# Visual and profilometric wear measurements

Karl-Johan M. Söderholm, Marion J. Roberts, Donald E. Antonson, Kenneth J. Anusavice, André P. Mauderli, David C. Sarrett and John W. Warren

Department of Dental Biomaterials, College of Dentistry, University of Florida, Gainesville, Florida, USA

Söderholm K-JM, Roberts MJ, Antonson DE, Anusavice KJ, Mauderli AP, Sarrett DC, Warren JW. Visual and profilometric wear measurements. Acta Odontol Scand 1992;50: 121–127. Oslo. ISSN 0001-6357.

Wear of composites can be estimated by the degree of marginal discrepancy between the prepared cavity wall and the occlusal margins of composites. Such evaluations are done on casts by comparing and rating the marginal discrepancy with those on standard casts. We analyzed the reliability of this technique on metal and stone specimens. These specimens contained grooves of different width and depth. For the visual comparison we used stone casts of machined standards of known groove depth. We measured the depths of the metal specimens with a profilometer and made stone casts of these original specimens. Using the stone casts of the standards, five dentists estimated the unknown groove depths on the remaining stone casts. These estimates were done under standardized conditions and repeated by each dentist on five different occasions. The results showed that visual depth evaluations of die stone specimens underestimated the depths when compared with the values measured with a profilometer on the original metal models. One investigator gave significantly different ($p < 0.05$) groove depth estimates at different occasions. □ *Composites; dental materials; marginal discrepancy; operative dentistry*

*Karl-Johan Söderholm, Department of Dental Biomaterials, College of Dentistry, Gainesville, FL 32610-0446, USA*

According to previous guidelines for the acceptance of a posterior composite, the wear loss should be less than 250 µm after 5 years of clinical service (1). In 1989 these guidelines were revised, and posterior composites were classified into two categories (2). Category A composites are for unrestricted use in posterior teeth, suitable for all class-I and class-II restorations, whereas category-B composites are for restricted use in posterior teeth. The restricted use implies that category-B composites are suitable for small, conservative cavity preparations in which the occlusal stress is limited. To qualify as a category-A composite, no more than 50 µm will wear during a 2-year period measured from 6 months after insertion, and no more than 100 µm will wear during 4 years measured from 6 months after insertion. For category-B composites the corresponding wear numbers are 125 µm and 175 µm. The material loss mentioned in the old and the new acceptance programs can be determined by comparing replicas of the restorations under investigation with replicas of restorations with known material loss. The estimated material loss is recorded close to the margins, and the mean value of repeated estimates is used to calculate the wear rate (3, 4). Naturally, such a method has several limitations. First, it can be described as a quantitative method with low resolution power. Second, the technique cannot identify and measure localized wear, which is often much more pronounced than the general wear (5–8). Third, the technique requires fixed reference lines (that is, margins) for comparison, making this approach more suitable for butt joint cavity preparations than for beveled preparations.

Other techniques, such as use of profilometers (7–9), measuring microscopes (5, 6), interferometers (10), and stereoplotters (11), have also been explored. The advantages with the latter techniques are that they offer superior resolution and can

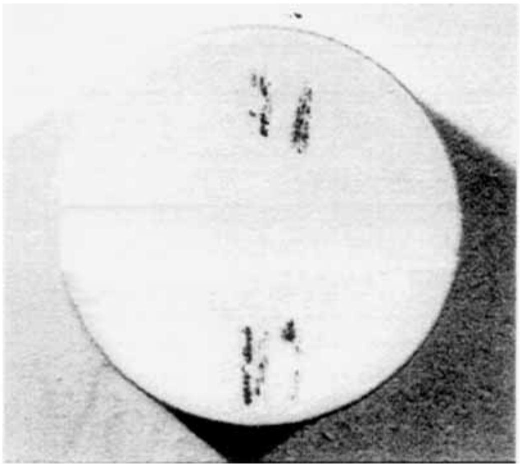

Fig. 1. A stone cast with unknown groove depth. The pencil marks on the two sides of the groove mark the band region within which the estimates were conducted.

record localized wear. The disadvantages of these techniques are that they are more complex and time-consuming than the visual estimation technique.

Despite the limitations of the visual evaluation technique, one can assume that this technique will persist as the most popular wear evaluation technique for years to come. Therefore, it is important to find the reliability and identify sources of error associated with this technique. Errors could include differences in wear depth and differences among investigators.

The objective of the present study was to estimate the magnitude of errors associated with the visual wear estimation technique (3). The evaluation consists of comparing visually estimated values on stone casts with values generated with a profilometer on the original metal specimens. This comparison simulates wear evaluations at various locations relative to the margins and also estimates marginal discrepancy differences within and among evaluators.

## Materials and methods

Fourteen grooves ranging from 2 to 4 mm in width and from 40 to 500 μm in depth were cut with a handpiece and carborundum stones into 10-mm-wide and 5-mm-high brass cylinders. The grooves were intentionally given irregular shapes and surface textures to cause detraction during the visual evaluation. Besides these nonstandardized specimens, standards machined with flat grooves 3.0 mm wide and 100, 200, 300, 400, or 500 μm deep with a precision lathe into similar brass cylinders were prepared. On each specimen, including the standards, we marked a local region, approximately 1 mm wide, on each edge of the groove. These markers were placed to make it possible to identify a particular location for the simulated wear evaluation. In addition, on one side of the groove the specimen was marked for identification purposes. The margin on the marked side was labeled 'the top margin', and the opposite margin was labeled 'the bottom margin'. One impression was made of each specimen with a poly(vinyl)-siloxane impression material (Mirror 3, Sybron/Kerr, Romulus, Mich., USA), from which we made one die stone cast (Silky-Rock, Whip Mix Corp., Louisville, Ky., USA) per impression (Fig. 1).

Five dentists conducted the visual evaluation of the 14 samples. Before they started the evaluation, each investigator was instructed how to conduct the evaluation in a standardized manner. In addition, they also received a protocol each time they conducted an evaluation, describing the evaluation procedure. They estimated the depth of each specimen within 1 mm of each edge of the groove (top and bottom margin) and at the deepest location between the two edges (center). The evaluators used a ×2 binocular magnifier while evaluating the specified regions marked on the casts. The five dentists repeated the evaluations on five different occasions. At least 2 weeks passed between each of the evaluations.

The profilometric measurements (Surfanalyzer 150 System, Clevite Corp., El Monte, Calif., USA) were performed on the original metal dies rather than on the gypsum dies. We recorded the marked regions on five different occasions. During each of these occasions, we repositioned the die and measured each marked area five times. The

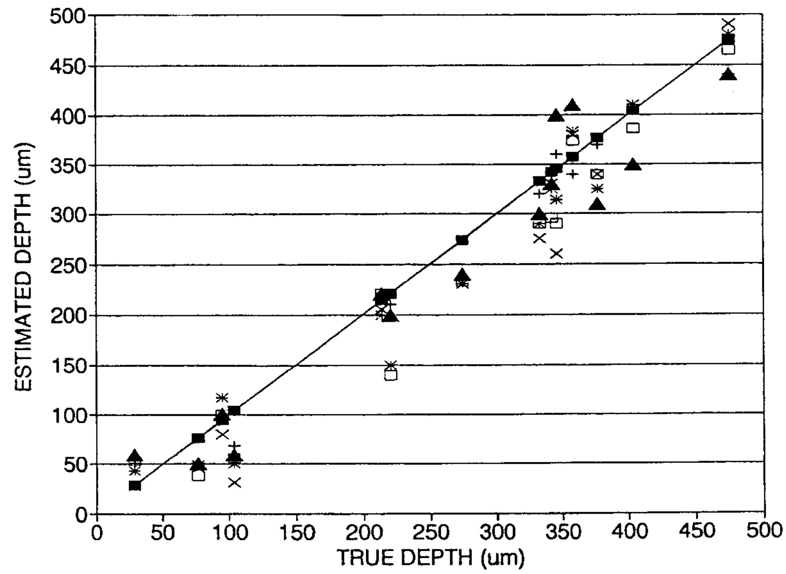


Fig. 2. The estimated depth values adjacent to the top margin of the groove (the margin most distant away from the investigator and closest to the light source). The diagonal line represents the estimated versus the true depth when measured with the profilometer. The filled rectangles represent the values of the 14 specimens determined with the profilometer. The other 5 symbols represent each evaluator and the mean depth estimate given for each of the 14 samples.

deepest measured value obtained within the top, bottom, and center regions was noted as the numerical value for the profilometric determination.

The pooled $5 \times 5 \times 14$ visual data and the corresponding $5 \times 5 \times 14$ profilometric measurements from the top, center, and bottom sites were compared. After these evaluations we pooled all top, center, and bottom values per investigator and compared the values to decide whether significant differences existed between the five investigators. The values for each investigator were then compared by pooling all values that each particular investigator generated on each of the evaluation occasions.

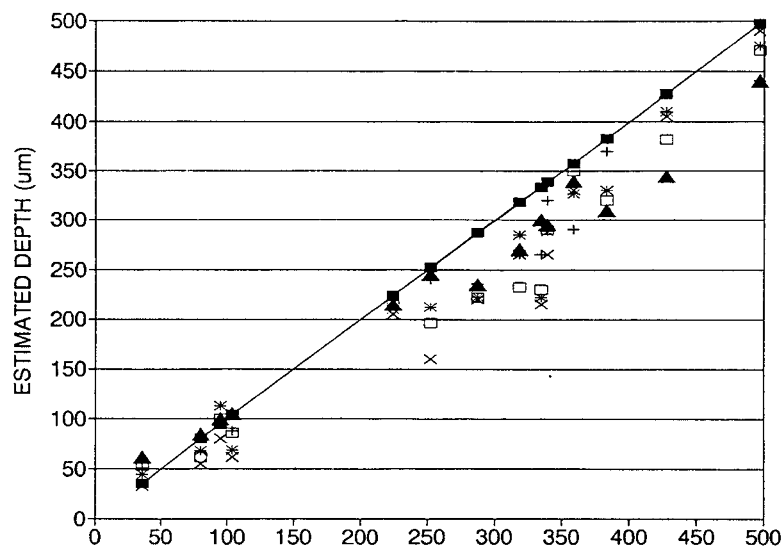


Fig. 3. The estimated depth values for the central part of the groove.

We then compared these pooled values among the different evaluation occasions.

The Kruskal–Wallis test was selected for the statistical analyses because we suspected a large variability among investigators and also because the collected data for each specimen could deviate from a normal distribution.

## Results

The pooled mean value for the deepest location, which was the central location, was 266 μm (SD, 137.1 μm) when measured with the profilometer. The pooled mean visual value for the same location was 235 μm (SD, 131.7 μm). These values were significantly different ($P < 0.01$).

Comparing the pooled data for the different investigators showed no significant differences. However, when the estimates of individual dies were compared, significant differences could be found between the different investigators.

As a general finding, the profilometer values were the highest, whereas the lowest depth estimates were randomly distributed among the five investigators (Figs. 2–4). By plotting the visually estimated values as a function of the depth generated by the profilometer (Figs. 2–4) and using linear regression analysis for the visually estimated values, the error at different depths could be estimated. By using this approach and separating the different estimates into subgroups—the top, center, and bottom values—additional information was obtained. Thus, at the top margin the depth was underestimated by approximately 15 μm independent of sample depth (Fig. 2). Rather than having a fixed depth error independent of groove depth, the linear regression suggested that the central and the bottom regions were underestimated by 12–13% (Figs. 3 and 4). Moreover, large variations existed between estimates made on different occasions by the different investigators (Tables 1–3). One of the five investigators gave significantly ($P < 0.05$) different readings on each of the different recording occasions.

## Discussion

This study shows that visual depth estimate on gypsum casts of simulated wear results in

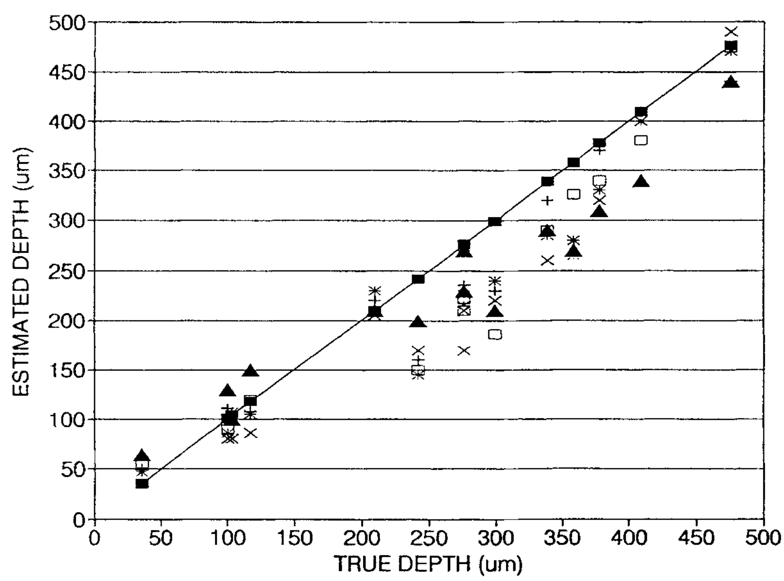


Fig. 4. The estimated depth values adjacent to the bottom part of the groove (the margin most distant away from the light source and closest to the investigator).

Table 1. Depth estimations at the top location of a specimen with a depth close to the accepted level of the acceptance programs

| Investigator | Mean (μm) | Median (μm) | SD (μm) | Min. (μm) | Max. (μm) |
|---|---|---|---|---|---|
| 1 | 210 | 200 | 89.4 | 100 | 300 |
| 2 | 150 | 150 | 35.4 | 100 | 200 |
| 3 | 140 | 100 | 54.8 | 100 | 200 |
| 4 | 150 | 150 | 61.2 | 100 | 250 |
| 5 | 200 | 200 | 70.7 | 100 | 300 |
| Profilometer | 219 | 218 | 9.4 | 210 | 235 |

$n = 5 \times 5$ per investigator.

Table 2. Depth estimations at the center location of a specimen with a depth close to the accepted level of the acceptance programs

| Investigator | Mean (μm) | Median (μm) | SD (μm) | Min. (μm) | Max. (μm) |
|---|---|---|---|---|---|
| 1 | 240 | 200 | 82.2 | 150 | 350 |
| 2 | 213 | 200 | 62.0 | 150 | 300 |
| 3 | 196 | 200 | 36.5 | 150 | 250 |
| 4 | 160 | 175 | 37.9 | 100 | 200 |
| 5 | 245 | 225 | 62.42 | 200 | 350 |
| Profilometer | 252 | 250 | 8.4 | 245 | 265 |

$n = 5 \times 5$ per investigator.

Table 3. Depth estimations at the bottom location of a specimen with a depth close to the accepted level of the acceptance programs

| Investigator | Mean (μm) | Median (μm) | SD (μm) | Min (μm) | Max (μm) |
|---|---|---|---|---|---|
| 1 | 270 | 300 | 97.5 | 150 | 400 |
| 2 | 226 | 200 | 78.3 | 150 | 350 |
| 3 | 210 | 200 | 74.2 | 100 | 300 |
| 4 | 170 | 200 | 44.7 | 100 | 200 |
| 5 | 270 | 300 | 44.7 | 200 | 300 |
| Profilometer | 276 | 275 | 6.5 | 270 | 285 |

$n = 5 \times 5$ per investigator.

underestimation of the true wear value. A possible explanation as to why the top margin locations seemed to give a fixed depth error of about 15 μm and the other two locations an underestimate of 12–13% could relate to the location of the light source during the evaluation. In this study the light source was located so as to cast a sharp shadow from the top margin on the surface of the grove. Since the light contacted the surface at an angle of 45° and was directed toward the evaluator, it did not produce any shadow on the margin closest to the investigator (the 'bottom margin'). The central part also did not produce any shadow because of its geometry. Because of these differences in shadowing ability of the different regions, we suspect that the eye interpreted the depth of these regions differently. This possibility suggests that future studies determine how a specimen under investigation should be held relative to the light source.

Whether the tendency to underestimate the depth only relates to the visual estimation or to the entire procedure including replication technique has not been resolved in this study. Since we only used the profilometer on metal dies, the possibility that the values differed between metal and stone cast specimens cannot be excluded. The reason for measuring the metal dies only with the profilometer was our concern that the sliding stylus could cause wear of the gypsum dies, especially in regions with fragile gypsum edges. In addition, by making the measurements on the metal dies, we felt that it would be possible to characterize the accuracy of the entire visual estimation procedure. With this approach, even errors introduced during replication would be included in the total error of the visual wear estimation.

This study showed that there were no significant differences in pooled mean values between the different investigators. This finding is explained by considering the variability between different readings for a particular specimen (Tables 1–3). Still, despite large variations, it could be shown that one investigator gave significantly different estimates on different occasions. This finding suggests that visual wear evaluation is not suitable for longitudinal comparisons when the evaluations are made on different occasions. Thus, during a longitudinal study there is a risk that overestimation and underestimation of wear may occur compared with previous or later observations. Such errors could lead to misinterpretations of wear patterns—that is, whether the wear rate increases or decreases with time. To reduce these potential errors, we suggest that visual evaluations be made retrospectively rather than longitudinally and that magnifying lenses are used to improve the visual resolution. In addition, these retrospective evaluations should be repeated at several different occasions. With such considerations, it is possible to reduce the final error of a visual wear evaluation to a magnitude of 12–13%.

No investigator gave consistently too low or too high values. The standard deviations of the estimated values were for some specimens as much as 20 times higher for the visual measurements than for the profilometric measurements. The range of these readings sometimes exceeded 100 μm (Tables 1–3). Such a large variation results in poor discrimination between the different examiners. Poor discrimination explains why no significant differences could be detected between the investigators.

The findings presented in this study raise serious doubts about the reliability of clinical results generated with this method. These concerns relate to the following points. First, the area under investigation in our study was clearly defined when the study was initiated (Fig. 1). In most clinical studies predetermined locations are not identified at base line. In addition, during clinical wear, material loss occurs at different locations and at different wear rates. Under such conditions, one must question which region will be selected for the wear evaluation. These factors complicate a longitudinal visual evaluation. Second, the human eye has a resolution power no better than 100 μm when an object is 25 cm away from the viewer (12). With the use of a ×2 binocular magnifier, we could improve the resolution to 50 μm. Third, it is not a general practice in clinical wear evaluations to repeat the estimates by the same individual investigator as many times as required in this study. Our findings suggest that future clinical wear estimates should be based on repeated estimates made by the same investigator on completely different occasions.

Considering the inherent limitation with the visual evaluation, we suggest that wear results generated with this technique should not be reported to single micrometers. Our findings further suggest that the large variability in the visual wear estimate technique should be considered in any future revision of the acceptance program guidelines for posterior composites. In such program documents, information should be given regarding the magnitude of error acceptable for the evaluation technique.

*Acknowledgement.*—This study was supported by NIH-NIDR grant DE07578.

## References

1. ADA council on dental materials, instruments, and equipment. Acceptance program for composite resin materials for occlusal class I and class II restorations. Chicago: American Dental Association, 1984.

2. ADA council on dental materials, instruments, and equipment. Guidelines for submission of composite resin materials for posterior restorations. Chicago: American Dental Association, 1989.
3. Leinfelder KF, Barkmeier WW, Goldberg AJ. Quantitative wear measurements of posterior composite resins [abstract]. J Dent Res 1983;62 Spec issue: 671.
4. Leinfelder KF. Evaluation of clinical wear of posterior composite resins. In: Vanherle G, Smith DC, editors. Posterior composite resin dental restorative materials. Utrecht: Peter Szulc Publishing Co, 1985:501–9.
5. Lambrechts PM, Vuylsteke G, Vanherle G, Davidson CL. Quantitative in vivo wear of posterior dental restorations: Four year results [abstract]. J Dent Res 1985; 64 Spec issue: 370.
6. Lambrechts PM, Braem M, Vanherle G. Accomplishments and expectations with posterior composite resins. In: Vanherle G, Smith DC, editors. Posterior composite resin dental restorative materials. Utrecht: Peter Szulc Publishing Co, 1985:521–40.
7. Roulet JF, Reich TH, Lutz F. High precision occlusal mapping. A new method for measuring wear of posterior composites [abstract]. J Dent Res 1983; 62 Spec issue 220.
8. Roulet JF. In vivo wear measurements of composite resins. In: Vanherle G, Smith DC, editors. Posterior composite resin dental restorative materials. Utrecht: Peter Szulc Publishing Co, 1985:365–72.
9. DeLong R, Pintado M, Douglas WH. Measurement of change in surface contour by computer graphics. Dent Mater 1985;1:27–30.
10. Atkinson JT, Gloves D, Lacor MJ, Cunningham J, Williams DF. The measurement of wear in dentin restorations using laser dual source contouring. Wear 1982;76:91–104.
11. Hintz RJ, Söderholm K-JM, Sarrett DC. Tooth morphology quantification using non-metric photography in an analytical system. ASPR-ACSM Annual Convention. Non-Topogr Photogram 1987;7:71–8.
12. Overheim RD, Wagner DL. Light and color. New York: John Wiley and Sons, Inc, 1982:166.


Received for publication 3 May 1991