Bilateral Clubfeet Are Highly Correlated: A Cautionary Tale for Researchers

Kelly Gray B App Sc, Paul Gibbons MBBS,
David Little MBBS, PhD, Joshua Burns PhD

Received: 6 February 2014 / Accepted: 20 June 2014 / Published online: 15 July 2014
© The Association of Bone and Joint Surgeons® 2014

Abstract

Background Congenital talipes equinovarus, or clubfoot, is a common pediatric orthopaedic condition of unknown origin. In many clubfoot clinical trials, interventions are assigned to a patient, but response to treatment is assessed separately in each foot. Trials commonly report x patients with y feet where y is greater than x (eg, 35 patients with 56 feet). However, common statistical tests assume that each data point is independent. Although data from

unilateral cases of clubfoot are independent, it is unknown if each foot of patients with bilateral clubfeet are correlated.

Questions/purposes The purpose of this study was to assess the correlation in the feet of patients with bilateral clubfeet by (1) evaluating the degree of severity between lower limbs of each patient with bilateral clubfeet at baseline; (2) determining if right and left feet of each patient responded to intervention in the same way; (3) determining the proportion of bilateral relapse; and (4) determining the proportion of right and left feet which required the same intervention to correct bilateral relapse.

Methods We performed a chart review of the records of 33 patients with bilateral clubfeet (66 feet). Baseline severity was assessed using the Pirani score. The number of Ponseti serial casts to correct the deformity, the proportion of patients who underwent bilateral Achilles tenotomy, the proportion of bilateral relapse, and the treatment to correct bilateral relapse were examined.

Results The degree of severity between right (Pirani score mean, 5.2; SD, 0.8) and left (Pirani score mean, 5.2; SD, 0.5) feet for each patient at baseline was highly correlated ($r = 0.76$, $p < 0.001$). Response to intervention between lower limbs was highly correlated for the number of Ponseti casts required for initial correction (right mean, 5.2, SD, 1.1; left mean, 5.2, SD, 1.3) ($r = 0.89$, $p < 0.001$) and the proportion of patients who underwent bilateral Achilles tenotomy (right, 17/18; left, 16/18) ($r = 0.94$, $p < 0.001$). In the nine patients who experienced relapse, eight experienced bilateral involvement. In all cases of bilateral relapse, the right and left foot of each patient required the same intervention to correct the relapse.

Conclusions In patients with bilateral clubfeet, baseline severity, response to initial Ponseti treatment, Achilles tenotomy, and relapse outcomes were highly correlated in

Each author certifies that he or she, or a member of his or her immediate family, has no funding or commercial associations (eg, consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article.

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research* editors and board members are on file with the publication and can be viewed on request. Each author certifies that his or her institution approved the human protocol for this investigation, that all investigations were conducted in conformity with ethical principles of research, and that informed consent for participation in the study was obtained.

K. Gray (✉)
Physiotherapy Department, The Children's Hospital at
Westmead, Locked Bag 4001, Westmead, Sydney, NSW 2145,
Australia
e-mail: kelly.gray@health.nsw.gov.au

P. Gibbons, D. Little
Department of Orthopaedic Surgery, The Children's Hospital at
Westmead, Sydney, NSW, Australia

J. Burns
Paediatric Gait Analysis Service of NSW, The Children's
Hospital at Westmead, Sydney, NSW, Australia

K. Gray, P. Gibbons, D. Little, J. Burns
The University of Sydney, Sydney, NSW, Australia

the right and left feet of each patient. Pooling clinical results of patients who present with bilateral clubfoot is statistically inappropriate, since results in two limbs of the same patient do not represent independent observations. These results support analogous work in other specialties suggesting that patients with bilateral presentations should not be analyzed as independent data points.

Level of Evidence Level IV, therapeutic study. See the Instructions for Authors for a complete description of levels of evidence.

Introduction

Orthopaedic conditions, such as clubfoot often present bilaterally [1, 4, 6, 14, 19]. Orthopaedic research commonly involves analysis of populations of patients with bilateral presentations. Often bilateral presentations (eg, two feet) are pooled as though they are independent observations [10]. Trials commonly report outcomes for y feet from x patients (eg, 31 feet from 20 patients).

It is unknown whether pooling of these data is appropriate. Many statistical analyses depend on independence; indeed, that is a foundational assumption for many common tests such as Student's t-tests and Mann-Whitney U tests [13]. It would not be assumed that the contralateral breast in someone with BRCA(+) breast cancer were independent data points, and it has been shown that they are not independent [5]. It has also been shown that multiple measurements taken from one patient (eg, two feet in bilateral cases) are more likely to be correlated than measurements taken from single limbs of multiple patients [2, 13]. To determine the correct analysis for clubfoot trials, any correlation between the feet of bilateral cases first needs to be determined.

We therefore sought to evaluate patients with bilateral presentations of clubfoot to (1) evaluate the degree of severity between lower limbs of each bilateral case at baseline; (2) determine if right and left feet of each patient responded to intervention in the same way; (3) determine the proportion of relapse in right and left feet of each patient; and (4) in bilateral relapse, determine if the intervention was similar in right and left feet of each patient.

Patients and Methods

A retrospective chart review of all patients with bilateral isolated clubfeet who underwent treatment at The Children's Hospital at Westmead, Sydney, Australia, between 2006 and 2013 were evaluated. Patients who had received treatment at another institution either before or during

followup, those with a presence of comorbidities, or those with feet that were classified as syndromal or atypical were excluded. Atypical presentation was determined clinically by the treating orthopaedic surgeon (PG). Institutional review board approval was obtained and parents provided consent for data to be kept for review on our Westmead Clubfoot Database. All children underwent treatment using the Ponseti method [15] at the dedicated clubfoot clinic at The Children's Hospital at Westmead. Baseline severity was established by the treating orthopaedic surgeon (PG) or senior physiotherapists using the valid and reliable Pirani score [8, 12, 17]. This score comprises six items, summed together to describe severity of the clubfoot deformity. Scoring ranges from 0 to 6 with 6 representing the most severe score. Long-leg plaster of Paris casts were applied by senior physiotherapists as per the Ponseti protocol [15] and changed every 7 days. Indication for Achilles tenotomy was determined by the orthopaedic surgeon (PG); during the study period the indication for Achilles tenotomy was 15° or less dorsiflexion at the completion of serial casting. At our institution, Achilles tenotomies are performed in the operating room with the patient under local anesthetic; therefore, the number of casts to correction was defined as the time when Achilles tenotomy was booked. If only one foot was corrected, casting continued on both feet until a decision regarding tenotomy/correction on both feet was established. After tenotomy, they wore long-leg plaster of Paris casts as per the Ponseti protocol for 3 weeks. After removal of the last cast, children wore boots and bar abduction bracing with 60° external rotation on both sides, full-time for 3 months and then during nights and naps until they were 4 years old. All patients were followed up in the clubfoot clinic by the treating orthopaedic surgeon and senior physiotherapist. When a foot showed evidence of relapse repeat serial casting using the Ponseti protocol was initiated. Relapse was determined by the treating orthopaedic surgeon (PG) and was defined as those with either dynamic or passive recurrence of any of the components of the deformity, necessitating repeat manipulation, casting, and/or surgery including tibialis anterior tendon transfer. Severity of relapse was determined by the type of treatment required to correct the relapse (repeat Ponseti serial casting $\pm$ surgery) and, where indicated, the type of surgery required. Minimum followup was classified as the time in which boots and bar were initiated. Although no child was lost to followup, the length of followup varied among children.

Thirty-three children (66 feet) with bilateral clubfeet met the inclusion criteria. The minimum followup was 2.5 months (mean $\pm$ SD, 25 $\pm$ 17 months; range, 2.5–60 months). Twenty-four patients had at least 12 months followup and 14 had at least two years followup (Fig. 1). Twenty-eight of 33 (85%) patients were male. The mean

Fig. 1 A graph shows the amount of followup time for patients after commencement of boots and bar bracing.

Fig. 2 A graph shows the high correlation between the number of initial Ponseti plaster casts for right and left feet ($r = 0.89$, $p < 0.001$). The numbers in the circles represent the number of patients.

age at presentation was 16 days (SD, 10 days; range, 5–52 days).

Data were analyzed using SAS Version 9.3 (Cary, NC, USA). Normality of data distribution was assessed using the Kolmogorov-Smirnov test. Because data were normally distributed, the correlations between feet were determined using Pearson product-moment correlation coefficient. Response to treatment was determined by the number of casts required to correct the initial deformity and the need for Achilles tenotomy. A p value less than 0.05 was considered significant.

Results

The degree of severity at baseline between the right and left foot of each patient was highly correlated. Twenty-eight of 33 (85%) patients presented with the same Pirani score in both feet and 30 of 33 (91%) recorded a maximum Pirani score difference of 0.5 between right and left lower limbs (right mean, 5.2, SD, 0.8; left mean, 5.2, SD, 0.5) ($r = 0.76$, $p < 0.001$).

Response to intervention was highly correlated in the right and left foot of each patient.

The average number of Ponseti plasters required to correct the deformity was 5.2 (right mean, 5.2, SD, 1.1; left mean, 5.2, SD, 1.3). There was a high correlation between the number of casts for each foot of each patient ($r = 0.89$, $p < 0.001$) with 29 of 33 (88%) requiring the same number of casts on right and left feet and 32 of 33 (97%) recording a maximum difference of one cast difference between both lower limbs (Fig. 2). Eighteen of 33 (55%) patients underwent an Achilles tenotomy on at least one foot. Seventeen of 18 (94%) patients underwent an Achilles tenotomy on both sides ($r = 0.94$, $p < 0.001$).

In patients who experienced relapse, a high proportion had bilateral involvement. Nine patients (27%) experienced a relapse during their followup period with eight experiencing bilateral involvement.

In all patients with bilateral relapse, both feet underwent the same intervention to correct the relapse. Six patients experienced bilateral relapse with all feet responding to repeat Ponseti serial casting. Four of those patients had a second relapse. All were bilateral relapses. Two patients (four feet) responded to repeat Ponseti serial casting, whereas one patient (two feet) required Ponseti serial casting and tibialis anterior tendon transfer on both feet. One patient (two feet) underwent bilateral Achilles tenotomy and medial stabilization for midfoot breaks. Three patients (three feet) experienced unilateral relapses; however, two patients experienced subsequent involvement of the other foot. After a unilateral relapse, corrected with Ponseti serial casting, one patient experienced bilateral relapse and underwent repeat Ponseti serial casting and tibialis anterior tendon transfer on both feet. One patient experienced a unilateral relapse, which corrected with repeat Achilles tenotomy and medial stabilization. This patient experienced a second unilateral relapse on the other side and underwent the same intervention as the first foot. Only one patient had a unilateral relapse without involvement of the other side. This was corrected with an Achilles tenotomy.

Discussion

Orthopaedic conditions commonly present bilaterally [1, 6, 14, 19]. In clinical trials, data from each limb are commonly treated as independent [10, 13]. However the results of our study show that patients with bilateral clubfeet have a high correlation between right and left feet for baseline severity, response to intervention including number of Ponseti plaster casts and Achilles tenotomies, relapse of right and left feet, and intervention required to correct bilateral relapse.

There are limitations to this study. Our trial is limited by small patient numbers; however, high correlations were

noted even with this small group. Larger trials are required. This trial is limited by its retrospective design. In particular, the definition of relapse was not predetermined. It therefore is unknown if in each case, one foot had a more significant relapse than the other foot. However as all patients with bilateral relapse underwent the same intervention on both feet, it is likely that severity between feet was similar. Prospective trials will improve precision of findings. While the average followup of 24.9 months may not have captured all relapses, a pattern of bilateral involvement was seen in this short time; longer followup is important.

Several areas of bias may have influenced outcomes in this study. First, without blinding of feet, assessment bias may have influenced the high correlation seen in baseline Pirani scores and diagnosis of bilateral relapse. However, 15% of our cohort presented with different Pirani scores in each foot and three of nine patients had unilateral relapse. Second, transfer bias may have led to higher correlation in the number of Ponseti plaster casts and Achilles tenotomies in each patient. In 12% of this cohort, it was documented that right and left feet did not require the same number of plaster casts. Furthermore, one patient underwent an Achilles tenotomy on one side only. Finally selection bias may have influenced results. As syndromal and atypical feet have been shown to be more resistant to treatment it is unknown how inclusion of these feet may have affected results. Future studies are required to address these subgroups.

The results of this study have important research implications as correlated data not analyzed correctly can significantly influence the outcome of trials [13]. First, including both feet from patients with bilateral clubfeet may artificially inflate the sample size. Consider a patient in a trial who is prescribed a medication for a fever. It would be incorrect to include measurements of temperature from two sites (eg, axilla and oral) because they are likely to be highly correlated. That is, as the temperature reduces in one site, it is also highly likely to reduce in the second site. In this example, the sample size has been incorrectly doubled. Similarly, this trial showed a large correlation between two sites (each foot) of each patient. When a sample size is inflated, the ability to detect a significant outcome increases. Therefore, if the sample size is artificially inflated, false (incorrect), significance may be incorrectly inferred. During study design, sample size adjustments can be used to account for known correlations, increasing the sample size and preserving power.

Second, biological factors result in interdependence. Intrinsic factors such as underlying ligamentous laxity may affect the response to manipulation and serial casting, whereas extrinsic factors such as parental compliance may affect the ability to follow bracing protocols. Stratification during study design can minimize these effects by ensuring a similar proportion of patients with bilateral clubfeet in each arm.

Third, common statistical tests assume multiple limbs from each participant are independent. Without accounting for correlations an over or under estimation of findings can occur. Simple (eg, excluding data from one limb) and sophisticated statistical models (eg, boot strap or generalized estimating equations) are able to account for known and unknown correlations, improving precision of findings [10, 11, 20].

This issue was identified in body-specific research including eyes, knees, and feet [9, 13, 18]. In a systematic review of orthopaedic literature, Bryant et al. [2] found that 42% of clinical studies in high-impact-factor orthopaedic journals had inappropriately used multiple observations from single individuals. In clubfoot research, Duffy et al. [7] combined multiple data from single individuals noting that an earlier study had observed no difference in findings when excluding data of one foot in bilateral clubfeet cases; however, this study had used statistical methods that assumed that each data point was independent [16]. In a trial comparing the French functional technique and Ponseti technique, Richards et al. [16] reported that the feet of patients with bilateral clubfeet were almost identical. These studies show that independence of multiple limbs from each participant in a trial cannot be assumed. Proof of independence is required to confidently conclude outcomes in trials which make this assumption.

Despite a high correlation in our cohort, we noted a small proportion of patients with differences in outcomes between feet. Similar to our findings, a minority of patients with differences between feet was in children with Charcot-Marie-Tooth disease [3]. Despite showing a high correlation of foot alignment between feet, Burns et al. [3] reported a small percentage of patients who had large differences between right and left feet. They concluded that although a high correlation existed in the majority of cases, this minority necessitated the ongoing need to clinically address each foot individually. Our findings suggest that clinically, this practice should continue with bilateral clubfeet.

In our cohort, bilateral clubfeet were highly correlated in initial severity, the number of Ponseti casts required to correct initial deformity, the need for Achilles tenotomy and the rate of relapse. These results have important implications in clubfoot trials in which data from bilateral cases are included. Pooling clinical results of patients who present with bilateral clubfeet is statistically inappropriate, since results in two limbs of the same patient are not independent observations. These results support analogous work in other specialties [9] and in other orthopaedic conditions [13, 18], suggesting that limbs in patients with bilateral presentations of musculoskeletal conditions should not be analyzed as independent data points unless they are proven to be independent. The inclusion of

bilateral presentations should be considered during study design, through sample size adjustments and stratification; and by using statistical analysis which can account for known and unknown correlations.

References

1. Barker SL, Macnicol MF. Seasonal distribution of idiopathic congenital talipes equinovarus in Scotland. *J Pediatr Orthop B*. 2002;11:129–133.
2. Bryant D, Havey TC, Roberts R, Guyatt G. How many patients? How many limbs? Analysis of patients or limbs in the orthopaedic literature: a systematic review. *J Bone Joint Surg Am*. 2006;88:41–45.
3. Burns J, Ouvrier R, Estilow T, Shy R, Laura M, Eichinger K, Muntoni F, Reilly MM, Pareyson D, Acsadi G, Shy ME, Finkel RS. Symmetry of foot alignment and ankle flexibility in paediatric Charcot-Marie-Tooth disease. *Clin Biomech (Bristol, Avon)*. 2012;27:744–747.
4. Byron-Scott R, Sharpe P, Hasler C, Cundy P, Hirte C, Chan A, Scott H, Baghurst P, Haan E. A South Australian population-based study of congenital talipes equinovarus. *Paediatr Perinat Epidemiol*. 2005;19:227–237.
5. Campeau PM, Foulkes WD, Tischkowitz MD. Hereditary breast cancer: new genetic developments, new therapeutic avenues. *Hum Genet*. 2008;124:31–42.
6. Chung CS, Nemechek RW, Larsen IJ, Ching GH. Genetic and epidemiological studies of clubfoot in Hawaii: general and medical considerations. *Hum Hered*. 1969;19:321–342.
7. Duffy CM, Salazar JJ, Humphreys L, McDowell BC. Surgical versus Ponseti approach for the management of CTEV: a comparative study. *J Pediatr Orthop*. 2013;33:326–332.
8. Dyer PJ, Davis N. The role of the Pirani scoring system in the management of club foot by the Ponseti method. *J Bone Joint Surg Br*. 2006;88:1082–1084.
9. Ederer F. Shall we count numbers of eyes or numbers of subjects? *Arch Ophthalmol*. 1973;89:1–2.
10. Gray K, Pacey V, Gibbons P, Little D, Frost C, Burns J. Interventions for congenital talipes equinovarus (clubfoot). *Cochrane Database Syst Rev*. 2012;4:008602.
11. Halanski MA, Davison JE, Huang JC, Walker CG, Walsh SJ, Crawford HA. Ponseti method compared with surgical treatment of clubfoot: a prospective comparison. *J Bone Joint Surg Am*. 2010;92:270–278.
12. Harvey N, Daley D, Mudge A, Sims S, Adams R. Reliability of physiotherapists using the Pirani scoring system for clubfoot. *Int J Ther Rehabil*. 2012;19:439–445.
13. Menz HB. Two feet, or one person? Problems associated with statistical analysis of paired data in foot and ankle medicine. *Foot*. 2004;14:2–5.
14. Pavone V, Bianca S, Grosso G, Pavone P, Mistretta A, Longo MR, Marino S, Sessa G. Congenital talipes equinovarus: an epidemiological study in Sicily. *Acta Orthop*. 2012;83:294–298.
15. Ponseti IV, Morcuende J, Mosca V, Pirani S, Dietz F, Herzenberg J, Weinstein S, Penny N, Steenbeek M. *Clubfoot: Ponseti Management*. 2nd ed. Seattle, WA: Global-HELP Organization; 2005.
16. Richards BS, Faulks S, Rathjen KE, Karol LA, Johnston CE, Jones SA. A comparison of two nonoperative methods of isolated clubfoot correction: the Ponseti method and the french functional (physiotherapy) method. *J Bone Joint Surg Am*. 2008;90:2313–2321.
17. Shaheen S, Jaiballa H, Pirani S. Interobserver reliability in Pirani clubfoot severity scoring between a paediatric orthopaedic

- surgeon and a physiotherapy assistant. *J Pediatr Orthop B*. 2012;21:366–368.
18. Sutton AJ, Muir KR, Jones AC. Two knees or one person: data analysis strategies for paired joints or organs. *Ann Rheum Dis*. 1997;56:401–402.
  19. Wallander H, Hovelius L, Michaelsson K. Incidence of congenital clubfoot in Sweden. *Acta Orthop*. 2006;77:847–852.
  20. Zeger S, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*. 1986;42:121–130.