Influences of Statistical Analyses on Result Presentations of Oral Implant Treatment

Irene Herrmann

GÖTEBORGS
UNIVERSITET

Sahlgrenska akademin
Department of Oral and Maxillofacial Surgery

Göteborg, Sweden

2007

Abstract

Evaluation of oral implants involves at least three major components: statistical methods, study design and success criteria.

The aims of the current thesis were to investigate how different statistical methods affect the outcome of oral implant treatment and to statistically determine if any dependence among implants placed in the same jaw exists. If so - how would that affect the outcome and the type and amount of missing data (withdrawn patients)? Furthermore, the aims were to evaluate patient, implant and treatment characteristics to find possible prognostic factors for implant failure and to study the impact of variances with or without handling dependence using the Jackknife technique.

Four prospective multi-centre studies, involving 487 patients and 1738 implants, were pooled to create a database for these elaborations. The database was divided into subgroups based on significant different outcomes regarding implant failures. Four jaw-bone combinations (Combination I-IV) were established, and in study IV - Combinations I-III were pooled and compared with Combination IV. Statistical methods used were: life table analyses, confidence intervals, chi^2 tests, step by step multivariate analyses, post hoc analyses, log rank tests and the Jackknife technique.

The result of the current statistical investigations demonstrated that dependence among implants placed in the same jaw existed. The impact of missing data was shown to depend on, if the patients were selectively or randomly withdrawn. A random selection could reach at the most 50% without affecting the CSR. However, if the selection was based on e.g. treated jaw, the outcomes were significantly different. Patient characteristics such as jaw, jaw-bone quality, jaw-shape and combination of these factors, and also on implant length and treatment protocol, showed significant differences. Patients with jaw shape D or E and bone-quality 4 were e.g. identified with a significantly higher risk for implant failure than all of the other combinations. Both life table analyses, using CI and log rank tests, demonstrated after the variances were calculated via the Jackknife technique significantly lower success rates for Combination IV. The p-values were “inadequately” stronger, however, when using a log rank test and ignoring the established dependence.

Based on these results it was stated that, following established dependence among implants within the same jaw, this should never be ignored when evaluating oral implant outcomes. Two methods were found possible to use to handle this dependence, the “one implant per patient” or the Jackknife technique if variances are part of the evaluation. Missing data is inevitable and will affect the outcome, and therefore a description of the characteristics of withdrawn patients should be presented. Jaw-combination IV showed the lowest success rate and would therefore be the most appropriate population to use, when evaluating new improved oral implant systems in order to prove significantly different follow-up outcome.

Keywords: Statistical methods, oral implants dependence, prognostic risk factors, multivariate analysis, bone-combination, Jackknife, log rank test, variance.
ISBN:978-91-628-7206-9

Correspondence: Irene Herrmann, Södra Småskolevägen 41, 429 44 Särö Sweden.; e-mail: irene.herrmann@lifecore.com

List of Papers

This thesis is based on the following papers, which will be referred to in the text by their Roman numbers:

I. **Herrmann I, Lekholm U, Holm S, Karlsson S.** Impact of implant interdependency when evaluating success rates: A statistical analysis of multicentre results.
Int J Prosthodont 1999; 12: 160-166.

II. **Herrmann I, Lekholm U, Holm S.** Statistical outcome of random versus selected withdrawal of dental implants.
Int J Prosthodont 2003; 16: 25-30.

III. **Herrmann I, Lekholm U, Holm S, Kultje C.** Evaluation of patient/implant characteristics as potential prognostic factors for oral implant failures.
Int J Oral Maxillofac Implants 2005; 20: 220-30.

IV. **Herrmann I, Kultje C, Holm S, Lekholm U.** A study on variances in multivariate Analyses of Oral Implant Outcome.
Clin Implant Dent Relat Res 2007; 9: 6-14.

Contents

Introduction

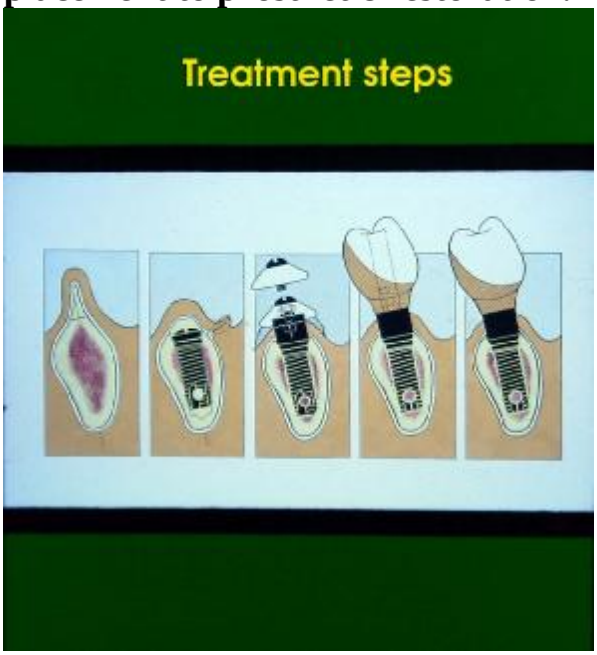
The use of oral implants

The modern era of oral implants started more than 40 years ago, when Professor Per-Ingvar Brånemark almost accidentally discovered that titanium components could integrate into living bone (Brånemark, 1969), a condition he referred to as osseointegration. Subsequently, this finding was developed into a clinical procedure, internationally presented in Toronto in 1982 (Zarb, 1983); the consensus of that meeting has with time become a milestone for many of the oral implant treatments used today.

The osseointegration procedure

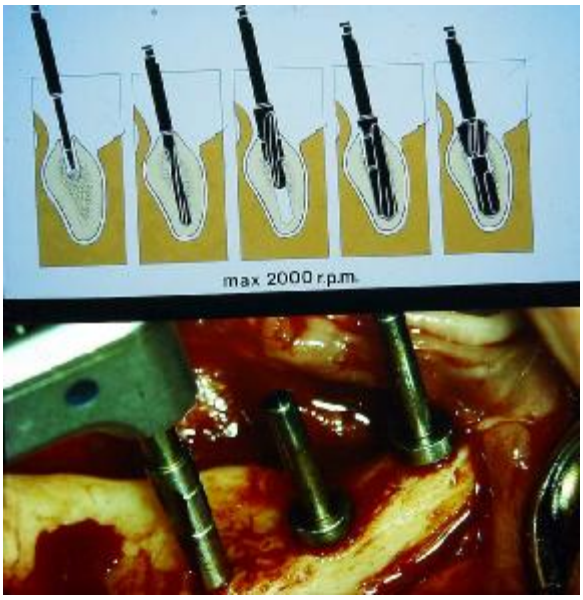
The original clinical procedure *ad modum* Brånemark consists of two surgical steps followed by the prosthetic rehabilitation (Brånemark et al., 1969, 1977; Figure 1).

Figure 1
Treatment steps from implant placement to prosthetic restoration.



During stage one surgery, a full-thickness muco-periostal flap is elevated under strict sterile conditions. The implant sites are then prepared by successive drills of gradually increasing diameter (Figure 2). The individual sites are subsequently countersunk and generally threaded before the titanium implants are placed into the bone-sites, using a non-traumatic surgical technique

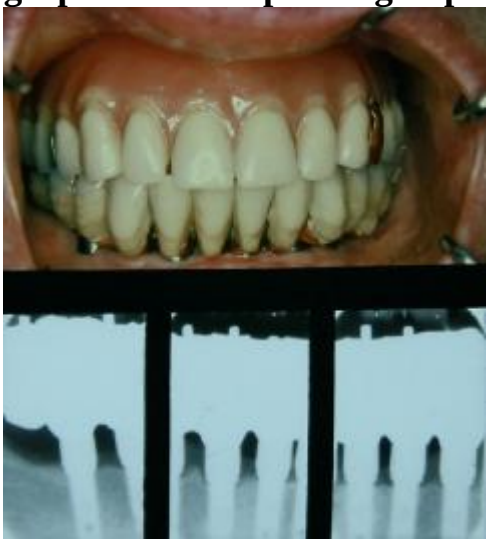
Figure 2
Preparation of implant sites by successive incremented drill diameters.



All drilling is performed under extensive cooling with saline to avoid overheating the bone (Adell et al., 1985; Sutter et al., 1992; Brägger et al., 1995). Following implant placement, the muco-periostal flap is re-adapted over the implant sites. Second stage surgery - abutment connection - takes place after undisturbed healing of the implants (three months for lower and six months for upper jaws). During this procedure, the mucosa perforating extensions are attached to the implants.

Finally, the prosthetic treatment is completed when the full bridge construction is fabricated and attached to the abutments without imposing any static forces on the anchoring units (Rangert B et al., 1989; Alkan et al., 2004; Natali et al., 2006). A well-balanced occlusion of the construction is ensured and maintained against the dentition of the opposing jaw (Figure 3).

Figure 3
Lower full arch bridge and radio-graphs of corresponding implant.



Since the technique was internationally introduced in 1982, it has been extensively used all over the world, and the outcome has been profoundly evaluated, using strict follow-up protocols (Albrektsson & Lekholm, 1989; Henry, 1999; Naert et al., 2001; Esposito et al., 2001:a, 2005; Gapski et al., 2003).

Consequently, treatment *ad modum* Brånemark has become a "gold standard" for oral implant treatment protocols used today (Millennium research group, 2005).

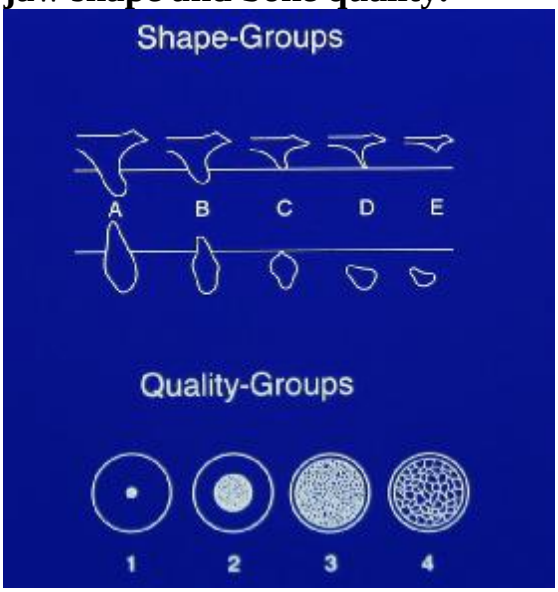
Possible prognostic factors

The outcome of oral implant treatment may depend on several technique related factors (Brånemark et al., 1977); such as patient characteristics and selection, implant components and skill of the performing clinicians. Six considerations have been identified by Albrektsson et al. (1981) as important for the establishment of implant integration in bone. The influence of these factors on treatment outcome has subsequently been confirmed by others (e.g. Friberg et al., 1991; Wennerberg et al., 1996; Ivanoff et al., 2001; Gapski et al., 2003).

Patient characteristics

As always, patient selection and examination are important for the outcome of the treatment (Adell et al., 1986; Lekholm, 1998; Andersson et al., 1995; van Steenberghe et al., 2002; Sugermann & Barber, 2002). The patient related characteristic specifically highlighted by Albrektsson et al. (1981) is the *status of the bone* of the implant site. This was later developed into an index (Lekholm & Zarb, 1985), related to bone quality (cortical versus marrow bone) and jaw-shape (degree of resorption).

Figure 4
Lekholm & Zarb index related to jaw shape and bone quality.



This Lekholm & Zarb index (Figure 4) has frequently been applied, due to its ease of use and strong predictive value for implant failures (Friberg et al., 1991; Truhlar et al., 1994; Sennerby & Roos, 1998; Hutton et al., 1995). Smoking is another patient related factor strongly correlated with implant failures (Bain & Moy, 1993; DeBruyn & Collaert, 1994), due to the influence of nicotine on the bone vascularisation.

Technique related factors

Two factors related to the clinical performance have been specifically pointed out by Albrektsson et al. (1981) as important for enabling a good treatment

outcome. The first is the use of a non-traumatic *surgical technique* to minimize tissue violence (Eriksson & Albrektsson, 1984; Eriksson et al., 1984). This has been confirmed in several other reports as one of the most important factors in decreasing the frequency of early implant failures (Cordioli et al., 1997; Ercoli et al., 2004). The second technique related factor put forward by Albrektsson et al. (1981) is *implant loading conditions,* i.e. the need for an undisturbed healing period of 3 months for lower and 6 months for upper jaws. Today, the actual healing times have been somewhat reduced for routine cases (Schnitman et al., 1997; Ericsson et al., 2000; Roccuzzo & Wilson, 2002; Cochran et al., 2002; Bischof et al., 2004). However, even micro movements during the primary healing period of the implants are still not accepted (Gapski et al., 2003; Cochran et al., 2004).

Implant characteristics

With regard to implant characteristics, three aspects were identified by Albrektsson et al. (1981), starting with *(1) the implant material* - where the authors recommended the use of pure non-alloyed titanium, followed by *(2) the implant design* – for which they suggested use of threaded implants to minimize implant movements and to enhance the implant surface area. The third implant related variable mentioned was *(3) the implant finish,* where a turned surface was recommended in order to improve cellular contacts with the implant. The ideal surface was, therefore, identified as - *neither polished nor too rough*.

Skill of the clinicians

Of course, all the technique and patient-selection related aspects may be associated with the experience and skill of the treating doctor, and reports indicate that a learning curve exists as doctors start to use oral implant treatment protocols (Lambert et al., 1997; Albrektsson, 2001).

Consequently, the bone status of the implant sites, the skill of the clinicians and the biocompatibility of the components used are all factors that may influence the treatment result and may thus also be possible prognostic risk factors for the clinical outcome. Therefore, these factors, one by one or together, ought to be assessed when performing result evaluations. However, most authors routinely assess only the implant related factors when performing studies on oral implant systems – i.e. - *(1) the implant material - (2) the macro design of the implant and - (3) the finish or micro design of the implant surface* (Piattelli et al., 2002; Bolind et al., 2005; Rasmusson et al., 2005; Shalabi, 2006).

Introduction of concurrent oral implant systems

At about the same time as the Brånemark procedure became public, researchers in Germany and Switzerland also presented concurrent implant innovations (Albrektsson and Wennerberg, 2005). The development of the Frialit technique,

a German implant system (Tubingen, Heidelberg, Germany), was begun in 1974 by Professor Schulte (Schulte, 1981). This system has become known for its root shaped implants with step wise increasing diameters, from the beginning produced in a ceramic material (Al_2O_3) (Schulte & Heimke, 1980). The system has been further developed for more than two decades since its introduction. Today Dentsply International Inc. (Pennsylvania, US), the current owner of the Frialit implant system, markets three different implant lines in titanium (Ankylos, Frialit, Xive), of which two are derived from the original Frialit implant. The Ankylos implant system has a progressive thread-design, while Frialit and Xive still have the root-analogue form combined with an internal abutment locking connection (Dentsply home page, 2007-01-31).

In 1980, an implant system introduced by the Institute Straumann AG (Basel, Switzerland) became publicly available (Schroeder et al., 1976). The system was developed and introduced by Professor Schroeder in collaboration with the ITI research institute in Bern. The Straumann oral implant system was introduced as the first one-stage implant system, eliminating the second surgical phase. The implants have a 2-3 mm band of smooth surface intended to be placed just caudal to the soft tissue margin (Straumann home page, 2007-01-31). The implants initially had a TPS-surface, which was later replaced by a SLA-surface, both of which have been extensively documented (Eckert et al., 2005). However, a new surface texture called SLA-active is currently used for three ITI implant-design options (Standard, Standard Plus and Tapered) (Straumann home page, 2007-01-31). The Institute Straumann and Nobel Biocare are presently the leading implant manufacturers on the market (Millenium report, 2005).

Nobel Biocare AB (currently Nobel Biocare Holding AG, Zurich, Switzerland), the manufacturer of the original Brånemark implant and previously known for its conservative approach and extensive clinical documentation, has recently developed a wide range of new implant lines. The classical two-stage threaded design with an external hexagonal connection has partly been replaced by new designs like Nobel Direct; a one-piece implant sculptured directly in the patients mouth and loaded immediately (Nobel Biocare home page, 2007-01-31).Via purchase in the mid-nineties Nobel Biocare also introduced a completely new implant line called Nobel Replace (Sullivan, 2001). This system has an internal hexed-abutment connection and is also available in a range of designs such as tapered, straight or with the marginal part sculptured (Emms, 2007). Regarding surface texture, Nobel Biocare has moved from their original turned-surface, pure titanium grade 1 (ad modum Brånemark) to a current rougher one called TiUnite, titanium grade 4 (Rocci et al., 2003; Friberg et al., 2005).

During the eighties several more implant brands were introduced, not all of which became successful and lasting and some even had to be removed from the

market, such as the IMZ (Fredensfeld, Germany) and Core-Vent (California, USA) (Esposito et al., 1997). These two implant systems had a hollow basket form instead of the full threaded design, and the IMZ implant was HA coated (Esposito et al., 1997). Unfortunately, this surface treatment did not withstand the demands of long-term bonding of the implant to the surrounding bone. Instead, the surface material frequently fractured and subsequently caused infections in the surrounding tissues and resorption of the bone (Esposito et al., 1997). The Core-Vent design also induced similar bone reactions and with time these implants often tended to fail (Malmqvist & Sennerby, 1990).

Astra Tech AB (Mölndal, Sweden) started a clinical research program in the mid-eighties to document their implant system, known for its internal conical connection between implant and abutment, which is called the conical seal design (Astra Tech home page, 2007-01-31). The system has a roughened implant surface called TiOblast, with over 10-years of clinical follow-up experience (Rasmusson et al., 2005). However, the Astra Tech company has also introduced new features in their components; such as micro threads and a connective contour design, which are aimed to enhance the marginal bone profile (Astra Tech home page, 2007-01-31; Rasmusson et al., 2001). The new implant surface texture OsseoSpeed, a fluoride-modified surface, is gradually replacing the TiOblast-surface. Furthermore, Astra tech has recently redesigned the implant-abutment connection and the company is therefore in a transitional phase regarding compatibility between older and new components (Astra Tech home page, 2007-01-31).

The basic principles of the Brånemark system have been adopted by many other implant systems (Misch & Misch, 1992; Lazzarra et al., 1996; U. S. Food and Drug Administration, 2007). For instance, 3i (Implant Innovation inc., Florida, USA) and Lifecore Dental (Lifecore Biomedical inc., Minnesota, USA), both companies on the top ten list by market share (Annual industry report, 2005), introduced generic products. They are designed to mimic the original Brånemark implant system, referring to the long-term clinical documentation available for turned and threaded implants (Adell et al., 1981; Ekelund et al., 2003; Lekholm et al., 2006). However, both companies now also have their own development of new oral implant systems with company specific macro and micro designs (surface technique; Lifecore Biomedicals home page 2007-01-31; Implant innovations home page, 2007-01-31).

Consequently, there is currently a broad variety of manufactures and implant systems on the market. Many of these are, however, plain copycats referring to old documentation of generic techniques. Other systems are constantly introducing components with new features and thus miss the link to long-term clinical documentation of previous implant lines. The need for sound evaluation

of new components and systems is of course obvious and has to be performed in a structured and scientific way, including a correct statistical handling of the long-term follow-up data.

Application possibilities

From the beginning (1965-1980), oral implants were mainly used to treat totally edentulous patients to give them better chewing and speech function and improve their social lives (Brånemark et al., 1977; Albrektsson et al., 1987). The most common condition to be treated with rigid fixed bridgework was in mandibles (Lindquist & Carlsson, 1985). Maxillae were of course also treated with the same type of prosthetic rehabilitation (Taylor, 1991; Desjardins, 1992). In some cases, the rigid full fixed bridges were replaced by a more cost effective overdenture treatment protocol, often requiring only 2 implants (Zitzmann et al., 2006). In the Netherlands, overdentures even became the only implant treatment procedure supported by their national dental health insurance system (Cune et al., 1997). However, when overdentures were used in upper jaws, typically presenting poor bone qualities and small jawbone volumes, the success rates were often less favourable (Jemt el al., 1992; Hutton et al., 1995; Jemt et al., 1996). Still, the successful treatment outcome of totally edentulous and orally handicapped patients in time further opened up the treatment panorama.

Therefore, dentists around the world started to use osseointegrated implants to treat all types of edentulism, i.e. from single to multiple tooth loss (Lorenzoni et al., 2003; Levin, 2006; Palmer, 2000; Lekholm et al., 2006). In single tooth replacement, the patients are often young and their treatment result must have a lifetime perspective. When using single implants in the incisor region the aesthetic is of course also very demanding (Kourtis, 2007). In partially edentulous jaws, where more than one tooth has to be replaced, the implants are generally connected by a fixed framework. Such constructions can be either screw retained or cemented (Keith et al., 1999; Uludag & Celic, 2006). A method by which partial fixed bridges could be supported by connection to natural teeth has also been described; however variable outcomes for both the teeth involved and the implant bridges have been reported (Åstrand et al., 1991; Schlumberger, 1998).

In cases with insufficient jawbone volume for implant placement, additional advanced surgery may be required using various grafting protocols prior to implant placement (Adell et al., 1990:a; Köndell et al., 1996; Kahnberg et al., 1989, 1999). The grafting can be performed as a one- or two-stage procedure, depending on the severity of the bone resorption and the health of the patient (Lundgren et al., 1997; Nyström et al., 2002). Recently, a method has become available as a complement to bone grafting, using oversized implants (Zygomaticus implant, developed by Nobel Biocare AB, Göteborg, Sweden)

placed in the zygomatic bone, a few conventionally sized implants placed within the alveolar bone, and a fixed construction (Higuchi, 2000; Brånemark et al., 2004; Malevez et al., 2004).

Finally, implants can be used in orthodontic treatment, where secure fixation is required in order to move misaligned teeth without displacing those already correctly aligned (Oncag, 2007). The "Onplant" implant system, developed and introduced by Nobel Biocare AB (Göteborg, Sweden), was designed for such treatment purposes (Ödman et al., 1988).

Consequently, oral implants are today used over a broad range of applications which must be individually and scientifically assessed before the outcomes can be claimed to be safe and predictable.

Follow up aspects

Guidelines for the performance and evaluation of clinical trials have been available and used for decades (Pocock, 1983) for introducing pharmaceutical products. However, the evaluation procedure, used for pharmaceutical products, is not suitable for oral implants; one of many products within the family of medical devices. Instead detailed evaluation protocols are needed in these cases to give researchers and authorities useful comparable data. The first attempt to introduce such a protocol was at a Harvard consensus meeting in 1978 (Schnitman & Schulman, 1980).

Basically, study design and statistical evaluations are considered of utmost importance in all clinical research, whether for pharmaceutical or biomedical products. The "gold rule" demands a clinical trial to be: controlled, double-blinded, randomised and prospective (Pocock et al., 1983; Herrmann, 2003). However, when evaluating medical devices such as oral implants, the clinical testing was not very demanding until the Medical Device Directive of 1993 (Liedström, 1995); the Food and Drug Administration's (FDA:s) stricter policy of registration applications (510(k) or PMA) was also introduced in the 90's (U. S. Food and Drug Administration home page 2007-02-02). The authorities are still quite liberal, when it comes to clinical documentation on oral implants. A full quality assurance certification (Liedström, 1995) gives the manufacturers the right to put a device on the market without any new data, if generic design and treatment modality can be referred to. However, the certification is normally only based on the production of the devices rather than on the clinical performance and the outcome of the treatment. Furthermore, only when the manufacturer is claiming a superior design or completely new features, laboratory or clinical studies are required (U. S. Food and Drug Administration).

Consequently, the decision to document new products or changed features of new components today seems to be made by manufacturers and/or researchers mainly randomly. Unfortunately, this means that risks of severe clinical problems may only be exposed after the products have been introduced onto the market and after placed in patients, which of course is not acceptable.

Study designs

Study designs can be divided into retrospective and prospective protocols, with or without using a control group. *Case reports,* the least advanced method of presenting oral implant results, have been and are often still used when introducing new methods and components (James et al., 1974; Dilek et al., 2007). This way of presenting outcomes can be valuable for example when presenting a complication or a solution to a problem (Strietzel et al., 2006). The Journal of oral and maxillofacial implants (JOMI) is one of the peer-reviewed journals that also present case reports in almost every number (Bousdras et al., 2006). However, for presenting results of a treatment procedure *per se*, case reports are of very restricted value.

Retrospective studies, based on the experience of one clinic using a new treatment protocol or component, were frequently published in the eighties (Cox et al., 1987; Adell et al., 1981; Dalise, 1988). This type of study design is seldom used today because it is statistically untidy. Instead, *retrospective multicentre-studies*, where more than one clinic has been responsible for the treatment and evaluation of outcome, are currently being published (Ko et al., 2006; Misch et al., 2006). The disadvantage with retrospective studies is that different follow up times have often been used by the different centres, as some patients may have been treated years ago while others have been treated just recently (Babbush & Shimura, 1993). Subsequently, the evaluation time may be long and the patient population large at the start, whereas the number of patients evaluated at the end of the full time period might be low. Therefore, the cumulative success rates for the longest follow-up periods may be based on just a few percent of the entire study population (Ferringo et al., 2002); which of course is not acceptable.

Prospective multicenter-trials, in which all clinicians follow a strict study protocol with well-defined success criteria and the statistical evaluations decided from the start, became a “gold standard” for oral implants when introduced in the 90´s (Lekholm et al., 1994; Jemt et al., 1996; Henry et al., 1996; Friberg et al., 1997). Such studies can also be used to identify risk factors for variables tested, such as bone quality and quantity (Lekholm et al., 1994; Jemt et al., 1996; Friberg et al., 1997). When comparing different implant systems, treatment modalities, or side effects such as marginal bone-loss, it has been recommended to use a controlled, randomised and prospective multicenter-study

design (Olsson et al., 1995; Örtorp & Jemt, 2006; Iacono & Cochran, 2007). However, to compare the success rates of two different implant systems would require thousands of patients, since the expected success rates may be 95-99 % (Pocock, 1983). For natural reasons, double-blinded studies are not an easy option in these situations.

In order to identify success rates and to establish reasons for failures, a few researchers have tried to combine the outcomes from different reports using *meta-analyses* (Esposito et al., 2001a; Shalabi et al., 2006; Proskin et al., 2007). However, problems with these techniques do occur - when the study design, success criteria and statistical methods have not been the same for all the studies included in the evaluation.

Consequently, studies on oral implants most often use a diversity of study designs and statistical methods (Babbush et al., 1986; Adell et al., 1990:b; Arvidson et al., 1992; Buchs et al., 1995; Wheeler et al., 1996; Lazzara et al., 1996; Åstrand et al., 2002; Esposito, 2001a; Shin et al., 2006; Jemt & Hager, 2006). The number of oral implant reports is also rapidly increasing (2838 in PubMed 2006-12-03), but only a few of these are based on 5 years or longer follow-up periods (Heydenrijk et al., 1998; Ekelund et al., 2003; Eckert et al., 2003; Attard & Zarb, 2004; Hallman et al., 2005; Lekholm et al., 1999, 2006). The purpose of every study should be clearly stated in advance, and the purpose should subsequently influence the choice of study design and statistical techniques used to enable reliable and correct assessments of the data.

Success criteria

The intention when placing implants into the body is, of course, that they should last a lifetime. Therefore, it is obvious that before implants are placed it should also be known that they will function for such a long period. Consequently, there is a need for establishing how to measure the success of new oral implant systems before their introduction to the market. An early attempt to present some basic guidelines for success criteria for oral implant treatment was suggested by Schnitman & Schulman (1980). Other researchers (Smith & Zarb, 1989; Buser et al., 1997) have subsequently presented their own criteria, of which those of Albrektsson et al. (1986) have become the most frequently used.

The Albrektsson success criteria are summarized as follows:

- an individual, unattached implant should be immobile when tested clinically.
- a radiograph should not demonstrate any evidence of peri-implant radiolucency,
- the vertical bone loss should not exceed 0.2 mm annually, following the implant's first year of service.
- no signs or symptoms such as pain, infection, neuropathies, parestesia, or harm to the mandibular canal, should be reported.
- an overall success rate for the system should be at least 85% after a 5-year observation period and 80% after 10-years.

Lekholm et al. (1994) suggested some minor adjustments to the Albrektsson success criteria, in order to make them suitable for individual implant evaluations. Basically, the modified criteria focused on the definition of failures as follows: implants found to be mobile or affected by persisting and incurable soft tissue or mechanical problems, should be regarded as failures. An implant was also considered as a failure when the marginal bone loss reached the apical third of the implant during the follow-up period. Other factors, such as interrupted marginal bone resorption, persistent pain, restorative complications and/or paresthesia were reported separately, but without influencing the implant success rate *per se*.

A further development of the Albrektsson and associates success criteria has been suggested by those authors (Albrektsson, 1993). They allocated the outcome of the implants into four groups: success, survival, failure or "unaccounted for". By doing so, all implants could be individually characterized and evaluated.

Consequently, it is important that success criteria are defined and agreed upon before any follow-up study is performed, as this may facilitate the researchers to comply with the study protocols and may enable comparison between different study outcomes later on.

Objectives of different statistical methods

In order to confine the scope of this thesis, the statistical terms and analyses commonly used in oral implant treatment studies are reviewed.

Descriptive statistics

All statistical questions have aspects to be considered before choosing the most appropriate evaluation model. Mean and standard deviation (spread) have often been referred to as sufficient and helpful statistics to describe the distribution of variables assessed (Altman, 1991; Gellerstedt, 2006). Defining a distribution for a specific variable will, however, help to choose between a parametric or non-parametric statistical model. Parametric statistics require that the data be normally distributed or similar, and are generally considered to be more efficient; to give greater power to the analyses. Non-parametric methods are more robust and can, on the other hand, be used on all types of distributions; but are generally regarded as more complex procedures. In order to select the correct statistical method, it is also important to determine which type of measurements needs to be collected. Parametric and correlation methods are preferably used in analysing quantitative and numerical measurements, while qualitative data, i.e. nominal and ordinal measurements, demand a non-parametric evaluation (Siegel & Castellan, 1988). Furthermore, data without internal order of the variables would be best handled via contingency tables (Altman, 1991), while data with any type of order should be evaluated by using rank order and/or correlation methods (Kendall, 1962).

Dependent or independent data

Several statistical tests used on oral implants have been based on the hypothesis that the variables tested are independent. The statistical method should instead be selected after determining weather the variables are independent or not. Methods to test for dependence such as Kendall's rank correlation (1948), can provide a distribution free test of independence and measure the strength of dependence between two variables. Testing a null hypothesis of independence between two variables, such as with Spearman's rank correlation (Best & Roberts, 1975), could be another alternative. These two correlation methods were originally designed for quantitative continuous variables (Lehmann & D'Abrera, 1975; Holm, 2007), but the only requirement is in fact that the data exhibits ordering to some extent.

The outcome of dental implant studies is usually based on the number of implants inserted, failed, withdrawn or judged as successful. Each patient could have had 1 - 7 implants inserted from the start, and consequently everything from zero to several implants could have failed during the follow-up period. The different numbers of implants at commencement and different numbers of failed implants make a dependence calculation complex (Holm, 2007). Furthermore, variables in oral implant treatment are often qualitative data displaying a categorical distribution, and the measurements have no obvious ordering, i.e. they represent a nominal condition. Consequently, evaluations of dependence

between oral implants are not easily performed with either of the two dependence tests mentioned (Holm, 2007).

To select the correct statistical method for evaluating the outcome of oral implant treatment, the purpose of the actual study has to be considered and clarified. Three different methods are eligible: survival data analysis; comparison of groups; or multivariate analysis.

Survival data analyses

Follow-up studies on oral implants, evaluated with survival data analyses, could analyse the number of successful subjects (patients) but could also give information on when failures occur and give a prognosis for the implants tested (Cutler & Ederer, 1958.). The assumption of survival analysis (Cox & Oakes, 1984), both for Kaplan & Meier (1958) and log rank tests (Bewick, 2004), is that patients withdrawn or excluded will have the same prospect of success as those who continue in the study. Patients who are followed for a short or a long period of time will also be expected to have the same survival probabilities, independent of when events happen, i.e. no time dependence. However, when the number of patients has declined, late failures will adversely affect the result more than early failures (Herrmann, 2003).

Life-table models

Life-table analyses, a development of the classic survival curve, are used to give a prognostic figure for the survival of a certain treatment protocol (Cutler & Ederer, 1958). The treatment could for example be the replacement of lost teeth with dental implants. The analysis then accounts for patients (implants) who have died (failed) or have been withdrawn (censored). A life-table presents the proportion of surviving, the cumulative success rates, and the proportion of "patients at risk". Subsequently, all events are supervised over time. A life-table can also be used to compare two groups of treatments by adding standard errors and calculated confidence intervals (Altman, 1991). The total success rates are not compared, but rather the success rates at preset time intervals. However, this outcome can be somewhat confusing if treatments give a better success rate at one preset time period but not at another.

Log rank test

A better alternative to the life-table technique is a log rank test (Azen et al., 1977), when survival analyses are performed to compare two (or more) test groups. The log rank test is based on the null hypothesis that there is no difference between the groups compared. Furthermore, one of the survival curves is compared with the expected survival curve of the whole time span. The difference in survival curves will be described by *P*-values to ascertain any significance between the compared groups. The log rank test is often used to

detect a difference between test groups when the risk of an event is consistently greater in one group than the other, and the relative risk does not cross over at any time period. Therefore, when analyzing survival data using the log rank test, the survival curves should always be plotted first (Gellerstedt, 2006). As the log rank test is purely a test of significance, it does not provide an estimate of the difference between the groups or a confidence interval. An alternative method would be to use a hazard ratio analysis, such as the Cox proportional hazard model (Altman, 1991).

Contingency Tables

When the purpose of a study is to compare groups without considering the time aspect, contingency tables (Altman, 1991) could be the method of choice. In connection with survival data, contingency tables are also used to compare test groups at a fixed time-period. The tables are then referred to as frequency tables. Each factor (dimension) is divided into a number of levels, and the frequencies of their combinations are registered. Two different correlation tests are available:

Pearson's chi-squared test

The Pearson's chi-squared test (Altman, 1991) is basically the sum of squares of the differences between the observed and expected frequencies. Then, each squared difference is divided by the corresponding expected frequency. The test is normally used when more than two research groups are to be compared.

Fisher's exact test

The Fisher's exact test (Altman, 1991) can be used when only two data groups are compared using two by two contingency tables. Furthermore, the test calculates the actual probability of the observed two by two contingency tables, with respect to all other possible two by two contingency tables within the test. The sum of the probabilities for more extreme cases will then be the *P*-value.

If the purpose of a study, on the other hand, is to detect specific factors responsible for the failures, the problem is more complex due to the nature of the data. The methods used are multivariate or logistic regression analyses which however are still not designed for this kind of complex condition and have therefore some limitations (Holm, 2007).

Multivariate analyses

Multivariate analysis techniques (Marda et al., 1979) are now commonly used to evaluate the outcome of oral implant treatment (Wennström et al., 2004; Noguerol et al., 2006). These techniques are normally used when several variables are present, which are presumed to give a cumulative effect on the outcome. The variables may interact independently or may be dependent on each other. However, multivariate analyses are mainly complements to

univariate tests (Herrmann, 2003) for which multiple interferences and the risk of cumulative errors might add up to a total enlarged risk of $P>0.05$.

In order to identify which group of background factors may have the strongest effect on the outcome of a study, a step-by-step multivariate analysis could also be used (Herrmann, 2003). In connection with this, a post hoc analysis might be used to further determine which specific variable is responsible for the significant outcome at each level studied (Hochberg & Tamhane, 1987).

Statistical terms and tests used in oral implant studies

As seen from the above review some improvements in statistical handling seem to be needed, when evaluating oral implant treatment results (Herrmann, 2003).

Descriptive statistics

Descriptive data for oral implants could be most elusive (Altman, 1991). In many medical and dental publications the description of the population studied, with regard to hypotheses used and distribution of the population, is often lacking. However, number of females and males and age ranges are information frequently presented in studies of oral implant outcomes (Buser et al., 1990; DeBruyn & Collaert, 1994; Chuang, 2002).

Dependent or independent data

Several authors have acknowledged dependence, while many others have ignored it completely (Bahat, 1993; Buchs et al., 1995; Engquist, 2002; Romeo et al., 2006). The existence of a presumed dependence among implants within the same jaw is obvious once a prosthetic construction has been attached, a fact first recognised by Adell et al. (1990:b). Statistical analysis of the outcome can, of course, be started at "the prosthetic level" as suggested by many researchers (e.g. Adell et al., 1990:b; Jemt et al., 1993; Leimola-Virtanen et al., 1995). However, starting at this level with follow-up evaluations should be avoided since all early implant failures would be excluded from the analyses. The first researcher to randomly select one implant per patient for calculating the statistical outcome on "implant level", and correctly taking care of a possible dependence, was Mau (1993). His method has since been followed by others (Cune & de Putter, 1996; Haas et al., 1996). A different way of handling dependence using regression analysis has been recommended by Hutton et al. (1995) and others (Eckert & Wollan, 1998; Eckert et al., 2001; Chuang et al., 2002). Recently other methods, such as the Bootstrap or Jackknife techniques, have been seen in studies on basic medicine (Brunelli & Rocco, 2006; Liu et al., 2004; Chen et al., 2006). Only one study on dental implants using the Jackknife technique (Yerit et al., 2006), but none with Bootstrap, has been found in the PubMed database (2007-01-10). However, the statistical method as such is not

discussed as the only information available is that a Jackknife variance estimate was used to handle the correlation of implant survivals within the same patient.

Comparisons

Statistical methods for comparing test-groups are available (Altman, 1991). However, sometimes each CSR may be calculated separately before being compared (Chiapasco et al., 2001; Pinholt, 2003) in order to identify risk factors or to evaluate individual implant systems.

In meta-analyses (Hunter et al., 1982), for which data is combined from different studies, dependence amongst the original sample populations may not have been fully evaluated (Esposito et al., 2001:a; Hinode et al., 2006). Some outcomes included in the meta-analysis may, for example, have been calculated with methods handling the dependence while others have not (Esposito et al., 2001:a; Henry et al., 1996). Furthermore, even if dependence has been acknowledged, the statistical methods used may not always account for such dependence.

Life-table analyses

Life-table analyses, such as described by Kaplan & Meier (1958) or Cutler & Ederer (1958), are often used when evaluating medical devices such as oral implants (Lekholm et al., 1994; Lazzara et al., 1996; Jemt et al., 1996; Lekholm et al., 1999; Shin et al., 2006; Romeo et al., 2006). However, it is important to acknowledge that life-table analyses are designed to evaluate prognoses for fatal diseases, as the name implies (Cutler & Ederer, 1958). Therefore, presenting a CSR without exploring the data in the form of a life-table (e.g. Buchs et al., 1995; Levine et al., 2002) has clear disadvantages. Many five or ten year reports have also been based on only a few patients followed the full time-period (Lazzara et al., 1996; Wheeler et al., 1996; Testori et al., 2001; Ferrigno 2002). Detailed information on the patients withdrawn or dropped-out is most often also missing when the outcome of retrospective studies are presented (Misch et al., 2006; Babbush & Shimura, 1993).

Log rank test

In a retrospective study of patients treated with endosseous implants, comparisons between subgroups of patients were made using the log-rank statistical test (Mundt et al., 2006). In another study, immediately loaded implants and submerged implants were compared (Schnitman et al., 1997), also using this test. The log rank test has the advantage of giving a *P*-value which is often desired by the researcher, especially in medical and dental research.

Withdrawal analysis
In long-term follow-up studies, some patients will be withdrawn or dropped-out before the final examination. Albrektsson (1993) described these as "unaccounted for" patients and also showed that they exhibit approximately twice as many failures as patients who remain within the study. Furthermore, he found that withdrawn or dropped-out patients were not interested in being re-evaluated or were not healthy enough to be re-examined. The reasons for patient withdrawals may be documented (van Steenberghe et al., 1990; Jemt et al., 1991; Johns et al., 1992; Lekholm et al., 1999) or not. However, the patients withdrawn influence the treatment outcome even when the number "unaccounted for" is relatively low (Albrektsson, 1993).

Lekholm et al. (1994) described a technique based on the assumption that all implants not eligible for check-up were failures, i.e. worst-case analysis, instead of assuming that they had the same fate as those being followed. The purpose of their protocol was to show that there is a discrepancy between patients/implants evaluated and the statistically calculated CSR. Later, Mau et al., (2002) developed the worst-case protocol further by introducing one more level called the best-case analysis, i.e. the opposite, assuming that all implants not being followed were successful. Thereafter, they compared the worst or best-case assumptions, demonstrating the uncertainty that may occur when too few patients were followed.

Considerations on commonly used implant statistics

From the above review it is clear that the statistical methods currently being used in studies of oral implant treatment are not always the most appropriate to ensure valid results. Important statistical considerations include:

- A dependence between implants placed within the same jaw has not yet been statistically proven to exist before the implants are connected via a fixed construction,

- Neither is it known how the outcome of studies which ignore dependence relate to the results of studies addressing dependence and handling the data accordingly,

- Survival analyses are normally based on the assumption that withdrawn or censored patients (implants) would have the same fate as those being followed. However, it is still not fully known how the fate of withdrawn or dropped-out patients/implants relates to the outcome of those followed,

- The reasons (selective or random) for withdrawals or exclusion of patients/ implants, and also the status of these patients/implants, are most

often not discussed in oral implant reports. It has not been shown either whether the result would be interpreted differently, if more information on withdrawn or dropped-out patients/ implants were known,

- Neither has it been shown whether the risk of implant failure is related to patient status, the type of implant or the treating doctor – or a combination of these factors,

- Neither is it known whether correct statistical analyses can before treatment identify patients with an increased risk of implant failures,

- Several generally different methods for handling dependence are available but there is little information on which would be preferable to display implant results. The particular situation of unequal numbers of implants per patient requires special statistical methods to handle the variance calculations,

- Finally, it does not seem to be known whether ignoring dependence would affect variances, confidence intervals and *P*-values, calculated for statistical analyses.

The importance of all these factors has not been clarified, and consequently more research on these topics is needed and will be the theme of this thesis.

Aims of thesis

The overall objective of this thesis was to study how different statistical tests and procedures affect the evaluation of oral implant treatment results. The specific objectives of the four investigations included within the thesis were to assess, in relation to oral implant treatment:

- whether any dependence exists among individual implants, before connecting the implants to a bridgework (study I),

- whether the outcome of life-table analyses is affected by the random selection of one implant per patient (study I),

- whether the success rates differ for followed patients or those withdrawn or dropped-out from the studies (study II),

- whether there is a threshold for the number of randomly withdrawn or dropped-out patients (implants) still allowing valid analysis (study II),

- whether random versus selective withdrawal of patients (implants) exerts any influence on the statistical outcome (study II),

- whether any individual patient, clinic or implant characteristics influence the implant failure rate (study III),

- whether any combinations of the above individual characteristics influence the outcome and could thereby identify high risk patients (study III),

- whether a Jackknife method, in conjunction with confidence intervals and life-table analyses, reveals the same bone-combination (i.e. high risk patient factors) as in study III (study IV),

- whether any combinations of the above individual characteristics display a statistical difference (i.e. high risk patient factors) when using the Jackknife method in conjunction with a log rank test (study IV).

Materials and methods

Original multicentre population

Four 5-year prospective multicentre (MC) studies (Lekholm et al., 1994; Jemt et al., 1996; Henry et al., 1996; Friberg et al., 1997) of oral implant treatment (Brånemark implant system by Nobel Biocare AG) concerning the success of turned threaded titanium implants constituted the base for the current research project. The original purposes of the four investigations were to analyse separately the specific treatment protocol by evaluating all implants at regular intervals, until the prostheses had served in clinical function for 5 years. All studies followed a similar program of research and follow-up protocols. Totally edentulous patients were restored with either fixed full dentures or overdentures, while partially edentulous patients were treated with either partial replacements or single tooth crowns, depending on the number of teeth missing. The four studies also followed the guidelines for clinical research extant at the time, including the Declaration of Helsinki.

All patients signing up at selected clinics during a pre-specified time-period under rigorous inclusion and exclusion criteria were consecutively included in the studies. The patients also had to be willing to comply with the treatment protocol in question and the intensive follow-up program, altogether taking about 6 years. According to dental status and treating centre the patients were enrolled for the corresponding treatment protocol. Several questions about the overall health of the patients focused on conditions known or suspected to influence the treatment were noted. However, details of smoking or alcohol usage were not registered. At the start of the studies, all included patients were reported to be mentally and physically healthy and when appropriate to have controlled medications.

The patient inclusion period was approximately 6 months for each study. The first surgical session, implant placement, was performed according to routine procedures (Adell et al., 1985; Lekholm, 1993). Second stage surgery, abutment connection, took place after a healing period of at least 3 months for lower jaws and 6 months for upper jaws. The prostheses were thereafter attached within a month, again following standardized protocols (Adell et al., 1985). However, postponements did occur at patients' request, delaying the treatment phase, so baseline was set when the prostheses were in position.

Table 1
Distribution of patients by MC-study and treatment outcome.

MC- study	Patients			Supra-structures at start
	Included	Withdrawn/ dropped-out	Successful	
Fixed partial bridges	159	26	133*	197
Overdentures	133	41	92	127
Single tooth Replacements	92	17	75	106
Fixed full bridges	103	20	83*	101
Total	487	104	383*	531

* A total of 13 bridges were not removed, but the supporting implants were still considered successful

Fifty -five percent of the patients included were females and the mean age of the entire group was 51 years (range 15-84 years). Two hundred and fifty-one patients had one or several missing teeth needing replacement, while the remaining 236 patients were completely edentulous from the start. Of the patients treated, 57% received implants in their lower jaw. The total number of patients, suprastructures and outcome per study can be seen in Table 1.

During the 5-year follow-up period after attachment of the prosthetic suprastructures, the patients were recalled at preset time-intervals. The patients included in three of the studies were examined after 1 week, 1 and 6 months, and thereafter yearly. Patients in the fixed full bridge study had their first evaluation after 1 year of function. Clinical examinations were combined with radiographic evaluations at baseline and after 1, 3 and 5 years. Radiographs were used (Gröndahl et al., 1996) as an objective and repeatable confirmation of the implant status during the entire study period. To further ensure objectivity, independent dental radiologists (at the University of Gothenburg) assessed marginal bone heights of each implant and also evaluated whether any signs of pathological changes were present.

Whether the implant treatment could be regarded as successful or not was, in principle, based on Albrektsson et al. (1986). However, minor modifications according to Lekholm et al. (1994) and Herrmann (2003) were also introduced, without altering the demand of implant stability, preset maximum of marginal bone-loss and absence of complications or pathogeneses.

To fulfil the criteria at the final 5-year examination most suprastructures, except single tooth restorations, were removed to be able to individually test the stability of each implant. A total of 13 suprastructures in the fixed full and fixed partial bridge studies were not removed but the supporting implants were considered successful due to the absence of clinical or radiological signs of integration loss. Another 3 patients suffering from parestesia were also regarded as successful implant outcomes. Besides the implant success variables, plaque index, gingival index, probing depth, bleeding on probing and stomatognatic function were also assessed. Prosthetic component failures were also noted during the follow-up period. Thus, all complications from first stage surgery and throughout the entire treatment period were included in the statistical evaluations.

Details of the treatment and follow-up protocols of each MC-study were described in the individual 1, 3 and 5-year reports: fixed partial bridges (van Steenberghe et al., 1990; Henry et al., 1993; van Steenberghe et al., 1993; Gunne et al., 1994; Lekholm et al., 1994; Higuchi et al., 1995), overdentures (Johns et al., 1992; Hutton et al., 1995; Jemt et al., 1996), single tooth replacements (Jemt et al., 1991; Laney et al., 1994; Henry et al., 1996) and fixed full bridges (Olsson et al., 1995; Friberg et al., 1997).

Pooled multicentre study population

The patients from the four multicentre studies, 487 patients with 531 suprastructures and 1738 implants, were pooled to create a combined database for statistical analyses (Tables 1 and 2). Among these patients, 323 implants were withdrawn or dropped-out from the study before the final check-up, and an additional 110 implants were regarded as failures, giving in total 1305 successfully functioning implants (Table 2).

Table 2
Distribution of implants by MC-study and treatment outcome.

MC- studies	Patients	Implants			
		Inserted	Failed	Withdrawn/ dropped-out	Successful
Fixed partial bridges	159	558	36	84	438
Overdentures	133	510	44	127	339
Single tooth replacements	92	107	3	18	86
Fixed full bridges	103	563	27	94	442
Pooled population	487	1738	110	323	1305

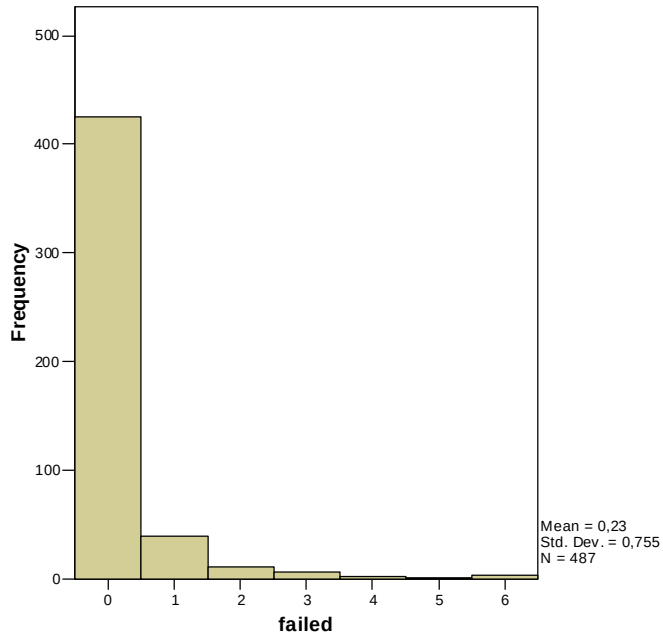
Furthermore, of the 487 patients enrolled, 104 failed to complete the entire follow-up program. The main reasons for this are presented in Table 3. Primarily, 73 of the 104 patients did not complete the studies due to death/hospitalization or non-compliance during the follow-up period, including persons refusing radiographic and/or clinical examinations at the final follow-up. Thirty-one additional patients did not complete the study, whereof 25 were considered to have a negative change of therapy. Twenty patients experienced complete failures of their initial implant rehabilitation and consequently had to be retreated with conventional prostheses, while another five had involuntarily changed prosthetic solutions. Finally, six more patients had to be excluded, since they had requested a new prosthetic rehabilitation not corresponding to the originally planned treatment. The majority of these were overdenture restorations in which a fixed full bridge had to replace the original implant supported denture.

Table 3
Distribution of reasons for withdrawal of patients within the four MC-studies.

MC-studies	Reasons for withdrawal of patients				
	Death or hospital-ization	Total failure	Non Comp-liance	Changed therapy (complication)	Changed therapy (requested)
Fixed partial bridges	9	7	8	2	0
Over-dentures	6	9	20	1	5
Single tooth replacements	2	1	11	2	1
Fixed full bridges	11	3	6	0	0
Pooled population	28	20	45	5	6

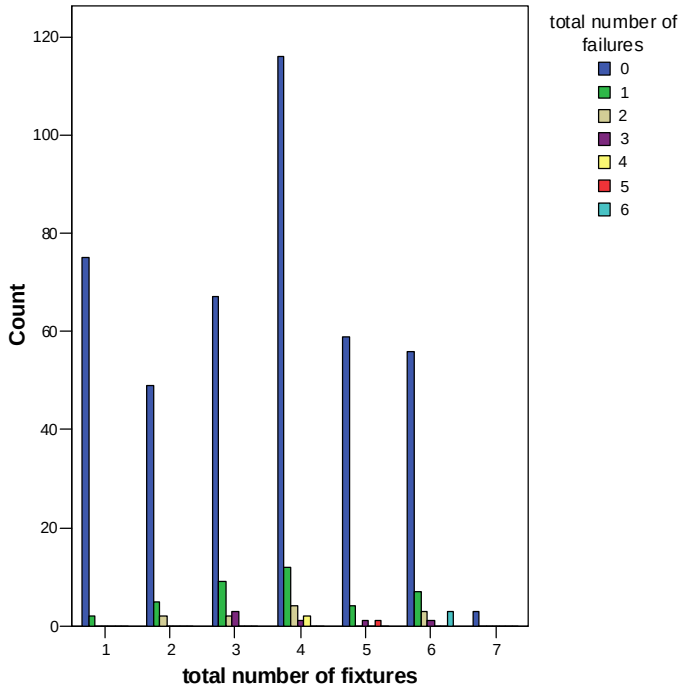
Thirteen percent or 62 of the 487 patients treated experienced implant failures, while the remaining 425 patients had no failures (Figure 5). The mean number of implants inserted per patient was 3.57 (1738/487) and the corresponding failure rate was 0.23 with a standard error of 0.755, representing everything from single to multiple implant failures in the patients followed.

Figure 5
Distribution of implant failures across patient population.



The number of failures against the number of implants placed per patient can be seen in Figure 6.

Figure 6
Distribution of implant failures against number of implants inserted.



Methods and statistical analyses used

Study I – Evaluation of dependence

Dependence between implants within the same jaw was calculated over the pooled population using the procedure described below. Single implants and patients withdrawn or dropped-out prior to abutment connection were excluded from this analysis, giving a total of 1639 implants for the dependence assessment. To evaluate any dependence among implants within the same jaw, before the implants were connected to a prosthetic construction, the following estimates were created:

- the independent variance (V_0) was estimated from a binomial distribution assuming no dependence between the implants.
- the dependent variance (V) was estimated from a general discrete distribution.

The following statistical formulae were then used to determine whether any dependence existed:

$$\mathbf{V_0 = mq\,(1-q),}$$

$$\mathbf{V = m^2\,(\text{å}\,y^2_i) - (\text{å}\,m_i^2)\,(\text{å}\,y_i)^2 / m^2 - (\text{å}\,m_i^2),}$$

where:

m = total number of implants,

q = frequency of failures

m_i = number of implants placed per patient

y_i = number of failed implants per patient (before loading)

y^2_i = squared number of failed implants per patient, respectively.

If $V > V_0$ - it was concluded that a dependence was substantiated.

In order to handle the dependent data, one implant per patient was selected. Manual randomisations to select one implant from each of the 487 patients were repeated five times using a randomisation chart (Pocock, 1983). Five unique "one implant per patient" groups were constructed, by beginning at different starting points on the chart. Individual life-tables (Kaplan & Meier, 1958) were then calculated for each of the five unique groups and a range of Cumulative Success Rates (CSR:s) for these were established. Finally, CSR:s were also calculated for the "all inserted implants/per patient" population, for comparison with the five individual CSR:s described above.

Study II – Analyses of withdrawn or dropped-out patients

A new "one implant per patient" population was constructed out of the pooled population for study II, by the previously described random selection technique (Pocock, 1983). However, information provided after the original multicentre studies and study I had been concluded was then included in the pooled

population. These updates did not change the number of implants in any of the groups – placed (n=487); withdrawn or dropped-out 16% (n=80); successfully functioning 76% (n=371); and/or failed implants 7% (n=36).

In order to study the effects of incomplete data, the following three different approaches were used to investigate how well the outcome of the patients withdrawn or dropped-out corresponded with the results of patients followed:

Originally withdrawn or dropped-out patients

To study if the withdrawn or "unaccounted for" patients had the same success rates as those evaluated - a hypothesis used in a standardized life-table analysis - attempts were made to again recall the "unaccounted for" patients. All former responsible investigators from the MC- studies were contacted and asked to re-examine previously withdrawn or dropped-out patients for the outcome of the implant treatment. If the patients did not return for a new check-up, the investigators were asked to find from their charts, if the patients had been examined regarding the implant outcome at any time after the MC-studies closed, and information from that examination was added to the pooled database. Repeated letters and e-mails requesting the information were sent to optimise the answering frequency.

Random withdrawals

To study whether there was a definite borderline percentage of the "one implant per patient" population that could be disregarded without affecting the outcome, 25% of the patients were successively and randomly withdrawn or dropped-out. Thus, three separate subgroups, called A, B and C, were created from the pooled population. Each random selection was performed starting from a new patient number in the current population to avoid selecting the same patients for all 3 groups. The same randomisation chart as mentioned before was utilized until only 75% (A); 50% (B); 25% (C) of the patients remained, respectively (Table 4).

Table 4
Distribution of random selections of the patients in study II.

Populations	Patients	Implants			
		At start	Failed	Withdrawn /dropped-out	Successful
Total	487	487	36	80	371
Group A	362	362	25	58	279
Group B	244	244	18	42	184
Group C	119	119	5	25	89

Life-table analyses (Kaplan & Meier, 1958) and chi-squared tests (Altman, 1991) were then used to study the implant outcomes of the randomly withdrawn or dropped-out patient groups A, B and C.

Withdrawals based on patient characteristics

In order to evaluate whether selected withdrawals based on patient characteristics interfered with the outcome, new subgroups were created out of the "one implant per patient" population. This time the patients were first selected by age, i.e. less than 51 years of age or over 59 years of age. Patients between 51 and 59 were disregarded during this selection to separate the younger and older patient groups. Secondly, the patients were selected by gender and jaw type. Thereafter, subgroups - for females or males and treatment in upper or in lower jaws - were created out of the pooled population (Table 5). Both Chi-square calculations and life-table analyses were then performed with the selected populations. However, CSR:s were not calculated for the age groups, since the middle age-group was disregarded.

Table 5
Distribution of selections of patient characteristics.

Patient		Implants		
		At start	**Failed**	**Successful**
Gender	**Female**	216	15	201
	Male	271	21	250
Age	**a <50**	211	12	199
	b 51-59	99	Excluded	
	c >59	168	16	152
Jaw	**Upper**	208	24	184
	Lower	279	12	267

Study III – Evaluation of potential prognostic risk factors

The pooled population (487 patients) of study II was also used as a basis for the analyses in study III. However, some minor changes of the database were then made to facilitate the separate statistical analyses, as described below for evaluating potential prognostic factors for oral implant failures:

Individual patient characteristics (level one in the multilevel analyses)

1. *Gender:* Male or female.
2. *Age-group:* Three groups were analysed: a) patients younger than 51 years; b) middle aged patients being between 51 and 59 years; and c) patients 60 years of age or older.
3. *Treated jaw:* Maxilla or mandible.
4. *Bone-quality:* Four classifications were used, where bone-quality 1 mainly represented cortical bone, and qualities 2, 3 and 4 represented gradually decreasing quantities of cortical bone being replaced by increasing amounts of bone marrow (Lekholm & Zarb, 1985). To evaluate whether any significant differences in failure frequencies could be detected between the 4 bone-qualities, Fishers exact test was used (Altman, 1991). The first variable to be tested was bone-quality 1 against qualities 2, 3 and 4. Thereafter, qualities 1 and 2 were compared with 3 and 4. Finally, qualities 1, 2 and 3 were tested against bone-quality 4. A boundary between the compared groups was declared, when the very first statistical difference between the compared groups was established. Two subgroups with a significantly different outcome were thereby established, which were later used for bone combination analyses.
5. *Jaw-shape:* The population was divided into five groups (A, B, C, D, E), where jaw-shape A represented the least resorbed jaws and jaw-shape E the most extremely resorbed ones (Lekholm & Zarb, 1985).

The same procedure as described under point 4 above was also used to identify a boundary between the 5 jaw-shape groups. It should be noted that all single tooth patients were excluded, giving a total of 395 patients remaining for this analysis.

6. *Treatment protocol:* Four types of prosthetic treatment were compared; fixed partial bridges; overdentures; single tooth replacements and fixed full prostheses.

When testing the variables, by using conditional binomial tests (Herrmann, 2003), multiple analyses were applied to evaluate each individual case with the mean of the others. A 499-sized bootstrap simulation was used in order to achieve a correct multiple significance-level (Efron & Tibshirani, 1993).

Clinic/treating doctor characteristics (level one in the multilevel analyses)

7. *Supporting implants/prosthesis:* The original numbers of implants placed and construction-supporting were used when evaluating this variable. When a patient had two or more prostheses, only the restoration being supported by the randomly selected implant(s) was analysed. Twenty-one patients had to be excluded because of implant failure or non- compliance. Thereby, 466 remaining patients participated in the "implants/prosthesis" calculation. A Pearson chi-square test (Altman, 1991) was used to evaluate whether the number of implants, from single implants up to 7 implants supporting the prosthesis, had any significant effect on the outcome.
8. *Responsible clinic:* Each clinic was regarded as one unit, without differentiating between treating surgeon and prosthodontist. However, more than one team of surgeons and prosthodontists could have been involved in the study but still would all be evaluated as one unit.

Individual implant characteristic (level one in the Multilevel analyses)

9. *Implant length:* Five groups of implant lengths, i.e. 7; 10; 13; 15; and 18 mm or longer, were analysed. Fourteen implants, wider in diameter than the commonly used Ø 3.75 mm implants, were excluded from the individual implant length evaluations.

Combinations of individual jaw-bone characteristics (level two in the multilevel analyses)

10. *Combinations of bone-qualities and jaw-shapes:* Four different jaw-shape/quality combinations (Table 6) were constructed for the 395 patients for whom these two variables were originally identified (point 4 and 5 - individual patient variables). Combination I represented implants placed in shapes/qualities A, B, C / 1, 2, 3. The shapes A, B,

C and qualities 1, 2, 3 have previously shown the lowest failure rates when evaluated on the first level of the multilevel analyses. Combination II consisted of implants placed in jaw-shapes and bone-qualities D, E / 1, 2, 3. Combination III consisted of implants placed in jaw-shapes A, B, C combined with bone quality 4. Finally, combination IV consisted of implants placed in bone types (D, E/ 4), which showed the highest failure rates on level one (points 4 and 5).

Table 6

Distribution of bone-qualities and jaw-shapes with regard to the four subgroups forming the different bone-combinations.

<table>
<tr><th rowspan="2">Bone-quality*</th><th colspan="5">Jaw-shape*</th></tr>
<tr><th>A</th><th>B</th><th>C</th><th>D</th><th>E</th></tr>
<tr><td>1</td><td colspan="3" rowspan="3">Combination I</td><td colspan="2" rowspan="3">Combination II</td></tr>
<tr><td>2</td></tr>
<tr><td>3</td></tr>
<tr><td>4</td><td colspan="3">Combination III</td><td colspan="2">Combination IV</td></tr>
</table>

***according to Lekholm & Zarb (1985)**

Implant lengths within constructed bone-combinations (level three in the multilevel analyses)

11. *Influences of implant length within the jaw-shape/-quality combinations (I; II; III; IV):* Each of the four combinations mentioned above (point 10 - combinations of individual jaw-bone characteristics) was divided into two new subgroups by the length of the implants inserted (7 and 10 mm versus 13 mm or longer), all of 3.75 mm diameter.

Analyses

Multilevel analyses were performed as a step-by-step method using chi-squared tests (Altman, 1991) on three levels. Furthermore, post hoc analyses (Hochberg & Tamhane, 1987) were used to identify whether an individual value differed within any of the samples tested; treatment protocol, jaw-bone quality, jaw-shape, implant length, number of implants per prosthesis, bone-combinations, and implant length within the jaw-bone combinations. The post hoc analyses were based on a multiple significance level of less than 0.05 for the detailed analyses.

Study IV - Evaluation of variances using the Jackknife method

The same, pooled population was employed as in studies II and III, but this time all implants inserted were analyzed. As mentioned before, the single tooth patients from one of the original MC-studies were only categorised with regard to bone-quality (density). To enable these patients to be included in the analysis of bone related combinations (I–IV), the following approximations were made and added to the pooled database. When a 7 mm implant was placed, the jaw shape (quantity) was assumed to be group E (extremely resorbed bone), 10 mm implants were assessed to group D, 13 mm to group C, 15 mm to group B, and 18 mm or longer to group A (most bone available), respectively.

Combinations of jaw-shapes and bone-qualities were then constructed in the same way as described in study III (point 10 – combinations of bone-qualities and jaw-shapes). However, this time patients from the single tooth study were also included, giving a total of 487 patients.

To prepare the population for the Jackknife method (Quenouille, 1956), the total population was divided into equally sized subgroups. One single tooth patient representing a successful outcome was excluded (n=487-1) to obtain an equal number of patients (n=54) in the subgroups (9). Thereafter, an adjusted randomization chart (Pocock, 1983) was used to give an almost equal distribution of patients from the various MC-studies in the 9 subgroups, allowing a maximum discrepancy of two patients per MC-study. Each patient was randomly allocated to one of the 9 subgroups, regardless of the number of implants inserted in the patient. Furthermore, one by one a different subgroup (n=54) was excluded giving 9 unique Jackknife samples based on the remaining 8 subgroups (n=432).

Standard errors were estimated (Altman, 1991) for the four bone-qualities (I-IV) using the nine Jackknife subgroups in relation to the preset time-periods. The standard errors were then used to emphasize uncertainties of the CSR:s for the six time-periods.

Calculation of CSR:s was performed for all 486 patients (1737 implants) and also for the four combinations (I-IV) separately, using the pooled population without Jackknife rearrangement of the data. For this purpose the standard life-table principles were utilized (Kaplan & Meier, 1958). Moreover, as previously stated, (study III - point 10) combination IV differed significantly from the three other combinations, the latter three groups were amalgamated and analysed as one sample (combinations I-III). Finally, confidence intervals (CI:s) were calculated and used to test whether the difference between the CSR:s (combinations I-III versus combination IV) was significant.

A log rank test (Cox & Oakes, 1984; Altman, 1991) was used as an alternative statistical test for differences in survival distribution between the two groups of combinations (I-III vs IV). An ordinary log rank test was computed using all data. In the log rank test a ratio of excess failure number and standard error was computed for the preset time-periods. The test was then repeated after the Jackknife re-sampling method to again estimate the difference between the results, but this time taking dependence into account.

Overall statistical principles

In order to study the statistical queries presented in the aims and material & methods, the following conditions were applied:

- Implants not followed throughout the entire study-period were included in the evaluations for as long as they were surveyed. Detailed information on withdrawn or dropped-out patients was published in the separate 5-year reports and has been presented in Table 3.
- The patients included in the pooled populations were assumed to be independent of each other.
- Chi-square tests used were either Pearson's or Fishers exact tests depending on the number of groups to be compared (Altman, 1991).
- When performing the significance analyses, withdrawn implants were included.
- Multiple interferences, including separate evaluations, were not formally taken into consideration for post hoc analyses.
- SPSS (Statistical computer program from SPSS Inc, Chicago, Il., USA) and Mathematica (Wolfram Research Inc, Champaign, Il. USA) were used for the statistical evaluations.
- The statistical program R (Swedish University Computer Network, Uppsala) was utilized, and also the New S Language: A Programming Environment for Data Analysis and Graphics, (Becker R, Chambers JM, Wilks AR, Chapman Hall NY (1988).

Results

Study I – Evaluation of dependence

The distribution of implant failures seen in Figures 5 and 6 and the distribution of individual implant failures indicated a dependence existed prior to any statistical calculations were performed.

The variances calculated were 60.1 for the discrete distribution (V) and 41.9 for the binomial distribution (V_0). Thus, V was found to be bigger than V_0, indicating that a dependence among the implants within the same jaw clearly existed prior to loading.

The CSR:s of the 5 manually randomised "one implant per patient" samples were 91.4%, 91.8%, 92.2%, 93.1% and 92.2%, after 5 years of clinical function. The results indicated that the "one implant per patient" technique could be a valid method when evaluating data affected by dependence.

Finally, the 5-year CSR for all inserted implants was 92.7%, which is within the range of the 5 manually randomised "one implant per patient" values (91.4-93.1%), despite the fact that established dependence was ignored.

Study II – Analyses of withdrawn or dropped-out patients

Originally withdrawn or dropped-out patients

In all clinical trials, some of the patients included will fail to complete the entire study protocol. In the current MC-studies, 21% or 104 patients were not evaluated or had failed completely after 5 years, whereof 15% or 73 patients failed due to lack of cooperation.

The answering frequency to update requests for the withdrawn or dropped-out patients was very low and only 3 of the centres did respond. One main reason for this was that 69% of the doctors previously responsible for the study subjects were no longer working at those clinics. The response from the treating centres stated that 9 of the 56 withdrawn or dropped-out patients had attended for a check-up after the MC- studies had been concluded. The fate of the implants in 47 "unaccounted for" patients remained unknown. Consequently, no conclusions regarding the true fate of the implants in the originally withdrawn or dropped-out patients could be drawn.

Random withdrawals

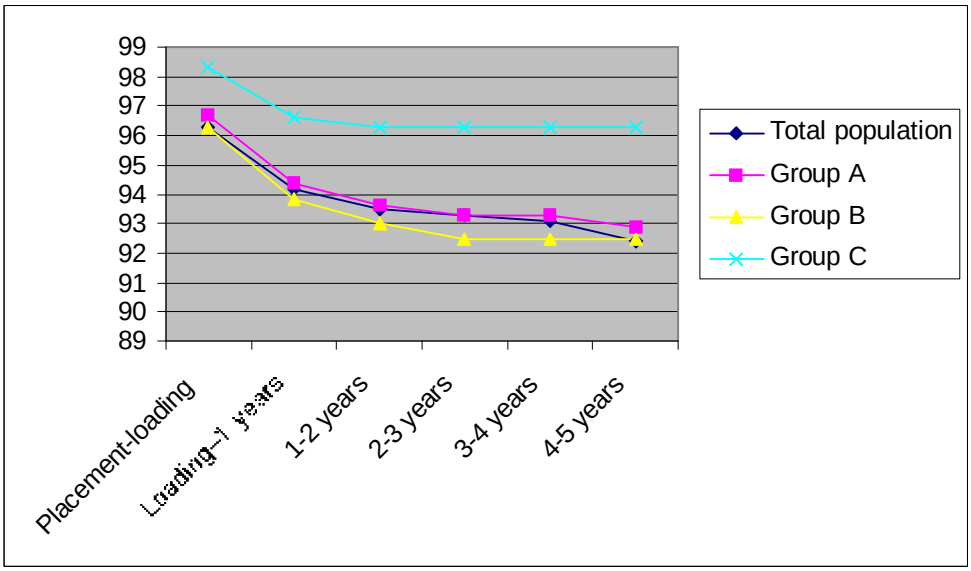
The CSR:s of the randomly withdrawn or dropped-out patient groups (Table 7; Figure 7) after 5 years of follow-up were calculated as the following: group A (25%) 92.9%, group B (50%) 92.5% and group C (75% withdrawn or dropped-

out) 95.7%. When all patients (100%) were evaluated the comparative 5-year CSR was 92.4%. When the number of withdrawn or dropped-out patients reached 50% (as in group B) the CSR:s remained the same (92.5%) during the last three years of follow-up. Corresponding figures for group C were 95.7% for the last four years. Unchanged CSR:s were obtained for the last three to four time-periods in groups B and C. However, the corresponding figures constantly decreased when the entire population was analysed (96.3, 94.2, 93.5, 93.3, 93.1, 92.4%). This showed that the risk for missing identification of late implant failures increased when the test samples became too small. Consequently, it was concluded that only 25% of the patients could be "unaccounted for" to maintain the same pattern of survival as in a certain full population. However, when chi-squared tests were performed on the same subgroups (A, B and C), no statistically significant differences (P>0.05) could be demonstrated.

Table 7
Distribution of CSR:s (%) for the total population and with 75% (A), 50% (B) and 25% (C) of the patients remaining in the groups, respectively.

Time-period	Total population	Group A	Group B	Group C
Placement-loading	96.3	96.7	96.3	98.3
Loading–1 years	94.2	94.4	93.8	96.6
1-2 years	93.5	93.6	93.0	96.3
2-3 years	93.3	93.3	92.5	96.3
3-4 years	93.1	93.3	92.5	96.3
4-5 years	92.4	92.9	92.5	96.3

Figure 7
Distribution of CSR:s (%) for the total population and with 75% (A), 50% (B) and 25% (C) of the patients remaining in the groups, respectively.



Withdrawals based on patient characteristics

The selective withdrawals, based on treated jaw, resulted in a strong statistically significant difference ($P<0.01$) in implant outcome, when chi-squared tests were used. Thereafter, the CSR:s were computed to 88.3% and 95.5% for upper versus lower jaws. By contrast, no significant differences ($P>0.05$) were observed for gender; the corresponding CSR-values were 92.0% for females and 92.8% for males.

Study III – Evaluation of potential prognostic risk factors

Over-all results

Significant or strongly significant differences were observed with potential prognostic factors for implant failures: jaw; jawbone quality; jaw-shape; treatment protocol; implant length and jawbone related combinations. However, no significant differences in implant failures were observed for; gender, age groups and responsible clinic or number of implants supporting the restoration (Table 8).

Table 8
Parameters contributing to risk of implant failure, studied by chi-square tests and post hoc analyses (multiple *P* level .05).

Parameters evaluated	X^2	Simulated limit of individual *P*-value	Outstanding sample	*P*-value for outstanding sample
Level One				
Patient characteristics:				
Gender: Male or female	*P*>0.05	Not tested		
Age group: <51, 51-59 or >59 years	*P*>0.05	Not tested		
Jaw: Maxilla or Mandible	*P*<0.01	Not tested		
Bone-quality: 1, 2, 3 or 4	*P*<0.001	0.0226	Bone-quality 4	0.00013
Jaw-shape: A, B, C, D or E	*P*<0.001	0.017	Jaw-shapes D, E	0.00009
Treatment protocol: Fixed partial bridges, overdentures; single tooth replacements or fixed full prostheses	*P*<0.05	0.0146	overdenture	0.0029
Treating doctor/clinic characteristics:				
Supporting implants/prosthesis	*P*>0.05	0.015	none	0.29
Responsible clinic	*P*>0.05	Not tested	-	-
Individual implant Characteristics:				
Implant length	*P*<0.001	0.018	7 mm implant	0.0004
Level Two				
Combinations of individual jaw-bone characteristics:				
Combination I, II, III or IV	*P*<0.001	0.0083	Combination IV	0.0006
Level Three				
Implant lengths (long or short) within the jaw-bone combinations I, II, III or IV	*P*>0.05	>.0125	none	0.05

Individual patient characteristics

Neither gender nor age influenced the outcome (P >0.05). However, when testing the outcome by jaw treated (upper versus lower), and the four different treatment protocols, significant differences (P <0.05) could be seen. With the post hoc analyses, overdenture (OD) rehabilitation was the one treatment that differed significantly from all other protocols
(P =0.0029). The OD study had 17 failures of the 133 implants placed (12.8%), while failures were 3.3% in the single tooth study, and 6.9% in the fixed partial and 4.9% in the fixed full prosthesis studies.
Strongly significant differences (P <0.001) were found for both bone-quality and jaw-shape analyses. Post hoc analysis showed that the highest failure rate and lowest P value (24.5% and P = 0.00013) occurred for jawbone quality 4. A similar high frequency of failure and low P value (21.0% and P = 0.00009) was demonstrated for jaw-shapes D and E together.

Regarding clinic/treating doctor characteristics

The "supporting implants/prosthesis" evaluation indicated the highest failure rate (13.0%) for prostheses supported by four implants. However, the failure frequency decreased when the restorations were supported via 3 or less implants or by 5 or more implants. No significant difference (P > 0.05) could be detected for the "supporting implants/prosthesis" analysis *per se*. Neither could the post hoc analyses detect any group that significantly differed (lowest
P = 0.29) from any of the other groups (1-7 implants/restoration).

Furthermore, when evaluating the influence of the treating clinic, the descriptive statistics confirmed that some of the clinics had experienced no failures at all, while others had 10-30% implant failures. Despite this, no significant differences (P >0.05) were observed for the responsible clinics.

Individual implant characteristics

Implant length was one factor that demonstrated a strongly significant difference (P <0.001). The post hoc analysis showed that the 7 mm long implants had the highest failure rate (21% and P = 0.0004). Combining 7 mm and 10 mm implants still produced a high failure rate (13.1%) and the p-value was even stronger (P = 0.00003). This combined group was also significantly different from the outcome of the 13-20 mm long implants. When the 7 and 10 mm implants were compared with each other, a significant difference (P-value <0.05) was also noted.

Furthermore, when evaluating whether the implant length corresponded with the given jaw-shape structure, it was demonstrated that in jaw-shape E was 83% of the placed implants 7 mm long. The corresponding figure for jaw-shape D was 57% for implants with a length of 10 mm.

Regarding combinations of individual jaw-bone characteristics

Strongly significant differences (P <0.001) were also established for the 4 bone-combinations. Looking into details - combination I (implants placed in jaw-shapes and bone-qualities A, B, C / 1, 2, 3) had the highest success rate of 95% (15 failures out of 296 inserted implants) Seventy-five percent of the patients belonged to this group. For combination II (representing 13% of the patients), 6 of 51 implants inserted failed (12%) in jaw-shapes known to have low success rates and in bone-qualities having high success rates (D, E / 1, 2, 3). Combination III, (with 9% of the patients) consisted of implants placed in areas where adequate jaw-bone was available (jaw-shapes A, B or C) but the jaw-bone quality (4) was poor. In this group 5 of 37 implants failed (14%). Finally, 3% of the patients belonged to combination IV, which had the highest failure rate (7 of 11 implants failed – 64%), consisting of implants inserted in jaw-shape and quality (D, E / 4). The post hoc analyses confirmed that the worst prognosis (P = 0.0006) would be with the latter combination.

Regarding implant lengths within constructed bone-combinations

Adding implant length (7 and 10 mm versus 13 mm or longer) to the bone combination analyses, as the last (third) level in the multilevel analyses, no significant difference could be detected.

Therefore, the multilevel analyses identified bone-combination IV as the condition most strongly correlated with a high failure rate while implant lengths could be disregarded. Hence, this bone combination can be considered the most valuable variable for predicting failure risk in this patient group.

Study IV - Evaluation of variances using the Jackknife method

The total number of inserted implants in the nine Jackknife subgroups (each of 54 patients) varied between 184 and 204. The mean failure rate was 6.3% and individual failure rates for the nine Jackknife subgroups varied from 2.0 to 9.6%. Furthermore, the majority of the implants (1311) were inserted in combination I (130-173 implants/subgroup). Only 40 implants were placed in combination IV, with a range from 0 to 14 implants per subgroup.

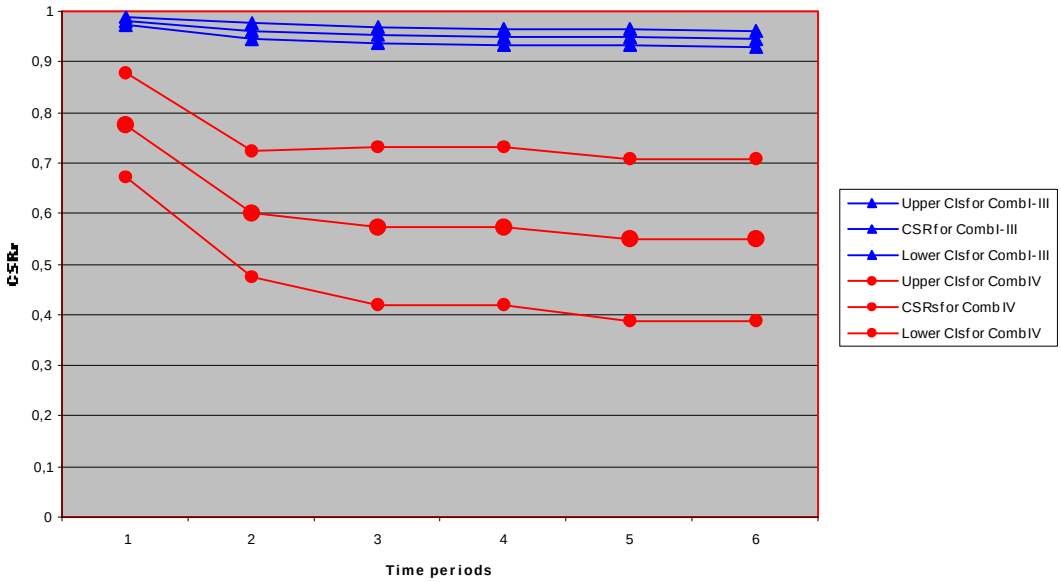
Standard errors (SE:s) varied by more than a power of 10 between the Jackknife samples (each of 432 patients) across the preset time-periods and jawbone combinations. Over all follow-up periods, the lowest standard errors were found in combination I (0.0037/time-period 1; and 0.0088/time-period 6), and the highest in Combination IV (0.0510 for time-period 1 and 0.079 for time-period 6).

Life-table analysis calculated a CSR of 93.4% after 5 years of clinical function for the 1737 implants placed in the 486 patients of the pooled population. The

CSR for combination IV was 54.8%, while the CSR:s for the other three bone combinations were 95.3% (I); 92.1%
(II) and 90.2% (III).

The 95% confidence interval on CSR:s for combinations I-III ranged from 97.9% +/- 1.5 to 94.4% +/-1.6 after 5 years. The corresponding figures for combination IV were 77.5% +/- 12.2 and 54.8% +/-15.9. Furthermore, the CSR for combination IV was about 20% lower than for the other combinations, once the prostheses were attached to the abutments and decreased a further 22% during the 5-year follow-up period. No overlapping of CI:s was noted at any time, indicating a statistical difference (Figure 8).

Figure 8
Diagram showing the distribution of CSR:s and corresponding CI:s for combinations I-III versus IV in relation to time-periods studied.



Log rank test

The standardized log rank test showed an excess failure number of 15.86 at the preset time-periods. Together with a standard error of 1.43, a ratio of 11.09 was computed. Subsequently, a strongly significant difference was established ($P < 0.001$) between combinations I-III and combination IV. A different result was seen with the standard error calculated from the nine Jackknife samples. The standard error for the nine Jackknife samples was 6.02 while the excess failure number remained the same i.e. 15.86. Thereby, the ratio decreased to 2.63, giving a less significant difference (*P*-value = 0.015). Consequently, when dependence was ignored it gave too strong a significance level.

Discussion

Population considerations

The purpose of this thesis was to study possible influences of some statistical aspects on the outcome of long term evaluations of oral implant treatment, identifying factors to be considered before new findings are given any rightful acceptance in the clinical practice.

In the mid-eighties the research division of Nobel Biocare AB (Göteborg, Sweden) introduced clinical trials in the field of oral implants, making it possible to undertake this study, as the trials produced a unique database of documentation. Four of the prospective MC-studies initiated at the time used the design of simplified pharmaceutical trials, but only one included a control group (Friberg et al., 1997). Furthermore, qualified statisticians were responsible not only for the statistical evaluation but also for the supervision and analysis of the data collected. Steering committee meetings were also held throughout the studies, in order to verify that all conclusions were correctly interpreted, with focus on evaluated variables, before publishing the results.

However, after pooling the 4 MC-studies mentioned, the information of which has been used in this thesis, it was possible to identify both similarities and differences between the 4 parts building the database, differences that have to be considered when assessing the outcome:

- The patients included in the 4 different studies had, for natural reasons, different numbers of implants supporting the constructions, ranging from one single implant to seven implants. This created a challenge of course, since the number of failure/s also varied, which must be taken into account when conducting statistical evaluations.

- The different prosthetic constructions also divided the material, since the outcome of the overdenture protocol significantly differed from the other treatments. However, it should be noted that the majority of the failures that were observed occurred early, prior to the prosthetic treatment or during the first year of function, and mainly in maxillas having compromised bone status.

- The full bridge study did not include any early follow up visits, i.e. after 1 week, 1 month and 6 months, in order to spare the patients unnecessary x-rays and excessive clinical evaluations. However, if a complication occurred during this period that information was still included in the registrations.

- Variables such as stomatognathic and periodontal evaluations, including oral hygiene and bleeding indices, pocket depth measurements and bleeding on probing registrations, were not recorded in the full bridge study since the other MC-reports included had not statistically associated those variables with documented failures.

- From the start, the experience of individual surgeons and prosthodontists was not regarded as a potential risk factor, as the selected clinics were acknowledged to be experienced in the implant treatment protocols being tested.

With all differences identified within the population studied, it is also important to emphasize the similarities that were observed.

- The 4 MC- studies in general used the same inclusion and exclusion criteria for patient selection.

- The sample sizes of the test-groups were decided with the aim of having at least 100 patients in each group at the end of the follow-up period.

- The very same success criteria were used in all the 4 MC-reports.

- The study design regarding treatment and yearly follow up sequences was the same in all the 4 MC-protocols.

- The individual patient health registrations did include information on diabetes, hypertension, allergies, colitis, and autoimmune diseases and medications, but not detailed information on smoking or drinking habits, previous infections of implant sites and/or on possible traumas.

- No recommendations on how many teeth each implant could or should support were given in any of the studies either.

Thus, the pooled patient population used in the current thesis was in many ways well defined but from other aspects some variations did exist. However, none of the differences mentioned were regarded as of any hindrance to a satisfactory performance of the current statistical evaluations. Consequently, the database was considered well suited for the elaborations that were to be carried out.

Study I – Evaluation of dependence

The results of study 1 showed from a mathematical point of view that a clear dependence existed among implants placed in the same jaw. The dependence identified was demonstrated to exist even before the prosthetic constructions had been attached. This outcome was obtained after looking upon the population in two different ways. First, it was calculated how many of the inserted implants had failed, assuming that implant failures had a binomial distribution. Secondly, the distribution of implant failures was studied, assuming the population was generally discretely distributed. However, in the current population, patients with only one implant failure were less common than those having no or more than one implant loss. This fact strongly indicates that if one implant has failed, the risk for more failures will increase for the same patient, i.e. dependence among the implants placed in the same jaw thus exists.

In order to handle the established dependence in the current study population, the "one implant per patient" technique (Mau, 1993) was utilized. This procedure was repeated 5 times and corresponding life-tables were constructed to calculate the 5-year CSR:s. By doing the randomisation manually, it was possible to see which specific implant was randomly selected, each time. The CSR:s ranged from 91.4 to 93.1 %, which indicated that the "one implant per patient" technique could give repeatable outcomes. When evaluating all the implants inserted, disregarding the established dependence, the CSR was 92.7%, i.e. within the range of the previously established CSR:s. However, the established dependence should still not be ignored in future evaluation studies.

No mathematical formulae to test for an anticipated dependence have been published (PubMed 2007-02-10), even although other authors have discussed potential causes of dependence among implants placed in the same jaw (Adell et al., 1990:b; Hutton et al., 1995; Cune & de Putter, 1996).

Adell et al. (1990:b) e.g. were the first authors to raise the subject of dependence in the oral implant literature and they presented a protocol that has been favoured by many others (Jemt et al., 1993; Lekholm et al., 1994; Leimola-Virtanten et al., 1995). However, Adell et al. (1990a) used the prostheses as the evaluation level instead of the implants, a correct method from a statistical point of view, but seen from the researcher's side, has some drawbacks. One is that all implant failures occurring before the prostheses have been attached would be disregarded, and most implant failures do occur prior to the connection of the prosthetic construction (Friberg et al., 1991; Snauwaert et al., 2000). Another drawback is that before the patient would be scored as a failure, all implants supporting the prosthesis have to fail. This indicates that the "Adell method" (Adell et al., 1990:a) is a rather insensitive technique, as complete treatment failures do rarely occur (Degidi et al., 2006).

Mau (1993) also stressed the challenge of dependence among implants placed in the same jaw. He, on the other hand, suggested performing the evaluation on the implant level and introduced the "one implant per patient" technique. The procedure to randomly select one of the patient's implants for the statistical evaluations has since been used by many others (Cune & de Putter, 1994, 1996; Haas et al., 1996, Chuang et al., 2001). Haas et al. (1996) investigated the repeatability of the technique. However, they used a computer program to repeat the procedure 500 times, instead of using manual selections. They found that by randomly selecting one of the inserted implants it was possible to achieve trustworthy CSR:s, each time. In the study by Chuang et al. (2001), three alternatives for handling dental implant data associated with dependence were used. The first alternative was the "one implant per patient" protocol; the second was testing the population for the hypothesis of independence, and the third one utilized all implants when testing the hypothesis that dependence existed. However, the authors did not find any differences between the outcomes of the three methods, and therefore these studies confirm the outcome of the current paper I.

Some researchers have disregarded the existence of any dependence, when selecting statistical evaluation methods (Engquist et al., 2002; Payne et al., 2002). One reason for this could be the lack of clarity over how to best handle dependence. Another reason might be that the final outcomes do not always show great differences, regardless of whether dependence is ignored or not (Chuang et al., 2002). A third reason could be the uncertainty that is created when evaluating only one of the implants inserted; thereby losing information about all other implants. Therefore, the question on how to handle dependence has led several research centres to use more and more sophisticated methods to solve the problem. Hutton et al. (1993) used a logistic regression model with backward selection for testing confounding influences and interactions. Eckert & Wollan (1998) on the other hand used the Cox proportional hazard model, while Chuang et al. (2002) used a clustered failure-time multivariate model, a method that they subsequently recommended for evaluating oral implants. However, none of the authors mentioned have demonstrated mathematically the existence of the dependence or have statistically shown that their method would give significantly different results, if dependence were to be acknowledged.

Consequently, study I is the first report in which dependence between implants has been mathematically demonstrated and acknowledged. The "one implant per patient" technique was also proven to be a simple and repeatable method for handling this dependence.

Study II – Analyses of withdrawn or dropped-out patients

The result of study II showed that random exclusion of patients did not affect the outcome much as long as 50% of the treated patients were followed the entire study period. If the excluded patients, on the other hand, were selected and not randomly chosen, the results could be significantly affected, depending on which of the selective variables was used for the sampling.

The outcome was obtained by randomly excluding groups of 25% of the population 3 times, leaving either 75% of the followed patients for group A, 50% for group B or 25% for group C. To exclude 25% each time may be regarded as an imprecise way of studying the effect of sample sizes. However, the purpose was not to find an exact borderline, but to study possible influences on the final outcome of portions of missing data.

From the life-table analyses used, the 5-year CSR:s for groups A and B showed results that corresponded well with the CSR of the entire population. However, when looking at the failure-times, implant losses were missed in all groups already after the second follow-up year, more frequently in groups B and C than in group A. This indicates that when the failures are few, which is often the case after the 2nd year of follow-up in oral implant studies (Adell et al., 1990:b; Zarb & Schmitt, 1990; Esposito et al., 1997), the risk for randomly excluding the few failures occurring is considerable, something that of course cannot be ignored.

No corresponding differences were detected, however, when using the paired chi-square test (Altman, 1991) for the same groups of patients (A, B and C). One reason for this could be the sample size in relation to the numbers of failures occurring. Another could be that the point of time for implant failure does not influence the result of the chi-square test. A third reason could be that the subgroups were created from one and the same database, and therefore some of the randomised patients could have belonged to more than one of the groups A, B and C. However, just because no differences between the subgroups were detected does not prove that the outcomes were equal.

When selectively excluding patients on the bases of gender, age or treated jaw, and still using the life-table approach and the “one implant per patient” technique (Mau, 1993), completely different outcomes were reached. For example, selection by jaw treated clearly showed a dramatic decrease in CSR, if the patients excluded were treatments in lower jaws. However, if females/males or different age groups were selected, no differences between these aspects could be detected. Also when the three variables were evaluated individually by chi-square tests (Altman, 1991), the only statistically significant difference demonstrated was between maxillas and mandibles. This emphasises the importance of always separating the results with regard to jaw type treated. On

the other hand, age and gender, which are often addressed in oral implant studies (Babbush et al., 1986; Buchs et al., 1995; Snauwaert et al., 2000), do not seem to be so important to the final implant outcome.

Before starting the elaborations with the current database, several attempts were made to recall the withdrawn or dropped-out patients of the 4 MC-studies, in order to evaluate the status of this missing population. However, very few patients did return for re-examination. One reason for this could be that the 4 studies were conducted more then 10 years earlier, and consequently several of the originally responsible investigators of the MC-studies were no longer working at the clinics involved in the reports. Another reason could be that the withdrawn patients were true drop-outs and could not or did not want to be re-examined. The fact that withdrawn or dropped-out patients remain "unaccounted for" has previously been confirmed by Roos et al. (1997).

No statistical evaluations of random or selective exclusion of patients in connection to follow-up of oral implant treatment have been found in the literature (PubMed 2007-02-15). However, some studies have selected or excluded patients on the basis of, for example, excluding maxillary treatments (Arvidson et al., 1992; Leimola-Virtanen et al., 1995), whilst other studies with broader recruitment have confirmed more frequent failure in the treatment of maxillae, especially when bone quality and quantity are poor ((Buser et al., 1997; Sennerby & Roos, 1998; Henry, 1999; van Steenberghe et al., 2002; Jemt & Hager, 2006).

Consequently, it is obvious that patient/research populations must be properly presented regarding current characteristics, and at least 50 %, but preferably 75%, of patients included should be followed the entire study period, in order to obtain trustworthy CSR:s and reliable times for implant losses. Furthermore, if a study population does not represent all types of patients, but only selections of these with features that may significantly affect the statistical outcome (such as good prognostic factors); this should be recognised in follow-up reports.

Study III – Evaluation of potential prognostic risk factors

Potential risk factors for implant losses were analyzed in study III, using a multilevel approach and post hoc analyses to identify patients at risk of implant failure.

The elaboration started by testing the research variables individually to search for significantly better or worse prerequisites for implant failures. The most important individual patient characteristics found to be related to a significantly higher failure rate were: maxillae, bone quality type IV, jaw shapes D and E and overdenture treatment. These factors have previously been associated with

decreasing implant success rates also by others (Babbuch et al., 1986; Hutton et al., 1995; Truhlar et al., 1994; Lazzara et al., 1996; Sennerby & Roos, 1998; Snauweat et al., 2000; Kourtis et al., 2004).

Variables showing no correlation to implant failures were: age-group and gender (as also discussed in study II) clinic/ treating team, and the number of implants supporting the prosthetic restoration. However, no detailed information was available for deeper analysis of the effect of clinic/treating doctor. This was because all information on specific surgeon or prosthodontist treating patients was unfortunately missing. A difference in success rates has previously been shown between surgeons (Lambert et al. (1997), which has been referred to as the learning curve.

In the current study, the risk of implant failures was highest with constructions supported by 4 implants. However, neither more nor fewer implants per construction showed significantly different outcome. Certainly, the risk of overloading the implants ought to increase when the number of implants per construction is decreasing. The opposite effect, with a greater risk of introducing static forces due to misfit seems likely, when more implants are connected to the construction (Rangert et al., 1989).

The only individual implant-related factor tested, (implant length) demonstrated significantly better results for longer implants than shorter ones. The 7-mm long implants showed the worst prognosis, (a failure rate of more than 20%) closely followed by the 10-mm long implants, with a 10% failure rate. There was also a significant difference between the results of these 2 implant lengths. Furthermore, a strongly significant difference was demonstrated between implants longer than 10-mm compared to the 7- and 10-mm long ones. Short implants have previously been associated with high failure rates (Snauwaert et al., 2000; Weng et al., 2003; Jemt & Hager, 2006). However, the new finding from this study was that the 10mm long implants caused a significantly higher failure rate than the 13-mm or longer implants. It should be noted, however, that long implants can not and should not be placed if insufficient bone is present. This was also illustrated by the fact that in jaw-shape E mainly 7-mm long implants were placed, indicating that even implant length is a kind of patient related characteristic.

When looking for prognostic risk factors, combinations of the worst conditions were tested in study III. At the first evaluation level in the multilevel analyses, the border lines between the best and the worst conditions for bone quality and jaw shape were statistically established. Thereafter, and based on the outcome, the second level tested statistical differences between the four newly created bone combinations (I-IV). At the third evaluation level, implant lengths were

added to the 4 bone combinations as the only individual implant characteristic to be tested. However, no statistical differences were seen between the shorter and the longer implants, even though 7 of the 9 short implants failed in bone combination IV. Probably no significant differences were found because too few implants in this combination were available for analysis.

Consequently, the most important risk factor demonstrated was the presence of both poor bone quality and inadequate jaw-bone volume as in combination IV. Fortunately, this combination was only seen in 3% of the total patient population. When the bone-related combinations were good, on the other hand, the failure rate dropped dramatically as in combination I, where only one of 20 patients experienced an implant failure. Seventy-five percent of all patients belonged to that combination. The high failure rate seen in patients with poor jaw-bone quality and extremely resorbed jaw-bone volumes has also been reported by others (Friberg et al., 1991; Jaffin & Berman, 1991; Lekholm et al., 1999; Kourtis et al., 2004; Friberg et al., 2005). However, the combination of these two bone-related factors identifying high risks of implant losses has not been reported previously.

Study IV - Evaluation of variances using the Jackknife method

The result of study IV clearly demonstrated a statistical difference between the amalgamated sample based on jaw-quality/-shape combinations I-III and combination IV. When performing the calculations, the Jackknife re-sampling method was used together with life-table analyses and the log rank test. The outcome coincided well with the outcome of study III, in which the "one implant per patient technique" and the life-table analyses were utilized. Consequently, it was possible via three different statistical evaluation procedures to show an increased failure risk for patients with jawbone quality 4 and jaw-shape D or E, i.e. the combination IV according to Herrmann et al. (2005).

In order to include the patients from the single tooth study in the current report, the missing information regarding jaw shape was adduced from the length of implant placed. Of course, this approximation was a rough way of judging the jaw-volume, as a 7-mm implant could have been placed in any type of jaw shapes. However, looking at implants placed in jaw shape E, in the 3 other MC-reports (van Steenberghe et al., 1990; Johns et al., 1992; Olsson et al., 1995), the majority were found to be 7mm long.

The Jackknife re-sampling technique restructured the implant sample sizes for the 9 Jackknife subgroups, each consisting of 432 patients. The failure rates varied from 2% to almost 10% for the subgroups, which was interpreted as showing that dependence did exist between implants placed in the same jaw. Furthermore, the standard errors for the 4 bone combinations studied and the 6

time-periods evaluated varied over the Jackknife samples. The greatest variances were seen for the combination IV group, which however only consisted of 40 patients, when being compared to combination I (n=1311). Combination IV also had the highest standard error (0.08), i.e. a power of ten times higher than combination I (0.008). Of course, the unequal sample sizes of the two combinations contributed to this difference. In order to simplify the test-model, the combinations I, II and III were amalgamated into one group when testing for the patients with the highest risks of implant failure. The Jackknife technique used could have been replaced by other techniques such as the Bootstrap procedure, when performing the current calculations. However, this would not have added any more information about the evaluation of variances. Besides, only one study of oral implant outcomes with the Jackknife, and none using the Bootstrap procedure, have been found in the literature by PubMed 2007-02-15, yet these two methods are often used in medical studies to calculate variances (Svensson, 1993). Therefore, using the Jackknife technique in the present report was an innovation in assessing oral implant treatment results.

Calculated CSR:s together with corresponding confidence intervals clearly demonstrated discrepancies between the two populations of combinations I-III and IV, as seen in Figure 8. It is important to note that the CSR:s obtained included all implants without first re-arranging these, while the CI:s were calculated using the Jackknife technique only. However, success rates of oral implants are most frequently calculated by life-table analyses, so if a comparison between groups should be conducted, CI:s ought to be added to the CSR:s. Alternatively, the log rank test could have been utilized, but both these methods still require some re-arrangement of the data. To illustrate the importance of correct consideration of the dependence, the standard errors used in the ordinary log rank test were compared to those calculated with the Jackknife re-sampling method. The log rank test values were approximately 25 per cent of the current Jackknife values, indicating that if the SE:s were underestimated, an inaccurate strong significance level might be reached. That is of course unacceptable. Alternative methods for handling dependence based on regression models or multivariate analyses have also been suggested (Eckert et al., 2003; Chuang et al., 2002; Eckert et al., 2001).

Consequently, the two statistical methods used in the current paper showed similar outcomes for oral implant treatment results as the “one implant per patient technique”. It is important to consider, though, that the established dependence has to be addressed in some way when dealing with follow-up data. The Jackknife re-sampling procedure combined with the log rank test or the life-table analysis seem possible ways of handling this problem.

Study design and statistical considerations

The declaration of Helsinki was written to protect human beings from unnecessary clinical tests and experiments (World Medical Association, WMA). According to the declaration, clinical trials should only be conducted when the study outcomes are expected to give new and valuable (for the patient) information about the variables being tested. One way to meet this demand is to perform sample size calculations prior to conducting the trials, to ensure statistically valid results can be achieved (Altman, 1991). The goal for each of the MC-studies, included in the present thesis, was also to have at least 100 patients to be evaluated at the 5-year final examination level and to be equally distributed among the participating clinics. It has previously been suggested (Albrektsson & Zarb, 1993) that 50 patients from at least 2 centers should have been followed to be able to claim valid long-term follow-up results. The database of the current thesis fulfilled this requirement well, as the pooled material used consisted of around 400 patients from several different clinics evaluated after 5 years. Mau et al. (2002) suggested that the sample size should be increased by 10-15% to ensure an acceptable test population at the end of the follow-up period. However, 15% of the current patients were lost to follow-up during the five-year research period, despite several attempts made to re-examine the lost patients both during the MC-study period and when conducting study II. An over sized research population of 115-120% therefore ought to be planned at commencement, at least for five year follow-up studies.

When the current MC-studies started in the mid $80^{\text{-ies}}$, the purpose was not to compare different oral implant systems but to evaluate the success rates of four different treatment protocols, using one and the same implant technique – the original Brånemark implant procedure. At the time, it was not fully understood and documented, though, which variables influenced the outcome achieved; which was first evaluated later in study III. However, the MC-reports were the very first prospective MC-studies conducted according to defined and reproducible principles. For example, the MC-studies used the same study design, inclusion and exclusion criteria and success and failure criteria, (Lekholm et al., 1994; Jemt et al., 1996; Henry et al., 1996; Friberg et al., 1997). The effect of different study designs has been discussed by several authors (Eckert et al., 2003, 2005; Esposito et al., 2001:b, 2005; Iacono & Cochran, 2007), of whom Eckert et al. (2003) noted the importance of having a clear purpose and correct study design from the start in order to achieve a significant follow-up outcome.

Clinical trials are often conducted to acknowledge or reject a hypothesis, e.g. if a new treatment procedure or device is tested, the hypothesis is that the new one is significantly better than a control procedure. An established protocol or product could then be used as the control. However, only one of the current MC-studies

included a control group, since Nobel Biocare AB was conducting the 4 MC-reports on the Brånemark system. The exception was the full fixed bridge-study (Friberg et al., 1997), in which a new self-tapping implant design (from Nobel Biocare AB) was compared with the original non-self-tapping Brånemark implant. Comparisons with any concurrent systems were thus not performed at the time. It would also have been possible to use traditional well-documented crown or bridge treatments of teeth (Karlsson, 1981) for the control purpose. However, the research committees responsible abandoned this possibility due to the completely different evaluation methods and success criteria being used for teeth and implant studies.

Due to the nature of the implant data, the results of the separate MC-studies were calculated with survival analysis (Altman, 1991). This is a technique which works very well for oral implants as long as the inclusion period is limited. If inclusion continues throughout the entire follow-up period, the number of implants at the start might be much higher than the number actually followed up. The literature frequently shows retrospective and even prospective reports based on less than 10% of the patients treated from the start, i.e. 90% or more of the subjects used for the calculations can be "unaccounted for" data (Babbush & Shimura, 1993; Ferringo et al., 2002). This means of course that substantial information regarding the fate of these patients/implants must be lacking. All life table analyses have, though, a rescue calculation built into the model to handle the "missing data", assuming that the "unaccounted for" have the same fate as those who have been followed (Kaplan & Meier, 1958). It should be noted that life table analysis was originally designed for fatal diseases, as the name suggests; not the same as oral implants. To some extent, worst case or best/worst case analyses could be used as complements in these situations to reveal the strength of the outcome (Mau et al., 2002). However, a better alternative seems to be to only refer to the CSR:s when 75% (or at least 50%) of the population have been evaluated, as also suggested in paper II. In some cases this would mean that only the outcome of the first follow-up years can be regarded as reliable, even if longer follow up periods are available (Herrmann, 2003).

As early as 1993, Albrektsson et al. (1993) shared the concern of including too many "unaccounted for" subjects when using life table analyses and therefore suggested the use of the four-field presentation protocol. However, this method does not facilitate statistical evaluations such as life-table analysis using CI:s. The four-field technique was therefore not used in the current thesis. Exploring the full life table, and not only the CSR:s, and using strict individual success criteria, also reveals the fate of the individual implants being followed, thereby eliminating a presentation of survival based solely on the group outcome.

Oral implant studies display a clustering pattern as discussed in paper I, as acknowledged and addressed by many authors (Hutton et al., 1995; Eckert & Wollan, 1998; Eckert et al., 2001; Chuang et al., 2002). Poor bone quality in combination with small bone volume was identified in the current thesis as the major reason for implant failure with expected failure rates of 30-60%. Therefore, patients with these characteristics would make a much better test population for evaluating implant design changes (such as surface texture alterations) than the standard population with expected success rates of 90% or more. However, no controlled studies are available in the literature in which combination IV has been used to evaluate new surfaces or other implant design features. One exception is Hallman et al. (2005), who used compromised jawbones needing bone grafting in such an approach. In their study they compared implants with a turned surface with those having a roughened texture, but the study included also a grafting protocol as mentioned.

In order to evaluate the outcome of oral implants, some basic information regarding the studied population needs to be addressed. It seems reasonable to use a chi^2 -test, e.g. the Pearson test, when evaluating the implant failure rates, as in papers II and III, since the Chi^2 test is used for categorical data and tests each group against the other in a four field table. Other authors have previously used the chi^2-tests for oral implant data (Morris et al., 2004; Fransson et al., 2005). However, figures and diagrams can also give information about the population and the strength of the outcome without focusing on significant differences (figure 5-7). One example of this was in study IV, where a diagram demonstrated the difference between two groups and the strength of this difference was calculated from the variances for the different time periods studied (Figure 8). A much stronger difference was seen in paper IV with the log rank test before the data were rearranged with the Jackknife technique. This indicates that obtaining *P*-values whilst ignoring the dependence may wrongly conclude a statistical difference, which of course is not acceptable.

In paper III, the purpose was to look for variables causing clustering effects as reasons for implant losses. Each research level was then evaluated separately, using a multivariate step-by-step technique. Such analyses have previously been used in other reports for similar purposes (Ferreria et al., 2006; Noguerol et al., 2006). Logistic regression analysis has been used when searching for prognostic indices in patients assumed to have a cluster pattern in their implant failures (Laine et al., 2006). Chuang et al. (2002) also found logistic regression analysis served this purpose. However, an important difference from logistic regression analysis, is that multivariate analysis handles both dependent background variables and dependent result variables, while the logistic regression only takes care of the dependent background variables. Evaluations of dental implants are extraordinarily complex, since the number of implants inserted can vary from 1

to 7 or more, and the number of failures may vary from zero to all of the implants. Furthermore, implants inserted within one patient are dependent regarding their background factors, and the result variables may also be dependent (Holm, 2007). Consequently, when only searching for background factors, multivariate analyses such as the step-by-step technique and/or the logistic regression analysis, seem to be suitable procedures to use. The dental implant data should then first be re-arranged, using a Bootstrap or Jackknife technique, which will improve the outcome of the logistic regression model. However, when searching for risk factors, multivariate analyses, life tables with confidence intervals and/or log rank tests were used in the current report, and they seemed to work well as long as the dependence was acknowledged and addressed accordingly.

Implant related variables, such as implant surface texture, implant diameters and designs, were not analyzed in the present study, since the current data only related to the original Brånemark implant system. This is because in the 1980^{ies}, the turned (machined) implant of one diameter dominated the placements performed and therefore only implant length was available as an implant related characteristic in the current report.

Retrospective publications and available long-term controlled prospective studies have drawbacks when it comes to comparing new implant features with well known ones. One reason for this is that most research activities at present only focus on implant design and/or surgical and prosthetic procedures. A second reason is that some of the comparative studies extant do not have implant success as the main variable evaluated but assess only one of the other success criteria, such as marginal bone loss (Balshi et al., 1996; Lindquist et al., 1997). A third reason is that the follow-up time in many situations is not long enough, while published data have shown implant failures do still occur after several years of function (Esposito et al., 1997; Snauwaert et al., 2000). This, of course, further complicates the evaluation of new surface textures, when only short follow up periods are available (Albrektsson & Wennerberg, 2004; Friberg et al., 2005). A fourth reason is that the sample size is not large enough, when success rates of both tests and controls are expected above 90%. Besides, the product cycle is becoming shorter and shorter (Millenium report, 2005), and consequently products may already be replaced before long-term follow-up data in scientific reports, companies' individual warranty systems, or the FDA's or other authorities failure reports, are available.

Recently, Eckert et al. (2005) asked six of the major implant companies for their recommended documentation on clinical long-term success of oral implant treatment in order to compare different implant systems. Fifty-nine of the sixty-nine reports sent in did not, however, qualify as valid clinical long-term

documentation. Furthermore, Eckert et al., (2005) stated that the remaining studies demonstrated no obvious differences in success rates. The pooled material of the Eckert study (2005) showed a 96% success rate with a confidence interval from 93 to 98%. No individual implant brand demonstrated success rates statistically better than the total success rate. Therefore, the studies should be designed so a statistical difference can be identified; for instance by using patients with compromised bone. Otherwise the Helsinki declaration is contravened in that unnecessary studies on humans should not be conducted. Equally good success rates would only be acceptable, if some negative side effects can simultaneously be reduced. For example, if the cost could be reduced to increase the number of patients who can benefit from implant treatment by using generic products that would be one such extra positive effect, provided that the same quality of treatment is reached. Fewer component complications (for example due to a higher grade of titanium), decreased pain and discomfort and/or shorter treatment times would be other possible effects that could count. Therefore, changes in product lines should give the treating doctors and their patients some benefits instead of just a marketing activity, i.e. there should always be a patient-value to newly introduced components and/or treatment procedures.

In summary, the scope of this thesis was not to cover all statistical aspects of validating oral implants. Rather it was to enlighten some challenges in using statistical methods without full understanding of the nature of the data studied or the statistical formulae used. Furthermore, to compare the outcomes of different statistical analyses, some basic statistical rules were violated such as using several statistical methods on the same material. However, the purpose was not then to search for significances, but rather to evaluate the strength and repeatability of the different outcomes. It has also been the hope that the outcomes of the thesis will help other researchers and clinicians to conduct their follow-up studies in an improved way in the future, using a better study design and the best statistical procedures.

Summary of Results

The most important finding of this thesis overall was that implants within the same jaw demonstrated an interdependence and that this dependence could be handled via suitable statistical methods for dependence analysis combined with well established statistical methods.
The separate findings were:

- A clear dependence among implants within the same jaw existed even before the attachment of the prosthetic construction (study I),

- The "one implant per patient" randomisation technique did not affect the outcome of the life-table analyses of implant treatment results (study I),

- Information about the true withdrawals was too limited to draw any conclusions regarding their correlation to the outcome of implants followed (study II),

- A cut-off limit for the percentage of randomly withdrawn or dropped-out patients could be set at approximately 25%, before any effects on the statistical evaluation of the results were observed (study II),

- The selective non-random withdrawal of patients did influence the statistical outcome, when evaluating dental implant outcomes, and therefore, the reasons for the withdrawals need to be acknowledged (study II),

- Patient related characteristics were identified to be the most important risk factors, by using a multilevel approach based on chi-squared and post hoc tests (study III),

- High-risk patients were found when the jaw bone-combination was D, E/4 - bone types, which, however, were only observed in a few percents of the population (study III),

- The same risk bone-combinations D, E/4 were also identified by the Jackknife re-sampling method combined with life-table analyses and confidence intervals (study IV),

- A statistical difference was also established for these risk bone-characteristics when studied by the Jackknife method with the log rank test (study IV).

Future considerations

Based on the present thesis there are some important issues to be addressed and to consider, when in the future designing or evaluating clinical trials on oral implant treatments.

The currently observed established dependence among implants within the same jaw should never be ignored when evaluating oral implant outcomes. From this thesis it was clear, that two methods can be used to handle this dependence, i.e. the "one implant per patient" or the Jackknife technique. The former procedure seems easy to handle and understand, while the Jackknife method is more efficient since all observations are included in the estimates.

Furthermore, the Jackknife technique, when used in combination with the log rank test, showed that the variances were more pronounced than if dependence was ignored and an ordinary log rank test was used. Therefore, the use of log rank tests, assuming independence, could underestimate the obtained standard errors, which may incorrectly lead to too high p-values and too strong significance levels. The recommendation is, consequently, to always re-arrange the data with the Jackknife method whenever variances are part of the evaluation.

In every prospective study it is inevitable that patients will be withdrawn or dropped-out. Consequently, the withdrawn and dropped-out levels should be addressed in the result presentation and should be allowed for at study start by increasing the number of patients to be included by 15-20% to ensure the desired sample size after 5 years of follow-up.

As currently also shown the CSR:s, calculated from life table analyses based on decreasing numbers of followed patients, may not always give a true picture. The long-term outcome, presented as CSR:s, should therefore instead be based on the result of at least 50%, and preferably 75%, of the patients included from the start. Consequently, the referred follow-up time should also be limited to periods, for which at least 50% of the patients are available for evaluation. This condition is even more important if the characteristics of the unexamined patients are not known, since some patient characteristics to be associated with higher failure frequencies.

The most important prognostic risk factors associated with patient characteristics found in this thesis were jaw treated, bone quality, jaw shape and combinations of the two latter variables. Three percent of the currently studied patients represented Combination IV, i.e. jawbone quality 4 and resorption degree D or E, and in this group two out of three of these patients experienced

one or more implant failures. Thus, this combination constituted the most potent risk variable. Consequently, it is necessary to always report the jaw bone characteristics when addressing oral implant outcomes as an otherwise good implant result might be hidden behind too good jaw bone variables, i.e. guided patient selection.

Finally, as jaw-combination IV resulted in the most failures, it would be more appropriate to use patients with this bone status when evaluating new products and/or treatment procedures, as a significance could be detected and still having a realistic sample size. One prerequisite is of course that there is a difference between the products.

To introduce new components with similar success rates when evaluated on an average population regarding bone status is insufficient unless other positive side effects accrue. However, if a new surface treatment designed for poor bone-quality can improve the outcome for patients with jaw-combination IV it would give these patients added value and be accepted on the market.

Acknowledgements

To finish this project required the contribution from many people and took a long time and I extend my sincere thanks to everyone involved.

Ulf Lekholm, my mentor, supervisor, co-author and friend for his never-ending support with great knowledge of the clinical aspects, the English and scientific language required.

Sture Holm my statistical guru, co-author and friend for helping me to understand the statistical aspects.

Christina Kultje and Stig Karlsson my co-authors.

Karl-Erik Kahnberg and Stig Karlsson for letting me belong to their departments during this time period.

All of the clinicians who participated in the 4 MC-studies for giving me their trust to continue evaluating their patients´ data.

Nobel Biocare for giving me continues access to their files.

To Professor PI Brånemark for letting my use the schematic illustrations of figures I, II and IV, drawn by Mr Fredrik Johansson.

To my family and friends for their support especially those how read my text over and over again Chanelle Stirling, Heinz Ingvarsson, Christina Kultje and Sally Bright.

Managements, friends and colleagues from Nobel Biocare; Claes Holmberg, Christina Bergström, Christina Kultje, Camilla Billström, Ellen Lund, Carina Persson and from Lifecore Dental; Dennis Alingham, Mary Beth Heffernan, and of course the staff and customers at Lifecore Biomedical AB.

The departments of Prosthodotic and Oral and Maxillofacial surgery, Sahlgrenska akademin, Göteborg University.

This project has indirectly been founded by the medical industries where I have been working since they gave me their consent to spend time on this project, Nobel Biocare AB, Biolin AB and Lifecore Biomedical AB.

References

Adell R, Lekholm U, Rockler B, Brånemark PI (1981). A 15-year study of osseointegrated implants in the treatment of the edentulous jaw. Int J Oral Surg; 10: 387-416.

Adell R, Lekholm U, Brånemark PI (1985). Surgical procedures. In: Brånemark PI Zarb G, Albrektsson T, (eds): Tissue-integrated prostheses: Osseointegration in clinical dentistry. Chicago, Quintessence Publishing Co; pp 211-232.

Adell R, Lekholm U, Rockler B, Lindhe J, Eriksson B, Brånemark PI, Shordone L (1986). Marginal tissue reactions at osseointegrated titanium fixtures (I). A 3-year longitudinal prospective study. Int J Oral Maxillofac Surg; 15: 39-52.

Adell R, Lekholm U, Gröndahl K, Brånemark PI, Lindström J, Jacobsson M (1990:a). Reconstruction of severely resorbed edentulous maxillae using osseointegrated fixtures in immediate autogenous bone grafts. Int J Oral Maxillofac Implants; 5: 233-46.

Adell R, Eriksson B, Lekholm U, Brånemark PI, Jemt T (1990:b). A long-term follow-up study of osseointegrated implants in the treatment of totally edentulous jaws. Int J Oral Maxillofac Implants; 5: 347-359.

Albrektsson T, Brånemark PI, Hansson BO, Lindström J (1981). Osseointegrated titanium implants. Require-ments for ensuring a long-lasting, direct bone-to-implant anchorage in man. Acta Orthop Scand; 5: 155-70.

Albrektsson T, Zarb G, Worthington P, Eriksson AR (1986). The long-term efficacy of currently used dental implants: a review and proposed criteria of success. Int J Oral Maxillofac Implants; 1: 11-25.

Albrektsson T, Blomberg S, Brånemark PI, Carlsson GE (1987). Edentulousness an oral handicap. Patient reactions to treatment with jawbone-anchored prostheses. J Oral Rehabil; 14: 503-11.

Albrektsson T, Lekholm U (1989). Osseointegration: current state of the art. Dent Clin North Am; 33: 537-54.

Albrektsson T (1993). On long-term maintenance of the osseointegrated response. Aust Prosthodont J; 7 Suppl: 15-24.

Albrektsson T, Zarb GA (1993). Determinants of correct clinical reporting. Ont Dent; 76: 29-33.

Albrektsson T (2001). Is surgical skill more important for clinical success than changes in implant hardware? Clin Implant Dent Relat Res; 3: 174-5.

Albrektsson T, Wennerberg A (2004). Oral implant surfaces: Part 2--review focusing on clinical knowledge of different surfaces. Int J Prosthodont; 17: 544-64.

Albrektsson T, Wennerberg A (2005). The impact of oral implants - past and future, 1966-2042. J Can Dent Assoc; 71: 327.

Alkan I, Sertgoz A, Ekici B (2004). The influence of occlusal loading location on stresses transferred to implant-supported prostheses and supporting bone: A three-dimensional finite element study. J. Prosthet. Dent; 91: 319-25.

Altman DG (1991). Practical statistics for medical research. London: Chapland & Hall.

Andersson B, Ödman P, Carlsson GE (1995). A study of 184 consecutive patients referred for single-tooth replacement. Clin Oral Implants Res; 6: 232-7.

Annual industry report (2005). http://www.opec.org/library/Annual

Arvidson K, Bystedt H, Frykholm A, von Konow L, Lothigius E (1992). A 3-year clinical study of Astra dental implants in the treatment of edentulous mandibles. Int J Oral Maxillofac Implants; 7: 321-329.

Astra Tech home page (2007). www.astratech.com

Attard NJ, Zarb GE (2004). Long-term treatment outcomes in edentulous patients with implant-fixed prostheses: the Toronto study. Int J Prosthodont.; 17: 417-24.

Azen SP, Roy S, Pike MC, Casagrande J, Mishell DR jr (1977). A new procedure for the statistical evaluation of intrauterine contraception. Am J Obstet gynecol; 128: 329-35.

Babbush CA, Kent JN, Misiek DJ (1986). Titanium plasma-sprayed (TPS) screw implants for the reconstruction of the edentulous mandible. J Oral Maxillofac Surg; 4: 274-282.

Babbush CA & Shimura M (1993). Five-year statistical and clinical observations with the IMZ two-stage osteointegrated implant system. Int J Oral Maxillofac Implants; 8: 245-253.

Bahat O (1993). Treatment planning and placement of implants in the posterior maxillae: Report of 732 consecutive Nobelpharma implants. J Oral Maxillofac Implants; 8: 151-161.

Bain C, Moy P (1993). The association between the failure of dental implants and cigarette smoking. J Oral Maxillofac Implants; 8: 609-615.

Balshi TJ, Hernandez RE, PryszlaK MC, Rangert B (1996). A comparative study of one implant versus two replacing a single molar. Int J Oral Maxillofac Implants; 11: 372-8.

Best DJ, Roberts DE (1975). Upper Tail Probabilities of Spearman's Rho. Applied Statistics; 24.

Bewick V, Cheek L, Ball J (2004). Statistics review 12: survival analysis. Crit Care; 8: 389-94.

Bischof M, Nedir R, Abi Najm S, Szmukler-Moncler, Samson J (2004). Implant stability measurement of delayed and immediately loaded implants during healing. Clin Oral Implants Res; 15: 529-39.

Bolind P, Johanson CB, Balshi T, Langer B, Albrektsson T (2005). A study of 275 retrieved Branemark oral implants. Int J Periodontics Restorative Dent; 25: 425-37.

Bousdras VA, Aghabeigi B, Griffiths B (2006). Management of apical bone loss around a mandibular implant: a case report. Int J Oral Maxillofac Implants; 21: 439-44.

Brunelli A, Rocco G (2006). Internal validation of risk models in lung resection surgery: bootstrap versus training-and-test sampling. J Thorac Cardiovasc Surg; 131: 1243-7.

Brånemark PI, Adell R, Breine U, Hansson BO, Lindström J, Ohlsson A (1969). Intra-osseous anchorage of dental prostheses. I. Experimental studies. Scand J Plast Reconstr Surg; 3: 81-100.

Brånemark PI, Hansson BO, Adell R, Breine U, Lindström J, Hallen O, Ödman A (1977). Osseointegrated oral implants in the rehabilitation of the edentulous patient. Experience from a 10-year period. Scand J Plast Reconstr Surg;11 (suppl 16).

Brånemark PI, Gröndahl K, Öhrnell LO, Nilsson P, Petruson B, Svensson B, Engstrand P, Nannmark U (2004). Zygoma fixture in the management of advanced atrophy of the maxilla: technique and long-term results. Scand J Plast Reconstr Surg Hand Surg; 38: 70-85.

Brägger U, Wermuth W, E Török (1995). Heat generated during preparation of titanium implants of the ITI® Dental Implant System: an *in vitro* study Clinical Oral Implants Research; 6: 254–259.

Buchs AU, Hahn J, Vassos DM (1995). Interim clinical study report: a threaded, hydroxylapatite-coated implant-five-year post-restoration safety and efficacy. J Oral Implantol; 21: 266-274.

Buser D, Weber HP, Brägger U (1990). The treatment of partially edentulous patients with ITI hollow-screw implants: presurgical evaluation and surgical procedures. Int J Oral Maxillofac Implants; 5: 165-75.

Buser D, Mericske-Stern R, Bernard JP, Behneke A, Behneke N, Hirt HP, Belser UC, Lang NP (1997). Long-term evaluation of non-submerged ITI implants. Part 1: 8-year life-table analysis of a prospective multi-center study with 2359 implants. Clin Oral Implants Res; 8: 161-72.

Chen CF, Feng X, Szeto J (2006). Identification of critical genes in microarray experiments by a Neuro-Fuzzy approach. Comput Biol Chem; 30: 372-81.

Chiapasco M, Abati S, Romeo E, Vogel G (2001). Implant-retained mandibular overdentures with Branemark System MKII implants: a prospective comparative study between delayed and immediate loading; 16: 537-46.

Chuang SK, Tian L, Wei LJ, Dodson TB (2001). Kaplan-Meier analysis of dental implant survival: a strategy for estimating survival with clustered observations. J Dent Res; 80: 2016-2020.

Chuang SK, Wei LJ, Douglass CW, Dodson TB (2002). Risk factors for dental implant failure: a strategy for the analysis of clustered failure-time observations. J Dent Res; 81: 572-577.

Cochran DL, Buser D, ten Bruggenkate CM, Weingart D, Taylor TM, Bernard JP, Peters F, Simpson JP (2002). The use of reduced healing times on ITI implants with a sandblasted and acid-etched (SLA) surface: early results from clinical trials on ITI SLA implants. Clin Oral Implants Res; 13: 144-53.

Cochran DL, Morton D, Weber HP (2004). Consensus statements and recommended clinical procedures regarding loading protocols for endosseous dental implants. Int J Oral Maxillofac Implants; 19 Suppl: 109-13.

Cordioli G, Mayzoub Z (1997). Heat generation during implant site preparation: an in vitro study. Int J Oral Maxillofac Implants; 12: 186-93.

Cox JF, Zarb GA (1987). The longitudinal clinical efficacy of osseointegrated dental implants: A 3-year report. Int J Oral Maxillofac Implants; 2: 91-100.

Cox DR, OakesD (1984). Analyses of survival data. Chapman & Hall, London.

Cune MS, de Putter C (1994). A comparative evaluation of some outcome measures of implant systems and suprastructures types in mandibular implant-overdenture treatment. Int J Oral Maxillofac Implants; 9: 548-555.

Cune MS, de Putter C (1996). A single dimension statistical evaluation of predictors in implant-overdenture treatment.J Clin Periodontol; 23: 425-431.

Cune MS, de Putter C, Hoogstraten J (1997). A nationwide evaluative study on implant-retained overdentures. J Dent; 25 Suppl 1: 13-19.

Cutler SJ, Ederer F (1958). Maximum utilization of the life-table method in analysing survival. J Chron Dis; 8: 699-713.

DeBruyn H, Collaert B (1994). The effect of smoking on early implant failure. Clin Oral Implants Res; 5: 260-4.

Dalise DD (1988) Utilization of the PS bar system for subperiosteal implants. J Oral Implantol; 14: 455-66.

Degidi M, Piattelli A, Gehrke P, Felice P, Carinci F (2006). Five-year outcome of 111 immediate nonfunctional single restorations. J Oral Implantol; 32: 277-85.

Desjardins RP (1992). Prosthesis design for osseointegrated implants in the edentulous maxilla. Int J Oral Maxillofac Implants; 7: 311-20.

Dentsply home page (2007). www.dentsply.com

Dilek OC, Tezulas E, Dincel M (2007). A mini dental implant-supported obturator application in a patient with partial maxillectomy due to tumor: case report. Oral Surg Oral Med Oral Pathol Oral Radiol Endod; 103: 6-10.

Eckert SE, Wollan PC (1998). Retrospective review of 1170 endosseous implants placed in partially edentulous jaws. J Prosthet Dent; 79: 415-421.

Eckert SE, Meraw SJ, Weaver AL, Lohse CM (2001). Early experience with Wide-Platform Mk II implants. Part I: Implant survival. Part II: Evaluation of risk factors involving implant survival. Int J Oral Maxillofac Implants; 16: 208-216.

Eckert SE, Choi YG, Koka S (2003). Methods for comparing the results of different studies. Int J Oral Maxillofac Implants; 18: 697-705.

Eckert SE, Choi YG, Sanchez R, Koka S (2005). Comparison of dental implant systems: quality of clinical evidence and prediction of 5-year survival. Int J Oral Maxillofac Implants; 20: 406-15.

Efron B, Tibshirani RJ (1993). An introduction to the bootstrap. Chapman & Hall, London.

Ekelund JA, Lindquist LW, Carlsson GE, Jemt T (2003). Implant treatment in the edentulous mandible: a prospective study on Branemark system implants over more than 20 years. Int J Prosthodont; 16: 602-8.

Emms M, Tredwin CJ, Setchell DJ, Moles DR (2007). The effects of abutment wall height, platform size, and screw access channel filling method on resistance to dislodgement of cement-retained, implant-supported restorations. J Prosthodont; 16: 3-9.

Engquist B, Åstrand P, Dahlgren S, Engquist E, Feldmann H, Gröndahl K (2002). Marginal bone reaction to oral implants: a prospective comparative study of Astra Tech and Branemark System implants. Clin Oral Implants Res; 13: 30-37.

Ercoli C, Funkenbusch PD, Lee HJ, Moss ME, Graser GN (2004). The influence of drill wear on cutting efficiency and heat production during osteotomy preparation for dental implants: a study of drill durability. Int J Oral Maxillofac Implants; 19: 335-49.

Ericsson I, Randow K, Nilner K, Peterson A (2000). Early functional loading of Branemark dental implants: 5-year clinical follow-up study. Clin Implant Dent Relat Res; 2: 70-7.

Eriksson RA, Albrektsson T (1984). The effect of heat on bone regeneration: an experimental study in the rabbit using the bone growth chamber. J Oral Maxillofac Surg; 42: 705-11.

Eriksson AR, Albrektsson T, Albrektsson B (1984). Heat caused by drilling cortical bone. Temperature measured in vivo in patients and animals. Acta Orthop Scand; 55: 629-31.

Esposito M (1997). Failure patterns of four osseointegrated oral implant systems. J Mater Sci Mater; 8: 843-7.

Esposito M, Worthington HV, Coulthard P (2001:a). In search of truth: the role of systematic reviews and meta-analyses for assessing the effectiveness of rehabilitation with oral implants. Clin Implant Dent Relat Res; 3: 62-78.

Esposito M, Coulthard P, Worthington HV, Jokstad A (2001.b). Quality assessment of randomized controlled trials of oral implants. Int J Oral Maxillofac Implants; 16: 783-792.

Esposito M, Coulthard P, Thomsen P, Worthington HV (2005). Interventions for replacing missing teeth: different types of dental implants. Cochrane Database Syst Rev; 25: 1.

Ferreria SD, Silva GL, Cortelli JR, Costa JE, Costa FO (2006). Prevalence and risk variables for peri-implant disease in Brazilian subjects. J Clin Periodontol.; 33: 929-35.

Fransson C, Lekholm U, Jemt T, Berglundh T (2005) Prevalence of subjects with progressive bone loss at implants. Clin Oral Implants Res; 16: 440-6.

Ferrigno N, Laureti M, Fanali S, Grippaudo G (2002). A long-term follow-up study of non-submerged ITI implants in the treatment of totally edentulous jaws. Part I: Ten-year life-table analysis of a prospective multicenter study with 1286 implants. Clin Oral Implants Res; 13: 260-273.

Friberg B, Jemt T, Lekholm U (1991). Early failures in 4,641 consecutively placed Branemark dental implants: a study from stage 1 surgery to the connection of completed prostheses. Int J Oral Maxillofac Implants; 6: 142-6.

Friberg B, Nilson H, Olsson M, Palmquist C (1997). Mk II: the self-tapping Brånemark implant: 5-year results of a prospective 3-center study. Clin Oral Implants Res; 8: 279-285.

Friberg B, Dahlin C, Widmark G, Östman PO, Billström C (2005). One-year results of a prospective multicenter study on Branemark System implants with a TiUnite surface. Clin Implant Dent Relat Res; 7 Suppl 1: 70-5.

Gapski R, Wang HL, Mascarenhas P, Lang NP (2003). Critical review of immediate implant loading. Clin. Oral Impl. Res; 14: 515-27.

Gellerstedt M (2006). M^{12} Medicinsk statistik. Merck Sharp Dohme. Holmbergs, Malmö AB Sweden ISBN 91-631-6163-X.

Gröndahl K, Ekestubbe A, Gröndahl H-G (1996). Radiography in oral endosseous prosthetics. Nobel Biocare AB, Göteborg, Sweden ISBN 91-630-4471-4: 111-125.

Gunne J, Jemt T, Linden, B (1994). Implant treatment in partially edentulous patients: a report on prostheses after 3 years. Int J Prosthodont; 7: 143-8.

Haas R, Mensdorff-Pouilly N, Mailath G, Watzek G (1996). Survival of 1920 IMZ implants followed for up to 100 months. Int J Oral Maxillofac Implants; 11: 581-588.

Hallman M, Mordenfeld A, Strandkvist T (2005). A retrospective 5-year follow-up study of two different titanium implant surfaces used after interpositional bone grafting for reconstruction of the atrophic edentulous maxilla. Clin Implant Dent Relat Res; 7: 121-6.

Henry PJ, Tolman DE, Bolender C (1993). The applicability of osseointegrated implants in the treatment of partially edentulous patients: three-year results of a prospective multicenter study. Quintessence Int; 24: 123-129.

Henry PJ, Laney W R, Jemt T, Harris D, Krogh P, Polizzi G, Zarb G, Herrmann I (1996). Osseointegrated implants for single-tooth replacement: A prospective 5-study multicenter study. Int J Oral Maxillofac Implants; 11: 450-455.

Henry PJ (1999). Clinical experiences with dental implants. Adv Dent Res; 13: 147-52.

Herrmann I (2003). Influences of study designs and statistical variations on the outcome of dental implant evaluation. Academic dissertation for the Licentiate degree in Odontology No 36, The Sahlgrenska Academy at Göteborg University Faculty of Odontology.

Heydenrijk K, Batenburg RH, Raghoebar GM, Meijer HJ, van Oort RP, Stegang B (1998). Overdentures stabilised by two IMZ implants in the lower jaw--a 5-8 year retrospective study. Eur J Prosthodont Restor Dent; 6: 19-24.

Higuchi K (2000). The zygomaticus fixture: an alternative approach for implant anchorage in the posterior maxilla. Ann R Australas Coll Dent Surg; 15: 28-33.

Higuchi K, Filmer T, Kultje C (1995). Implant survival rates in partially edentulous patients: a 3-year prospective multicenter study. J Oral Maxillofac Surg; 53: 264-8.

Hinode D, Tanabe S, Yokoyama M, Fujisawa K, Yamauchi E, Myamoto Y (2006). Clin Oral Implants Res; 17: 473-8.

Hochberg Y, Tamhane AC (1987). Multiple comparison procedures. John Wiley and Sons, New York City.

Holm S (2007). Personal communication.

Hunter J, Schmidt F, Jackson G (1982). Meta analysis. Cumulating research findings across studies Beverly Hills CA: Sage.

Hutton JE, Heath MR, Chai JY, Harnett J, Jemt T, Johns RB, McKenna S, McNamara DC, van Steenberghe D, Taylor R, Herrmann I (1995). Factors related to success and failure rates at 3-year follow-up in a multicenter study of overdentures supported by Brånemark implants. Int J Oral Maxillofac Implants; 10: 33-44.

Iacono VJ, Cochran DL (2007). Int J Oral Maxillofac Implants: 2007; 22 suppl: 11-18.

Implant Innovation homepage (2007). www.3i-online.com

Ivanoff CJ, Widmark G, Hallgren C, Sennerby L, Wennerberg A (2001). Histologic evaluation of the bone integration of TiO(2) blasted and turned titanium micro implants in humans. Clin Oral Implants Res; 12: 128-34.

James RA, Kelln EE (1974). A histopathological report on the nature of the epithelium and underlying connective tissue which surrounds oral implants. J Biomed Mater Res; 8: 373-83.

Jemt T, Laney WR, Harris D, Henry PJ, Krogh PH Jr, Polizzi G, Zarb GA, Herrmann I (1991). Osseointegrated implants for single tooth replacement: a 1-year report from a multicenter prospective study. Int J Oral Maxillofac Implants; 6: 29-36.

Jemt T, Book K, Linden B, Urde G (1992). Failures and complications in 92 consecutively inserted overdentures supported by Branemark implants in severely resorbed edentulous maxillae: a study from prosthetic treatment to first annual check-up. Int J Oral Maxillofac Implants; 7: 162-7.

Jemt T (1993). Implant treatment in resorbed edentulous upper jaws. Clin Oral Implants Res; 4: 187-194.

Jemt T, Chai J, Harnett J, Health R Hutton J, Johns R, McKenna S, McNamara D, van Steenberghe D, Watson R, Herrmann I (1996). A 5-year prospective multicenter follow-up report on overdentures supported by osseointegrated implants. Int J Oral Maxillofac Implants; 11: 291-298.

Jemt T, Hager P (2006). Early complete failures of fixed implant-supported prostheses in the edentulous maxilla: a 3-year analysis of 17 consecutive cluster failure patients. Clin Implant Dent Relat Res; 8: 77-86

Johns RB, Jemt T, Heath MR, Hutton JE, McKenna S, McNamara DC, van Steenberghe D, Taylor R, Watson RM, Herrmann I (1992). A multicenter study of overdentures supported by Brånemark implants. Int J Oral Maxillofac Implants; 7: 513-522.

Kahnberg KE, Nyström E, Bartholdsson (1989). Combined use of bone grafts and Branemark fixtures in the treatment of severely resorbed maxillae. Int J Oral Maxillofac Implants; 4: 297-304.

Kahnberg KE, Nilsson P, Rasmusson L (1999). Le Fort I osteotomy with interpositional bone grafts and implants for rehabilitation of the severely resorbed maxilla: a 2-stage procedure. Int J Oral Maxillofac Implants; 14: 571-8.

Kaplan EL, Meier P (1958). Non-parametric estimation from incomplete observations. J Am Stat Assoc; 53: 457-481.

Karlsson S (1986). A clinical evaluation of fixed bridges, 10 years following insertion. J Oral Rehabil. Sep; 13: 423-32.

Kendall MG (1948). Rank correlation method. Griffin. London.

Kendall PH (1962). Local corticosteroid injection therapy. II. Injection technique. Ann Phys Med; 6: 359-74.

Keith SE, Miller BH, Woody RD, Higgenbottom FL (1999). Marginal discrepancy of screw-retained and cemented metal-ceramic crowns on implants abutments. Int J Oral Maxillofac Implants; 14: 369-78.

Ko SM, Lee JK, Eckert Se, Choi YG (2006). Retrospective multicenter cohort study of the clinical performance of 2-stage implants in South Korean populations. Int J Oral Maxillofac Implants; 21: 785-8.

Kourtis SG, Sotiriadou S, Voliotis S, Challas A (2004). Private practice results of dental implants. Part I: survival and evaluation of risk factors-Part II: surgical and prosthetic complications.Implant Dent; 13: 373-85.

Kourtis S, Psarri C, Andritsakis P Doukoudakis A (2007). Provisional restorations for optimizing esthetics in anterior maxillary implants: a case report. J Esthet Restor Dent; 19: 6-17.

Köndell PA, Nordenram A, Moberg LE, Eliasson Snyberg B (1996). Reconstruction of the resorbed edentulous maxilla using autogenous rib grafts and osseointegrated implants. Clin Oral Implants Res; 7: 286-90.

Lambert PM, Morris HF, Ochi S (1997). Positive effect of surgical experience with implants on second-stage implant survival. J Oral Maxillofac Surg; 55: 12-8.

Laine P, Salo A Contio R, Yljoki S, Lindquist C, Suuronen R (2005). Failed dental implants - clinical, radiological and bacteriological findings in 17 patients. J Craniomaxillofac Surg; 33: 212-7.

Laney WR, Jemt T, Harris D, Henry PJ, Krogh PH, Polizzi G, Zarb GA, Herrmann I (1994). Osseointegrated implants for single-tooth replacement: progress report from a multicenter prospective study after 3 years. Int J Oral Maxillofac Implants; 9: 49-54.

Lazzara R, Siddiqui AA, Binon P, Feldman SA, Weiner R, Phillips R, Gonshor A (1996). Retrospective multicenter analysis of 3i endosseous dental implants placed over a five-year period. Clin Oral Implants Res; 7: 73-83.

Lehmann E L, D'Aabrera H J M (1975). Non-parametrics, statistical methods based on ranks Holden Day. San Francisco.

Leimola-Virtanen R, Peltola J, Oksala E, Helenius H, Happonen RP (1995). ITI titanium plasma-sprayed screw implants in the treatment of edentulous mandibles: a follow-up study of 39 patients. Int J Oral Maxillofac Implants; 10: 373-378.

Lekholm U, Zarb GA (1985). Patient selection and preparation. In:Brånemark P-I, Zarb GA, Albrektsson T (eds): Tissue-integrated Prostheses: Osseointegration in Clinical Dentistry. Chicago: Quintessence; 199-209.

Lekholm U (1993). New surgical procedures of the osseointegration technique A.M. Branemark. Aust Prosthodont J; 7 Suppl: 25-32.

Lekholm U, van Steenberghe D, Herrmann I, Bolender C, Folmer T, Gunne J, Henry P, Higuchi K, Laney W, Lindèn U (1994). Osseointegrated implants in the treatment of partially edentulous jaws: A Prospective 5-year multicenter study. Int J Oral Maxillofac Implants; 9: 627-635.

Lekholm U (1998). Surgical considerations and possible shortcomings of host sites. J Prosthet Dent; 79: 43-8.

Lekholm U, Gunne J, Henry P, Higuchi K, Linden U, Bergström C, van Steenberghe D (1999). Survival of the Brånemark implant in partially edentulous jaws: a 10-year prospective multicenter study. Int J Oral Maxillofac Implants; 14: 639-645.

Lekholm U, Gröndahl K, Jemt T (2006). Outcome of oral implant treatment in partially edentulous jaws followed 20 years in clinical function. Clin Implant Dent Relat Res; 8: 178-86.

Levin L, Sadet P, Grossmann Y (2006). A retrospective evaluation of 1,387 single-tooth implants: a 6-year follow-up. J Periodontol; 77: 2080-3.

Levine RA, Clem D, Beagle J, Ganales J, Johnson P, Solnit G, Keller GW (2002). Multicenter retrospective analysis of the solid-screw ITI implant for posterior single-tooth replacements. Int J Oral Maxillofac Implant; 17: 550-6.

Liedström G (1995). Lagen om medicintekniska produkter, bilaga 10 Fritzes förlag AB, Stockholm; pp 220-221.

Lindquist LW, Carlsson GE (1985). Long-term effects on chewing with mandibular fixed prostheses on osseointegrated implants. Acta Odontol Scand; 43: 39-45.

Lindquist LW, Carlsson GE, Jemt T (1997). Association between marginal bone loss around osseointegrated mandibular implants and smoking habits: a 10-year follow-up study. J Dent Res; 76: 1667-74.

Lifecore Biomedical homepage (2007). www.lifecore.com

Liu Y, Smith M, Rangayyan R (2004). The application of Efron's bootstrap methods in validating feature classification using artificial neural networks for the analysis of mammographic masses. Conf Proc IEEE Eng Med Biol Soc; 2: 1553-6.

Lorenzoni M, Pertl C, Zhang K, Wimmer G, Wegscheider WA (2003). Immediate loading of single-tooth implants in the anterior maxilla. Preliminary results after one year. Clin Implant Res; 14: 180-7.

Lundgren S, Nyström E, Nilsson H, Gunne J, Lindhagen O (1997). Bone grafting to the maxillary sinuses, nasal floor and anterior maxilla in the atrophic edentulous maxilla. A two-stage technique. Int J Oral Maxillofac Surg; 26: 428-34.

Malmqvist JP, Sennerby L (1990). Clinical report on the success of 47 consecutively placed Core-Vent implants followed from 3 months to 4 years. Int J Oral Maxillofac Implants; 5: 53-60.

Malevez C, Abarca M, Durdu F, Daelemans P (2004). Clinical outcome of 103 consecutive zygomatic implants: a 6-48 months follow-up study. Clin Oral Implants Res; 15: 18-22.

Marda KV, Kent JT, Bibby JM (1979). Multivariate analysis. Academic press ISBN 0124712525.

Mau J (1993). On statistics of success and loss for dental implants. Int Dent J; 43: 254-261.

Mau J, Behneke A, Behneke N, Fritzemeier CU, Gomez-Roman G, d'Hoedt B, Spiekermann H, Strunz V, Yong M (2002). Randomized multicenter comparison of two coatings of intramobile cylinder implants in 313 partially edentulous mandibles followed up for 5 years. Clin Oral Implants Res; 13: 477-487.

Millennium research group, (2005). 175 Bloor Street East, South Tower Suite, 701 Toronto, Ontario CANADA M4W 3R8.

Misch CM , Misch CE (1992). Generic terminology for endosseous implant prosthodontics. J Prosthet Dent; 68: 809-12.

Misch CE, Steigna J, Barboza E, Misch-Dietsh F, Cianciola LJ, Kazor C (2006). Short dental implants in posterior partial edentulism: a multicenter retrospective 6-year case series study. J Periodontol; 77: 1340-7.

Morris HF, Ochi S, Orenstein IH, Petrazzuolo V (2004). AICRG, Part V: Factors influencing implant stability at placement and their influence on survival of Ankylos implants. J Oral Implantol; 30: 162-70.

Mundt T, Mack F, Schwahn C, Biffar R (2006). Private practice results of screw-type tapered implants: survival and evaluation of risk factors. Int J Oral Maxillofac Implants; 21: 607-14.

Natali AN, Peyan PG, Ruggero AL (2006). Evaluation of stress induced in peri-implant bone tissue by misfit in multi-implant prosthesis. Dent Mater; 22: 388-95.

Naert IE, Duyck JA, Hosny MM, van Steenberghe D (2001). Freestanding and tooth-implant connected prostheses in the treatment of partially edentulous patients. Part I: An up to 15-years clinical evaluation. Clin Oral Implants Res; 12: 237-44.

Nobel Biocare home page (2007). www.nobelbiocare.com

Noguerol B, Munoz R, Mesa F, de Dios Luna J, O'Valle F (2006). Early implant failure. Prognostic capacity of Periotest: retrospective study of a large sample. Clin Oral Implants Res; 17: 459-64.

Nyström E, Ahlqvist J, Legrell PE, Kahnberg KE (2002). Bone graft remodelling and implant success rate in the treatment of the severely resorbed maxilla: a 5-year longitudinal study. Int J Oral Maxillofac Surg; 31: 158-64.

Olsson M, Friberg B, Nilson H, Kultje C (1995). MkII-a modified self-tapping Brånemark implant: 3-year results of a controlled prospective pilot study. Int J Oral Maxillofac Implants; 10: 15-21.

Oncag G, Akyalcin S, Arikan F (2007). The effectiveness of a single osteointegrated implant combined with pendulum springs for molar distalization. Am J Orthod Dentofacial Orthop; 131: 277-84.

Palmer RM, Palmer PJ, Smith BJ (2000). A 5-year prospective study of Astra single tooth implants. Clin Oral Implants Res; 11: 179-82.

Payne AG, Tawse-Smith A, Duncan WD, Kumara R (2002). Conventional and early loading of unsplinted ITI implants supporting mandibular overdentures. Clin Oral Implants Res; 13: 603-9.

Piattelli M, Scarano A, Paolantonio M, Iezzi G, Petrone G, Piattelli A (2002). Bone response to machined and resorbable blast material titanium implants: an experimental study in rabbits. J Oral Implantol; 28: 2-8.

Pinholt EM (2003). Branemark and ITI dental implants in the human bone-grafted maxilla: a comparative evaluation. Clin Oral Implants Res; 14: 584-92.

Pocock S.J (1993). In Clinical trials: A practical approach. Chichester, UK. J Wiley.

Proskin HM, Jeffcoat RL, Catlin A, Campbell J, Jeffcoat MK 2007. A Meta-analytic approach to determine the state of science on dental implant dentistry. Int J Oral Maxillofac Implants:; 22 suppl: 11-18.

Quenouille MH, (1956) Notes on bias in estimation. Biometrika; 43:353-360.

Rangert B, Jemt T, Jörneus L (1989). Forces and moments on Branemark implants. Int J Oral Maxillofac Implants; 4: 241-7.

Rasmusson L, Kahnberg K, Tan A (2001). Effects of implant design and surface on bone regeneration and implant stability: an experimental study in the dog mandible. Clin Implant Dent Relat Res; 3: 2-8.

Rasmusson L, Roos J, Bystedt H (2005). A 10-year follow-up study of titanium dioxide-blasted implants. Clin Implant Dent Relat Res; 7: 36-42.

Rocci A, Martignoni M, Gottlow J (2003). Immediate loading of Branemark System TiUnite and machined-surface implants in the posterior mandible: a randomized open-ended clinical trial. Clin Implant Dent Relat Res; 5 Suppl 1: 57-63.

Roccuzzo M, Wilson T (2002). A prospective study evaluating a protocol for 6 weeks' loading of SLA implants in the posterior maxilla: one year results. Clin Oral Implants Res; 13: 502-7.

Romeo E, Ghisolfi M Rozza R, Chiapasco M, Lops D (2006). Short (8-mm) dental implants in the rehabilitation of partial and complete edentulism: a 3- to 14-year longitudinal study. Int J Prosthodont; 19: 586-92.

Roos J, Sennerby L, Lekholm U, Jemt T, Gröndahl K, Albrektsson T (1997). A qualitative and quantitative method for evaluating implant success: a 5-year retrospective analysis of the Brånemark implant. Int J Oral Maxillofac Implants; 12: 504-514.

Schlumberger TL, Bowley JF, Maze GI, (1998). Intrusion phenomenon in combination tooth-implant restorations: a review of the literature. Int J Prosthodont; 80: 199-203.

Schnitman PA, Schulman LB (1980). Recommendations of the consensus development conference on dental implants. J Am Dent Assoc; 98: 373-7.

Schnitman PA, Wohrle PS, Rubenstein JE, DaSilva JD, Wang NH (1997). Ten-year results for Branemark implants immediately loaded with fixed prostheses at implant placement. Int J Oral Maxillofac Implants; 12: 495-503.

Schroeder A, Pohler O, Sutter F (1976). Tissue reaction to an implant of a titanium hollow cylinder with a titanium surface spray layer. SSO Schweiz Monatsschr Zahnheilkd; 86: 713-727.

Schulte W (1981). Tubingen implant made of Frialit: 5 years of experience. Dtsch Zahnarztl Z; 36: 544-50.

Schulte W, Heimke G (1980). Ceramic endosseous implants with an aluminum oxide base (Tubingen-type Frialit) Chir Dent Fr; 11: 50: 41-5.

Sennerby L, Roos J (1998). Surgical determinants of clinical success of osseointegrated oral implants: a review of the literature. Int J Prosthodont; 11: 408-20.

Shalabi MM, Manders P, Mulder J, Jansen JA, Creugers NH (2006). A meta-analysis of clinical studies to estimate the 4.5-year survival rate of implants placed with the osteotome technique. Int J Oral Maxillofac Implants; 22: 110-6.

Shin YK, Han CH, Heo SJ, Kim S, Chun HJ (2006). Radiographic evaluation of marginal bone level around implants with different neck designs after 1 year. Int J Oral Maxillofac Implants; 21: 789-94.

Siegel S, Castellan JN (1988). Nonparametric statistics for the behavioral sciences. McGraw-Hill, New York.

Smith DE, Zarb GA (1989). Criteria for success of osseointegrated endosseous implants. J Prosthet Dent; 62: 567-72.

Snauwaert K, Duyck J, van Steenberghe D, Quirynen M, Naert I (2000). Time dependent failure rate and marginal bone loss of implant supported prostheses: a 15-year follow-up study. Clin Oral Investig; 4: 13-20.

van Steenberghe D, Lekholm U, Bolender C, Folmer T, Henry P, Herrmann I, Higuchi K, Laney W, Linden U, Åstrand P (1990). Applicability of osseointegrated oral implants in the rehabilitation of partial edentulism: a prospective multicenter study on 558 fixtures. Int J Oral Maxillofac Implants; 5: 272-281.

van Steenberghe D, Klinge B, Linden U, Quirynen M, Herrmann I, Garpland C (1993). Periodontal indices around natural and titanium abutments: a longitudinal multicenter study. J Periodontol; 64: 538-541.

van Steenberghe D, Jacobs R, Desnyder M, Maffei G, Quirynen M (2002). The relative impact of local and endogenous patient-related factors on implant failure up to the abutment stage. Clin Oral Implants Res; 13: 617-22.

Straumann home page (2007). www.straumann.com

Strietzel FP, Rothe S, Reichart PA, Schmidt-Westhausen A (2006). Implant-prosthetic treatment in HIV-infected patients receiving highly active antiretroviral therapy: report of cases. Int J Oral Maxillofac Implants; 21: 951-6.

Sugermann PB, Barber MT (2002). Patient selection for endosseous dental implants: oral and systemic considerations. Int J Oral Maxillofac Implants; 17: 191-201.

Sullivan RM (2001). Implant dentistry and the concept of osseointegration: a historical perspective. J Calif Dent Assoc; 29: 737-45.

Sutter F, Krekeler G, Schwammberger AE, Sutter FJ (1992). Atraumatic surgical technique and implant bed preparation. Quintessence Int; 23: 811-6.

Svensson E (1993). Analysis of systematic and random differences between paired categorical data Göteborg University, Göteborg.

Taylor (1991). Fixed implant rehabilitation for the edentulous maxilla. Int J Oral Maxillofac Implants; 6: 329-37.

Testori T, Szmukler-Moncler S, Francetti L, Del Fabbro M, Scarano A, Piatelli A, Weinstein RL (2001). Immediate loading of Osseotite implants: a case report and histologic analysis after 4 months of occlusal loading. Int J Periodontics Restorative Dent; 21: 451-9

Truhlar RS, Morris HF, Ochi S, Winkler S (1994). Second-stage failures related to bone quality in patients receiving endosseous dental implants: DICRG Interim Report No. 7. Dental Implant Clinical Research Group. Implant Dent; 3: 252-5.

Uludag,B, Celic G (2006). Fabrication of a cement- and screw-retained multiunit implant restoration J Oral Implantol; 32: 248-50.

U. S. Food and Drug Administration (2007). www.fda.gov

Weng D, Jacobson Z, Tarnow D, Hurzeler MB, Faehn O, Sanevi F, Barkvoll Pstach RM (2003). A prospective multicenter clinical trial of 3i machined-surface implants: results after 6 years of follow-up. Int J Oral Maxillofac Implants; 18: 417-23.

Wennerberg A, Albrektsson T, Andersson B (1996). Bone tissue response to commercially pure titanium implants blasted with fine and coarse particles of aluminum oxide. Int J Oral Maxillofac Implants; 11: 38-45.

Wennström JL, Ekestubbe A, Gröndahl K, Karlsson S, Lindhe J (2004). Oral rehabilitation with implant-supported fixed partial dentures in periodontitis-susceptible subjects. A 5-year prospective study. J Clin Periodontol; 31: 713-24.

Wheeler SL, Holmes RE, Calhoun CJ (1996). Six-year clinical and histologic study of sinus-lift grafts. Int J Oral Maxillofac Implants; 11: 26-34.

WMA World Medical Association (2004). www.wma.net/e/ethicsunit/helsinki.htm

Yerit KC, Posch M, Seemann M, Hanisch S, Dortbudak O, Turhani D, Ozyuvaci H, Watzinger F, Ewers R (2006). Implant survival in mandibles of irradiated oral cancer patients. Clin Oral Implants Res; 17: 337-44.

Zarb GA (1983). Proceedings of the Torornto conference on osseointegration in clinical dentistry. St Louis MO: Mosby; Reprinted from the Journal of Prosthetic dentistry, vol 49 and 50; 1-84.

Zarb GA, Schmitt A (1990). The longitudinal clinical effectiveness of osseointegrated dental implants: the Toronto study. Part III: Problems and complications encountered. J Prosthet Dent; 64: 185-94.

Zitzmann NU, Marinello CP, Sendi P (2006). A cost-effectiveness analysis of implant overdentures. J Dent Res; 85: 717-21.

Åstrand P, Borg K, Gunne J, Olsson M (1991). Combination of natural teeth and osseointegrated implants as prosthesis abutments: a 2-year longitudinal study. Int J Oral Maxillofac Implants; 6: 305-12.

Åstrand P, Engquist B, Anzen B, Bergendal T, Hallman M, Karlsson U, Kvint S, Lysell L, Rundcrantz T (2002). Nonsubmerged and submerged implants in the treatment of the partially edentulous maxilla. Clin Implant Dent Relat Res; 4: 115-127.

Ödman J, Lekholm U, Jemt T, Brånemark PI, Tilhander B (1988). Osseointegrated titanium implants-a new approach in orthodontic treatment. Eur J Orthod; 10 : 98-105.

Örtorp A, Jemt T (2006). Clinical experiences with laser-welded titanium frameworks supported by implants in the edentulous mandible: a 10-year follow-up study. Clin Implant Dent Relat Res; 8: 198-209.

Abbreviations and Explanations

A

B

Binomial distribution: is the simplest distribution to be used for discrete data, representing the number of times a certain event occurs.

Bootstrap simulation: a simulation method to estimate statistical properties.

C

CI: Confidence interval

CSR: Cumulative success rate

Categorical distribution: means that the observations have qualitative values.

Censored: the subject is no longer followed, i.e. the endpoint.

Chi-square test: is a non parametric test for comparing distributions.

Cohort: is a statistical term for a limited group that has been followed.

Confidence intervals: are interval estimates including uncertainty of calculated values.

Contingency table: tables with two or more categorical characterisations.

Cumulative success rates (CSR:s): are calculated by taking the success rate (SR) of an actual time-interval multiplied with the cumulative success rate of the previous interval.

D

Dependence: the opposite to independence.

Discrete distribution: means that the observations take on separate numerical values, e.g. integers.

E

Endpoint: is often equal to death or a failed treatment (e.g. implant loss).

Exclusion criteria: a reason for not including as patient in a study.

F

Fisher's exact test: Fisher's exact test is used when 2 groups are compared by a 2x2 contingency table.

G

H

Hypothesis: is a specific assumption about a population. It concerns the value of a specific parameter, which characterises the population.

I

IMZ: Cylindrical implant with an HA coat.

Independence: exists if the joint probability density is the product of the individual probability densities, regarding two random variables, e.g. - A and B - are independent, when the probability is that any given value of A is not affected by value B.

Inclusion criteria: a reason for including a patient in a study.

J

Jackknife re-sampling method: is used to estimate statistical properties for data.

L

Life-table: Life-tables for survival analyses are methods to be used on cohort mortality populations.

Log rank test: is a statistical test for survival data.

M

MC-study: Multicentre-study.

Multiple tests: consist of a number of tests, with a given total multiple level of significance, which protects against wrongly rejecting any true hypothesis.

Multifactor analysis: means a test considering several background factors (covariates).

Multiple inference: means the prescribed total risk of error for several tests and confidence intervals.

Multivariate analyses: involve considerations of subjects, which can be characterised by several variables.

N

Nominal data: are more than two categories but no obvious ordering of the categories, e.g. implant surfaces.

Non-parametric tests: are used when qualitative data are to be measured, or limited assumptions about the distribution are at hand.

Normal distribution: is the most common distribution (also called Gaussian distribution), which is defined by mean and standard deviation.

Null hypothesis: is a basic assumption that there is no statistical difference, e.g. that two population means are equal.

O

Ordinal data: Non- numerical data defined by verbally described borders between the groups.

P

Parameter: is a description of the whole population, e.g. mean age, median blood pressure etc. A more statistical definition is a theoretical characterisation of the population, e.g. the probability (p) for a certain outcome in that population.

Parametric tests: are used when a specific assumption about the distribution (e.g. normal distribution) is available.

"Patients at risk" is a term used to describe - all patients (or e.g. successful implants) being alive at the start of the next time interval together with half of the withdrawn or dropped-out patients (implants) during that time interval.

Pearson's chi-squared test: The test is used to compare observed frequencies with those expected, where the null hypothesis of the row and column classification factors is independent.

Post hoc analysis: is used, when an overall hypothesis has been rejected, to identify the reasons for the rejection, or to identify which subgroup is responsible for the significant outcome.

Probability (p): is an estimate of how likely an event is to occur.

Prospective: means to have an agreement, prior to the start of the study (e.g. a study protocol), on how the upcoming events should be analysed.
***P*-value:** is the probability that obtained outcome or more extreme outcomes will occur, calculated under the null hypothesis.

Q

Qualitative: means a type of data that can also be called **categorical** (e.g. female/male).
Quantitative: means a type of data that can be counted or physically measured, i.e. numerical.

R

Regression analysis: is an analysis to study influence of factors on a variable.
Randomization: means a random allocation of subjects into different treatment groups to avoid bias.
Random Effect Model: is a statistical test used to handle data that is not necessarily independent.

S

SLA: Sandblasted and acid-etched surface.
Selected: is the opposite to random subject allocation.
Significance level: is the pre-selected probability of rejecting the null hypothesis, when it IS in fact true.
Standard deviation: the mean size of the variation of a random variable.
Standard error: the mean size of the variation of an estimate of a parameter.
Statistical method: concerns the mathematical evaluation of models of observed data e.g. a statistical test.
Statistical model: is a mathematical picture of the reality.
Statistical formulas: are the mathematical calculations to a statistical model.
Step by step method: is performed by adding one background factor at a time, e.g. patient-, jaw- and/or implant- characteristics.

T

TPS: Titanium plasma-sprayed surface.
Treatment protocols: are bridges or overdenture for edentulous patients and single tooth replacement or partial bridges.

U

"Unaccounted for": i.e. withdrawn or dropped–out patients.

V

Variable: is a description of a subject, e.g. gender, age, health status etc. A more statistical definition is an empirical characterisation of a subject.
Variance: Square of the standard deviation.

W

Withdrawn/Dropped-out: means subjects who have not completed the entire follow-up program.