Technical Overview

Scientific and Engineering Summary of VALITICA's Injury Prediction Framework

Contents

1 Introduction

Professional football generates vast quantities of physical, medical and contextual data every day. While clubs analyse these datasets to understand training responses and performance trends, the biggest challenge remains identifying early patterns that precede soft-tissue injuries.

VALITICA provides a scientific framework that transforms multi-source data into daily, individualised risk indicators for intrinsic muscle injuries.

This document describes the principles behind the system's data pipeline, modelling approach, explainability layer and internal validation methodology.

The information presented here is high-level and intentionally does not reveal proprietary algorithms or intellectual property. Its purpose is to help medical, performance and technical staff understand the scientific foundations that support VALITICA's injury prediction model.

2 Data Pipeline and Input Requirements

VALITICA integrates seamlessly into the existing data ecosystem of each Club. No additional hardware is required.

2.1. Data Sources Used

The model ingests all relevant datasets available in professional football environments, including:

- External load (GPS / EPTS): distance, HSR, sprinting, accelerations, decelerations.
- Internal load: heart-rate response, HRV trends when available.
- Metabolic and mechanical indicators.
- Kinematic data.
- Strength and power testing.
- Wellness and questionnaire inputs.
- Injury history and recovery cycles.
- Medical records and treatment sessions.
- Match and training schedule.
- Environmental or contextual factors (e.g., fixture congestion).

2 Data Pipeline and Input Requirements

2.2. Data Quality Requirements

The accuracy of the model improves significantly when:

- Measurements are consistent over time.
- Variables remain stable (no frequent changes of technology/provider).
- Testing protocols are standardised.
- Gaps in data collection are minimised.

We also support Clubs in auditing and improving their data collection processes.

3 Variable Clusters and Multi-Source Integration

Soft-tissue injuries do not originate from a single factor. They emerge from interactions between physiological load, mechanical stress, individual history and contextual conditions.

VALITICA processes variables in clusters, enabling the model to detect weak signals that only become meaningful when combined.

3.1. Variable Clusters Included

- Metabolic
- Mechanical
- Kinematic
- Personal characteristics
- Injury history
- Injury treatment & recovery patterns
- Biomarkers (when provided)
- Internal load
- External load
- Cumulative load
- Thermal load
- Foster Index
- Borg / RPE Index
- VAS pain scale
- Club-specific or context-specific variables

3 Variable Clusters and Multi-Source Integration

3.2 Multi-Cluster Integration

Instead of analysing metrics independently, VALITICA evaluates:

- How clusters interact.
- How relationships change over time.
- How deviations from expected patterns map to injury risk.

This approach captures complexity that is invisible through single-variable analysis.

4 Player Baseline and Individual Profiling

Each player responds differently to identical workloads. Therefore, VALITICA develops a personalised baseline for every player based on historical data.

4.1. Baseline Components

- Typical physiological response to load.
- Expected mechanical patterns.
- Recovery behaviour after matches.
- Individual tolerance to load variability.
- Positional norms (full-backs vs midfielders vs forwards).

4.2 Detecting Deviations

Once the baseline is established, the model identifies:

- Overload exposure.
- Underload patterns.
- Atypical internal load responses.
- Declines in wellness or recovery indicators.
- Unusual trends during congested schedules.
- Inconsistencies during RTP reintegration.

Player-specific analysis is essential to VALITICA's accuracy.

5 Model Architecture

Our predictive engine uses a hybrid modelling approach combining Machine Learning, statistical inference and time-series analysis. Although the exact architecture remains proprietary, the principles governing it are transparent.

5.1. Core Principles

- Multi-source feature integration.
- Temporal modelling across training cycles.
- Identification of complex, non-linear relationships.
- Individualised prediction rather than population-level generalisation.
- Continuous model updating as new data is collected.

5.2. Why a Hybrid Approach?

Soft-tissue injury events are:

- Rare.
- Multifactorial.
- Influenced by individual history.
- Dependent on contextual dynamics.

A single-method model (e.g., purely machine learning or purely statistical) is insufficient. VALITICA combines several to optimise sensitivity and specificity.

6 Injury Risk Score and Explainability Layer

The output presented to staff is not a black box. Every risk score is accompanied by explainable indicators.

6.1. Output Components

- Current probability of intrinsic muscle injury.
- Trend line vs baseline..
- Most influential variable clusters.
- Alerts on atypical responses.
- Contribution summary (high-level, non-mathematical)

6.2. Explainability Principles

The system ensures:

- Transparent logic for elevated risk.
- Clear language aligned with medical workflows.
- No exposure of proprietary algorithmic structure.
- Clinically interpretable patterns rather than mathematical features.

7 Internal Validation and Performance Metrics

VALITICA has been validated internally using anonymised datasets from elite-level professional environments.

7.1. Core Performance Indicators

- Predictive effectiveness up to 90% when data quality is high.
- Higher precision as long-term history accumulates.
- Reduced false positives through baseline refinement.
- Improved risk differentiation during congested fixture cycles.

7.2. Validation Methodology

- Multi-season historical analysis.
- Cross-user and cross-context verification.
- Continuous error reduction as more data is ingested.

7.3. Interpretation Guidelines

A high risk score does not guarantee that an injury will occur. It indicates:

- Elevated probability.
- Need for clinical assessment.
- Consideration for load adjustment.
- Monitoring across the training week.

8 Ethical, Regulatory and Security Foundations

VALITICA complies fully with GDPR and the EU AI Act, and has built-in safeguards for ethical and secure data processing.

8.1. Core Principles

- Data minimisation.
- Purpose limitation.
- Full club ownership of data.
- Multi-layer authentication.
- Strict access control.
- Complete action logging.

8.2. Anonymisation and Privacy

- All data used for model training is anonymised.
- No club-specific data is shared across environments.
- Federated learning methods ensure privacy when required

8.3. Deployment and Hosting

- Cloud-native infrastructure.
- Encrypted data transfer and storage.
- High availability for daily operational use.

9 Limitations and Appropriate Use Conditions

No injury prediction system is infallible. VALITICA establishes clear limitations to ensure scientific integrity.

- Predictions depend on the quality and consistency of club data.
- The model supports clinical decision-making; it does not replace medical judgement.
- Extreme contextual circumstances (e.g., acute illness, trauma injuries) fall outside the model scope.
- Risk outputs should always be interpreted within the player's holistic medical and performance context.

10 Contact and Support

For further technical information, integration support or scientific queries, please contact us.

VALITICA
Email: valitica@valitica.com
Website: valitica.com

VALITICA INTELLIGENCE
Hermosilla 48, 1D
28001 Madrid
Spain