5

The impact of compliance on sports injury prevention effect estimates in randomized controlled trials: looking beyond the CONSORT statement

Hupperets MDW, Verhagen EALM, Finch CF, van Mechelen W. The impact of compliance on sports injury prevention effect estimates in randomized controlled trials: looking beyond the CONSORT statement. Submitted

ABSTRACT

Objectives The level of compliance with a trialled intervention can influence estimates of its effectiveness. This paper investigates the difference in outcome effect of a proprioceptive training programme aimed at preventing ankle sprain recurrences when analysed by means of intention-to-treat (ITT) versus per-protocol (PP) principles.

Methods 522 male and female athletes, with a recent lateral ankle sprain, were followed prospectively for one year. Athletes were randomized after usual care treatment to an intervention group (256 athletes) or a control group (266 athletes). Intervention athletes followed a prescribed proprioceptive training programme; control athletes only underwent usual care. Self-rated compliance with the training programme was recorded. Clinical outcome was ankle sprain incidence. Separate Cox regression analyses were performed according to the ITT and the PP approaches.

Results Twenty-three percent of the intervention group indicated to have fully complied with the proprioceptive training programme, whereas 35% was classified as not compliant. Compliance to the training programme was unknown for 13% of the intervention group. Significantly fewer recurrent ankle sprains were found in the fully compliant group compared to the group that was not compliant (relative risk 0.63; 95% confidence interval 0.43 to 0.99). Significantly fewer recurrent ankle sprains were found in the intervention group compared to the control group when analysed according to the ITT principle (0.63; 0.45 to 0.88). A PP analysis showed a stronger effect in favour of the intervention group (0.19; 0.07 to 0.53).

Conclusions A PP analysis on fully compliant athletes versus control group athletes showed that ankle sprain recurrence risk was over threefold higher in favour of the intervention group compared to when ITT analysis was utilized for the complete intervention group.

Trial registration ISTRCN34177180

INTRODUCTION

There has been increasing recognition of the need to formally evaluate the preventive capabilities of a range of sports injury preventive measures. The sequence of prevention[1] has been widely used in sports injury prevention research. More recently, the model has been extended to emphasise the need for both efficacy and effectiveness studies as well as an increased focus on actual behaviours within the context of implementing preventive measures.[2] From a theoretical and quality evidence perspective, the best way to evaluate the effect of any preventive measure is by performing a randomized controlled trial (RCT).[3] Although the RCT is considered the gold standard in research[4], there are some important considerations and limitations when the methodology is applied to sports injury prevention interventions. First, usually RCTs in a clinical setting involve diseased or injured people being allocated to the respective study groups and the outcome of interest is recovery, rehabilitation or reduction of the severity of symptoms of the disorder. In the sports injury prevention context the study population is presumed healthy or injury-free at baseline and the outcome of interest is becoming injured, with the expectation of there being fewer injured athletes in the intervention group(s). This requires an adequate number of injured cases in all study groups to be able to confidently test hypotheses involving comparisons between groups. Given that injury is largely a rare event and large amounts of exposure time are needed to observe a sufficient number of injuries, this implies that large cohorts and a relative long follow-up are needed for injury prevention studies.

Second, it is not always easy to blind the trial participants to the intervention being trialled. For example, one RCT required players to wear protective headgear or mouthguards[5], highly visible both to the user and any observer. Similarly for trials of new exercise modalities to be adopted during warm-up sessions it is often obvious who is and is not undertaking the intervention protocol.[6] Unfortunately, this limitation means that the

methodological quality of many studies can be compromised, when studies are assessed against the CONSORT guidelines[3], which give most weight to double-blinded RCTs. In order for RCTs to be adequately reported and to alleviate the problems arising from inadequate reporting, the CONSORT group developed the CONSORT statement.[3]

Third, there is existing adoption of many sports safety measures to varying degrees already in athlete populations. This means that getting a true control group may not be possible, especially when there is good evidence of a likely positive benefit. In the aforementioned protective equipment trial[5] it was not possible to have a true control group where the football players were not permitted to use any headgear or mouthguards. For ethical reasons this study was confined to using a control group that corresponded to "usual behaviours" on the part of the players randomized to that group.[7] This would have the effect of diluting any real group differences.

Finally, most sports injury prevention RCTs require athletes to adopt some form of behaviour change for injury reduction benefits. The long follow-up time that is required to demonstrate such behaviour change can result in significant loss to follow-up.[8-10] Another important consideration that determines outcome of a RCT is compliance to the intervention or the required behaviour change. Compliance in sports injury prevention research is a term used to indicate the athlete's correct following of the prescribed intervention. Large cohorts are harder to control and it can be difficult to directly influence the behaviours of individuals. This can lead to compromised compliance, as there is increased dependence on the athletes' own motivation levels. Sports medicine research can be used to exemplify the importance of intervention compliance on RCT outcome. For example, in a variety of sports, multiple studies have studied the effectiveness of exercise training programmes for the prevention of injuries. Many such studies showed that these training programmes can prevent injuries.[11-16] In contrast, two recent studies have failed to show a

preventive effect, but did report suffering from poor compliance with the prescribed training programme.[17,18] The impact on outcome of 'non-compliance' with specific interventions on injury reductions has also been highlighted in other RCTs of lower limb injury prevention measures in soccer[19-21] and for protective equipment in Australian Football.[5]

With respect to analysis of RCT study results, the CONSORT statement favours intention-to-treat (ITT).[22] By using an ITT approach, RCTs aim to remove potential for selection bias and placebo type effects influencing the conclusions. However, poor compliance and adherence to the intervention tend to dilute treatment effects when analyses go by the ITT approach.[23] A different approach that has a stronger focus on the efficacy of an intervention is analysis by treatment actually received, i.e. per-protocol analysis (PP).[24] By solely analysing compliers to the intervention the maximal achievable effect of the intervention is shown. Recently, studies on protective equipment in football[7,25] and studies on injury prevention in soccer[17,18] have adopted the PP approach.

The aim of the present study was to investigate the difference in outcome effect of a proprioceptive training programme aimed at preventing ankle sprain recurrences[26] when analysed by means of ITT versus PP analysis.

MATERIALS AND METHODS

A complete overview of the study design can be found in Hupperets et al.[26]

Population and randomization

A total of 522 athletes between 12 and 70 years, who were actively participating in sports, and who had recently suffered an acute lateral ankle sprain were included in the study. Athletes were randomized to the intervention group or the control group, with stratification for gender, source of enrolment (emergency room, general practice, physical therapy office, media), and ankle sprain treatment (non-medical treatment, paramedical treatment, intramural treatment, or extramural treatment).

This resulted in intervention and control groups consisting of 256 and 266 athletes, respectively.

Intervention

The intervention consisted of an eight-week proprioceptive training programme, of which a detailed description is given elsewhere.[26]

Athletes of the intervention group received a balance board (Avanco AB, Sweden), and general written and visual (DVD) information on the programme. All information was also provided on a dedicated website, only accessible for subjects of the intervention group.

The programme prescribed three training sessions per week, with a maximal duration of 30 minutes per session. Exercises gradually increased in difficulty and training load during the eight-week programme. All athletes trained individually, without supervision of a coach or medical practitioner.

Data collection

During the one-year follow-up, athletes reported all sudden inversions of the same ankle and details of their sports participation for each training session and match on a monthly basis. Athletes who reported an ankle sprain recurrence completed a web-based questionnaire. Furthermore, athletes who had sustained an ankle sprain recurrence received a cost-diary, which registered all health care costs and costs due to productivity loss from the moment of injury until full recovery.

During the eight-week proprioceptive training programme, the intervention group received two questionnaires, after four and eight weeks, which consisted of statements on the self-rated compliance with the proprioceptive training programme. Statements were answered on a five-point Likert Scale, ranging from complete disagreement (score=1) to complete agreement (score=5) with the statement. The following statement was used to measure compliance: “The last four weeks I

performed the exercises of the programme as prescribed". Athletes rating the statement with a four or a five on both occasions (i.e. after four and eight weeks) were considered fully compliant. The term partially compliant was used for athletes who scored a four or a five on the Likert Scale in only one out of the two occasions. Athletes who gave scores of three or lower on both occasions were considered not to have complied with the prescribed training programme. These three categories (i.e. fully compliant, partially compliant, and not compliant) were chosen arbitrarily, but differentiate between athletes who did comply with the programme and those who did not. To rule out spill over or contamination between intervention and control group athletes, control group athletes were asked monthly whether they had participated in proprioceptive training during the past month.

Outcome measures

Compliance numbers were presented as the absolute number of athletes (and percentages) in each category. Control group athletes who performed some form of proprioceptive training during the one-year follow-up were presented as a total number and a percentage of the total number of athletes in that group. In the analyses, it was assumed that 100% of the controls 'complied' with their usual treatment regimens, which were expected not to include the exercises used by the intervention group, as these are not part of usual treatment.

Clinical outcome was ankle sprain incidence and its 95% confidence interval (CI), calculated as the number of incident ankle sprains per 1,000 hours of exposure. Follow-up of athletes ceased after injury recurrence.

Analysis

Cox regression analysis was used to compare ankle sprain recurrence risk within the intervention group, using compliance as an independent variable. It was decided a priori to adjust for age, type of sport, and level

of sports. Other variables were checked for confounding and/or interaction, but none where found. Differences were considered statistically significant at $P<0.05$.

Further Cox regression analyses on overall effects were performed from an ITT as well as a PP perspective. In the ITT analysis both groups were compared without taking into account compliance with the prescribed treatment. In contrast, PP analysis was performed only on athletes who were fully compliant to the entire intervention.

RESULTS

Compliance

Compliance numbers are presented in Table 1. A total of 58 (23%) athletes in the intervention group indicated to have fully complied with the eight-week proprioceptive training programme; 75 (29%) intervention group athletes indicated to have been partially compliant with the training programme; 89 (35%) responded in such a way that they were classified as not compliant. Compliance to the training programme was unknown for 34 (13%) athletes, since they did not complete the questionnaires. Five out of 266 control group athletes (2%) reported to have performed some sort of proprioceptive training exercises during the one-year follow-up. These athletes performed proprioceptive training exercises as part of medical treatment of an ankle sprain recurrence and were not incorporated in the analysis.

Cox regression analysis adjusted for age, type of sport, and level of sports showed that the risk for an ankle sprain recurrence was significantly lower in the fully compliant group compared to the group that was not compliant (relative risk 0.63; 95% confidence interval 0.43 to 0.99).

Table 1 Compliance data (percentage) and ankle sprain incidences (95% CI) for the intervention group.

	n	Incidence
Full compliance	58 (23%)	**0.52** (0.01 to 1.04)
Partial compliance	75 (29%)	**2.02** (1.16 to 2.89)
No compliance	89 (35%)	**3.14** (2.04 to 4.25)

A total of 34 (13%) athletes did not complete the questionnaire and were omitted from the analysis

Ankle sprain incidence

The overall ankle sprain incidences in the intervention and control group were respectively 1.86 per 1,000 hours of sports (95% confidence interval 1.37 to 2.34), and 2.90 per 1,000 hours of sports (2.30 to 3.50). Cox regression analysis according to the ITT principle showed that the risk for ankle sprain recurrences was significantly lower in the intervention group compared to the control group (0.63; 0.45 to 0.88).

A PP analysis on ankle sprain recurrences puts this treatment effect in another perspective. Ankle sprain recurrence incidence for fully compliant athletes was 0.52 (0.01 to 1.04), whereas non-compliant athletes showed an incidence of 3.14 (2.04 to 4.25) (Table 1). Cox regression analysis showed that the risk for ankle sprain recurrences was significantly lower in compliant intervention athletes compared to control group athletes (0.19; 0.07 to 0.53).

DISCUSSION

Although the present RCT suffered from relatively poor levels of full compliance to the intervention, a preventive effect on ankle sprain recurrences of proprioceptive training was found. The present study

showed that with a PP analysis on fully compliant athletes versus control group athletes, ankle sprain recurrence risk was over threefold higher in favour of the intervention group compared to when ITT analysis was utilized for the complete intervention group. No significant differences in baseline characteristics were found between complete intervention group and complete control group, between fully compliant intervention athletes compared to complete control group, and between fully compliant and non-compliant athletes (data not shown).

A PP analysis is justified if a preventive measure is shown to be applicable in preventing injuries to further investigate which intervention component creates what effect. The present study on the effectiveness on ankle sprain recurrences of proprioceptive training is not the first to establish this effect. Research has shown that a proprioceptive training programme is effective in reducing ankle sprain recurrences[11-16], therefore compliance to the programme is a key to establishing proper results. A recent RCT on wearing padded headgear in rugby[25] illustrated that another reason to adopt a PP analysis is when an intervention is critical to injury reduction. By not wearing headgear, there is no physical barrier between the rugby player and the ground or another player during a collision and the transfer of energy to the head cannot be prevented. This applies to all sports injury prevention studies in which a 'material' intervention is utilized.

Methodological considerations

The benefit of the ITT principle is that it gives a pragmatic estimate of the benefit of a change in treatment policy[27] and that it minimises selection bias.[24] Furthermore, by including all study participants in the analysis, power of the trial is improved. However, despite these benefits and despite the ITT principle being widely used in injury preventive RCTs, this method has some disadvantages. It tends to dilute the treatment effect and leads to a loss of power compared with a situation of full compliance of a total population. The problem is that the ITT analysis includes people

who did not apply the intervention at all. Conversely, perhaps the ITT principle is a sensible approach, since it is best to be conservative when publishing effect sizes.

Although ITT analysis removes or adjusts for a placebo effect, for some injury interventions, such as wearing protective equipment, a placebo effect cannot be possible. If people do not wear protective equipment, they can never be prevented from sustaining significant impact, because there is no physical barrier to absorb the energy exchange.

The PP analysis requires strong assumptions about the comparability at baseline of compliers and non-compliers within different randomized groups and is known to be subject to selection bias. It should be noted that this same bias may be introduced in ITT analysis when there is significant loss to follow-up.[23] Furthermore, PP restricts the analysis to a self-selected patient population, precluding a more generalisable assessment of therapy in larger patient groups. Since it is unlikely that all patients will comply to an intervention all of the time, ITT analysis reflects a pragmatic treatment effect in a 'standard' group of patients. The PP analysis restricts comparison of the treatments to ideal intervention patients, that is, those who adhered highly to the clinical trial instructions as stipulated in the protocol. Hence, a PP analysis is only possible if there is diversity in the intervention group regarding compliance. A PP analysis more clearly demonstrates the maximal achievable effect of an intervention. It attempts to measure efficacy and is sometimes used as a means of regaining some of the power lost.[24]

In the present study, analysis from an ITT perspective as well as from a PP perspective was possible and the opportunity was taken to directly compare the two approaches. It is possible that the adequacy of our measure of compliance could have contributed to the difference between the two analyses, but this is unlikely. It should be noted that compliance was measured through only one statement, which made it impossible to rule out social desirability bias. Nevertheless, athletes were encouraged to

answer questions honestly and they were made aware beforehand that answers would never be related to an individual athlete. Despite this shortcoming, a large effect on ankle sprain recurrences was shown for athletes who had reported to comply to the intervention. Occurrence of spill over or contamination between intervention and control groups is unlikely. It was shown that only 2% of the control group athletes performed some sort of proprioceptive training during the one-year follow-up. However, these athletes all endured an ankle sprain recurrence and underwent (para-)medical treatment for this re-injury. All five athletes indicated to have performed proprioceptive training exercises as part of this medical treatment. Thus, the reported proprioceptive training by these control athletes did not affect results as the outcome measure was injury incidence and the intervention ceased after injury recurrence.

Compliance is the term most commonly used in clinical RCTs, but in the sports injury behavioural studies other terms such as adoption[28], uptake[2], or sustainability[29] are used as well. In essence, these terms are often used as synonyms for compliance, since they cover the same construct. An intervention, whether it is wearing protective equipment or a training programme, is either 'used' by the athlete or not. This is particularly true in protective equipment interventions, but as was shown in the present study and in other exercise training programme interventions[17,18], there can be degrees of usage.

Assessment of compliance to an intervention programme and the subsequent reporting of compliance is of critical importance for RCTs. Since compliance is such a critical issue in reporting RCTs, there should be consensus on how to measure compliance and how to report it. It is therefore striking that the CONSORT guideline does not contain a section on compliance measurement and reporting. This has led to differing approaches on how to measure and report compliance.[8,17,18] Preferably, compliance to sports injury preventions is assessed by an independent person. For all RCTs of injury interventions, we strongly recommend that

compliance is assessed at least through multiple compliance questions and at different times points throughout the study. This would help to ascertain the accuracy of the information. One way to achieve this would be by requiring athletes complete a compliance log throughout the entire intervention period. A five-point Likert Scale as used in the present study could be used as a means of scoring the abovementioned questions. Differentiation between full compliance, partial compliance, and no compliance as presented in this study could be adopted in the reporting of compliance.

Low compliance can make statistical power too low to assess effects of the training programme in intervention groups through a PP analysis. A recent study on injury prevention in soccer suffered from this shortcoming.[17] Since many cohorts are already large and maximum sample size capacity is often reached, efforts should be made to minimize non-compliance. The importance of maximizing compliance is illustrated by the fact that the approximate sample size required for a given power is inversely proportional to the square of the compliance rate.[30] Strategies to increase compliance significantly need to be explored before the results of ITT studies can effectively be applied in the real world.

Considerations for the design of RCTs with >1 intervention

In this study, we have considered the specific case when the RCT comprised one intervention and one control group. The issues relating to compliance (or non-compliance) with interventions become even more complex when RCTs aim to compare more than one intervention group, relative to controls. There is potential for all data from such studies to be of poor quality due to unanticipated low compliance rates, so that no conclusions can be drawn at all about any of the interventions.

A major way to avoid this possibility is in the design phase of the trial and to adopt a factorial study design.[31] An example of the successful use of this design approach was a two intervention RCT aimed at determining the

effects of protective headgear and custom-fitted mouthguards for preventing orofacial injuries in Australian football players.[5] This study was designed as a four arm factorial trial: headgear only, mouthguards only, headgear+mouthguard, control (usual behaviour). Unfortunately, however, so few players wore the assigned headgear[7] that it was not possible to assess its effectiveness in this trial. Because of the factorial design, it was possible (without loss of power) to collapse the final study to a two arm factorial trial and to still formally test the hypotheses relating to the preventive potential of mouthguards.[5] This example also highlights the importance of only conducting RCTs of interventions which are likely to have good levels of acceptance, and hence compliance, by the trial participants.

CONCLUSIONS

Although the CONSORT statement[3] is widely used to guide the reporting of RCTs in many major journals, in the specific example of sports injury prevention research additional analysis strategies might also apply in addition to the favoured ITT. From the present study it can be concluded that differences in effect sizes can be found depending on the type of analysis that is chosen. In determining effect size of an intervention relative to control it is best to adopt the ITT principle, since an intervention is eventually applied to a general public, in which not every patient will be compliant to a preventive programme. Nevertheless, a PP analysis should be presented alongside ITT analysis to provide more insight into the effectiveness of the intervention in only those who adopt the intervention as per the protocol specifications. A PP analysis is only possible if measurement of compliance to the intervention programme is undertaken. Therefore, it is highly recommended that RCTs incorporate proper compliance measurement strategies.

REFERENCES

1. Van Mechelen W, Hlobil H, Kemper HCG. Incidence, severity, aetiology and prevention of sports injuries. Sports Med 1992; 14(2): 82-99
2. Finch C. A new framework for research leading to sports injury prevention. J Sci Med Sport 2006; 9: 3-9
3. Altman DG. Better reporting of randomised controlled trials: the CONSORT statement. BMJ 1996; 313: 570-1
4. Agabegi SS, Stern PJ. Bias in research. Am J Orthop 2008; 37(5): 242-8
5. Finch C, Braham R, McIntosh A, et al. Should Australian football players wear custom-made mouthguards? Results from a group-randomised controlled trial. Inj Prev 2005; 11: 242-6
6. Finch CF, Lloyd D, Elliott B. The preventing Australian football injuries with exercise (PAFIX) study – a group randomised controlled trial. Inj Prev. In Press.
7. Braham R, Finch C. Do community football players wear allocated protective equipment? Descriptive results from a randomised controlled trial. J Sci Med Sport 2004; 7: 216-20
8. Verhagen EALM, Van der Beek AJ, Twisk JWR, et al. The effect of a proprioceptive balance board training program for the prevention of ankle sprains: a prospective controlled trial. Am J Sports Med 2004; 32(6): 1385-93
9. Lun V, Meeuwisse WH, Stergiou P, et al. Relation between running injury and static lower limb alignment in recreational runners. Br J Sports Med 2004; 38: 576-80
10. Stevenson MR, Hamer P, Finch CF, et al. Sport, age, and sex specific incidence of sports injuries in Western Australia. Br J Sports Med 2000; 34: 188-94
11. Bahr R, Bahr IA. Incidence of acute volleyball injuries: a prospective cohort study of injury mechanisms and risk factors. Scand J Med Sci Sports 1997; 7: 166-71

12. Verhagen EALM, Van der Beek AJ, Bouter LM, et al. A One-Season Prospective Cohort Study of Volleyball Injuries. Br J Sports Med 2004: 38(4); 477-81
13. Stasinopoulos D. Comparison of three preventive methods in order to reduce the incidence of ankle inversion sprains among female volleyball players. Br J Sports Med 2004; 38(2): 182-5
14. Petersen W, Hohmann G, Pufe T, et al. A controlled perspective case control study of a prevention training program in female team handball players: the German experience. Arch Orthop Trauma Surg 2005; 125(9): 614-21
15. Wedderkopp N, Kaltoft M, Lundgaard B, et al. Prevention of injuries in young female players in European team handball. A prospective intervention study. Scand J Med Sci Sports 1999; 9(1): 41-7
16. Tropp H, Askling C, Gillquist J. Prevention of ankle sprains. Am J Sports Med 1985; 13(4): 259-62
17. Engebretsen AH, Myklebust G, Holme I, et al. Prevention of injuries among male soccer players: a prospective, randomized intervention study targeting players with previous injuries or reduced function. Am J Sports Med 2008; 36: 1052-60
18. Steffen K, Myklebust G, Olsen OE, et al. Preventing injuries in female youth football – a cluster-randomized controlled trial. Scand J Med Sci Sports 2008; 36(4): 700-8
19. Kilding A, Tunstall H, Kuzmic D. Suitability of FIFA's "the 11" training programme for young football players – impact on physical performance. Clin J Sports Med 2008; 7: 320-6
20. Reider B. A popular cult. Am J Sports Med 2008; 36: 1467
21. Croisier JL, Ganteaume S, Binet J, et al. Strength imbalances and prevention of hamstring injury in professional soccer players: a prospective study. Am J Sports Med 2008; 36: 1469-75
22. The CONSORT statement website: http://www.CONSORT-statement.org/index.aspx?o=1089. Visited at 12/12/08

23. Montori VM, Guyatt GH. Intention-to-treat principle. CMAJ 2001; 165(10): 1339-41
24. Becque T, White IR. Regaining power lost by non-compliance via full probability modeling. Statist Med 2008; 27: 5640-63
25. McIntosh AS, McCrory P, Finch CF, et al. Does padded headgear prevent head injury in rugby union football? Med Sci Sports Excerc; In press
26. Hupperets MDW, Verhagen EALM, Mechelen W van. The 2BFit study: is an unsupervised proprioceptive balance board training programme, given in addition to usual care, effective in preventing ankle sprain recurrences? Design of a Randomized Controlled Trial. BMC Musc Dis 2008; 9: 71
27. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ 1999; 319: 670-4
28. De Meij JS, Chinapaw MJ, Kremers SP, et al. Promoting physical activity in children: the stepwise development of the primary school-based JUMP-in intervention applying the RE-AIM evaluation framework. Br J Sports Med 2008; 19: epub ahead of print
29. Eime R, Finch C, Wolfe R, et al. The effectiveness of a squash eyewear promotion strategy. Br J Sports Med 2005; 39: 681-5
30. Lachin JM, Foulkes MA. Evaluation of sample size and power for analyses of survival with allowance for non-uniform patient entry, losses to follow-up, noncompliance and stratification. Biometrics 1986; 42: 507–19
31. Murray D. Design and analysis of group-randomized trials. New York: Oxford University Press, 1998