L.E. Johnston Jr

A comparative analysis of Class II treatments: a retrospective/prospective alternative

Authors' affiliations:
Lysle E. Johnston Jr., Department of Orthodontics and Pediatric Dentistry, The University of Michigan, Ann Arbor, MI, USA

Correspondence to:
Lysle E. Johnston, Jr., DDS, MS, PhD
Department of Orthodontics and Pediatric Dentistry
The University of Michigan
Ann Arbor, MI 48109-1078
USA
Tel: +1 734 747 3814
Fax: +1 734 763 8100
E-mail:lejjr@umich.edu

Abstract: Because of a perceived need to examine problems more pressing than 'now versus later' or 'headgear versus functionals', a retrospective/prospective alternative to the randomized clinical trial has been developed. In this approach, discriminant analysis is used to identify samples of ex-patients who, prior to treatment, were equally susceptible to various contemporary treatment alternatives: extraction, non-extraction; orthodontics, surgery; fixed versus functional. This method of 'confounder summarization' has made it possible for us to conduct a wide variety of long-term comparisons and, in the process, to assess treatment efficiency, stability, esthetics, functional impact, and interaction with the pattern of facial growth. These studies are described in broad outline and serve as examples of the type of data that can be generated in a timely fashion by carefully planned, quasi-experimental investigations.

Key words: Class II treatment; clinical trial; discriminant analysis; prospective; retrospective

Introduction

Throughout its long existence, the orthodontic specialty has been criticized, generally to little lasting effect, by a succession of vocal fringe organizations. Of late, however, a mixed bag of special-interest groups, loosely united under the banner of 'functional orthodontics', has managed to reel in a surprisingly large following of generalists and even the occasional specialist. What is remarkable about this strange little revolution is that it is, at bottom, almost entirely devoid of theoretical basis and support in the refereed literature. Unfortunately, much that is done by 'mainstream' orthodontists is equally as un-characterized. As a result, it has proved surprisingly difficult to

defend many of our most time-honored procedures against attacks that, on the face of it, seem patently outrageous and self-serving.

Given the specialty's apparent need for a bird in the hand, the Saint Louis University/University of Michigan response to the Request for Applications was designed to provide timely, relatively bias-free long-term data on the effects of bicuspid extraction, surgery, and two-phase treatments (and their therapeutic alternatives). It took the form of a *retrospective/prospective*, quasi-experimental design in which recall data are obtained from statistically defined, equally susceptible sub-samples of 'borderline' patients (1, 2). Because the subjects already would have been treated, this design minimizes problems associated with informed consent, Hawthorne effect, sample attrition, blinding, 'data-peeking', etc. Our method of sample selection uses discriminant analysis as a 'confounder summary' and thus is similar to an approach described earlier by Miettinen (3).

Discriminant analysis

Discriminant analysis is a multivariate technique that can be used to screen pre-treatment descriptive data (cephalograms, models, examination forms, etc.) to ferret out the anatomical differences that seem to have determined the original choice of treatments. These 'discriminating variables' are then weighted and combined to yield one or more discriminant functions, formulae by which a so-called 'discriminant score' can be assigned to each subject. On the basis of these discriminant scores, one can select samples of patients who, at the outset, were similar with respect to the characteristics that drove the treatment plans.

It has been our experience that the resulting groups of 'borderline' patients tend to have been so similar prior to treatment that, at least to a first approximation, any long-term differences can be attributed largely to the treatments and not to susceptibility bias. These similarly susceptible, borderline ex-patients are then recalled for records. Because sample selection is based solely on pre-treatment records, many common sources of retrospective bias are avoided. Moreover, because most of the patients we have studied were treated in university clinics by a wide variety of residents and instructors, both proficiency bias and detection bias should be minimal. Thus, in contrast to the usual clinical study in

which treatment is rendered by one or two workers and only the best outcomes are documented, the present university-based design provides a relatively conservative and unbiased characterization of the effects of treatment.

In a university clinic, all results are documented; however, the best records tend to walk out the door to assist the new graduate in passing specialty boards and in subsequent 'patient education'. Thus, in any sort of management trial, the use of university records may tend to underestimate slightly the goodness of the treatments under investigation. Indeed, there are only two obvious filters at work in this retrospective/prospective design: all patients had to have finished treatment (i.e. they had to have qualified for final records) and all subjects had to have expressed a willingness to be recalled in the event that an analysis of their pre-treatment records placed them in the borderline stratum. A simple example can be used to illustrate the process.

We know intuitively that women and men tend to differ in terms of weight and height. Based on these known differences, we can construct a simple discriminant function: add your weight in pounds to your height in inches and subtract 220. If the result – a 'discriminant score' – is negative, it is probable that you are female; if positive, that you are male. Scores in the neighborhood of zero would define a borderline stratum peopled by smaller men and larger women. Orthodontic treatment assignments, however, are too complex to be modeled by intuition. To this end, discriminant analysis serves not only to determine which variables are significant discriminators, but also to specify how they are to be weighted to yield the discriminant scores.

Clinical studies

To date, discriminant analysis has been used to conduct quasi-experimental long-term comparisons of extraction and non-extraction edgewise in both European and African American patients, to compare the medium-term impact of surgery and adult orthodontics, and to compare the immediate and long-term impact of one-stage and two-stage treatment protocols. In the course of these studies, we have recalled nearly 300 ex-patients. The goal of these recall examinations is to gather data

that bear on the various claims and counter-claims that bedevil the practicing orthodontist. On average, the post-treatment intervals for our recall studies ranged from about 4 years for the surgery patients to 16 years for the white orthodontic patients.

In addition to an analysis of the lateral cephalograms (4) and study models, the recall protocol featured a detailed periodontal examination (including probing and periapical radiographs), a detailed assessment of the functional impact of treatment (5, 6), and, finally, an analysis of the perceived impact of treatment by way of visual analog scales. The collected visual analog scales were also submitted to a variety of panels (European and African–American orthodontists, surgeons, laypersons) for evaluation.

Although a detailed discussion of these clinical studies is both beside the intent and beyond the scope of this brief summary, several key findings/conclusions may be listed.

Early treatment

1. Two-stage treatment (bionator/edgewise, activator/ edgewise) produced results in terms both of growth and stability that could not be distinguished from those achieved in a single, significantly shorter, more efficient phase of edgewise therapy (7–9).
2. Headgear/edgewise two-stage protocols aimed at the midface produced results that were largely the same as those achieved by two-stage treatments whose first phase was a 'functional' treatment designed to produce mandibular growth modification (10).

Extraction versus non-extraction

1. Of 125 white Class II extraction and non-extraction orthodontic patients recalled, on average, nearly 15 years after treatment, the vast majority – 9 in 10 – demonstrated a pattern of growth during treatment in which the mandible advanced relative to the maxilla; this excess mandibular growth was a major factor in the molar and overjet corrections (11).
2. Much of the post-treatment relapse/change seen in these 125 ex-patients appeared to be a dentoalve-

olar compensation for the excess mandibular growth that continued during and after retention (12–17).
3. On average, first bicuspid extraction patients – both black and white – ended up only about 2 mm 'flatter' than comparable non-extraction patients (12–14, 18).
4. Both white and black patients and observers seem to desire the profile flattening characteristic of bicuspid-extraction treatments; black patients, however, tend to prefer profiles that are several millimeters more protrusive than those favored by white patients (19).
5. Throughout a broad spectrum of malocclusions, bicuspid extraction had no discernible functional impact on the temporomandibular joint (TMJ) and the muscles of mastication in both black and white patients (18, 20).
6. In both black and white patients, bicuspid extraction had little if any impact on the health of the periodontium (21, 22).
7. The extraction and non-extraction patients (185 in all) were treated more or less without expansion, and perhaps as a result, their treatments proved considerably more stable (average irregularity less than 3.5 mm) than has been reported elsewhere in the literature (12–18).
8. Of the two basic orthodontic strategies examined here, only extraction treatments had any marked capacity to 'cure' crowding or to flatten the profile (12, 18).

Adult orthodontics and adjunctive surgery

1. Medium-term recall of adult orthodontic and surgical patients revealed a surprisingly high incidence of condylar resorption in the surgical patients – something on the order of 1 in 9 (23).
2. Neither adult orthodontics nor surgery appear to have a discernible functional impact on either the TMJ or the muscles of the head and neck (23).
3. Adult orthodontics and surgery were both seen by the patients as having a salutary effect on facial appearance; therefore, the greater risk of clinical misadventure that accompanies surgery argues that orthodontics has the greater expected utility, at least for the 'borderline' patient (23).

Abstrakt

Aufgrund des Bedarfs an einer Untersuchung von schwerwiegenderer Problematik als 'jetzt oder später' oder 'Headgear oder Funktionsapparat' wurde eine retrospektive/prospektive Alternative zur klinischen Zufallsstudie entwickelt. Bei dieser Methode wird eine Diskriminanten-analyse verwendet, um Ex-Patienten zu identifizieren, deren Qualifikationen für die verschiedenen neuartigen Behandlungsalternativen vor Behandlungsbeginn gleich waren: Extraktion, keine Extraktion; kieferorthopädische oder kieferchirurgische Behandlung; festsitzende oder Funktionsapparatur. Diese Methode der 'Confounder Summarization' hat es uns ermöglicht, eine große Anzahl von langfristigen Vergleichen anzustellen und dabei die Wirksamkeit der Behandlung, Stabilität, Ästhetik, Auswirkung auf die Gebißfunktionsfähigkeit und das Verhalten im Zusammenhang mit dem Gesichtswachstum zu beurteilen. Diese Studien werden im Groben beschrieben und gelten als Beispiele für die Art von Daten, die durch sorgfältig geplante quasi-Experimentaluntersuchungen zeitgerecht gewonnen werden können.

Resumen

Debido a la nececidad percibida de evaluar problemas mas urgentes 'ahora y no despúes' o 'aparatos cebenza en vez de aparatos funcionales,' se a deserallado una alternativa retrospectiva/prospectiva al uso de investigaciones clínicas que experimentan de manera aleatoria. En este ensayo, el análisis discriminiatorio se utiliza para identificar muestras de ex-pacientes que antes del tratamiento tenian igual susceptibilidad a varias alternivas de tratameintos contemporaneos: por ejemplo, extracción, no extracción; ortodoncia, cirugía; fijo contra funcional. Este método de resumir ha hecho posible el llevar a cabo una variedad de comparaciones a largo plazo y, en el proceso, evaluar la eficacia del tratamiento, estabilidad, estética, impacto funcional, e interacción con el modelo de crecimiento facial. Los estudios dan una idea general en resúmen y sirven como ejemplo del tipo de datos que se pueden generar a tiempo en investigaciones cuasi-experimentales planificadas cuidadosamente.

抄録

「今か後か」あるいは「ヘッドギアか機能的矯正装置か」などという問題よりもっと火急の問題を検討する必要性が認識されているので、無作為臨床試験に代わる後向き／前向き試験が開発されてきた。本アプローチでは判別分析を用いて、治療前に、抜歯、非抜歯、矯正治療、手術、固定式装置対機能装置といった、様々な現行の治療方法が等しく可能であった過去の患者集団を特定した。この "confounder summerization" の方法により、様々な期間の長期比較を行い、その過程において治療の効率、安定性、審美性、機能的影響及び顔面の成長パターンとの相互作用について評価することが可能となった。これらの研究の概要を述べ、注意深く計画した疑似実験的研究により、タイミングよく作成することができるデータの種類の実例を供する。

Structured Abstract

Author – Johnston Jr, LE.

Objectives – To describe a retrospective/prospective alternative to the randomized clinical trial and to illustrate its use in controlling susceptibility bias in a variety of long-term orthodontic treatment comparisons.

Design – A quasi-experimental design in which discriminant analysis is used to identify equally susceptible, morphologically comparable patients who can be recalled for the purpose of conducting long-term treatment comparisons.

Setting and Sample Population – White patients from the Department of Orthodontics, Saint Louis University, The Orofacial Pain Center, St. Louis, MO, and various private practices; black patients from various American orthodontic programs and private practices.

Experimental Variable – Contemporary strategic options: orthodontics versus surgery; extraction versus non-extraction; one-stage versus two-stage.

Outcome Measure – Descriptive cephalometric and study–model measures; clinical assessments of periodontal and craniomandibular health; various measures of the esthetic impact of treatment.

Results – By using discriminant analysis to control for the dentoskeletal characteristics that seem to have determined the choice of treatments, it was possible to select, in retrospect, samples of patients who were morphologically similar prior to treatment – exactly the type of patient one would want to enroll in a prospective trial. By identifying and recalling these 'borderline' patients, it proved possible to conduct relatively bias-free long-term comparisons of treatments that cannot easily be randomized. These studies are summarized in broad detail.

Conclusion – Discriminant analysis can serve as the basis of timely, relatively bias-free comparisons of basic clinical strategies that would be difficult or impractical to randomize.

References

1. Johnston Jr LE, Paquette DE, Beattie JR, Cassidy Jr DW, McCray JF, Killiany DM. The reduction of susceptibility bias in retrospective comparisons of alternative treatment strategies. In: Dryland Vig K, Vig PS, editors. *Clinical Research as the Basis of Clinical Practice. Monograph 25, Craniofacial Growth Series.* Ann Arbor: Center for Human Growth and Development, The University of Michigan; 1991. pp. 155–77.

2. Johnston Jr LE. Clinical studies in orthodontics: taking the low road to Scotland. In: Trotman CA, McNamara Jr JA, editors. *Orthodontic Treatment: Outcome and Effectiveness. Monograph 30,*

Craniofacial Growth Series. Ann Arbor: Center for Human Growth and Development, The University of Michigan; 1995. pp. 21-41.

3. Miettinen OI. Stratification by a multivariate confounder score. *Am J Epidemiol* 1976;104:609-20.

4. Johnston Jr LE. Balancing the books on orthodontic treatment: an integrated analysis of change. *Brit J Orthod* 1996;23:93-102.

5. Fricton JR, Schiffman EL. Reliability of a craniomandibular index. *J Dent Res* 1986;65:1359-64.

6. Fricton JR, Schiffman EL. The craniomandibular index: validity. *J Prosthet Dent* 1987;58:222-8.

7. Livieratos FA, Johnston Jr LE. A comparison of one- and two-stage non-extraction alternatives in matched Class II samples. *Am J Orthod Dentofac Orthop* 1995;108:118-31.

8. Johnston Jr LE. Functional appliances: a mortgage on mandibular position. *Aust Orthod J* 1996;14:154-6.

9. McKnight MM, Daniels CP, Johnston Jr LE. A retrospective study of two-stage treatment outcomes assessed with a modified PAR index. *Angle Orthod* 1998;68:521-4.

10. Chu JW. A comparison of two non-extraction Class II treatment strategies [Master's thesis]. Ann Arbor: The University of Michigan; 1997.

11. Johnston Jr LE. Growth and the Class II patient: rendering unto Caesar. *Semin Orthod* 1998;4:59-62.

12. Paquette DE, Beattie JR, Johnston Jr LE. A long-term comparison of non-extraction and bicuspid-extraction edgewise therapy in 'borderline' Class II patients. *Am J Orthod Dentofac Orthop* 1992;102:1-14.

13. Paquette DE, Beattie JR, Johnston Jr LE. Nichtextraktionsbehandlung vs. Prämolaren-extraktion in der Edgewise-Technik bei Klasse-II-'Grenzfall'-Patienten: ein Langzeitvergleich. *Informationen Orthod Kieferorthop* 1992;24:489-512.

14. Johnston Jr LE. Terapia estrattiva o non estrattiva: un confronto a lungo termine quasi-sperimentale. *Proc XIII SIDO [Italian Orthodontic Society] Congresso Nazionale* 1995;13:19-25.

15. Luppanapornlarp S, Johnston Jr LE. The effects of premolar-extraction treatment: a long-term comparison of outcomes in 'clear-cut' extraction and non-extraction Class II patients. *Angle Orthod* 1993;63:257-72.

16. Luppannapornlarp S, Johnston Jr LE. Gli effetti della estrazione dei premolari: confronto a lungo termine dei risultati tra i pazienti di II Classe 'chiaramente' estrattivi e non estrattivi. *Ortognatodonzia Italiana* 1995;4:831-47.

17. Luppannapornlarp S, Johnston Jr LE. Wirkungen der Prämolar-enextraktion: eine vergleichende Langzeitstudie der Behandlungsergebnisse bei eindeutigen Extraktionsfällen und Klasse-II-Nonextraktionsfällen. *Informationen Orthod Kieferorthop* 1996;28:351-70.

18. Hagler BL, Lupini J, Johnston Jr LE. Long-term comparison of extraction and non-extraction alternatives in matched samples of African American patients. *Am J Orthod Dentofac Orthop* 1998;114:393-403.

19. Scott SH. A comparison of the esthetic impact of extraction and non-extraction orthodontic in African-Americans [Master's thesis]. Ann Arbor: The University of Michigan; 1996.

20. Beattie JR, Paquette DE, Johnston Jr LE. The functional impact of extraction and non-extraction treatments: a long-term comparison in 'borderline,' equally-susceptible Class II patients. *Am J Orthod Dentofac Orthop* 1994;105:444-9.

21. Soeters PA. The long-term effect of extraction treatment on the periodontium [Master's thesis]. St. Louis: Saint Louis University; 1992.

22. Reed PW, Johnston Jr LE. Periodontal effects of extraction and non-extraction therapy in African American patients of comparable susceptibility: a long-term comparison. In: Sachdeva RCL, Bantleon HP, White LW, Johnson J, editors. *Orthodontics for the Next Millennium.* Dallas: Ormco; 1997. pp. 385-94.

23. Cassidy Jr DW, Herbosa EG, Rotskoff KS, Johnston Jr LE. A comparison of surgery and orthodontics in 'borderline' Class II, division 1 adults. *Am J Orthod Dentofac Orthop* 1993;104:455-70.