Correspondence

Hamstring tendon autograft better than bone–patellar tendon–bone autograft in ACL reconstruction—a cumulative meta-analysis and clinically relevant sensitivity analysis applied to a previously published analysis

Sir—We have read with interest the work by Poolman et al. (2007), which uses the data from our previously published meta-analysis (Biau et al. 2006) on the comparison of hamstring tendon autografts versus bone—patellar tendon—bone autografts for reconstruction of the anterior cruciate ligament. In their article the authors develop two ideas: the use of cumulative meta-analyses and the use of sensitivity analyses. Although the authors are undoubtedly willing to give readers more information on the topic we believe their conclusions may be flawed.

Cumulative meta-analyses have been developed to look for a trend in treatment effect based on a covariate. The classical funnel plot of the meta-analysis is therefore ordered by increment of the variable studied, for instance chronologically by date of publication as in the seminal report by Lau et al. (1992), or by duration of exercise therapy as in the article by Kwakkel et al. (2004). The interest of a cumulative meta-analysis ordered by date of publication, as indicated by the authors, is that it may show the time at which sufficient evidence of the superiority of one treatment over another has accumulated. At that point two conclusions may be drawn: first, randomised controlled trials conducted after that time have unnecessarily exposed participants to the risks and burden of medical research and should not have been performed, and second, there should have been a change in surgical practice. In their article, Poolman et al. technically showed that randomised controlled trials after the year 2001 did not change the conclusions regarding the outcome "anterior knee pain". However, we would like to notify that this result does not preclude the conduction of randomised controlled trials or of meta-analyses following that date. First, the basis for conducting randomised controlled trials in the context of anterior cruciate ligament reconstruction is not just anterior knee pain, but also knee stability. Despite improvements in surgical techniques investigators have not proved the superiority of the hamstring tendon autograft reconstruction as yet and many still believe that the bone—patellar tendon—bone yields better stability and consider it an attractive option. Therefore, randomized controlled trials are still necessary, and welcome, to clarify this issue and there are no reasons to believe that trials performed after the year 2001 were superfluous. Second, the evidence obtained from a meta-analysis is not definite nor should it be considered as doubtless, and surgeons preferring the bone—patellar tendon—bone autograft for reconstruction of the anterior cruciate ligament could argue that the evidence of the superiority of the hamstring tendon autograft with regard to anterior knee pain accumulated by the end of the year 2001 was not sufficient to change their practice. The quality of the trials available for a meta-analysis has a major influence on the strength of evidence a meta-analysis yields, and it is believed by some researchers that the quality of a meta-analysis is that of the worst trial included. By the year 2001, none of the seven randomised controlled trials reported had a high quality score and readers could have doubted the conclusions of any

DOI 10.1080/17453670710014455

meta-analysis conducted at that time, and could still do at present. Moreover, the rationale for initiating a meta-analysis at a time 't' is only based on the available evidence at that specific time. Whether or not it is the best time, we don't control time and it makes no difference that enough evidence could have been accumulated earlier if it had not been reported. Last, the authors wrongly conclude that their "cumulative meta-analysis strengthens the evidence for reduced morbidity using hamstring tendon autograft": a cumulative meta-analysis does not give more evidence than a standard meta-analysis including the same trials, and it yields exactly the same results. However, because successive tests are being performed, a cumulative meta-analysis conveys an increased risk for type I error, and the authors have not accounted for this important statistical issue.

The authors also stated that we had not used sensitivity analyses optimally and they compared the treatment effects of studies using an Endobutton for femoral fixation to studies using other types of fixation. Based on the fact that the treatment effect was not significant in the former group and significant in favour of the bone—patellar tendon—bone in the latter group of studies, the authors concluded that "sensitivity analysis focusing on state-of-the-art hamstring graft fixation techniques further weakens the evidence that bone—patellar tendon—bone autografts provide better stability". However, there are many reasons to believe their conclusions are flawed. First, in the trials included in our meta-analysis patients were not randomised upon the type of fixation employed for the reconstruction procedure but on the type of autografts used, and given the limited trust one should have on the results of subgroup analyses (Brookes et al. 2001), as a rule, the conclusion of a trial, and even more of a meta-analysis, should not be upon subgroup findings. Second, Poolman et al. rationale for choosing this subgroup was adapted from the conclusions of another previously published meta-analysis, namely the work by Prodomos et al. (2005). However, this meta-analysis does not give enough methodological certainty to support their rationale and was indeed recently rated by the same authors as one the poorest meta-analysis ever published in the domain (Poolman et al. 2007). Third, another reason we did not perform subgroup analyses based on the type of fixation is that we had no way to differentiate between fixation effect and other effects such as centre effect, period effect, patients characteristics, etc, and we decided that the conclusions we would obtain would not be supported enough. Therefore, we did not plan these analyses a priori and having no statistical heterogeneity there was no reason to perform subgroup analyses a posteriori. Poolman et al. stated that although we "did not find statistical heterogeneity, clinically relevant differences between studies did skew the results". It is true that our results were possibly skewed but despite the limited place we were allowed in the discussion we did state that "technical issues (such as cycling of the graft, degree of knee flexion and graft tension when securing the graft, and fixation devices) cannot be analysed in a meta-analysis based on aggregate patient data, and ideally data from individual patients should be analysed". Fourth, and more importantly, the authors wrongly conclude that, because the treatment effect in the group of studies using other types of fixation than the Endobutton is significantly in favour of the bone—patellar tendon—bone autograft when it is not significantly different in the group of studies using Endobutton, "state-of-the-art hamstring graft fixation techniques further weakens the evidence that bone—patellar tendon—bone autografts provide better stability". However, the authors have not used interaction tests as we did in the original manuscript, and their conclusion is not statistically founded. Based on the interaction test by Gail and Simon (1985) there is no evidence for variation in the treatment effect upon the fixation technique: the P-value is 0.34, namely, assuming there is no effect of the fixation technique on the treatment effect, the authors had 34% chances to observe the results they obtained or more extreme results. Therefore, the authors have failed to give further evidence that bone—patellar tendon—bone autografts provide better stability. Finally, aggregate patient data meta-analyses have shown limitations and ideally meta-analyses on patient level data should be preferred to carry out detailed data checking, adjust for potential confounding factors, and better estimate the effect of treatment (Stewart and Parmar 1993). As a matter of fact, we have recently submitted a meta-analysis based on patient level data of seven randomised clinical trials where we found

that the bone – patellar tendon – bone autograft provides better stability than the hamstring tendon autograft and that it remains an attractive option for reconstruction of the anterior cruciate ligament reconstruction (Biau et al. 2007).

In conclusion, the more meta-analyses are performed, the more scepticism they encounter among the medical community and because with the development of statistical softwares it has become easier to perform analyses, we should all make sure that the results and conclusions we present are strongly supported by sound statistics and methodology. As the authors rightfully remind us at the end: "data dredging can lead to flawed conclusions and can mislead readers".

David Jean Biau, Sandrine Katsahian, and Rémy Nizard
Département de Biostatistique et Informatique Médicale (DBIM), Hôpital Saint-Louis, Assistance Publique-Hôpitaux de Paris, 1, avenue Claude Cellefaux, FR-75475 Paris, France
djmbiau@yahoo.fr

Sir—We are glad to see our provocative manuscript raised discussion. We thank Biau et al. for their critical remarks and are happy to clarify the issues raised. Unfortunately Biau et al. missed the purpose of our manuscript: acknowledging feasibility of subgroup analysis based on a priori hypothesis. Thus, we did not embark on a data fishing expedition, as well, we warn readers for a sensible use of sensitivity analysis. We merely used Biau and co-workers' work as a compelling example. We choose two illustrations helpful to exemplify our aim: cumulative meta-analysis and a subgroup analysis based on surgical technique.

Cumulative meta-analysis

Our cumulative meta-analysis clearly shows a trend, this trend only illustrates the sample size needed to reach conclusions and does not undermine the need for new RCTs finally clarifying optimal graft choice for ACL reconstruction based on other, clinically relevant outcomes. Clearly, with accumulating numbers anterior knee pain became more evident in the Bone Patellar Tendon Bone group. Although we argue that statistical correction for multiple outcomes is not needed for this occasion, even with a Bonferonni correction (Bland and Altman 1995) our results stand.

It was never our intent to advise against future RCTs comparing hamstring tendon autograft with bone patellar tendon bone autograft ACL reconstruction, we even encouraged embarking on such a trial in our review of systematic reviews on ACL reconstruction (Poolman et al. 2007). Moreover, our cumulative meta-analysis certainly justifies the conduct of an RCT and may help focussing the research question on a functional outcome measure. This future study should not be powered on anterior knee pain, but on patient orientated outcome including stability. Interestingly, Biau and co-workers decided not to publish their meta-analysis on functional outcome in their BMJ manuscript (Biau et al. 2006), but in a separate publication (Biau et al. 2007).

Sensitivity analysis

First, Biau et al. confuse clinical heterogeneity and statistical heterogeneity. This is a key element of our manuscript. Although statistical testing may not reveal statistical heterogeneity, a strong biological rationale can easily identify sources for inter study differences (Hopayian 2001). We analyzed one example discussed by Biau et al., but not analyzed in their meta-analysis: graft femoral fixation technique (Biau et al. 2006). Our analysis illustrated the complicated nature of conducting meta-analyses that leaves it open for discussion. We agree with Biau et al. that many confounding factors, such as differences between centres, exist. The most sound scientific way to definitively answer the optimal graft choice and fixation in ACL reconstruction will be a sufficiently and appropriately powered RCT resulting in pure data and unbiased findings overcoming the issues discussed and analysed in our manuscripts.

Second, Biau et al.'s interaction test suggest only 34% chance of proving our result in error, thus 76% chance that we did conclude the correct conclusion form our sensitivity analysis. Still, our choice remains a sensible subgroup hypothesis and

point estimates suggest a possible difference that requires confirmation in future randomized controlled studies.

Patient-based meta-analysis

We applaud Biau and co-workers' endurance embarking on a patient based meta-analysis. This type of meta-analysis has certainly advantages, but poor design in primary study always stays the weakest link (Hopayian 2001, Lyman and Kuderer 2005). Only 4 of the 18 studies included in Biau's meta analyses had optimal study design, the mainstay had considerable methodological shortcomings. Therefore, the results of a patient based meta-analysis drawn from these poor studies are likely subject to considerable amount of biases and will only add confusion among clinicians. "A meta-analysis is like a Mediterranean bouillabaisse—in concert, all ingredients will enhance its delightful flavour but, no matter how much fresh fish is added, one rotten fish will make it stink." (Messerli 1985, Hopayian 2001).

The only conclusion that can be drawn from the currently available ACL bouillabaisse is, as we stated in our review of overlapping systematic reviews: "The issue of graft choice will ultimately be resolved with a large, appropriately powered, randomized trial comparing alternative approaches to anterior cruciate ligament reconstruction utilizing a validated patient-orientated outcome instrument to measure a clinically relevant treatment effect. Ideally, this would be a multi-center, expertise-based trial to facilitate generalizability of its results." (Poolman et al. 2007). Any meta-analysis based on the currently available data, literature based or patient based, will not change this conclusion and may certainly have a sub-optimal flavour.

Rudolf W. Poolman, Forough Farrokhyar and Mohit Bhandari

Department of Orthopaedic Surgery,
Onze Lieve Vrouwe Gasthuis, Postbus 95500,
NL-1090 HM Amsterdam, The Neterhlands
poolman@trauma.nl

Biau D J, Tournoux C, Katsahian S, Schranz P J, Nizard R S. Bone-patellar tendon-bone autografts versus hamstring autografts for reconstruction of anterior cruciate ligament: meta-analysis. Bmj 2006; 332 (7548): 995-1001.

Biau D J, Tournoux C, Katsahian S, Schranz P, Nizard R. ACL reconstruction: a meta-analysis of functional scores. Clin Orthop 2007; (458): 180-7.

Biau D J, Katsahian S, Kartus J, Harilainen A, Feller J, Sajovic M, Ejerhed L, Zaffagnini S, Röpke M, Nizard R. Bone–patellar tendo –bone versus hamstring tendon autografts for reconstruction of the anterior cruciate ligament: a meta-analysis on patient level data. Submitted 2007.

Bland J M, Altman D G. Statistics notes: Multiple significance tests: the Bonferroni method. BMJ 1995; 310: 170.

Brookes S T, Whitley E, Peters T J, Mulheran P A, Egger M, Davey Smith G. Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives. Health Technol Assess 2001; 5 (33): 1-56.

Gail M, Simon R. Testing for qualitative interactions between treatment effects and patient subsets. Biometrics 1985; 41 (2): 361-72.

Hopayian K. The need for caution in interpreting high quality systematic reviews. BMJ 2001; 323: 681-4.

Kwakkel G, van Peppen R, Wagenaar R C, Wood Dauphinee S, Richards C, Ashburn A, Miller K, Lincoln N, Partridge C, Wellwood I, Langhorne P. Effects of augmented exercise therapy time after stroke: a meta-analysis. Stroke 2004; 35 (11): 2529-39.

Lau J, Antman E M, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers T C. Cumulative meta-analysis of therapeutic trials for myocardial infarction. N Engl J Med 1992; 327 (4): 248-54.

Lyman G H, Kuderer N M. The strengths and limitations of meta-analyses based on aggregate data. BMC Med Res Methodol 2005; 5 (1): 14.

Messerli F. Metaanalysis. Are calcium antagonists safe? Lancet 1985: 7678.

Poolman R W, Abouali J A, Conter H J, Bhandari M. Overlapping systematic reviews of anterior cruciate ligament reconstruction comparing hamstring autograft with bone-patellar tendon-bone autograft: why are they different? J Bone Joint Surg (Am) 2007; 89 (7): 1542-52.

Poolman R W, Farrokhyar F, Bhandari M. Hamstring tendon autograft better than bone patellartendon bone autograft in ACL reconstruction A cumulative meta-analysis and clinically relevant sensitivity analysis applied to a previously published analysis. Acta Orthop 2007; 78 (3): 350-4.

Prodromos C C, Joyce B T, Shi K, Keller B L. A meta-analysis of stability after anterior cruciate ligament reconstruction as a function of hamstring versus patellar tendon graft and fixation type. Arthroscopy 2005; 21 (10): 1202.

Stewart L A, Parmar M K. Meta-analysis of the literature or of individual patient data: is there a difference? Lancet 1993; 341 (8842): 418-22.