ORIGINAL ARTICLE

*International Dental Journal* 2014; 64: 187–194

doi: 10.1111/idj.12104

# The Caries Assessment Spectrum and Treatment (CAST) instrument: its reproducibility in clinical studies

Ana Luiza de Souza[1,2], Ewald M. Bronkhorst[3], Nico H. J. Creugers[4], Soraya C. Leal[2] and Jo E. Frencken[1]

*[1]Department of Global Oral Health, College of Dental Sciences, Radboud University Medical Center, Nijmegen, the Netherlands; [2]Department of Paediatric Dentistry, School of Health Sciences, University of Brasília, Brasília, Brazil; [3]Department of Preventive and Restorative Dentistry, College of Dental Sciences, Radboud University Medical Center, Nijmegen, the Netherlands; [4]Department of Oral Function and Prosthetic Dentistry, College of Dental Sciences, Radboud University Medical Center, Nijmegen, the Netherlands.*

A new caries assessment instrument, the Caries Assessment Spectrum and Treatment (CAST), was developed. It covers carious lesion progression from no lesion, sealants and restorations to lesions in enamel and dentine, advanced stages in pulpal and tooth-surrounding tissues, and tooth loss owing to dental caries, in nine codes. The objective of this study was to determine the reproducibility of the CAST instrument in primary and permanent dentitions, using three age groups. Two epidemiological surveys were conducted in Brazil, covering three age groups: 2–6-year-old and 6–9-year-old children and 19–30-year-old adults. Four trained and calibrated examiners performed the examinations. Reproducibility was calculated for intra- and inter-examiner at surface and tooth levels and expressed as unweighted kappa-coefficient value (κ) and percentage of agreement (Po) for CAST codes (0–7) and for the categories healthy (0–2) *versus* diseased (3–7), and non-cavitated (0–3) *versus* cavitated (4–7) teeth. Using CAST codes (0–7) for the 2–6-year-old age group in primary dentitions, inter-examiner consistency was $\kappa = 0.74$ and Po was 98.3%. In the 6–9-year-old age group in primary dentitions, inter-examiner consistency ranged from $\kappa = 0.68$ to $\kappa = 0.86$ and Po was ≥93.7%. In the 19–30-year-old age group inter-examiner consistency was $\kappa = 0.87$ and Po was 94.1%. The reproducibility of the CAST instrument for use in the primary dentition of 2–6-year olds and of 6–9-year olds was 'substantial' to 'almost perfect'. The reproducibility for its use in the permanent dentition of 19–30-year olds was 'almost perfect'. The CAST instrument can reliably be applied in epidemiological studies covering these ages.

***Key words:*** *Caries assessment spectrum and treatment (CAST), dental caries, DMF index, caries epidemiology, validation studies*

## INTRODUCTION

Various diagnostic instruments have been developed for detecting clinical signs of carious lesion progression, distinguishing its different stages in enamel and dentine[1] and, more recently, determining the consequences of untreated dental caries in pulp and tooth-surrounding tissues[2]. The detection of carious lesions in epidemiological surveys needs to be carried out adequately. Therefore, a detection instrument that is uncomplicated, validated and can meet the study's objectives should be used.

Following the reduction of dental caries in the Western World in the last decades of the 20th century, emphasis on caries epidemiology shifted from, in most cases, merely counting dentine cavities to assessing carious lesions in enamel and dentine and, in particular, to determining the level of caries lesion activity in enamel. The most frequently used assessment instrument for the latter purpose is that developed by Nyvad[3].

With the purpose of creating one suitable dental caries assessment instrument for use in the 21st century, a group of researchers developed ICDAS (International Caries Detection and Assessment System)[4]. This is a two-digit system in which the first digit refers to the presence and condition of sealants and restorations, and the second refers to various stages of carious lesion progression. It was developed for use in dental practice, education, research and public health (epidemiology). However, its use in epidemiological surveys has been criticised. Some authors have had to simplify ICDAS either during the examinations[5] or during analysis and presentation of results[6,7]. Furthermore, ICDAS does not differentiate between the

stages of dental caries progression confined to dentine and those reaching the pulp, and tooth-surrounding tissues [pufa(pulp, ulceration, fistula, abscess)-index][2].

The ICDAS Coordinating Committee (ICDAS CC) has accepted some of the criticisms of ICDAS and proposed a merged codes system, termed the International Caries Classification and Management System (ICCMS)[8]. This system consists of one-coding scheme and the suggestion is that caries codes be merged; from the original seven into four codes. Apparently, this adjustment is an effort to overcome the difficulties reported when ICDAS (II) has been used in epidemiological surveys. However, instructions about how to operate and report this merged-system (ICDAS/ICCMS) are currently (autumn 2013) unclear. Therefore, ICDAS CC intends to produce an updated version of the learning programme, book and software to clarify its use in the coming years[8].

The frequently used caries assessment criteria, described by the World Health Organisation (WHO), merely distinguish absence and presence of obvious dentine cavities[9]. Data regarding the presence of carious lesions in enamel and dentine, and those involving the pulp and those not involving it, cannot therefore be obtained through use of the WHO criteria.

As a consequence of the difficulty in using ICDAS and the limitations of the WHO criteria, a pragmatic, reliable and internationally accepted caries assessment instrument for use in epidemiological surveys was still needed. An attempt to develop such an instrument was made by the research group that had used ICDAS II[7] and the pufa index[10] for assessing the oral health of schoolchildren. It resulted in the development of the Caries Assessment Spectrum and Treatment (CAST) instrument[1,11], which covers the complete range of carious lesions: from 'no carious lesions', to 'caries protection' (sealants), 'restorations', 'lesions in enamel', 'dentine', 'advanced stages in pulpal and tooth-surrounding tissues' and 'teeth lost due to the caries process' (*Table 1*). Use of CAST makes calculation of a DMF (decayed, missing and filled) score possible, thus allowing comparison of CAST results and the DMF scores available in the many publications worldwide[11].

The CAST instrument has been validated for face and content by a group of 56 epidemiologists from 24 countries, using the e-Delphi consensus method[12]. Its construct validity has been obtained[13]. Complementary to the validation process, the reproducibility of the CAST instrument needs to be determined to ensure its reliability for use in epidemiological surveys. The objective of this paper was to present the level of reproducibility of applying the CAST instrument to primary and permanent dentitions of three different age groups.

## METHODS

The reproducibility of the CAST instrument was assessed among Brazilians of three age groups: children aged 6–9 years (Study 1), infants aged 2–6 years and their 19–30 year-old mothers (Study 2). The Ethics Committee of the University of Brasília approved the study protocols (CEP-FM 014/2011 and 047/2011). Being also a controlled clinical study, study 1 was registered at the Dutch Trial Registration Centre, as number 1699. Parents and legal guardians (Study 1) and mothers (Study 2) were given a consent form explaining the nature of the study. Only those who returned the duly signed form and agreed to be examined were included in the study. The research was been conducted in full accordance with the World Medical Association Declaration of Helsinki.

### Sample

#### *Study 1*

In 2009, a 4-year mixed-longitudinal study covering 6–7-year-old children in Paranoá, a district of Brasília,

**Table 1** The Caries Assessment Spectrum and Treatment (CAST) instrument codes and descriptions validated for face and content[12]

| Characteristic | Code | Description |
|---|---|---|
| Sound | 0 | No visible evidence of a distinct carious lesion is present |
| Sealant | 1 | Pits and/or fissures are at least partially covered with a sealant material |
| Restoration | 2 | A cavity is restored with an (in)direct restorative material |
| Enamel | 3 | Distinct visual change in enamel only. A clear caries-related discolouration is visible, with or without localised enamel breakdown |
| Dentine | 4 | Internal caries-related discolouration in dentine. The discoloured dentine is visible through enamel, which may or may not exhibit a visible localised breakdown of enamel |
| | 5 | Distinct cavitation into dentine. The pulp chamber is intact |
| Pulp | 6 | Involvement of the pulp chamber. Distinct cavitation reaching the pulp chamber or only root fragments are present |
| Abscess/fistula | 7 | A pus-containing swelling or a pus-releasing sinus tract related to a tooth with pulpal involvement |
| Lost | 8 | The tooth has been removed because of dental caries |
| Other | 9 | Does not correspond to any of the other descriptions |

Brazil was introduced[7]. Two years later these children were re-examined on their school premises, together with a new birth cohort of 6–7-year-old children, by three trained and calibrated examiners using the CAST instrument, portable equipment and artificial light.

### *Study 2*

A primary oral health care programme, aimed at reducing early childhood caries in infants from deprived areas and at improving the oral health of their mothers, was delivered at the dental clinic of the School of Dentistry of the University of Brasília. Two trained and calibrated examiners using the CAST instrument evaluated the programme between October and December 2011.

### Examiner training

The training sessions were conducted under the supervision of a senior epidemiologist (JEF). These comprised a theoretical explanation about the CAST instrument (1.5 hours) and a practical session (2 hours) in which a total of 20 extracted teeth were examined and scored according to CAST by each of the examiners. Individual scores were compared and, where there was a difference, examiners discussed the scores until consensus was reached. This process was repeated until good agreement among examiners was reached. After being trained, examiners of Study 1 ($n = 3$) were calibrated for 8 hours under the supervision of JEF. They examined 14 children of the same age as those included in the main study and from a similar socioeconomic background. The two examiners of Study 2 (one examiner also participated in Study 1) were calibrated in two afternoons spent in examining 10 infants and their mothers. The kappa-coefficient values for the inter-examiner agreement at the end of the calibration session for studies 1 and 2 were 0.75 and 0.81, respectively.

### Clinical examination

The oral examination started with assessment of the presence of toothache, plaque[14] and gingival bleeding[15]. The examiners then brushed the patients' teeth, without toothpaste, to facilitate clear visibility of tooth surfaces. Dental floss and gauzes were used to remove any remaining plaque. Thereafter, the CAST instrument was used in assessing dental caries status. A mirror handle with a battery-powered built-in light source illuminated the tooth to be examined (MirrorLite®; Kudos, Hong Kong). No air-drying was applied. If necessary, excess saliva was removed, using a cotton wool roll or gauzes. A trained recorder assisted each examiner. In both studies the senior epidemiologist was present during the first week of examinations to assist the examiners in case of doubts, to answer questions and discuss cases.

### Reproducibility test

Reproducibility of CAST codes was obtained through an intra- and inter-examiner consistency test and expressed as percentage of agreement (Po) and unweighted kappa coefficient ($\kappa$) with standard error. The percentage of agreement summarises the total number of units in which there was agreement between the first and second observations of one observer or between two observers[9,16]. The kappa coefficient expresses the agreement between two observations corrected for chance[17]. The commonly used categorisation of Landis and Koch[18] was used for classifying the kappa-coefficient values.

### Statistical analysis

Data were analysed by an oral statistician using the statistical package IBM SPSS (version 20.0 for Windows, IBM, Chicago, IL, USA). Results are presented by study, examiners and type of dentition (primary or permanent) at surface and tooth level. As the CAST instrument is ordered hierarchically, and as a tooth lost because of dental caries is not considered a carious lesion anymore, the prevalence of dental caries in the present studies included only codes 0–7[11]. The hierarchical order permits the calculation of a maximum CAST score per mouth. Agreement was calculated for CAST codes 0–7 and by determining two sets of categories. The cut-off points for calculating agreement between examiners were determined as 'healthy' *versus* 'diseased' (codes 0–2 *vs.* 3–7) and 'non-cavitated' *versus* 'cavitated' (codes 0–3 *vs.* 4–7) teeth.

## RESULTS

### *Study 1*

The sample consisted of 1,617 children, 957 aged 6–7 years old and 660 aged 8–9 years old. A total of 171 children (10.6%) were examined twice or three times, resulting in 349 duplicate examinations. *Table 2* shows the frequency distribution of CAST codes (0–8), maximum CAST code per mouth and cumulative percentages of CAST codes by dentition. For the primary dentition, prevalence of enamel and dentine lesions was 52.4% (codes 3–7) and 43.5% (codes 4–7) of dentine lesions only. The most prevalent caries code was code 5 (21.1%). For the permanent dentition, the prevalence of enamel and dentine lesions (codes 3–7) was 14.1% and for dentine lesions the prevalence was 8.0% (only code 5).

**Table 2** Number of tooth surfaces, maximum Caries Assessment Spectrum and Treatment (CAST) code per mouth and cumulative percentages of CAST codes by dentition amongst 6–9 year-olds

| CAST code | *n* | Max | Cumulative% |
|---|---|---|---|
| Primary dentition | | | |
| 0 | 10,488 | 58 | 34.1 |
| 1 | 2 | 1 | 34.7 |
| 2 | 147 | 5 | 37.6 |
| 3 | 192 | 15 | 46.5 |
| 4 | 15 | 1 | 47.1 |
| 5 | 435 | 36 | 68.2 |
| 6 | 363 | 32 | 87.1 |
| 7 | 44 | 5 | 90 |
| 8 | 105 | 17 | 100 |
| Total | 11,791 | 170 | 100 |
| Permanent dentition | | | |
| 0 | 6,062 | 134 | 82.2 |
| 1 | 1 | 1 | 82.8 |
| 2 | 8 | 5 | 85.9 |
| 3 | 51 | 10 | 92 |
| 4 | 4 | 0 | 92 |
| 5 | 15 | 13 | 100 |
| Total | 6,141 | 163 | 100 |

*n*, Number of surfaces; Max, maximum caries score per mouth.

Results regarding the reproducibility test of the CAST instrument (codes 0–7) at surface level for both dentitions are presented in *Table 3*. The kappa value for the inter-examiner consistency tests in primary and permanent dentitions ranged from 0.68 to 0.86 and from 0.28 to 0.67, respectively. The corresponding Po values ranged from 93.7% to 97.0% and from 97.7% to 98.9%. Examiner 2 had lower intra-examiner kappa and Po values than examiners one and three.

Results of the reproducibility test at tooth level for the categories 'healthy' *versus* 'diseased' teeth and for 'non-cavitated' *versus* 'cavitated' teeth for both dentitions are presented in *Table 4*. Inter-examiner kappa-value in primary dentitions for 'healthy' *versus* 'diseased' teeth, and for 'non-cavitated' *versus* 'cavitated' teeth ranged from 0.54 to 0.86, and from 0.78 to 0.88, respectively. The corresponding Po values ranged from 88.5% to 96.6% and from 96.8% to 97.9%, respectively. For the permanent dentition, kappa values for 'healthy' *versus* 'diseased' teeth ranged from 0.24 to 0.58 with corresponding Po values ranging from 95.9% to 97.2%. For 'non-cavitated' *versus* 'cavitated' teeth kappa values could not be calculated for two of the three inter-examiner consistency tests, because of the absence of variation within scores.

### *Study 2*

The sample consisted of 177 children (2–6 years old) and their mothers (19–30 years old). Re-examination of 24 infants (13.5%) and 24 mothers (13.5%) produced 96 duplicates. *Table 5* shows the number of surfaces scored by CAST code (0–8), maximum CAST code per mouth and cumulative percentages of CAST codes by dentition. The prevalence of enamel and dentine lesions for the primary dentition was 14.6% (codes 4–7). The most prevalent caries code for the primary dentition was code 5 (8.3%). For the permanent dentition, the prevalence of enamel and dentine lesions was 62.5%. The most prevalent caries code was code 6 (33.3%).

Results regarding the reproducibility test of the CAST instrument (codes 0–7) at tooth level for both dentitions are presented in *Table 6*. The kappa and Po values for the inter-examiner consistency tests in assessing CAST codes in primary dentitions were 0.74 and 98.3%, respectively. The kappa and Po values for inter-examiner consistency tests were 0.87 and 94.1%, respectively, for assessing the CAST codes in permanent dentition.

*Table 7* presents the results of the reproducibility test at tooth level for the categories 'healthy' *versus* 'diseased' teeth and for 'non-cavitated' *versus* 'cavitated' teeth for both dentitions. The kappa and Po values for inter-examiner consistency test in primary dentition for the category 'healthy' *versus* 'diseased teeth' were 0.65 and 96.7%, and for the category

**Table 3** Intra- and inter-examiner consistency test of assessing primary and permanent dentition at surface level amongst 6–9-year-olds with the Caries Assessment Spectrum and Treatment (CAST) instrument (0–7)

| | Primary dentition | | | | Permanent dentition | | | |
|---|---|---|---|---|---|---|---|---|
| | *n* | κ | SE | Po (%) | *n* | κ | SE | Po (%) |
| Intra-examiner | | | | | | | | |
| Examiner 1 | 2,227 | 0.86 | 0.02 | 97.0 | 949 | 0.67 | 0.10 | 99.1 |
| Examiner 2 | 1,220 | 0.80 | 0.02 | 93.9 | 523 | 0.33 | 0.13 | 97.1 |
| Examiner 3 | 1,793 | 0.82 | 0.02 | 96.8 | 1082 | 0.56 | 0.10 | 98.8 |
| Inter-examiner | | | | | | | | |
| Examiner 1–2 | 1,220 | 0.86 | 0.03 | 97.0 | 441 | 0.67 | 0.11 | 97.7 |
| Examiner 2–3 | 994 | 0.68 | 0.03 | 93.7 | 608 | 0.28 | 0.13 | 98.4 |
| Examiner 1–3 | 4,933 | 0.86 | 0.02 | 96.3 | 1812 | 0.67 | 0.11 | 98.9 |

*n*, Number of surfaces; κ, kappa-coefficient value; SE, standard error; Po, percentage of agreement.

**Table 4** Intra- and inter-examiner consistency test for primary and permanent dentition for categories healthy (0–2) *versus* diseased (3–7) and for non-cavitated (0–3) *versus* cavitated (4–7) teeth among 6–9-year-olds

| | Primary dentition | | | | | | Permanent dentition | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | 0–2 *vs.* 3–7 | | | 0–3 *vs.* 4–7 | | | *n* | 0–2 *vs.* 3–7 | | | 0–3 *vs.* 4–7 | | |
| | | κ | SE | Po (%) | κ | SE | Po (%) | | κ | SE | Po (%) | κ | SE | Po (%) |
| Intra-examiner | | | | | | | | | | | | | | |
| Examiner 1 | 493 | 0.84 | 0.03 | 95.3 | 0.87 | 0.03 | 97.1 | 218 | 0.65 | 0.13 | 97.2 | 0.32 | 0.25 | 98.1 |
| Examiner 2 | 272 | 0.82 | 0.04 | 93.0 | 0.93 | 0.03 | 98.1 | 121 | 0.40 | 0.14 | 90.0 | 1.00 | 0.00 | 100 |
| Examiner 3 | 395 | 0.89 | 0.03 | 97.2 | 0.84 | 0.04 | 96.9 | 250 | 0.61 | 0.13 | 97.2 | 0.60 | 0.16 | 98.0 |
| Inter-examiner | | | | | | | | | | | | | | |
| Examiner 1–2 | 271 | 0.86 | 0.05 | 96.6 | 0.86 | 0.05 | 97.7 | 105 | 0.48 | 0.22 | 96.1 | *0.00 | 0.00 | 99.0 |
| Examiner 2–3 | 219 | 0.59 | 0.07 | 88.5 | 0.78 | 0.08 | 96.8 | 148 | 0.58 | 0.19 | 97.2 | *0.00 | 0.00 | 97.9 |
| Examiner 1–3 | 1099 | 0.79 | 0.03 | 94.9 | 0.88 | 0.02 | 97.9 | 418 | 0.24 | 0.12 | 95.9 | 0.49 | 0.18 | 98.5 |

*n*, Number of surfaces; κ, kappa-coefficient value; SE, standard error; Po, percentage of agreement.
*No statistics were computed because of the absence of variation. In the first case the examiners disagreed about one tooth and, in the second, they disagreed about three teeth.

**Table 5** Number of tooth surfaces, maximum Caries Assessment Spectrum and Treatment (CAST) code per mouth and cumulative percentages of CAST codes by dentition amongst 2–6-year-olds (primary dentition) and 19–30-year-olds (permanent dentition)

| CAST code | *n* | Max | Cumulative% |
|---|---|---|---|
| Primary dentition | | | |
| 0 | 1,908 | 38 | 79.2 |
| 2 | 12 | 3 | 85.4 |
| 3 | 31 | 0 | 85.4 |
| 4 | 1 | 1 | 87.5 |
| 5 | 22 | 4 | 95.8 |
| 6 | 23 | 0 | 95.8 |
| 7 | 8 | 2 | 100.0 |
| Total | 2,005 | 48 | |
| Permanent dentition | | | |
| 0 | 2,167 | 2 | 4.2 |
| 1 | 1 | 0 | 4.2 |
| 2 | 442 | 16 | 37.5 |
| 3 | 74 | 0 | 37.5 |
| 4 | 16 | 1 | 39.6 |
| 5 | 46 | 11 | 62.5 |
| 6 | 39 | 16 | 95.8 |
| 7 | 5 | 2 | 100.0 |
| Total | 2,790 | 48 | |

*n*, Number of surfaces; Max, maximum caries score per mouth.

'non-cavitated' *versus* 'cavitated' teeth they were 1.00 and 100%. For the permanent dentition, the kappa and Po values for the category 'healthy' *versus* 'diseased' teeth were 0.82 and 95.3%, and those for the category 'non-cavitated' *versus* 'cavitated' teeth were 0.88 and 98.6%.

## DISCUSSION

### Methodology

Kappa statistics are commonly used to determine the reproducibility of a diagnostic test. They function well if the codes that comprise the test are frequently scored. This usually allows for sufficient variation in scores per code. However, if the frequency of scoring a code is low, one variation from the diagonal in a $2 \times 2$ table can lower the kappa-coefficient value tremendously. The same is true when the prevalence of a condition among the population is high[19]. In low-prevalence populations the unit of analysis (surface, tooth, mouth) also influences the kappa-coefficient value. At surface level, the agreement between examiners is expected to be higher than at tooth level. This is because a tooth without the study condition counts for four or five correctly determined scores, increasing the prevalence of agreed scores considerably[16]. As a result of the effect of the prevalence on the kappa-coefficient value, it is scarcely possible to compare kappa coefficients between different population groups or characteristics[20].

In contrast to the decision made by the ICDAS CC, who instructed users of the ICDAS to apply the weighted kappa coefficient, in the present investigation the unweighted, simple kappa was applied. The CAST codes, like many other caries lesion codes included in other assessment instruments, are of a categorical and not of an ordinal nature, which is a requisite for application of a weighted kappa statistic[21,22].

Although frequently done, merely reporting kappa coefficient values is uninformative and can be misleading[22]. Therefore, it was recommended that information about systematic disagreement among examiners, marginal homogeneity of the sample, bivariate symmetry in misclassification and the underlying prevalence of the condition in the population should be reported together with the kappa-coefficient value[22]. Because of the absence of a gold standard (benchmark examiner) in a sample size such as ours (total of 219 subjects), it was not possible to adhere to some of these recommendations in the present study. As there are at present no standards for reporting the existence

**Table 6** Intra- and inter-examiner consistency test of assessing primary (2–6-year-olds) and permanent (19–30-year-olds) dentition at tooth surface level with the Caries Assessment Spectrum and Treatment (CAST) instrument

| | Primary dentition | | | | Permanent dentition | | | |
|---|---|---|---|---|---|---|---|---|
| | *n* | κ | SE | Po (%) | *n* | κ | SE | Po (%) |
| Intra-examiner | | | | | | | | |
| Examiner 1 | 998 | 0.61 | 0.11 | 87.0 | 1,275 | 0.89 | 0.01 | 95.5 |
| Examiner 2 | 616 | 0.87 | 0.03 | 99.3 | 1,011 | 0.91 | 0.01 | 96.4 |
| Inter-examiner | | | | | | | | |
| Examiner 1–2 | 400 | 0.74 | 0.10 | 98.3 | 750 | 0.87 | 0.02 | 94.1 |

*n*, Number of surfaces; κ, kappa-coefficient value; SE, standard error; Po, percentage of agreement.

**Table 7** Intra- and inter-examiner consistency test of assessing primary (2–6-year-olds) and permanent (19–30-year-olds) dentition for the categories healthy (0–2) *versus* diseased (3–7) and for non-cavitated (0–3) *versus* cavitated (4–7) teeth

| | Primary dentition | | | | | | Permanent dentition | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | 0–2 *vs.* 3–7 | | | 0–3 *vs.* 4–7 | | | *n* | 0–2 *vs.* 3–7 | | | 0–3 *vs.* 4–7 | | |
| | | κ | SE | Po (%) | κ | SE | Po (%) | | κ | SE | Po (%) | κ | SE | Po (%) |
| Intra-examiner | | | | | | | | | | | | | | |
| Examiner 1 | 228 | 0.60 | 0.16 | 97.8 | 1.00 | 0.00 | 100 | 260 | 0.76 | 0.05 | 93.0 | 0.82 | 0.06 | 96.9 |
| Examiner 2 | 138 | 0.88 | 0.05 | 96.3 | 0.91 | 0.05 | 97.8 | 208 | 0.86 | 0.05 | 97.1 | 0.88 | 0.07 | 98.5 |
| Inter-examiner | | | | | | | | | | | | | | |
| Examiner 1–2 | 92 | 0.65 | 0.19 | 96.7 | 1.00 | 0.00 | 100 | 151 | 0.82 | 0.06 | 95.3 | 0.88 | 0.08 | 98.6 |

*n*, Number of surfaces; κ, kappa coefficient value; SE, standard error; Po, percentage of agreement.

of agreement coefficients[22], the decision was made to present not only the prevalence of the condition, kappa coefficient and standard error, but also the percentage of agreement per code and per code-combination. The last measurement is easy to calculate and has a strong intuitive appeal[23]. However, it is also influenced by the unit of analysis (tooth, surface level) and by the prevalence of the condition. Thus, it is important to report the percentage of agreement together with other agreement measurements when the prevalence of non-diseased cases is high[16]. Therefore, the prevalence, kappa coefficient and percentage of agreement were combined as complementary measurements to overcome the paradoxes faced by these measurements when used alone.

The kappa-coefficient value for the inter-examiner agreement in identifying cavitated *versus* non-cavitated permanent teeth could not be calculated in two out of three occasions. These results were caused by the absence of sufficient variation in scores and because the prevalence of dentine carious lesions in permanent teeth in the age group of 6–9 years old was very low (only 8%). Even when enamel carious lesions were combined with dentine carious lesions, increasing the prevalence to 14.1%, the kappa-coefficient value remained low. Nevertheless, the percentage of agreement among examiners was high, as examiners disagreed in only a few cases. These paradoxes – low kappa-coefficient values and high percentage of agreement, and kappa-coefficient value variation or no computation – are consequences of the very low prevalence of CAST codes that are different from code 0 in the population (*Table 2*). This situation may have occurred because the study children were only beginning their mixed dentition, had relatively few recently erupted permanent teeth, in most cases unsealed and unrestored, and only a few with a carious lesion. Therefore, the outcomes of the reproducibility test for the permanent dentition in this age group should be considered with extreme care.

## Findings

The reproducibility test for the use of CAST in primary teeth was carried out on a group of young and somewhat older children originating from a low and higher prevalence of CAST codes, respectively. The intra- and inter-examiner agreement for identifying all CAST codes – those for the categories 'healthy' *versus* 'diseased' and 'non-cavitated' *versus* 'cavitated' tooth surfaces – in both groups were 'substantial' to 'almost perfect'. The level of intra-examiner agreement was higher in the older group of children and was generally higher than the inter-examiner agreement. The former might result from the higher prevalence of the individual CAST codes present in that group than in the younger group of children. The latter indicates that the calibration exercise could have been more

intense. As all percentage of agreement values for the three conditions under analysis were high, the indication is that the reproducibility of CAST in these two age groups reached a substantial to high level.

In adults, the examiner agreement regarding the identification of CAST codes 0–7, those for the categories 'healthy' *versus* 'diseased' and 'non-cavitated' *versus* 'cavitated' surfaces – were almost perfect, but for the intra-examiner agreement of one of the examiners that showed a substantial level of agreement. This finding and the high values for the corresponding percentage of agreement indicate a near to almost perfect reproducibility of CAST codes in this adult group of 19–30 years old.

Measuring the reproducibility of a newly developed diagnostic instrument with more than one measurement *in vivo* is part of the evaluation process that eventually determines the quality of the instrument and, consequently, the quality of the data collected. As this is the first time that the reproducibility of the CAST instrument has been measured, it is not possible to relate the findings of the present study to those of others. Therefore, there is a need to measure the reproducibility of CAST in the same age groups as reported here, and in others, with varying prevalence of CAST codes and by other researchers.

If the manner in which the level of reproducibility of other recently introduced caries assessment instruments is analysed, it emerges that the kappa-coefficient was the only measurement used to express reproducibility among examiners in the majority of epidemiological studies that had used the ICDAS for the first time. These studies used the ICDAS as if the level of reproducibility in different age groups had already been assessed and been considered to be sufficiently high. Some studies reported weighted kappa-coefficient values[6,24] or unweighted kappa-coefficient values[6] and others did not specify whether the kappa-coefficient used had been weighted or not[5,7]. Only one study reported, in addition to the kappa-coefficient values, the percentage of agreement for two sets of categories of ICDAS codes when reporting on its reproducibility[7].

The level of reproducibility, using the kappa statistics, of the other recently developed caries assessment instrument, PUFA/pufa, was tested in three studies[10,25,26]. Hence, limited data regarding the *in vivo* reproducibility of ICDAS and PUFA/pufa indices are available.

In similarity to the present study, the reproducibility of the Nyvad criteria was measured *in vivo* among the 9–14-year-old[3] and 3–7-year-old children[27] using the kappa-coefficient and corresponding percentage of agreement for different categories of instrument codes.

We conclude that the reproducibility of the CAST instrument for use in the primary dentition of 2–6-year-olds and of 6–9-year-olds was 'substantial' to 'almost perfect'. Its reproducibility for use in the permanent dentition of 19–30-year-olds was almost perfect. Studies by other research groups on various age groups are needed to further test the reproducibility of CAST. The CAST instrument can be applied in epidemiological studies for the age groups studied.

## Acknowledgements

We are grateful to the Department of Education (Brasília, DF), school principals, teachers and children from the Public Schools of Paranoá. Our appreciation goes to the mothers and the children from the Oral Health Care Programme of the Paediatric Dentistry Department of the University of Brasília. We thank the examiners Maria José Figueiredo, Juliana Bittar and Patrícia Medeiros, and the dental assistants who kindly assisted us. Our gratitude goes to Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Radboud University Medical Center for supporting this study financially.

## Conflicts of Interests

None declared.

## REFERENCES

1. Frencken JE, de Amorim RG, Faber J *et al.* The caries assessment spectrum and treatment (CAST) index: rational and development. *Int Dent J* 2011 61: 117–123.
2. Monse B, Heinrich-Weltzien R, Benzian H *et al.* PUFA An index of clinical consequences of untreated dental caries. *Community Dent Oral Epidemiol* 2010 38: 77–82.
3. Nyvad B, Machiulskiene V, Baelum V. Reliability of a new caries diagnostic system differentiating between active and inactive caries lesions. *Caries Res* 1999 33: 252–260.
4. Pitts N. "ICDAS"– an international system for caries detection and assessment being developed to facilitate caries epidemiology, research and appropriate clinical management. *Community Dent Health* 2004 21: 193–198.
5. Cadavid AS, Lince CMA, Jaramillo MC. Dental caries in the primary dentition of a Colombian population according to the ICDAS criteria. *Braz Oral Res* 2010 24: 211–216.
6. Agustsdottir H, Gudmundsdottir H, Eggertsson H *et al.* Caries prevalence of permanent teeth: a national survey of children in Iceland using ICDAS. *Community Dent Oral Epidemiol* 2010 38: 299–309.
7. de Amorim RG, Figueiredo MJ, Leal SC *et al.* Caries experience in a child population in a deprived area of Brazil, using ICDAS II. *Clin Oral Invest* 2012 16: 513–520.
8. Pitts NB, Ekstrand KR. The ICDAS Foundation: international caries detection and assessment system (ICDAS) and its international caries classification and management system (ICCMS) – methods for staging of the caries process and enabling dentists to manage caries. *Community Dent Oral Epidemiol* 2013 41: e41–e52.
9. World Health Organization. *Oral health survey. Basic methods*, 4th ed. Geneva: World Health Organization; 1997.
10. Figueiredo MJ, de Amorim RG, Leal SC *et al.* Prevalence and severity of clinical consequences of untreated dentine carious

lesions in children from a deprived area of Brazil. *Caries Res* 2011 45: 435–442.

11. Frencken JE, de Souza AL, van der Sanden WJM *et al.* The caries assessment and treatment (CAST) instrument. *Community Dent Oral Epidemiol* 2013 41: e71–e77.
12. de Souza AL, van der Sanden WJM, Leal SC *et al.* The caries assessment spectrum and treatment (CAST) index: face and content validation. *Int Dent J* 2012 62: 270–276.
13. de Souza AL, Leal SC, Bronkhorst EM *et al.* The caries assessment spectrum and treatment (CAST) instrument: construct validation. *Eur J Oral* 2014 (in press).
14. Alaluusua S, Malmivirta R. Early plaque accumulation – a sign for caries risk in young children. *Community Dent Oral Epidemiol* 1994 22: 273–276.
15. Ainamo J, Bay I. Problems and proposals for recording gingivitis and plaque. *Int Dent J* 1975 25: 229–235.
16. Agbaje JO, Mutsvari T, Lesaffre E *et al.* Measurement, analysis and interpretation of examiner reliability in caries experience surveys: some methodological thoughts. *Clin Oral Invest* 2012 16: 117–127.
17. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960 20: 37–46.
18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977 33: 159–174.
19. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol* 1990 43: 543–549.
20. Thompson WD, Walter SD. A reappraisal of the kappa coefficient. *J Clin Epidemiol* 1988 41: 949–958.
21. Tooth LR, Ottenbacher KJ. The κ statistic in rehabilitation research: an examination. *Arch Phys Med Rehabil* 2004 85: 1371–1376.
22. Altarakemah Y, Al-Sane M, Lim S *et al.* A new approach to reliability assessment of dental caries examinations. *Community Dent Oral Epidemiol* 2013 41: 309–316.
23. Stemler SE. A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Pract Assess Res Eval* 2004 9: 4.
24. Jablonski-Momeni A, Winter J, Petrakakis P *et al.* Caries prevalence (ICDAS) in 12-year-olds from low caries prevalence areas and association with independent variables. *Int J Paediatr Dent* 2014 24: 90–97.
25. Gradella C, Bernabé E, Bönecker M. Caries prevalence and severity, and quality of life in Brazilian 2- to 4-year-old children. *Community Dent Oral Epidemiol* 2011 39: 498–504.
26. Monse B, Benzian H, Araojo J *et al.* A silent public health crisis untreated caries and dental infections among 6- and 12-year-old children in the Philippine national oral health survey 2006. *Asia Pac J Public Health* 2012 [Epub ahead of print]
27. Séllos MC, Soviero VM. Reliability of the Nyvad criteria for caries assessment in primary teeth. *Eur J Oral Sci* 2011 119: 225–231.

*Correspondence to:*
Jo E. Frencken,
Department of Global Oral Health,
College of Dental Sciences,
Radboud University Medical Center,
PO Box 9101, NL–6500 HB, Nijmegen,
the Netherlands.
Email: Jo.Frencken@radboudumc.nl