Biodegradable versus titanium osteosynthesis in maxillofacial traumatology
Gareb, B; van Bakelen, N B; Dijkstra, P U; Vissink, A; Bos, R R M; van Minnen, B

Published in: International Journal of Oral and Maxillofacial Surgery

DOI: 10.1016/j.ijom.2019.11.009

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Final author's version (accepted by publisher, after peer review)

Publication date: 2020

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, 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), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 02-05-2021
Biodegradable versus titanium osteosyntheses in maxillofacial traumatology: a systematic review
with meta-analysis and trial sequential analysis

Barzi Gareb, Nico B. van Bakelen, Pieter U. Dijkstra, Arjan Vissink, Ruud R.M. Bos, Baucke van Minnen

Short title
Biodegradable vs. titanium fixation

Institution
Department of Oral and Maxillofacial Surgery
University Medical Center Groningen
University of Groningen
Hanzeplein 1, 9713 GZ Groningen, The Netherlands
P.O. Box 30001, 9700 RB Groningen, The Netherlands

Corresponding author
Barzi Gareb
P.O. Box 30001, 9700 RB Groningen, The Netherlands
Telephone number: +31 503611054
Fax number: +31 503612831
E-mail address: b.gareb@umcg.nl

Sources of support
None.

Key words
Biodegradable; titanium; fixation; craniofacial ; trauma; osteosynthesis; plates
ABSTRACT

Titanium osteosynthesis is currently the fixation system of choice in maxillofacial traumatology. Biodegradable osteosynthesis systems have the ability to degrade in the human body. The aim of this study was to conduct a systematic review, with meta- and trial sequential analyses, to assess the efficacy and morbidity of biodegradable versus titanium osteosyntheses after maxillofacial trauma.

MEDLINE, EMBASE, and CENTRAL were searched for randomized controlled trials, and prospective and retrospective controlled studies. Five time periods were studied: perioperative, short-term (0-4 weeks), intermediate (6-12 weeks), long-term (>12 weeks) and overall follow-up. After screening 3542 records, 24 were included. All had a high risk of performance and detection bias due to the nature of the interventions. Meta-analysis showed no differences in efficacy and morbidity between biodegradable and titanium osteosyntheses. Risk of perioperative screw breakage was significantly higher (RR 17.13, 95% CI: 2.19;34.18) and the symptomatic plate removal rate lower in the biodegradable group (RR 0.11; 95% CI 0.02;0.57), which was confirmed by the trial sequential analysis. The quality of evidence ranged from very low to moderate. Based on both narrative review and meta-analyses, current evidence shows that biodegradable osteosyntheses are a viable alternative to titanium osteosyntheses when applied in the treatment of maxillofacial trauma with similar efficacy but significantly lower symptomatic plate removal rates. Perioperative screw breakage occurred significantly more often in the biodegradable compared to the titanium group.
INTRODUCTION

Titanium osteosynthesis systems are considered the gold standard in maxillofacial fracture treatment and orthognathic surgery. Titanium plates and screws combine excellent mechanical and handling properties, providing adequate bone stability. The disadvantages of titanium osteosyntheses include palpability, sensitivity to temperature changes, stress shielding of the underlying bone, growth restrictions, interference with radiographic imaging and radiotherapy, titanium particles in the soft tissue and regional lymph nodes, and possible mutagenic effects. As a consequence, titanium plates and screws are removed in a second operation in 0-33% of the cases with the associated burdens and costs.

Currently, the most commonly used biodegradable osteosynthesis systems are made of resorbable polymers (e.g., poly-DL-lactic acid), whose properties might eliminate the need to remove implants in a second operation, thereby avoiding the accompanied additional risks, costs, and burdens of a second operation. Additionally, the other disadvantages associated with titanium osteosynthesis are avoided. The limitations of biodegradable osteosynthesis systems include less favourable mechanical properties, which could potentially lead to mobility or malunion of bone segments, and possible adverse tissue reactions. Biodegradable implants have to be removed in 0-17% of the cases.

A systematic review focusing on the efficacy and safety of these interventions in maxillofacial traumatology was published in 2009, but could not include any studies because none met the inclusion criteria. It was concluded that there was insufficient evidence to support or refute the use of biodegradable osteosynthesis. Since then, many studies comparing biodegradable versus titanium osteosyntheses have been published, but the results of these solitary studies remain controversial. Our randomized controlled trial showed an unexpected higher symptomatic plate removal rate in the biodegradable compared to the titanium group after trauma and orthognathic surgery. To place these results in the literature context, we looked for systematic reviews addressing efficacy and morbidity of these interventions. The most recent systematic review...
comparing both systems in maxillofacial surgery was published in 2013\textsuperscript{[15]}. However, it only focused on complications and failed to account for clinical or methodological heterogeneity. Therefore, there is still a need for a systematic review that adequately assesses the efficacy and safety of biodegradable versus titanium systems in trauma patients, including all the relevant endpoints for clinicians, and which takes the methodological heterogeneity of the studies into account, thereby enabling well informed and evidence based decisions.

The aim of this study was to conduct a systematic review, with meta- and trial sequential analyses, of randomized controlled trials, prospective controlled cohort studies, and retrospective controlled cohort studies examining the efficacy (i.e., bone healing and occlusion) and morbidity of biodegradable (i.e., composed of (co-)polymers) versus titanium osteosyntheses in patients with maxillofacial fractures.

\textbf{MATERIAL AND METHODS}

This systematic review and meta-analysis was conducted following the recommendations of the \textit{Cochrane Handbook for Systematic Reviews of Interventions, Risk Of Bias In Systematic Reviews tool (ROBIS)} and \textit{A Measurement Tool to Assess Systematic Reviews 2 (AMSTAR 2)}, and is reported according to the \textit{Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)} statement to ensure quality and completeness\textsuperscript{[16-19]}. This study’s protocol was registered in PROSPERO prior to the systematic literature search (registration number CRD42018086477).

\textit{Study identification}

A systematic literature search of three electronic databases (MEDLINE (1964-2019), EMBASE (1947-2019), and the Cochrane Central Register of Controlled Trials (CENTRAL; inception to 2019) was conducted. The sensitive search strategy consisted of medical subject heading terms and free-text words (Table S1). The search strategy also included orthognathic populations as some studies include both populations in a single study. Data of trauma patients were derived from the authors of those
studies and were included while data of orthognathic patients were excluded. The complete search
was performed in January 2018 and was updated on 20 April, 2019. Additionally, the reference lists
of the included studies and leading oral and maxillofacial journals were screened for relevant studies
and maxillofacial surgery experts in biodegradable and titanium osteosynthesis (RRMB and NBvB)
were asked if any relevant studies were missing which should have been included in this review. No
language or period restrictions were applied.

Study selection
The inclusion criteria were formulated using the PICOS format. The population (P) included all the
patients who had been treated for maxillofacial fractures, i.e., Le Fort I, Le Fort II, Le Fort III, cranial,
zygomaticomaxillary complex and mandibular fractures. The intervention group (I) was treated
surgically with biodegradable fixation (i.e., plates and/or screws/pins) that consisted of (co-
polymers. The control group (C) received surgical treatment with titanium fixation (i.e., plates
and/or screws). The primary outcomes (O) were efficacy of the fixation method, i.e., adequate bone
healing with the absence of malunion of bone segments, clinical mobility of bone segments, and
objective and subjective malocclusion. Secondary outcomes were related to morbidity, i.e.,
symptomatic plate removal rate (i.e., routinely removed asymptomatic plates were excluded), pain,
analgesia usage, maximal mouth opening (MMO), mandibular function impairment questionnaire
(MFIQ; lower score equals better function), temporomandibular joint dysfunction (TMJ-dysfunction),
infection, swelling, wound dehiscence, plate exposure, palpability of plates and/or screws, the
patient’s satisfaction with the performed surgery, and revision surgery (e.g., abscess incision and
drainage; plate removal was excluded). Additionally, the handling of the osteosynthesis systems by
the surgeons, plate and screw breakage, and total costs (i.e., direct and indirect costs) of both groups
were evaluated. The included study types (S) were randomized controlled trials (RCTs), prospective
studies with a control group, and retrospective studies with a control group. The RCT is the highest
quality of evidence of an original manuscript, while the latter two designs are useful for adverse
events assessment. The follow-up (FU) of each corresponding endpoint is described below (see Data collection).

Exclusion criteria consisted of patients with syndromic disorder(s), patients with cleft lip or palate, multiple publications of the same study and endpoints, case reports, case series with fewer than 10 cases, experts’ opinions, letters to the editor, review articles, and conference abstracts.

Two reviewers (BG and NBvB) independently assessed the titles and abstracts for eligibility for inclusion. If the title and abstract provided insufficient information or in case of any doubt, they were included for full text assessment. The full text articles of included titles and abstracts were independently assessed by the same two reviewers for final inclusion using the above mentioned in- and exclusion criteria. Any disagreement was resolved by a discussion. If no consensus was found, a third reviewer (PUD) was asked to give a final decision.

After each selection stage, the inter-observer agreement was expressed as Cohen’s kappa and percentage of agreement. Studies written in languages that the observers were not competent in were translated by researchers fluent in both that language and English. Subsequently, these translated studies underwent the same review process.

Assessment of methodological quality

The risk of bias of all the included studies was independently assessed by two reviewers (BG and NBvB). Trials performed by the author’s research group were assessed by two independent researchers not involved in those studies (PUD and SJvdG; see acknowledgement) to avoid conflict of interests.

Randomized controlled trials were assessed using the The Cochrane Collaboration’s tool for assessing risk of bias\textsuperscript{20}, including 7 domains: sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective outcome reporting and ‘other issues’. The domains were graded low risk, unclear risk or high risk of bias.
The nonrandomized studies’ risk of bias was assessed using The Methodological Index for Non-Randomized Studies (MINORS)\textsuperscript{21}. The MINORS is a valid and reliable instrument for quality assessment\textsuperscript{21}. It includes 8 items which are applicable to all nonrandomized studies, and an additional 4 applicable to comparative studies. Each item was scored either 0 (not reported), 1 (reported but inadequate), or 2 (reported and adequate).

The quality of the body of evidence for each outcome was graded by two independent reviewers (BG and NBvB) as high, moderate, low, or very low quality using the Grades of Recommendation, Assessment, Development and Evaluation Working Group system (GRADE system). The grades can be increased or decreased based on the underlying methodology depending on the presence of certain factors (e.g., downgrading studies with a high risk of bias)\textsuperscript{22}.

**Data collection**

The data was extracted using a standardised, pre-defined form. Two reviewers (BG and NBvB) extracted data from a sample (10%) of eligible studies. If an agreement of ≥80% was achieved, the remainder of the data was extracted by one reviewer (BG). The collected data included: first author and year of publication, country in which the study was conducted, study design, number of patients, gender, age, tobacco and alcohol usage, surgical procedures, types of osteosynthesis systems used, intra-operative switching to another osteosynthesis system, osteosynthesis principle, duration of maxillomandibular fixation (MMF), duration of FU, and conflict of interests. The endpoints were collated for 5 time periods: perioperative, short-term FU (i.e., 0-4 weeks; soft tissue healing), intermediate FU (i.e., 6-12 weeks; bone healing), long-term FU (i.e., >12 weeks; degradation effects), and overall FU (i.e., the endpoints of the longest FU; Table S2).

If the relevant data could not be extracted, the authors of the studies were contacted by email from May - November 2018 and April – July 2019. Data were not included in the analyses if the authors could not provide the relevant data or did not respond despite a minimum of three email attempts.
Statistical analysis

The inter-observer agreement was calculated using IBM SPSS Statistics 23 (SPSS, Chicago, IL, USA).

Regarding binary variables, the events and totals were used to calculate the risk ratio (RR) and 95% confidence intervals (CI). The standardised mean difference (SMD), with 95% CI, was calculated for continuous variables. Statistical heterogeneity was regarded substantial if \( I^2 > 50\% \). The meta-analysis was performed in R-met{a}23, version 3.5.3, using a random-effects model because of clinical heterogeneity (e.g., different polymer compositions).

Separate analyses were conducted for the study designs. A summary effect estimate was calculated if ≥2 studies with the same study design could be pooled. Also, a subgroup analysis of low risk versus high risk bias RCTs was performed as well as subgroup analyses of the primary endpoints and plate removal rate of paediatric patients (<16 years) versus adults, and mandibular versus other fractures. Plate removal rate was also analysed according to the FU of the included studies, i.e., ≤1 year FU and >1 year FU. A narrative synthesis was performed if only a single study per study design or subgroup was available.

Since a conventional meta-analysis excludes studies with zero events in both treatment groups, a sensitivity analysis was performed, including those studies with a reciprocal continuity correction of the opposite arm24. A meta-regression analysis with a random-effects model evaluated the effect of the study design and items of methodological quality on each primary endpoint and plate removal. Reporting bias was assessed through funnel plots if >10 studies were available per endpoint and study design, and did not have clinical heterogeneity16. Funnel plots with ≤10 studies are underpowered and the presence of clinical or statistical heterogeneity results in inconclusive funnel plots16,25–27. P<0.05 was considered statistically significant. The meta-regression was conducted using Comprehensive Meta-Analysis, version 3 (Biostat, Englewood, NJ, USA).

As traditional meta-analyses are prone to type-I errors (i.e., false positive findings) due to random error and repeated significance testing after each additional trial is published28,29, trial
sequential analyses (TSA), including RCTs, were performed for each endpoint. TSA reduces the risk of type-I errors by combining information size estimations with trial sequential monitoring boundaries and provides information on how many patients are required in the meta-analysis to sufficiently support the conclusions (i.e., equivalent to a sample size calculation in RCTs). An explanation of TSA, with an example and the interpretation of the data, is shown in Figure S1. The TSA, which included the random-effects (DerSimonian-Laird) model based on the observed relative risk reduction (RRR) and diversity ($D^2$) of RCTs, and an overall type I error ($\alpha$) of 0.05 and a type II error ($\beta$) of 0.20, was performed using Trial Sequential Analysis Viewer, version 0.9.5.10 beta (Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet).

RESULTS

Study identification and selection

The search resulted in 5479 potentially eligible papers. After excluding duplicates, 3542 papers were screened by title and abstract (Figure 1). The percentage of agreement and kappa were 99% and 0.91, respectively. The full text manuscripts of the remaining 80 papers were screened for inclusion. Fifty-six studies were excluded due to not fulfilling the inclusion criteria (n=47), fulfilling the exclusion criteria (n=6), providing insufficient details (n=2), or due to including the same study population and endpoints with a shorter FU (n=1) (Table S3). The percentages of agreement and kappa were 100% and 1.0, respectively. The remaining 24 publications were included in the qualitative synthesis of this review, and 21 of them were included in the quantitative synthesis. There was no need to consult the third reviewer in any phase of the identification and selection of a study.

Assessment of methodological quality

The included studies consisted of 7 publications of RCTs, of which 4 were publications of a single RCT, each with a different FU, 4 prospective cohort studies, and 13 retrospective cohort studies. Low risk of bias was observed in the ‘random sequence generation’
domain in all but one of the included RCTs (Table 1). High risk of performance and detection bias was observed in all the included RCTs. ‘Other sources of bias’ were assessed as high risk in four publications of a single RCT due to the fact that, whenever the surgeon deemed it necessary, the surgeon chose to switch perioperatively from biodegradable to titanium systems. As none of the included RCTs were assessed as low risk of bias, no subgroup analyses could be performed between high and low risk of bias.

None of the cohort studies had undertaken an adequate unbiased assessment of the study endpoints (Table 1). All of them had an adequate control group as this was one of the inclusion criteria. Seventy-five percent of the included studies had adequate contemporary groups. Two studies declared funding from research programmes and one from the armed forces. Six studies did not mention funding or conflict of interest. All the remaining studies declared no funding or conflict of interest.

Patient characteristics

The number of patients in the studies ranged from 12 to 1122, resulting in a total of 2450 patients (Table S4). Of these, 1639 patients received titanium and 811 patients received biodegradable osteosynthesis systems. The majority of patients were male. Four studies just had male patients in the biodegradable group. Ages ranged from 4 to 83 years. Two studies only included paediatric patients. The most common types of fractures were mandibular, zygomatic and maxillary fractures. Ten studies solely included patients with mandibular fractures, while six studies included patients with only zygomatic fractures. The remaining studies included various types of fractures (e.g., Le Fort or orbital fractures). Comminuted fractures were excluded in 16 studies, while two studies did not exclude these type of fractures. The remaining six studies did not report specific in- or exclusion criteria regarding comminuted fractures. Four studies included both orthognathic and trauma...
patients, but only the trauma patients’ data have been included in this review\textsuperscript{2,33,35,36}. None of the included studies reported information regarding tobacco or alcohol usage by the patients.

Procedural characteristics

The procedural characteristics of the included studies are presented in Table S4. In one study, the procedure was endoscopically assisted\textsuperscript{8}. The most commonly used titanium osteosynthesis systems were manufactured by KLS Martin\textsuperscript{2,11,33,35,36,41}, Synthes\textsuperscript{14,37,40,42,47}, and Stryker\textsuperscript{13,43}. Twelve studies reported details regarding the size of the titanium plates and the screws\textsuperscript{2,8,40,43,47,11,13,14,33,35–38}. The screw diameter ranged from 1.3 to 2.0 mm with corresponding plates, depending on the location of the fracture.

The most frequently used biodegradable osteosynthesis systems were the Inion CPS (79/15/6 poly-L-lactic acid (PLLA)/poly-DL-lactic acid (PDLLA)/trimethylene carbonate)\textsuperscript{2,11,14,33,35,36,44,46,47} and the BioSorb FX (self-reinforced 70/30 PLLA/PDLLA)\textsuperscript{38,41,42,46,51} (Table S4). The screw diameters ranged from 1.5 to 2.5 mm, being generally larger compared to the titanium systems for similar fracture types. Five articles reported intra-operative switches from a biodegradable to a titanium osteosynthesis system\textsuperscript{2,14,33,35,36}. Of these, one RCT\textsuperscript{14} reported 1 intra-operative switch (5%). The other 4 articles were publications of the same RCT with different FUs\textsuperscript{2,33,35,36} and reported 4 intra-operative switches (40%) in the trauma patients. The main reason for switching material was inadequate fixation due to non-grip screws or inadequate stability of the bone segments after fixation of the osteosynthesis plates\textsuperscript{52}.

Nine studies followed Champy’s principle\textsuperscript{2,11,14,33,35,36,42,46,47} and one the Association for Osteosynthesis/Association for the Study of Internal Fixation (AO/ASIF) principle\textsuperscript{40} for osteosynthesis of mandibular fractures. Six studies did not report the osteosynthesis principle\textsuperscript{8,13,34,38,48,49}. MMF was used in 14 studies, of which 5 studies used soft guiding elastics in both groups\textsuperscript{2,8,33,35,36}, 3 studies used rigid MMF in both groups\textsuperscript{34,42,46}, 2 studies just used MMF in the biodegradable group\textsuperscript{13,14}, and 3
studies only used MMF whenever this was deemed necessary\(^{37,40,49}\), although no details regarding this clinical decision were reported (Table S4).

**Primary endpoints**

All the pooled endpoints are reported as RR or SMD (95% CI), with the quality of the evidence. A total of 16 studies reported data regarding malunion (Table S5)\(^{8,11,43–47,49,13,33,34,37,38,40–42}\). In 14 of these studies, no malunion was found in either the titanium or the biodegradable groups. Malunion, assessed after 6-12 weeks FU, was present in two retrospective studies and pooling of the data showed no significant differences between both groups (RR 0.93 (0.15;5.75), very low quality, Figure 2A).

The mobility of bone segments was assessed in 5 of the studies after 6-12 weeks FU\(^{14,33,34,38,49}\). Two studies reported no mobility of bone segments\(^{33,49}\). One prospective study assessed that 4% and 13% of the patients had mobile bone segments after biodegradable and titanium osteosyntheses, respectively\(^{38}\). Data derived from two RCTs showed no significant differences between both groups (RR 2.11 (0.32;13.79), very low quality, Figure 2B). No subgroup analysis could be performed.

Malocclusion within 4 weeks FU was assessed in 7 studies\(^{11,13,14,38,44,47,49}\). Three of them reported zero events in both groups\(^{13,38,44}\). One RCT found similar rates of short-term objective malocclusion in both groups (24%)\(^{14}\). One prospective study reported objective malocclusion in 41% and 21% of the cases in the biodegradable and titanium group, respectively\(^{11}\). Data derived from two retrospective studies showed no significant difference in objective malocclusion between both groups (RR 0.51 (0.06;4.68), very low quality, Figure 2C). Both of these two retrospective studies only included patients with mandibular fractures. Subgroup analysis between paediatric patients and adults showed no significant difference in the estimate between both subgroups (adults: RR 0.91 (0.29;2.83); paediatric: RR 1.83 (0.81;4.11), very low quality, Figure S2).
Eight studies documented malocclusion after 6-12 weeks FU\(^{11,13,14,33,34,38,47,48}\). Three of these studies reported no objective malocclusion in both groups\(^{13,38,47}\). Pooling of the data from the RCTs showed no significant differences between both groups (RR 1.01 (0.21;4.81), very low quality, Figure 2D). One prospective study mentioned 3% and 7% of the patients had objective malocclusion\(^{11}\), while one retrospective study found subjective malocclusion in 17% and 10% of the cases in the biodegradable and titanium groups, respectively\(^{48}\). No subgroup analysis could be performed.

Six studies assessed malocclusion after >12 weeks FU\(^{2,11,13,35,38,41}\). One RCT reported one case (13%) of objective malocclusion in the titanium group after 1-year of FU\(^{35}\). Another RCT with >5 years FU reported two cases (50%) of subjective malocclusion\(^{2}\) (Table S5). Both these RCTs included the same study population with different FU moments. No subgroup analysis could be performed.

**Secondary endpoints**

Focusing on perioperative endpoints, the occurrence of plate breakage ranged from 0 to 6% and 0 to 2% plates in the biodegradable and titanium groups, respectively (Table S5). Breakage of screws occurred in 0-7% of the biodegradable screws, while only one study reported a single broken titanium screw\(^{48}\). The RCTs showed that biodegradable screws broke more often compared to titanium screws (RR 17.13 (2.19;134.18), moderate quality) while the retrospective studies showed no significant difference between both groups (Figure S3). The mean operative time in the biodegradable and titanium groups ranged between 119-169 and 94-127 minutes, respectively. Data derived from the retrospective studies did not result in a significant difference in operation time between both groups (SMD 0.72 (-0.17;1.61), very low quality, Figure S4). Plate and screw handling, as assessed by surgeons, was only reported in one RCT and was similar for both groups\(^{33}\).

Infection within 4 weeks FU occurred in 0-8% and 0-10% in the biodegradable and titanium groups, respectively, and did not differ significantly between both groups in all the study designs (RCTs: RR 0.26 (0.03;2.26), very low quality, Figure 3A). Short-term swelling was assessed in one RCT\(^{37}\), one prospective study \(^{11}\), and two retrospective studies\(^{41,44}\). One of the retrospective studies
reported swelling in all the included patients after 1 week FU. Therefore, it was not possible to pool this study’s data. Abscess formation at short-term FU was assessed in one study and was not present in either group. Pain within 4 weeks FU ranged from 10-71% in the biodegradable group, while 0-65% of the patients treated with titanium presented with pain. No study reported analgesic usage. MMO was assessed in three studies. One study reported a similar postoperative MMO in both groups while another study reported a higher postoperative MMO in the biodegradable group. One study only gave bar graphs and could not provide numbers for the data synthesis. Dehiscence ranged between 0-37% and 0-38% in the biodegradable and titanium groups, respectively. The RCTs and retrospective studies did not show statistical differences between both groups (RCTs: RR 1.68 (0.56;5.00), very low quality; Figure S5). Finally, plate exposure after short-term FU did not differ significantly on pooling the retrospective studies’ data (RR 0.79 (0.23;2.71), very low quality, Figure S6).

Secondary endpoint data from 6-12 weeks FU were scarce (Table S5). Pain was reported in two RCTs but the studies measured pain differently. MMO was only presented as bar graphs in one study, while another study reported similar postoperative MMOs in both groups. TMJ-dysfunction was assessed in two studies and occurred in 7-8% and 7-16% of the patients after biodegradable and titanium osteosynthesis, respectively.

At long-term FU, the presence of pain was scarce in both groups (Table S5). Pooling of the retrospective data did not result in significant differences between both groups (RR 0.40 (0.10;1.68), very low quality, Figure S7). TMJ-dysfunction was assessed in one study with a FU of 1 year. MFIQ was assessed in two publications of one RCT. The MFIQ was better after >5 years FU in the biodegradable compared to the titanium group (17 (interquartile range 17-17) and 35 (21-41), respectively). Three retrospective studies reported abscess formation after 1-year and 2-years FU. No significant difference between both treatment groups was found (RR 2.37 (0.42;13.23), very low quality, Figure 3B). Long-term swelling assessment was generally scarce. One RCT with a FU >5 years reported 20% (1/5) and 25% (1/4) of cases with swelling in the biodegradable and titanium
groups, respectively\textsuperscript{2}. The retrospective studies showed no significant differences between both groups regarding long-term swelling (RR 4.55 (0.78;26.68), very low quality; Figure 3C). Palpability of plates and screws after long-term FU occurred only in the titanium group, but did not differ between both groups based on the data derived from the retrospective studies (RR 0.30 (0.07;1.37), very low quality, Figure 3D). Both groups' patients were similarly satisfied with the result after 1-year (prosp. CS: SMD -0.20 (-0.92;0.52), very low quality, Figure 3E;\textsuperscript{13,39}) and >5-years FU\textsuperscript{2}.

Symptomatic titanium and biodegradable plate removal rates ranged from 0-39\% and 0-17\%, respectively (Table S5). The FU ranged from 8 weeks to >5 years (Table S4). The main reason for plate removal was chronic infection or disturbed wound healing. The data of one study was not included in the analysis as the authors could not provide the symptomatic plate removal rates and all the titanium plates were removed after 6-8 months due to possible growth disturbances\textsuperscript{48}. Although the RCTs data showed a significant difference in plate removal rate in favour of the biodegradable group (RR 0.11 (0.02;0.57), moderate quality), the prospective and retrospective studies did not demonstrate any significant differences (Figure 3F). Subgroup analyses showed that the symptomatic plate removal rate did not differ significantly between the paediatric titanium and biodegradable groups (RR 1.11 (0.36;3.45). However, all the titanium plates were eventually removed from the paediatric patients due to possible growth disturbances, while only symptomatic biodegradable plates were removed in both studies which included paediatric patients. In adult patients, the symptomatic plate removal rate was significantly lower in the biodegradable group (RR 0.33 (0.13;0.84), Figure S8). Subgroup analyses of plate removal rates comparing mandibular versus other fractures showed no differences (mandibular fractures: RR 0.41 (0.13;1.34); other fractures: RR 0.56 (0.11;2.96), Figure S9). Comparing plate removal rates between ≤1 year and >1 year FU did not display any significant differences between different FU and treatment groups (Figure S10).

One RCT assessed total costs (i.e., direct and indirect costs) after 2 years FU, and found mean costs of 6137 ± 2980 and 8128 ± 5453 euros after biodegradable and titanium osteosynthesis, respectively\textsuperscript{36}. The higher total costs in the titanium group was mainly due to a second operation for
symptomatic plate removal. Finally, revision surgery (i.e., no plate removal) was performed in 0-8% and 0-7% of the patients after biodegradable and titanium osteosynthesis, respectively (retrospective studies: RR 1.16 (0.33;4.06), very low quality, Figure S11). The FU ranged from 8 weeks up to 74 months and the most common indication for revision surgery was abscess formation. The summary of the findings, including the quality of evidence of all the endpoints, is shown in Table 2.

Additional analyses

The results of a sensitivity analysis, including both-armed zero event studies, were not significantly different than the above mentioned analyses (available via the corresponding author). In the meta-regression analysis, study design had no effect on malocclusion in the intermediate FU (P>0.05) but had an effect on the reported risk ratios of plate removal (P=0.03). The prospective cohort studies had a significantly higher log risk ratio (2.61), whereas the retrospective studies did not (1.27) compared to the RCTs (-2.21; Table S6). No other meta-regression analyses could be performed. No funnel plots were constructed as none of the endpoints included >10 studies per study design.

The TSA showed that the required information size (RIS) for the infection and mobility of bone segment endpoints were not achieved and no boundaries were crossed (Table S7). Thus, based on the currently available evidence, TSA could not support the conclusions derived of conventional meta-analyses for these endpoints. Regarding the endpoints dehiscence and malocclusion at intermediate FU, the included patients made up <5% of the RIS and therefore a TSA could not be performed. The RIS for plate removal was achieved and the conventional test and the O’Brien-Fleming test boundary for benefit were crossed. Therefore, the provided evidence suggests that less symptomatic plate removal of biodegradable osteosynthesis occurred (Table S7). TSA could not be performed on all the other endpoints as these endpoints were assessed in no or in only a single randomized controlled trial, or were only assessed in total zero-event trials.

DISCUSSION
The present meta-analysis shows that the performance of biodegradable osteosynthesis is similar to titanium osteosynthesis regarding malunion, mobility of bone segments, and malocclusion after fixation of non-comminuted maxillofacial fractures. Additionally, no differences were found between both types of osteosyntheses regarding infection, dehiscence, plate exposure, pain, abscess formation, swelling, palpability of plates and/or screws, satisfaction, operative time, and revision surgery (i.e., no plate removal) at the predefined follow-up moments. The TSA showed that the required information size was not reached and thus the data remain inconclusive for these endpoints (i.e., may be false neutral). However, perioperative screw breakage during application occurred significantly more often in the biodegradable group compared to the titanium group. The symptomatic plate removal rate was significantly lower (i.e., 89% risk difference) in the biodegradable compared the titanium group. The TSA confirmed a true positive effect regarding plate removal, although only high risk of biased RCTs could be included. Finally, the meta-regression analysis showed that prospective cohort studies had significantly higher effect estimates of plate removal rate (i.e., in favour of the titanium group) compared to the RCTs and retrospective cohort studies.

Malunion was scarce in both intervention groups. Since pooled data derived from total zero-event studies is not available, the data from the RCTs and prospective cohort studies could not be synthesized. These outcomes, accompanied with the data on low mobility of bone segments and objective malocclusion, emphasise that both interventions are adequate for the fixation of maxillofacial fractures. This review focused on the objective and subjective malocclusion assessments by healthcare professionals or patients themselves, respectively. Although objective assessment of malocclusion is preferred over subjective ones for literature comparison purposes, we also feel that the patient’s opinion regarding occlusion is of high importance. Three studies assessed subjective malocclusion\textsuperscript{2,11,41}, of which one small RCT assessed subjective malocclusion after >5-years FU\textsuperscript{2}. In this latter study, subjective malocclusion was present in 50% of the titanium group compared to 0% in the biodegradable group. Also, the former group had worse mandibular function, as assessed by
the MFIQ, even though these patients were not assessed as having an objective malocclusion at the 2
year FU\textsuperscript{35}. Researchers should therefore also focus on long-term (i.e., \textgreater 5-years FU) objective and
subjective assessments of malocclusion and mandibular function as there may be discrepancy
between both assessments and after long-term follow-up.

An essential aspect of biodegradable osteosynthesis is its ability to degrade and be resorbed
in the human body, which may eliminate the need to remove implants in a second operation. Second
plate removal operations are accompanied with an additional risk of complications\textsuperscript{11}. The present
review shows that biodegradable osteosyntheses are removed significantly less often compared to
titanium ones due to symptoms. Although the subgroup analysis shows that symptomatic plate
removal did not differ significantly between both interventions in paediatric patients, all the titanium
plates were eventually removed (i.e., 100\% of plates) due to possible growth disturbances, while only
symptomatic biodegradable plates were removed from those patients (i.e., 12\% of plates; Figure S9).
Thus, titanium osteosyntheses will also eventually result in more re-operations compared to
biodegradable osteosyntheses in paediatric patients. The present review also performed a subgroup
analysis of plate removal rate between mandibular and other fractures. The biomechanical forces
acting on the mandible are considerably higher compared to fractures elsewhere hence, this could
result in loosening of the screws and subsequently to inflammation\textsuperscript{5}. Only three of all the
biodegradable osteosynthesis systems used in the included studies are certified to be used in the
mandible, namely the Inion CPS (Inion Oy, Tampere, Finland), GrandFix (Gunze, Kyoto, Japan), and
OsteotransMX (Teijin Medical Corp., Osaka, Japan)\textsuperscript{53-55}. All the instructions for the other
biodegradable systems explicitly state that these are contraindicated for use in load-bearing areas in
adults, including the mandible\textsuperscript{56-58} and yet, several studies implanted biodegradable osteosyntheses
off-label\textsuperscript{13,34,38,42}. Furthermore, the morphology and lesser vascularization of the mandible could
negatively influence fixation and degradation of biodegradable osteosyntheses\textsuperscript{2}. These factors have
been suggested to contribute to higher symptomatic plate removal rate in the mandible compared to
other facial fractures in both biodegradable and titanium osteosyntheses\textsuperscript{2}. The current meta-analysis
did not find significant differences between both osteosynthesis systems regarding symptomatic
plate removal rate when mandibular and other fractures were compared separately. Finally, most of
the included studies note a FU of up to 1 to 2 years. However, different studies have reported
titanium and biodegradable plate removal rates, during maxillofacial surgery, of up to 19% after a 5-
year FU\(^2,4,5\), while no plates were removed between 1- and 5-years FU\(^5\). Therefore, future research
should extend the FU beyond 2-years in order to assess the plate removal rate adequately in both
intervention groups.

Foreign-body reactions after implantation of biodegradable osteosynthesis systems have
been reported and remain a concern in the usage of such systems\(^2,9,10\). The present review did not
find any differences, regarding the presence of swelling or abscess formation, between both
interventions after short- and long-term FUs, although it must be noted that only two studies had
included patients with \(>3\) years follow-up\(^2,40\). Also, revision surgery (i.e., non-plate removal) was
scarce and there was no difference between both groups. The factors that are known to influence
foreign-body reactions are implant related (i.e., polymer composition, plate size and shape, surface
texture), recipient related (i.e., blood supply, temperature), and related to the location of plate
placement (i.e., subcutaneous, epiperiosteal, subperiosteal). Of these factors, polymer composition
has been studied the most\(^60\)–\(^62\). The reported foreign body reactions occur predominately in
biodegradable osteosyntheses with high proportion of PLLA (i.e., \(>70\%\)) composition\(^2,9,10,63\). PLLA
degrades in two phases to eventually form CO\(_2\) and H\(_2\)O as final products: early degradation via
hydrolysis produces crystalline structures which undergo secondary hydrolysis. Secondary hydrolysis
is the rate-limiting step and depends highly on the crystallinity and hydrophobicity of the
intermediate products. L-isomers form crystalline products that are highly hydrophobic and
therefore more resistant to degradation and resorption compared to D-isomers\(^60\). PLLA crystalline
particles have been identified intra-cellularly up to 5.7 years after fixation of zygomatic fractures in
patients\(^10\). Only one of the included studies reported sterile abscess formation which was incised and
drained during a second operation\(^40\). That study used a 70%-30% PLLA/PDLLA biodegradable
osteoosynthesis system. More amorphous (co-)polymer compositions such as polyglycolide (PGA), poly(lactic-co-glycolic acid) (PLGA), or PDLLA are more hydrophilic and undergo degradation and resorption more quickly\textsuperscript{60}. Tissue response to PLLA has been extensively studied in animals and patients, with a long-term follow-up (i.e., up to 6 years), whereas no long term data is currently available for PGA, PLGA, or PDLLA (co-)polymer compositions. Current in vivo studies including these biodegradable systems have been performed with a follow-up up to 18 months\textsuperscript{60,62,64}. Long-term in vivo degradation of these (co-)polymer compositions are currently being investigated by our research group and the results are eagerly being awaited. Additionally, future research should preferably incorporate the other factors that contribute to foreign-body reactions.

Data about analgesia usage, MMO, MFIQ, TMJ-dysfunction, handling of osteosynthesis systems by surgeons, perioperative plate breakage, and total costs could not be synthesised due to a lack of studies which had (adequately) assessed these endpoints. Analgesia usage was not assessed in any of the included studies and TMJ-dysfunction was only noted in one recent study\textsuperscript{8}. Data of (postoperative) MMO could not be synthesized on account of only a few studies reporting MMO or because the authors could not provide the data. MFIQ was only assessed in two publications consisting of the same study population\textsuperscript{2,35}. Thus, there is currently insufficient evidence to provide conclusions regarding mandibular and TMJ-function after both interventions. Although pre-operative endpoint data are preferred in order to assess the effect of the osteosynthesis system on these endpoints, the patients presenting with maxillofacial fractures often have restricted MMO and impaired mandibular function as a consequence of the trauma. It is unlikely that any data will be at hand regarding mandibular function before the fracture. Therefore, future researchers should collect post-operative data regarding TMJ-function and MMO or use validated questionnaires (e.g., MFIQ) to make adequate assessments of mandibular function and to enable comparisons with healthy subjects.

Total costs were assessed in only one small RCT and titanium osteosyntheses were associated with higher costs compared to biodegradable ones, mainly due to the additional costs of a
second operation for symptomatic plate removal\textsuperscript{36}. Finally, only a small RCT reported the handling of osteosynthesis systems by surgeons\textsuperscript{33}. The differences between both systems were small and the authors report that more exposure to biodegradable systems by surgeons could diminish this difference.

The meta-regression analysis showed that the effect estimates of plate removal rates by prospective studies were significantly higher compared to randomized controlled trials and retrospective studies. One of the prospective studies included in this analysis allowed the patients to voluntarily choose the fixation material\textsuperscript{13}. The patient’s choice is always dependent on the provided information, and therefore dependent on the healthcare professional. The other study could not randomize the patients due to the occasional unavailability of the required plating systems\textsuperscript{11}. Both studies are therefore prone to selection bias. Selection bias has been shown to exaggerate effect estimates\textsuperscript{16} and, thus, this could explain the difference in the effect estimates between the different study designs.

Comparison to other systematic reviews

A systematic review in 2013, comparing complications after fracture fixation between five studies, showed that biodegradable osteosyntheses had lower overall complication rates compared to titanium osteosyntheses (RR 0.71, 95% CI 0.52;0.97)\textsuperscript{15}. A subgroup analysis of these complications indicated that only the palpability of the plates remained significantly lower in the biodegradable group (RR 0.38, 95% CI 0.22;0.68). However, that review used a fixed-effects model, while methodological and clinical heterogeneity was clearly present (e.g., different study designs, composition of biodegradable plates), and it did not perform an assessment of the endpoints in relation to follow-up. Additionally, the difference in palpability was based on a single small retrospective study\textsuperscript{65}. The results of the present review show that, according to current evidence, there is no significant difference in complications between both interventions. In particular, there is no difference in long-term palpability between both interventions. Furthermore, the aforementioned
review concluded that no publication bias was present by using funnel plots, although only five
studies were included and the endpoints were only assessed based on one (e.g., palpability) to four
studies (e.g., infection). Funnel plots with ≤10 studies are underpowered and inconclusive, and thus,
their usage is discouraged if insufficient studies could be included for a meta-analysis\textsuperscript{16,25–27}. Finally,
the authors do not provide any data regarding inter-observer agreement and do not incorporate risk
of bias in the interpretation of the results. We therefore express our concerns about the conclusions
drawn in that particular review.

\textit{Quality of the evidence}

All the studies considered had two or more domains assessed as high risk of bias owing to the nature
of the intervention. Biodegradable plates and screws are easily distinguished by surgeons (i.e., no
blinding possible) and are not visible on radiographs while titanium osteosyntheses are visible (i.e.,
no blinding of the outcome assessment is possible). Therefore, these two domains do not result in
differences in quality between the included studies.

The evidence was of very low or moderate quality as assessed by the GRADE system. The
main reasons of downgrading the quality of evidence was high risk of bias, indirectness, and
imprecision of the data. Moderate quality evidence was found for perioperative screw breakage and
plate removal rate. Infection (<4 weeks FU), dehiscence (<4 weeks FU), mobility of bone segments
(6-12 weeks FU), and malocclusion (6-12 weeks FU) were assessed as very low quality. The quality of
evidence of the endpoints malunion and pain (6-12 weeks FU), and MFIQ, swelling, and palpability of
plates/screws (>12 weeks FU) could not be assessed due to zero event studies, different outcome
measures, or studies that consisted of the same study population with different follow-up moments.
Also, the RCTs data regarding revision surgery could not be pooled due to zero event studies.

The data derived from the prospective and retrospective cohort studies were assessed as
very low quality. Endpoints based on very low quality evidence cannot be used to make
recommendations to surgeons and should therefore be interpreted with caution\textsuperscript{16}. 

22
Strengths and limitations

The strengths of the current meta-analysis are: the transparent and robust methodology used, based on a pre-specified protocol, the PRISMA statement, and the Cochrane Handbook. Also, a comprehensive and up-to-date literature search was performed without language or period restrictions. A range of relevant endpoints with predefined follow-up moments were included. Furthermore, study eligibility, data-extraction, and risk of bias assessment were performed independently by two reviewers with excellent inter-observer agreement. Also, we used TSA to increase the reliability of our data and to determine the required information size of each endpoint. Finally, certainty of evidence was assessed in duplicate using GRADE.

The limitations of this review include the low quality of the studies due to high risk of bias. Therefore, we cannot exclude a biased effect estimate. Additionally, clinical heterogeneity could not be excluded due to the inclusion of studies with different biodegradable and titanium systems (i.e., different compositions), different sized osteosynthesis systems, and the differences in the application and duration of the MMF. Subgroup analysis (i.e., mandibular versus other fractures) of the primary endpoints could not be performed due to a lack of studies. Finally, some data could not be retrieved from the authors of the original manuscripts despite multiple efforts and could therefore not be included in this review.

Implications for future research

This review shows that the quality of the current evidence ranges from very low to moderate and high quality research is therefore necessary. The main reason for downgrading the evidence was the high risk of bias in all of the included studies. Although blinding the surgeons and the outcome assessors is not possible due to the nature of the intervention, and thus contributes substantially to the risk of bias, none of the studies could be assessed as low risk of bias when these two domains of blinding were excluded. We, therefore, suggest that future RCTs should be performed with long-term
follow-up using pre-specified and well-defined protocols. The pre-specified protocol should pay particular interest to: (i) well-defined endpoints to minimize reporting bias, (ii) adequate follow-up of the corresponding endpoints to minimize attrition bias, and (iii) well-defined indications for plate removal to minimize detection bias. Also, more patient reported outcomes (e.g., subjective malocclusion, MFIQ) are preferred. Additionally, the reporting of patient characteristics, surgical procedures, and outcomes should be improved. In particular, researchers should include details regarding the osteosynthesis systems used (i.e., composition, sizes, osteosynthesis principle), alcohol and tobacco usage, as these factors are known to compromise wound healing and decrease vascularization intra-orally which may affect degradation and resorption rates, and the use of, reasons for, and duration of, the MMF. We advocate that future studies should comply with the CONSORT guidelines to ensure high quality reporting of all aspects of the methodology and results. This enables appraisal, interpretation, and pooling of future data. Finally, future studies should focus on the cost-effectiveness of biodegradable systems, including direct (i.e., perioperative costs) and indirect costs (e.g., second operations, absence from work).

Based on all currently available evidence after both narrative review and meta-analyses, biodegradable and titanium osteosyntheses are similar regarding the efficacy and morbidity of fixation of non-comminuted maxillofacial fractures. However, perioperative screw breakage occurred significantly more often in the biodegradable compared to the titanium group. The symptomatic plate removal rate was significantly lower after biodegradable compared to titanium fixation in this population. Combining these aspects, current available evidence shows that biodegradable osteosynthesis is a viable alternative to titanium osteosynthesis after maxillofacial trauma. Due to the low to moderate quality of the included studies, the results of this systematic review should be interpreted with caution.
ACKNOWLEDGMENT

The authors would like to thank Ms. S. van der Werf, biomedical information specialist at the University of Groningen, for her assistance in developing the search strategy. The authors would also like to thank Ms. X. Wang and Ms. S.J. van der Geer, research scientists, for translating a Chinese article and performing the risk of bias assessment of trials performed by the author’s research group, respectively.

FUNDING AND CONFLICT OF INTERESTS

No funding was received to conduct this study. The authors state that they have no conflict of interests.
REFERENCES


7. Jorgenson DS., Mayer MH., Ellenbogen RG., Centeno JA., Johnson FB., Mullick FG., Manson


29. Wetterslev J., Thorlund K., Brok J., Gluud C. Trial sequential analysis may establish when firm
evidence is reached in cumulative meta-analysis 2008;61:64–75. Doi:
10.1016/j.jclinepi.2007.03.013.

information size and potentially false positive results in many meta-analyses 2008;61. Doi:
10.1016/j.jclinepi.2007.10.007.

31. Wetterslev J., Jakobsen JC., Gluud C. Trial Sequential Analysis in systematic reviews with meta-

Sequential Analysis (TSA). *Copenhagen Trial Unit, Cent Clin Interv Res Copenhagen, Denmark*

33. Buijs GJ., Van Bakelen NB., Jansma J., de Visscher JGAM., Hoppenreijs TJM., Bergsma JE.,
Stegenga B., Bos RRM. A randomized clinical trial of biodegradable and titanium fixation
10.1177/0022034511434353.

Bioresorbable versus titanium plates for mandibular fractures. *J Coll Physicians Surg Pakistan*

35. van Bakelen NB., Buijs GJ., Jansma J., de Visscher JG., Hoppenreijs TJ., Bergsma JE., Stegenga
B., Bos RR. Comparison of biodegradable and titanium fixation systems in maxillofacial
Doi: 10.1177/0022034513508953 [doi].

36. Van Bakelen NB., Vermeulen KM., Buijs GJ., Jansma J., De Visscher JGAM., Hoppenreijs TJM.,
Bergsma JE., Stegenga B., Bos RRM. Cost-effectiveness of a biodegradable compared to a
titanium fixation system in maxillofacial surgery: A multicenter randomized controlled trial. 


Wu CM., Chen YA., Liao HT., Chen Chao., Pan CH., Chen CT. Surgical treatment of isolated zygomatic fracture: Outcome comparison between titanium plate and bioabsorbable plate.


55. Teijin Medical Technologies. OsteoTrans-MX Product Overview.


60. Xue AS., Koshy JC., Weathers WM., Wolfswinkel EM., Kaufman Y., Sharabi SE., Brown RH.,
Hicks MJ., Hollier LHJ. Local foreign-body reaction to commercial biodegradable implants: an
0033-1364199.

61. Suuronen R., Pohjonen T., Hietanen J., Lindqvist C. A 5-year in vitro and in vivo study of the
2391(98)90461-X [pii].

62. Landes CA., Ballon A., Roth C. In-Patient Versus In Vitro Degradation of P(L/DL) LA and PLGA

63. Jeon HB., Kang DH., Gu JH., Oh SA. Delayed Foreign Body Reaction Caused by Bioabsorbable

Miniplates (LactoSorb): Long-Term Results in Infant Minipigs and Clinical Results. *J Craniofac

65. Menon S., Chowdhury SKR. Evaluation of bioresorbable vis-à-vis titanium plates and screws
10.1016/S0377-1237(07)80008-6.

# Table 1: Risk of bias assessment of all included studies.

<table>
<thead>
<tr>
<th>Study name (year)</th>
<th>Random sequence generation (selection bias)</th>
<th>Allocation concealment (selection bias)</th>
<th>Blinding of participants and researchers (performance bias)</th>
<th>Blinding of outcome assessment (detection bias)</th>
<th>Incomplete outcome data (attrition bias)</th>
<th>Selective reporting (reporting bias)</th>
<th>Other bias</th>
<th>Cochrane Collaboration’s tool for assessing risk of bias</th>
<th>MINORS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bhatt et al. (2010)</td>
<td>Low</td>
<td>Unclear</td>
<td>High</td>
<td>High</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buljs et al. (2012)</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ahmed et al. (2013)</td>
<td>Low</td>
<td>Unclear</td>
<td>High</td>
<td>High</td>
<td>Unclear</td>
<td>Low</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bakelen et al. (2013)</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bakelen et al. (2015)</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Type</td>
<td>Year</td>
<td>Country</td>
<td>Study Design</td>
<td>Relative Risk</td>
<td>Diagnosis</td>
<td>Matched Cases</td>
<td>Matched Controls</td>
<td>Study Design</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------------------</td>
<td>------</td>
<td>---------</td>
<td>------------------------</td>
<td>---------------</td>
<td>-------------------------</td>
<td>---------------</td>
<td>------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Sukegawa et al. (2016)</td>
<td>Unclear Unclear</td>
<td>2016</td>
<td>Japan</td>
<td>Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gareb et al. (2017)</td>
<td>Low Low High High</td>
<td>2017</td>
<td>Egypt</td>
<td>Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leonhardt et al. (2008)</td>
<td>Prospective cohort studies</td>
<td>2008</td>
<td>Germany</td>
<td>Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mahmoud et al. (2016)</td>
<td>High High Low Low</td>
<td>2016</td>
<td>Egypt</td>
<td>Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leno et al. (2017)</td>
<td>Low Low High High</td>
<td>2017</td>
<td>France</td>
<td>Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bell et al. (2006)</td>
<td>Low Low Low Low</td>
<td>2006</td>
<td>Germany</td>
<td>Retrospective cohort studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wittwer et al. (2006)</td>
<td>High High Low Low</td>
<td>2006</td>
<td>Germany</td>
<td>Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee et al. (2010)</td>
<td>High High Low Low</td>
<td>2010</td>
<td>Germany</td>
<td>Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Park et al. (2011)</td>
<td>High High Low Low</td>
<td>2011</td>
<td>Germany</td>
<td>Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Menon et al. (2012)</td>
<td>High High Low Low</td>
<td>2012</td>
<td>Germany</td>
<td>Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tripathi et al. (2013)</td>
<td>High High Low Low</td>
<td>2013</td>
<td>Germany</td>
<td>Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kang et al. (2014)</td>
<td>High High Low Low</td>
<td>2014</td>
<td>Germany</td>
<td>Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lim et al. (2014)</td>
<td>High High Low Low</td>
<td>2014</td>
<td>Germany</td>
<td>Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bhatt et al. (2015)</td>
<td>High High Low Low</td>
<td>2015</td>
<td>India</td>
<td>Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burlini et al. (2015)</td>
<td>High High Low Low</td>
<td>2015</td>
<td>Italy</td>
<td>Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Filinte et al. (2015)</td>
<td>High High Low Low</td>
<td>2015</td>
<td>Italy</td>
<td>Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wu et al. (2017)</td>
<td>High High Low Low</td>
<td>2017</td>
<td>China</td>
<td>Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim et al. (2018)</td>
<td>High High Low Low</td>
<td>2018</td>
<td>South Korea</td>
<td>Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
MINORS: Methodological index for non-randomized studies. High: high risk of bias; Low: low risk of bias; Unclear: unclear risk of bias; 0: not reported; 1: reported but inadequate; 2: reported and adequate. Empty cells: not applicable
**Table 2:** Summary of findings with quality of evidence assessment.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Randomized controlled trials</th>
<th>Prospective cohort studies</th>
<th>Retrospective cohort studies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subjects, N (studies)</td>
<td>RR or SMD (95% CI)</td>
<td>Subjects, N (studies)</td>
</tr>
<tr>
<td>Plate</td>
<td>Two studies, of which 1 had zero events (see Table S5)</td>
<td>Single study (see Table S5)</td>
<td>Four studies, of which 3 had zero events (see Table S5)</td>
</tr>
<tr>
<td>Screw</td>
<td>718 (2)</td>
<td>17.13 (2.19; 134.18)</td>
<td>748 (3)</td>
</tr>
<tr>
<td>Operation time</td>
<td>Single study (see Table S5)</td>
<td>Single study (see Table S5)</td>
<td>165 (3)</td>
</tr>
<tr>
<td>Handling by surgeon</td>
<td>Single study (see Table S5)</td>
<td>No studies</td>
<td>No studies</td>
</tr>
</tbody>
</table>

**Perioperative endpoints**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Randomized controlled trials</th>
<th>Prospective cohort studies</th>
<th>Retrospective cohort studies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subjects, N (studies)</td>
<td>RR or SMD (95% CI)</td>
<td>Subjects, N (studies)</td>
</tr>
<tr>
<td>Plate</td>
<td>Two studies, of which 1 had zero events (see Table S5)</td>
<td>Single study (see Table S5)</td>
<td>Four studies, of which 3 had zero events (see Table S5)</td>
</tr>
<tr>
<td>Screw</td>
<td>718 (2)</td>
<td>0 per 1000</td>
<td>No studies</td>
</tr>
<tr>
<td>Operation time</td>
<td>Single study (see Table S5)</td>
<td>Single study (see Table S5)</td>
<td>165 (3)</td>
</tr>
<tr>
<td>Handling by surgeon</td>
<td>Single study (see Table S5)</td>
<td>No studies</td>
<td>No studies</td>
</tr>
</tbody>
</table>

**Short-term follow-up**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Randomized controlled trials</th>
<th>Prospective cohort studies</th>
<th>Retrospective cohort studies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subjects, N (studies)</td>
<td>RR or SMD (95% CI)</td>
<td>Subjects, N (studies)</td>
</tr>
<tr>
<td>Malocclusion</td>
<td>Single study (see Table S5)</td>
<td>Three studies of which 2 had zero events (see Table S5)</td>
<td>91 (2)</td>
</tr>
<tr>
<td>Infection</td>
<td>Single study (see Table S5)</td>
<td>Three studies of which 2 had zero events (see Table S5)</td>
<td>194 (3)</td>
</tr>
</tbody>
</table>

---

*a* Includes multiple studies with varying outcomes; *b* Includes multiple studies with varying outcomes; *c* Includes multiple studies with varying outcomes; *d* Includes multiple studies with varying outcomes; *e* Includes multiple studies with varying outcomes; *f* Includes multiple studies with varying outcomes; *g* Includes multiple studies with varying outcomes; *h* Includes multiple studies with varying outcomes; *i* Includes multiple studies with varying outcomes; *j* Includes multiple studies with varying outcomes; *k* Includes multiple studies with varying outcomes; *l* Includes multiple studies with varying outcomes; *m* Includes multiple studies with varying outcomes; *n* Includes multiple studies with varying outcomes; *o* Includes multiple studies with varying outcomes; *p* Includes multiple studies with varying outcomes.
<table>
<thead>
<tr>
<th>Condition</th>
<th>Studies/Event Rate</th>
<th>Studies/Event Rate</th>
<th>Studies/Event Rate</th>
<th>Studies/Event Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Swelling</strong></td>
<td>Single zero-event</td>
<td>No studies</td>
<td>Two studies, of which one had 100% event rate in both groups</td>
<td>(see Table S5)</td>
</tr>
<tr>
<td></td>
<td>study (see Table S5)</td>
<td></td>
<td>(see Table S5)</td>
<td></td>
</tr>
<tr>
<td><strong>Abscess</strong></td>
<td>Single study</td>
<td>No studies</td>
<td>No studies</td>
<td>(see Table S5)</td>
</tr>
<tr>
<td></td>
<td>(see Table S5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Pain</strong></td>
<td>Single study</td>
<td>No studies</td>
<td>Single study (see Table S5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(see Table S5)</td>
<td></td>
<td>(see Table S5)</td>
<td></td>
</tr>
<tr>
<td><strong>Analgesics</strong></td>
<td>No studies</td>
<td>No studies</td>
<td>No studies</td>
<td></td>
</tr>
<tr>
<td><strong>MMO</strong></td>
<td>No studies</td>
<td>Two studies, of which one only reported postoperative MMO (see Table S5)</td>
<td>Single study (see Table S5)</td>
<td></td>
</tr>
<tr>
<td><strong>Dehiscence</strong></td>
<td>126 (2)</td>
<td>1.68</td>
<td>126 per</td>
<td>Four studies, of which 3 had zero events (see Table S5)</td>
</tr>
<tr>
<td></td>
<td>(0.56;5)</td>
<td>75 per</td>
<td>126 per</td>
<td>(see Table S5)</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>1000</td>
<td>Very low</td>
<td>123 (3)</td>
</tr>
<tr>
<td></td>
<td>(0.32;1)</td>
<td>(42;375)</td>
<td>1.94</td>
<td>0.58</td>
</tr>
<tr>
<td><strong>Plate exposure</strong></td>
<td>Single zero-event</td>
<td>Single study (see Table S5)</td>
<td>Single study (see Table S5)</td>
<td>(see Table S5)</td>
</tr>
<tr>
<td></td>
<td>study (see Table S5)</td>
<td></td>
<td>(see Table S5)</td>
<td></td>
</tr>
<tr>
<td><strong>Malunion</strong></td>
<td>Three zero-event</td>
<td>Three zero-event</td>
<td>Very low</td>
<td>312 (2)</td>
</tr>
<tr>
<td></td>
<td>studies (see Table S5)</td>
<td>(see Table S5)</td>
<td>1.24</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>(0.15;1000)</td>
<td>0.32</td>
<td>16 per</td>
<td>(see Table S5)</td>
</tr>
<tr>
<td><strong>Mobility bone</strong></td>
<td>100 (2)</td>
<td>2.11</td>
<td>21 per</td>
<td>Single study (see Table S5)</td>
</tr>
<tr>
<td></td>
<td>(0.32;1)</td>
<td>1000</td>
<td>44 per</td>
<td>(see Table S5)</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>Very low</td>
<td>1.34</td>
<td></td>
</tr>
<tr>
<td>Condition</td>
<td>Two studies, different outcome measures (see Table S5)</td>
<td>No studies</td>
<td>No studies</td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>--------------------------------------------------------</td>
<td>------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>Malocclusion</td>
<td>Three studies of which 2 had zero-events (see Table S5)</td>
<td>Two studies of which 1 had zero-events (see Table S5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td>Two studies, of which one only reported postoperative MMO (see Table S5)</td>
<td>No studies</td>
<td>No studies</td>
<td></td>
</tr>
<tr>
<td>MMO</td>
<td>No study</td>
<td>No studies</td>
<td>No studies</td>
<td></td>
</tr>
<tr>
<td>TMJ-dysfunction</td>
<td>Single study (see Table S5)</td>
<td>Single study (see Table S5)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Long-term follow-up**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Two studies with the same study population (see Table S5)</th>
<th>Three zero-event studies (see Table S5)</th>
<th>Single zero-event study (see Table S5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malocclusion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td>Two zero-event studies with the same study population (see Table S5)</td>
<td>No studies</td>
<td>194 (3) 0.40 44 per 18 per Very low (^{3,4} )</td>
</tr>
<tr>
<td>MMO</td>
<td>No studies</td>
<td>Two studies, of which one only reported postoperative MMO (see Table S5)</td>
<td>Single study with only postoperative data (see Table S5)</td>
</tr>
<tr>
<td>TMJ-dysfunction</td>
<td>No studies</td>
<td>No studies</td>
<td>Single study (see Table S5)</td>
</tr>
<tr>
<td>MFIQ</td>
<td>Two studies with the same study population (see Table S5)</td>
<td>No studies</td>
<td>No studies</td>
</tr>
<tr>
<td>Abscess</td>
<td>Single study (see Table S5)</td>
<td>No studies</td>
<td>391 (3) 2.37 4 per 9 per Very low (^{3,4} )</td>
</tr>
</tbody>
</table>

\(^{3,4} \)
### Swelling

<table>
<thead>
<tr>
<th>Two studies with the same study population (see Table S5)</th>
<th>No studies</th>
<th>363 (3)</th>
<th>4.55</th>
<th>0 per</th>
<th>NA</th>
<th>Very low&lt;sup&gt;1–4&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.78;26.68)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Palpability

<table>
<thead>
<tr>
<th>Three studies, of which two had the same study population and one had zero events (see Table S5)</th>
<th>Single zero-event study (see Table S5)</th>
<th>188 (4)</th>
<th>0.30</th>
<th>60 per</th>
<th>18 per</th>
<th>Very low&lt;sup&gt;1–4&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.07;1.3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Satisfaction

<table>
<thead>
<tr>
<th>No studies</th>
<th>Single study (see Table S5)</th>
<th>71 (2)</th>
<th>-0.20</th>
<th>NA</th>
<th>NA</th>
<th>Very low&lt;sup&gt;1–4&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-0.92;0.52)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Overall follow-up

<table>
<thead>
<tr>
<th>Symptomatic plate removal*</th>
<th>Moderate&lt;sup&gt;1&lt;/sup&gt;</th>
<th>104 (2)</th>
<th>1.51</th>
<th>80 per</th>
<th>121 per</th>
<th>Very low&lt;sup&gt;1–4&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.49;1.000;1.000)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Downgraded one level due to high risks of bias identified across studies: the majority of studies had high or unclear risk of bias in at least two of the domains assessed; <sup>1</sup> Downgraded one level for inconsistency: substantial methodological or clinical heterogeneity that could not be accounted for in analyses; <sup>2</sup> Downgraded one level for indirectness: the evidence of the original manuscripts were more restrictive than the review question; <sup>3</sup> Downgraded one level for imprecision: limits of effect estimate confidence interval are not consistent (i.e., cover both benefit and harm); <sup>4</sup> Upgraded one level due to large effect (i.e. RR<0.5 or RR>2.0, or SMD<−0.8 or SMD>0.8).
1 FIGURE CAPTIONS

2 Figure 1: Flowchart of study identification and selection progress.

3 Figure 2: Forest plots of the primary endpoints (A) malunion (6-12 weeks FU), (B) mobility of bone segments (6-12 weeks FU), (C) malocclusion (<4 weeks), and (D) malocclusion (6-12 weeks FU) stratified by study design. FU, follow-up; Retrosp. CS, retrospective cohort studies; RCT, randomized controlled trials; RR, risk ratio; 95%-CI, 95% confidence interval, NA: not applicable.

4 Figure 3: Forest plots of the secondary endpoints (A) infection (<4 weeks FU), (B) abscess (>12 weeks FU), (C) swelling (>12 weeks FU), (D) palpability of plates/screws (>12 weeks FU), (E) satisfaction (>12 weeks FU), and (F) symptomatic plate removal (overall FU) stratified by study design. FU, follow-up; RCT, randomized controlled trials; Prospec. CS, prospective cohort studies; Retrosp. CS, retrospective cohort studies; RR, risk ratio; SMD, standardised mean difference; 95%-CI, 95% confidence interval.