



Assessment of the Accuracy of the AO Spine-TL Classification for Thoracolumbar Spine Fractures Using the AO Surgery Reference Mobile App

Global Spine Journal
2021, Vol. 11(2) 187-195
© The Author(s) 2020
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2192568220901694
journals.sagepub.com/home/gsj



Allan Hiroshi de Araujo Ono, MD, MSc¹ , Verônica Yulin Prieto Chang, MD¹,
Erico Myung Rodenbeck, MD¹, Alex Oliveira de Araujo, MD¹,
Rafael Garcia de Oliveira, MD¹ , Raphael Martus Marcon, MD, MSc, PhD¹,
Alexandre Fogaça Cristante, MD, PhD¹, and
Tarcisio Eloy Pessoa Barros Filho, MD, PhD¹

Abstract

Study Design: Cohort study.

Objectives: This study aimed to evaluate the accuracy of the AO Surgery Reference mobile app in the diagnosis of thoracolumbar fractures of the spine according to the AO TL classification, and to discuss the usefulness of this app in the teaching and training of the resident physicians in orthopedics and traumatology area.

Methods: The 24 residents of Orthopedic and Traumatology program assessed 20 cases of thoracolumbar fractures selected from the hospital database on 2 different occasions, with a 30-day interval, and they classified these cases with and without using the AO Surgery Reference app. A group of spine experts previously established the gold standard and the answers were statistically compared, with the inter- and intraobserver reliability evaluated by the kappa index.

Results: The use of the AO Surgery Reference app increased the classification success rate of the fracture morphology (from 53.4% to 72.5%), of the comorbidity modifier (from 61.4% to 77.9%) and of the neurological status modifier (from 55.1% to 72.9%). In addition, the mobile app raised the classification agreement and accuracy. The kappa index increased from 0.30 to 0.53 regarding the morphological classification of fractures.

Conclusions: The residents improved their ability to recognize and classify thoracolumbar spine fractures, which reinforces the importance of this tool in medical education and clinical practice.

Keywords

data accuracy, spine, spinal fracture, classification

Introduction

Smartphones and tablets have made medical information easily accessible mainly because of mobile apps.¹⁻⁴

Several studies have demonstrated the growth in the use of apps among physicians and students, in areas of orthopedic teaching, emphasizing its importance and usefulness in medical education and clinical practice.⁵⁻⁷

Orthopedic and medical apps are considered to be user-friendly, with a good intra- and interobserver agreement, and should preferably support in the medical conduct. Related to

thoracolumbar fractures, the classification of the disease aims to facilitate the evaluation and the adequate diagnosis, the communication between the expert doctor and, essentially,

¹ Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, Sao Paulo, Brazil

Corresponding Author:

Allan Hiroshi de Araújo Ono, Rua Dr. Ovídio Pires de Campos, 333, Cerqueira Cesar, Cep: 05403-010, Sao Paulo (SP), Brazil.
Email: allanmedcoluna@gmail.com



the management of the patient with traumatic spinal column injury.⁸

In the beginning, the classifications presented a descriptive morphological profile, with the only purpose of associating certain fracture patterns with the trauma mechanism, being described as compression, flexion-distraction, and rotation, according to the classification shown by Magerl et al⁹ in 1994.

However, they did not consider other important factors in the determination of the treatment, such as the presence of ligament injuries, neurological deficit, comorbidities, and did not report the good intra- and interobserver agreement.¹⁰

In 2005, Vaccaro et al¹¹ proposed the TLISS (Thoracolumbar Injury Severity Score), a scale based on 3 main characteristics of the thoracolumbar lesion: (a) trauma mechanism, (b) integrity of the posterior ligament complex, and (c) state of the patient. However, despite their validity as a method of classification, studies showed inconsistency among the observers, which led to the creation of TLICS (Classification of Thoracolumbar Injury and Severity Score), in which the mechanism of trauma was replaced by fracture morphology.¹²

Considering the problems associated with the classifications cited above, in 2013, a new classification called AO Spine TL in order to adapt the international needs has presented, with alphanumeric sequence and ascending order of severity, reducing inconsistencies between observers and allowing lines of treatments from newly established parameters.¹³

The new order of severity was classified by (a) the fracture morphology, with order of severity from A0 to A4, fractures B1 to B3 and type C; (b) the neurological status of the patient being N0 the absence of neurological deficit and N4 the complete deficit; and (c) the clinical and radiological determinants classified by the M1 and M2 modifiers.

The authors of the new rank participated in the development of a new mobile app on Android and iOS platforms. The AO Surgery Reference mobile app, released by the AO Foundation, contains all the published fractures and are validated by the AO Foundation, as well as the access routes and possible treatments for appendicular and spinal fractures, been accessible by any device with internet access.¹⁴

Taking into consideration the orthopedic injuries in the spine, the therapeutic decision is often linked to the correct interpretation of the lesion pattern and its classification.

Due to the multiple details and variety of the classifications, the AO Surgery Reference app has the advantage of assisting the traumatologist in the correct evaluation of the presented case. In this mobile app, the physician can simulate the classifications for various types of fractures, through self-explanatory drawings and textual information such as thoracolumbar fractures, which makes it possible to identify fracture types and classify them for morphology and its associated modifiers.

Due to the great amount of knowledge that the resident physicians in orthopedic and traumatology must acquire during their training, strategies are sought that facilitate the learning and the correct interpretation of the presented clinical cases.

The aim of this study was to evaluate the accuracy of the AO Surgery Reference mobile app in the diagnosis and treatment of thoracolumbar fractures as proposed by the AO TL classification, as well as to discuss the usefulness of this app in the teaching and training of the resident physicians in the orthopedics and traumatology area.

Material and Methods

A diagnostic study was carried out at the Orthopedics and Traumatology Service of a hospital school in the year 2016 with resident physicians in the Orthopedics and Traumatology Service.

We selected 20 de-identified thoracolumbar cases from our medical database that were representative of a multitude of thoracolumbar trauma fracture patterns. Pathological fractures secondary to tumor or infection were excluded.

The originally published classification system served to aid in this process. The cases were classified in a collective meeting by an expert session, where 6 of the authors who could consult the mobile app, as well by consulting the original article where the classification was published.¹³ Of the authors, 2 were senior spine surgeons and former members of AO Spine, 2 were AO Spine Fellowship-trained doctors and, 2 current spine fellows.

It was selected 20 cases from 40 cases of our data bank to be used in this study, with unanimous classification among the researchers, establishing a standard.

The remaining cases were excluded. Ten cases were excluded because of the similarity with other cases; 4 cases were excluded because the patient has more than 1 spinal fracture, and 6 cases that were not classified the same by the authors as consensus.

These preselected 20 cases were then presented to all 30 Orthopedic and Traumatology resident physicians regularly enrolled in the Orthopedics and Traumatology Service in a 2-hour theoretical lecture on thoracolumbar fractures, to present the AO Spine TL classification, as well as to discuss teaching, training, and debate.

The lecture was presented and projected on a big screen and was accompanied by radiographic images, computed tomography, magnetic resonance imaging when pertinent to the case, and other relevant clinical information such as neurological deficit and comorbidities.

Just after the lecture, all the residents were submitted to an individual test, including questions with selected cases of thoracolumbar fractures presented one by one projected on the screen. Each participant had 3 minutes to evaluate each of 20 cases, classify and answer the questions without the use of any support material.

At the end of the session, resident physicians were instructed to download the AO Surgery Reference mobile app on their own cell phones and a short presentation on how to use the app was made. For 1 month, all the resident physicians were solicited to use this mobile app in their daily practice.

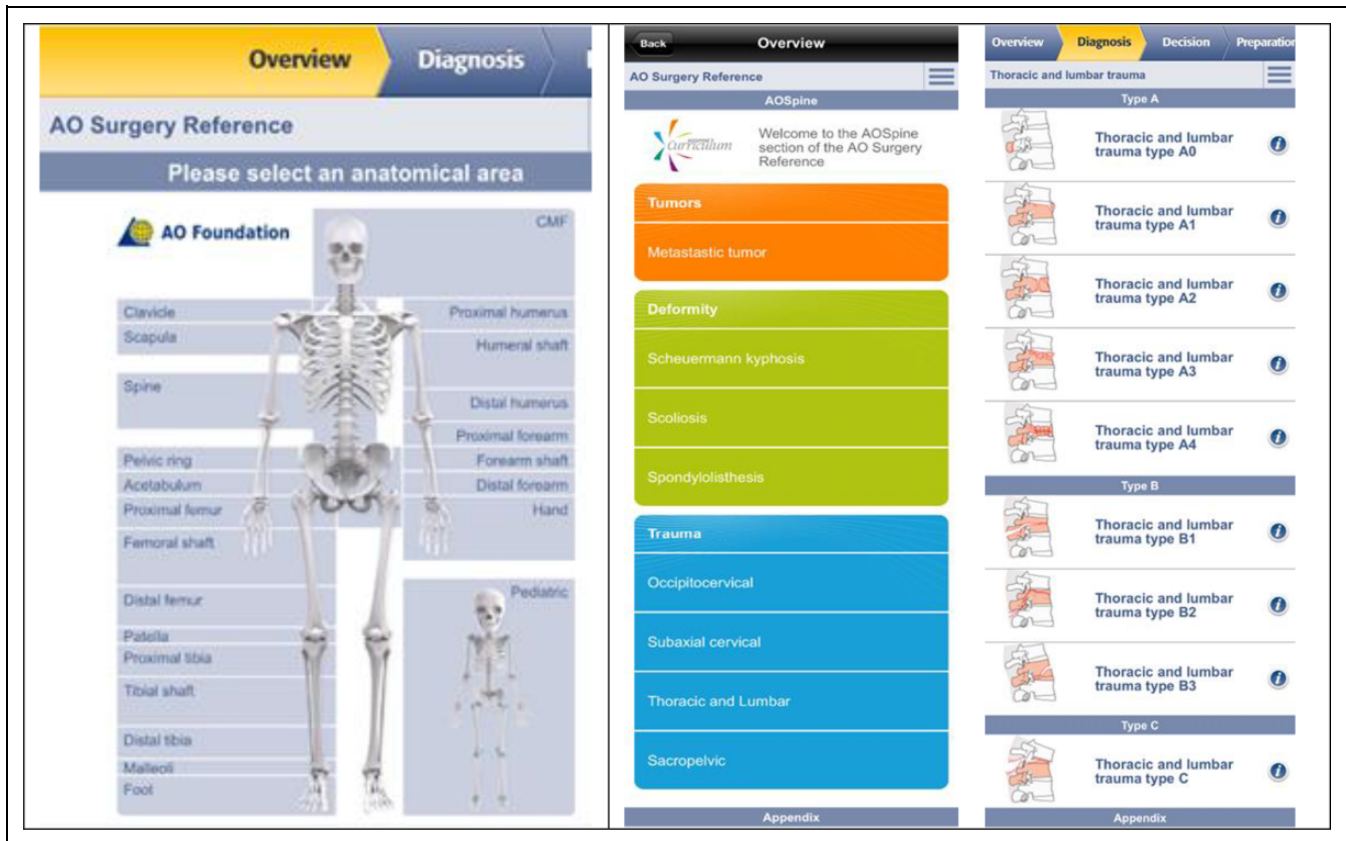


Figure 1. Initial screens of AO Surgery Reference mobile app.

All the 30 resident physicians were invited to participate voluntarily in the case presentations and training and classification sessions of the AO Surgery Reference mobile app and sign an informed consent form.

One month later, the same resident physicians were submitted to a second theoretical lecture, in which the same previous thoracolumbar fractures cases were presented in random order and they had 3 minutes to evaluate, classify, and answer the questions, but this time with the use of the mobile app.

The resident physicians were not allowed to discuss the cases at the time of the classifying session with one another and did not have access to the images during the time period of 1 month between the first and second lecture.

The correct answers precisely debated by consensus among the researchers and authors of this study formed the standard template for the evaluation of the answered questionnaires and the number of correct answers at each moment of the evaluation.

All the difficulties of each physician were recorded at the end of the lecture for further consideration.

In addition, it was solicited to all resident physicians to answer a questionnaire about the use of the mobile app, where it was asked if the app was studied as proposed, if the use of the app facilitated in the classification of fractures and if the use of this tool modified the management or treatment of the cases.

The blank or erased questionnaires were excluded. Those who did not have a smartphone or did not download the mobile app or those who did not participate in the theoretical lecture/case presentation and the ones who refused to sign the free informed consent form were also excluded.

Figures 1 to 4 include the main initial screens of the AO Surgery Reference mobile app used in this study and downloaded by all participants.

Statistical Analysis

The results were presented as tables, figures, and graphs. All demographic and characteristic data at the baseline were summarized and associated.

The accuracy, as well as the interobserver agreement analysis, were performed using the Fleiss kappa index, with values between 0 and 0.19 been considered as weak agreement; 0.20 and 0.39 as a regular agreement; between 0.40 and 0.59, a moderate agreement; between 0.60 and 0.79 as a substantial agreement and the values between 0.80 and 1.00 being near-perfect.¹⁵

The McNemar test was used to access accuracy when we compared the residents' answers with the standard established by the expert session.¹⁶

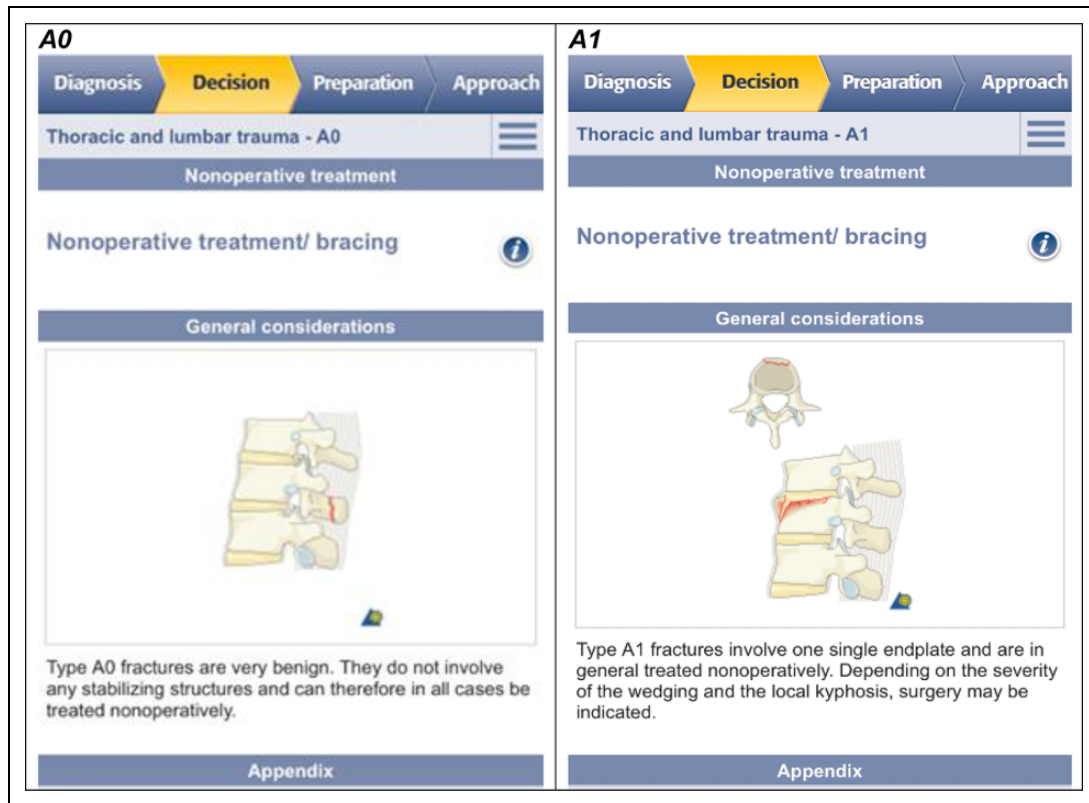


Figure 2. A0 and A1: Thoracic and lumbar trauma initial screens of AO Surgery Reference mobile app.

Results

Twenty-four resident physicians, at different stages of their 3 years of medical residency of the orthopedic service, agreed to participate in the study and met the inclusion criteria. Six residents were excluded because they did not participate in both sessions.

Despite all participants having downloaded the mobile app immediately after the first session and being encouraged by the researchers to study and practice during the period of 1 month, not a single resident physician claimed to have studied the app, and 100% stated that they did not use the app for theoretical study, as shown in the answers to the questionnaires at the end of the second theoretical lecture.

Even without the resident physicians having studied previously the tool, the answers in the questionnaire has shown that 95.65% stated that the app facilitated the understanding on the classification of thoracolumbar fractures, 91.3% stated that the app helped in the conduct of the case, and 60.86% stated that the use of the app modified their conduct in most of the cases.

Related to the morphological classification of fractures, there was an improvement in the accuracy rate from 53.4% to 72.5%. Overall, there was an increase of 194 cases classified correctly in the pre-app test and 263 correctly classified cases in the post-app test, with an increase in the fracture morphology accuracy of 35.5% ($P < .001$). The distribution by morphological type of fractures before and after the use of this app is shown in Figure 5.

Altogether, there was an increase of correctly classified cases from 174 before the use of the mobile app to 230 cases after the use of the mobile app, an increase of 32.2% ($P = .009$). The stratification of the modifier “neurological status” before and after the use of this app is demonstrated in Figure 6.

Regarding the modifier “comorbidities” (Figure 7), there was an increase in the number of correct answers from 61.4% to 77.9%. The number of correctly classified cases increased from 153 in the first theoretical lecture to 194 in the second theoretical lecture, an increase of 26.8% ($P < .001$).

The classification of the modifier “comorbidities”, pre- and post-use of the app is demonstrated in Figure 7.

The results of the morphological classification questionnaires pre and post the app use, using the McNemar test, have shown that there was an increase in the accuracy of responses by 39.8% ($P < .001$).

The interobserver concordance assessed before the intervention was 0.3 (95% CI 0.29-0.31) calculated by the Fleiss kappa test and after the intervention, the agreement index increased to 0.53 (95% CI 0.52-0.54).

The results of the neurological status modifier in the pre and post questionnaires using the McNemar test has shown an improvement in the accuracy of the responses by 45.2% ($P = .01$). The index of interobserver agreement between the answers evaluated using the Fleiss kappa test showed a correlation of 0.23 (95% CI 0.21-0.24) in the first evaluation and increased to 0.53 (95% CI 0.51-0.55) in the second evaluation after app consulting.



Figure 3. A2, A3, and A4: Thoracic and lumbar trauma initial screens of AO Surgery Reference mobile app.

The results of the classification of the modifier “comorbidities” in the pre and post questionnaires using the McNemar test, has shown that there was an increase in the number of correct answers by 58.1% ($P < .001$).

The agreement index between the results, calculated by the Fleiss kappa test evaluated before the intervention showed concordance of 0.04 (95% CI 0.01-0.07) and after the intervention, the agreement index increased to 0.29 (95% CI 0.27-0.32).

Discussion

This study specifically evaluates the accuracy of classification of thoracolumbar fractures using a mobile app in the field of orthopedics and traumatology, including the classification of trauma and its complications and treatment.

The present study has demonstrated that the use of the AO Surgery Reference app helps in the classification of clinical cases, increasing the accuracy in all parameters (morphology, comorbidities, neurological) and, consequently, interobserver agreement and accuracy.

The present study demonstrated that the AO Surgery Reference mobile app increases the accuracy and the comparison indices, functioning as an effective consultation tool, being used by young physicians even without in-depth training on the use of the app.

In this study, the accuracy of correct identification of the fracture morphology increased from 53.4% to 72.5%; modifier “comorbidity” increased from 61.4% to 77.9% and the modifier “neurological condition” increased from 55.1% to 72.9%.

The use of the AO Surgery Reference app also has increased the agreement and accuracy in the classifications of the assessed group.

The kappa index, which measures the agreement among a population studied, increased from 0.3 to 0.53 for the morphological classification of fractures, approaching the index evidenced by Sadiqi et al,¹⁷ between spine surgeons with experience up to 10 years (0.69 [0.44-0.91]), 11 to 20 years (0.69 [0.22-1]), and older than 20 years (0.67 [0.31-0.85]).¹⁷

However, for type A3 and A4 fractures, Sadiqi et al¹⁷ observed a small interobserver agreement, contrary to what was observed in this study, which showed a significant increase in



Figure 4. B2, B3 and C: Thoracic and lumbar trauma initial screens of AO Surgery Reference mobile app.

the concordance rate in the classifications of these fracture subtypes after the app assistance.

This agreement was also demonstrated for fracture modifiers: The kappa index for “neurological status” varied from the agreement of 0.23 to 0.53, and for “comorbidity” ranged from 0.04 to 0.29.

The results showed, therefore, that the mobile app improved the residents understanding, showing its important effect on the classification of thoracolumbar fractures. This strengthens the importance that smartphone devices can be a useful tool in medical teaching and education.

Health professionals and academics have used the technology on daily basis as it is an extremely useful tool for getting information for diagnosis, treatment, medications, and updates in the area of interest.¹⁸

This phenomenon occurs in other areas.^{7,19,20} The use of mobile apps in medicine has been increasing significantly. There are tens of thousands of medical apps with multispecialty content, but quality and reliability have been questioned.^{7,19,20}

These studies have shown the great need and desire for apps that assist in daily medical practice.

Payne et al⁵ have analyzed questionnaires that evaluated trained physicians and medical students and found that 52% of students used more than one medical request for clinical consultation and 51% of resident physicians used more than one mobile app, most of which were consulted more than once a day.²⁰

Grasso et al¹⁸ have evaluated the use of mobile devices like palmtops and tablets in the teaching and medical practice of medical students, and observed that 52% of individuals used mobile devices for consultation and study, and 74% of the physician students used the devices in clinical practice.

Although most of the residents evaluated in this study did not use the mobile application during their daily clinical practice, the results showed that the instrument enabled a large increase in the number of correct fracture classification, which could generate direct benefits in orthopedics teaching and in daily clinical practice with patients.

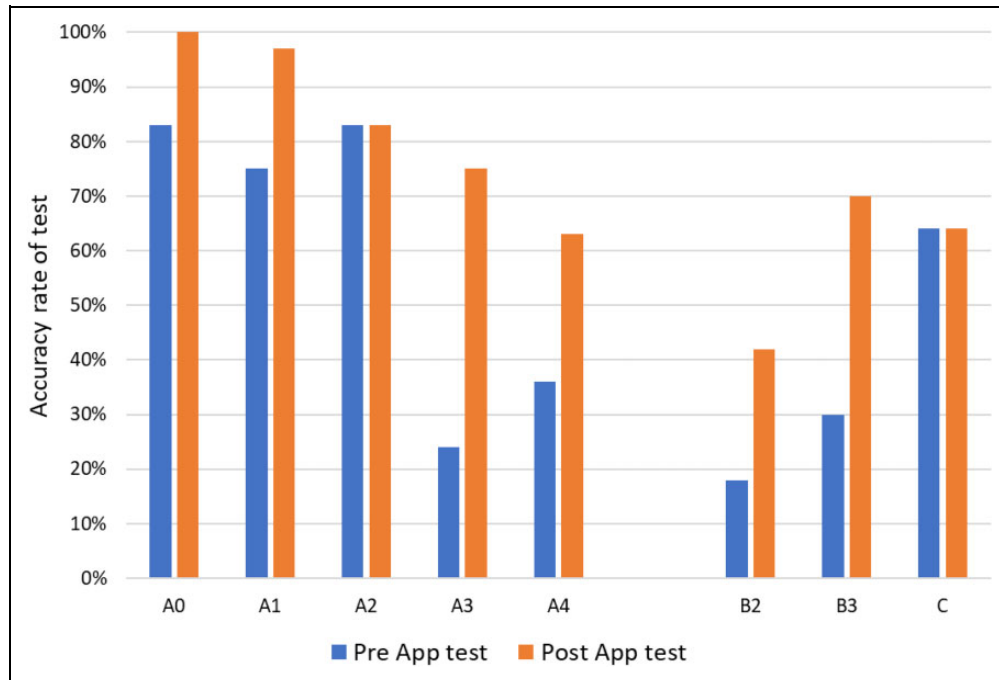


Figure 5. Distribution of morphological type of fractures before and after the use of AO Surgery Reference mobile app.

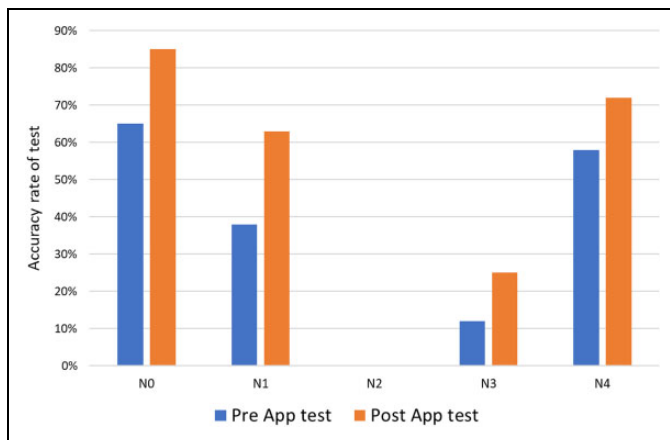


Figure 6. Stratification of the neurological status before and after the AO Surgery Reference mobile app.

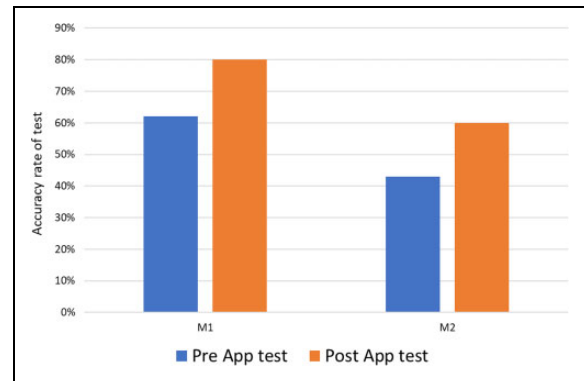


Figure 7. Stratification of the comorbidities status before and after the AO Surgery Reference mobile app.

However, well-defined criteria and detailed methodology must be meticulously studied for the creation and validation of mobile apps, assessing their quality and avoiding misinformation, which can be dangerous for professionals and patients. Within the context of limitations and concerns in the use of the mobile apps in medicine, the US Food and Drug Administration has recently released recommendations for device software functions.^{21,22}

Following these concerns, Buijink et al²³ have proposed the creation of institutes and mechanisms that regulate the quality of the medical information contained in the mobile apps, wrong medical information could be dangerous to professionals and patients. Recently, a mobile app developed by a

pharmaceutical company was removed from the internet by providing a dose calculator, different from the original formula.

Within the orthopedic area, there are several available mobile apps. Franko et al have conducted a study in the United States with orthopedists and found that 84% of respondents had smartphones, 53% use apps in their medical practice, and 96% would like to see more orthopedic-related apps.^{1,7}

Thus, in our study, the AO Surgery Reference mobile app has been shown to be a medical tool in improving the classification of thoracolumbar fractures even for orthopedic physicians and resident traumatologists.

This corroborates the results of Sadiqi et al¹⁷ who showed that years of experience as spine surgeons did not affect the improvement of levels of agreement.

The present study was able to statistically demonstrate a significant improvement in the recognition and classification of cases of thoracolumbar spine injuries.

Limitations

Some limitations of this study may be considered.

The assessment of participants at 2 or 3 points could be considered using the app at all times to verify and validate the intra- and interrater reliabilities.

The lack of a control group can also be considered as a limitation. However, we consider that it can be justified by analyzing possible benefits in clinical practice and the learning of surgeons and residents using this classification tool. Ultimately, all benefits are accrued to the patients treated, which would not exist if access to information provided by the app were denied to a control group without the app.

Another limitation of the study was that it did not consider possible effects of a learning curve or knowledge progression over the period of one month between assessments. Despite the guidance given to the residents to use the study app, we realize that many still prefer to use printed study material, this can be a bias.

In future studies, we intend to evaluate if the same happens with the other classifications, as for example in the cervical lesions. We also intend to use other orthopedics applications as a study tool.

Conclusion

The AO Surgical Reference app has been demonstrated to improve the accuracy and reliability interobserver in the classification of thoracolumbar fractures.

The physicians who have participated in this study have used this app in several clinical situations, with an improvement of accuracy and interobserver reliability of the group, which reinforces the importance of this tool not only in teaching but also in everyday life, as a quick and practical source of information that benefits not only the doctor, but also the patient. Due to its high availability and easy access, the use of this technology in the medical environment is widespread.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Allan Hiroshi de Araujo Ono  <https://orcid.org/0000-0003-2112-4654>

Rafael Garcia de Oliveira  <https://orcid.org/0000-0002-9160-9246>

References

1. Andrawis JP, Muzykewicz DA, Franko OI. Mobile device trends in orthopedic surgery: rapid change and future implications. *Orthopedics*. 2016;39:e51-56. doi:10.3928/01477447-20151228-01
2. Duncan SFM, Hendawi TK, Sperling J, Kakinoki R, Hartsock L. iPhone and iPad use in orthopedic surgery. *Ochsner J*. 2015;15:52-57.
3. Blaya JA, Fraser HS, Holt B. E-health technologies show promise in developing countries. *Health Aff (Millwood)*. 2010;29:244-251. doi:10.1377/hlthaff.2009.0894
4. Alexander I. The impact of future trends in electronic data collection on musculoskeletal research and evidence-based orthopaedic care. *Arthroscopy*. 2003;19:1007-1011. doi:10.1016/j.arthro.2003.09.024
5. Payne KFB, Wharrad H, Watts K. Smartphone and medical related App use among medical students and junior doctors in the United Kingdom (UK): a regional survey. *BMC Med Inform Decis Mak*. 2012;12:121. doi:10.1186/1472-6947-12-121
6. Burdette SD, Herchline TE, Oehler R. Surfing the web: practicing medicine in a technological age: using smartphones in clinical practice. *Clin Infect Dis*. 2008;47:117-122. doi:10.1086/588788
7. Franko OI. Smartphone apps for orthopaedic surgeons. *Clin Orthop Relat Res*. 2011;469:2042-2048. doi:10.1007/s11999-011-1904-0
8. Marcon RM, Cristante AF, Teixeira WJ, Narasaki DK, Oliveira RP, de Barros Filho TE. Fractures of the cervical spine. *Clinics (Sao Paulo)*. 2013;68:1455-1461. doi:10.6061/clinics/2013(11)12
9. Magerl F, Aebi M, Gertzbein SD, Harms J, Nazarian S. A comprehensive classification of thoracic and lumbar injuries. *Eur Spine J*. 1994;3:184-201. doi:10.1007/BF02221591
10. Oner FC, Ramos LM, Simmermacher RK, et al. Classification of thoracic and lumbar spine fractures: problems of reproducibility. A study of 53 patients using CT and MRI. *Eur Spine J*. 2002;11:235-245. doi:10.1007/s00586-001-0364-8
11. Vaccaro AR, Zeiller SC, Hulbert RJ, et al. The thoracolumbar injury severity score: a proposed treatment algorithm. *J Spinal Disord Tech*. 2005;18:209-215. doi:10.1097/01.bsd.0000164608.63526.56
12. Vaccaro AR, Lehman RA Jr, Hurlbert RJ, et al. A new classification of thoracolumbar injuries: the importance of injury morphology, the integrity of the posterior ligamentous complex, and neurologic status. *Spine (Phila Pa 1976)*. 2005;30:2325-2333.
13. Vaccaro AR, Oner C, Kepler CK, et al. AOSpine Thoracolumbar Spine Injury Classification System: fracture description, neurological status, and key modifiers. *Spine (Phila Pa 1976)*. 2013;38:2028-2037. doi:10.1097/BRS.0b013e3182a8a381
14. AO Foundation. Trauma, Spine, VET, CMF, Neuro, Recon. <https://www.aofoundation.org/Structure/Pages/default.aspx>. Accessed June 14, 2019.
15. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*. 1977;33:363-374. doi:10.2307/2529786

16. McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*. 1947;12:153-157. doi:10.1007/BF02295996
17. Sadiqi S, Oner FC, Dvorak MF, Aarabi B, Schroeder GD, Vaccaro AR. The influence of spine surgeons' experience on the classification and intraobserver reliability of the novel AOSpine Thoracolumbar Spine Injury Classification System—an international study. *Spine (Phila Pa 1976)*. 2015;40:E1250-E1256. doi:10.1097/BRS.0000000000001042
18. Grasso MA, Yen MJ, Mintz ML. Survey of handheld computing among medical students. *Comput Methods Programs Biomed*. 2006;82:196-202. doi:10.1016/j.cmpb.2006.03.005
19. Warnock GL. The use of apps in surgery. *Can J Surg*. 2012;55:77-78. doi:10.1503/cjs.006412
20. Carey E, Payne KF, Ahmed N, Goodson A. The benefit of the smartphone in oral and maxillofacial surgery: smartphone use among maxillofacial surgery trainees and iPhone apps for the maxillofacial surgeon. *J Maxillofac Oral Surg*. 2015;14:131-137. doi:10.1007/s12663-013-0550-3
21. Cortez NG, Cohen IG, Kesselheim AS. FDA regulation of mobile health technologies. *N Engl J Med*. 2014;371:372-379. doi:10.1056/NEJMhle1403384
22. US Food and Drug Administration. Policy for device software functions and mobile medical applications: guidance for industry and Food and Drug Administration staff. <https://www.fda.gov/vaccines-blood-biologics/guidance-compliance-regulatory-information->. Published September 25, 2013. Accessed October 31, 2019.
23. Buijink AWG, Visser BJ, Marshall L. Medical apps for smartphones: lack of evidence undermines quality and safety. *Evid Based Med*. 2013;18:90-92. doi:10.1136/eb-2012-100885