Content validation of the progressive collapsing foot deformity classification

Mingjie Zhu1, Moustafa A. Maher1,2, Wanjun Gu3, Kenneth J. Hunt1, Mark S. Myerson1,4, Shuyuan Li1,4

1. Department of Orthopaedic Surgery, University of Colorado School of Medicine, Denver, CO, USA.
2. Cairo University Kasr Alainy Faculty of Medicine, Orthopaedic Surgery Department, Cairo, Egypt.
3. University of California San Diego School of Medicine, La Jolla, CA, USA.
4. Steps2Walk, Inc, Denver, CO, USA.

Abstract

Objective: The aim of this study was to validate the content accuracy of the PCFD classification.

Methods: A survey-based study distributed through international foot and ankle programs among surgeons with vast experience in practice to analyze the terminology and interpretations used in the PCFD classification. A returned survey with completion of all questions filled out was considered a valid record. Descriptive statistical analysis was applied using SAS version 9.4 for data processing, statistical analysis, and visualization.

Results: Eighty-two valid returned surveys from surgeons in 22 countries with a mean of 16 years in clinical practice were included. Among them, 80.5% of the participants considered the PCFD classification helpful in guiding decision-making, 79.3% thought it helped facilitate diagnosis and documentation, 58.5% found it easy to use, 30.5% were unlikely to use the classification, and 29.3% noted that the interpretation of the classification was not clear. Regarding the accuracy, clarity, and clinical relevance of terminology, 42.7% had difficulty in using increased foot and ankle offset, 35.4% had difficulty in using increased hindfoot moment arm, 19.5% found peritalar subluxation not clear, 13.4% found the term sinus tarsi impingement an unclear description, and 8.5% found forefoot varus difficult to diagnose.

Conclusions: This international survey-based study provides readers with insights into the content of the PCFD classification. The findings indicate that some terminologies used in the PCFD classification are not universally understood. The authors recommend that modifications may be beneficial to enhance the accuracy and user-friendliness of the PCFD classification.

Level of Evidence II; Retrospective study.

Keywords: Adult; Data accuracy; Flatfoot; Foot deformities; Posterior tibial tendon dysfunction.

Introduction

The exact etiology of the adult-acquired flatfoot deformity (AAFD) remains unclear. However, it was initially described as a result of posterior tibial tendon failure, hence the original term posterior tibial tendon dysfunction7-9. Johnson and Strom were the first to classify this disease process as a 3-stage classification system10, modified by Myerson11 and subsequently by Bluman et al.12. To improve standardization of diagnosis and treatment, a consensus group recently introduced the terminology and classification of Progressive Collapsing Foot Deformity (PCFD). This classification includes two stages (stage 1 flexible deformity and stage 2 rigid deformity) and five classes (A = hindfoot valgus deformity; B = midfoot/forefoot abduction deformity; C...
The PCFD classification has been found to have moderate interobserver reliability, very good intraobserver agreement, and good diagnostic accuracy. The aim of this study is to validate the content accuracy of the PCFD classification system, an important step when a newly introduced classification system is being widely adopted for clinical use.

**Methods**

**Study design**

This study was exempted from Institutional Review Board (IRB) review. An online survey was distributed through REDCap to foot and ankle fellowship training programs worldwide to analyze the terminology and interpretations associated with the PCFD classification (Appendix 1). Foot and ankle consultants from 22 countries who willingly agreed to participate in this study were included. The demographic data from the participants were collected, including practice country/region, years of clinical practice since fellowship, and surgical volume per month before the pandemic in treating foot/ankle and flatfoot deformities. All data was anonymous, and no personal information was collected (Figure 1).

Each participant was provided with detailed instructions on using the PCFD classification system with the original paper (Figure 2), a video presentation, and one case example to demonstrate how to use the classification. Following the instructions, three clinical cases were presented with a combination of history, physical examination videos,

---

**Figure 1.** Demographic data of the surgeon’s participant.

**Figure 2.** The progressive collapsing foot deformity (PCFD) classification.

**Source:** Myerson et al. (2020).
photographs, and radiographic images\(^2\) to test participants’ understanding of the PCFD classification. The full survey was then provided to gain participants’ feedback on the PCFD classification system, including each stage and class. For general opinions on the PCFD classification, the questions focused on necessity, applicability, and ease of use. For the content of each stage and class, questions focused on ease of use, accuracy, and clarity. A rating scale of 1–5 (“1” for strongly disagree, “2” for disagree, “3” for neutral, “4” for agree, and “5” for strongly agree) was provided to assess the degree to which a participant agreed or disagreed about an item, then multiple-answer questions followed to investigate the reasons for agreeing or disagreeing. Open-answer questions were also provided to collect subjective opinions not covered by answer options in the closed-ended questions.

Survey evaluation and data processing

Participants without foot and ankle fellowship training or uncompleted surveys missing information in one or more sections/questions were considered invalid responses. The online link for the survey was kept active for three months and then closed for data processing and evaluation. Invalid responses were excluded during the data cleaning process.

Statistical analysis

The SAS software version 9.4 (SAS Institute Inc., 100 SAS Campus Drive, Cary, North Carolina 27513, USA) was used for data processing, statistical analysis, and visualization. Descriptive statistical analysis was applied. Means with standard deviations (SDs) were used to describe numerical data. Bar charts were utilized to describe categorical data, and pie charts to describe proportional data. Each selection from multiple-answer questions was counted independently to reflect actual weight and option percentage. All numbers were accurate to one decimal place. Subjective opinions and evaluations were analyzed separately.

Results

Eighty-two valid anonymous surveys were received, meeting the recommended sample size for an e-survey. This number was considered statistically sufficient for inquiring into PCFD content\(^8\). The 82 surgeon participants had been in practice for a mean of 15.9 years, treating 17.5 cases of foot and ankle deformities, including 3.42 cases of flatfoot monthly (Table 1). Over 90% of the participants agreed with the importance of a classification system for the flatfoot for clinical and research purposes, 73.2% thought they did not need a classification system for diagnostic and treatment purposes, 69.5% would use the classification in their daily practice while 14.7% would not. The latter group would not use the PCFD classification; 66.7% found it difficult to use, 41.7% found it not advantageous, and 16.7% stated they had never used a classification system (Figure 3). Out of the total, 46.3% agreed that weight-bearing computed tomography (WBCT) would benefit their patients, 22% did not agree with using WBCT either due to lack of access or thought that WBCT was unsuitable for incorporation into clinical practice, and 29.3% thought it would be beneficial to have independent diagnostic criteria for flatfoot based on clinical and radiographic findings, and WBCT results depending on accessibility to these three types of examination, instead of mixing the different types of exams.

Regarding the five classes, for classes A, B, C, and E, over 90% of the participants were familiar with using it and comfortable incorporating it into their practice, while the familiarity and comfort in class D were 87.8% and 80.5%, respectively. Among the 19.5% who were not comfortable with using it in class D, 14.6% considered the definition (peritalar subluxation) unclear, 3.7% were uncomfortable with the concept, 1.2% had no idea what the term “peritalar subluxation” itself was, and 2.4% had no idea how to diagnose it (Figure 5).

---

**Table 1.** Demographic data of participant’s clinical experience.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sample Size</th>
<th>Range</th>
<th>SD</th>
<th>Mean (95%CI)</th>
<th>Median (Q1, Q3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years of clinical practice since fellowship</td>
<td>82</td>
<td>0-40</td>
<td>8.60</td>
<td>15.90 (14.01, 17.79)</td>
<td>15 (10, 21)</td>
</tr>
<tr>
<td>Flatfoot deformities/ Month</td>
<td>82</td>
<td>0-15</td>
<td>3.26</td>
<td>3.42 (2.71, 4.14)</td>
<td>2 (1, 4)</td>
</tr>
<tr>
<td>Foot &amp; ankle deformities/ Month</td>
<td>81</td>
<td>1-60</td>
<td>14.09</td>
<td>17.57 (14.45, 20.68)</td>
<td>15 (5, 25)</td>
</tr>
</tbody>
</table>

**Figure 3.** Reasons not to use the PDCF classification.
The ease of use for diagnosis criteria of each class was investigated, especially concerning accuracy, clarity, and relevance. Among the total, 35.4% reported that “increased hindfoot moment arm” in class A was difficult to use (19.5% considered the description unclear, 11% did not know how to diagnose it, 5% considered the term inaccurate, and 5% considered this concept was not suitable for clinical use), 42.7% reported that “increased foot and ankle offset” in class A was difficult to use (14.6% considered the description unclear, 13.4% did not know the term, 8.5% did not know how to diagnosis it, and 61% considered this concept was not suitable for clinical use) (Figure 6), 13.4% reported that the term “Sinus tarsi impingement” in class B was difficult to use (4.9% considered “sinus tarsi impingement” had an unclear description with a missing reference range for diagnosis, 4.9% did not know how to diagnose it, 3.7% considered the term inaccurate, and 1.2% were not familiar with the term itself) (Figure 7).

**Discussion**

Reliability (i.e., precision) and validity (i.e., accuracy) are required for a classification system to be able to guide users (10-12). The precision of a classification system is generally evaluated by testing inter- and intra-observer reliabilities (10, 13), while accuracy applies to the content of the classification, such as terminology and its interpretation. In other words, does the instrument measure what it is intended to do (14)? A good classification for clinical practice and research is expected to categorize and interpret the attribute of
interest with precision and accuracy\(^{(16,46)}\). Lee et al.\(^{(15)}\) tested the reliability, Li et al.\(^{(9)}\) evaluated the diagnostic accuracy, while our study investigated the content validity of the PCFD classification.

Most participants in this study agreed that a new classification system was needed for the flatfoot deformity, consistent with the statements of the consensus group for introducing the PCFD classification\(^{(4)}\). Regarding the “easy to use” aspect of this new classification, 14.7% reported “difficult to use,” with 75% indicating difficulty, and the primary reason for these difficulties was the diagnostic criteria.

For the five classes, 100% had no difficulty with class A (hindfoot valgus), a very commonly applied term\(^{(2)}\) supporting the previous finding that Class A had a high diagnostic rate (96.8%)\(^{(5)}\). However, regarding how to evaluate hindfoot valgus, the “hindfoot moment arm” and “foot and ankle offset” metrics had a high difficult-to-use rate of 35.4% and 42.7%, respectively. To date, there is no consensus on the best method for radiographic hindfoot measurement\(^{(17-20)}\). Saltzman and Khoury first described the hindfoot moment arm in 1995\(^{(29)}\) using the lowest point of the calcaneus as the landmark to measure the hindfoot alignment. Arena et al.\(^{(22)}\) reported that the hindfoot moment arm was a valid method with high intra- and inter-observer reliability with adequate delineation of the anatomical landmarks through a simple measurement, avoiding angular measurements, which are still imprecise. Neri et al.\(^{(23)}\) compared multiple hindfoot radiographic alignment methods and reported that the hindfoot moment arm was subject to poor reproducibility among observers when trying to reach a consensus on the location of the lowest point of the calcaneus. The high susceptibility of hindfoot radiograph analysis to multiple errors led research to explore more accurate analysis on WBCT scans\(^{(9,20,24)}\). As a result, new measurements like “foot and ankle offset” and “calcaneal moment arm” were introduced\(^{(25)}\). “Foot and ankle offset” was developed as a 3D biometric WBCT measurement to represent an optimized biomechanical assessment of the relationship between the tripod of the foot and the center of the ankle joint\(^{(25,26)}\). It is a measurement exclusively used on WBCT scans\(^{(26-28)}\).

Regarding the concepts of “sinus tarsi impingement” in class B and “subtalar joint subluxation/subfibular impingement” in class D, 13.4% and 8.5% considered the description unclear. Johnson and Strom, in their early descriptions of tibialis posterior tendon rupture, showed that pain could develop in the lateral tarsal region, worsening the flatfoot deformity and causing bony contact between the inferior aspect of the talus and the dorsal aspect of the calcaneus\(^{(2)}\). Malicky et al.\(^{(20)}\) studied both concepts and reported the coexistence of calcaneofibular and sinus tarsi impingement in 100% of cases. However, Jeng et al.\(^{(23)}\) stated that calcaneofibular and sinus tarsi impingement could occur separately in a reasonable percentage of cases. Moreover, Lavelee et al.\(^{(31)}\) studied the correlation of peritalar subluxation with calcaneofibular and sinus tarsi impingements and reported the correlation between the peritalar subluxation and sinus tarsi impingement was statistically significant. Recently, Kim et al.\(^{(32)}\) showed a strong “predictive value” correlating sinus tarsi impingement with talonavicular coverage and calcaneofibular impingement with hindfoot valgus in weight-bearing radiographs compared with WBCT\(^{(33,34)}\). There are no cut-off points to diagnose sinus tarsi and subfibular/calcaneofibular impingement clinically, radiographically, or on WBCT\(^{(31,32,35)}\). Li et al.\(^{(9)}\) reported that users had difficulties with identifying and diagnosing subtalar subluxation and those impingements, explaining why there was a low misdiagnosis rate in class B (17.48%) and class D (26%)\(^{(8)}\).

To incorporate the new WBCT technology in the diagnostic criteria of the PCFD classification, 46.3% of participants strongly agreed. WBCT has proved its significance in clinical practice by offering unique advantages, including improved spatial resolution, multiplanar 3-D assessments, minimizing rotational and positional bias, and bony superimposition under physiologic loading\(^{(26-39)}\). However, 22% of participants had doubts about the use of WBCT. In the PCFD classification, class D must mainly be diagnosed on WBCT scans; 19% of the participants in our study had problems with class D itself. Li et al.\(^{(9)}\) reported a 26% misdiagnosis rate in class D, most of which were underdiagnosed\(^{(45)}\). It might be possible that participants were avoiding the diagnosis of this class because they may not have been familiar with it or were uncertain about how to apply it. Lee et al.\(^{(7)}\) demonstrated that class D was the least agreed upon among observers when relying solely on weight-bearing radiographs. They emphasized the necessity of using WBCT to diagnose this class accurately\(^{(7)}\). In Lavelee’s comparison with the refined classification, class D also received a lower interobserver reliability consensus as it was difficult to diagnose clinically and radiographically\(^{(40)}\).

These results corroborate the findings of our study. In addition, foot and ankle offset\(^{(22,28)}\), sinus tarsi impingement\(^{(29-31)}\), and subfibular impingement\(^{(29,30,31,44)}\) are concepts diagnosed mainly on WBCT. About one-third of the participants (29.3%) suggested that in the PCFD classification, those diagnosis criteria using WBCT should be separated from others depending on clinical examinations and weight-bearing radiographs.

Class C represents both forefoot varus and medial column instability and participants in our study had no difficulty with this class. Lavelee et al.\(^{(42)}\) reported lower interobserver reliability in class C than other classes, which might be explained as forefoot varus and medial column instability being two separate entities described in the refined classification by Bluman et al.\(^{(5)}\). As a result, class C had a 9.52% underdiagnosis rate compared to a 1.6% overdiagnosis rate, that is, among 100 PCFD surgeon users in this study, 9.52% surgeons might miss a class C deformity. Moreover, the diagnostic cut-off point for medial column instability is obscure, making it susceptible to interobserver disagreement on clinical and radiological diagnostic reliability\(^{(42-46)}\).

Class E refers to “ankle instability” and, in our study, had a high percentage of acceptance (Figure 8), and others had a high interobserver agreement\(^{(7,40)}\). Evidence, however,
supports the author’s opinion that “ankle instability” is not an accurate term for a valgus ankle associated with a flatfoot since the ankle is not necessarily unstable. The authors suggest to use “ankle valgus deformity” for class E(5,47-49).

There are limitations in this study. Firstly, it did not cover extensive numbers of foot and ankle fellowship training programs worldwide. Secondly, a language barrier could limit some non-native English speaker’s understanding of the PCFD classification and the survey used in this study. Thirdly, there was an inevitable overlap in this study population with that used in previous studies. The authors believe that repeating similar survey investigations in the same study group is not ideal since screening the target population multiple times could increase bias and decrease the participation rate.

**Conclusion**

This international survey-based study provides readers with insights into the content of the PCFD classification. The findings indicate that some terminologies used in the PCFD classification are not universally understood. The authors recommend that modifications may be beneficial to enhance the accuracy and user-friendliness of the PCFD classification.

**Authors’ contributions:** Each author contributed individually and significantly to the development of this article: MZ (https://orcid.org/0000-0002-9685-4048) Conceived and planned the activities that led to the study, data collection, statistical analysis, interpreted the results of the study, bibliographic review, participated in the review process, formatting of the article, approved the final version. MAM (https://orcid.org/0000-0001-7031-5137) interpreted the results of the study, bibliographic review, participated in the review process, formatting of the article; WG (https://orcid.org/0000-0002-7342-7000) Statistical analysis, interpreted the results of the study, participated in the review process; KJH (https://orcid.org/0000-0002-8369-8744) Participated in the review process, formatting of the article, approved the final version; MSM (https://orcid.org/0000-0003-1238-8455) Conceived and planned the activities that led to the study, clinical examination, participated in the review process, approved the final version. All authors read and approved the final manuscript. ORCID (Open Researcher and Contributor ID).

**References**
