

Onderwijs

DE TOEPASBAARHEID VAN UNIVERSELE CRITERIA VOOR COMPOSITPREPARATIES IN FRONTELEMENTEN

SAMENVATTING

In onderstaand artikel wordt verslag gedaan van een onderzoek naar de toepasbaarheid van universele criteria voor composietpreparaties in frontelementen. Hiertoe is de inter- en intra-beoordelingsbetrouwbaarheid bepaald bij vier stafleden en tien studenten.

Er waren geen significante verschillen tussen de twee groepen. De gemiddelde inter- en intra-beoordelaarsbetrouwbaarheid voor de 11 items afzonderlijk was 88%. De gemiddelde intra-beoordelaarsbetrouwbaarheid voor alle items was eveneens 88%. De gemiddelde inter-beoordelaarsbetrouwbaarheid berekend over alle items bedroeg 85%.

De itemlijst bleek toepasbaar zonder training of hulpapparatuur.

WIEGMAN JE, PURDEL LEWIS DJ, OEOLOFF-KOOY A, VAN DE POEL ACM. De toepasbaarheid van universele criteria voor composietpreparaties in frontelementen. Tijdschr Tandheelkd 1987; 94: 250-3.

J. E. Wiegman, tandarts
D. J. Purdell Lewis, tandarts
A. Oeloff-Kooy, tandarts
A. C. M. van de Poel, tandarts

Uit de vakgroep Parodontologie, Prothetodontie en Sosiodontie van de rijksuniversiteit te Groningen.

Trefwoorden: Conserverende tandheelkunde – Composietpreparatie – Prekliniek

Datum acceptatie: 6 april 1987.

Adres: J. E. Wiegman, Ant. Deusinglaan 1, 9713 AV Groningen.

1. INLEIDING

De patiënt vertrouwt erop dat tandheelkundige behandelingen volgens hoge kwaliteitsnormen worden uitgevoerd. De student moet daarom al tijdens zijn opleiding leren, dat deelaspecten van een behandeling zoals bijvoorbeeld de preparatie en restauratie aan bepaalde eisen moeten voldoen.^{1, 2} Het blijkt helaas niet eenvoudig deze eisen zodanig te formuleren, dat ze eenvoudig en hanteerbaar blijven en ook nog door een ieder eenduidig en goed reproduceerbaar kunnen worden gehanteerd. Bovendien zouden deze eisen of wel criteria zó moeten zijn geformuleerd, dat alle typen preparaties voor een bepaald restauratiemateriaal ermee kunnen worden geëvalueerd.³⁻⁷ Dergelijke criteriumlijsten zullen moeten worden bijgesteld wanneer dit restauratiemateriaal door nieuwe en verbeterde eigenschappen een ontwikkeling doormaakt. De student van nu, de tandarts van morgen, zal voortdurend worden geconfronteerd met nieuwe materialen en verbeteringen van bestaande producten. Dit vraagt soms een nieuwe wijze van verwerken met als gevolg veranderde preparatievormen en aangepaste criteria. De student/tandarts zal niet alleen met criteria moeten kunnen omgaan maar ook in staat moeten zijn, aan de hand van de eigenschappen van de materialen zelf, nieuwe criteria te formuleren. Dit wordt nog complexer wanneer ook de resultaten uit bijvoorbeeld onderzoek ten aanzien van cariës en belasting op gerestaureerde gebitsdelen in de te formuleren criteria moeten worden meegenomen.⁸

De geschetste situatie geldt reeds nu ten aanzien van de composietmaterialen, die een snelle ontwikkeling doormaken.⁹

Bij de vakgroep Parodontologie, Prothetodontie en Sosiodontie bestond dan ook behoefte aan universeel toepasbare criteria voor composietpreparaties in frontelementen, gebaseerd op de huidige kennis van zaken.

Een dergelijke lijst met criteria zou moeten worden bijgesteld en aangepast als nieuwe ontwikkelingen in het composietmateriaal c.q. nieuwe onderzoekresultaten dit mogelijk en noodzakelijk maken.

In dit artikel zal verslag worden gedaan van de ontwikkeling en toetsing van deze criteriumlijst. Tevens wordt de mate van toepasbaarheid en betrouwbaarheid beschreven waarmee deze criteria, zonder speciale training, door staf en studenten kunnen worden gehanteerd.

2. MATERIAAL EN METHODE

2.1. De itemlijst met criteria

De criteria zijn als volgt tot stand gekomen. Aan vijf docenten, die al jarenlang betrokken waren bij het onderwijs in het vervaardigen van met plastisch vulmateriaal gemaakte tandkleurige restauraties, werd gevraagd 20 preparaties te beoordelen en daarvan de goede en foute aspecten op te noemen. Per docent werd van de genoemde aspecten een lijst samengesteld. Vervolgens werd een ieder verzocht bij ieder item op deze lijst de criteria te vermelden waaraan dit item moet voldoen om als goed te worden beoordeeld. De aldus verkregen vijf itemlijsten werden daarna op basis van overeenkomst en volledigheid tot één lijst samengevoegd. De lijst werd aan de vijf docenten afzonderlijk voorgelegd met de vraag of ze deze zo acceptabel

vonden en, voor zover dat niet het geval was, het verzoek aanvullingen en/of veranderingen aan te geven. Deze procedure leverde voor de composietpreparaties een lijst op van 11 items met criteria (tabel I).

2.2. De proefgroepen

De itemlijst werd door twee proefgroepen bij het beoordelen van preparaties van composietmateriaal gebruikt. De ene proefgroep bestond uit tien willekeurige studenten uit een jaargroep (N=60) die twee jaar lang patiënten hadden behandeld bij wie zij onder andere composietrestauraties hadden gemaakt. De tweede groep omvatte vier willekeurige tandartsen uit een groep van 19 docenten die waren betrokken bij de begeleiding van de 'Integrale patiëntenbehandeling'. Zij waren hierin gemiddeld al zes jaar gedurende ± 12 uur per week werkzaam.¹⁰ De deelnemers uit beide groepen waren op geen enkele wijze betrokken bij het ontwikkelen van de itemlijst.

2.3. De preparaties

Voor het onderzoek zijn 15 preparaties gemaakt in geëxtraheerde elementen met carieuze laesies. De elementen werden zodanig gekozen, dat een composietrestauratie een goede behandeling van het defect was. Deze elementen met hun preparaties werden in de normale anatomische positie tussen twee gave buurelementen in een kaak van een fantoomkop geplaatst. Door de plaats en grootte van de carieuze laesies varieerden de preparaties voor wat betreft plaats, vorm, uitgebreidheid en diepte. Drie preparaties lagen op het buccale vlak en de overige 12 preparaties waren in het

Tabel I. De itemlijst met criteria voor composietpreparaties.

Item	Criterium
1. Perifeer cariës	Binnen de glazuur-dentine- en/of cement-dentinegrens is een minimaal 1 mm brede zone, die: - zo hard is dat een scherpe sonde niet 'pikt' na druk; - niet bruin of geel verkleurd is ten gevolge van het cariësproces.
2. Centraal cariës	Al het carieuze dentine wordt in principe verwijderd tot het dentine hard is, tenzij dan een pulpapronatie verwacht wordt. In dit geval mag enigszins zacht dentine aanwezig zijn.
3. Glazuur in cervicaal deel van de preparatie	Het glazuur wordt ondersteund door hard gezond dentine (als de cervicale rand in het cement ligt, dan 'goed' scoren).
4. Glazuur in het palatinale deel van de preparatie	Het palatinale glazuur hoeft niet door dentine ondersteund te zijn, tenzij dit gebied tijdens de articulatie belast wordt. (Geldt voor proximale caviteiten, anders 'goed' scoren.)
5. Glazuur in het incisale deel van de preparatie	Het incisale glazuur hoeft niet door dentine ondersteund te zijn, tenzij: - er zichtbare fractuurlijntjes in het glazuur te zien zijn; - de dikte van het glazuur minder dan 2 mm is en er slechte gewoonten bestaan die abnormale belasting geven, bijvoorbeeld abrasie, pijproken enz.
6. Plaats van de preparatierand	De rand bevindt zich op een zodanige plaats dat een matrixband aangebracht kan worden.
7. Kwaliteit glazuur op de preparatierand	Er zijn geen tekenen van initiële demineralisatie aanwezig.
8. Afwerking preparatierand	Met een sonde afgetast, is de rand glad en er zijn geen scherpe 'hakkels' aanwezig.
9. Plaats van de preparatie-opening	De glazuur-dentine- en cement-dentinegrenzen zijn met direct of indirect zicht te bekijken en met een sonde af te tasten.
10. Grootte van de preparatie-opening	Een onderlaag kan op het dentine worden aangebracht en de composiet kan worden ingebracht.
11. Het te etsen glazuur	Binnen de preparatie-outline is een 1 mm brede glazuurwand die een hoek van tenminste 95° maakt met het glazuur-buitenoppervlak. Als de cervicale wand zich in het cement bevindt, moet deze hoek 90° of minder zijn.

proximale vlak gesitueerd, al of niet in combinatie met grotere of kleinere uitbreidingen naar palatinaal en/of buccaal en/of incisaal. Het materiaal bevatte, volgens de

indeling van Black, klasse III-, IV- en V-preparaties. De preparaties waren bewust zó gemaakt dat niet alle items van de lijst als goed zouden kunnen worden beoor-

deeld.

2.4. De uitvoering

Alle proefpersonen beoordeelden, met een week tussenruimte, de preparaties twee keer. Zo kon de intra-beoordelingsbetrouwbaarheid worden bepaald. De beoordeling was zó georganiseerd, dat dezelfde onderzoeksbegeleider gedurende beide beoordelingsrondes bij elke preparatie de items met hun criteria oplaas en de beoordeling noteerde. De proefpersonen moesten als beoordeling kiezen tussen 'goed' of 'fout'. Zij werden niet in de gelegenheid gesteld terug te gaan naar een reeds beoordeeld item of om hun scores te bekijken.

2.5. De analyse

In de analyse werd vooral aandacht besteed aan de intra- en inter-beoordelaarsbetrouwbaarheid.

2.5.1. De intra-beoordelaarsbetrouwbaarheid geeft aan in welke mate er door een en dezelfde beoordelaar op verschillende tijdstippen consistent wordt gescoord. Omdat het gebruik van de ontwikkelde itemlijst in

Tabel II. De intra- en inter-beoordelaarsbetrouwbaarheid voor de 11 items gebruikt door de twee proefgroepen.

Item	A Intra-beoordelaars- betrouwbaarheid		B Inter-beoordelaars- betrouwbaarheid	
	Studenten	Docenten	Studenten	Docenten
1	83.3	78.3	81.3	81.7
2	90.0	91.7	87.3	96.7
3	86.0	93.3	79.3	96.7
4	95.3	96.7	95.3	98.0
5	86.0	90.0	82.0	93.3
6	94.7	96.7	96.0	96.7
7	86.0	80.0	80.7	81.7
8	87.0	93.3	82.0	95.0
9	88.7	95.0	90.7	91.7
10	94.0	93.3	89.3	93.3
11	84.0	71.7	84.7	76.7
Gemiddeld	87.8	89.1	86.2	91.0

Tabel III. De intra- en inter-beoordelaarsbetrouwbaarheid van de docenten en studenten berekend over het totale aantal itemscores.

Kolom 1. De procentuele score van intra-beoordelaarsbetrouwbaarheid.

Kolom 2. De procentuele score van de inter-beoordelaarsbetrouwbaarheid.

Kolom 3. Het gemiddelde aantal goed gescoorde items per preparatie.

	1	2	3
Docenten			
A	91.5	90.0	9.4
B	91.5	87.0	8.5
C	87.3	88.2	8.7
D	86.1	89.4	9.2
Gemiddelde	89.1	88.7	8.95
Studenten			
I	87.9	82.6	8.9
II	84.8	83.2	9.0
III	84.8	80.6	9.2
IV	92.7	84.2	9.7
V	84.8	76.4	6.7
VI	87.3	82.0	8.0
VII	91.5	84.2	8.6
VIII	79.4	85.6	8.7
IX	89.7	79.1	7.7
X	88.5	82.9	8.9
Gemiddelde	87.1	82.1	8.5

dit onderzoek centraal staat is de intra-beoordelaarsbetrouwbaarheid voor alle proefpersonen voor elk item en voor de totaalscores bepaald.

Per item

Voor beide proefgroepen werd de intra-beoordelaarsbetrouwbaarheid berekend door voor elk item het totaal aantal identieke beoordelingen, in de twee beoordelingsronden, te bepalen. Voor item 1 is bijvoorbeeld voor de proefgroep van tien studenten het totaal bepaald van het aantal keren twee maal goed plus twee maal fout scores. Dit was 125 voor item 1 en dit is 83,3% van de maximaal te behalen score. Voor de docentengroep bedroeg dit

$$\frac{47}{15 \times 4} \times 100\% = 78,3\% \text{ (Tabel II, kolom A).}$$

De intra-beoordelaarsbetrouwbaarheid berekend over alle items

Vervolgens is berekend hoeveel de intra-beoordelaarsbetrouwbaarheid bedraagt als het totaal aantal gelijklopende itembeoordelingen als uitgangspunt wordt genomen. Maximaal kon $11 \times 15 = 165$ keer twee maal dezelfde score (++) of (--) gegeven worden. Docent A blijkt bijvoorbeeld 151 keer twee maal dezelfde score te hebben. Dit is 91,5% van het maximum van 165. (Tabel III, kolom 1)

2.5.2. De inter-beoordelaarsbetrouwbaarheid geeft de mate van overeenstemming aan tussen de beoordelingen van de verschillende beoordelaars. Evenals bij de intra-beoordelaarsbetrouwbaarheid wordt bij het berekenen van de inter-beoordelaarsbetrouwbaarheid het aantal gelijklopende itembeoordelingen als uitgangspunt genomen.

Per item

Voor beide proefgroepen werd de inter-beoordelaarsbetrouwbaarheid voor elk van de 11 items van de 15 preparaties bepaald. Een maximale overeenstemming van tien bestond per item per preparatie bij de studentengroep en vier bij de docentengroep. Voor elk item varieert het getal van maximale overeenstemming voor de studentengroep van tien (maximale overeenstemming) tot vijf (maximaal verschillend). Per item is het aantal keren bepaald dat in de 15 preparaties het getal tien werd gehaald (N10 genoemd), negen werd gehaald (N9 genoemd), enz. De inter-beoordelaarsbetrouwbaarheid is voor de studentengroep met de volgende formule te berekenen:

$$\frac{(N10 \times 10) + (N9 \times 9) + (N8 \times 8) + (N7 \times 7) + (N6 \times 6) + (N5 \times 5)}{15 \times 10} \times 100\%$$

Voor de docentengroep is dezelfde methode gebruikt uitgaande van de maximum overeenstemming (4) en maximaal ver-

schillend (2). De berekeningen werden dan uitgevoerd met de formule:

$$\frac{(N4 \times 4) + (N3 \times 3) + (N2 \times 2)}{15 \times 4} \times 100\%$$

Per beoordelaar

De inter-beoordelaarsbetrouwbaarheid van de individuele beoordelaars ten opzichte van de andere leden van zijn proefgroep werd op een wijze als boven beschreven berekend.

Per preparatie

De inter-beoordelaarsbetrouwbaarheid voor de totalen per preparatie werd berekend door een variantie-analyse uit te voeren.¹¹

3. RESULTATEN

Er werd een intra-beoordelaarsbetrouwbaarheid voor de afzonderlijke items gevonden van gemiddeld $87.8 \pm 9.8\%$ voor de studentengroep en van $89,1 \pm 17.4\%$ voor de docentengroep (tabel II, kolom A). De intra-beoordelaarsbetrouwbaarheid berekend over alle itemscores bedroeg voor de studenten gemiddeld $87.1 \pm 7\%$ en voor de vier docenten $89.1 \pm 3\%$ (tabel III, kolom 1).

Voor wat betreft de inter-beoordelaarsbetrouwbaarheid werden voor het gebruik van de 11 items afzonderlijke waarden gevonden van gemiddeld $86.2 \pm 9.8\%$ voor de studenten en $91.0 \pm 14.4\%$ voor de docenten (tabel II, kolom B). En de waarden voor de 'betrouwbaarheid' tussen de waarnemers was gemiddeld $82.1 \pm 5.7\%$ voor de studenten en $88.7 \pm 1.7\%$ voor de docenten (tabel III, kolom 2). De gemiddelden van de totalen van goed gescoorde items voor ieder van de 15 preparaties waren voor de studenten $8.5 \pm 1.8\%$ en voor de docenten $8.95 \pm 0.45\%$ (tabel III, kolom 3). De variantie-analyse van inter-beoordelaarsbetrouwbaarheid van de totaal goed gescoorde items voor iedere preparatie door de tien studenten en vier docenten gaf aan dat er geen significante verschillen binnen de beide proefgroepen aanwezig is ($p < 0.05$).

4. DISCUSSIE

Mackenzie stelt, dat men tevreden kan zijn wanneer inter-beoordelaarsbetrouwbaarheidwaarden tussen 80 en 90% worden behaald. Het kost veel moeite om deze waarden met extra training, herschrijven

van de criteria of het gebruik van betere meetmethoden te verhogen en de toename blijkt dan veelal gering te zijn.¹ Ook de

relevantie van de betrouwbaarheidsgraad ten aanzien van hetgeen geëvalueerd wordt dient in overweging genomen te worden bij een beslissing om beoordelingscriteria wel of niet te accepteren. Zo zullen bijvoorbeeld waarden van 95% haalbaar zijn als twee beoordelaars speciaal worden getraind in het gebruik van zeer specifieke criteria.¹³ Indien de criteria echter door grote groepen gebruikers in een klinische situatie moeten worden gebruikt, dan moeten de criteria eenvoudig en snel te gebruiken zijn en zal een lager percentage acceptabel zijn. In dit opzicht kunnen de in dit onderzoek gevonden waarden van 85-90% als ruim voldoende worden beschouwd, daar de deelnemers niet waren getraind, de itemlijst snel, eenvoudig en in klinische situaties toepasbaar is en geen speciale meetapparatuur nodig is.

Elk item wordt beoordeeld op een tweepuntsschaal (goed of fout), hetgeen tot gevolg heeft dat de gevonden waarden voor een deel door kans verkregen worden. Er mag zelfs worden aangenomen dat alleen door kans in deze beoordelingsopzet een overeenstemming van $\pm 50\%$ gehaald wordt. Het is dus nodig om na te gaan of de gevonden waarden significant van de 50%-score verschillen. Dit laatste werd op 5%-niveau gedaan met behulp van de formule:¹¹

$$(P_1 - P_2) \pm 1.96 \sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}$$

In deze formule werd $P_2 = 50\%$ gesteld en dan bleek bij de intra-beoordelaarsbetrouwbaarheid voor het gebruik van de items door de studenten en docenten hun laagste waarden van respectievelijk: 83.3 (item 1) en 71.7 (item 11) nog significant van kans te verschillen ($p < 0.05$). Ook voor de inter-beoordelaarsbetrouwbaarheid kon op dezelfde wijze worden aangetoond dat de gevonden scores van de tien studenten en vier docenten allemaal significant van kans verschilden ($p < 0.01$). De gevonden waarden behaald voor de items door de studenten en docenten, bleken niet significant te verschillen ($p > 0.05$). De docenten scoren gemiddeld hoger dan de studenten. Dit is misschien te verklaren uit het feit dat de docenten vertrouwd zijn met het gebruik van criteriumlijsten. De studenten daarentegen traden voor het eerst als beoordelaar op. Wél waren zij gewend om met dit soort lijsten te worden beoordeeld.

Een vergelijking van de waarden met andere gepubliceerde data is moeilijk, daar de resultaten sterk beïnvloed worden door de soort criteria en preparatievormen. Ook kunnen de preparaties in het onderzoek de resultaten hebben beïnvloed. In dit onderzoek zijn daarom de resultaten voor elke

preparatie nog eens nagegaan. Toen bleek de correlatie tussen inter- en intra-beoordelaarsbetrouwbaarheid voor de docenten .60 en voor de studenten .70 te bedragen. Dit geeft wel aan dat, hoewel de totaal scores redelijk overeenstemmen, de beoordeling per item door de beoordelaars per preparatie toch aanzienlijk verschilt. Er bleek echter géén correlatie tussen specifieke typen van preparaties en de betrouwbaarheidsscores, hetgeen aangeeft dat de itemlijst voor elke preparatie te gebruiken is. Uit de waarden in tabel III, kolom 2 kan worden afgeleid dat de docenten een significant hogere inter-beoordelaarsbetrouwbaarheid hebben dan de tien studenten ($p < 0.001$). Er was echter een kleiner verschil tussen de twee groepen wanneer werd gekeken naar de intra-beoordelaarsbetrouwbaarheid ten aanzien van de individuele items. Deze elkaar tegensprekende waarden zijn mogelijk te

verklaren door de mate van homogeniteit binnen de twee groepen. De studentengroep is minder homogeen, hetgeen duidelijk door student no. V wordt gedemonstreerd (tabel III).

Als conclusie kan worden gesteld dat de resultaten in dit onderzoek aantonen, dat het mogelijk is een criteriumlijst voor composietpreparaties voor de klinische situatie te ontwikkelen. De criteria zijn gebaseerd op de huidige kennis op het gebied van composietmateriaal en blijken gemakkelijk en zonder training, uitgebreide beschrijvingen of meetinstrumenten te gebruiken. De beschreven itemlijst lijkt daarom geschikt om te worden toegepast in zowel de preklinische als de klinische (postacademische) opleiding. Soortgelijke itemlijsten zouden op dezelfde wijze voor andere gebieden in het tandheelkunde-onderwijs kunnen worden ontwikkeld.

SUMMARY

A CRITERIA LIST FOR EVALUATING COMPOSITE PREPARATIONS

Keywords: Education – Conservative dentistry – Composite – Preclinic

An eleven point criteria list for cavity preparation for composite restorative materials was tested in this study. The reliability with which it could be used by experimental staff and student groups without any descriptive material or training was determined. The members of both groups judged 15 preparations made in extracted carious human teeth placed in the correct anatomical position in a phantom head.

There was an inter- and intra group reliability for the eleven items of approximately 88%. The mean intra-observer reliability when using these items was 88%, with a lower mean inter-observer reliability of 85%. There were no significant differences between the two observer groups.

These results show that it is possible to develop a highly reproducible criteria list which can be used without training, additional descriptive material or measuring instruments.

LITERATUUR

- MACKENZIE RS. Defining clinical competence in terms of quality, quantity and need for performance criteria. *J Dent Educ* 1973; 37(9):37-44.
- PRATRIDGE MA, MAST TA. Dental clinical evaluation: A review of the research. *J Dent Educ* 1978; 42(6):300-5.
- VANN WF, MACBE JB, HAMSELL PB. Effects of criteria and checklists on reliability in preclinical evaluation. *J Dent Educ* 1983; 47(10):671-80.
- EDWARDS WG, MORSE PK, MITCHELL RJ. A practical evaluation system for preclinical restorative dentistry. *J Dent Educ* 1982; 46(12):693-96.
- FEIL PH. An analysis of the reliability of a laboratory evaluation system. *J Dent Educ* 1982; 46(8):489-94.
- GROEPFERD SJ, KERBER PE. A comparison of two methods for evaluating primary class II cavity preparations. *J Dent Educ* 1980; 44(9):537-42.
- O'CONNOR P, LOREY RE. Improving interrater agreement in evaluation in dentistry by use of comparison stimuli. *J Dent Educ* 1978; 42(4):174-79.
- PURDELL LEWIS DJ. A modern restorative dentistry programme. *Neth Dent J* 1982; 89:29-34.
- BOWEN RL. Composite and sealant resins – past, present and future. *Pediatr Dent* 1982; (4)10-15.
- WIEGMAN JE, CORBA NHC. Beschrijving van een onderwijsopzet voor het geïntegreerd behandelen van patiënten. *Ned Tijdschr Tandheelk* 1983; 90(12):613-18.
- WONNACOT TH, WONNACOTT RJ. *Introductory statistics*, New York, Santa Barbara, London, Sydney, Toronto, John Wiley & Sons, 1977.
- MACKENZIE RS, ANTONSON DE, WELDY PL, WELSCH BB, SIMPSON WJ. Analysis of disagreement in the evaluation of clinical products. *J Dent Educ* 1982; 46(5):284-89.
- BACKER DIRKS O, VAN AMERONGEN J, WINKLER KC. A reproducible method for caries evaluation. *J Dent Res* 1951; 30:346-59.