BMJ Open Sport & Exercise Medicine

Effect of an unsupervised multidomain intervention integrating education, exercises, psychological techniques and machine learning feedback, on injury risk reduction in athletics (track and field): protocol of a randomised controlled trial (I-ReductAI)

Spyridon Iatropoulos ,[1] Pierre-Eddy Dandrieux ,[1] David Blanco ,[2] Alexis Ruffault,[3] Estelle Gignoux,[1] Constance Mosser,[1] Karsten Hollander ,[4] Laurent Navarro,[5] Pascal Edouard [1,6]

To cite: Iatropoulos S, Dandrieux P-E, Blanco D, *et al*. Effect of an unsupervised multidomain intervention integrating education, exercises, psychological techniques and machine learning feedback, on injury risk reduction in athletics (track and field): protocol of a randomised controlled trial (I-ReductAI). *BMJ Open Sport & Exercise Medicine* 2025;**11**:e002501. doi:10.1136/bmjsem-2025-002501

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/bmjsem-2025-002501).

Accepted 7 February 2025

For numbered affiliations see end of article.

Correspondence to
Dr Spyridon Iatropoulos; Spyridon.Iatropoulos@univ-st-etienne.fr

ABSTRACT

The primary aim is to assess the impact of a multidomain intervention that integrates education, exercise, psychological techniques and machine learning feedback on the duration athletes remain free from injury complaints leading to participation restriction (ICPR) during a 20-week summer competitive athletics season. The secondary aims are to assess the intervention's effect on reducing (i) the incidence, (ii) the burden, (iii) the period prevalence and (iv) the weekly prevalence of ICPR during the same timeframe. We will perform a two-arm randomised controlled trial. This study will involve an intervention group and a control group of competitive athletes licensed with the French Federation of Athletics, aged between 18 and 45, over an outdoor athletics competitive season lasting 20 weeks (March to July 2025). Data will be collected before the start (demographic, training and injury history) and one time per day (training and competition volume/intensity, perceived physical and psychological state, and illness and injury incidents) for both groups. The intervention group will be required to (i) view a series of 12 educational videos on injury prevention, (ii) engage in discipline-specific exercise programmes, (iii) implement stress and anxiety management techniques and (iv) view daily the injury prognostic feedback generated by the athlete's collected data based on machine learning. Outcomes will be analysed over the final 14 weeks of follow-up to allow time for the intervention to establish any potential efficacy. The primary outcome will be the time-to-event for each ICPR. Secondary outcomes will include (i) incidence, (ii) burden, (iii) period prevalence and (iv) weekly prevalence of ICPR. The primary outcome will be analysed using a Prentice–Williams–Peterson gap-time model. In contrast, the secondary outcomes will employ Poisson (i, ii), logistic (iii) and generalised estimating equations (iv) regression models, respectively.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Strategies to reduce the risk of injury are essential for decreasing the likelihood, severity and consequences of injuries in athletics.
⇒ Strategies for reducing injury risk should be developed using a holistic approach that reflects the complex aetiology of sports injuries.
⇒ Different injury risk reduction strategies, such as education, exercises, psychological techniques and machine learning-based injury prognostic feedback, have been researched, but each one in isolation.

WHAT THIS STUDY ADDS

⇒ This protocol for a randomised controlled trial develops a multidomain unsupervised intervention consisting of education, exercises, psychological techniques and injury prognostic feedback based on machine learning, following the complex aetiology of injuries.
⇒ This study hypothesises that the proposed multidomain intervention will reduce injury risk in athletics athletes over a 20-week summer competitive season.

INTRODUCTION

In athletics (track and field), over half of athletes sustain at least one injury during a season, which can impact participation,[1] performance,[2 3] career length[4] and athletes' health.[1] Therefore, strategies to reduce injury risk are essential to lower the injury's probability, severity and/or consequences. To our knowledge, only two randomised controlled trials (RCTs) have been published that evaluate injury risk reduction strategies in athletics: an exercise-based programme[5] and

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This protocol presents a comprehensive approach to developing strategies for preventing sports injuries.
⇒ This protocol boasts several methodological strengths related to the statistical analysis of injury risk data, which are not yet widely adopted in sports medicine.
⇒ The results of this study can directly impact current injury prevention practices in athletics and lay the foundation for developing multi-domain injury risk reduction strategies in other sports.

an educational programme.[6] Consequently, there is a need to continue developing injury risk reduction strategies in athletics and assessing them through RCTs.

The development of effective injury risk reduction strategies requires the understanding of the aetiology of injuries.[7] In the last decade, there has been a paradigm shift in the research of sports injury causality, from identifying single risk factors[8 9] to investigating the additive risk accumulation of multiple risk factors,[10] and finally to acknowledging the complexity of a system of factors that interact with each other and jointly influence the injury risk.[11–13] This implies that injury risk reduction strategies should be developed following a holistic approach corresponding to sports injuries complex aetiology.[14]

However, to our knowledge, research studies on injury risk reduction strategies in any sport have yet to apply a holistic approach. Despite many different injury risk reduction strategies being investigated in sports,[15] each has been evaluated in isolation, encompassing education,[6 16] exercise-based programmes[5 17] and psychological techniques.[18 19] Concurrently, with the rapid advancements and widespread adoption of artificial intelligence (AI), particularly machine learning (ML) techniques, injury prognostic feedback (IPF) through ML modelling could also form part of a holistic injury risk reduction strategy.[20 21]

In this context, we hypothesised that providing athletes with a multidomain unsupervised intervention that integrates education, exercise-based programmes, psychological techniques and IPF delivered via a website reduces the risk of injury during a summer competitive season in athletics, compared with standard athletics activities. While analysing the effect of each component of such a multidomain strategy individually may be suitable for accurately estimating their exposure-outcome relationship, this approach fails to fully capture the complexity of the injury problem.[11 12]

OBJECTIVES

Given that various epidemiological metrics can indicate the injury risk within a population,[22] we will evaluate our hypothesis through multiple, hierarchically ordered objectives.

The primary objective will be to assess the impact of the multi-domain intervention on the duration that athletes stay free from injury complaints leading to participation restriction (ICPR) during a summer competitive season in athletics.

The secondary objectives are to assess the impact of the multidomain intervention during a summer competitive athletics season on:

1. The total number of ICPR sustained by athletes per 1000 hours of athletic activity (ie, incidence);
2. The total time (in days) of athletes with ICPR per 1000 hours of athletics activity (ie, burden);
3. The overall proportion of athletes sustaining at least one ICPR (ie, period prevalence);
4. The weekly proportion of athletes sustaining at least one ICPR (ie, weekly prevalence).

METHODS

Study design and overall procedure

We will conduct a RCT, titled Injury risk Reduction in Athletics through Integration (I-ReductAI) of education, exercise-based programmes, psychological techniques and ML feedback, over one competitive athletics (track and field) outdoor season lasting 20 weeks, from March 2025 to July 2025. This competitive period is crucial for athletes to remain injury-free. The study will be divided into a 6-week intervention and data collection period, followed by a 14-week phase dedicated to statistical analysis, during which the intervention and data collection will continue. This study will involve competitive athletes registered with the French Federation of Athletics (Fédération Française d'Athlétisme (FFA), https://www.athle.fr). The flowchart of the study is illustrated in figure 1.

The study protocol was reviewed and approved by the Committee for the Protection of Persons (Comité de protection des personnes Ouest VI n° 2024-A02281-46). This study protocol is reported following the *Standard Protocol Items: Recommendations for Interventional Trials involving Artificial Intelligence* guidelines.[23]

Trial registration number

It is registered on ClinicalTrials.gov (https://clinicaltrials.gov, Identifier: NCT06805162).

Patient and public involvement

There was no direct patient or public involvement in the design and development of the study protocol. However, most co-authors are athletes and routinely work as physicians, physiotherapists or coaches within the athletics community. Additionally, the conception and development of the study's intervention were informed by an online survey on athletes' preferences concerning injury prevention education[24] and another study on athletics stakeholders' perceptions regarding injury prevention.[25]

Equity, diversity and inclusion statement

All FFA-licensed athletes will be eligible and may be included in this study without any restrictions based on gender, race/ethnicity/culture, socioeconomic status or representation of marginalised groups. The only

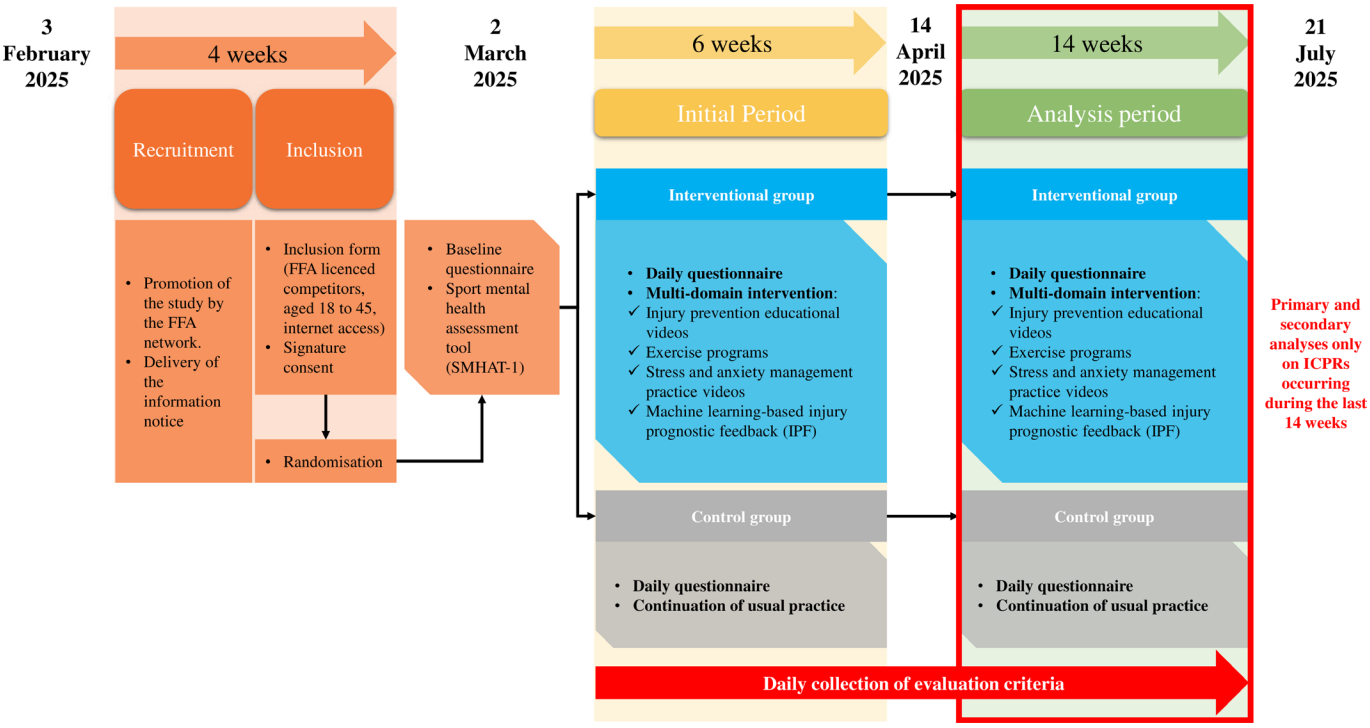

Figure 1 The Injury risk Reduction in Athletics through Integration (I-ReductAI) study timeline and flowchart. FFA, French Federation of Athletics.

restrictions will be age (between 18 and 45 years) and country (France). The research team comprises four junior and five senior researchers, including two women and seven men, from various disciplines (medicine, sports medicine, physical medicine and rehabilitation, exercise physiology, physiotherapy, psychology, sports science, sports epidemiology, data science, statistics and research methodology) and four different countries in Europe (France, Germany, Greece and Spain).

Population recruitment and inclusion and exclusion criteria

The recruitment and inclusion period will occur from 3 February 2025 to 2 March 2025 (4 weeks). The FFA will distribute information regarding this study through individual emails to their licensed athletes and athletics clubs and via their website and social media platforms. These emails will contain an invitation for individual participation in this study, along with a description of the study's purpose and procedure, participation rights, an information letter and a link to register on the website application (IPrevApp, https://iprevapp.emse.fr). Once registered, athletes must complete an eligibility survey (see the inclusion criteria below). If deemed eligible, they must provide their free, informed and immediate consent by clicking on 'I want to participate in this study'.

The inclusion of athletes will be based on the following criteria: athletes must (i) be licensed at the FFA for competition in sprints, hurdles, jumps, throws, combined events or endurance disciplines, without any contraindications for competitive athletics activity attested by the license at the FFA; (ii) be aged between 18 and 45 years old; and (iii) have daily access to a digital device (smartphone, computer, tablet) with a network connection (public or private). We will not exclude athletes based on their baseline injury status and history.[5 20] Athletes will be excluded if they are deprived of liberty, subject to legal protection (guardianship, curatorship, legal protection) or are pregnant.

Materials

The recruitment, intervention and data collection will be done through an online website (IPrevApp, https://iprevapp.emse.fr) accessible via mobile and computer devices. The website will be hosted on a local host server, AlmaLinux V.8.6 (Sky Tiger), with a database (MariaDB V.10.8.8).

Randomisation, allocation and blinding

A random allocation sequence will be generated before the start of the recruitment phase to assign athletes to the intervention and control groups with a 1:1 allocation ratio. As new eligible athletes are added, each will be allocated to one of the groups by matching their order of registration to the allocation sequence. Athletes will be informed about their allocation group on 2 March 2025. At the end of the study, if the intervention is deemed efficacious, athletes of the control group will have full access to it.

The research team will be unaware of the allocation for each participant, apart from one investigator responsible for the allocation process, who will not participate in the statistical analysis. Given the nature of the intervention, the participating athletes will not be blinded to

their allocation. There will be no face-to-face interaction between the research team and the athletes.

Data collection

The athletes will self-report all collected data through questionnaires on the study website. From 2 March 2025, included athletes will be requested to complete a baseline questionnaire about their characteristics: sex, date of birth, height, body mass, primary athletics discipline, number of years practising athletics, average weekly hours of athletics and non-athletics sport activity, their illness and injury history during the winter period of the 2024–2025 athletics season and their mental health status using the sports mental health assessment tool-step 1 (SMHAT-1),[26] developed by the International Olympic Committee (table 1).

During the 20-week follow-up period (from 3 March 2025 to 21 July 2025), all athletes will be asked to complete a daily questionnaire each evening, in which they will provide information about their athletic activity (ie, duration and intensity), psychological state (ie, motivation, stress, anxiety), physiological condition (ie, fatigue), sleep patterns (ie, duration and perceived quality), and any illnesses or injuries occurring on the same day (see table 1). Completing each daily questionnaire will take approximately 1 min and will be available on the study website from 15:00 on the same day for the following 72 hours. After this period, any incomplete questionnaires will be regarded as missing data. The individual response rate will be calculated as the ratio of completed daily questionnaires to the total number of questionnaires expected (n=140).

For secondary analyses beyond this study's objectives, in the last week of the follow-up, athletes from both groups will be asked to complete again the SMHAT-1 questionnaire and a short questionnaire about sociocognitive aspects that might be related to the adoption of injury risk reduction programmes inspired by Ruffault *et al*.[27]

A detailed form of all questionnaires is presented in the online supplemental file 1.

Intervention

The intervention will be accessible via the study website. The intervention group will be asked to engage with all components of the multi-domain intervention for 20 weeks, while the control group will continue their regular athletic activities, without any access to the multi-domain intervention. Each component of the intervention is detailed below.

Education

A series of 12 short videos (2–3 min each) will be available on the study website. These videos will cover various topics related to injury prevention in athletics: the anatomy of frequently injured areas, definitions and pathophysiology of injuries, injury epidemiology, risk factors for injury, strategies for injury risk reduction and injury management. The videos will be designed to be accessible to an audience without specialised knowledge in health and medicine. Athletes in the intervention group can adjust their practice according to the information provided in the videos. They will not receive individual recommendations on how to manage injury risk. Therefore, the responsibility for injury risk management will rest solely with them. Access and viewing time for each video will be monitored. Whenever a video is accessed and viewed for at least 50% of its total length, it will automatically count as one completed session for this intervention component. Athletes may complete more than one session daily and can watch the same session multiple times.

Exercise-based programmes

The research team have selected a collection of 30 exercises aiming at improving the flexibility, stability and/or strength of the most frequently injured anatomical locations in athletics[1] based on available published[5] and anecdotal evidence from practitioners. Each exercise has two levels of difficulty (easier or harder). Since injury patterns differ between disciplines,[1 28] we have created three specific programmes in the form of videos for each of the five discipline groups (sprints/hurdles, jumps, throws, combined events and endurance events including middle/long distances, road/trail running and race walking). Each programme is available in two difficulty levels and consists of seven exercises. Each programme is approximately 7 to 10-min long. The athletes in the intervention group will only have access to the three programmes specific to their discipline, which will be determined according to what athletes reported in their baseline questionnaire. The programmes will be available at any time throughout the day. Athletes will receive no individual guidance on how to perform the programmes (eg, location, time of the day, frequency of sessions per week, and selection between easy and hard sessions). However, athletes will be informed through the education component of the intervention that they should perform an exercise-based programme at least two times per week and decide between easy and hard programmes based on their capabilities and their usual athletics activities. The access and watching time of each video will be tracked. Whenever a video is accessed and watched for at least 50% of its total length, it will automatically count as one completed session for this intervention component. Athletes can complete more than one session per day and the same session more than once.

Psychological techniques

A series of 12 audio recordings that athletes can follow to practise stress and anxiety management techniques (ie, mindfulness, meditation and/or breathing exercises) will be available on the study website. The recordings will vary in duration from 1 to 10 min. Individual guidance will not be provided on where, when, how often and how to perform the techniques while watching the videos. However, athletes will be informed through the educational component of the intervention that they should

Table 1 Variables collected with the baseline and monitoring (daily) questionnaires. The type of answer for each variable and whether or not it is included as a predictor in the injury prognostic feedback based on machine learning are also indicated.

Questionnaire		Variable	Type of answer	IPF input
Baseline	1	Sex	Binary	x
	2	Date of birth	Discrete	x
	3	Height	Continuous	x
	4	Bodyweight	Continuous	x
	5	Years of athletics practice	Discrete	x
	6	Average weekly hours of athletics activity	Continuous	x
	7	Average weekly hours of sport activity outside athletics	Continuous	x
	8	Primary athletics discipline	Categorical	x
	9	Injury history since September 2024	Binary	x
	10	Illness history since September 2024	Binary	x
	11	Sports mental health assessment tool-step 1 (SMHAT-1)	Ordinal	
Monitoring	1	Time of falling asleep	Continuous	x
	2	Time of waking up	Continuous	x
	3	Sleep quality	Continuous	x
	4	Sense of fatigue	Continuous	x
	5	Sense of stress	Continuous	x
	6	Sense of anxiety	Continuous	x
	7	Motivation to train	Continuous	x
	8	Level of stress during the day	Continuous	x
	9	Sense of satisfaction at the end of the day	Continuous	x
	10	Total training duration	Continuous	x
	11	Total training intensity	Discrete	x
	12	Training details	Optional	
	13	Total duration of competitions	Continuous	x
	14	Total intensity of competitions	Discrete	x
	15	Total duration of sports activity outside athletics	Continuous	
	16	Total intensity of sports activity outside athletics	Continuous	
	17	Illness event	Categorical	x
	18	Illness detail(s)	Short response	
	19	Injury event	Categorical	x
	20	New injury event	Binary	x
	21	New injury place of appearance	Categorical	
	22	Mode of onset	Categorical	
	23	New injury mechanism of occurrence	Categorical	
	24	New injury body laterality	Categorical	
	25	New injury body locality	Categorical	
	26	Impact of new injury on the training	Categorical	
	27	Tracking of past injury(ies)	Categorical	x
	28	Impact of past injury(ies) on training	Categorical	

ICPR, injury complaints leading to participation restriction; IPF, ICPR prognostic feedback.

engage in a mindfulness or meditation technique at least once a week and a breathing exercise at least once a week. Access and listening time for each audio will be tracked. Whenever audio is accessed and listened to for at least 50% of its total length, it will automatically count as one completed session for this intervention component. Athletes can complete more than one session daily and may listen to the same session more than once.

Injury complaints leading to participation restriction (ICPR) prognostic feedback

An individualised ICPR prognostic feedback (IPF) for the subsequent day will be available daily to athletes on the study website immediately after completing that day's monitoring questionnaire.[20 21] The data collected from the baseline and the daily questionnaires up to a specific day (*D*) will be used as input for the ML model (table 1). The predicted probabilistic output of this model will correspond to the risk of ICPR for the following day (*D+1*).[20 21] Specifically, the IPF will consist of a probability (ranging from 0% to 100%) that corresponds to the risk of sustaining an ICPR the following day, accompanied by (i) the discrimination score of the model (ie, area under the curve of the receiver operating characteristic), which reflects the confidence of the estimated probability, and (ii) the contribution of each input variable (eg, sleep duration, training load) to the estimated probability, based on the Shapley values derived from the model.[29] If the daily questionnaire has not been completed, the IPF for the following day will not be generated.

First, the ML model will be pretrained on similar data collected with the same methodology from a previous study.[20 21] The parameters of the ML model will be updated weekly throughout the study and as more data is collected. Preprocessing will be applied by encoding each categorical variable to several binary variables and by scaling continuous variables if it improves model performance. For model selection and hyperparameter tuning, as well as the updated models every week, we will compare a range of algorithms and combinations of hyperparameters. The algorithms explored will be: (i) logistic regression, (ii) decision tree, (iii) support vector machine, (iv) random forest and (v) extreme gradient boosting. The hyperparameters explored will be: (i) parameter C (ie, weight of penalisation), (ii) class weight (ie, relative weight of the injured over the uninjured class), (iii) number of estimators (only for tree-based algorithms) and (iv) tree depth (only for tree-based algorithms). Each developed model will undergo an internal validation process using non-parametric bootstrapping (500 samples), described in detail by Collins *et al*.[30] The criterion for selecting the best-performing model will be the optimism-corrected area under the receiver operating characteristic.[31] The sci-kit-learn library[32] of the Python programming language (https://www.python.org) will be used. The exact code of the model development process, as published for another study currently under review for publication, can be found in: https://github.com/spyrosiatrop/ML_Injury_Prediction_Prevathle_RCT.

Access to the IPF page of the study website will be monitored. Visiting the page displaying the generated IPF at least once will be considered one completed session. Further visits on the same day will not count as additional sessions. Athletes in the intervention group will be free to adapt their practice based on the IPF. They will not receive individual recommendations on how to manage injury risk. Consequently, managing injury risk will be their sole responsibility.

Adherence

The adherence to the entire intervention over a specified period will be calculated as the ratio of the total number of completed sessions for all intervention components divided by the duration of the period in days. The adherence to a single intervention component over a specified period will be calculated as the ratio of the total number of completed sessions for this specific component divided by the duration of the period in days. Recommendations through the educational intervention component and regular automated reminders via the study website will be arranged to promote athletes' engagement with each of the intervention components.

Injury definition

An injury is 'a pain, physical complaint or musculoskeletal lesion sustained by an athlete regardless of whether it received medical attention or its consequences concerning impairments in competition or training'.[5 22 33] Athletes reporting an injury complaint in the daily questionnaire will also be asked to report (i) the circumstance of injury occurrence (training, competition, unrelated to athletics activity), (ii) the mode of onset (sudden, gradual), (iii) the injury location and (iv) the consequences of this injury on athletics participation, classified as (a) full participation without discomfort, (b) full participation with discomfort, (c) partially restricted participation due to injury complaint and (d) completely restricted participation due to injury complaint.[5 20] The last two categories (ie, injury complaints with partially (*c*) or completely (*d*) restricted participation) define the 'injury complaints leading to participation restriction' (ICPR).[5 20]

Outcomes

The primary outcome will be the time to each ICPR during the last 14 weeks of follow-up. For the first ICPR, the time to the event will be calculated from the start of the seventh week of follow-up until the first ICPR. For all subsequent ICPRs, the time to the event will be calculated from the end of the previous ICPR to the start of the next ICPR.

The secondary outcomes, which will also be analysed during the final 14 weeks of follow-up, will be:

1. Incidence of ICPR, defined as the total number of ICPR per 1000 hours of athletics activity.
2. Burden of ICPR, defined as the number of days with ICPR per 1000 hours of athletics activity;
3. Period prevalence, defined as the proportion of athletes sustaining at least one ICPR;
4. Weekly prevalence defined as the weekly proportion of athletes sustaining at least one ICPR.

We do not expect the proposed intervention to be effective as soon as it is implemented. Indeed, we hypothesise that based on the characteristics of the intervention,

there will be a period needed before the intervention could become efficacious to reduce the risk of injury.[34] Given the lack of a pilot study aiming to determine the exact length of this period for this study's intervention and the need to declare our statistical analysis plan in advance, we based our choice on available evidence. Other studies in football have also implemented their intervention for 8 and 13 weeks before analysing the effect of their intervention on the injury risk reduction in the follow-up thereafter, but without giving a justification for the selection of this period's length.[35 36] A previous study on an exercise-based programme intervention in athletics showed that 6 weeks were necessary to gain efficacy.[34] Therefore, to determine the efficacy of the intervention, only the last 14 weeks of the follow-up will be included in the analysis.

Statistical analysis

We will initially carry out a descriptive analysis of the population characteristics, the primary and secondary outcomes, the exposure to athletic activity, the individual response rate during the follow-up, the adherence to the intervention components for the intervention group and other collected variables. We will employ frequencies and relative frequencies (n (%)) for categorical variables and means with SD for normally distributed continuous variables or medians with interquartile ranges for non-normally distributed continuous variables. This analysis will be conducted separately for the first 6 weeks, the last 14 weeks and the overall follow-up period.

For the primary objective, we will perform a survival analysis using an adjusted Prentice–Williams–Peterson gap-time model (extension of the Cox proportional hazards model for recurrent event analysis),[37 38] stratified by the order of ICPR occurrence (first injury, second injury, etc). Athletes will be right-censored at (a) their first day of missing data, (b) the day they report an ICPR that is non-related to athletics activity or (c) the end of the 20-week follow-up if they report no ICPR. The independent variable of interest will be the group allocation, while the analyses will be adjusted for some independent variables: (i) ICPR history during the last winter season, (ii) ICPR history during the first 6 weeks of follow-up, (iii) the interaction term between the variables 'ICPR history during the first 6 weeks of follow-up' and 'group allocation', and (iv) exposure (ie, hours of athletics activity). We will consider adjusting for additional variables (eg, age, sex, athletic discipline) if this improves the model's goodness-of-fit.

For each of the secondary objectives, the statistical analyses will be:

1. Adjusted Poisson regression analysis with the number of new injuries as dependent variable;
2. Adjusted Poisson regression analysis with the days with ICPR as dependent variable;
3. Logistic regression with the occurrence or not occurrence of ICPR during the analysis period as a binary dependent variable.
4. Generalised estimating equation model with the occurrence or not occurrence of ICPR each week as a binary dependent variable.

The independent variable of interest will be the group allocation, while the analyses will be adjusted for the same independent variables as in the primary objective (see above). For objectives *i* and *ii*, exposure (ie, hours of athletic activity) will be included as an *offset* term in the models.[39] For these analyses, we will consider only athletes with at least a 50% response rate during the analysis period.[40]

We will conduct a sensitivity analysis based on the response rate to the daily questionnaires by repeating the secondary analyses, varying the individual response rate threshold from 0% to 100%.[21 41] We will also consider conducting a sensitivity analysis based on best-case and worst-case scenarios, if appropriate, given the extent of missing data. Furthermore, we will analyse the dose-response relationship of the intervention from the first to the 12th weeks (ie, testing shorter and longer periods than the preselected 6 weeks) using both primary and secondary outcomes.[34] This analysis aims to retrospectively explore the appropriateness of selecting 6 weeks to allow the intervention to achieve efficacy before the start of the analysis period and to obtain information on the time required for this intervention to gain efficacy. The analyses of the between-group comparison for the primary and secondary outcomes will be conducted according to the intention-to-treat principle, meaning that we will not adjust the analyses based on the adherence of each athlete to their intervention. Nonetheless, we will repeat these analyses while adjusting for each athlete's level of adherence to the intervention.

Sample size justification

An a priori sample size calculation for the primary objective was performed. Recurrent event analysis is a relatively new extension of the Cox proportional hazards analysis, so an exact sample size calculating tool could not be found. Nevertheless, recurrent event analyses for the same sample size generally demonstrate higher power levels than Cox proportional hazards analysis.[42] Consequently, we estimated the required sample size for a Cox proportional hazards analysis, which reflects the upper limit of our necessary sample size. As this multidomain intervention has not been previously studied, we could not estimate its expected effect size directly. Instead, there is some evidence for each element individually:

- An RCT analysing the effect of an educational online platform in youth athletics found an approximate 40% risk reduction.[6]
- A secondary analysis of an RCT analysing the effect of an exercise-based injury prevention programme in athletics showed an approximate 60% risk reduction after the first 6 weeks of intervention.[34]

- A systematic review analysing the effect of psychological techniques on sports injury risk reduction found an overall small to large effect size (Cohen's d=0.2–1.2).[18]
- A prospective study analysing the association between the use of IPF and the reduction of sports injuries showed a small but significant association.[21]

Given that the interaction of the four components is unknown, we conservatively anticipated that our intervention would reduce injury risk by at least 30% to avoid underestimating the sample size. Based on data from a previously conducted RCT in athletics,[5] we estimate (i) the median time-to-event of the control group to be 4 weeks (28 days) and (ii) the average dropout rate during the first 14 weeks to be approximately 1% per day. Given these parameters and using alpha=0.05, power=0.9, allocation ratio 1:1 and expected risk reduction of 30% in the intervention group (relative hazard=0.7), it was estimated that 330 events were needed. Therefore, for a follow-up of 14 weeks, the estimated total sample size was 516 athletes (258 per group). However, to account for the dropout during the first 6 weeks (42 days) of intervention, we estimated that we should increase the sample size to $516 \cdot 1.01^{42}$=784 athletes (392 per group).

Strategies to limit the bias

Sampling bias will be controlled by offering participation in the study to all licensed athletes of the FFA, regardless of their characteristics. Selection bias will be avoided by randomising and concealing the allocation of participants in the two groups from the researchers leading and conducting the study. Evaluation bias will be avoided by using remote and self-reported data collection by the participants. Attrition bias will be avoided by implementing strategies to reduce the amount and impact of missing data: (i) use only mandatory close-ended questions (ie, accepting a predetermined range of possible answers) in the questionnaires; (ii) select a time-to-event outcome for our primary objective, which is analysed with survival analysis techniques that treat missing data as right-censored data (see Statistical analysis); and (iii) facilitating athletes' engagement in completing the questionnaires. The latter will be achieved through (a) an automatic reminder sent daily at 20:00, (b) posts on social media (Instagram, https://www.instagram.com/iprevapp; and YouTube, https://www.youtube.com/@iprevapp) and periodic newsletters via the study website that present and discuss the study aims and the usefulness of the study website, (c) access to a visual dashboard displaying the individual data provided over time and (d) gamification[43] through a progressive 'medal-awarding' system (ie, from 'regional bronze' to 'Olympic gold') that allows athletes to advance by earning points for each questionnaire completed. No imputation strategy will be applied for missing data. Finally, reporting bias will be controlled by registering the study protocol in a public registry of randomised clinical trials (ClinicalTrials.gov, https://clinicaltrials.gov, Identifier: NCT06805162) before the enrolment of the first participant. Only SI, P-ED, PE, LN and the engineers who developed the application and managed the CHU server will have access to the database before, during and after the trial. Any extraction from this database will necessarily involve the pseudo-anonymisation of the subjects. The raw database will not be made available online.

The study results will be communicated through articles in peer-reviewed journals, following the *Consolidated Standards of Reporting Trials for interventions involving Artificial Intelligence*[44] guidelines for manuscript preparation and at international scientific congresses. Individuals who have contributed to the design and implementation of the protocol will be eligible for inclusion in publications as co-authors. The study findings will also form part of SI's doctoral thesis. Additionally, participants in this study will be informed about the study results. The dissemination of results to end users will occur to facilitate knowledge translation.

Author affiliations
[1]Laboratoire Interuniversitaire de Biologie de la Motricité (EA 7424), Université Jean Monnet Saint-Etienne, Lyon 1, Université Savoie Mont-Blanc, Saint-Etienne, France
[2]Department of Physiotherapy, Universitat Internacional de Catalunya, Sant Cugat del Vallès, Barcelona, Spain
[3]Laboratory Sport, Expertise, and Performance (EA 7370), Institut National du Sport de l'Expertise et de la Performance (INSEP), Paris, France
[4]Institute of Interdisciplinary Exercise Science and Sports Medicine, MSH Medical School Hamburg, Hamburg, Germany
[5]Centre CIS, Laboratoire Interuniversitaire de Biologie de la Motricité, Univ Lyon, Univ Jean Monnet, Mines Saint-Etienne, Saint-Etienne, France
[6]Department of Clinical and Exercise Physiology, Sports Medicine Unit, University Hospital of Saint-Etienne, Saint-Etienne, France

X Spyridon Iatropoulos @spyros_iatrop, Pierre-Eddy Dandrieux @PE_Dandrieux, David Blanco @david_blanco91, Constance Mosser @ConstanceMosser, Karsten Hollander @k_hollander_, Laurent Navarro @LaurentNavarro5 and Pascal Edouard @PascalEdouard42

Acknowledgements The authors thank Diana Rimaud, Arnauld Garcin and the University Hospital of Saint-Etienne for their help acquiring approval from the Committee for the Protection of Persons and Colin Riviere for developing the website application.

Contributors All co-authors contributed to conceiving and designing the protocol. SI wrote the initial draft, and all co-authors contributed to and approved the final manuscript. SI is the guarantor of the manuscript.

Funding The authors have not declared any specific grant for this research from funding agencies in the public, commercial, or not-for-profit sectors. This research is part of a PhD scholarship funded by the University of Saint-Etienne; this funding source did not influence the design, conduct, analysis, interpretation of the data, or the decision to present the results. This research is supported by the University Hospital Centre of Saint Etienne (France). DB was funded by the Ministerio de Ciencia e Innovación (Spain) [PID2023-148033OB-C21].

Competing interests None declared. PE is an Associate Editor for the British Journal of Sports Medicine, BMJ Open Sport & Exercise Medicine, and the Scandinavian Journal of Medicine & Science in Sports. KH is the head team physician for the German Athletics Federation, the German Journal of Sports Medicine Editor, and an Associate Editor for BMJ Open Sport & Exercise Medicine. PED is a member of the Student Panel of BMJ Open Sport & Exercise Medicine.

Patient and public involvement Patients and/or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; internally peer-reviewed.

Data availability statement No data are available.

ORCID iDs
Spyridon Iatropoulos http://orcid.org/0000-0003-3381-2403
Pierre-Eddy Dandrieux http://orcid.org/0000-0001-7230-6728
David Blanco http://orcid.org/0000-0003-2961-9328
Karsten Hollander http://orcid.org/0000-0002-5682-9665
Pascal Edouard http://orcid.org/0000-0003-1969-3612

REFERENCES

1 Edouard P, Dandrieux P-E, Iatropoulos S, *et al*. Injuries in Athletics (Track and Field): A Narrative Review Presenting the Current Problem of Injuries. *DtschZSportmed* 2024;75:132–41.
2 Chapon J, Navarro L, Edouard P. Relationships Between Performance and Injury Occurrence in Athletics (Track and Field): A Pilot Study on 8 National-Level Athletes From Sprints, Jumps and Combined Events Followed During at Least Five Consecutive Seasons. *Front Sports Act Living* 2022;4:852062.
3 Raysmith BP, Drew MK. Performance success or failure is influenced by weeks lost to injury and illness in elite Australian track and field athletes: A 5-year prospective study. *J Sci Med Sport* 2016;19:778–83.
4 Edouard P, Mosser C, Chapon J, *et al*. Understanding the first injury in athletics and its effect on dropout from sport: an online survey on 544 high-level youth and junior athletics (track and field) athletes. *BMJ Open Sport Exerc Med* 2024;10:e001767.
5 Edouard P, Steffen K, Peuriere M, *et al*. Effect of an Unsupervised Exercises-Based Athletics Injury Prevention Programme on Injury Complaints Leading to Participation Restriction in Athletics: A Cluster-Randomised Controlled Trial. *Int J Environ Res Public Health* 2021;18:11334.
6 Jacobsson J, Kowalski J, Timpka T, *et al*. Universal prevention through a digital health platform reduces injury incidence in youth athletics (track and field): a cluster randomised controlled trial. *Br J Sports Med* 2023;57:364–70.
7 van Mechelen W, Hlobil H, Kemper HC. Incidence, severity, aetiology and prevention of sports injuries. A review of concepts. *Sports Med* 1992;14:82–99.
8 van der Worp H, van Ark M, Roerink S, *et al*. Risk factors for patellar tendinopathy: a systematic review of the literature. *Br J Sports Med* 2011;45:52:446–52:.
9 Freckleton G, Pizzari T. Risk factors for hamstring muscle strain injury in sport: a systematic review and meta-analysis. *Br J Sports Med* 2013;47:351–8.
10 Meeuwisse WH, Tyreman H, Hagel B, *et al*. A dynamic model of etiology in sport injury: the recursive nature of risk and causation. *Clin J Sport Med* 2007;17:215–9.
11 Hulme A, Finch CF. From monocausality to systems thinking: a complementary and alternative conceptual approach for better understanding the development and prevention of sports injury. *Inj Epidemiol* 2015;2:31.
12 Bittencourt NFN, Meeuwisse WH, Mendonça LD, *et al*. Complex systems approach for sports injuries: moving from risk factor identification to injury pattern recognition-narrative review and new concept. *Br J Sports Med* 2016;50:1309–14.
13 Pol R, Hristovski R, Medina D, *et al*. From microscopic to macroscopic sports injuries. Applying the complex dynamic systems approach to sports medicine: a narrative review. *Br J Sports Med* 2019;53:1214–20.
14 Jacobsson J, Bergin D, Timpka T, *et al*. Injuries in youth track and field are perceived to have multiple-level causes that call for ecological (holistic-developmental) interventions: A national sporting community perceptions and experiences. *Scandinavian Med Sci Sports* 2018;28:348–55.
15 Stephenson SD, Kocan JW, Vinod AV, *et al*. A Comprehensive Summary of Systematic Reviews on Sports Injury Prevention Strategies. *Orthop J Sports Med* 2021;9.
16 Hespanhol LC Jr, van Mechelen W, Verhagen E. Effectiveness of online tailored advice to prevent running-related injuries and promote preventive behaviour in Dutch trail runners: a pragmatic randomised controlled trial. *Br J Sports Med* 2018;52:851–8.
17 Lauersen JB, Andersen TE, Andersen LB. Strength training as superior, dose-dependent and safe prevention of acute and overuse sports injuries: a systematic review, qualitative analysis and meta-analysis. *Br J Sports Med* 2018;52:1557–63.
18 Gledhill A, Forsdyke D, Murray E. Psychological interventions used to reduce sports injuries: a systematic review of real-world effectiveness. *Br J Sports Med* 2018;52:967–71.
19 Ericksen S, Dover G, DeMont R. Psychological Interventions Can Reduce Injury Risk in Athletes: A Critically Appraised Topic. *J Sport Rehabil* 2022;31:224–9.
20 Dandrieux P-E, Navarro L, Blanco D, *et al*. Relationship between a daily injury risk estimation feedback (I-REF) based on machine learning techniques and actual injury risk in athletics (track and field): protocol for a prospective cohort study over an athletics season. *BMJ Open* 2023;13:e069423.
21 Dandrieux P-E, Navarro L, Blanco D, *et al*. Association between the use of daily injury risk estimation feedback (I-REF) based on machine learning techniques and injuries in athletics (track and field): results of a prospective cohort study over an athletics season. *BMJ Open Sport Exerc Med* 2025;11:e002331.
22 Bahr R, Clarsen B, Derman W, *et al*. International Olympic Committee consensus statement: methods for recording and reporting of epidemiological data on injury and illness in sport 2020 (including STROBE Extension for Sport Injury and Illness Surveillance (STROBE-SIIS)). *Br J Sports Med* 2020;54:372–89.
23 Cruz Rivera S, Liu X, Chan A-W, *et al*. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Lancet Digit Health* 2020;2:e549–60.
24 Edouard P, Ruffault A, Bolling C, *et al*. French Athletics Stakeholders' Perceptions of Relevance and Expectations on Injury Prevention. *Int J Sports Med* 2022;43:1052–60.
25 Edouard P, Dandrieux PE, Tondut J, *et al*. Injury Risk Reduction Perceptions in Athletics: Survey on Elite Athletes and Stakeholders Participating at the Munich 2022 European Championships. *Deutsche Zeitschrift Für Sportmedizin* 2023;74:194–204.
26 Gouttebarge V, Bindra A, Blauwet C, *et al*. International Olympic Committee (IOC) Sport Mental Health Assessment Tool 1 (SMHAT-1) and Sport Mental Health Recognition Tool 1 (SMHRT-1): towards better support of athletes' mental health. *Br J Sports Med* 2021;55:30–7.
27 Ruffault A, Sorg M, Martin S, *et al*. Determinants of the adoption of injury risk reduction programmes in athletics (track and field): an online survey of 7715 French athletes. *Br J Sports Med* 2022;56:499–505.
28 Edouard P, Navarro L, Branco P, *et al*. Injury frequency and characteristics (location, type, cause and severity) differed significantly among athletics ('track and field') disciplines during 14 international championships (2007-2018): implications for medical service planning. *Br J Sports Med* 2020;54:159–67.
29 Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems; 2017. Available: http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf
30 Collins GS, Dhiman P, Ma J, *et al*. Evaluation of clinical prediction models (part 1): from development to external validation. *BMJ* 2024;384:e074819.
31 Leckey C, van Dyk N, Doherty C, *et al*. Machine learning approaches to injury risk prediction in sport: a scoping review with evidence synthesis. *Br J Sports Med* 2024.
32 Pedregosa F, Varoquaux G, Gramfort A, *et al*. Scikit-learn: Machine Learning in Python. *J Mach Learn Res* 2011;12:2825–30.
33 Timpka T, Alonso J-M, Jacobsson J, *et al*. Injury and illness definitions and data collection procedures for use in epidemiological studies in Athletics (track and field): consensus statement. *Br J Sports Med* 2014;48:483–90.

34 Iatropoulos S, Dandrieux PE, Navarro L, *et al*. The Dose–Response Relationship of an ExerciseRandomized Controlled Trial on Athletics (Track39 Follow. *Scand J Med Sci Sports* 2024.
35 van der Horst N, Smits D-W, Petersen J, *et al*. The preventive effect of the nordic hamstring exercise on hamstring injuries in amateur soccer players: a randomized controlled trial. *Am J Sports Med* 2015;43:1316–23.
36 Harøy J, Clarsen B, Wiger EG, *et al*. The Adductor Strengthening Programme prevents groin problems among male football players: a cluster-randomised controlled trial. *Br J Sports Med* 2019;53:150–7.
37 Ullah S, Gabbett TJ, Finch CF. Statistical modelling for recurrent events: an application to sports injuries. *Br J Sports Med* 2014;48:1287–93.
38 Yang W, Jepson C, Xie D, *et al*. Statistical Methods for Recurrent Event Analysis in Cohort Studies of CKD. *Clin J Am Soc Nephrol* 2017;12:2066–73.
39 Frome EL. The analysis of rates using Poisson regression models. *Biometrics* 1983;39:665–74.
40 Andersson SH, Bahr R, Clarsen B, *et al*. Preventing overuse shoulder injuries among throwing athletes: a cluster-randomised controlled trial in 660 elite handball players. *Br J Sports Med* 2017;51:1073–80.
41 Edouard P, Dandrieux P, Blanco D, *et al*. How do sports injury epidemiological outcomes vary depending on athletes' response rates to a weekly online questionnaire? An analysis of 39-week follow-up from 391 athletics (track and field) athletes. *Scandinavian Med Sci Sports* 2024;34.
42 Claggett B, Pocock S, Wei LJ, *et al*. Comparison of Time-to-First Event and Recurrent-Event Methods in Randomized Clinical Trials. *Circulation* 2018;138:570–7.
43 Mazeas A, Duclos M, Pereira B, *et al*. Evaluating the Effectiveness of Gamification on Physical Activity: Systematic Review and Meta-analysis of Randomized Controlled Trials. *J Med Internet Res* 2022;24:e26779.
44 Liu X, Rivera SC, Moher D, *et al*. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI Extension. *BMJ* 2020;m3164:m3164.