APPENDICES

Appendix Table 1. Selection criteria and search terms using the PICOS framework.

PICOS	Selection Criteria		Search terms
Population	Female football players of any age and playing at any level	#1	Female OR women OR girls
		#2	Athlete OR sport OR player OR players
Intervention	Any type of injury prevention program	#3	Prevent OR prevention OR preventing OR injury OR injuries OR risk reduction OR program OR programme OR intervention
		#4	neuromuscular OR exercise OR safe OR prophylactic OR FIFA OR balance training OR strength OR proprioceptive OR movement training OR conditioning OR warm-up OR warm up OR therapy OR load OR nutrition OR diet OR sleep OR educate OR education OR advice OR strategy OR strategies OR wellness OR psychological OR cognitive OR bracing OR strapping OR taping
Comparison	Control group with 'usual training' or no intervention		
Outcome	Injury incidence, risk or burden	#5	Injury rate OR injury risk OR incidence OR burden OR prevalence
Study Design	Randomised controlled trials	#6	Randomised controlled trials OR randomized controlled trials OR RCT OR randomized OR randomised OR controlled trials

Note: Databases searched as (#1 AND #2) AND (#3 AND #4) AND #5 AND #6. These exact same search terms were used across all databases, with 'all fields' searched for AMED, CINAHL, EMBASE, and SPORTDiscus. The 'title, abstract, keywords' field was searched in Cochrane CENTRAL and Scopus, with the 'title/abstract' field searched in PubMed. The 'title' and 'topic' fields were searched in the core collection database of Web of Science. No other filters/restrictions/limiters (e.g. language, publication type, publication date) were placed on the search in any database. These search terms were not used in the PEDro database, instead three searches in 'clinical trials' were performed using the terms 'athlete injury', 'player injury', and 'sport injury'.

Appendix Table 2. Training component definitions for exercise-based interventions.

Training component	Definition
Agility	Activities aimed to promote the ability to move, cut, turn, or change direction quickly and effectively, while under control.
Balance	Activities that involved single- or double-legged stance activities specifically designed to challenge balance and proprioceptive awareness.
Mobility	Included any static or dynamic stretching/mobility activities.
Plyometric	Activities that utilised powerful dynamic movements, such as jumping, hopping, landing, or bounding.
Running	Any activity that was primarily running in a straight line, may be a warm-up, or warm-down, activity or interspersed with other components of the programme.
Strength	Any activity used to improve muscular capacity (either strength or endurance) through the use of resistance, including the use of bodyweight, free weights, bands, or machines.

Note: agility, balance, mobility, plyometric, and strength components were adapted from previous definitions.[38]

Appendix Table 3. Results of the methodological quality assessment using the Cochrane Collaboration Risk of Bias Tool (version 1).

	Selection bias		Performance bias	Detection bias	Attrition bias	Reporting bias	Other bias
Article	**Random sequence generation**	**Allocation concealment**	**Blinding of participants and personnel**	**Blinding of outcome assessment**	**Incomplete outcome data**	**Selective reporting**	
Barber Foss et al. (2018)[49]	Unclear	Low	Unclear	High	Unclear	Low	Unclear
Emery and Meeuwisse (2010)[50]	Unclear	Low	Unclear	Low	Unclear	Low	Unclear
Espinosa et al. (2015)[51]	Unclear	Unclear	High	Low	Low	Low	Low
Gilchrist et al. (2008)[52]	Unclear	High	High	High	High	Unclear	Low
Heidt Jr et al. (2000)[53]	Unclear	High	High	Low	Unclear	High	Unclear
LaBella et al. (2011)[54]	Low	Low	High	High	Low	Low	Low
Rössler et al. (2018)[55]	Low	Low	High	High	High	Low	Low
Söderman et al. (2000)[56]	Unclear	Low	High	High	High	Low	Unclear
Soligard et al. (2008)[57]	Unclear	Low	High	Unclear	Unclear	Low	Low
Steffen et al. (2008)[58]	Low	Low	High	Low	Low	Low	Unclear
Waldén et al. (2012)[59]	Low	Low	High	Low	Unclear	Low	Low
Zebis et al. (2018)[60]	Low	Low	High	Low	Low	Low	Unclear

Random sequence generation: selection bias due to inadequate generation of a randomised sequence; **Allocation concealment:** selection bias due to inadequate concealment of allocations prior to assignment; **Blinding of participants and personnel:** performance bias due to knowledge of the allocated interventions by participants and personnel during the study; **Blinding of outcome assessment:** detection bias due to knowledge of the allocated interventions by outcome assessors; **Incomplete outcome data:** attrition bias due to amount, nature or handling of incomplete outcome data; **Selective reporting:** reporting bias due to selective outcome reporting; **Other bias:** lack of reporting of or not accounting for poor adherence to the intervention program, or unequal balance of participants in intervention and control groups across different age levels.

Appendix Table 4. Summary of adherence to the intervention in each article for the intervention group for all studies.

Article	Adherence to the intervention	Education and resources provided to teams
Barber Foss et al. (2018)[49]	95% adherence (all sports included)	Not reported
Emery and Meeuwisse (2010)[50]	Warm-up completed at all practice and games[#]	Physiotherapist taught coaches at baseline and reviewed 3- and 6-weeks later. Provided supplementary written materials
Espinosa et al. (2015)[51]	80% of training sessions attended	Not reported
Gilchrist et al. (2008)[52]	Average of 25.8 sessions out of 36 (range of 12 to 37)	Teams provided with video and supplemental written materials
Heidt Jr et al. (2000)[53]	Not reported	Not reported
LaBella et al. (2011)[54]	Team adherence: 1425 of 1773 practices (80.4%)	Coaches had 2hr course and received a DVD, laminated card, and printed educational materials.
Rössler et al. (2018)[55]	Completed sessions per week ranged from 0.3 to 2.9 (all participants including males and females)	Study assistants visited clubs and gave coaches instruction session, detailed manual, and summary.
Söderman et al. (2000)[56]	Average of 65±19 sessions (range of 36 to 97)	Provided supplementary written materials
Soligard et al. (2008)[57]	Average of 44±22 (77%) sessions (range of 11 to 104)	Coaches and team captains had a 3hr course. Coaches and all players received a poster. Coaches received a DVD, exercise book, and small exercise cards. Coaches regularly contacted.
Steffen et al. (2008)[58]	Average of 23±9 (52%) sessions (range of 2 to 42)	All coaches and players received detailed brochure. Instructors visited the teams 3x/week initially, plus a booster after the summer break. Coaches regularly contacted.

Waldén et al. (2012)[59] *	Team adherence: 5244 of 6632 training sessions (79%)	Coach and player from each team had session with study therapist. Coach received CD-ROM and leaflet.
Zebis et al. (2018)[60]	93.2% and 84.2% of 159/332 players reported using the football during training and matches respectively.	Not reported

Note: *, data retrieved from the secondary analysis of this study published by Hägglund, et al. [105]. [#], the authors report that they cannot be certain that all components of the program were completed for all sessions, and the data for the home-based balance program only had <15% of players returning self-report adherence, with a median of 23 sessions (range of 2 to 26).

105. Hägglund M, Atroshi I, Wagner P, et al. Superior compliance with a neuromuscular training programme is associated with fewer ACL injuries and fewer acute knee injuries in female adolescent football players: Secondary analysis of an RCT. Br J Sports Med 2013;47:974-79.

Appendix Table 5. Grades of Recommendation, Assessment, Development and Evaluation (GRADE) quality of evidence.

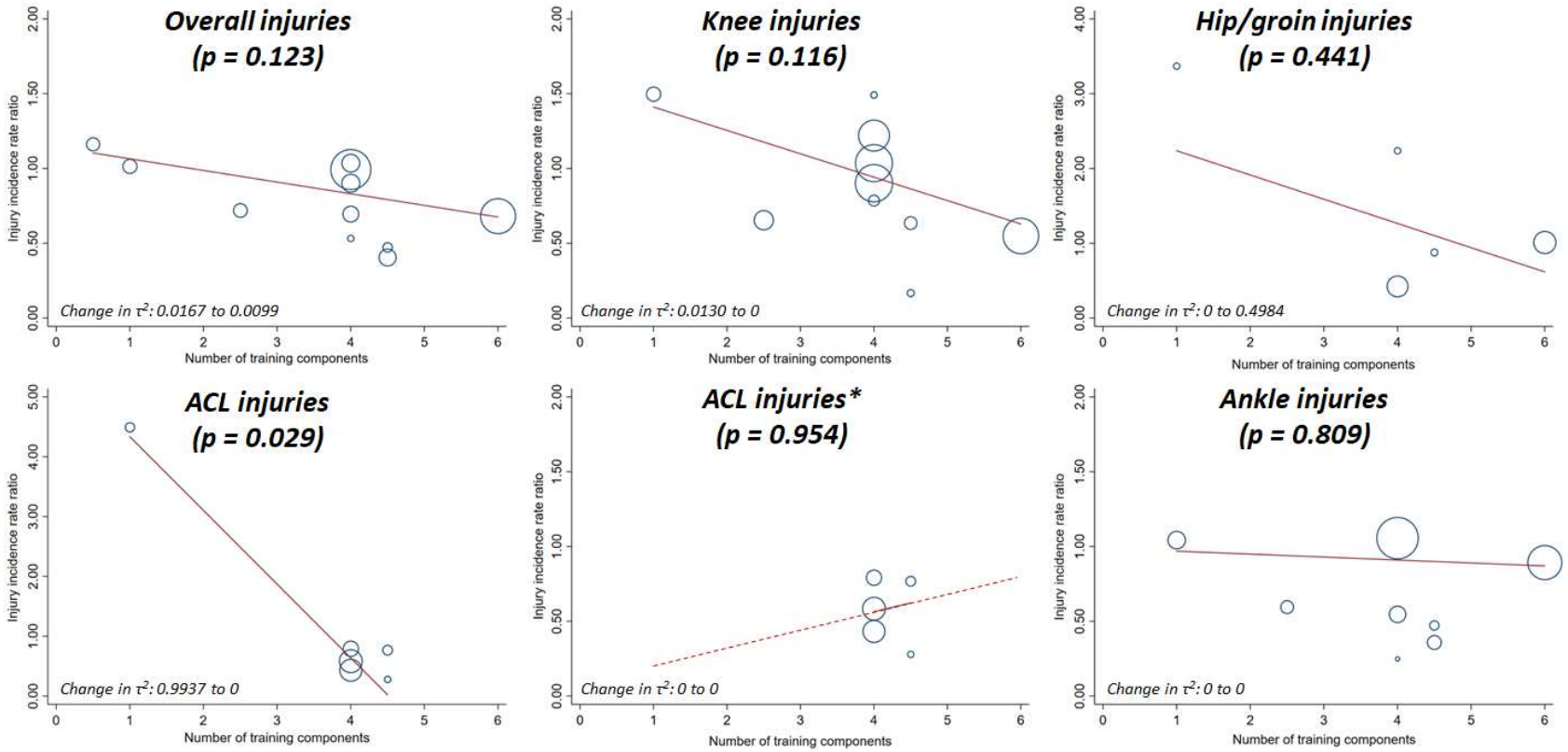
Meta-analysis outcome	*Number of RCTs*	*Risk of bias*	*Inconsistency*	*Indirectness*	*Imprecision*	*Publication bias*	GRADE quality
Overall injuries							
All studies: IRR 0.78, 95%CI 0.64 to 0.95, I^2 = 60.3%	11	-1 [a]	-1 [b]	0	0	0	Low
Multi-component studies: IRR 0.73, 95%CI 0.59 to 0.91, I^2 = 64.7%	9	-1 [a]	-1 [b]	0	0	0	
ACL injuries							
All studies: IRR 0.62, 95% CI 0.37 to 1.05, I^2 = 0%	6	-1 [a]	0	0	-1 [c]	0	Low
Multi-component studies: IRR 0.55, 95% CI 0.32 to 0.92, I^2 = 0%	5	-1 [a]	0	0	-1 [c]	0	
Knee injuries							
All studies: IRR 0.85, 95% CI 0.67 to 1.09, I^2 = 21.9%	10	-1 [a]	0	0	-1 [c]	0	Low
Multi-component studies: IRR 0.83, 95% CI 0.65 to 1.06, I^2 = 23.2%	9	-1 [a]	0	0	-1 [c]	0	
Ankle injuries							
All studies: IRR 0.83, 95% CI 0.65 to 1.07, I^2 = 13.1%	8	-1 [a]	0	0	-1 [c]	0	Low
Multi-component studies: IRR 0.78, 95% CI 0.58 to 1.05, I^2 = 23.2%	7	-1 [a]	0	0	-1 [c]	0	
Hip/groin injuries							
All studies: IRR 0.75, 95% CI 0.41 to 1.40, I^2 = 0%	5	-1 [a]	0	0	-1 [c]	0	Low
Multi-component studies: IRR 0.71, 95% CI 0.38 to 1.33, I^2 = 0%	4	-1 [a]	0	0	-1 [c]	0	
Hamstring injuries							
All studies: IRR 0.40, 95% CI 0.17 to 0.95, I^2 = 0%	4	-1 [a]	0	0	-1 [c]	0	Low
Multi-component studies: IRR 0.60, 95% CI 0.21 to 1.71, I^2 = 0%	2	-1 [a]	0	0	-1 [c]	0	

ACL, anterior cruciate ligament; IRR, injury incidence rate ratio; CI, confidence interval.

[a] Downgraded as the majority of trials scored an overall high risk of bias on the Cochrane Collaboration Risk of Bias Tool.

[b] Downgraded as there was greater than low levels of statistical heterogeneity between studies (I^2 > 40%).

[c] Downgraded as the clinical course of action would differ if the upper versus lower confidence interval represented the truth; or downgraded as the upper and lower confidence intervals had >0.5 difference.



Appendix Figure 6. Meta-regressions examining the relationship between the number of exercise-based training components and injury incidence rate ratio.

Note: The size of each bubble is proportional to the weight of each study. A reduction in tau squared (τ^2) may indicate the ability of the covariate (number of training components) to explain the heterogeneity in the injury incidence rate ratio. *, when removing the study by Söderman, et al. [56] from the ACL injury meta-regression, the relationship became non-significant, indicating this study had a large effect on the relationship, most likely due to the large injury incidence rate ratio of 4.49 that skewed the results of the analysis.

Appendix Table 7. Injury incidence for women's football, pooling control group data.

	Injury incidence per 1000 exposure hours - all studies	Injury incidence per 1000 exposure hours - adolescent teams
Overall	3.42 (3.19 to 3.67)	3.39 (3.16 to 3.65)
Ankle	0.97 (0.85 to 1.11)	0.95 (0.83 to 1.09)
Knee	0.57 (0.51 to 0.64)	0.58 (0.50 to 0.66)
Hamstring	0.22 (0.15 to 0.32)	0.14 (0.08 to 0.23)
Hip/groin	0.15 (0.10 to 0.21)	0.15 (0.11 to 0.22)
ACL	0.12 (0.09 to 0.16)	0.10 (0.07 to 0.14)

Note: Data reported as injury incidence per 1000 exposure hours, with the 95% confidence intervals in parentheses. Control group data from Table 3 in the manuscript were used to calculate injury incidence per 1000 exposure hours for studies that reported each injury (e.g. studies by Gilchrist, et al. [52] and Waldén, et al. [59] were not included in the overall incidence as they only reported knee and ACL injuries). Exposure hours were calculated from athletic exposures in four studies (1 athletic exposure equivalent to 2 exposure hours).[49,52-54] Training and match injuries were combined due to a lack of studies reporting these separately. Two studies examined senior teams (which also were the one training component studies) and one study examine college teams (average age >18 years), and were removed from the reporting in the right hand column of adolescent teams (<18 years).[51,52,56]