http://dx.doi.org/10.2106/JBJS.19.00955

February 21, 2020

Accurate Prediction of Non-union Risk in Displaced Midshaft Clavicle Fractures?

Andreas Haubjerg Qvist
MD
Aalborg University Hospital, Denmark

Other Contributors:

Steen Lund Jensen
MD. PhD.
Aalborg University Hospital, Denmark

We have with great interest read the study by Nicholson et al (1). The authors conclude that a QuickDASH score of 40 points or above, absent radiographic callus formation and fracture movement on examination at six weeks were significant predictors of nonunion. This is an interesting finding, as we believe the treatment of displaced midshaft clavicular fractures should be individualized, based on predicted outcome and patient information, since the traditional treat-all/treat-none approach results in either surgical overtreatment or undertreatment.

In our opinion, however, there are some methodologic issues that need attention:

1: Selection of predictors.

The authors used logistic regression to test possible predictors, which were identified by univariate analysis. The exact number of possible predictors included is not mentioned, but from evaluation of table II alone, it seems that more than 15 variables were tested. The authors give no sample size calculation, despite existing guidelines concerning the minimum required number of events(non-unions) per possible predictors (event-per-variable, EPV). In general an EPV lower than 10 is not advisable (2) and if stepwise predictor selection is performed (as in the present study) an EPV of more than 50 may be needed (3). If 15 possible predictors were examined, the EPV in this study would be 1.8 given the observed number of nonunions (n=27). Low EPV values in the prediction model development could lead to poor predictive performance upon validation (3). This potential problem is not addressed in the present study.

2: The dichotomizing of a continuous predictor.

Dichotomizing of a continuous predictor is generally not recommended, as this could lead to serious bias, and risk of type 1 error (4,5). Furthermore, the statistical power to detect a relation between the variable and patient outcome is reduced (6). In the present study the continuous variable QuickDASH was dichotomized, without a discussion of the potential implications. The disadvantage of dichotomizing of a continuous predictor may be exemplified using table V. A patient with a QuickDASH score of 41 points combined with movement at the fracture site and no callus on radiographs would have a predicted non-union rate of 85.7%. A change in QuickDASH score to 39 points reduces the predicted non-union rate to 50.2%. This two point reduction in QuickDASH (0-100 scale) score results in an absolute reduction of 35.5% in predicted non-union rate.

3: Validation.

Two important steps in prediction modeling are internal and external validation (7-11). Prediction models perform best in the dataset on which they are developed. This performance may further be overestimated when the training dataset is small and the number of possible predictors is high. Internal validation is used to evaluate and correct the overfitting/optimism of the model (7–11). The internal performance (calibration, discrimination, and clinical usefulness) of the model should be examined on a resampled dataset (8,11). Numerous methods for resampling exists, but in general the principle is to split the data into a training set(on which the model is developed) and a test set(on which the model is validated).In the present study no method for resampling is described and it seems that the model was developed and tested on the same dataset. This practice could lead to severe overfitting (7–11). In this dataset measurement of discrimination, and a crude evaluation of clinical usefulness were performed. No calibration of the model was performed. The issue of potential overfitting is not discussed in the study.

After sufficient internal validation, external validation should be performed to assess the generalizability of the predictor model. External validation is performed by testing a developed model on a new external dataset (7–11). In the present study no external validation of the ”At 6 weeks” model was performed, and there is no mention of the need or the intent to perform such an analysis.

In table VII the ”At 6 weeks” model is compared with ”At the time of injury” model. As expected the model developed on the dataset performs best, and this comparison between models is of little value overall. Unrecognized by the authors, the true value of table VII lies in the clarification of why external validation is important. The ”At the time of injury” model behaves poorly in the new dataset, and performs only marginal better than coin flipping. This illustrates why no model should be used in a clinical setting

before external validation. As no external validation of the ”At 6 weeks” model was performed, it is in fact possible, that the ”At 6 weeks” model would behave just as poorly as the ”At the time of injury” model when tested on a new dataset.

Despite the lack of internal and external validation, the authors conclude that “delayed assessment at 6 weeks following displaced midshaft clavicle fractures can reliably predict union”. We do not believe this conclusion is valid until sufficient internal and external validation has been performed. The use of unvalidated models could lead to wrongful distribution of health resources and the application of potential harmful procedures in false positive cases.

Disclaimer: e-Letters represent the opinions of the individual authors and are not copy-edited or verified by JBJS.

References

1. Nicholson JA, Clement ND, Clelland AD, MacDonald D, Simpson AHRW, Robinson CM. Displaced Midshaft Clavicle Fracture Union Can Be Accurately Predicted with a Delayed Assessment at 6 Weeks Following Injury: A Prospective Cohort Study. J Bone Joint Surg Am. January 2020. doi:10.2106/JBJS.19.00955
2. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol. 1996;49(12):1373-1379. doi:10.1016/s0895-4356(96)00236-3
3. van Smeden M, Moons KG, de Groot JA, et al. Sample size for binary logistic prediction models: Beyond events per variable criteria. Stat Methods Med Res. 2019;28(8):2455-2474. doi:10.1177/0962280218784726
4. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. Stat Med. 2006;25(1):127-141. doi:10.1002/sim.2331
5. Austin PC, Brunner LJ. Inflation of the type I error rate when a continuous confounding variable is categorized in logistic regression analyses. Stat Med. 2004;23(7):1159-1178. doi:10.1002/sim.1687
6. Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ. 2006;332(7549):1080. doi:10.1136/bmj.332.7549.1080
7. Shipe ME, Deppen SA, Farjah F, Grogan EL. Developing prediction models for clinical use using logistic regression: an overview. J Thorac Dis. 2019;11(Suppl 4):S574-S584. doi:10.21037/jtd.2019.01.25
8. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J. 2014;35(29):1925-1931. doi:10.1093/eurheartj/ehu207
9. Lee YH, Bang H, Kim DJ. How to Establish Clinical Prediction Models. Endocrinol Metab (Seoul, Korea). 2016;31(1):38-44. doi:10.3803/EnM.2016.31.1.38
10. Steyerberg EW, Harrell FEJ. Prediction models need appropriate internal, internal-external, and external validation. J Clin Epidemiol. 2016;69:245-247. doi:10.1016/j.jclinepi.2015.04.005
11. Cowley LE, Farewell DM, Maguire S, Kemp AM. Methodological standards for the development and evaluation of clinical prediction rules: a review of the literature. Diagnostic Progn Res. 2019;3:16. doi:10.1186/s41512-019-0060-y

Conflict of Interest: None Declared

Article Author Response

9 March 2020

Article Author(s) to Letter Writer(s)

Response to JBJS e letter 'Accurate Prediction of Non-union Risk in Displaced Midshaft Clavicle Fractures?'

We thank Dr Qvist and colleagues for their comments and critique of our paper which we would like to address.

Selection of predictors

As outlined in our discussion, we recognize the potential limitations with the sample size of our study and the number of predictors included in the model. Given the challenges of recruitment of a prospective cohort and ensuring high follow-up we had only a specific time period of patient recruitment. Despite the large confidence intervals observed, the overall model fit was strong and we did find significant predictors with good biological plausibility.

Dichotomizing of continuous predictor

The QuickDASH cut-off may have limitations as Dr Qvist states and borderline scores around 40 could be problematic.

The ROC curve cut-off at 40 was superior to that of a continuous QuickDASH score based on our pilot work with the data. Unfortunately, we removed this from the discussion in the manuscript given word limitations for publication. As shown in figure 4, patients with a nonunion clustered towards the higher end of the QuickDASH score at six weeks. Self-assessment of pain and functional limitations as judged by the QuickDASH will vary greatly between patients. Rather than the absolute score being predictive it is likely a minimum 'threshold limit' is required to represent a significantly impaired return of function which in turn raises the risk of nonunion. This is very similar finding to the previous estimate of the DASH score of 35 as established by ROC curve modelling from a post-hoc analysis of the UK clavicle trial (1).

However, over two thirds of all patients will have a QuickDASH score well below 40, as found in our

study. This is possibly the main finding of our work given that almost half the cohort are strongly predicted to unite, with approximately 97% accuracy, based on a low QuickDASH score along with the absence of fracture mobility and callus on radiograph at six weeks.

In the presence of one or more of the clinical predictors being identified, an honest discussion of the potential increased risk of nonunion can be estimated to the patient which may provide a useful aid to guide treatment. Accurate prediction of fracture union is challenging and to date is almost exclusively based on factors available at time of presentation. The most objective predictor model to date is based on fracture displacement, comminution and smoking by Murray as we used as a comparison in our study (2).

We suspect, however, that the majority of acute fixations performed in clinical practice are undertaken solely on judgement from displacement on radiographs. This is despite accurate judgement of fracture displacement being affected by projection, patient position and timing of the scan (3, 4). The use of easily identifiable clinical predictors which could be performed in clinic may hold potential for better nonunion risk identification and we believe this may be a novel approach for others to validate.

Validation

The authors suggest several comprehensive strategies for model validation. With regards to interval validation, we did not feel we had a sufficient size dataset to make this worthwhile. The development of a ‘training and test set’ would have further reduced the powering of our model and increased the risk of chance findings.

With this in mind we would welcome external validation with a different data set or with a prospective evaluation of this model to assess the benefit compared to routine non-operative management of all displaced fractures.

We agree that a targeted fixation for clavicle fractures only at high risk of nonunion would appear to be both cost-effective (5) and would reflect the advancing evidence based approach to managing the majority of these injuries (6, 7). This is reinforced by the increasing consensus that a good outcome can be generally expected when union occurs despite the inevitable malunion (8–11).

The recommendation to perform acute fixation is currently based on patient preference and potentially by the treating surgeon’s bias for treatment. We are aligned with Dr Qvist’s belief that “treatment of displaced midshaft clavicular fractures should be individualized, based on predicted outcome and patient information, since the traditional treat-all/treat-none approach results in either surgical overtreatment or

under treatment." This was the aim of the study to help tailor treatment to each patient using an objective scoring system to predict union. We fully appreciate that this model needs to be validated in other centres and would welcome collaboration for that.

We hope our findings help surgeons to consider tailoring their treatment of these contentious injuries and provide a basis for future work in this challenging area of nonunion risk prediction.

Yours sincerely

Jamie A Nicholson

On behalf of all authors

MBChB (Hons), BScMedSci (Hons), PG Cert MRS, PG Cert CAP, FHEA, MRCS (Ed) and MFSTEd
Specialist Trainee and Honorary Clinical Teaching Fellow
Trauma and Orthopaedics
South East Scotland Rotation

References

1. Clement ND, Goudie EB, Brooksbank AJ, Chesser TJS, Robinson CM. Smoking status and the Disabilities of the Arm Shoulder and Hand score are early predictors of symptomatic nonunion of displaced midshaft fractures of the clavicle. Bone Jt J. 2016 Jan 1;98B(1):125–30.
2. Murray IR, Foster CJ, Eros A, Robinson CM. Risk factors for nonunion after nonoperative treatment of displaced midshaft fractures of the clavicle. J Bone Jt Surg - Ser A. 2013 Jul 3;95(13):1153–8.
3. Plocher EK, Anavian J, Vang S, Cole PA. Progressive Displacement of Clavicular Fractures in the Early Postinjury Period. J Trauma Inj Infect Crit Care. 2011 May 70(5):1263–7.
4. Backus JD, Merriman DJ, McAndrew CM, Gardner MJ, Ricci WM. Upright versus supine radiographs of clavicle fractures: Does positioning matter? In: Journal of Orthopaedic Trauma. 2014. 28 (11). 636-41
5. Nicholson JA, Clement N, Goudie E, Robinson CM. TRAUMA Routine fixation of displaced midshaft clavicle fractures is not cost-effective: A cost analysis from a randomized controlled trial. Bone Jt J. 2019 Jul 31;101 B(8):995–1001.
6. Lenza M, Buchbinder R, Johnston R V, Ferrari BA, Faloppa F. Surgical versus conservative interventions for treating fractures of the middle third of the clavicle. Cochrane Database Syst Rev. 2019 Jan 22 ;2019(1).
7. Amer K, Smith B, Thomson JE, Congiusta D, Reilly MC, Sirkin MS, et al. Operative Versus Nonoperative Outcomes of Middle-Third Clavicle Fractures: A Systematic Review and Meta-Analysis. J Orthop Trauma. 2020;34(1):e6–13.
8. Rasmussen J V., Jensen SL, Petersen JB, Falstie-Jensen T, Lausten G, Olsen BS. A retrospective study of the association between shortening of the clavicle after fracture and the clinical outcome in 136 patients. Injury. 2011;42(4):414–7.
9. Figueiredo GSDL, Tamaoki MJS, Dragone B, Utino AY, Netto NA, Matsumoto MH, et al. Correlation of the degree of clavicle shortening after non-surgical treatment of midshaft fractures with upper limb function. BMC Musculoskelet Disord. 2015 Jun 17;16(1).
10. Stegeman SA, De Witte PB, Boonstra S, De Groot JH, Nagels J, Krijnen P, et al. Posttraumatic midshaft clavicular shortening does not result in relevant functional outcome changes. Acta Orthop. 2015 Sep 3;86(5):545–52.

11. Goudie EB, Clement ND, Murray IR, Lawrence CR, Wilson M, Brooksbank AJ, et al. The Influence of Shortening on Clinical Outcome in Healed Displaced Midshaft Clavicular Fractures after Nonoperative Treatment. J Bone Jt Surg - Am Vol. 2017;99(14):1166–72.