Pressure Cooker

FC Barcelona Defensive Analysis — More than a Hack 2026

We built a system that pinpoints when and how Barcelona's defence breaks down. It processes Metrica Sports Smart Tagging data and tracking feeds from 11 matches, converts events into a continuous risk score, flags danger moments, and uses an LLM to explain what went wrong in plain tactical language.

> **Tactical Question:** *When and how does Barcelona's defensive structure break down, and which vulnerability patterns recur across matches?*

Scope: Scope 2 (Game modeling and pattern detection) with elements of Scope 1 (Interpretation) and multi-match comparative analysis.

Contents

1. Data Sources and Limitations

1.1 Available Data Per Match

Source	Format	Content	Coverage
Smart Tagging	*_pattern.xml	Team-level tactical phases	Full match, both teams
ATD Metadata	*_FifaData.xml	Team names, player IDs, pitch dims	Complete metadata
ATD Positions	*_FifaDataRawData.txt	Per-frame x/y/speed	Ball: good. Outfield: partial
Parsed Tracking	player_positions.csv + ball_positions.csv	Per-match player/ball positions with team ID mapping	All 11 matches (custom parser)
Broadcast Video	*.mp4	Full match recording	All 11 matches

1.2 Critical Limitation: Tracking Sparsity

The ATD tracking feed comes from broadcast camera footage. Because the camera follows the ball, players frequently go off-screen, producing NaN values. The ATD also has identity fragmentation (40–60 extra IDs per match). In practice this means:

- Individual player metrics (progressive carries, press resistance) are **not feasible**
- Defensive shape, compactness, or line height **cannot be reliably computed** from tracking alone
- Ball position is usable and provides limited spatial context
- The risk scoring pipeline is built on **team-level event phase data**

Tracking data recovery: We built a custom batch parser (tracking_batch_parser.py) that extracts usable ball positions and partial player positions from the raw ATD feed across all 11 matches. The parsed data is stored as player_positions.csv and ball_positions.csv per match, with a team_map.json file mapping track IDs to Barcelona vs. opponent.

1.3 Smart Tagging Event Codes and Weights

Opponent events increase risk (positive weights). Barcelona possession events reduce risk (negative weights). Events not listed carry zero weight.

Event Code	Description	Weight
BALL IN THE BOX	Ball enters penalty area	+1.55
DEFENSIVE TRANSITION	Losing possession, recovering	+1.35
COUNTER ATTACK	Counter-attack initiated	+1.35
FAST BREAK	Fast break opportunity	+1.30
ATTACKING TRANSITION	Transition to attack (opponent)	+1.25
PLAYERS IN THE BOX	Attackers in penalty area	+1.15
BALL IN FINAL THIRD	Ball in attacking third	+1.10
PLAYERS IN FINAL THIRD	Players in final third	+1.05
SET PIECES	Corner, free kick, throw-in	+0.70
POSSESSION (Barca)	Barcelona in possession	−0.35
SUSTAINED ATTACK (Barca)	Barcelona sustained attack	−0.30
BUILD UP (Barca)	Barcelona building from back	−0.25
PROGRESSION (Barca)	Barcelona advancing ball	−0.20
FINAL THIRD (Barca)	Barcelona in final third	−0.15

Smart Tagging is not manually ground-truthed. It may contain false positives and negatives.

1.4 Project Pivot

The original plan was individual player performance analysis. Tracking identity fragmentation and sparse position data made that impractical, so we pivoted to team-wide defensive analysis — a question that the Smart Tagging data can answer cleanly.

2. Methodology: Risk Scoring Pipeline

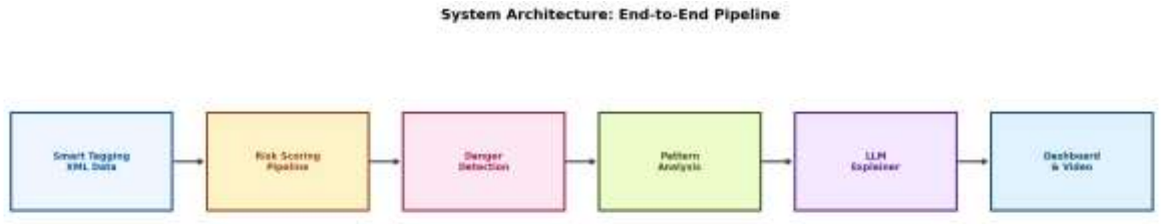

Figure 1: System Architecture — End-to-End Pipeline

The pipeline converts raw Smart Tagging annotations into a continuous risk score (0–100) for each match. The flow is: Smart Tagging XML → Risk Scoring → Danger Detection → Pattern Analysis → LLM Explanation → Dashboard with Video Seek.

2.1 Time Grid and Raw Score

The time grid is built at 0.25-second intervals (4 Hz) from the first to the last event timestamp. For each grid point, all overlapping events are identified and their weights summed. Opponent attacking events add positive risk; Barcelona possession events subtract risk. When multiple events overlap, their weights compound — capturing the combined danger of simultaneous attacking pressure and defensive disorganisation.

2.2 Smoothing and Normalization

Raw scores are smoothed with a 3-second moving average (12 samples at 4 Hz) to reduce noise from single-frame event boundaries. The smoothed signal is then min-max scaled to 0–100: the minimum is shifted to zero and the maximum maps to 100. This means each match is self-calibrated — the highest-pressure moment always peaks at 100, making it easy to compare across matches.

2.3 Design Rationale

The weight system is intentionally transparent. Each weight reflects how close the tactical phase is to a goal-scoring opportunity: events near the box (BALL IN THE BOX at +1.55) score higher than upstream events (SET PIECES at +0.70). Barcelona possession events carry negative weights because they reduce defensive risk. The resulting scores track closely with what a coach watching the match would consider dangerous.

2.4 Spatial Feature Extraction from Tracking Data

For each danger moment, the system extracts a spatial snapshot from the parsed tracking data (tracking_features.py). At the peak timestamp of each danger window, it computes:

- **Team shape metrics:** Width, length, and centroid position for both Barcelona (defending) and the opponent (attacking), based on the 5th–95th percentile spread of player positions.
- **Ball proximity:** Distance from the nearest defender and nearest attacker to the median ball position within the window.
- **Ball-side overload detection:** Within a configurable radius around the ball, the system counts defenders vs. attackers to flag numerical disadvantages.
- **Coverage diagnostics:** Each snapshot includes a tracking_coverage_warning flag when fewer than 6 players per team are tracked, so the LLM does not over-interpret sparse data.

These spatial features are passed as a structured JSON evidence pack alongside the event codes to the LLM, enabling spatially-grounded tactical explanations. The LLM system prompt includes interpretation rules (e.g. distance thresholds: ≤0.05 = very tight, ≤0.12 = close, ≤0.25 = moderate, >0.25 = far) to keep claims grounded in the data.

3. Danger Moment Detection

Danger moments are detected by finding continuous segments where the risk score exceeds a threshold, then merging nearby segments into sustained spells.

3.1 Detection Parameters

Parameter	Value	Purpose
Threshold	45.0	Risk score must exceed this to start a danger window
Min duration	5.0 sec	Windows shorter than this are discarded as noise
Min gap	12.0 sec	Windows closer than this are merged into one spell

These values were tuned iteratively across all 11 matches, balancing sensitivity (catching real threats) against specificity (not flagging routine possession changes). A threshold of 45 roughly corresponds to the point where multiple opponent attacking events overlap with Barcelona defensive events.

3.2 Severity Classification

Severity	Score Range	Meaning
High	80–100	Major danger — goal conceded or clear-cut chance
Moderate	50–79	Significant pressure, multiple threatening events
Low	25–49	Elevated risk, territorial opponent advantage
Very Low	< 25	Minor pressure, mostly controlled

3.3 Merge Logic

When two above-threshold segments are separated by less than 12 seconds, they are merged into a single sustained pressure spell. The merged moment keeps the highest peak score, the widest time window, and the union of all active event codes.

4. Match-by-Match Findings

4.1 Overview

The pipeline processed 11 FC Barcelona matches, detecting **144 danger moments** in total.

Statistic	Value
Total matches analyzed	11
Total goals: Barca scored	32
Total goals: Opponents scored	17
Total danger moments detected	144
Critical severity	21 (14.6%)
High severity	22 (15.3%)
Moderate severity	101 (70.1%)
Goal-anchored moments	17/17 (100% coverage)
Average danger moments per match	13.1

4.2 Match-by-Match Breakdown

#	Opponent	Score	Dangers	Crit	High	Mod	Goal-Anch	Avg Risk
1	AC Milan (A)	1-0	15	0	2	13	0	10.7
2	Arsenal (A)	3-5	22	6	1	15	5	12.6
3	AC Milan (H)	2-2	10	2	1	7	2	6.5
4	AS Monaco (H)	0-3	12	5	3	4	3	9.4
5	Como (H)	5-0	13	0	2	11	0	8.3
6	Man City (H)	2-2	20	3	5	12	2	14.5
7	Real Madrid (H)	3-0	21	0	3	18	0	13.3
8	Daegu FC (A)	5-0	3	0	1	2	0	3.0
9	FC Seoul (A)	7-3	10	3	1	6	3	6.4
10	Real Madrid (A)	2-1	15	1	2	12	1	9.3
11	Vissel Kobe (A)	3-1	3	1	1	1	1	4.8

(H) = Home, (A) = Away

4.3 Key Observations

High-danger matches: Arsenal (5-3) and Manchester City (2-2) produced the most danger moments (22 and 20). Both were open, transitional games against top-tier opponents.

Monaco (0-3) — worst defensive performance: Only 12 danger moments, but 5 were critical and all 3 goals were captured. The high critical-to-total ratio (42%) points to concentrated, lethal attacks rather than sustained pressure.

Dominant wins still produce risk: Even in the 5-0 wins (Como, Daegu), the system picked up moderate danger moments — every team faces occasional counter-attacks, even when dominating.

Real Madrid (3-0 win) — high moderate count: 21 danger moments with 0 critical suggests Barca controlled the outcome but faced steady mid-level pressure.

4.4 Visualizations

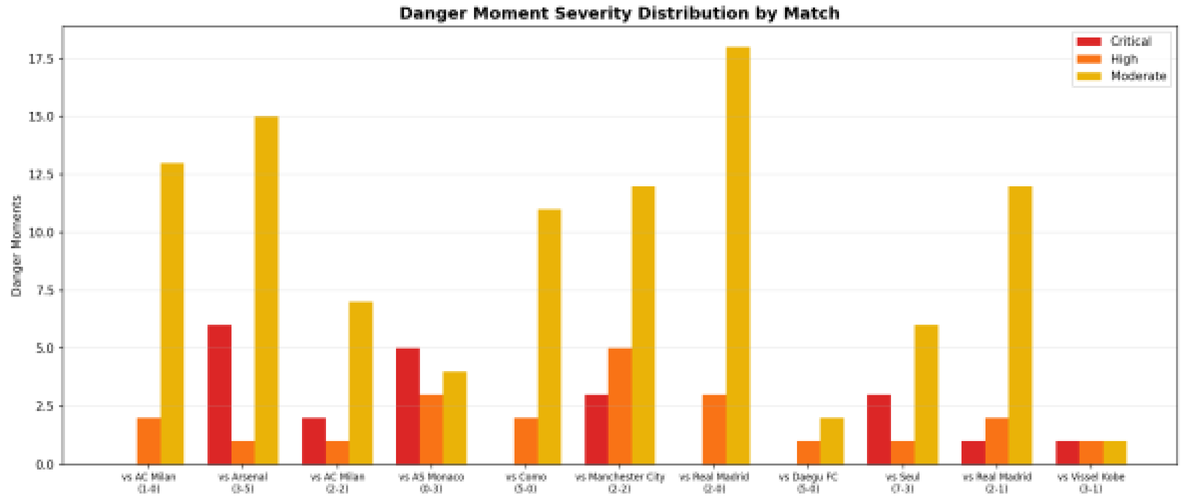

Figure 2: Danger Moment Severity Distribution by Match

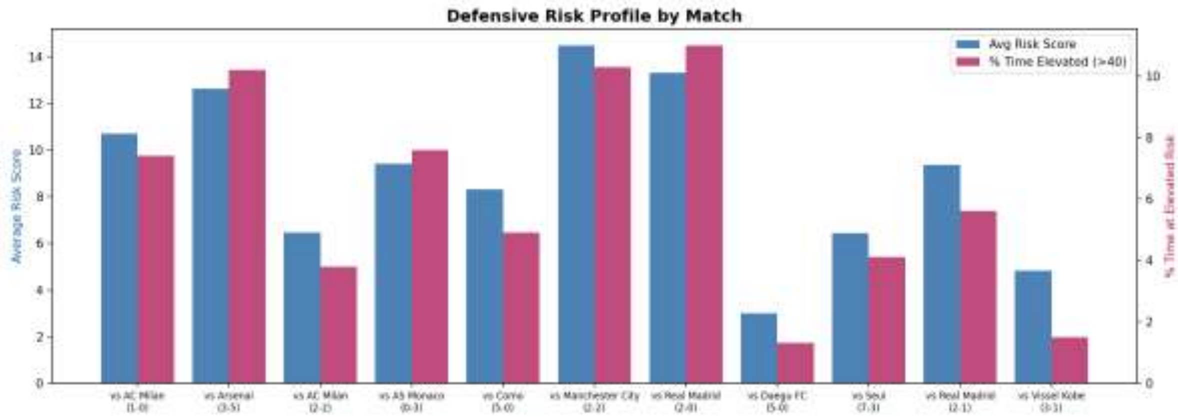

Figure 3: Defensive Risk Profile by Match — Average Risk and % Time at Elevated Risk

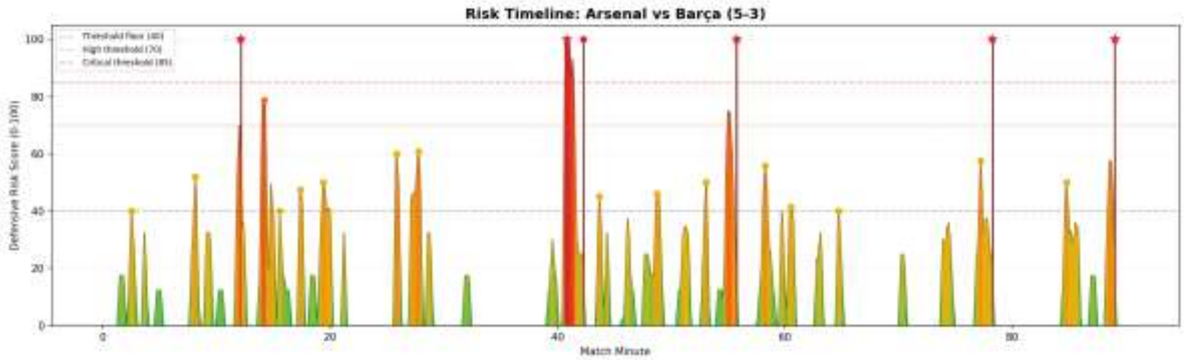

Figure 4: Risk Timeline — Arsenal vs Barca (5-3). Stars = goal-anchored moments.

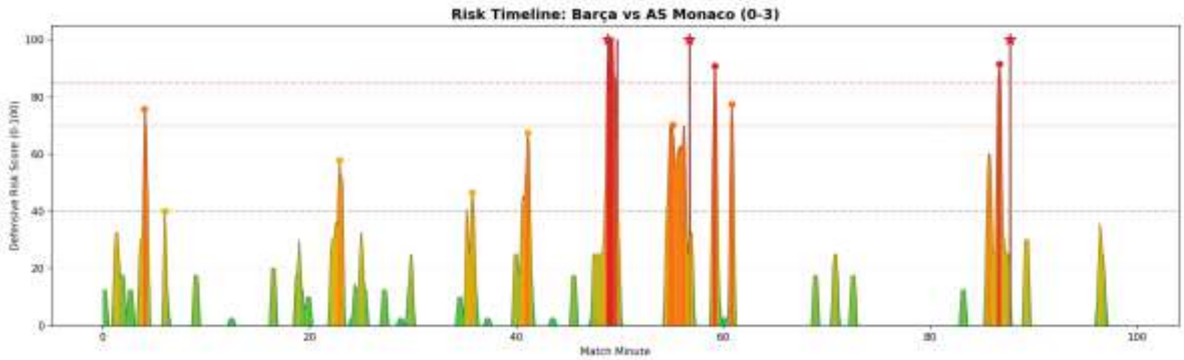

Figure 5: Risk Timeline — Barca vs AS Monaco (0-3). All 3 goals in the second half.

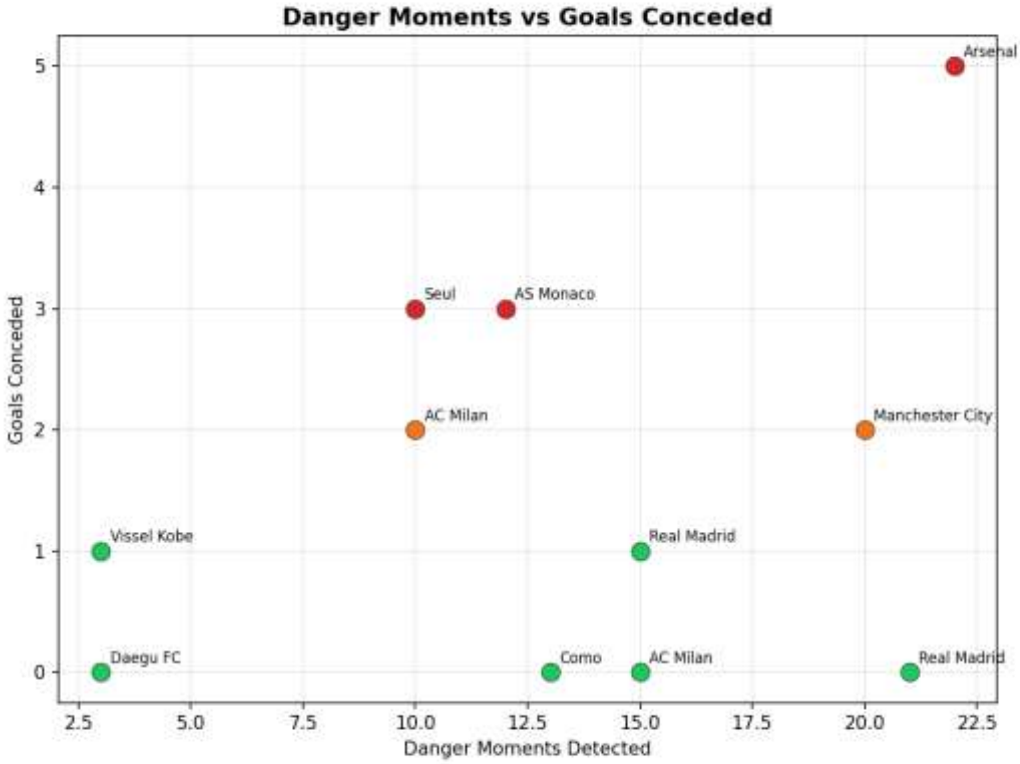

Figure 6: Danger Moments vs Goals Conceded — correlation between system output and results.

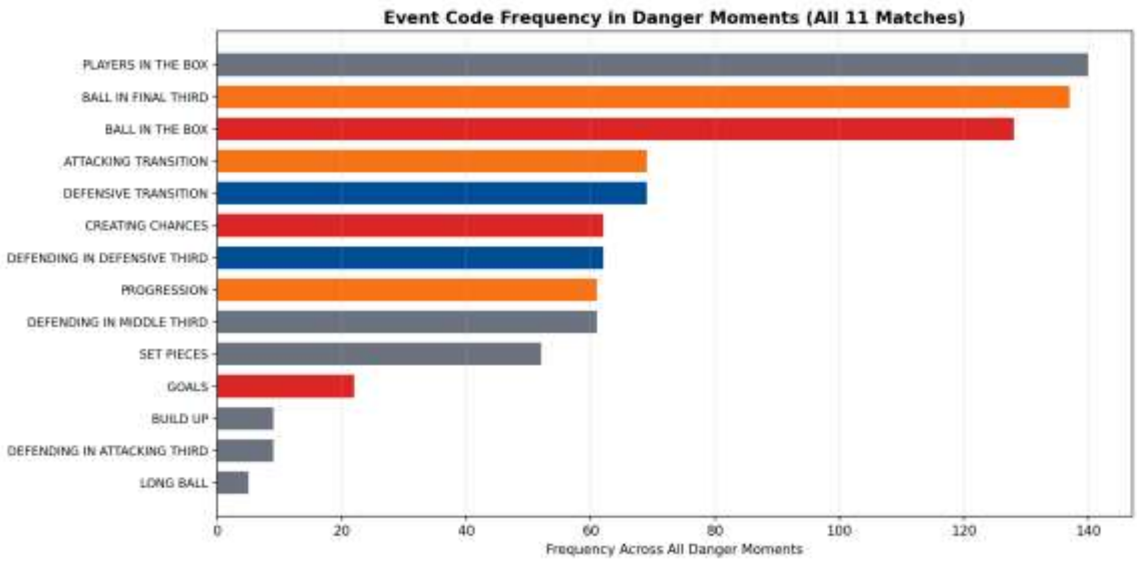

Figure 7: Event Code Frequency in Danger Moments (All 11 Matches)

4.5 Event Code Frequency in Danger Moments

Event Code	Frequency	% of Danger Moments
PLAYERS IN THE BOX	140	97.2%
BALL IN FINAL THIRD	137	95.1%
BALL IN THE BOX	128	88.9%
ATTACKING TRANSITION	69	47.9%
DEFENSIVE TRANSITION	69	47.9%
CREATING CHANCES	62	43.1%
DEFENDING IN DEF. THIRD	62	43.1%
PROGRESSION	61	42.4%
DEFENDING IN MID. THIRD	61	42.4%
SET PIECES	52	36.1%

PLAYERS IN THE BOX and BALL IN FINAL THIRD show up in almost every danger moment, which validates the weighting approach. The 48% co-occurrence of transitions is the most interesting finding — it points to transition play as a primary vulnerability vector.

4.6 Half-by-Half Vulnerability Distribution

Period	Danger Moments	Critical	Goals Conceded
1st Half	76 (52.8%)	10	8
2nd Half	68 (47.2%)	11	9

Worth noting: Monaco went from 0 critical in H1 to 5 critical in H2 (all 3 goals conceded). Whatever Monaco changed at half-time worked.

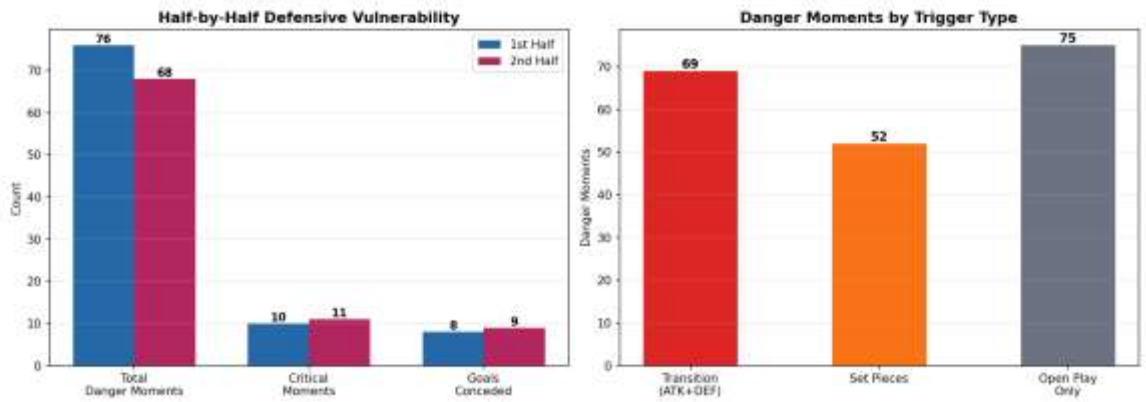

Figure 8: Half-by-Half Vulnerability and Danger Trigger Type Distribution

4.7 Transition Play: The Primary Vulnerability Vector

Transition moments appear in **47.9% of all danger moments** (69/144) and account for **29.4% of goals conceded** (5/17). Set pieces show up in **36.1% of danger moments** (52/144) and **23.5% of goals** (4/17).

4.8 Deep Dive: Most Dangerous Matches

Arsenal (5-3): 22 moments, 6 critical. Danger spread across both halves (12 in H1, 10 in H2) — sustained vulnerability throughout, not a single-half collapse.

AS Monaco (0-3): 42% critical rate. Everything escalated in the second half (0 critical H1, 5 critical H2). Monaco's half-time adjustment was clearly decisive.

Manchester City (2-2): 20 danger moments with the highest average risk in the dataset (14.5). First half was especially intense (16 of 20 dangers).

5. Cross-Match Pattern Analysis

5.1 How It Works

Each danger moment has a **signature**: the sorted set of active event codes at the peak timestamp. These signatures are compared across matches. A pattern must appear in at least 2 different matches to count. We compute prevalence in goal-producing danger moments vs. all danger moments and derive a lift score to see which combinations are disproportionately associated with goals conceded.

5.2 Confidence Scoring

Each pattern gets a composite confidence score (0–100) based on two factors: how many matches it appears in (prevalence) and how much more common it is in goal moments than in all danger moments (lift). The scoring is deliberately conservative — with only 11 matches, we want to avoid overstating patterns that might be coincidental.

Tier	Score	Coaching Guidance
High	≥ 70	Recurring vulnerability. Address in tactical sessions.
Medium	45–69	Notable pattern. Monitor in upcoming matches.
Low	< 45	Candidate theme. Not enough evidence to act on yet.

5.3 Detected Patterns

The pattern analyzer found **3 recurring vulnerability patterns** across the 11 matches:

Pattern 1: ATTACKING TRANSITION → DEFENSIVE TRANSITION

Confidence: 0.601 (Medium) | **Lift:** 2.66× | **Occurrences:** 10 (3 goals) | **Matches:** Arsenal, AC Milan (H), FC Seoul

Barcelona loses the ball during an attacking move and the opponent immediately launches a counter-attack. Players are caught upfield and the defensive shape is not set. The 2.66× lift means this sequence is nearly 3 times more likely to end in a goal than a random danger moment.

Pattern 2: CREATING CHANCES → DEFENDING IN DEFENSIVE THIRD

Confidence: 0.471 (Medium) | **Lift:** 1.97× | **Occurrences:** 9 (2 goals) | **Matches:** Arsenal, AS Monaco

The opponent creates a chance and Barcelona drops straight into deep defending — the mid-block gets bypassed entirely. The opponent goes from chance creation to box-area threats without an intermediate defensive phase, pointing to a structural gap between the lines.

Pattern 3: PROGRESSION → DEFENDING IN MIDDLE THIRD

Confidence: 0.463 (Medium) | **Lift:** 1.48× | **Occurrences:** 12 (2 goals) | **Matches:** AS Monaco, FC Seoul

The opponent progresses through midfield while Barcelona defends in the middle third. Both heavy-defeat matches feature this pattern, suggesting that when the midfield press fails, the defensive line gets exposed.

5.4 Baseline Statistics

Metric	Value
Total danger moments with valid signatures	144
Baseline goal rate (any danger moment → goal)	11.8%
Pattern 1 goal rate	30.0% (2.66× baseline)
Pattern 2 goal rate	22.2% (1.97× baseline)
Pattern 3 goal rate	16.7% (1.48× baseline)

6. LLM Integration

The system uses GPT-4o-mini (via OpenRouter) to turn structured evidence packs into tactical explanations. The LLM never sees raw data — only curated context: event codes, risk scores, severity, and (where available) spatial tracking summaries with team shape, ball proximity, and overload metrics.

6.1 Prompt Architecture

Template	Input	Output
Moment prompt	Active codes, risk score, severity, goal flag, tracking summary	3–5 sentence tactical explanation with [CODE] refs
Window prompt	Events in 5-min window, avg risk, team breakdown	Tactical summary of passage of play
Pattern prompt	Sequence, frequency, confidence stats	Structural vulnerability explanation + recommendations

6.2 Example LLM Outputs

Example 1 — High severity, open play: AC Milan vs Barca (0-1), risk 77.17

> *Barcelona's defensive structure broke down primarily through poor organisation during the [DEFENSIVE TRANSITION]. As AC Milan initiated an [ATTACKING TRANSITION], Barcelona's players were slow to regroup, leaving unmarked opponents in the [FINAL THIRD]. The failure to close down space in [DEFENDING IN DEFENSIVE THIRD] allowed Milan to create multiple [BALL IN THE BOX] scenarios.*

Example 2 — Moderate severity, set piece: AC Milan vs Barca (0-1), risk 50.0

> *Barcelona faced a clear threat due to failures in managing [DEFENSIVE TRANSITION] and positioning during AC Milan's attacks. The combination of [BALL IN THE BOX] situations and [SET PIECES] points to a lack of effective marking and organisation during dead-ball scenarios.*

6.3 Hallucination Mitigation

- **Data-limitation-aware system prompt:** The LLM is told what data is and is not available, and instructed not to invent details.
- **Code citation requirement:** Event codes must be cited in [BRACKETS], so every claim is verifiable against the evidence pack.
- **Confidence-gated language:** Only patterns with confidence ≥ 0.60 are described as "recurring." Below that, the prompt says "not enough evidence."
- **Tracking-aware interpretation rules:** The system prompt includes explicit distance thresholds and overload criteria. The LLM must flag when tracking coverage is sparse (<6 players tracked).
- **Post-processing:** Outputs are stripped of timestamps (regex), bullets/numbering are removed, and length is capped at 3–5 sentences to prevent rambling.
- **Caching:** SHA-256 of prompt → JSON file. Identical evidence packs always produce the same explanation.

7. Interactive Dashboard

A React + FastAPI dashboard lets coaching staff explore the analysis interactively:

- **Match selector:** Choose from all 11 matches
- **Risk timeline chart:** Colour-coded risk score over match time with goal markers
- **Danger moment list:** Clickable cards ranked by severity with LLM explanations
- **Video seek:** Click a danger moment to jump to the corresponding broadcast timestamp
- **Custom window analysis:** Click two points on the timeline for on-demand LLM explanation

Video offset calibration (pre-match broadcast time, halftime extra time) is stored per match for accurate video seeking.

8. Limitations and Honest Disclosure

8.1 What the System CAN Do

- Identify specific time windows where Barca was most defensively vulnerable
- Quantify danger severity on a 0–100 scale with four tiers
- Detect recurring event-code combinations that precede danger across matches
- Provide confidence scores for pattern recurrence based on prevalence and lift
- Generate LLM explanations constrained to the available evidence
- Link danger moments to broadcast video timestamps
- Extract spatial snapshots from tracking data (team shape, ball proximity, overload) to ground LLM explanations

8.2 What the System CANNOT Do

- Attribute failures to **individual players** — events are team-level
- Reliably compute defensive shape or compactness (tracking too sparse for full team)
- Compute player-level xG, progressive carries, or dribble success rates
- Determine **causal** relationships — patterns are correlational
- Replace expert coaching judgment — the system is an analytical aid, not a replacement

8.3 LLM Output Caveats

LLM explanations are constrained to the evidence pack and post-processed to remove timestamps and enforce length limits, but they are not infallible. Subtle inference errors can occur. All outputs should be reviewed by coaching staff before informing tactical decisions.

8.4 Sample Size

11 matches is a limited sample for pattern analysis. The confidence scoring accounts for this (small samples produce lower scores), but coaches should treat medium/low confidence patterns as hypotheses to track, not confirmed weaknesses.

8.5 Smart Tagging Accuracy

Smart Tagging data is not manually ground-truthed. It is inferred through automated processes and may contain both false positives and false negatives.

9. Coaching Recommendations

- **Transition vulnerability (Pattern 1, confidence 0.60):** Work on defensive recovery speed and positional discipline when committing players forward. The 2.66× lift makes this the highest-priority finding.
- **Mid-block bypass (Pattern 2, confidence 0.47):** Review pressing triggers and the gap between midfield and defence — opponents are skipping the mid-block entirely.
- **Midfield progression control (Pattern 3, confidence 0.46):** The midfield press needs to be more effective at preventing opponent progression. Both heavy defeats featured this pattern.
- **Second-half resilience:** Monaco went from 0 to 5 critical moments in H2. Opponent half-time adjustments deserve specific attention in match preparation.
- **Use the video linkage:** Every danger moment maps to a broadcast timestamp. Review footage alongside the analysis.
- **Track patterns over more matches:** All three patterns are at medium confidence. They need more data to confirm or rule out.

10. Suggestions for Metrica Nexus

- **Smart Tagging timeline offset tool:** Let users offset timestamps so they align with in-game time.
- **Player ID merging:** Let users merge fragmented player IDs in tracking data and connect them to known player names.
- **Player recognition confidence scores:** Include confidence values when correlating different player IDs.

11. Conclusion

The Pressure Cooker system turns 11 matches of Smart Tagging annotations and tracking data into a continuous risk timeline, detects 144 danger moments with 100% goal coverage, identifies 3 recurring vulnerability patterns, and generates LLM-assisted tactical explanations grounded in verifiable evidence.

The main tactical takeaway is that transition play is Barcelona’s primary defensive weakness, showing up in nearly half of all danger moments and producing the highest lift in the pattern analysis. Every insight the system produces is traceable to specific match moments, event codes, and broadcast timestamps, so coaching staff can verify anything against the footage.

Tactical explanations in Sections 6.2 are generated by GPT-4o-mini (via OpenRouter). All statistical analysis, risk scoring, pattern detection, and verification logic is deterministic Python code.