfield J, Nelson L, et al. The role of motion analysis in elite soccer: contemporary performance measurements techniques and work rate data. *Sports Med* 2008; 38: 839–862.
12. Moura FA, Martins LEB, Anido RdO, et al. Quantitative analysis of Brazilian football players' organisation on the pitch. *Sports Biomech* 2012; 11: 85–96.
13. Fernandes T, Camerino O, Garganta J, et al. How do elite soccer teams perform to ball recovery? Effects of tactical modelling and contextual variables on the defensive patterns of play. *J Hum Kinet* 2020; 73: 165–179.
14. Hewitt A, Greenham G and Norton K. Game style in soccer: what is it and can we quantify it? *International Journal of Performance Analysis in Sport* 2016; 16: 355–372.
15. Goes FR, Kempe M, Meerhoff LA, et al. Not every pass Can Be an assist: a data-driven model to measure pass effectiveness in professional soccer matches. *Big Data* 2019; 7: 57–70.
16. Gudmundsson J and Wolle T. Football analysis using spatiotemporal tools. In: Proceedings of the 20th International Conference on Advances in Geographic Information Systems, Redondo Beach, California, pp.566–569.
17. Rein R, Raabe D and Memmert D. Which pass is better?" novel approaches to assess passing effectiveness in elite soccer. *Hum Mov Sci* 2017; 55: 172–181.
18. Spearman W, Basye A, Dick G, et al. Physics-Based Modeling of Pass Probabilities in Soccer. In: MIT SLOAN Sports Analytics Conference, 2017. Hynes Convention Center.
19. Chawla S, Estephan J, Gudmundsson J, et al. Classification of passes in football matches using spatiotemporal data. *ACM Trans Spatial Algorithms Syst* 2017; 3: 1–30.
20. Goes F, Kempe M and Lemmink KAPM. Predicting match outcome in professional Dutch football using tactical performance metrics computed from position tracking data. In: Conference: MathSport International, Athens, Greece, 2019.
21. Taki T and Hasegawa J. Visualization of dominant region in team games and its application to teamwork analysis. In: Proceedings Computer Graphics International 2000, Geneva, Switzerland, 2000, pp.227–235.

22. Moura FA, Martins LEB, Anido RO, et al. A spectral analysis of team dynamics and tactics in Brazilian football. *J Sports Sci* 2013; 31: 1568–1577.
23. Toda K, Teranishi M, Kushiro K, et al. Evaluation of soccer team defense based on prediction models of ball recovery and being attacked. *Computer Science ArXiv* 2021: 1–12.
24. Maneiro R, Casal CA, Álvarez I, et al. Offensive transitions in high-performance football: differences between UEFA euro 2008 and UEFA euro 2016. *Front Psychol* 2019; 10: 1230–1240.
25. Winter C, Rasche C and Pfeiffer M. Linear vs. non-linear classification of winners, drawers and losers at FIFA world Cup 2014. *Science and Medicine in Football* 2017; 1: 164–170.
26. Lepschy H, Woll A and Wäsche H. Success factors in the FIFA 2018 world Cup in russia and FIFA 2014 world Cup in Brazil. *Front Psychol* 2021; 12: 1–9.
27. Georgievski B, Labadze L and Aboelsoud ME. Comparative advantage as a success factor in football clubs: evidence from the English premier league (EPL). *Journal of Human Sport & Exercise* 2019; 14: 292–314.
28. Bosca JE, Liern V, Martinez A, et al. Increasing offensive or defensive efficiency? An analysis of Italian and spanish football. *Omega-International Journal of Management Science* 2009; 37: 63–78.
29. Gavião LO, Sant'Anna AP, Alves Lima GB, et al. Evaluation of soccer players under the moneyball concept. *J Sports Sci* 2020; 38: 1221–1247.
30. Hirotsu N and Wright M. An evaluation of characteristics of teams in association football by using a markov process model. *Journal of the Royal Statistical Society Series D-The Statistician* 2003; 52: 591–602.
31. Lepschy H, Waesche H and Woll A. Success factors in football: an analysis of the German bundesliga. *International Journal of Performance Analysis in Sport* 2020; 20: 150–164.
32. Santos PM and Lago-Penas C. Defensive positioning on the pitch in relation with situational VAriables of a prodessional football team during regaining possession. *Human Movement* 2019; 20: 50–56.
33. Matsuoka H, Tahara Y, Ando K, et al. Development of criterion-referenced measurement items for soccer defensive tactical play from tracking data. *Football Science* 2020; 17: 29–40.
34. Stöckl M, Seidl T, Marley D, et al. Making Offensive Play Predictable -Using a Graph Convolutional Network to Understand Defensive Performance in Soccer. In: MIT SLOAN Sports Analytics Conference, Virtual, 2021, pp.1–19.
35. Tricco A, Lillie E, Zarin W, et al. PRISMA Extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018; 169: 467–473.
36. Rico-González M, Pino-Ortega J, Clemente F, et al. Guidelines for performing systematic reviews in sports science. *Biol Sport* 2021; 39(2): 463–471. DOI: 10.5114/biolsport.2022.106386
37. Andrienko G, Andrienko N, Budziak G, et al. Visual analysis of pressure in football. *Data Min Knowl Discov* 2017; 31: 1793–1839.
38. Goes F, Schwarz E, Elferink-Gemser M, et al. A risk-reward assessment of passing decisions: comparison between positional roles using tracking data from professional men's soccer. *Science and Medicine in Football* 2021: 1–9.
39. Lago-Ballesteros J, Lago-Peñas C and Rey E. The effect of playing tactics and situational variables on achieving score-box possessions in a professional soccer team. *J Sports Sci* 2012; 30: 1455–1461.
40. Castellano J and Casamichana D. What are the differences between first and second divisions of spanish football teams? *International Journal of Performance Analysis in Sport* 2015; 15: 135–146.
41. Clemente FM, Martins FML, Couceiro MS, et al. Developing a tactical metric to estimate the defensive area of soccer teams: the defensive play area. *Journal of Sports Engineering and Technology* 2016; 230: 124–132.
42. Folgado H, Duarte R, Marques P, et al. Exploring how movement synchronization is related to match outcome in elite professional football. *Science and Medicine in Football* 2018; 2: 101–107.
43. Goes FR, Brink MS, Elferink-Gemser MT, et al. The tactics of successful attacks in professional association football: large-scale spatiotemporal analysis of dynamic subgroups using position tracking data. *J Sports Sci* 2021; 39: 523–532.
44. Bartlett R, Button C, Robins M, et al. Analysing team coordination patterns from player movement trajectories in soccer: methodological considerations. *International Journal of Performance Analysis in Sport* 2012; 12: 398–424.
45. Welch M, Schaerf TM and Murphy A. Collective states and their transitions in football. *PLoS One* 2021; 16: 1–20.
46. Frencken W, Poel Hd, Visscher C, et al. Variability of inter-team distances associated with match events in elite-standard soccer. *J Sports Sci* 2012; 30: 1207–1213.
47. Fernandez-Navarro J, Ruiz-Ruiz C, Zubillaga A, et al. Tactical variables related to gaining the ball in advanced zones of the soccer pitch: analysis of differences Among elite teams and the effect of contextual variables. *Front Psychol* 2019; 10: 1–14.
48. Santos P, Lago-Penas C and Garcia-Garcia O. The influence of situational variables on defensive positioning in professional soccer. *International Journal of Performance Analysis in Sport* 2017; 17: 212–219.
49. Le HM, Carr P, Yue Y, et al. Data-Driven Ghosting using Deep Imitation Learning. In: MIT SLOAN Sports Analytics Conference, 2017, pp.1–15: Hynes Convention Center.
50. Faber I, Bustin P, Oosterveld F, et al. Assessing personal talent determinants in young racquet sport players: a systematic review. *J Sports Sci* 2015; 34: 1–16.
51. Clemente FM, Couceiro MS, Martins FML, et al. Measuring tactical behaviour using technological metrics: case study of a football game. *Int J Sports Sci Coach* 2013; 8: 723–739.
52. Castellano J, Casamichana D and Lago C. The Use of match statistics that discriminate between successful and unsuccessful soccer teams. *J Hum Kinet* 2012; 31: 139–147.
53. Sarmento H, Marcelino R, Campanico J, et al. Match analysis in football: a systematic review. *J Sports Sci* 2014; 32: 1831–1843.