

ORIGINAL ARTICLE

Volume 18 Issue 2 2023

DOI: 10.21315/aos2023.1802.OA03

ARTICLE INFO

Submitted: 03/07/2023

Accepted: 10/12/2023

Online: 20/12/2023

Assessing Inter-rater Reliability of the Dental Practicality Index and the American Association of Endodontists Endodontic Case Difficulty Assessment Form among Undergraduates

Amy Kia Cheen Liew^a, Dalia Abdullah^{b*}, Eason Soo^b, Norziha Yahaya^b

^aDepartment of Family Oral Health, Faculty of Dentistry, Universiti Kebangsaan Malaysia, 50300 Kuala Lumpur, Malaysia

^bDepartment of Restorative Dentistry, Faculty of Dentistry, Universiti Kebangsaan Malaysia, 50300 Kuala Lumpur, Malaysia

*Corresponding author: daliaabdullah@ukm.edu.my

To cite this article: Liew AKC, Abdullah D, Soo E, Yahaya N (2023). Assessing inter-rater reliability of the Dental Practicality Index and the American Association of Endodontists Endodontic Case Difficulty Assessment form among undergraduates. *Arch Orofac Sci*, **18**(2): 153–165. <https://doi.org/10.21315/aos2023.1802.OA03>

To link to this article: <https://doi.org/10.21315/aos2023.1802.OA03>

ABSTRACT

Dental Practicality Index (DPI) and American Association of Endodontists Endodontic Case Difficulty Assessment (AAECDA) form potentially can guide clinicians in making clinical decisions and triaging in large practices and academic settings. Nonetheless, the reliability and validity should be evaluated before institution-wide implementation. This study aimed to evaluate the inter-rater reliability of the DPI and AAECDA forms. Ten randomly selected, trained students rated 25 cases with both forms. The item-by-item inter-rater and overall reliability were estimated with Gwet's agreement coefficient (AC2) and intraclass correlation coefficient (ICC), respectively. The association between clinical decisions and the scores was analysed with the Generalised Estimating Equation. The inter-rater reliability of DPI was generally very good (AC2 = 0.81–1.00), except context (good; AC2 = 0.718; 95% confidence interval [CI] = 0.575–0.861). The inter-rater reliability of AAECDA was generally very good (AC2 = 0.81–1.00) and good (AC2 = 0.61–0.80), except the radiographic appearance of the canal(s) (fair; AC2 = 0.424, 95% CI = 0.263–0.585). Moderate overall inter-rater reliability of AAECDA (ICC = 0.53, 95% CI = 0.38–0.70) and DPI (ICC = 0.62, 95% CI = 0.48–0.77) was observed. Referral to an endodontist was positively associated with AAECDA score (odds ratio [OR] = 1.323, 95% CI = 1.145–1.52, $p < 0.001$). The decision of tooth extraction was positively associated with the DPI score (OR = 1.983, 95% CI = 1.539–2.555; $p < 0.001$). In conclusion, DPI and AAECDA are methods with moderate inter-rater reliability when used among dental students.

Keywords: *Dental education; dental student; referral; root canal; tooth extraction*

INTRODUCTION

When a patient presents with a tooth potentially requiring endodontic treatment, a clinician needs to make several decisions. Messer (1999) listed three important considerations. The first consideration is the appropriateness of endodontic treatment, including objective clinical findings, health conditions, and patient motivation. If the answer is affirmative, the second consideration is whether the clinician can competently provide standard care or whether a referral to a specialist is required. The third consideration concerns the prognostication of the tooth, which includes the endodontic treatment and factors such as restorability and periodontal status.

The Dental Practicality Index (DPI) was developed to assess the tooth's structural integrity, endodontic and periodontal treatment needs, and contextual aspects for treatment planning (Dawood & Patel, 2017). In contrast, the American Association of Endodontists Endodontic Case Difficulty Assessment (AAECDA) form focuses on a single tooth needing endodontic treatment, taking patient, diagnostic, treatment, and other factors into consideration for case selection (AAE, 2019). Both instruments are potentially useful for guiding individual clinicians in making clinical decisions and triaging in large practices and academic settings. Nonetheless, the prerequisites of reliability and validity should be evaluated before institution-wide implementation.

The percentage agreement between dental students trained with DPI and a consensus expert panel ranged from 31% to 91%, depending on the cases (Hamer *et al.*, 2021). The intra-rater reliability of DPI was reported at $\kappa = 0.59$, suggesting some degree of subjectivity (Tifooni *et al.*, 2019). AAECDA is taught in a quarter of all United States-accredited dental schools, surpassing other assessment methods (Kim *et al.*,

2019). The intra- and inter-rater reliability of AAECDA between two calibrated researchers were reported at $\kappa = 0.90$ (Fezai & Al-Salehi, 2019). However, in a multi-centre study among undergraduate dental students, the overall reliability of AAECDA was reported at $\kappa = 0.176$ (Shah *et al.*, 2020).

Accessibility, cost-effectiveness, and efficiency are the main advantages of fulfilling restorative needs at the primary care level. Nonetheless, inappropriate care provided by unskilled clinicians may consume additional resources in the long run than would have been required if the care had been provided by a properly trained and supported clinician from the start. Therefore, communication between primary and secondary care regarding agreed-upon referral criteria is critical (Alani & Bishop, 2012). This study aimed to evaluate the inter-rater reliability of each individual item within the DPI and AAECDA, as well as the overall reliability, for assessing teeth requiring endodontic treatment. The study explored the extent to which the raters' decision agreed with the DPI and AAECDA recommendations regarding treatment and referral. Although it seems ideal to implement both the DPI and AAECDA for screening, there would be substantial overlap in terms of restorability, periodontal and endodontic assessment. Implementing both methods for each case in a busy practice would be redundant. Hence, the clinicians' likelihood of using either of these assessment methods was also evaluated.

MATERIAL AND METHODS

The conduct of the study is summarised in Fig. 1. Ethics approval was granted by the Research Ethics Committee (UKM PPI/111/8/JEP-2021-196). Permission to use the assessment tools was obtained from the developers.

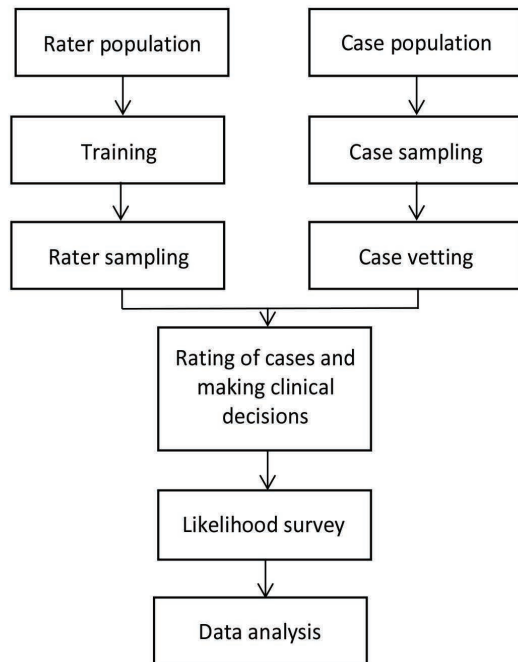


Fig. 1 Flowchart of the study – Raters: Trained to analyse cases; Cases: Selected for rating based on specific criteria; Rating: Decisions made based on case analysis; Survey: Raters indicated their likelihood on using the forms, and data were analysed.

Rater Selection

The rater population consisted of all final year students ($n = 42$), cohort 2021–2022, in the Doctor of Dental Surgery programme, Universiti Kebangsaan Malaysia. All the students had completed the theoretical didactic teaching, the pre-clinical endodontic exercises, and competency tests and passed the endodontic course in Year 3. The additional didactic and clinical endodontic practice continued in Year 4 and was ongoing during the study period. In the pilot phase, a three-hour training session consisting of a lecture on DPI and AAECDA was conducted for the students. They were then required to assess a patient's case with both instruments and decide on a treatment or referral plan. The written assessment and decision were submitted. Discussion ensued to clarify further the items listed in both instruments. Cognizant of the material presented, the students verbally acknowledge their understanding through self-reporting.

Only students who completed the training were eligible for rater selection. The sample size of the raters (n_r) was estimated as follows (Gwet, 2014):

$$n_r = 2/cv,$$

whereby cv is the anticipated coefficient of variation, determined at 20%. Hence, the required number of raters was 10. The students were selected by drawing lots. Written consent was obtained from the raters.

Case Selection

The case population comprised all the cases from patients who attended the consultation sessions at the Endodontic Clinic from 2015 to 2020, as listed in the appointment spreadsheet. Only cases with retrievable and complete preoperative clinical information were included. Cases diagnosed as previously treated were excluded because these always warrant a referral to endodontists or postgraduate trainees in our institution. Each case contributed only one tooth for assessment. The sample size of cases required (n_c) was computed using the following equation (Gwet, 2014):

$$n_c = 1/E^2,$$

whereby E represents the acceptable margin of error, determined at 20%. Hence, the minimal sample size required was 25 cases.

Purposive sampling was performed to ensure a variety of cases with different levels of complexity. Permission to retrieve the paper-based dental records of the selected cases was obtained from the Clinical Service Unit. Personal identifiers and clinical outcomes were not extracted or included in the case scenarios (Fig. 2). Intraoral photos and radiographs were digitised. Two endodontists and a prosthodontist vetted the case scenarios.

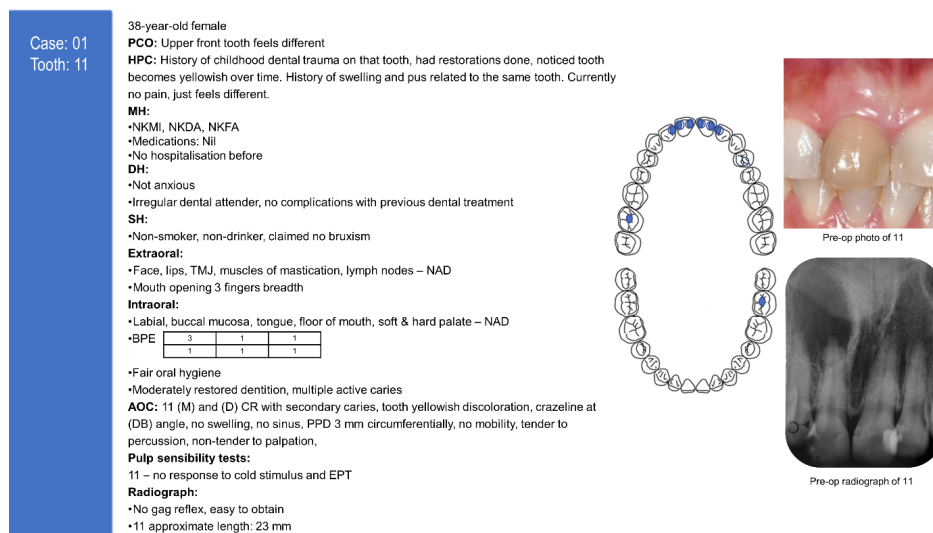


Fig. 2 One of the clinical case scenarios used for rating.

Rating of Cases

The rating of cases was conducted at the information technology laboratory with the raters allocated a computer each, distanced one meter apart. Each rater rated all 25 cases consecutively using both DPI and AAECDA. They were also asked to answer multiple-choice questions on each case's appropriate treatment and treatment provider. Each case was given an allotted time of approximately 10 mins for rating, but the raters were free to move through the case scenarios and complete the rating before the allotted time. Discussions and access to external information were prohibited. After rating all the cases, the raters were asked to respond to two questions with a five-point Likert scale on how likely they are to use the DPI and the AAECDA form in their clinical practice.

Data Analysis

The item-by-item inter-rater reliability using AAECDA and DPI was assessed using AgreeStat 360 to yield Gwet's agreement coefficient (AC_2). Altman's interpretation (Altman, 1991) was used to report the strength of agreement. The scoring system for AAECDA involved assigning scores of

1, 2, and 5 to indicate minimal, moderate, and high difficulty for each item. The total score was determined by summing these individual scores (AAE, 2005). In contrast, DPI used a scoring system where 0, 1, 2, and 6 were assigned to represent no treatment needed, simple treatment required, complex treatment required, and impractical to treat, respectively. The total DPI score was calculated by summing the scores in each of the four specified categories (Dawood & Patel, 2017). The overall inter-rater reliability of DPI and AAECDA was assessed with the two-way random rater effects analysis of the variance model to compute the intraclass correlation coefficient (ICC). The association of clinical decision and the assessment categorisation was assessed with the Generalised Estimating Equation (GEE) in SPSS 22.0.

RESULTS

Item-by-item Gwet's AC_2 ranged from 0.718 to 0.917 for DPI and from 0.424 to 1.000 for AAECDA (Table 1). Moderate inter-rater reliability was observed in the total AAECDA and DPI scores (Table 2).

Table 1 Item-by-item inter-rater reliability for DPI and AAECDA

	Gwet's AC2	95% CI	Altman's interpretation
DPI			
Structure integrity	0.824	(0.764, 0.883)	Very good
Periodontal treatment need	0.917	(0.837, 0.997)	Very good
Endodontic treatment need	0.817	(0.743, 0.892)	Very good
Context	0.718	(0.575, 0.861)	Good
AAECDA			
Medical history	0.968	(0.926, 1.000)	Very good
Anaesthesia	1.000	-	-
Patient disposition	1.000	-	-
Ability to open mouth	1.000	-	-
Gag reflex	1.000	-	-
Emergency condition	0.878	(0.789, 0.967)	Very good
Diagnosis	0.930	(0.869, 0.992)	Very good
Radiographic difficulties	0.950	(0.901, 0.998)	Very good
Tooth type	1.000	-	-
Inclination	0.683	(0.515, 0.851)	Moderate
Rotation	0.955	(0.905, 1.000)	Very good
Tooth isolation	0.772	(0.643, 0.900)	Good
Crown morphology	0.777	(0.691, 0.863)	Good
Canal and root curvature	0.817	(0.683, 0.952)	Good
Canal and root morphology	0.957	(0.915, 0.999)	Very good
Apical opening	0.937	(0.885, 0.989)	Very good
Radiographic appearance of canal(s)	0.424	(0.263, 0.585)	Fair
Resorption	0.742	(0.614, 0.870)	Good
Trauma history	0.952	(0.885, 1.000)	Very good
Endodontic treatment history	0.878	(0.782, 0.974)	Very good
Periodontal-endodontic condition	0.911	(0.815, 1.000)	Very good

Table 2 Overall inter-rater reliability for DPI and AAECDA

	DPI total score	AAECDA total score
Case variance	4.160	14.002
Rater variance	0.081	2.707
Error variance	2.486	9.697
Intraclass correlation (95% CI)	0.618 (0.48, 0.77)	0.530 (0.38, 0.70)
Interpretation	Moderate	Moderate

A total of 250 rater-case pairs were included in the GEE model testing for the association between the DPI total score and the decision to extract the tooth. The range of scores was 0 to 16. For every unit increase in DPI total score, there were 1.983 times increased odds of the decision to extract the tooth (95% CI = 1.539–2.555, $p < 0.001$).

The DPI endodontic treatment need score and AAECDA total score were dependent (Kendall’s coefficient = 0.366, $p < 0.001$). Hence, only the AAECDA total score was fitted in the GEE model. Excluding cases

regarded by the raters as needing extraction, 232 rater-case pairs were tested for the association between the total AAECDA score and the decision to refer to endodontists. The range of scores was 19 to 46. For every unit increase of the total AAECDA score, there were 1.323 times increased odds of the decision to refer to endodontists (OR = 1.323, 95% CI = 1.145–1.529, $p < 0.001$). The diverging stacked bar chart (Fig. 3) summarises the raters’ perceived likelihood to apply the assessment methods in clinical practice.

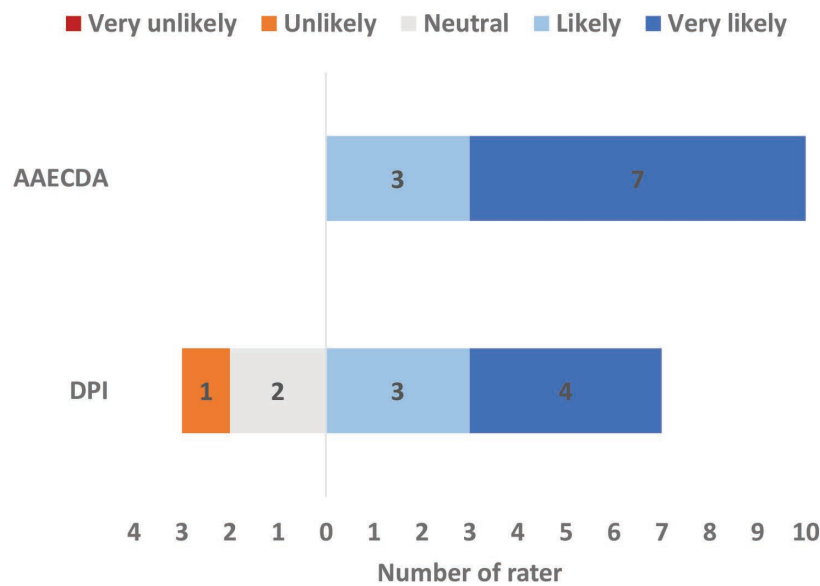


Fig. 3 Likelihood of the raters to apply AAECDA and DPI in clinical practice.

DISCUSSION

While metacognition and improvement of knowledge and skills remain central to clinical management, cognitive aids such as clinical checklists, guidelines, and algorithms assist clinicians in streamlining decision-making (Thammasitboon & Cutrer, 2013). Because the initial development of both DPI and AAECDA did not follow the conventional phases of concept identification, item construction, validity, and reliability testing (Davis, 1996), evidence

should be sought from later studies to support the appropriate interpretation of the information collected using the instruments. The interval estimation of the overall inter-rater reliability and the novel item-by-item inter-reliability in this study would add scientific rigour to these instruments.

Even though the clinical photos used in the present study were not to scale, the raters could reliably assess the “structure integrity” in DPI, primarily by the number of residual coronal walls. Teeth with less

than 30% of the original structure had unfavourable outcomes (Al-Nuaimi *et al.*, 2017), but restorability is also retroactively dependent on the type of final restoration. When extra-coronal restoration is required, the ferrule effect should be optimised for biomechanical function (Sorensen & Engelman, 1990). However, in structure integrity, whether “inadequate structure for ferrule” is “impractical to treat” could be subject to debate because the ferrule effect can be gained secondarily through orthodontic extrusion, surgical extrusion, and surgical crown lengthening. Poor restorability should be considered only if there is a lack of coronal structure, a favourable crown-to-root ratio, and adequate root length for extrusion (Sorensen & Engelman, 1990). Even then, the minimum root length required is controversial (Su *et al.*, 2014). Apart from the remaining tooth structure, the restorability of a tooth should be considered in the context of functional and lateral stresses (Dietschi *et al.*, 2008).

The inter-rater reliability for “periodontal treatment need” is very good. Based largely on the Basic Periodontal Examination (BPE), which has a long history of use in our institution, high inter-rater reliability was expected. However, the referral should not be based solely on the BPE score. Instead, it should take into account the complexity score, which considers factors such as medical history, concurrent mucogingival disease, drug interactions, and tobacco smoking. These factors can significantly increase the complexity of cases (British Society of Periodontology, 2011; 2019). Although high inter-rater reliability was observed for “endodontic treatment need”, it depended on AAECDA total score because both assessment methods were provided to the raters simultaneously.

Among the four constructs listed in DPI, “context” has the lowest inter-rater reliability. This may be due to the combination of several unrelated factors into one criterion. The combined scoring for the local context and the general context

may undermine the complexity when several criteria are fulfilled. It is also uncertain whether active periodontal disease reflects a general context or a periodontal treatment need. On one end of the spectrum, criteria that favour tooth retention, such as the history of intravenous bisphosphonates, are listed. However, it is perplexing that, on the other end, potentially life-threatening medical conditions are categorised as “impractical to treat”, despite the recommendation to manage such cases in tertiary care (Dawood & Patel, 2017).

While the item-by-item inter-rater reliability was generally good, the overall reliability of DPI was moderate. Each of the impractical to treat cells contributes six points to the total score; hence, it carries a lot of weight when one or more of these cells is checked. Despite that, the objective criteria for untreatable periodontal disease and untreatable root canal systems were not defined in the index table. Also, multiple selections in the context column were observed.

The medical history categorisation in AAECDA follows the American Society of Anesthesiologists (ASA) physical status classification system. Although very good inter-rater reliability was observed in this study, all the raters rated the two case scenarios of alcohol drinkers as ASA 1. This was against the recent improvisations of classifying current smokers, social alcohol drinkers, or pregnant people into ASA 2 (ASA, 2014). However, this would not be an issue with the latest AAECDA form revision because ASA 1 and ASA 2 are grouped as low difficulty (AAE, 2022). The case scenarios did not vary in terms of anesthesia, patient disposition, ability to open mouth, and Gag reflex. Hence, a conclusion cannot be made concerning the inter-rater reliability of these items.

Emergency condition, diagnosis, and radiographic difficulties achieved very good inter-rater reliability in this study because this information was provided in

the case scenarios. In practice, gathering the appropriate history and conducting a comprehensive examination and investigation are crucial skills that vary between clinicians. About 3.4% of nonodontogenic pain is identified in endodontically treated teeth with persistent pain (Nixdorf *et al.*, 2010). A significant association between female gender, preoperative pain, chronic pain history, orofacial pain history, and persistent pain was reported after successful root canal treatment (Polycarpou *et al.*, 2005).

Moderate agreement was observed for “inclination” because the raters could only assess the inclination of the tooth with intraoral photos and periapical radiographs. Ideally, a study model should be provided for a more accurate tooth inclination and rotation assessment. Good inter-rater reliability was observed for “tooth isolation”. Although heavily broken-down teeth often require pre-endodontic modification to allow rubber dam placement, such modifications can include split-dam technique, pre-endodontic build-up, doughnut technique, canal projection technique, deep margin elevation, gingivectomy, apically repositioned flap, orthodontic extrusion, surgical extrusion, and surgical crown lengthening (Gavriil *et al.*, 2021). The mastery (or the lack) of all these techniques by the clinicians means that there is a wide range of subjectivity on the simplicity or extensiveness of the pre-treatment modification required for rubber dam isolation.

Non-exhaustive choices were listed in crown morphology. It could only be assumed that a tooth with a small restoration falls into the “minimal difficulty” category. Overlapping choices were also observed because both “full coverage restoration” and “porcelain restoration” are applicable in cases with porcelain-fused metal crowns, thereby getting a two-point extra score. Access or removal of ceramics such as zirconia or lithium disilicate varied. Designs such as bonded inlay, onlay, overlay, veneer-lay, and endocrown also pose different challenges for removal and access cavity preparation, even though all of these

can be fabricated with ceramic. Despite these imperfections, a good inter-reliability was achieved for “crown morphology” in the present study.

There are numerous methods to measure root canal curvature. Good inter-rater reliability could be achieved because all the students were taught with only one method (Weine, 2004). Assessing the angle, the radius, and the length of the canal curvature with two radiographic views may reduce ambiguity (Schäfer *et al.*, 2002). Nevertheless, apical opening, as a distinct item with three exhaustive choices, has very good inter-rater reliability in the present study. However, the clinical relevance of the categories < 1 mm, 1 mm–1.5 mm, and > 1.5 mm could be questioned since the Mineral Trioxide Aggregate (MTA) apical plug is suggested when apical size is #60 or larger (Schäfer *et al.*, 2002).

Of all the criteria listed in AAECDA, the radiographic appearance of the canal(s) had the lowest Gwet’s AC₂. Superimposed structures and overlapping images are the inevitable flaws of periapical radiographs. Indirect digitised periapical radiographs, obtained by digital photography of the periapical radiographs, were used in the present study. Various pixel sizes, grey levels, image processes, and zoom factors were applied. Digital enhancement does not improve the locating of small endodontic file tips in relation to the radiographic apex (Peters & Arias, 2016). Likewise, the constricted canal may not be readily identifiable and categorisable in the present study. The potential spectral overlap of “canal(s) and chamber visible but reduced in size”, “indistinct canal path”, and “canal(s) not visible” may also increase inter-rater variability.

Similarly, detecting the small resorptive lesion depends on the resolution of the digitised periapical radiographs. When observed under a scanning electron microscope, many teeth with irreversible pulpitis or pulpal necrosis had internal root

resorption less than 100 μm in depth (Fuge *et al.*, 1998). Such superficial defects may not be detected with periapical radiographs. Also, grading on the severity of apical root resorption is highly variable when the measurement is made with periapical radiographs (Gabor *et al.*, 2012). The selections available for “trauma history” are not exhaustive. So, the raters were instructed to make a mark when trauma history was not applicable. This results in high inter-rater reliability. “No history of trauma” was added in the latest revision of the AAECDA form (AAE, 2022).

Very good inter-rater reliability was achieved for “endodontic treatment history”, suggesting clear-cut selections were provided. However, slight ambiguity is imaginable in categorising teeth that previously underwent vital pulp therapy or regenerative endodontics, especially when the canals contained filling materials such as MTA. This is because the definition for “previously initiated therapy” is applicable when “the tooth has been previously treated by partial endodontic therapy” (e.g., pulpotomy, pulpectomy), while “previously treated” is applicable when “the tooth has been endodontically treated, and the canals are obturated with various filling materials” (Ponder *et al.*, 2013). Likewise, in the AAECDA, the terms “previous access” and “previous endodontic treatment completed” were used to describe endodontic treatment history. In these situations, accurate history from the patient and the previous clinician is crucial in making the correct diagnosis and case difficulty assessment.

Very good inter-rater reliability was achieved for periodontal-endodontic condition, most likely because of the limited case variability. Nonetheless, it is somewhat disputable to use “mild”, “moderate”, and “severe periodontal disease” to classify clinical cases. In 2003, the Centers for Disease Control and Prevention collaborated with the American Academy of Periodontology to streamline measures for the population-based surveillance of periodontal disease.

Following that, case definitions for mild, moderate, and severe periodontitis involving clinical attachment loss and probing depth at two or more interproximal sites were listed, but these were not approved for clinical application (Eke & Genco, 2007). During the 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions, the classification of periodontitis and the endo-periodontal lesion was revamped (Papapanou *et al.*, 2018). Hence, an update of the AAECDA item may be required to be in line with this change.

The moderate overall inter-rater reliability of the AAECDA could be explained by the “check all that apply” nature of the form. For instance, when assessing teeth with multiple roots/canals, some raters would rate the highest complexity point applicable, but others would rate all the canals, resulting in a discrepancy in the total score. Missing data or failing to check one of the items would result in at least one point discrepancy. Although seemingly trivial, it could mean a difference between recommendations to treat or refer because the AAECDA Educator Guide score range recommendations have a point difference between categories (AAE, 2005).

Comparison to previous studies was not possible because inter-rater reliability was reported with Cohen’s kappa for the final categorical data in previous studies (Fezai & Al-Salehi, 2019; Shah *et al.*, 2020) while the ICC was calculated for the summed scores in the present study. Nonetheless, widely divergent inter-rater reliability between studies indicated the influence of the number of raters and single- or multi-centre study on inter-rater reliability.

There was a statistically significant association between the total DPI score and the decision to extract the tooth. This mirrors a previous validity study where teeth with DPI scores of six or more (18.8%) were more likely to be extracted than teeth with DPI scores of less than six (3.9%) after four years (Al-Nuaimi *et al.*, 2020). An alternative

perspective was that 81.2% of teeth deemed impractical to treat survived four years. Hence, the validity of the recommendations for extractions based on a cut-off at a DPI total score of six deserved further investigation. The DPI endodontic treatment need score and AAECDA total score were associated. As the cases were assessed with both methods by the same raters at the same time, it was apparent that the assessments were dependent. Hence, it is unclear how this would affect the inter-rater reliability of each method when used independently. There was a significant association between AAECDA's total score on the decision to refer to endodontists. Indeed, the use of AAECDA categorisation demonstrated that endodontic mishaps (Ungerechts *et al.*, 2014; Haug *et al.*, 2018) and compromised treatment quality (Fezai & Al-Salehi, 2019) were associated with difficult cases.

Direct comparison between DPI and AAECDA with inferential statistics is not sensible because the scope of both instruments varies. However, survey responses were more favourable towards AAECDA than DPI. These findings can be explained by the prior exposure of the raters to both methods because the AAECDA form was used in our institution before the study, whereas the DPI was only introduced to the raters during the study period. A focus group study is best suited to evaluate the raters' perceptions further.

It is impractical to delineate the minimum knowledge or skills required in endodontic education (McCaul *et al.*, 2001). In the present study, students were considered "satisfactorily trained" when the returned case study practice sheet had been answered without overt oddity. Because the students were trained in only one session, it was impossible to understand the biological and mechanical rationale of every item within the stipulated timeframe. Nonetheless, the students' exposure to both instruments in the present study is aligned with the holistic endodontic education that emphasises the deliberation of restorability, endodontic

treatment complexity, and referral need (De Moor *et al.*, 2013). It is known that undergraduate students and general dentists were more likely than endodontic postgraduate students and endodontists to offer a wide range of treatment options, including extraction (Dechouniotis *et al.*, 2010). Hence, findings in the present study could only be generalised to a population of dental students with similar training and experience.

During data entry and analysis, some data was reorganised according to the revised AAECDA form and guidelines (AAE, 2022) published during the writing of this manuscript. However, some changes could not be incorporated in the present study retrospectively, such as the re-categorisation of "medical history" and "periodontal-endodontic condition", the addition of "C-shaped morphology" and "radix ento/paramolaris", and a new construct for "proximity of the root apices to vital structures". It is unclear how these revisions would impact the inter-rater reliability.

CONCLUSION

The inter-rater reliability of several items of DPI and AAECDA is compromised by the lack of clearly defined and mutually exclusive categorisation. Nonetheless, moderate overall inter-rater reliability can be achieved with both instruments, even with limited training. DPI and AAECDA help dental students make clinical decisions on extraction and referral to endodontists, respectively.

REFERENCES

- Alani A, Bishop K (2012). Contemporary issues in the provision of restorative dentistry. *Br Dent J*, 213(4): 163–170. <https://doi.org/10.1038/sj.bdj.2012.725>

- Al-Nuaimi N, Patel S, Austin RS, Mannocci F (2017). A prospective study assessing the effect of coronal tooth structure loss on the outcome of root canal retreatment. *Int Endod J*, **50**(12): 1143–1157. <https://doi.org/10.1111/iej.12760>
- Al-Nuaimi N, Ciapryna S, Chia M, Patel S, Mannocci F (2020). A prospective study on the effect of coronal tooth structure loss on the 4-year clinical survival of root canal retreated teeth and retrospective validation of the Dental Practicality Index. *Int Endod J*, **53**(8): 1040–1049. <https://doi.org/10.1111/iej.13322>
- Altman DG (1991). *Practical Statistics for Medical Research*. London: Chapman & Hall, pp. 403–408.
- American Association of Endodontists (AAE) (2005). Use of Endodontic Case Difficulty Assessment Form: AAE Educator Guide. Retrieved 3 Jul 2023, from <https://www.aae.org/specialty/wp-content/uploads/sites/2/2017/10/educatorguidetocdaf.pdf>
- American Association of Endodontists (AAE) (2019). AAE Endodontic Case Difficulty Assessment Form and Guidelines; 2019. Retrieved 13 Feb 2022, from https://www.aae.org/specialty/wp-content/uploads/sites/2/2019/02/19AAE_CaseDifficultyAssessmentForm.pdf
- American Association of Endodontists (AAE) (2022). AAE Endodontic Case Difficulty Assessment Form and Guidelines; 2022. Retrieved 14 December 2023, from <https://www.aae.org/specialty/wp-content/uploads/sites/2/2022/01/CaseDifficultyAssessmentFormFINAL2022.pdf>
- American Society of Anesthesiologists (ASA) (2014). Statement on ASA Physical Status Classification System; 2014. Retrieved 3 Jul 2023, from <https://www.asahq.org/standards-and-practice-parameters/statement-on-asa-physical-status-classification-system>
- British Society of Periodontology (2011). Referral Policy and Parameters of Care; 2011. Retrieved 3 Jul 2023, from https://www.bsperio.org.uk/assets/downloads/Parameters_of_Care.pdf
- British Society of Periodontology (2019). Basic Periodontal Examination (BPE). Retrieved 3 Jul 2023, from https://www.bsperio.org.uk/assets/downloads/BSP_BPE_Guidelines_2019.pdf
- Davis AE (1996). Instrument development: Getting started. *J Neurosci Nurs*, **28**(3): 204–208. <https://doi.org/10.1097/01376517-199606000-00009>
- Dawood A, Patel S (2017). The Dental Practicality Index—assessing the restorability of teeth. *Br Dent J*, **222**(10): 755–758. <https://doi.org/10.1038/sj.bdj.2017.447>
- Dechouniotis G, Petridis XM, Georgopoulou MK (2010). Influence of specialty training and experience on endodontic decision making. *J Endod*, **36**(7): 1130–1134. <https://doi.org/10.1016/j.joen.2010.03.023>
- De Moor R, Hülsmann M, Kirkevang LL, Tanalp J, Whitworth J (2013). Undergraduate curriculum guidelines for endodontology. *Int Endod J*, **46**(12): 1105–1114. <https://doi.org/10.1111/iej.12186>
- Dietschi D, Duc O, Krejci I, Sadan A (2008). Biomechanical considerations for the restoration of endodontically treated teeth: A systematic review of the literature, Part II (Evaluation of fatigue behavior, interfaces, and in vivo studies). *Quintessence Int*, **39**: 117–129.
- Eke PI, Genco RJ (2007). CDC Periodontal Disease Surveillance Project: Background, objectives, and progress report. *J Periodontol*, **78**(7 Suppl): 1366–1371. <https://doi.org/10.1902/jop.2007.070134>
- Fezai H, Al-Salehi S (2019). The relationship between endodontic case complexity and treatment outcomes. *J Dent*, **85**: 88–92. <https://doi.org/10.1016/j.jdent.2019.05.019>

- Fuge KN, Stuck AM, Love RM (1998). A comparison of digitally scanned radiographs with conventional film for the detection of small endodontic instruments. *Int Endod J*, **31**: 123–126.
- Gabor C, Tam E, Shen Y, Haapasalo M (2012). Prevalence of internal inflammatory root resorption. *J Endod*, **38**(1): 24–27. <https://doi.org/10.1016/j.joen.2011.10.007>
- Gavriil D, Kakka A, Myers P, O'Connor CJ (2021). Pre-endodontic restoration of structurally compromised teeth: current concepts. *Br Dent J*, **231**: 343–349. <https://doi.org/10.1038/s41415-021-3467-0>
- Gwet KL (2014). *Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement among Raters*, 4th edn. Gaithersburg: Advanced Analytics LLC.
- Hamer S, Kanagasigam S, Sonde N, Mannocci F, Patel S (2021). The impact of the Dental Practicality Index on treatment planning. *Br Dent J*. <https://doi.org/10.1038/s41415-021-2948-5>
- Haug SR, Solfeld AF, Ranheim LE, Bårdsen A (2018). Impact of case difficulty on endodontic mishaps in an undergraduate student clinic. *J Endod*, **44**(7): 1088–1095. <https://doi.org/10.1016/j.joen.2018.03.012>
- Kim A, Candamo F, Wee AG, Yuan JCC, Garcia LT (2019). Predoctoral dental education in tooth prognosis and treatment complexity: A survey of US dental schools. *J Dent Educ*, **83**(10): 1151–1157. <https://doi.org/10.21815/JDE.019.122>
- McCaul LK, McHugh S, Saunders WP (2001). The influence of specialty training and experience on decision making in endodontic diagnosis and treatment planning. *Int Endod J*, **34**(8): 594–606. <https://doi.org/10.1046/j.1365-2591.2001.00433.x>
- Messer HH (1999). Clinical judgement and decision making in endodontics. *Aust Endod J*, **25**(3): 124–132. <https://doi.org/10.1111/j.1747-4477.1999.tb00119.x>
- Nixdorf DR, Moana-Filho EJ, Law AS, McGuire LA, Hodges JS, John MT (2010). Frequency of nonodontogenic pain after endodontic therapy: A systematic review and meta-analysis. *J Endod*, **36**(9): 1494–1498. <https://doi.org/10.1016/j.joen.2010.06.020>
- Papapanou PN, Sanz M, Buduneli N, Dietrich T, Feres M, Fine DH *et al* (2018). Periodontitis: Consensus report of workgroup 2 of the 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions. *J Periodontol*, **89**(Suppl 1): S173–S182. <https://doi.org/10.1002/JPER.17-0721>
- Peters OA, Arias A (2016). Shaping, disinfection, and obturation for molars. In: Peters OA (ed.), *The Guidebook to Molar Endodontics*. Berlin: Springer, pp. 153. https://doi.org/10.1007/978-3-662-52901-0_6
- Polycarpou N, Ng YL, Canavan D, Moles DR, Gulabivala K (2005). Prevalence of persistent pain after endodontic treatment and factors affecting its occurrence in cases with complete radiographic healing. *Int Endod J*, **38**(3): 169–178. <https://doi.org/10.1111/j.1365-2591.2004.00923.x>
- Ponder SN, Benavides E, Kapila S, Hatch NE (2013). Quantification of external root resorption by low-vs high-resolution cone-beam computed tomography and periapical radiography: A volumetric and linear analysis. *Am J Orthod Dentofacial Orthop*, **143**(1): 77–91. <https://doi.org/10.1016/j.jado.2012.08.023>
- Schäfer E, Diez C, Hoppe W, Tepel J (2002). Roentgenographic investigation of frequency and degree of canal curvatures in human permanent teeth. *J Endod*, **28**(3): 211–216. <https://doi.org/10.1097/00004770-200203000-00017>

- Shah PK, Duncan HF, Abdullah D, Tomson PL, Murray G, Friend TM *et al.* (2020). Comparison of two case difficulty assessment methods on cohorts of undergraduate dental students: A multi-centre study. *Int Endod J*, **53**(11): 1569–1580. <https://doi.org/10.1111/iej.13377>
- Tifooni A, Al-Nuaimi N, Dawood A, Mannocci F, Patel S (2019). Validation of the effectiveness of the Dental Practicality Index in predicting the outcome of root canal retreatments. *Int Endod J*, **52**(10): 1403–1409. <https://doi.org/10.1111/iej.13142>
- Sorensen JA, Engelman MJ (1990). Ferrule design and fracture resistance of endodontically treated teeth. *J Prosthet Dent*, **63**(5): 529–536. [https://doi.org/10.1016/0022-3913\(90\)90070-s](https://doi.org/10.1016/0022-3913(90)90070-s)
- Su H, Liao HF, Fiorellini JP, Kim S, Korostoff J (2014). Factors affecting treatment planning decisions for compromised anterior teeth. *Int J Periodontics Restorative Dent*, **34**(3): 389–398. <https://doi.org/10.11607/prd.1581>
- Thammasitboon S, Cutrer WB (2013). Diagnostic decision-making and strategies to improve diagnosis. *Curr Probl Pediatr Adolesc Health Care*, **43**(9): 232–241. <https://doi.org/10.1016/j.cppeds.2013.07.003>
- Ungerechts C, Bårdsen A, Fristad I (2014). Instrument fracture in root canals: Where, why, when and what? A study from a student clinic. *Int Endod J*, **47**(2): 183–190. <https://doi.org/10.1111/iej.12131>
- Weine FS (2004). Intracanal treatment procedures, Basic and Advance topics. In *Endodontic Therapy*, 6th edn. St. Louis: The CV Mosby Company, pp. 188–189.