Some issues in applied statistics in clinical restorative dental research
Tobi, H.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 7

Cost-effectiveness of composite resins and amalgam in the replacement of amalgam Class II restorations

H Tobi¹ ², CM Kreulen¹, H Vondeling², WE van Amerongen¹

¹ Department of Cariology Endodontology Pedodontontology, Academic Centre for Dentistry Amsterdam, The Netherlands
² Department of Epidemiology and Biostatistics, Vrije University Amsterdam, The Netherlands
Abstract

Objectives: The replacement of an old amalgam Class II restoration is a common treatment and will remain so for decades. In addition to effectiveness, possible adverse health effects and aesthetics, costs of the treatment options will play a role in the choice of material. The aim of this study was to yield information on the relative cost-effectiveness of the use of composite resins and amalgam for the rerestoration of amalgam Class II restorations.

Methods: As part of a larger randomized clinical trial, treatment effectiveness and treatment costs were estimated in 73 composite and amalgam Class II posterior rerestorations. The main treatment outcome was longevity. Secondary outcomes included need of repair and quality of the margin while in situ. Costs were analyzed from the perspective of dentistry, assuming a treatment strategy aimed at offering 'value for money'. From this perspective, differential costs were based on personnel costs as approximated by treatment time.

Results: Replacing an amalgam Class II restoration with amalgam is associated with lower costs than replacing with a composite resin. A sensitivity analysis, considering type of composite, increasing proficiency with the material, and time needed for future removal of material, demonstrated that these differences are fairly robust. The materials performed equally well for the first 5 years after placement with respect to longevity. Differences in secondary outcomes were minor and not all in favor of the same material.

Conclusions: It is tentatively concluded that amalgams are more cost-effective than composites for replacing existing Class II amalgam restorations.
Introduction

In clinical practice, the choice between amalgam and composite for replacement of an old posterior amalgam restoration is a daily occurrence. Until a decade ago, the durability of the materials at hand was decisive, and this favored amalgam. Nowadays, many factors influence and complicate the choice of restorative materials such as esthetics and alleged adverse health effects. Moreover, cost-effectiveness may play a role in treatment decisions. In economic evaluation, the inclusion of individual categories of costs depend on the perspective taken in the analysis (Gold, Siegel \textit{et al.}, 1996; Drummond, O'Brien \textit{et al.}, 1997). The perspective may be societal, to assist policy makers decide on reimbursement of a recently introduced treatment. Other possible perspectives are the patient, the insurance companies and, of course, the field of restorative dentistry. Some interesting studies on the relative cost-effectiveness of dental materials have been done. Smales and Hawthorne (1996) described the cost-effectiveness of different materials and types of restorations relative to Class I amalgam restorations, using fees as a cost-estimate and survival estimates for effectiveness. Mjör, Burke and Wilson (1997) reported relative cost of restoring and rerestoring teeth during a person's lifetime using amalgam Class I and MOD restorations as standard. Longevity is based on information provided by general practitioners and clinical studies. One important drawback of these studies is that they are cross-sectional and factors that have led to the choice of a particular restorative material in a certain situation confound and bias the cost-effectiveness analysis. In addition, fees do not accurately reflect the treatment time needed and the other expenses involved, but instead reflect the result of a negotiation process. Furthermore and partially as a consequence, the relative cost of restorations using fees differs between countries (Mjör, Burke \textit{et al.}, 1997).

Many studies on costs of dental restorative treatment compare treatment times (Hendriks, Letzel \textit{et al.}, 1985; Plasman and van't Hof, 1987; Dilley and Vann, 1990; Verzijden, Creugers \textit{et al.}, 1990; Kreulen van Amerongen \textit{et al.}, 1991a; Advokaat, van't Hof \textit{et al.}, 1992; Burgess, Haveman \textit{et al.}, 1992; Kreulen, van Amerongen \textit{et al.}, 1992; Meijering, Creugers \textit{et al.}, 1995). Most of these studies focused on the detection of factors determining treatment time, such as dentist and size of restoration. Although time studies in the literature aim to yield information on cost-effectiveness, they often fail to include information on treatment outcome or use treatment effectiveness as estimated in other studies from the literature. Gathering information on costs and effectiveness from the same population is extremely important because costs and effectiveness may differ between populations. Hence, it is strongly recommended that data on costs are collected alongside the clinical trial
studying the effectiveness of treatments, even though it may complicate the study (Drummond and Davies, 1991; Gold, Siegel et al., 1996; Drummond, O'Brien et al., 1997). This study reports on the cost-effectiveness of composite resins and amalgam used to replace old amalgam Class II restorations from the perspective of the field of dentistry. Both treatment costs and treatment effectiveness are estimated from the same Class II restorations made in the context of a randomized clinical trial.

Material and Methods

Materials

In a controlled clinical trial described in more detail elsewhere (Kreulen, van Amerongen et al., 1991b), 244 standard conservative Class II restorations (no cusps involved and cervical outline located in enamel) were made in 56 patients. These patients were primarily recruited by the out-patient clinic of the Academic Centre for Dentistry Amsterdam from half-year check-up appointments. They were required to be between 15 and 35 years of age and in good general health and to sign an informed consent. Each patient was given a series of four restorations, made of four different materials Herculite XR (Kerr), Clearfil Ray Posterior (Kuraray/Cavex Holland), Visiomolar (ESPE) and Tytin (Kerr). Which tooth received what material was randomly assigned. Of these 244 restorations, 101 were new fillings and 69 expanded on previous fillings. Seventy-four were complete rerestorations and except for one composite which is excluded, all these completely replaced restorations were amalgam. The remaining 73 existing amalgam Class II fillings were replaced by 53 composite restorations and 20 amalgam restorations. These 73 complete rerestorations involved 35 patients who were all available for the final evaluation after five years. In previous studies, the factors ‘number of surfaces’ and ‘type of tooth’ were found influence on the treatment time (Hendriks, Letzel et al., 1985; Kreulen, van Amerongen et al., 1991a; Advokaat, van’t Hof et al., 1992; Kreulen, van Amerongen et al., 1992). The frequency distribution for the 73 restorations with respect to number of surfaces and type of tooth is given in Table 1.

Costs and cost-analysis

The differential treatment costs are important for the purpose of comparison. Costs that do not depend on the method if treatment are therefore not incorporated in a so-called incremental cost-effectiveness study. For example, cost of administering anesthetics can be ignored because the decision to
apply anesthetics does not depend on the material and does not influence differential costs. The costs of the filling materials are negligible and hence do not add to the cost difference compared to the main determinant of costs, the personnel costs. Thus, in accordance with the literature, treatment time is used to approximate treatment costs.

Time was measured in five treatment stages by the assistant using a series of stopwatches. The treatment stages are preparation, preliminary work, application, finishing, and polishing, the latter during a separate visit (Table 2). These five stages include only those operations required for the placement of the restoration. Details on time per treatment stage are given elsewhere (Kreulen, van Amerongen et al., 1991a). For present purposes the total time in minutes, being the sum of the stages, is used.

**Effectiveness**

Longevity of the restoration served as primary treatment outcome. As secondary outcome variables the number and type of repairs needed while in situ were selected. Since dentists frequently report marginal adaptation and marginal discoloration as reasons for replacement (Mjör and Qvist, 1997), marginal adaptation and discoloration were added as secondary outcome variables. Two trained and calibrated dentists assessed marginal adaptation and marginal discoloration, independently of each other. The issue of observer variation will be addressed in the section statistical analysis.

For the assessment of quality, the restoration margin was divided into several sections (Figure 1). Marginal adaptation was scored for all sections of the restorations separately on a four point scale derived from USPHS criteria (Table 3). Because the treatment decision whether or not to replace the restoration was based on the worst section, the maximum score was chosen to represent the marginal adaptation of the restoration for each assessor. To obtain one score for each restoration, the two scores were averaged. Marginal discoloration was scored by categorizing the number of sections discolored (0, 1-4, 5-8, 9-12, more than 12). Polishing for the second time and adding material were agreed to be regarded types of mending.

---

Table 1. Frequency distribution of the rerestorations with respect to type of tooth and number of surfaces

<table>
<thead>
<tr>
<th>Type of restoration</th>
<th>MO/DO</th>
<th>MOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Material</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Premolar</td>
<td>35</td>
<td>10</td>
</tr>
<tr>
<td>Molar</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Composite Amalgam Composite Amalgam</td>
<td></td>
</tr>
</tbody>
</table>
### Table 2. Treatment stages *per* type of material

<table>
<thead>
<tr>
<th>Treatment Stage</th>
<th>Composite resins</th>
<th>Amalgam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preparation A</td>
<td>Removal of old material&lt;br&gt;Preliminary work A Application of rubber dam and nail polish</td>
<td>Removal of old material&lt;br&gt;Application of rubber dam</td>
</tr>
<tr>
<td>Preparation B</td>
<td>Selective beveling of the outline&lt;br&gt;Preliminary work B Application of Ca(OH$_2$) when necessary, glass ionomer cement base, matrix and wedges. Etching procedure.&lt;br&gt;Application of bonding agent</td>
<td>Application of Ca(OH$_2$) when necessary, matrix and wedges</td>
</tr>
<tr>
<td>Application</td>
<td>Application of resin composite and intermediate light-curing</td>
<td>Application and condensation of amalgam</td>
</tr>
<tr>
<td>Finishing</td>
<td>Removal of the excess of material and contouring; removal of rubber dam and verification of the occlusal height</td>
<td>Removal of the excess of material and contouring; removal of rubber dam and verification of the occlusal height</td>
</tr>
<tr>
<td>Polishing</td>
<td>Verification of marginal adaptation/occlusal height and adjustment if necessary; polishing</td>
<td>Verification of marginal adaptation/occlusal height and adjustment if necessary; polishing</td>
</tr>
</tbody>
</table>

#### Statistical analysis

Overall cumulative survival can be summarized by a Kaplan-Meier curve (Eriksson and Adell, 1994; Aalen, 1994). When estimating the effect of material on the risk of replacement, multiple restorations within patients need to be accounted for, which can be achieved by a logistic regression with random components (Tobi, Kreulen *et al.*, 1998).
Table 3. Description of four point scale for marginal adaptation.

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No visible evidence of a crevice along the section of the margin, no</td>
</tr>
<tr>
<td></td>
<td>catch of an explorer</td>
</tr>
<tr>
<td>2</td>
<td>Visible evidence of a crevice along the section of a margin, an explorer</td>
</tr>
<tr>
<td></td>
<td>catches slightly</td>
</tr>
<tr>
<td>3</td>
<td>An explorer penetrates into a crevice but no dentine or base is exposed</td>
</tr>
<tr>
<td>4</td>
<td>An explorer penetrates into a crevice, dentine or base exposed</td>
</tr>
</tbody>
</table>

Figure 1. Division of the margin in sections for molars and premolars.
Chapter 7

We evaluated observer variation before adopting marginal adaptation and marginal discoloration as secondary outcomes, to ascertain reliability. Kappa is a measure for agreement beyond chance with values between '0' (no agreement beyond chance) and '1' (complete agreement). A threshold is set for the lower limit of the 95% confidence interval of kappa at 0.60, a value which can be interpreted as good agreement beyond chance (Altman, 1991). When the lower limit of the 95% confidence interval of kappa was below this threshold, the assessment of that variable was considered unreliable and ignored in the further analysis. Two statistical tests were used to help interpret a specific value of kappa. McNemar’s test for symmetry around the main diagonal of a square table and the Test of Marginal Homogeneity were both regarded not significant at the P> 0.10 level.

In the current study design, each patient had a maximum of one amalgam restoration only. Many patients, however, have more than one composite restoration. This multiple units within patients phenomenon is a well-recognized problem in periodontology (Blomqvist, 1985; Imrey, 1986; Osborn, 1997). In clinical restorative dental research, the problem is often ignored although this results in underestimated standard errors, and hence, too narrow confidence intervals and too small p-values. With respect to the secondary outcomes, the impact of this phenomenon was studied as follows. For restorations in patients with more than one composite restoration, the average within patient variation was computed. Straightforward variations for these restorations individually was also estimated. If the restorations within a patient correlate significantly, the average variation within patients will be smaller than the variation over patients.

To test for an association between treatment time and secondary outcome, the Spearman rank correlation was used, which allows for non-normal distributions. All computations were done using the statistical package BMDP/Dynamic.

Results

Summary statistics of the distribution of treatment time needed for each type of restoration are given in Table 4. Because of skewed distributions, a Mann-Whitney test was performed to test for different (median) time needed between amalgam and composite restorations when at least five amalgam restorations were involved. All these differences were statistically significant (p<0.002). As is evident from Table 4, these comparisons are in favor of amalgam.
Table 4. Median treatment time (minimum, maximum) for the different types of restorations, in minutes

<table>
<thead>
<tr>
<th>Type of restoration</th>
<th>MO/DO</th>
<th>MOD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Composite</td>
<td>Amalgam</td>
</tr>
<tr>
<td>Premolar</td>
<td>39</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>(30, 60)</td>
<td>(15, 26)</td>
</tr>
<tr>
<td>Molar</td>
<td>52</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>(33, 69)</td>
<td>(18, 30)</td>
</tr>
</tbody>
</table>

* Mann-Whitney p<0.002.

Since none of the 73 restorations was replaced, the conclusion is that there is no significant difference between amalgam and composite restorations with respect to survival within five years after placement. Two out of the ten composite MO/DO restorations in molars had been repaired, the associated exact 95% Confidence Interval was 2.52% to 55.61% (Geigy, 1980). The repairs involved local preparation and the addition of composite, after chipping or appearance of a void.

Before the other secondary outcomes were compared, observer agreement on marginal adaptation and marginal discoloration was evaluated. Although one assessor seemed somewhat more likely than the other to score a '2' than a '3', the assessors did not differ on their perceived distribution of marginal adaptation (McNemar's test, p=0.04; Test for Marginal Homogeneity p=1.0). The 95% confidence interval for kappa (0.62 to 0.83) indicates that the assessments on marginal adaptation are sufficiently reliable.

Despite the training and calibration of the assessors, no sufficient observer agreement could be reached to guarantee a reliable judgment of marginal discoloration: the lower limit of the 95% confidence interval for kappa was 0.49. Dichotomizing discoloration into 'no section discolored' and 'at least one section discolored' showed to be no alternative (lower limit = 0.31). Since the kappa threshold could not be reached, sufficient reliability of assessment of marginal discoloration is not assured. Consequently, marginal discoloration is disregarded.

The distribution of the marginal adaptation of the restorations as scored approximately five years after construction, is depicted in Table 5. All restorations showed some evidence of a crevice along the margin. In general, the differences between marginal adaptation of composite and amalgam restorations are minor. MO/DO restorations in premolars seem to be the exception, with composite restorations performing somewhat better than amalgam with respect to marginal adaptation (Mann-Whitney p=0.04).
There is no evidence that higher costs are associated with higher quality: For neither amalgam nor composite was the association between treatment time and marginal adaptation statistically significant (p>0.05). Since no difference in longevity between amalgam and composite could be established the first five years after replacement of an old Class II amalgam restoration, the estimation of incremental costs per extra restoration year in situ is infeasible. Repairs were done for MO/DO composite restorations in molars only. Except for MO/DO restorations in premolars, no statistically significant differences in the quality of the marginal adaptation were found. So the two materials are more or less equally effective the first five years after placement. In the conceptual framework of economic evaluation, this means that when gathering more data on effectiveness is infeasible, the results of this study are appropriately expressed as a cost-minimization analysis.

Table 5. Marginal adaptation scores for composite and amalgam restorations

<table>
<thead>
<tr>
<th>Premolars</th>
<th>Type of restoration</th>
<th>MOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td>MO/DO</td>
<td>2</td>
</tr>
<tr>
<td>Material</td>
<td>Composite</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Amalgam</td>
<td>48</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Molars</th>
<th>Type of restoration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td>MO/DO</td>
</tr>
<tr>
<td>Material</td>
<td>Composite</td>
</tr>
<tr>
<td></td>
<td>Amalgam</td>
</tr>
</tbody>
</table>

The observers agreed that one amalgam restoration was impossible to score due to limited accessibility.

In a cost-minimization analysis, the treatment option associated with least cost is regarded as the treatment of choice in terms of 'value for money'. The results presented in Table 4 demonstrate that amalgam rerestorations are associated with about half the treatment time required for composite rerestorations. It is reasonable to assume that this time difference reflects a significant cost difference. So amalgam rerestorations are associated with least cost and therefore the treatment of choice.
In this section uncertainty relating to generalizability of results and uncertainty relating to analytical methods are investigated. With regard to generalizability it is important to mention that three composites were used in this study. These three were selected to represent a wide array of possible composites with respect to filler particle size and organic matrix. The filler particle size of Herculite is smaller than that of both Clearfil and Visiomial. The polymer structure of Visiomial is essentially different from those of Herculite and Clearfil. Treatment time did not differ between the types of composites and there is also no evidence for different performance with respect to survival times.

The secondary outcome variable 'need for repair' may differ between composite resins. The two repaired restorations were made from Clearfil Ray Posterior and Visiomial. Maybe between Herculite and amalgam there would be no statistically significant difference with respect to need to repair MO/DO restorations in molars.

Dentists who are nowadays taught to use posterior composite resins during their training may have an advantage over the dentists who were not. The dentists in this study had prior experience with composite resins through their regular construction of Class IV restorations. Nonetheless, in this study there is evidence that proficiency in the use of composite resins for Class II cavities would affect the length of treatment time. The last ten MO/DO composite resin restorations in premolars took 10% to 25% less time than the first ten. The same trend can be found for composite resin MO/DO restorations in molars. After a reduction of the treatment times by 25% for composite resin restorations, the statistical tests for comparing treatment times remained significant in favor of amalgam (both p<0.01).

The current treatment time for composite will be somewhat shorter due to the modified resin restorative procedure. The perceived need for applying glass-ionomer liner or base and making a bevel may have decreased (Wilson, Wilson et al., 1991; Opdam, 1998). In the present study, the resulting less elaborate construction would have resulted in a treatment time reduction of 6 to 10 minutes. After subtracting 6 or 10 minutes from the treatment time for all composite restorations, the treatment time distributions for amalgam and composite resins overlap. The Mann-Whitney tests, however, remained statistically significant (for a treatment time reduction of 10 minutes both p<0.01).

The two above sources of variation combined reduce the treatment times for composite resin restorations significantly. If, for example, the decreased need for applying glassionomer liner or base, results in 8 minutes less treatment time, and increased proficiency results in a reduction of 20%, then both Mann-
Chapter 7

Whitney tests would remain statistically significant in favor of amalgam (for MO/DO restorations in molars: \( p=0.0018 \), for MO/DO restorations in premolars: \( p=0.039 \)).

Uncertainty related to analytical methods encompasses the choice of costs and effects to be included in an evaluation (Briggs, Sculpher et al., 1994). The assumption that treatment time needed in the life cycle of a tooth is a good approximation of the real costs of treatment should also take into account the time needed for removal of the new material. Results from a study in a quasi-clinical setting indicate that the removal of totally bonded and tooth-colored posterior restorations made of composite and ceramic is more technically demanding and time-consuming than the removal of glassionomer and amalgam restorations (Krejci, Lieber et al., 1995). The authors found the mean removal time of composite restorations to be about 1.5 times the mean removal time of amalgam restorations (24.9 respectively 15.2 minutes). So the inclusion of removal time increases the difference in treatment time between composite resins and amalgam restorations.

Overall, although treatment times for composite resins may be reduced substantially the difference in treatment time remains in favor of amalgam.

Discussion

A full 100% five year survival is in its own value a positive result, but does not allow the detection of any difference in longevity. Any differences occurring after this period would, due to discounting, have a limited impact on both costs and effectiveness. With respect to quality of the margin, it was shown how difficult it is to evaluate discoloring. This is a disturbing result considering that general dentists reported marginal discoloring as an important reason for replacement (Mjör and Qvist, 1997; van Noort and Davis, 1993; Mjör and Toffenetti, 1992).

In this study a replacement with composite takes more time than a similar replacement with amalgam. At first sight, these results do not align with an earlier study on new Class I and Class II restorations (Hendriks, Letzel et al., 1985). That study included polishing time for amalgam only. If the composite restorations had been polished too, the treatment time for composite resins would have been longer than for amalgam. In the present study the difference between median polishing time for any type of restoration was less than four minutes. So the polishing of both composite resins and amalgam does not account for the different findings of the two studies.

On the basis of this study one would prefer for efficiency reasons the use of amalgam over composite resins when rerestoring an old amalgam Class II restoration. The choice of a restorative material in dental practice will not exclusively be driven by cost considerations. The choice depends on a large
extent on the demand of patients. Patients may put more emphasis on (alleged) differences in safety or on the esthetics of alternative restorative materials. Including the patients’ perspective would also have an impact on the design of the economic evaluation. In this case one does not only take into account direct costs of dental treatments but also costs borne outside the health care sector. These costs would include patients' out-of-pocket expenses for traveling and over-the-counter drugs, as well as time spent traveling and waiting (Drummond, O'Brien et al., 1997). Treatment effectiveness too, would have to incorporate more than the longevity and quality of the margin, especially since opinions on esthetics and possible adverse health effects, which are hard to quantify, influence the effectiveness as perceived by patients. So for an economic evaluation from the patients’ perspective, in addition to the performance of the restorations, an "oral" health profile which covers the aforementioned topics (Karlsson, Teiwik et al., 1995; Maas, 1996; Slade, 1997) is essential.

In conclusion, this study is not a final answer to the cost-effectiveness of amalgam and composite in replacing old amalgam restorations. However, our tentative conclusion, based on the considerations of both effectiveness and costs is that amalgam is to be preferred over composite when rerestoring Class II amalgam restorations because amalgam rerestorations are associated with the consumption of less resources than composite rerestorations.

Acknowledgments: We would like to thank the two anonymous reviewers and dr. PD Bezemer and dr. L Rekers from the working group “Statistical aspects of MTA” of the Department of Epidemiology and Biostatistics, Vrije University Amsterdam, for their helpful suggestions.

This study was financially supported by Kerr and Cavex Holland/Kuraray.

References


Maas WR. Demands and opportunities for development of self-reported assessments of oral health outcomes. *Journal of Dental Education* 60: 508-513.


