This is an author produced version of a paper published in Journal of Oral Rehabilitation. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the published paper:

URL: https://doi.org/10.1111/joor.12415

Publisher: Wiley

This document has been downloaded from MUEP (http://muep.mah.se).
Turned versus anodized dental implants: a meta-analysis

Running head: Turned vs. anodized implants: meta-analysis

REVIEW

Bruno Ramos Chrcanovic 1*
Tomas Albrektsson 1,2
Ann Wennerberg 1

1 Department of Prosthodontics, Faculty of Odontology, Malmö University, Malmö, Sweden
2 Department of Biomaterials, Göteborg University, Göteborg, Sweden

* Corresponding author:
Bruno Ramos Chrcanovic, Department of Prosthodontics, Faculty of Odontology, Malmö University, Carl Gustafs väg 34, SE-205 06, Malmö, Sweden. bruno.chrcanovic@mah.se; brunochrcanovic@hotmail.com Mobile: +46 725 541 545 Fax: +46 40 6658503
SUMMARY

The aim of this meta-analysis was to test the null hypothesis of no difference in the implant failure rates, marginal bone loss and postoperative infection for patients being rehabilitated by turned versus anodized-surface implants, against the alternative hypothesis of a difference. An electronic search without time or language restrictions was undertaken in November/2015. Eligibility criteria included clinical human studies, either randomized or not. Thirty-eight publications were included. The results suggest a risk ratio of $2.82$ (95%CI 1.95-4.06, $P<0.00001$) for failure of turned implants, when compared to anodized-surface implants. Sensitivity analyses showed similar results when only the studies inserting implants in maxillae or mandibles were pooled. There were no statistically significant effects of turned implants on the marginal bone loss (mean difference-MD 0.02, 95%CI -0.16-0.20; $P=0.82$) in comparison to anodized implants. The results of a meta-regression considering the follow-up period as a covariate suggested an increase of the MD with the increase in the follow-up time (MD increase 0.012 mm/year), however without a statistical significance ($P=0.813$). Due to lack of satisfactory information, meta-analysis for the outcome ‘postoperative infection’ was not performed. The results have to be interpreted with caution due to the presence of several confounding factors in the included studies.

KEYWORDS

Dental implants; turned implants; anodized implants; implant failure rate; marginal bone loss; meta-analysis
INTRODUCTION

Albrektsson et al. (1) proposed six factors that have been generally accepted as especially important for the establishment of a reliable osseointegration, surface quality being one of them. As the implant surface is the first component to interact with the host, several surface modifications have been extensively investigated in the search for improved bone healing (2). Alterations in surface texture and chemistry are modifications used commonly to increase the biological response to implants (3-6). Anodic oxidation, an electrochemical process that increases the TiO2 surface layer and roughness, is one of the methods currently used to modify the surface of a dental implant. In this process, the implant is immersed in a suitable electrolyte and becomes an anode in an electrochemical cell. When a potential is applied, an ionic transport of charge is transferred through the cell, and an electrolytic reaction takes place at the anode, resulting in the growth of an oxide film (7).

Some studies have compared the clinical outcomes of turned and anodized-surface dental implants. However, much of the research in the field is limited by small cohort size and short follow-up periods. In order to address this issue, meta-analyses are important, due to the increase of the sample size of individual trials to reach more precise estimates of the effects of interventions. We conducted a systematic review and meta-analysis to compare the survival rate of dental implants, marginal bone loss (MBL), and postoperative infection of turned and anodized-surface dental implants.

MATERIALS AND METHODS

This study followed the PRISMA Statement guidelines (8). A review protocol does not exist.

Objective

The purpose of the present review was to test the null hypothesis of no difference in the implant failure rates, MBL and postoperative infection for patients being rehabilitated by turned versus anodized-surface implants, against the alternative hypothesis of a difference. The focused question was elaborated by using the PICO format (Participants, Interventions, Comparisons and Outcomes): to compare three outcomes (implant failure rates, MBL, and postoperative infection) of clinical studies.
including patients undergoing implant-prosthetic rehabilitation comparing endosseous implants with turned and anodized surfaces.

**Search strategies**

An electronic search without time or language restrictions was undertaken in November 2015 in the following databases: PubMed/Medline, Web of Science, and the Cochrane Oral Health Group Trials Register. The following terms were used in the search strategy on PubMed/Medline, refined by selecting the term:

{Subject AND Adjective}

{Subject: (dental implant OR oral implant [all fields])}

AND

Adjective: (oxide-coated OR oxidized OR anodized OR anodization OR TiUnite [all fields])

The following terms were used in the search strategy on Web of Science, in all databases:

{Subject AND Adjective}

{Subject: (dental implant OR oral implant [topic])}

AND

Adjective: (oxide-coated OR oxidized OR anodized OR anodization OR TiUnite [topic])

The following terms were used in the search strategy on the Cochrane Oral Health Group Trials Register:

(((dental implant) OR oral implant)) AND (((oxide-coated) OR oxidized) OR anodized) OR anodization) OR TiUnite)


The reference list of the identified studies and the relevant reviews on the subject were scanned for possible additional studies. Moreover, online databases providing information about clinical trials in progress were checked (clinicaltrials.gov; www.centerwatch.com/clinicaltrials; www.clinicalconnection.com).

Inclusion and Exclusion Criteria

Eligibility criteria included clinical human studies, either randomized or not, comparing implant failure rates, MBL and/or postoperative infection in any group of patients receiving turned (machined) and anodized-surface (TiUnite) implants, both from the same implant manufacturer (Nobel Biocare AB, Göteborg, Sweden). Based on the choice of comparing only implants from the same manufacturer, the focus is set on whether the clinical outcome and failure rate in similarly shaped implants, but with different surface characteristics of clinical relevance. For this review, implant failure represents the complete loss of the implant. Exclusion criteria were case reports, technical reports, biomechanical studies, finite element analysis (FEA) studies, animal studies, in vitro studies, and review papers.

Study selection

The titles and abstracts of all reports identified through the electronic searches were read independently by the three authors. For studies appearing to meet the inclusion criteria, or for which there were insufficient data in the title and abstract to make a clear decision, the full report was obtained. Disagreements were resolved by discussion between the authors.

Quality assessment
Quality assessment of the studies was executed according to the Newcastle-Ottawa scale (NOS), which is a quality assessment tool to use when observational studies are also included in systematic reviews (9). The NOS assesses nine items of the studies, divided in three main categories: (I) the selection of the study groups ([a] representativeness of the exposed cohort, which assesses whether the representativeness of the exposed individuals are representative of the average from some general population; [b] selection of external control, which assesses whether the control group was drawn from the same community as the exposed cohort; [c] ascertainment of exposure, which assesses whether the data comes from a secure record, a structure interview, or a written self-report; [d] outcome of interest not present at start), (II) comparability of cohorts (the study controls for the [a] main factor; and the study controls for any [b] additional factor) and (III) the ascertainment of either the exposure or outcome of interest for case-control or cohort studies respectively ([a] assessment of outcome, which assesses whether the data comes from a secure record, a structure interview, or a written self-report; [b] follow-up long enough; [c] adequacy of follow-up, which assesses the follow-up of the exposed and control cohorts to ensure that losses are not related to either the exposure or the outcome). It assigns a maximum of 4 stars for selection, a maximum of 2 stars for comparability, and a maximum of 3 stars for outcome. According to that quality scale, a maximum of 9 stars/points can be given to a study, and this score represents the highest quality, where six or more points were considered high quality.

Concerning the item ‘follow-up long enough’ of the component ‘outcome’, 5 years of follow-up was chosen to be enough for the outcome ‘implant failure’ to occur. To allow the survival and success of implants to be analyzed appropriately, a minimum of 5 years of follow-up is necessary (10).

**Data extraction and meta-analysis**

At least two review authors independently extracted data using specially designed data extraction forms. The data extraction forms were piloted on several papers; these were modified as required before use. Any disagreements were solved by discussion and a third review author was consulted where necessary.

From the studies included in the final analysis, the following data was extracted (when available): year of publication, study design, unicenter or multicenter study, number of patients, patients’ age, follow-up, days of antibiotic prophylaxis, mouth rinse, implant healing period, failed and
placed implants, period of failure (before or after loading), postoperative infection, MBL, implant surface modification, type of prosthetic rehabilitation, jaws receiving implants (maxilla and/or mandible). Contact with authors for possible missing data was performed.

Implant failure and postoperative infection were the dichotomous outcomes measures evaluated. Weighted mean differences were used to construct forest plots of MBL, a continuous outcome. The statistical unit for ‘implant failure’ and ‘MBL’ was the implant, and for ‘postoperative infection’ was the patient. Whenever outcomes of interest were not clearly stated, the data were not used for analysis. The $I^2$ statistic was used to express the percentage of the total variation across studies due to heterogeneity, with 25% corresponding to low heterogeneity, 50% to moderate and 75% to high. The inverse variance method was used for random-effects or fixed-effects model. Where statistically significant ($P < .10$) heterogeneity is detected, a random-effects model was used to assess the significance of treatment effects. Where no statistically significant heterogeneity is found, analysis was performed using a fixed-effects model (11). The estimates of relative effect for dichotomous outcomes were expressed in risk ratio (RR) and in mean difference (MD) in millimeters for continuous outcomes, both with a 95% confidence interval (CI). Only if there were studies with similar comparisons reporting the same outcome measures was meta-analysis to be attempted. In the case where no events (or all events) are observed in both groups the study provides no information about relative probability of the event and is automatically omitted from the meta-analysis. In this (these) case(s), the term ‘not estimable’ is shown under the column of RR of the forest plot table. The software used here automatically checks for problematic zero counts, and adds a fixed value of 0.5 to all cells of study results tables where the problems occur.

In order to explore the possible heterogeneity of effect between studies, a meta-regression was performed in order to verify how a categorical study characteristic is associated with the intervention effects in the meta-analysis, but only when there were at least ten studies available with relevant variables.

A funnel plot (plot of effect size versus standard error) will be drawn. Asymmetry of the funnel plot may indicate publication bias and other biases related to sample size, although the asymmetry may also represent a true relationship between trial size and effect size.
The data were analyzed using the statistical software Review Manager (version 5.3.3, The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen, Denmark, 2014). Meta-regressions (when possible) were performed by using the software OpenMeta[Analyst] (12).

RESULTS

Literature search

The study selection process is summarized in Figure 1. The search strategy resulted in 1086 papers. A number of 312 articles were cited in more than one research of terms (duplicates). The three reviewers independently screened the abstracts for those articles related to the focus question. Of the resulted 774 studies, 736 were excluded for not being related to the topic, resulting in 38 entries. Additional hand-searching of the reference lists of selected studies yielded 8 additional papers. The full-text reports of the remaining 46 articles led to the exclusion of 8 because they did not meet the inclusion criteria (2 studies did not inform of the number of implants failures per group, 2 papers were earlier follow-up of the same study, 2 papers did not present separated numbers of the focused implants for one of the study groups, 1 paper was same study published in another journal, 1 study evaluated only replaced implants). Thus, a total of 38 publications were included in the review.

Description of the Studies

Four randomized clinical trials (13-16), six controlled clinical trials (7, 17-21), five prospective studies (22-26) and twenty-three retrospective analyses (27-49) were included in the meta-analysis. Detailed data of the 38 included studies are listed in Tables S1 and S2. Of the 38 studies comparing the procedures, a total of 43,680 dental implants were turned, with 3545 failures (8.11%), and 23,306 implants had an anodized surface, with 456 failures (1.96%). Nineteen studies (13, 15, 16, 20-22, 24, 25, 28, 32, 34, 35, 39, 41-45, 47) provided information about the MBL separately by implant type, with mean and standard deviation. Only four studies (7, 20, 26, 44) provided information about postoperative infection. However, two (7, 20) of them didn’t have any occurrences and only one study (44) provided information about which type of implants presented infections.

Quality Assessment

N/A
Quality assessment of the studies was executed according to the Newcastle-Ottawa scale (NOS), and the scores are summarized in Table S3. Twenty-three studies were of high quality and fifteen were of moderate quality.

*Meta-analysis and meta-regression*

In this study, a random-effects model was used to evaluate the implant failure in the comparison between the procedures, since statistically heterogeneity was found ($r^2 = 0.66; \chi^2 = 213.31; I^2 = 84\%; P < 0.00001$). The insertion of turned implants statistically affected the implant failure rates in comparison to implants with an anodized surface (RR $2.82$, 95% CI $1.95, 4.06$, $P < 0.00001$; Figure 2).

Since the effect size could differ depending on the insertion of implants in bone areas of different quality, a sensitivity analysis was performed. When only the studies inserting implants in maxillae were pooled, a RR of $2.54$ resulted (95% CI $1.32, 4.89$, $P = 0.005$; heterogeneity: $r^2 = 0.26, \chi^2 = 16.14, I^2 = 69\%, P = 0.006$, random-effects model; Figure 3), statistically affecting the implant failure rates ($P = 0.005$). When only the studies inserting implants in mandibles were pooled, a RR of $2.51$ resulted (95% CI $1.27, 4.97$, $P = 0.008$; heterogeneity: $r^2 = 0.50, \chi^2 = 18.00, I^2 = 44\%, P = 0.06$, random-effects model; Figure 4), also statistically affecting the implant failure rates ($P = 0.008$).

Another sensitivity analysis was performed taking into consideration the different scores for the NOS. When the 23 studies with 6 or more points in the NOS were pooled, a RR of $3.31$ resulted (95% CI $2.18, 5.02$, $P < 0.00001$; heterogeneity: $r^2 = 0.38, \chi^2 = 65.24, I^2 = 69\%, P < 0.00001$, random-effects model), in comparison to a RR of $2.09$ (95% CI $1.40, 3.14$, $P = 0.003$; heterogeneity: $r^2 = 0.26, \chi^2 = 39.34, I^2 = 64\%, P = 0.0003$, random-effects model) when the 15 studies getting until 5 points in the NOS were pooled.

There were no apparent significant effects of turned implants on the MBL (MD $0.02$, 95% CI $-0.16, 0.20$; $P = 0.82$; heterogeneity: random-effects model, $r^2 = 0.18; \chi^2 = 937.19; I^2 = 97\%; P < 0.00001$, Figure 5) in comparison to implants with an anodized surface. The same resulted for sensitivity analyses taking into consideration the different scores for the NOS. When the studies with 6 or more points in the NOS were pooled, the MD was $0.00$ (95% CI $-0.25, 0.25$; $P = 0.97$; heterogeneity: random-effects model, $r^2 = 0.30; \chi^2 = 757.49; I^2 = 97\%; P < 0.00001$), whereas the MD
was 0.10 (95% CI -0.15, 0.35; \( P = 0.42 \); heterogeneity: random-effects model, \( I^2 = 0.07 \); \( \chi^2 = 122.06 \); \( I^2 = 97\% \); \( P < 0.00001 \)).

When a plotting considering the follow-up period as a covariate was performed, it was observed an increase of the MD of MBL with the increase in the follow-up time (\( y = -0.008 + 0.001x \); Figure 6). According to this statistical model, an increase of each year in follow-up time increases the MD in 0.012 mm (12 x 0.001). However, the model was not statistically significant (\( P = 0.813 \)).

Due to lack of enough information, meta-analysis for the outcomes ‘postoperative infection’ was not performed.

**Publication bias**

The funnel plot did not show asymmetry when the studies reporting the outcome ‘implant failure’ were analyzed (Figure 7), indicating possible absence of publication bias.

**DISCUSSION**

The purpose of the present review was to compare the implant failure rates, MBL and postoperative infection between turned and anodized-surface implants. One previous review (50) tried to compare the clinical outcome between these two types of implants. However, as the review only included RCTs, the analysis was hindered by the limited number of included studies, and the authors concluded that there is no evidence for a better clinical outcome of one surface over the other. On the other hand, the results of the present study suggest that the use of turned implants statistically affected the implant failure rates in comparison to implants with an anodized surface.

The higher failure rate of turned implants is hypothesized to be related to the small differences in the osseointegration process. TiUnite implants utilize moderately roughened surfaces characterized by a micro-porous thickened oxide layer (1 to 10 µm thick), which is created through an electrochemical process. This open porous structure with various pits of variable dimensions (1 to 5 µm in diameter) creates a surface designed to allow greater bone-to-implant contact (33). A study (51) demonstrated a difference in bone healing between TiUnite and the turned Brånemark implants, with new bone formation directly on the surface of TiUnite implants, whereas it formed appositionally over osteotomy bone around the turned implants. These findings suggested an improved osseocostruction
process of bone healing around the moderately rough-surface TiUnite implants (33). The fact is that when a biomaterial is inserted into living body, it absorbs proteins before cells adhere to its surface (52). These proteins significantly affect the attachment, adhesion, and spreading of osteoblasts, the cells that form bony tissues (53). For such cells, the implant’s surface-charge influences their reactions to the implant, by affecting the type and amount of proteins attached on its surface (54). The enlarged surfaces, such as the anodized surface, provide better possibilities for microbiomechanical retention due to larger surface and thus more retention for proteins to attach and new bone formation. It is a matter of debate whether these differences in the early osseointegration process between these two surfaces may have significant impact on the long-term outcome of the implants.

The present results suggest that turned implants have a statistically significant higher failure rates in relation to anodized implants regardless whether the implants were placed in maxilla or mandible. The results are not in agreement with the results of Balshe et al. (33), who observed that turned implants performed better than anodized implants in the mandible, while the anodized implants performed better in the maxilla. However, a histological study of bone response between oxidized and turned titanium implants in human jawbone (55) supports the findings of the present meta-analysis concerning the similar outcome in both jaws. In this study, twenty patients received one test (TiUnite) and one control (turned Brånemark implant) micro-implant. After a healing time, the microimplants and the surrounding tissue were removed with a trephine bur and the histomorphometric evaluation demonstrated significantly higher bone-to-implant contact for the oxidized implants, both in the maxilla and in the mandible. It was also suggested that turned implants have a statistically significant higher failure rates in relation to anodized implants regardless whether only studies having or not having high NOS quality scores were pooled together.

Concerning MBL, the beneficial effect of rough implant surfaces on peri-implant bone formation is considered to be based on the changes in microtopography and subsequent alterations of surface energy that result in increased interaction with the adjacent biological environment by adsorption of proteins and blood components which in turn can enhance cell attachment and implant integration (56). However, the results of the present meta-analysis suggesting no apparent significant effects of turned implants on the MBL in comparison to anodized implants challenge this statement and the results of many clinical trials comparing MBL of turned and TiUnite implants (16, 17, 22, 24, 25, 28, 39, 41). A possible reason is the fact that some studies may lack statistical power, given the
small number of patients per group in the clinical trials comparing the techniques. Moreover, it was reported that a surface roughness of more than 2 µm (Sa) is associated with a higher risk of peri-implantitis (57). Rougher implant surfaces are more susceptible to accumulation of bacteria on hard surfaces (58, 59). Bacterial infection, characterized as bacterial colonization and biofilm formation on dental implants, is an important risk factor for peri-implantitis. A roughened surface does not only increase the susceptibility for peri-implantitis, but also reduces the treatment efficacy of the bacteria biofilm (60). The prevalence of peri-implantitis with a TiUnite surface is not higher than turned surface, but once there is a peri-implantitis the progression is increased compared to other surfaces (61). Hence, moderate surface modifications may improve implant therapy in terms of speeding up the treatment, but may be disadvantageous for the patients prone to peri-implantitis (39). It is important to stress that the difference in implant surface is not the only factor playing a role in MBL, but the mechanisms behind such bone loss are most likely multifactorial and may be also be explained by remodeling as part of implant healing, the response to loading, ongoing atrophy after tooth loss, infection, or by other factors (44).

The statistical heterogeneity stands for the variability in the intervention effects being evaluated in the different studies, and is a consequence of clinical or methodological diversity, or both, among the studies. The high level of heterogeneity observed when the outcome ‘implant failure’ was analyzed is not surprising, given the variability of the included studies. For this reason, a random-effects model was also used to incorporate heterogeneity among studies. However, it is important to stress that care must be taken in the interpretation of the chi-squared test, since it has low power in the (common) situation of a meta-analysis when studies have small sample size or are few in number. This means that while a statistically significant result may indicate a problem with heterogeneity, a non-significant result must not be taken as evidence of no heterogeneity (62). Some argue that, since clinical and methodological diversity always occur in a meta-analysis, statistical heterogeneity is inevitable (63). Thus, the test for heterogeneity is irrelevant to the choice of analysis; heterogeneity will always exist whether or not we happen to be able to detect it using a statistical test (62).

Limitations of the present study. The results of the present study have to be interpreted with caution because of its limitations. First of all, several confounding factors may have affected the outcomes and not just the fact that implants had turned and anodized surfaces. The impact of these variables on the implant survival rate, postoperative infection and MBL is difficult to estimate if these
factors are not identified separately between the two different implant types in order to perform a meta-
regression analysis. The lack of control of the confounding factors limited the potential to draw robust
conclusions. Second, most of the included studies had a retrospective design. As all data from a
retrospective study rely on the accuracy of the original examination and documentation, there are
problems were manifested by the gaps in information and incomplete records, because items may
have been excluded in the initial examination or not recorded in the medical chart. In a retrospective
study is difficult to assess the adverse effects of implant surface differences on the prognosis of
implants purely on the basis of implant failure because of the multifactorial genesis of implant failure
(64). Third, some of included studies are characterized by a low level of specificity, i.e. the assessment
of the implant surface as a complicating factor for dental implants was not the main focus of the
investigation. Fourth, much of the research in the field is limited by small cohort size and short follow-
up periods. Short follow-up periods might have led to an underestimation of actual failures, as longer
follow-up periods can lead to an increase in the failure rate, especially if it extended beyond functional
loading, because other prosthetic factors can influence implant failure from that point onward.
However, it is hard to define what it would be considered a short follow-up period to evaluate implant
failures. Furthermore, the quality assessment tool used here has received criticism (65, 66). Among
other issues, there is no clear explanation for the identification of the threshold score for distinguishing
the quality of the studies. However, one has to consider that a perfect quality assessment tool does
not exist. Scales vary considerably in dimensions covered and complexity. Many scales include items
for which there is little evidence that they are related to the internal validity of a trial (67). Still, the NOS
is an indicated quality assessment tool for use on nonrandomized studies included in systematic
reviews. The continued improvement of the scale by the authors of NOS would provide a better validity
assessment of it.

The authors believe that new research efforts should be concentrated in large cohort long-
term studies comparing failure rates and MBL between implants of moderately rough surfaces. The
use of turned implants has been abandoned in several places (46), even though it still has its place in
some countries and is considered as the ‘golden standard’ of comparison between implants of
different surface treatments. It is also valid to incorporate a clear distinction of multiple confounding
factors suggested to have some influence on the implant failure rates, such as smoking (68), bruxism
(69), and the history of periodontal disease (70), as the clinical outcome may vary when other factors are taken into consideration.

**CONCLUSION**

Within the limitations of the existing investigations, the present study suggests that turned implants have a statistically higher probability to fail than anodized-surface implants, regardless whether the implants were placed in maxilla or mandible, or when studies having or not having high quality scores were pooled together. There were no statistically significant effects of turned implants on the MBL in comparison to anodized implants. A comparison of postoperative infection between the implant types was not possible, due to lack of enough information. The reliability and validity of the data collected, the limitations of the quality assessment tool, and the potential for biases and confounding factors are some of the shortcomings of the present study.
ACKNOWLEDGEMENTS

The authors would like to thank Dr. Torsten Jemt for providing missing information about his study.

ETHICAL APPROVAL

Not applicable.

SOURCE OF FUNDING

None.

CONFLICT OF INTERESTS

There are no conflicts of interest.

REFERENCES


N/A
by Dental Implants Used in an Immediate-Provisionalization Protocol: Report on Long-Term

49. Maló P, de Araújo Nobre M, Lopes A, Queridinha B, Ferro A, Gravito I. Axial Implants in
Immediate Function for Partial Rehabilitation in the Maxilla and Mandible: A Retrospective Clinical
Study Evaluating the Long-Term Outcome (Up to 10 Years). Implant Dent. 2015;24:557-564.

50. Esposito M, Ardebili Y, Worthington HV. Interventions for replacing missing teeth: different types


Horbett TA, eds. Proteins at Interfaces: Physicochemical and Biochemical Studies. Washington


54. Hing KA. Bone repair in the twenty-first century: biology, chemistry or engineering? Philos Trans


clinical trial comparing one- and two-stage titanium screw-shaped fixtures with one-stage plasma-

58. Teughels W, Van Assche N, Sliepen I, Quirynen M. Effect of material characteristics and/or


FIGURE CAPTIONS

Figure 1. Study screening process.

Figure 2. Forest plot for the event ‘implant failure’.

Figure 3. Forest plot for the event ‘implant failure’, when only the studies evaluating implants inserted in maxillae only were pooled.

Figure 4. Forest plot for the event ‘implant failure’, when only the studies evaluating implants inserted in mandibles only were pooled.

Figure 5. Forest plot for the event ‘marginal bone loss’.

Figure 6. Scatter plot for the meta-regression with the association between the mean differences (in millimeters) of the marginal bone loss between the two implants (turned vs. anodized) and the follow-up time (in months). Circles indicate individual studies, and the size of the circles indicates the weight of each study.

Figure 7. Funnel plot for the studies reporting the outcome event ‘implant failure’ (RR – risk ratio; SE – standard error).
1081 records identified through database searching

5 additional records identified through other sources

774 records after duplicates removed

736 records excluded

38 records screened

8 records identified through hand searching

46 full-text articles assessed for eligibility

8 full-text articles excluded:
2 did not inform the number of failures per group
2 earlier follow-up
2 did not present separated numbers of the focused implants for one of the study groups
1 repeated
1 evaluated only replaced implants

38 studies included in qualitative synthesis

38 studies included in quantitative synthesis (meta-analysis)

Figure 1. Study screening process.
119x122mm (300 x 300 DPI)
<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Turned Events</th>
<th>Anodized Events</th>
<th>Total</th>
<th>Total Weight</th>
<th>Risk Ratio IV, Random, 95% CI Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vanden Bogueme et al.</td>
<td>4</td>
<td>124</td>
<td>1</td>
<td>111</td>
<td>1.9% 3.58 [0.41, 31.56] 2004</td>
</tr>
<tr>
<td>Attard et al.</td>
<td>3</td>
<td>180</td>
<td>1</td>
<td>70</td>
<td>1.8% 1.17 [0.52, 11.03] 2005</td>
</tr>
<tr>
<td>Adami and Nowzari</td>
<td>0</td>
<td>88</td>
<td>0</td>
<td>58</td>
<td>Not estimable</td>
</tr>
<tr>
<td>Calandrillo and Tomatis</td>
<td>2</td>
<td>11</td>
<td>0</td>
<td>46</td>
<td>1.2% 19.58 [0.1, 381.51] 2005</td>
</tr>
<tr>
<td>Jungner et al.</td>
<td>7</td>
<td>195</td>
<td>0</td>
<td>199</td>
<td>1.3% 15.31 [0.88, 266.18] 2005</td>
</tr>
<tr>
<td>Osterman et al.</td>
<td>0</td>
<td>11</td>
<td>1</td>
<td>232</td>
<td>1.1% 6.47 [0.28, 150.07] 2005</td>
</tr>
<tr>
<td>Remouard and Nisand</td>
<td>4</td>
<td>54</td>
<td>1</td>
<td>42</td>
<td>1.9% 3.11 [0.36, 26.81] 2005</td>
</tr>
<tr>
<td>Fröberg et al.</td>
<td>0</td>
<td>45</td>
<td>0</td>
<td>44</td>
<td>Not estimable</td>
</tr>
<tr>
<td>Watzak et al.</td>
<td>0</td>
<td>64</td>
<td>1</td>
<td>60</td>
<td>1.1% 0.31 [0.01, 7.53] 2006</td>
</tr>
<tr>
<td>Balshi et al.</td>
<td>6</td>
<td>20</td>
<td>2</td>
<td>24</td>
<td>2.8% 3.60 [0.81, 15.91] 2007</td>
</tr>
<tr>
<td>Alsaadi et al.</td>
<td>228</td>
<td>6088</td>
<td>24</td>
<td>606</td>
<td>5.0% 0.95 [0.63, 1.43] 2007</td>
</tr>
<tr>
<td>Malò et al.</td>
<td>12</td>
<td>150</td>
<td>2</td>
<td>283</td>
<td>2.8% 11.32 [2.57, 49.92] 2007</td>
</tr>
<tr>
<td>Alsaadi et al.</td>
<td>93</td>
<td>1223</td>
<td>8</td>
<td>190</td>
<td>4.4% 1.81 [0.89, 3.66] 2008</td>
</tr>
<tr>
<td>Fröberg and Jernt</td>
<td>18</td>
<td>750</td>
<td>0</td>
<td>450</td>
<td>1.3% 22.22 [1.34, 357.81] 2008</td>
</tr>
<tr>
<td>Osterman et al.</td>
<td>3</td>
<td>77</td>
<td>1</td>
<td>180</td>
<td>1.8% 7.01 [0.74, 65.36] 2008</td>
</tr>
<tr>
<td>Balshe et al.</td>
<td>111</td>
<td>2192</td>
<td>85</td>
<td>2425</td>
<td>5.2% 1.45 [1.10, 1.91] 2009</td>
</tr>
<tr>
<td>Eliassson et al.</td>
<td>9</td>
<td>253</td>
<td>7</td>
<td>117</td>
<td>3.9% 0.59 [0.23, 1.56] 2009</td>
</tr>
<tr>
<td>Lidbeck and Henry</td>
<td>3</td>
<td>8</td>
<td>0</td>
<td>25</td>
<td>1.3% 20.22 [1.15, 354.74] 2010</td>
</tr>
<tr>
<td>r-rod aand Jernt</td>
<td>11</td>
<td>111</td>
<td>0</td>
<td>291</td>
<td>1.8% 0.21 [0.06, 0.61] 2010</td>
</tr>
<tr>
<td>Fung et al.</td>
<td>2</td>
<td>22</td>
<td>0</td>
<td>20</td>
<td>1.2% 4.57 [0.23, 89.72] 2011</td>
</tr>
<tr>
<td>Balshi et al.</td>
<td>1</td>
<td>13</td>
<td>6</td>
<td>151</td>
<td>2.0% 1.94 [0.25, 14.86] 2011</td>
</tr>
<tr>
<td>Hatanos et al.</td>
<td>10</td>
<td>143</td>
<td>3</td>
<td>253</td>
<td>3.2% 5.80 [1.65, 21.08] 2011</td>
</tr>
<tr>
<td>Malò and Nobre</td>
<td>9</td>
<td>127</td>
<td>3</td>
<td>120</td>
<td>3.2% 2.83 [0.79, 10.22] 2011</td>
</tr>
<tr>
<td>Delicieux-Ouellet et al.</td>
<td>13</td>
<td>134</td>
<td>5</td>
<td>50</td>
<td>4.0% 1.13 [0.46, 2.88] 2012</td>
</tr>
<tr>
<td>Nicl et al.</td>
<td>1</td>
<td>48</td>
<td>1</td>
<td>48</td>
<td>1.3% 1.00 [0.06, 15.53] 2012</td>
</tr>
<tr>
<td>Van Assche et al.</td>
<td>2</td>
<td>43</td>
<td>0</td>
<td>42</td>
<td>1.2% 4.89 [0.24, 98.89] 2012</td>
</tr>
<tr>
<td>Arnhart et al.</td>
<td>2</td>
<td>52</td>
<td>2</td>
<td>136</td>
<td>2.1% 2.62 [0.38, 18.09] 2013</td>
</tr>
<tr>
<td>Rocchi et al.</td>
<td>8</td>
<td>55</td>
<td>3</td>
<td>66</td>
<td>3.2% 3.20 [0.89, 11.48] 2013</td>
</tr>
<tr>
<td>Delchi et al.</td>
<td>107</td>
<td>719</td>
<td>41</td>
<td>660</td>
<td>5.1% 5.09 [0.53, 50.77] 2010</td>
</tr>
<tr>
<td>Polizi et al.</td>
<td>19</td>
<td>257</td>
<td>4</td>
<td>243</td>
<td>3.7% 4.49 [1.55, 13.01] 2013</td>
</tr>
<tr>
<td>Sayardous et al.</td>
<td>13</td>
<td>144</td>
<td>4</td>
<td>108</td>
<td>3.6% 2.44 [0.82, 7.27] 2013</td>
</tr>
<tr>
<td>Jungner et al.</td>
<td>2</td>
<td>47</td>
<td>1</td>
<td>45</td>
<td>1.7% 1.91 [0.18, 20.39] 2014</td>
</tr>
<tr>
<td>Jungner et al.</td>
<td>7</td>
<td>133</td>
<td>1</td>
<td>154</td>
<td>2.0% 8.11 [1.01, 65.03] 2014</td>
</tr>
<tr>
<td>Malò et al.</td>
<td>10</td>
<td>170</td>
<td>15</td>
<td>424</td>
<td>4.2% 1.62 [0.79, 3.38] 2016</td>
</tr>
<tr>
<td>Malò et al.</td>
<td>89</td>
<td>620</td>
<td>84</td>
<td>3178</td>
<td>5.2% 4.11 [0.38, 5.48] 2015</td>
</tr>
<tr>
<td>Fröberg and Jernt</td>
<td>22</td>
<td>1088</td>
<td>9</td>
<td>750</td>
<td>4.3% 1.69 [0.78, 3.64] 2015</td>
</tr>
<tr>
<td>Malò et al.</td>
<td>5</td>
<td>107</td>
<td>3</td>
<td>374</td>
<td>3.0% 5.83 [1.41, 23.98] 2015</td>
</tr>
<tr>
<td>Jernt et al.</td>
<td>2719</td>
<td>27914</td>
<td>131</td>
<td>10774</td>
<td>5.3% 6.01 [0.73, 5.33] 2015</td>
</tr>
</tbody>
</table>

Total (95% CI): 43680 / 22306 / 100.0% 2.62 / 1.95, 4.06

Total events: 3545 / 456

Heterogeneity: Tau² = 0.66; Chi² = 213.31, df = 35 (P < 0.00001); I² = 84%

Test for overall effect: Z = 5.53 (P < 0.00001)

Figure 2. Forest plot for the event ‘implant failure’. 185x178mm (300 x 300 DPI)
Figure 3. Forest plot for the event 'implant failure', when only the studies evaluating implants inserted in maxillae only were pooled.
54x15mm (300 x 300 DPI)
Figure 4. Forest plot for the event ‘implant failure’, when only the studies evaluating implants inserted in mandibles only were pooled.

78x33mm (300 x 300 DPI)
Figure 5. Forest plot for the event 'marginal bone loss'.

125x78mm (300 x 300 DPI)

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Turned</th>
<th>Anodized</th>
<th>Mean Difference</th>
<th>Mean Difference</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vanden Bogarde et al.</td>
<td>1.6</td>
<td>1.3</td>
<td>0.8</td>
<td>0.80 [0.40, 1.20]</td>
<td>2004</td>
</tr>
<tr>
<td>Watczak et al.</td>
<td>1.42</td>
<td>0.13</td>
<td>63</td>
<td>1.17 [0.13, 60]</td>
<td>2006</td>
</tr>
<tr>
<td>Friberg et al.</td>
<td>0.89</td>
<td>1.4</td>
<td>15</td>
<td>0.89 [1.55, 35]</td>
<td>2006</td>
</tr>
<tr>
<td>Friberg et al.</td>
<td>0.6</td>
<td>0.8</td>
<td>29</td>
<td>0.73 [2.19, 35]</td>
<td>2017</td>
</tr>
<tr>
<td>Friberg et al.</td>
<td>1.2</td>
<td>0.9</td>
<td>106</td>
<td>1.11 [283, 35]</td>
<td>2007</td>
</tr>
<tr>
<td>Norsten et al.</td>
<td>0.5</td>
<td>0.8</td>
<td>74</td>
<td>0.78 [179, 4.2]</td>
<td>2008</td>
</tr>
<tr>
<td>Figl et al.</td>
<td>0.38</td>
<td>0.46</td>
<td>690</td>
<td>0.49 [355, 4.5]</td>
<td>2008</td>
</tr>
<tr>
<td>Elsæren et al.</td>
<td>0.21</td>
<td>0.53</td>
<td>106</td>
<td>0.54 [105, 4.2]</td>
<td>2009</td>
</tr>
<tr>
<td>Liddelol and Henry (left)</td>
<td>2.12</td>
<td>0.66</td>
<td>5</td>
<td>0.87 [259, 2.9]</td>
<td>2010</td>
</tr>
<tr>
<td>Liddelol and Henry (right)</td>
<td>2.3</td>
<td>0.74</td>
<td>5</td>
<td>0.01 [259, 2.7]</td>
<td>2010</td>
</tr>
<tr>
<td>Fung et al. (2y)</td>
<td>0.26</td>
<td>0.44</td>
<td>20</td>
<td>0.41 [20, 3.7]</td>
<td>2011</td>
</tr>
<tr>
<td>Fung et al. (1y)</td>
<td>1.06</td>
<td>0.01</td>
<td>20</td>
<td>0.02 [20, 3.7]</td>
<td>2011</td>
</tr>
<tr>
<td>Van Assche et al. (odontol.)</td>
<td>1.8</td>
<td>0.6</td>
<td>2</td>
<td>0.88 [21, 3.6]</td>
<td>2012</td>
</tr>
<tr>
<td>Van Assche et al. (gum)</td>
<td>0.06</td>
<td>0.09</td>
<td>23</td>
<td>0.36 [31, 9.4%]</td>
<td>2013</td>
</tr>
<tr>
<td>Rocci et al. (1y)</td>
<td>1.0</td>
<td>0.9</td>
<td>47</td>
<td>0.9 [63, 4.0]</td>
<td>2013</td>
</tr>
<tr>
<td>Rocci et al. (8y)</td>
<td>1.7</td>
<td>2.3</td>
<td>39</td>
<td>1.41 [51, 2.1]</td>
<td>2013</td>
</tr>
<tr>
<td>Sayardoust et al. (nonsmokers)</td>
<td>0.94</td>
<td>0.14</td>
<td>66</td>
<td>1.26 [52, 4.5]</td>
<td>2013</td>
</tr>
<tr>
<td>Sayardoust et al. (smokers)</td>
<td>1.54</td>
<td>0.21</td>
<td>78</td>
<td>1.16 [56, 4.5]</td>
<td>2013</td>
</tr>
<tr>
<td>Arnoux et al.</td>
<td>2.42</td>
<td>0.34</td>
<td>50</td>
<td>1.53 [134, 4.6]</td>
<td>2013</td>
</tr>
<tr>
<td>Pellegrini et al.</td>
<td>1.96</td>
<td>0.75</td>
<td>31</td>
<td>1.25 [31, 2.0]</td>
<td>2013</td>
</tr>
<tr>
<td>Jungner et al.</td>
<td>1.8</td>
<td>0.8</td>
<td>126</td>
<td>2</td>
<td>193</td>
</tr>
<tr>
<td>Jungner et al.</td>
<td>1.4</td>
<td>0.7</td>
<td>45</td>
<td>1.77 [44, 4.1]</td>
<td>2014</td>
</tr>
<tr>
<td>Friberg et al.</td>
<td>0.5</td>
<td>0.59</td>
<td>492</td>
<td>0.81 [238, 4.4]</td>
<td>2015</td>
</tr>
<tr>
<td>Malo et al. (5y)</td>
<td>1.1</td>
<td>0.7</td>
<td>100</td>
<td>1.06 [415, 4.4]</td>
<td>2015</td>
</tr>
<tr>
<td>Malo et al. (1y)</td>
<td>0.39</td>
<td>0.8</td>
<td>41</td>
<td>1.23 [373, 4.4]</td>
<td>2015</td>
</tr>
<tr>
<td>Malo et al. (10y)</td>
<td>1.38</td>
<td>0.09</td>
<td>93</td>
<td>2.12 [154, 2.2]</td>
<td>2015</td>
</tr>
</tbody>
</table>

Total (95% CI): 2573 3022 100.0% 0.02 [0.16, 0.20]

Heterogeneity: Tau² = 0.18; Chi² = 937.19, df = 26 (P < 0.00001); I² = 97%
Test for overall effect: Z = 0.23 (P = 0.82)
Figure 6. Scatter plot for the meta-regression with the association between the mean differences (in millimeters) of the marginal bone loss between the two implants (turned vs. anodized) and the follow-up time (in months). Circles indicate individual studies, and the size of the circles indicates the weight of each study.

108x56mm (300 x 300 DPI)
Figure 7. Funnel plot for the studies reporting the outcome event ‘implant failure’ (RR – risk ratio; SE – standard error).

65x42mm (300 x 300 DPI)