J. Biomechanics Vol. 22, No. 8/9, pp. 963-965, 1989.
Printed in Great Britain

0021-9290/89 $3.00+.00

TECHNICAL NOTE

COMMENT ON 'THE INFLUENCE OF RUNNING VELOCITY AND MIDSOLE HARDNESS ON EXTERNAL IMPACT FORCES IN HEEL-TOE RUNNING'

BARRY T. BATES

Biomechanics/Sports Medicine Laboratory, University of Oregon, Eugene, OR 97403, U.S.A.

Abstract—Researchers not sufficiently sensitive to statistical power may interpret non-significant results as demonstrating that the conditions made no difference. Related issues include performer variability, sample size and effect size as they relate to experimental design. Examination of these factors suggests that caution must be exercised when interpreting the results of experimental data.

INTRODUCTION

In the past, research studies in biomechanics have often been designed and implemented on the basis of available equipment, cost and time constraints. These factors, along with a general lack of statistical background on the part of many researchers, have resulted in numerous examples of poor research. In many instances, little attention has been paid to proper or even adequate statistical design and appropriate procedures. Designs are often incorrect or insufficient, the number of subjects is inadequate, and too few trials are used resulting in unreliable data. These factors often lead to inadequate statistical power and false support for the null hypothesis. To complicate matters further, studies that support the null hypothesis are often not repeated, and the results go unchallenged. In addition, it is not uncommon for authors to present 'absolute conclusions', and make sweeping generalizations on the basis of little 'real' data. Although the points identified are applicable to many biomechanics research studies, I have chosen to emphasize them using a recent *Journal of Biomechanics* article as an example.

Nigg *et al.* (1987) identified impact force as an important variable related to the etiology of selected running injuries. The authors identified running velocity and midsole hardness as two of the boundary conditions that might influence the impact force and present the results of an experiment designed to investigate the effect of these two parameters for heel-toe runners. Based upon the results of the experiment, the authors conclude that increasing running velocity increases impact forces and impact loading rates, while changes in midsole hardness '*do not change* impact force peaks' (p. 956). They further emphasize this latter point by stating that 'the common assumption that the hardness of the midsole can be used to reduce external impact forces ('cushioning') *is not correct*' (p. 957). Suggested reasons for this 'surprising' result include the possible use of a protective mechanism and a redistribution of the load internally. The purpose of this article is to present some additional information as to why these conclusions regarding the effects of midsole hardness should be viewed with some reservations.

DISCUSSION

The discussion to follow will be limited to a single variable, impact force peak, since it is the variable of primary concern. The theoretical constructs discussed, however, are equally applicable to the other data presented, although fewer data are available for some of these variables.

Variability

The protocol used to obtain the data in the Nigg *et al.* (1987) study requires one to assume that a single data trial provides representative information for each of the subject-condition combinations, or that each subject trial comes from a common population of trials representative of all subjects. This latter assumption would be dependent upon all subjects exhibiting similar variabilities and/or performing using the same response strategy.

The validity of using a single trial as representative of the total population of steps produced by a runner is questionable. The results from a number of studies (Norman, 1980; Bates *et al.*, 1983a; Bates *et al.*, 1983b; Kinoshita *et al.*, 1985; DeVita and Bates, 1989) have strongly suggested that single trial protocol is both invalid and unreliable. Norman (1980) initially indicated that more than three trials are required to improve estimates of true scores when evaluating shoe differences. The study by Bates *et al.* (1983b) determined that a minimum of eight trials was necessary to obtain sample means that were on the average within 1/4 standard deviation of the criterion value. DeVita and Bates (1989) recently indicated that 25 trials are necessary to detect the subtle real differences that exist between shoe conditions. These results are at least in part due to the fact that differences introduced by footwear represent only a small portion of the total variance inherent in the biomechanical system being evaluated. Without representative data, the accurate assessment of shoe effects is, at best, difficult.

The assumption of a common, singular response pattern must be viewed with caution. Each individual has numerous options available for controlling the kinematics of the link system responsible for the production of the ground reaction forces. Winter (1980) has shown that many different combinations of lower extremity joint moments are available for producing an appropriate support moment during gait. Bates *et al.* (1983b) reported a subject by shoe condition interaction for five runners performing in five different shoes. Subjects responded differently to the various shoes (with different midsole characteristics) resulting in no apparent differences between conditions when the data were collapsed across subjects. Simpson *et al.* (1989) and Bates *et al.* (1986) have shown that individuals use different strategies to accommodate to external loads attached to the legs during running. Response strategies included a Newtonian response where the effect of the added weight could be observed in increased impact forces and a protective neuro-muscular response

Received in final form 5 October 1988.

where the impact forces did not change. Different response strategies can therefore result in no apparent differences between conditions when data are collapsed across subjects, suggesting the need for a within subject design when investigating these types of phenomena.

Design

Statistical analysis of the data was accomplished using a one-way analysis of variance for each variable. For evaluation of the effects of velocity, the data for the three stiffness conditions were combined. Stiffness results were apparently evaluated separately for each running speed. I would suggest that the correct design for data evaluation is a two factor repeated measures analysis of variance (velocity × stiffness × subject). This approach allows for the analysis of a possible interaction between the two main effects (velocity and stiffness). If no interaction is present, then the main effects can be evaluated. However, if an interaction does occur, the simple main effects must be examined. The authors appear not to have followed this procedure.

The merit of combining the variables of speed and midsole stiffness appears to have some limitations also. Increases in peak impact forces with increases in running speed are well documented in the literature (Hamill *et al.*, 1983; Frederick and Hagey, 1986; Munro *et al.*, 1987). Subject–speed interactions are rarely, if ever, observed over wide ranges in running speed. On the other hand, shoe characteristics are far more subtle and can result in runner–shoe interactions as previously indicated. The risk of combining these two variables in the same study is that the more subtle effect (midsole stiffness) will be masked by the dominant speed effect and variable subject response patterns.

Power

Perhaps a more important issue than the research design is the statistical power of the experiment. A less than ideal experimental design having adequate power is more likely to result in correct statistical decisions, whereas poor design may only result in a loss of information. Regardless of the design, a *post hoc* estimate of power is important in terms of how the results of a study are interpreted. Researchers not sufficiently sensitive to power may interpret non-significant results as demonstrating that the conditions made no difference, when in reality it is more likely that an error in detection (type II) of true differences was made. Insufficient power may result from large between-subject variability, small sample size and/or small effect size.

A current problem facing biomechanists is the identification of biomechanically meaningful differences. Since considerable anecdotal evidence is available to suggest that shoe midsole materials and characteristics do cause running related injuries, it would appear that researchers should be sensitive to detecting rather small differences. Ten percent of body weight (72 N) does not seem like an unreasonable difference to be concerned about. One hundred N has been suggested as a biomechanically meaningful difference based upon a review of the literature on reported shoe differences (DeVita, 1986; DeVita and Bates, 1988). Seven of the 12 reported differences between hardness across speeds were greater than 100 N.

No *post hoc* power analysis was reported for the study. Although insufficient information is available to complete an actual power analysis, an estimate is possible using the information presented. Using the mean standard deviation of 345 N, an effect size of 100 N and $\alpha = 0.05$, the statistical power of the study was estimated to be approximately 0.25 (Keppel, 1982, pp. 70–73). This indicates a 75% chance of making a type II error (accepting a false null hypothesis). Using these same data to estimate the necessary sample size to obtain a minimal power of 0.70 for the same criteria indicates the need for a sample size of 49 subjects. This of course, assumes that the individual data items are valid, and that all subjects are, in fact, using a similar response strategy.

If, however, an effect size of only 200 N is to be detected, then the Nigg *et al.* (1978) study has a power of approximately 0.75, which is reasonable. Increasing effect size is always a way to improve statistical power, but needs to be justified in functional terms, and/or acknowledged in discussion and conclusions.

Other ways to improve power are to increase the α level of the study, or reduce the between subject variability. Increasing α may or may not be a viable alternative, depending upon the purpose of the study. However, an exploratory study could justify the use of a larger α. Reduced variability can be accomplished in some instances by more practice within the experimental setting or additional accommodating time between conditions (DeVita, 1987; DeVita and Bates, 1989). Selection of a more homogeneous subject pool is another alternative. I am not advocating any of these procedures, but simply identifying them as ways for increasing the power of a study.

COMMENTS

The ultimate test of any significant finding is in its repeatability. Unfortunately, studies that support the null hypothesis are often not repeated, and the results go unchallenged. Researchers must be responsible for their results and the conclusions drawn from them. Limitations of the study must be acknowledged and conclusions presented within the context of the acknowledged limitations. The presentation of 'absolute conclusions' without acknowledging any reservations appears unjustifiable and misleading in the case of the Nigg *et al.* (1987) study based upon the information and data presented in the article. If this is an incorrect conclusion, then information should have been included to address the points discussed.

REFERENCES

Bates, B. T., Hamill, J. and DeVita, P. (1986) The effects of additional load on impact force. *Proceedings of the 4th Biennial Conference of the Canadian Society for Biomechanics: Human Locomotion IV* **4,** 215–216.

Bates, B. T., Osternig, L. R., Sawhill J. A. and Hamill, J. (1983a) Identification of critical variables describing ground reaction forces during running. *Biomechanics VIII-B* (Edited by Matsui, H. and Koboyashi, K.), pp. 635–640. Human Kinetics, Champaign, IL.

Bates, B. T., Osternig, L. R., Sawhill, J. A. and James, S. L. (1983b) An assessment of subject variability, subject–shoe interaction and the evaluation of running shoes using ground reaction force data. *J. Biomechanics* **16,** 181–191.

DeVita, P. (1986) Intraday and interday reliability of ground reaction force data. Unpublished Doctoral dissertation, University of Oregon.

DeVita, P. and Bates, B. T. (1987) The effects of time on selected ground reaction force parameters. *Biomechanics X-B* (Edited by Jonsson, B.), pp. 909–912. Human Kinetics, Champaign, IL.

DeVita, P. and Bates, B. T. (1988) Intraday reliability of ground reaction force data. *J. Hum. Mvmt Sci.* **7,** 73–85.

DeVita, P. and Bates, B. T. (1989) Shoe evaluation methodology for ground reaction force data. *Biomechanics XI-B* (Edited by de Groot, G., Hollander, A., Huijing, P. and van Ingen Schenau, G.), pp. 705–709. Free University Press, Amsterdam.

Frederick, E. C. and Hagey, J. H. (1986) Factors affecting peak ground reaction forces in running. *Int. J. Sport Biomechanics* **2,** 41–49.

Hamill, J., Bates, B. T., Knutzen, K. M. and Sawhill, J. A. (1983) Variations in ground reaction force parameters at different running speeds. *Hum. Mvmt Sci.* **2,** 47–56.

Keppel, G. (1982) *Design and Analysis, A Researcher's Handbook.* Prentice-Hall, Englewood Cliffs, NJ.

Kinoshita, H., Bates, B. T. and DeVita, P. (1985) Intertrial variability for selected running gait parameters. *Biomechanics IX-A* (Edited by Winter, D., Norman, R., Wells, R., Hayes, K. and Patla, A.), pp. 499–502. Human Kinetics, Champaign, IL.

Munro, C. F., Miller, D. I. and Fuglevand, A. J. (1987) Ground reaction forces in running. *J. Biomechanics* **20,** 147–155.

Nigg, B. M., Bahlsen, H. A., Luethi, S. M. and Stokes, S. (1987) The influence of running velocity and midsole hardness on external impact forces in heel–toe running. *J. Biomechanics* **20,** 951–959.

Norman, R. W. (1980) Information content in biomechanical analyses of the effects of shoes on joggers. *Proceedings of the Special Conference of the Canadian Society for Biomechanics: Human Locomotion I* **1,** 126–127.

Simpson, K. J., Bates, B. T. and McCaw, S. T. (1989) Impact force accommodation to additional loads. *Biomechanics XI-B* (Edited by de Groot, G., Hollander, A., Huijing, P. and van Ingen Schenau, G.), pp. 701–704. Free University Press, Amsterdam.

Winter, D. A. (1980) Overall principle of lower limb support during stance phase of gait. *J. Biomechanics* **13,** 923–927.