Knees, Comrades and sample size

To the Editor: We wish to raise concerns with regard to the study published by Hagemann et al. (‘Do knees survive the Comrades Marathon?’).1 The design of the study is purported to be a prospective study of 10 randomly selected participants. Closer reading shows the sampling to be that of a convenience sample in which participants volunteered for the study.

No mention is made of how potential participants were approached or, later, how many patients were excluded from the study owing to pre-existing injury. This sampling technique is not statistically random and introduces serious selection bias. Factors such as age, weight or whether it was an uphill or downhill race are also not considered.

The second point of concern is the very small sample size used in this study. Small sample sizes in medical studies are often a result of necessity, but there are inherent dangers in making use of them. Over the 6-year period (1997 - 2002), there were over 90 000 entrants in the Comrades Marathon.2 Using an alpha value of 0.05 and a 95% confidence level, based on a population of 90 000 entrants, the recommended sample size is 383. A sample size of only 10 introduces a 30.99% margin of error. This study then becomes an example of a type II error contravening the conclusion that there was no difference between the two groups.

The study was designed to determine the effect of ultra-marathon running on the structures of the normal knee and any unknown pre-existing abnormalities of the knee. In selecting participants, knees that had previous surgery or documented injuries were excluded and, therefore, none were excluded later on the basis of any such pre-existing injury. No other qualifications (such as age, weight or gender) were considered as conditions for eligibility. As part of the recruiting protocol, all participants were volunteers. We disagree that this selection of knees is non-random or introduces a selection bias. We did not provide information on the specifics of the race as this is readily available.3

As is often the case with prospective MRI studies, the small sample size was the result of necessity. However, in designing this study, we purposely restricted our aims to avoid the dangers inherent to this small size by limiting the study to normal knees (including by necessity those with unknown abnormalities) and by only registering changes on follow-up scans. Specifically, we disagree with the view that 383 participants would be required for this study to meet the conditions of alpha equal to 0.05 and a 95% confidence level, based on 90 000 entrants. Such numbers of participants would have to be recruited if the purpose of this study were a complete inventory of all injuries, new and pre-existing, collected and followed up over the course of the three sequential MRI studies. However, by limiting the aims of this study, we have been able to demonstrate convincingly that there was no difference between the two groups. Researchers of similar prospective MRI studies on the knees of runners (references 5, 6, 7, 10, 11) had enrolled between 5 and 10 participants to arrive at their conclusions.


Drs Hagemann, Rijke and Corr reply: We thank Drs Rodseth and Geddes for their comments on our study. They focus on two aspects of the study: (i) its design – specifically on how the participants were selected in this prospective study; and (ii) the small sample size, which would introduce a margin of error contravening the conclusion that there was no difference between the two groups.

The study was designed to determine the effect of ultra-marathon running on the structures of the normal knee and any unknown pre-existing abnormalities of the knee. In selecting participants, knees that had previous surgery or documented injuries were excluded and, therefore, none were excluded later on the basis of any such pre-existing injury. No other qualifications (such as age, weight or gender) were considered as conditions for eligibility. As part of the recruiting protocol, all participants were volunteers. We disagree that the selection of knees is non-random or introduces a selection bias. We did not provide information on the specifics of the race as this is readily available.

As is often the case with prospective MRI studies, the small sample size was the result of necessity. However, in designing this study, we purposely restricted our aims to avoid the dangers inherent to this small size by limiting the study to normal knees (including by necessity those with unknown abnormalities) and by only registering changes on follow-up scans. Specifically, we disagree with the view that 383 participants would be required for this study to meet the conditions of alpha equal to 0.05 and a 95% confidence level, based on 90 000 entrants. Such numbers of participants would have to be recruited if the purpose of this study were a complete inventory of all injuries, new and pre-existing, collected and followed up over the course of the three sequential MRI studies. However, by limiting the aims of this study, we have been able to demonstrate convincingly that there was no difference between the two groups. Researchers of similar prospective MRI studies on the knees of runners (references 5, 6, 7, 10, 11) had enrolled between 5 and 10 participants to arrive at their conclusions.

1. http://results.comrades.com