Meta-Analysis in context: Chemotherapy versus Chemotherapy combined with Bisphosphonate Therapy in Multiple Myeloma Patients

Inaugural-Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry and Pharmacy
of Freie Universität Berlin

by

Jasmina Redzepovic
from Belgrade (Serbia)

February, 2009

Time period: **June, 2005 – February, 2009**

Freie Universität Berlin, Department of Biology, Chemistry and Pharmacy

In collaboration with

Bayer-Schering Pharma AG, Berlin
Department of Medical Communication, Oncology

1st Reviewer: **Prof. Dr. Ronald Gust**
2nd Reviewer: **Prof. Dr. Heinz Pertz**

Date of dissertation defence 4. June 2009

Acknowledgments

This PhD was made possible by the financial support of Bayer-Schering Pharma AG.

I am indebted to my supervisor at Schering AG (Berlin), Ms. May Gall, for providing me with the opportunity to work on this project, as well as her valuable advice and technical and financial support.

I would like to thank my supervisor at Freie Universität Berlin, Professor Ronald Gust, for his guidance, intelligent advice, unconditional support and ever-open door to students.

I am also indebted to my co-supervisor, Professor Heinz Pertz (Freie Universität Berlin), for his expertise in systematic reviewing and technical guidance in finishing my thesis. I am likewise indebted to Professor Ingo Ott (Technische Universität Braunschweig).

A thank to Dr. Mirsad Selimovic for his support.

I would also like to thank my colleagues from Bayer-Schering Pharma Gisela, Jacinta, Harald for their support and advices. A special thank to my friends Claudia, Martin and Tara and all others who encouraged and supported me. I greatly appreciate the guidance and wisdom of my husband Emanuele throughout this time.

I would also like to thank my parents, Svetlana and Nuho, and my sister Bisera, for their moral support and encouragement.

My daughter, Sara, although you are too young to know this, this is for you.

Introduction 6

Part 1. Methodology 12

1.1. Systematic Review Process 12

1.1.1. Introduction 12

1.1.2. Identification of a clinical problem - review question 12

1.1.3. Searching for studies 12

1.1.4. Selecting studies 14

1.1.5. Critical appraisal 15

1.1.6. Collection of data 16

1.1.7. Evaluation of evidence 16

1.2. Meta-analysis 18

1.2.1. Introduction 18

1.2.2. Steps involved in a meta-analysis 19

1.2.2.1. Research hypothesis and hypothesis testing 19

1.2.2.2. Effect Size Statistics 19

1.2.2.3. Inverse Variance Weights (ω) 21

1.2.2.4. Confidence Interval (CI) 22

1.2.2.5. Heterogeneity Analysis 23

Part 2. Multiple Myeloma Therapy and Management 25

2.1. Multiple myeloma therapy and management 25

2.1.1. Clinical symptoms 26

2.1.2. Prognostic factors 26

2.1.3. Epidemiology and risk factors 28

2.1.4. Therapy management 28

2.1.5. Therapy options 28

2.1.6. Evaluation of therapeutic outcome 30

2.2. Bone disease management: Bisphosphonates 30

2.2.1. Bisphosphonates: introduction 30

2.2.2. Risks/benefits of bisphosphonate therapy 32

Part 3. Results and Discussion 34

3.1. Systematic review of multiple myeloma clinical trials 34

3.1.1. Goals 34

3.1.2. Search strategy 34

3.1.3. Selection criteria 37

3.1.4. Included multiple myeloma trials 37

3.1.5. Excluded multiple myeloma trials 38

3.1.6. Included ONJ observational studies 40

3.1.7. Excluded ONJ studies 40

3.1.8. ONJ case reports 41
3.2. Meta-analyses of efficacy results 42
3.2.1. Meta-analysis of mortality reduction data 42
3.2.2. Meta-analysis of SRE reduction data 43
3.3. Bisphosphonate side-effect analysis 47
3.4. Discussion 52
3.4.1. Method 52
3.4.2. Multiple myeloma data evaluation and interpretation 54
3.4.2.1. Efficacy of bisphosphonates concerning measurement of outcomes 56
3.4.2.2. Efficacy of bisphosphonates concerning mortality reduction 57
3.4.2.3. Efficacy of bisphosphonates concerning skeletal related event reduction 58
3.4.2.4. Sensitivity analysis 59
3.4.2.5. Harms of bisphosphonate therapy 59
3.4.2.6. Clinical significance versus statistical significance 61
3.4.3. Limitations 63
3.5. Conclusion 65
Appendix. QUALITY EVALUATION OF THE INCLUDED MULTIPLE MYELOMA TRIALS 66
ABSTRACT 93
ZUSAMMENFASSUNG 94
GLOSSARY 96
ACRONYMS AND ABBREVIATIONS 100
REFERENCES 102

INTRODUCTION

This research develops a detailed systematic review of therapy evidence on bisphosphonate effects in multiple myeloma patients. The objectives of this work are threefold:

- To introduce and discuss the advantages and also the shortcomings of systematic reviews, providing insights into a still not especially diffused methodology in the healthcare decision-making process;
- To give an up-to-date overview of multiple myeloma therapy and management, focusing on bone disease management with bisphosphonates;
- To develop a systematic review, by means of a meta-analysis of study data, on bisphosphonate effects in multiple myeloma patients. In addition, data from observational studies and case reports on osteonecrosis of jaw (ONJ), a side-effect only in recent years associated with bisphosphonates, will be analyzed in order to obtain a new evidence for a risk-benefit reevaluation of bisphosphonate therapy.

A *systematic review* is defined as a method to synthesising and analysing the results of different research on a specific topic using a careful handling of data, mostly by means of statistical tools, called *meta-analysis*. This term encompasses all the methods and techniques of quantitative research synthesis used during the development of a systematic review. The aim is to reach conclusions that are stronger and more reliable for clinical practice than the results of individual studies. Systematic review and meta-analysis can both be understood as a type of survey in which research and study data, rather than people, are investigated. They presuppose the cumulative nature of research evidence. As such, they are part of the broader field of what is known as *evidence-based* medicine, which aims not only to use the *best evidence* [Sackett 1996] available to make better decisions in health cases, but also to provide methods to assess the quality and reliability of the evidence itself.

However, as a systematic review can reduce biases that may occur in single studies or experience-based evaluations by the determination of the effect size and the precision of study results, it could in some cases contradict individual studies or expert opinions.

A methodologically effective way to conduct a systematic review requires constant update of the analysis through the integration of data coming in from new studies and an attentive reviewing process. The Cochrane Collaboration has taken

up this challenge for health care issues. An international organisation founded in 1993 under the leadership of Iain Chalmers in the United Kingdom, the Cochrane Collaboration now includes over 11 500 researchers worldwide who apply a systematic process to review effects tested in biomedical randomised controlled trials. The results of these reviews are published in the Cochrane Library databases and are available on line.

Meta-analysis, the methodological backbone of systematic reviews, is not without drawbacks, however. These should be carefully considered at each stage of the reviewing process. Properly done, a meta-analysis takes considerably more effort and expertise than a conventional qualitative research review and requires the application of specialised knowledge of statistical tools.

Besides the problem of complexity, the most persistent criticism of meta-analysis is the *heterogeneity* of separate studies that are averaged together in a grand mean effect size. This criticism concerns not only differences in sampled populations across studies, but also differences in the methodological quality of study findings. In such cases, a good meta-analysis of poorly designed studies will still result in bad statistics. We could summarise these criticisms as the “mixing apples and oranges” and the “garbage-in, garbage-out” problems.

Regarding the first criticism, technical advances in meta-analysis now allow for statistical testing of homogeneity to decide wether a grouping of effect-sizes from isolated studies show a greater variation than would be expected from sampling error alone. This provides a test to determine if different results may or may not be comparable for the purposes of meta-analysis.

To address the second problem, one should keep strict methodological criteria for accepting studies which are to be included in the analysis. This assures the analysis is conducted on the “best evidence” only (that is, randomised controlled trials, or RCTs), but narrows the research domain, for instance, by relying on published studies alone. This may increase the reviewed effect, since it is very hard to publish studies that show no significant results. This is an example of the “publication bias” problem. There are also many other biases that may be introduced in the process of locating, selecting, and combining studies [Easterbrook 1991, Gøtzsche 1987, Egger 1997a].

However, when the quality of research has been adequate, research syntheses have had an important impact on policy and practice [Chalmers 2001, Chelimsky 1995].

This thesis aims to assist medical professionals, researchers, consumers and policy makers in updating current knowledge in the therapy management of bone disease in multiple myeloma patients and to address potential questions for further research.

Treatment of multiple myeloma, a B-cell cancer, is usually palliative. Bone disease affects 70% of multiple myeloma patients [Badros 2006] and causes several complications, such as pathologic fractures, severe bone pain, impaired mobility, spinal cord compression and hypercalcaemia, leading to greater morbidity and poorer quality of life (QOL) for patients. Management of these complications can include treatment with bisphosphonates.

Bisphosphonates are endogenous pyrophosphate analogues in which a carbon atom replaces the central oxygen atom with various side chains (P-C-P) [Rodan 1996]. The presence of a hydroxyl group (-OH) as a side chain of bisphosphonates enhances the capacity to chelate with calcium ions. This leads to their binding to hydroxyapatite bone mineral surfaces and their internalisation by bone-resorting osteoclast and osteoclast inhibition [Rogers 2000, Berenson 1998, Fleisch 1997]. There are eight bisphosphonates currently on the market. Out of these, the following five bisphosphonates have been tested through RCTs that aimed to improve the condition of patients with multiple myeloma: clodronate, etidronate, ibandrontate, pamidronate and zoldronate (Table 14, Table15). However, etidronate has not been proven beneficial in clinical trials on multiple myeloma [Belch 1991, Daragon 1993].

In order to identify all relevant studies for this thesis, MEDLINE, EMBASE and the Cochrane Controlled Trials Register (CCTR) were searched.

MEDLINE is probably the best known free access database. It is accessible via PubMed and alternative platforms that indexes approximately one-third of all medical and health-related literature. It was developed in the USA by the National Library of Medicine at the National Institutes of Health and uses a specific thesaurus for indexing. This is called Medical Subject Headings, or MeSH.

The EMBASE is the second best known indexing database. It requires a fee for use. It covers all MEDLINE database records and also allows access to over 1,800 biomedical journals, not covered in PubMed (approximately 25% more journals). EMBASE uses a thesaurus for indexing EMTREE to search MEDLINE and EMBASE records.

MeSH and EMBASE's EMTREE tools both allow for improved and broader searches, taking into account synonyms used both as text words and keywords (index terms) in the databases. By default, PubMed and EMBASE automatically “map” input to the appropriate subtree of MeSH/EMTREE synonyms. It is also possible to perform a specific MeSH/EMTREE terms search in order to identify all synonyms for related terms. When the search in PubMed is performed, it is possible to refine the search by clicking on subheadings, thereby restricting the MeSH terms to those found in specific contexts. The record of the subheadings is listed in Table 1. The use of this option should supplement rather than limit the search, so that improperly coded articles are not missed.

Table 1. List of subheadings available in PubMed to refine the MeSH term search.
Administration and dosage
Adverse effects
Analogs and derivatives
Analysis
Blood
Chemical synthesis
Chemistry
Diagnostic use
Economics
Immunology
Isolation and purification
Metabolism
Pharmacokinetics
Pharmacology supply and distribution
Therapeutic use
Toxicity
Urine

The Cochrane Controlled Trials Register (CCTR) is a bibliography of controlled trials, downloaded from databases such as MEDLINE and EMBASE or identified through manual searching of journals by Cochrane Collaboration participants.

Bisphosphonates were investigated in numerous preclinical and clinical trials for several indications. Their role in multiple myeloma concerning patients' mortality and QOL is still unclear. In this thesis efforts were made, if not to completely answer these questions, then at least to call attention to the need for further research and to offer some suggestions about how this could be done.

As in most clinical trials, safety data are generally assessed less rigorously than efficacy data [Lassere 2005]. Bisphosphonate RCTs also report exclusively GI side-effects. Other bisphosphonate side-effect evidence can be found in either descriptive studies such as case reports or case series, or observational studies without a control group. Their credibility depends on the number of cases reported. The reason behind this difference in their assessment in clinical trials is that the regulatory agencies (the US Food and Drug Administration or the FDA, the European Medicines Agency or the EMEA) give the drug market authorisations based primarily on determination of efficacy and not of safety. Furthermore, as there is a possibility that a rare side-effect may come to light after a drug is approved and

used by the larger population, post-marketing surveillance is a very important issue. However, safety reports to the agencies are made on a voluntary basis.

In order to show how seriously insufficient the existing monitoring systems of the post-marketing safety issue by national authorities are, a brief example will be described. This is, not only relevant to the issue of bisphosphonate therapy, but also the global health-care system.

In 2005, facing the large number of published case reports on ONJ, a side-effect, only in recent years associated with bisphosphonate use, the FDA's Oncologic Drugs Advisory Committee (ODAC) re-evaluated the risk/benefit ratio of intravenous bisphosphonates in multiple myeloma, breast cancer and prostate cancer patients being treated for metastatic bone disease. During the ODAC meeting, committee members (mostly independent experts) evaluated in an open forum presentations made by Novartis, the pharmaceutical sponsor of pamidronate and zoldronate, which were the drugs under review. FDA review staff and a third-party oncology expert, Dr. Brian Durie, also took part. The FDA review staff could not determine any proven occurrence of ONJ. A search carried out in the FDA`s Adverse Event Reporting System (AERS), known as MedWatch, shows that only 9 ONJ cases were voluntarily reported in 2002. Novartis was supporting a retrospective chart review of 2, 500 patients who had been treated with intravenous bisphosphonates at MD Anderson Clinical Center over the last ten years. 11 ONJ cases occurred in 631 breast cancer patients and 6 ONJ cases occurred in 148 multiple myeloma patients. Dr. Brian Durie presented an online survey on ONJ conducted by himself in collaboration with the International Myeloma Foundation. In total, 904 multiple myeloma and 299 breast cancer patients, responded to the survey, 116 ONJ cases were reported in the myeloma and 36 in the breast cancer patients. The meeting ended with the committee providing recommendations to the FDA suggesting that the benefits of pamidronate and zoldronate remained greater than the risks (It should be noted that warnings about ONJ were already stated on pamidronate and zoldronate product information labels).

Based on this example, it is clear that lack of evidence is a serious issue when the risks/benefits of therapies need be re-evaluated. This thesis has therefore undertaken a review of published ONJ case reports and observational studies to investigate the amount of evidence on this particular side-effect. A critical amount of evidence shows that ONJ is not a rare side-effect, leading to the conclusion that authorities require an in-depth investigation by manufacturers on this point. Finally, the general question will be addressed as to how the evidence of reported cases and observations should be assessed.

Ultimately, an analogous analysis for other bisphosphonate side-effects would also be necessary. This would, however, go beyond the scope of this work.

All bisphosphonates can cause hypocalcaemia, regardless of their method of administration, though this is infrequently found to be a clinically symptomatic problem. The most common side-effects with oral bisphosphonates (depending on whether an aminobisphosphonate or a non-aminobisphosphonate is being used) are upper gastrointestinal problems, such as gastritis [Van Holten-Verzantvoort 1993] and diarrhoea [Atula 2003]. IV infusions can be associated with injection site reaction and acute systematic inflammatory reactions [Tanvetyanon 2006].

Renal dysfunction is a particularly problematic adverse event which may also occur after infusion of IV bisphosphonates. However, the incidence may vary between agents, depending on renal uptake and elimination. The FDA reported that 72 patients suffered renal failure following zoledronate therapy [Chang 2003]. As a result, the product labels of pamidronate and zoledronate were amended to include additional nephrotoxicity warnings.

An oncologist's decision to use bisphosphonates for multiple myeloma patients is based on evidence about their efficacy and safety. A selection of which bisphosphonates must be used should ideally be based on evidence from comparative trials. Unfortunately, there is only one comparative Rosen [2004] study [Rosen 2001, 2003, 2004, Berenson 2001], and its data were not adequately reported. Comparative investigations should be urgently done in the future in order to decide on the best possible medical treatment.

The thesis is structured in the following way:

Part 1 is a general overview of the methodological approaches to the reviewing and meta-analysis process.

Part 2 is an overview of options in multiple myeloma therapy, including myeloma staging criteria, response criteria and compounds available. Additionally bone disease management with bisphosphonates is introduced.

Part 3 develops and discusses the results of the systematic review and meta-analysis conducted on efficacy data obtained from randomised trials in bisphosphonate therapy in multiple myeloma patients. In addition, a review of ONJ observational studies will be conducted, followed by a listekd summary of published case reports.

In this thesis, the quality of the studies referred to, was assessed by means of a checklist, in order to make the weighting decisions of each study more transparent. The Appendix includes copies of these checklists showing how each study evaluated according the quality criteria.

PART 1. METHODOLOGY

1.1. Systematic Review Process

1.1.1. Introduction

A systematic review of study data for a specific treatment involves careful and systematic data collection, quality data measurement and synthesis of the available information, either unpublished or published. The aim of systematic reviews in medical care is to answer a specific medical question based on all of the best evidence available. This kind of review is an essential tool for medical professionals who want to keep up with progress being made in their field.

This research used the formal method advised by the Cochrane Collaboration [Clarke 1999] to produce explicitly formulated, reproducible and up-to-date summaries of the treatment effects.

The process of systematic reviewing involves a number of steps: formulating of research questions, finding studies to potentially include, appraising and selecting of studies and summarising and synthesising relevant studies.

Before undertaking the systematic review, it is essential to develop a protocol outlining the question to be answered.

1.1.2. Identification of a clinical problem - review question

A detailed review protocol written in advance is important in order to avoid biases being introduced by decisions which could influence the data.
In the protocol for this research four evaluation tasks were set:

1. The data of all symptomatic patients regardless of their gender or age.
2. The treatment of interest: standard chemotherapy with bisphosphonates versus standard chemotherapy.
3. The eligibility criteria: randomised trials.
4. The outcomes to be looked for: skeletal related events, mortality and side-effects.

1.1.3. Searching for studies

Since there currently exist over 22. 000 journals and 10 million articles in the area of biomedical literature, a systematic searching approach is essential to identify the best evidence available to answer a clinical question [Pirozzo/Mayer 2004a]. In

order to identify all synonyms for related terms, the search strategy started with an initial search looking for MeSH terms in two major medical databases, MEDLINE and EMBASE. Limits were put in place to refine the search. Sensitivity and specificity of the research strategies were tested on the findings resulting from the search conducted by Cochrane Collaboration in their latest bisphosphonate therapy review [Djulbegovic 2002]. Additionally, a search was carried out in the Cochrane Controlled Trials Register using all identified synonyms for related terms. Subsequently, a search of the reference lists of all retrieved papers was conducted in order to identify any additional studies missed during the database searches.

As a positive outcomes are more likely to be accepted and published in journals than trials that fail. It would be biased, therefore, to include only published studies. Therefore, a search of unpublished studies was conducted especially through searches of databases listing conference proceedings. Literature in English and also publications in German and Italian have been considered.

The phases of the searching process are summarised in Table 2.

Table 2. Phases of the searching process (adopted from the open learning material of The Cochrane Collaboration).

Phase	Description	Strategies
Phase One	Initial search for literature	Searching in the Cochrane Database of Systematic Reviews (CDSR) for existing reviews. Determining what databases should be searched. Identifying key search terms by performing a MeSH/EMTREE terms search. Developing and documenting a search strategy.
Phase Two	Search for literature / publications	Searching in all databases using the identified search terms. Using inclusion criteria to determine which papers should be retrieved.
Phase Three	Bibliography search	Searching the reference lists and bibliographies of all relevant papers for additional studies.
Phase Four	Search for unpublished studies	Searching in the databases listing conference proceedings.

1.1.4. Selecting studies

Several authors have suggested methods for evaluating the quality of a clinical trial [Hayward 1995, Girling 2003, Altman 2001]. A hierarchy of study designs was first suggested by Campbell and Stanley in 1963 [Campbell 1963]. Levels of evidence based on study design were proposed by Fletcher and Sackett for the Canadian Taskforce on Periodic Health Examination in 1979 [Canadian Task Force 1979].

The level of evidence has been defined as a ranking in which study designs are classified according to their efficacy in eliminating biases. The level of evidence of the various types of study design commonly used to assess clinical and public health issues is shown in Table 3. As is commonly accepted, RCTs are considered the best evidence source for reviews that seek to evaluate effectiveness. This ranking was used as the inclusion criterion in the sections of this research evaluating treatment benefits. Further inclusion criteria are summarised in Table 4.

Table 3. Designation of levels of evidence (Source: Australian National Health and Medical Research Council, NHMRC, 1999).

Level of evidence	Study design
I	Evidence obtained from a systematic review of all relevant randomized controlled trials.
II	Evidence obtained from at least one properly-designed randomized controlled trial.
III-1	Evidence obtained from well-designed pseudorandomized controlled trials (alternate allocation or some other method).
III-2	Evidence obtained from comparative studies (including systematic reviews of such studies) with concurrent controls and allocation not randomized, cohort studies, case-control studies, or interrupted time series with a control group.
III-3	Evidence obtained from comparative studies with historical control, two or more single arm studies, or interrupted time series without a parallel control group.
IV	Evidence obtained from case series, either post-test or pretest/post-test.

Table 4. Inclusion criteria (Source: Australian National Health and Medical Research Council (NHMRC)1999).	
Inclusion Criteria	Description
Population	Adult symptomatic multiple myeloma patients.
Intervention	Bisphosphonate treatment in symptomatic multiple myeloma patients.
Comparison	Standard chemotherapy or another bisphosphonate
Outcome measures	Skeletal related events, mortality, side-effects.
Study design	Randomized controlled trial

However, there is a persistent controversy regarding reliance on the quality of study designs as the main credibility criterion for evidence concerning intervention effectiveness [Rychetnik 2002]. The debate concentrates on the primacy of the RCT for evaluating research findings, with respect to (a) the difficulty of conducting RCTs for complex programmatic interventions, (b) the difficulty of interpreting their results, and (c) the consequent tendency to downgrade the contribution of observational studies [Rychetnik 2002].

Since in clinical trials safety data are assessed less rigorously than efficacy data [Lassere 2005] (bisphosphonate RCTs also report almost exclusively gastrointestinal (GI) side-effects), it is necessary to obtain other bisphosphonate side-effect evidence from either descriptive studies such as case reports or case series, or observational studies without a control group. Their credibility depends on the number of cases reported.

This thesis, therefore, has thoroughly investigated the growing number of published reports on ONJ with the aim of assisting medical professionals, policy makers and authorities with the re-evaluation of the risk/benefit ratio of bisphosphonate therapy.

1.1.5. Critical appraisal

The preliminary critical appraisal aimed to evaluate the methods used by the investigators during a study, their impact on the research results and the subsequent quality of the evidence.

As part of the systematic review process, all included RCTs were first assessed for methodological rigour (Appendix Tables A1.1 to A9.3). The rigour refers to the methods used by investigators in the design of the study and its conduct in order to minimise the risk of biases.

The critical appraisal of RCTs aimed to identify sources of biases that may have resulted during the four main stages of research:

1. Selection of participants (number of participants, age, etc.)
2. Treatment provided to the study groups (is the study double blinded?)
3. Follow-up of participants
4. Measurement of outcomes

1.1.6. Collection of data

Data used by systematic reviews are the findings from individual studies reported in different statistical forms. A research finding is defined as a statistical representation of one empirical relationship among variables measured on a single subject sample [Lipsey 2001].

In this research, one of the findings of interest was the difference in mortality between the treatment group and control group, measured as the dependent variable representing a certain outcome construct. The data obtained from different studies should be comparable and have a similar statistical form or be configured in similar statistical forms in order to be carried forward for meta-analysis.

It is generally not appropriate to analyse and compare two study findings derived from different research designs and appearing in different statistical forms, even if they deal with the same topic. For example, two studies dealing with survival statistics may exist, one reporting on survival duration and the other on mortality rates. As it is often the case, many new medical interventions look fairly good when evaluated by survival time, but fail when evaluated by mortality rates [Woloshin/Schwarz 2007]. The reason is that an early diagnosis may lead to a longer survival time - but often with no improvement in mortality.

It is also important to note that one research study may represent more than one research finding. For example, this may occur if a study reports differences between male and female subjects, or within different time periods or perhaps mean differences between treatment and control groups in the periods immediately after treatment and in follow-up. A review could deal with all findings or only with selected findings, depending on the review question. The study eligibility criteria, upon which it was decided to include or exclude studies as appropriate for data retrieval, were set up in the review protocol (see 1.1.2).

1.1.7. Evaluation of evidence

The evaluation of evidence is based on a standardised approach following the Guidelines of the Australian NSW Health Department [Liddle 1996]. To assess the

quality of each included study, three checklists were used, which evaluate the following aspects:

1. *Descriptive study information*, covered authors and year of publication, a description of the study intervention, outcomes (both beneficial and harmful) other factors that might affect the outcome, characteristics of the study population, setting and the number of groups or sites in the study.
2. *Study evaluation criteria* were used to describe the main components of study quality. They provide information on whether the study was well conducted in order to prevent either overestimation or underestimation of the true effect of the intervention.
Table 5 summarises the codes used for the evaluation criteria. The codes were used as descriptive aids and not as quantitative scoring system.
3. *Study overall assessment* was used to assess and code the overall quality of the study using the codes in Table 6. The overall assessment of study quality was determined by the evaluation criteria and a judgment about the relative importance of each source of bias and the extent to which potential biases may have influenced results. Study quality codes A, B1, B2, C (Table 6) compatible with those of the Cochrane Collaboration [Oxman 1994] were used for overall quality assessment of study checklists.

Tables A1.1 to A9.3 in Appendix contain the complete evaluation checklists of the qualities of the studies used in the current research's analysis.

Table 5. Coding for evaluation criteria (Source: Australian National Health and Medical Research Council, NHMRC, 1999).

Evaluation criteria are coded according to the extent to which the criteria are fulfilled	Code
Criterion entirely fulfilled	a
Criterion mostly fulfilled	b1
Criterion mostly not fulfilled	b2
Criterion not at all fulfilled	c
Criterion not described adequately to classify as a, b1, b2 or c	?
Criterion not applicable	n/a

Table 6. Codes for overall quality assessment of study checklists (Source: Australian National Health and Medical Research Council (NHMRC)1999).		
Low risk of bias	A	All or most evaluation criteria from the checklist are fulfilled. Where evaluation criteria are not fulfilled, the conclusions of the study or review are thought very unlikely to alter.
Low to moderate risk of bias	B1	Some evaluation criteria from the checklist are fulfilled. Where evaluation criteria are not fulfilled or are not adequately described, the conclusions of the study or review are thought unlikely to alter.
Moderate to high risk of bias	B2	Some evaluation criteria from the checklist are fulfilled. Where evaluation criteria are not fulfilled or are not adequately described, the conclusions of the study or review are thought likely to alter.
High risk of bias	C	Few or no evaluation criteria fulfilled. Where evaluation criteria are not fulfilled or are not adequately described, the conclusions of the study or review are thought very likely to alter.

1.2. Meta-analysis

1.2.1. Introduction

Meta-analysis is a technique which uses special adaptations of conventional statistical methods to combine results of the selected set of studies in order to investigate, compare and interpret patterns of findings. This allows for making the best use of all the information gathered using a systematic reviewing process.

A critical step in a meta-analysis is to encode or “measure” study results on a numerical scale, so that that the resulting values can be meaningfully compared [Lipsey 2001]. The numerical measurement scale used in a meta-analysis is an effect-size statistic. Each research finding must be encoded as a value on the same effect size statistic. It involves transforming findings of individual studies into some common measure of treatment effect and then using a conventional statistical procedure to determine if there is an overall statistically significant effect.

The results of the meta-analysis can be displayed graphically, making interpretation easier for users of the review. This graphical display also allows a visual comparison of the findings of individual studies.

1.2.2. Steps involved in a meta-analysis

The following steps are involved in a meta-analysis: (a) defining the research hypothesis (association between exposure and outcomes believed to exist) (b) identifying a common effect size statistic for research findings of interest, (c) computing the weighted mean weighted by sample size, (d) determining the confidence interval for the mean and (e) testing for heterogeneity of the distribution with the aim of determining if combining the studies makes sense, (f) testing for statistical significance (hypothesis testing) aimed at answering the question concerning differences between variables, or outcome measures, in test and control groups.

1.2.2.1. Research hypothesis and hypothesis testing

Firstly, as with any other analytical investigation, a meta-analysis has to establish the research hypothesis which states that a difference exists between groups. Hypothesis testing determines if the difference between means of a variable in test and control groups occurs by chance. The customary scientific approach is to accept or reject the null hypothesis. The null hypothesis states that the findings from the study are the result of chance or random factors. Therefore the overall purpose of a typical trial is to reject the null hypothesis.

A statistical significant result means it is highly unlikely the difference found between groups could have occurred by chance alone. In a clinical research context, it is common to interpret a result as statistically significant if the difference between groups could have occurred by chance alone in less than 1 time in 20 (5% of the times). This is expressed as a p value lower than 0.05 ($p< 0.05$).

A study with a very large number of patients can show a statistical significance even when the actual clinical difference between the two groups is very small. The same could happen with the aggregated research findings of a meta-analysis. Clinicians must decide for themselves whether in these cases a statistically significant result has reasonable clinical significance [Mayer 2004].

1.2.2.2. Effect Size Statistics

The size of the effect in a meta-analysis is a statistical measure that represents each relevant treatment effect shown in the selected set of studies.
Different types of study outcomes, which are statistical representations of specific empirical relationships, require different effect size statistics. To combine identified research findings, a common effect size statistic must be determined.

The choice of the common effect size statistic depends on a number of parameters, the most important is the treatment effect measurement scale. The effect

of some outcomes will be measured on a continuous scale (e.g. blood pressure, serum cholesterol, QOL), while others are on a dichotomous or binary scale (e.g. improved/not improved, dead/alive). The most common measures of treatment effect are shown in Table 7.

The outcomes of interest for this thesis were reported as central tendency values, i.e. a mean or proportion, upon which the groups can be compared.

Table 7. Measures of treatment effect for continuous and binary outcomes (NHMR 2000).

Outcome measure	**Description**
	Continuous outcomes
Difference between group means	Difference between treatment and control groups in mean values of outcome variable.
Standardized difference	Differences between the treatment and control group means for each study, standardized by an estimate of the standard deviation of the measurements in that study. This removes the effect of the scale of measurement, but can be difficult to interpret.
Weighted difference in means	Average (pooled) difference between treatment and control groups in mean values across a group of studies using the same scale of measurement for the outcome (e.g. blood pressure measured in mmHg).
Standardized weighted mean difference	Average (pooled) standardized difference between treatment and control groups across a group of studies, where the outcome was measured using different scales with no natural conversion to a common measure (e.g. different depression scales or different quality-of-life instruments).
	Binary outcomes
Risk difference (RD)	Difference (absolute) between treatment and control groups in relation to the outcome. If the outcome represents an adverse event (such as death) and the risk difference is negative (below zero) this suggests that the treatment reduces the risk. In this situation the risk difference, without the negative sign, is called the absolute risk reduction.

Relative risk or risk ratio (RR)	Ratio of the risk proportions in the treatment and control groups in relation to the outcome. This expresses the risk of the outcome in the treatment group relative to that in the control group. For an adverse outcome, if the relative risk is below 1, this suggests that the treatment reduces the risk; its complement (1–relative risk) or relative risk reduction is also often used.
Odds ratio (OR)	Ratio of the odds of the outcome in the treatment group to the corresponding odds in the control group. Again, for an adverse outcome, an odds ratio below 1 indicates that the treatment reduces the risk. In some studies (e.g. population-based case-control studies) the odds ratio is a reasonable estimate of the relative risk. It is not a good estimate when the outcome is common or is measured as prevalence.
Hazard ratio (HR)	Ratio of the hazards in the treatment and control groups (when time to the outcome of interest is known); where the hazard is the probability of having the outcome at time t, given that the outcome has not occurred up to time t. Sometimes, the hazard ratio is referred to as the relative risk. For an adverse outcome, a hazard ratio less than1 indicates that the treatment reduces the risk of that outcome.
Number needed to treat (NNT)	The number of patients who have to be treated to prevent one event. It is calculated as the inverse of the risk difference without the negative sign (NNT = 1/RD). When the treatment increases the risk of the harmful outcome, then the inverse of the risk difference is called number needed to harm (NNH = 1/RD).

1.2.2.3. Inverse Variance Weights (ω)

A research finding represented as an effect size value based on the results of a large study is a more precise estimate of the corresponding therapy effect value than a finding based on the results of a smaller trial. Therefore, in a meta-analysis every effect size is weighted by its sample size and optimal weights are based on the standard error of the effect size. As a larger standard error corresponds to a less precise effect size value, the actual value is computed as the inverse of the squared standard error value (known in meta-analysis as the inverse variance weight). In practice, meta-analysis is typically conducted using a small number of effect size statistics with known standard errors (e.g. standardised mean difference, the odds-ratio).

A research finding in the form of the proportion of patients having a particular outcome, e.g. alive or dead, can be used as an effect size, as in following applicable statistics:

$$ES_p = p = \frac{k}{n}, \quad (1.1)$$

$$SE_p = \sqrt{\frac{p(1-p)}{n}}, \quad (1.2)$$

$$\omega_p = \frac{1}{SE_p^2} = \frac{n}{p(1-p)}, \quad (1.3)$$

where k is the number of patients with the characteristic of interest and n is the total number of patients in the observed group.
The mean effect size ($\overline{ES}$) of the research findings is computed by weighting each effect size (ES_{pi}) by the inverse of its variance (ω_{pi}), as following:

$$\overline{ES} = \frac{\sum(\omega_{pi} ES_{pi})}{\omega_{pi}}, \quad (1.4)$$

1.2.2.4. Confidence Interval (CI)

Confidence intervals (CIs) indicate the range within which the true value of an effect is likely to be. If CI does not include zero, then the mean effect size is statistically significant.

The CI for a mean effect size is based on the standard error of the mean and a critical value from z-distribution. The standard error of the mean ($SE_{\overline{ES}}$) is computed as the square root of the sum of the inverse variance weights ($\sum \omega_{pi}$) [Hedges and Olkin 1985] as shown in:

$$SE_{\overline{ES}} = \sqrt{\frac{1}{\sum \omega_{pi}}}, \quad (1.5)$$

To construct CI, the standard error is multiplied by the critical z-value representing the desired CI. The product is then added to the mean effect size for the upper limit ($\overline{ES}_U$), and subtracted for the lower limit ($\overline{ES}_L$) as shown in:

$$\overline{ES}_U = \overline{ES}_L + z_{(1-\alpha)}(SE_{\overline{ES}}), \quad (1.6)$$

$$\overline{ES}_L = \overline{ES} + z_{(1-\alpha)}(SE_{\overline{ES}}), \quad (1.7)$$

where $\overline{ES}$ is the mean effect size, $z_{(1-\alpha)}$ is the critical value for the z-distribution (1.96 for α= .05; 2.58 for α= 0.01), and $SE_{\overline{ES}}$ is the standard error of the mean effect size. If the CI does not include zero, the mean effect size is statistically significant at $p \leq \alpha$.

A direct test of significance of the mean effect size can be obtained by computing a z-test as shown:

$$z = \frac{\left|\overline{ES}\right|}{SE_{\overline{ES}}}, \quad (1.8)$$

The result of this formula is distributed as a standard normal variance. Therefore, if it exceeds 1.96 it is statistically significant with $p \leq .05$, two tailed and if it exceeds 2.58 it is significant with $p \leq .01$, two-tailed [Lipsey 2001].

1.2.2.5. Heterogeneity Analysis

To analyse the combined effect of a group of similar studies, it is necessary to check if their individual effects are similar enough to allow a meaningful combined estimate of the entire set.

Due to randomisation, a chance variation in the estimates is to be expected. It is thus necessary to test if there is more variation than expected by chance alone. When an excessive variation occurs, there is heterogeneity and, conversely, no homogeneity.

The heterogeneity analysis is a statistical test of the null hypothesis (H_0) assuming homogeneity. If the dispersion of effect sizes around their mean is greater than that expected from sampling error alone, then the H_0 stating there is homogeneity between groups is rejected. In this case, each effect size does not estimate a common effect mean and thereforea meta-analysis of them would be misleading.

The heterogeneity test is based on a chi-square (Chi^2) statistics, which is distributed as a Chi^2 with k -1 degrees of freedom where k is the number of effect sizes [Hedges, Olkin 1985]. The formula for Chi^2 is

$$Chi^2 = \sum \omega_{pi}(ES_{pi} - \overline{ES})^2, \qquad (1.9)$$

where ES_{pi} is the individual effect size for pi = 1 to k (the number of effect sizes), $\overline{ES}$ is the weighted mean effect size over the k effect sizes, and ω_{pi} is the individual weight for ES_{pi}. If Chi^2 exceeds the critical value for Chi^2 with k-1 degrees of freedom, then the H_0 of homogeneity is rejected.

With the (1.9) formula and a standard Chi^2 table from any ordinary statistics textbook, we can conduct a test for heterogeneity. A statistically significant Chi^2, indicates no heterogeneity among trial findings. No significant Chi^2, corroborates the assumption of homogeneity and therefore allows conduction of meta-analysis.

PART 2. MULTIPLE MYELOMA THERAPY AND MANAGEMENT

2.1. Multiple myeloma therapy and management

Multiple myeloma (also known simply as myeloma or plasma cell myeloma) is a B-cell cancer, or more precisely, plasma cell cancer. B-cell cancers are malignant clones of B-cells on various developmental states. Myeloma cells represent malignant plasma cells (malignant B-cells of the latest developmental state). As tumour cells, to a great extent, maintain the characteristics of the healthy cells from which they originate, multiple myeloma cells that originate from plasma cells migrate to the bone marrow (multiple myeloma is a bone marrow cancer) and produce antibodies [Janeway 2002]. Malignant plasma cells' abnormal antibodies are called paraproteins. There are various blood and urine tests available for quantitative measurement of paraprotein (also known as M protein, where M stands for monoclonal).

Malignant plasma cells continuously multiply and so the cancer grows inside or outside of the bones. Healthy bone marrow usually consists of less than 5% of plasma cells. The bone marrow of multiple myeloma patients will usually consist more than 30% of plasma cells. This percentage can increase to over 90%. Myeloma also stimulates bone-remodelling cells called osteoclasts. Simultaneously, it suppresses bone-building cells called osteoblasts. The pathogenese of the osteoclast bone resorption is understood as resulting from an abnormal cytokine signalling between malignant plasma cells, osteoclasts and osteoblasts. Increased levels of RANK-ligand produced by myeloma cells and marrow stromal cells coupled with a suppression soluble osteopetegrin (OPG) favour osteoclast bone resorption [Oxford handbook of oncology 2006]. Other cytokines such as interleukin 6 further support an excess of osteoclast activity [Oxford handbook of oncology 2006].

A slow and steady bone destruction caused by myeloma has symptoms which initially mimic those of osteoporosis. The bone damage, or *osteolytic lesions*, may lead to fractures of the long bones or compression fractures in the spine. Bone pain, especially in form of a severe back pain, is often a symptom of this disease. Furthermore, when myeloma cell growth occurs inside the marrow-producing bones, healthy cells (e.g. red blood cells, white blood cells, platelets) are crowded out by cancer cells, causing immune system impairment, and as such, increased susceptibility to infections, as well as tiredness and weakness.

2.1.1. Clinical symptoms

A mnemonic sometimes used to remember common myeloma-related organ or tissue impairment (end organ damage) is CRAB: C = calcium (elevated), R = renal failure, A = anemia, B = bone lesions. Patients with monoclonal plasma cell proliferation in their bone marrow without end organ damage can be considered asymptomatic. They do not require therapy but must be regularly monitored as they have a life-long risk of progressing to multiple myeloma or developing related malignancies. Common clinical features in multiple myeloma patients are the following:

The dissolution of bone (osteolysis) releases calcium. This causes the calcium levels in the blood to rise (hypercalcemia) and brings about the associated symptoms of thirst, polyuria, nausea, constipation, drowsiness, and even coma [Oxford handbook of oncology 2006].

The Bence-Jones protein is deposited in the renal tubules and leads to renal failure. Other factors contributing to renal failures are hypercalcemia, dehydratation, amyloid deposition and infection [Oxford handbook of oncology 2006]. Renal impairment occurs in up to 30% of patients at presentation and up to 50% of patients at some stage of the illness [Alexanian 1990] [Winearls 1995].

Anemia is found in almost two thirds (60%) of patients with multiple myeloma. Displacement of the healthy bone marrow cells by cancer cells can obstruct the production of normal red blood cells and thus lead to anemia [Desikan 2002].

Osteolytic destruction of the skeleton and hypercalcemia are the characteristic pathological features in the myeloma patient. There may be fractures of proximal long bones, ribs, sternum, and vertebral crush fractures [Oxford handbook of oncology 2006]. Compression of the spinal cord occurs in 5% of patients during the course of the disease [Kyle 2004]. Measures to reduce skeletal related events by multiple myeloma are important for optimising the patient's QOL.

2.1.2. Prognostic factors

The Durie staging system (Table 8) was introduced in 1975. This system shows three stages (I, II, and III) of the disease, each stage having a subclassification related to renal function. This staging system is based on measurements of levels of M-protein production (IgG, IgA, etc. and kappa, lambda values), the number of lytic bone lesions (bone x-ray results), hemoglobin (anemia parameter), serum calcium (bone destruction indicator) and myeloma cell mass (tumour burden).

Table 8. The Durie-Salmon staging system [Durie 1975].			
Stage	**I**	**II**	**III**
Criteria	All of the following: Hemoglobin > 10 g/dl Bone X-ray Normal bone structure (scale O) or solitary bone plasmacytoma only Low M-component production rates IgG < 5.0 g/dl IgA < 3.0 g/dl Urine light-chain M-component in electrophoresis <4 g/24 h	Fitting neither Stage I nor Stage II	One or more of the following: Hemoglobin < 8.5 g/dl Serum calcium > 12 mg/dl Advanced lytic bone lesions (scale 3) High M-component production rates IgG > 7.0 g/dl IgA > 5.0 g/dl Urine light chain M-component in electrophoresis > 12g/ 24h
Cancer cell mass	<0.6 x 1012/m^2 of the body surface	>0.6x 1012/m^2 of the body surface	>1.2 x 1012/m^2 of the body surface
Renal function	Serum creatinine value <2 mg/dl: **Stage A**	Serum creatinine value >2 mg/dl: **Stage B**	

However, several studies have identified serum beta-2-microglobulin as a more accurate prognostic factor and indicator for survival [Palumbo 2004 a] [Palumbo 2004 b] [Sonnenveld 2005].

This background provides the basis for the new international staging system ISS (International Staging System, Table 9) [International Myeloma Working Group 2003] which was compiled by an international cooperation in which clinical and laboratory parameters from 10,750 previously untreated, symptomatic patients with multiple myeloma were evaluated.

However, this system is based on only two factors, serum beta-2-microglobulin and serum albumin, and is independent of age, type of therapy and geographic region.

The analysis of prognostic factors is essential to compare outcomes within and between clinical trials. For individual patients the best staging system can predict survival outcome with a rate of around 70% sensitivity and specificity. Whether staging systems can beneficially influence choice of therapy is unproven [Smith 2005].

Table 9. New International Staging System (ISS) [International Myeloma Working Group 2003] [Greipp 2005].			
Stage	**I**	**II**	**III**
Criteria	beta-2-microglobulin <3.5 g/l Albumin >3.5 g/dl	beta-2-microglobulin <3.5 g/l Albumin <3.5 g/dl or beta-2-microglobulin 3.5-5.5 g/l	beta-2-microglobulin >5.5 g/l
Median survival (months)	62	45	29

2.1.3. Epidemiology and risk factors

With an estimated 86,000 new multiple myeloma cases per year worldwide [Parkin 2005], multiple myeloma is the second most prevalent blood cancer after non-Hodgkin's lymphoma. With its 62,546 deaths per year [Kamangar 2006] it accounts for approximately 1% of all cancers and 2% of all cancer deaths. Multiple myeloma is more common in the Afro-American population than in the white American population and occurs more often in men than in women (with a ratio of 3:2). This distribution probably reflects the fact that men are more likely to work in low-income industrial sectors associated with more risk factors and that lower socioeconomic status is associated with a higher risk [Baris 2000]. Recently, associations between obesity and multiple myeloma have been established [Calle 2003]. Finally, the risk of developing multiple myeloma increases with age [Schottenfeld 1996]. Multiple myeloma cannot be cured. The mean survival is 3 years and fewer than 10% of patients live longer than 10 years [Myeloma Trialists' Collaborative Group 1998].

2.1.4. Therapy management

The treatment strategy is determined mainly by the stage of the disease defined according to the ISS criteria and by the patient's age. Asymptomatic patients with multiple myeloma do not benefit from an early initiation of treatment [Hjorth1993] [Riccardi 2000] and drug therapy is therefore not indicated in this group. Patients with stage II and stage III disease should always start treatment.

2.1.5. Therapy options

The most common first-line of treatment is high-dose chemotherapy (HDT) with stem cell transplantation (SCT) or conventional chemotherapy. Patients may

also be given maintenance therapy. The treatment for relapsed or refractory multiple myeloma is not yet standardised but there are a number of effective treatment options available. The treatment goals are prolonged survival and improved quality of life. Drug compounds used in treatment of multiple myeloma are summarised in Table 10.

Table 10. Possible drug compounds available.

Pharmacological classification	Compounds
Alkylating agents	Melphalan Carmustine Cyclophosphamide Cisplatin in DTPACE (dexamethasone, thalidomide, cisplatin)
Anthracyclines	Doxorubicin or idarubicin
Alkaloids	Vincristine, vinorelbine in VAD (vincristine, doxorubicin and dexamethasone), dDV (vincristine, dexamethasone and liposomal doxorubicin) or VBCMP (vincristine, carmustine, cyclophosphamide, melphalan and prednisone)
Glucocorticoids	Prednisone, dexamethasone
Thalidomide and derivatives	Thalidomide, lenalidomide
Protease inhibitors	Bortezomib
Farnesyl transferase inhibitor	Tipifarnib
Cytokines	Interferons: IFN-α-2a and IFN-α-2b

2.1.6. Evaluation of therapeutic outcome

The criteria for measuring therapeutic response are declines in M-protein in serum and urine, normalisation of any anemia and hypoalbuminemia and osteolytic status. In current studies the response rates are evaluated according to the criteria of the EBMT (European Group for Blood and Bone Marrow Transplantation, Table 11).

Table 11. EBMT/IBMTR/ABMTR* criteria for response [Bladé 1998].

Response to Treatment	Description
Objective remission (OR)	Complete remission + partial response (CR+PR).
Complete remission (CR)	No M protein detectable in serum and urine by electrophoresis, normal serum calcium, stable skeletal status.
Near complete remission (nCR)	Positive results in immunofixation electrophoresis but no M protein detectable in serum and urine by less sensitive methods.
Partial response (PR)	M protein reduction 50% or more.
Very Good Partial Response (VGPR)	Protein reduction > 90%.
Minimal response (MR)	Protein reduction less than 50%.
Stable disease (SD))	No change in M protein.
Progressive disease (PD)	25% or more increase in M-protein. Increase in plasma cells in the bone marrow. New or larger bone lesions.

*EBMT, European Group for Blood and Marrow Transplantation; IBMTR, International Bone Marrow Transplant Registry; ABMTR, Autologous Blood and Marrow Transplant Registry.

2.2. Bone disease management: Bisphosphonates

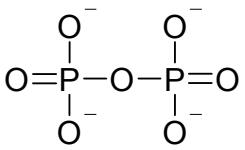
2.2.1. Bisphosphonates: introduction

Bisphosphonates are endogenous pyrophosphate analogues in which a carbon atom replaces the central oxygen atom with various side chains (P-C-P) [Rodan 1996]. The presence of a hydroxyl group (-OH) as a side chain of bisphosphonates enhances the capacity to chelate with calcium ions, which leads to a binding to hydroxyapatite bone mineral surfaces, their internalization by bone-resorting osteoclast and osteoclast inhibition [Rogers 2000, Berenson 1998, Fleisch 1997].

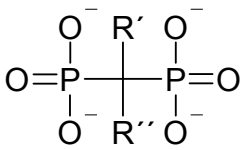
Alendronate, risedronate, ibandronate, pamidronate and zoledronate are all bisphosphonates containing nitrogen in a side chain and are called aminobisphosphonates (Figure 1).

In contrast, bisphosphonate without nitrogen in a side chain, such as clodronate, etidronate and tiludronate are called non-aminobisphosphonates (Figure 1). While nitrogencontaining bisphosphonates inhibit the mevalonate pathway (the main target being farnesyl diphosphate synthase), nonnitrogen- containing bisphosphonates are incorporated into hydrolytically stable analogues of adenosine triphosphate. Both events cause an impairment of osteoclast cell function and, ultimately, lead to osteoclast apoptosis [Brawn 2004], which indicates a therapeutic utilisation in multiple myeloma.

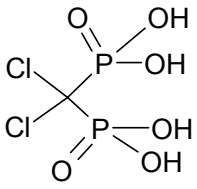
Figure 1. Bisphosphonate chemical structures.



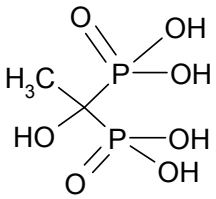
pyrophosphate



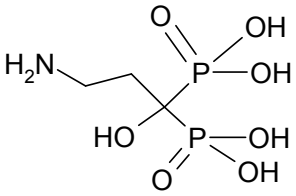
geminal bisphosphonate



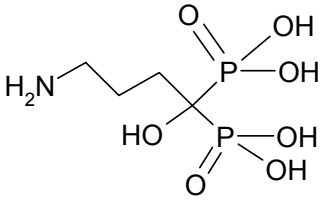
clodronate



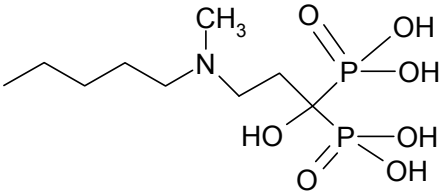
etidronate



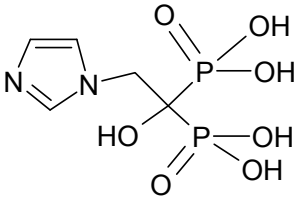
pamidronate



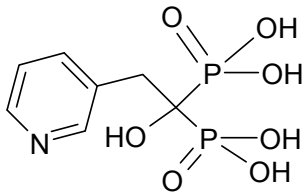
alendronate



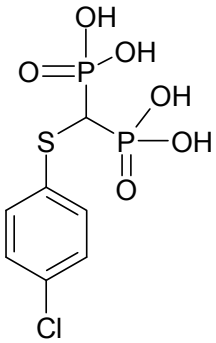
ibandronate



zoldronate



risedronate



tiludronate

The following five bisphosphonates have been tested through RCTs that aimed to improve outcomes of patients with multiple myeloma: clodronate, etidronate, ibandrontate, pamidronate and zoldronate (Table 14, Table15). Out of these five, etidronate was the only bisphosphonate not to have shown a benefit in clinical trials on multiple myeloma [Belch 1991, Daragon 1993].

2.2.2. Risks/benefits of bisphosphonate therapy

Bisphosphonates were investigated in numerous preclinical and clinical trials for several indications. Their role in multiple myeloma concerning patients' mortality and QOL is still unclear.

The myeloma trials do not report (or at least not adequately) on QOL data. The therapy efficacy represented as reduction of skeletal related events is, along with pain reduction, used as a surrogate end point of QOL instead of the actual QOL assessment. Therefore, future efforts should be made to appropriately use and standardize QOL measures in cancer in randomised controlled trials [Garratt 2009].

In this thesis, side-effects were also taken into account as an important part of therapy management. All bisphosphonates can cause hypocalcemia, regardless of their method of administration, though this is infrequently found to be a clinically symptomatic problem. The most common side-effects with oral bisphosphonates (depending on whether an aminobisphosphonate or a nonaminobisphosphonate is being used) are upper gastrointestinal troubles, such as gastritis [Van Holten-Verzantvoort 1993] and diarrhoea [Atula 2003]. IV infusions can be associated with injection site reaction and acute systematic inflammatory reactions [Tanvetyanon 2006].

Renal dysfunction is a particularly problematic adverse event which may also occur after infusion of IV bisphosphonates. However, the incidence may vary between agents, depending on renal uptake and elimination. The US Food and Drug Administration (FDA) reported that 72 patients suffered renal failure following zoledronate therapy [Chang 2003]. As a result, the product labels of pamidronate and zoledronate were amended to include additional nephrotoxicity warnings.

Even though only one randomised trial reports on ONJ as a side-effect [Attal 2006], a growing number of ONJ case reports and observational studies evaluating ONJ have been published in recent years (Table 19, 21, 22, 23). Typical symptoms for ONJ are pain, soft-tissue swelling and infection, loose teeth and exposed bone. Since 2003, when the first reports were published, several groups and organisations have developed or issued recommendations, position papers, or statements regarding bisphosphonate associated ONJ. An expert panel to look into the issue was sponsored by Novartis, the manufacturer of pamidronate and zoldonate. Its recommendations were first distributed as a white paper at the 2004 Annual ASCO Meet-

ing and later published [Ruggiero 2006]. In the following years, positioning papers by the American Academy of Oral Medicine and the American Academy of Oral and Maxillofacial Pathology [Migliorati 2005, Woo 2006], the American Association of Endodontists [AAE 2006] have been published. In 2007 the updated ASCO Guideline for the first time included recommendations regarding ONJ [ASCO 2007]. All these documents agree that prevention of bisphosphonate-associated ONJ is the best approach to the management of this complication. In effect this means, avoiding elective jaw procedures while undergoing bisphosphonate therapy, and having any routine dental exams and tooth extractions done prior to bisphosphonate therapy.

The side-effects experienced by patients in the trials included in the current analysis are the following:

Abdominal pain	Duodenal stomach ulcer	Nausea
Aggravation of tumor	Dysphagia/dyspepsia	Neutropenia
Allergic reaction	Dyspnea	Esophageal ulcer
Alopecia	Edema, lower limb	Osteonecrosis of the jaw
Anemia	Emesis	Pain in limb
Anorexia	Fatigue	Paresthesia
Arthralgia	Fever	Peripheral neuropathy
Back pain	Infection	Pyrexia
Bone pain	Insomnia	Renal
Cardiac	Hemorrhage	Thrombocytopenia
Cardiac arrhythmia	Headache	Thrombosis
Constipation	Heart failure	Vomiting
Cough	Hypocalcemia	Weakness
Depression	Mood change	Weight decrease
Diarrhea	Myalgia	
Dizziness		

Through systematic review and meta-analysis, this thesis made the effort, if not to make a complete risk/benefit assessment, to at least call attention to the need for further research and to give some suggestions about how this should be done.

PART 3. RESULTS AND DISCUSSION

3.1. Systematic review of multiple myeloma clinical trials

3.1.1. Goals

The primary goal is to determine whether adding bisphosphonates to standard chemotherapy in multiple myeloma patients decreases both mortality and the number of patients experiencing skeletal related events (defined as one or more manifestation of bone illnesses, pathological fractures, fractures and hypercalcemia) and to identify side-effects.

3.1.2. Search strategy

Two searching strategies were applied resulting in findings of different methodological quality. These must be analysed separately.

1) Search strategy aimed at identifying randomised multiple myeloma trials.

MEDLINE (1966 until January 2008), EMBASE (1974 until January 2008) and the Cochrane Controlled Trials Register (all years until October 2007) were searched to identify all randomised trials in multiple myeloma. The searched MeSH (Medical Subject Heading) and EMTREE terms for clodronate, pamidronate, etidronate, ibandronate, alendronate, risedronate, tiludronate and zoledronate were each cross-referenced with MeSH / EMTREE terms for multiple myeloma. The identified MeSH / EMTREE searched key words are listed in the Table 12. The searches were limited to reports of clinical trials in humans. Additionally, relevant references in articles published in peer-reviewed journals were also checked. A broad search of the Cochrane Controlled Trials Register and meeting proceedings of the American Society of Clinical Oncology from 2005 to 2006 was performed using following key words: bisphosphonates / disphosphonates, clodronate / clodronic acid, pamidronate / amidronate, ibadronate / ibadronic acid, alendronate, etidronate / etidronic acid, risedronate / risedronic acid, tiludronate / tiludronic acid, zoledronate / zoledronic acid, multiple, and myeloma / plasma cell.

Finally, after the searches were completed, the sensitivity and specificity of the search strategy was tested using the findings which resulted from the search conducted by Cochrane Collaboration for their last bisphosphonate therapy review in 2002 [Djulbegovic 2002].

2) Search strategy aimed at identifying observational studies and ONJ case reports.

Daily searches of MEDLINE and PubMed (January 2003 to October 2007) with the MeSH / EMTREE terms clodronate, pamidronate, etidronate, ibandronate alendronate, risedronate, tiludronate, zoledronate were each cross-referenced with MeSH / EMTREE identified key words for multiple myeloma, ostenecrosis and jaw diseases (Table 12, Table 13). A broad search of meeting proceedings of the American Society of Clinical Oncology from 2005 to 2006 was performed by using following key words: bisphosphonates / disphosphonates, clodronate / clodronic acid, pamidronate / amidronate, ibadronate / ibadronic acid, alendronate, etidronate / etidronic acid, risedronate / risedronic acid, tiludronate / tiludronic acid, zoledronate / zoledronic acid, multiple, myeloma/plasma cell, jaw, diseases, mandible, maxilla, osteonecrosis and necrosis. Additionally, important references to ONJ reviews published in peer-reviewed journals were taken into account.

Table 12. Identified MeSH / EMTREE searched key words.

Key term	Entry term
Clodronate	Acid, Clodronic Dichloromethane Diphosphonate Diphosphonate, Dichloromethane Dichloromethylenebisphosphonate Dichloromethanediphosphonic Acid Acid, Dichloromethanediphosphonic Dichloromethylene Biphosphonate Biphosphonate, Dichloromethylene Dichloromethylene Diphosphonate Diphosphonate, Dichloromethylene Cl2MDP Dichloromethanediphosphonate Clodronate Clodronate Disodium Disodium, Clodronate Clodronate Sodium Sodium, Clodronate Bonefos
Pamidronate	3-amino-1-hydroxypropylidene)-1,1-bisphosphonate 1-hydroxy-3-aminopropane-1,1-diphosphonic acid AHPrBP amino-1-hydroxypropane-1,1-diphosphonate aminohydroxypropylidene diphosphonate aminopropanehydroxydiphosphonate APD (3-amino-1-hydroxypropylidene)-1,1-biphosphonate amidronate pamidronate monosodium pamidronate disodium Aredia Novartis brand of pamidronate disodium salt pamidronate calcium
Ibandronate	ibandronic acid Bondronat Roche brand of ibandronic acid, sodium salt, monohydrate Bonviva ibandronic acid, sodium salt, monohydrate RPR 102289A RPR-102289A Boniva BM 21.0955 BM-21.0955 BM-210955 BM 210955 ibandronate 1-hydroxy-3-(methylpentylamino)propylidenebisphosphonate (1-hydroxy-3-(methylpentylamino)propylidene)bisphosphonate

Etidronate	Etidronic Acid Hydroxyethylidene Diphosphonic Acid Diphosphonic Acid, Hydroxyethylidene (1-hydroxyethylene)diphosphonic acid Didronel Etidronate HEDP Hydroxyethanediphosphonate Ethanehydroxydiphosphonate 1-Hydroxyethane-1,1-Diphosphonate 1 Hydroxyethane 1,1 Diphosphonate 1-Hydroxyethylidene-1,1-Bisphosphonate 1 Hydroxyethylidene 1,1 Bisphosphonate EHDP Ethanehydroxyphosphate 1,1-hydroxyethylenediphosphonate 1,1 hydroxyethylenediphosphonate Etidronate, Tetrapotassium Salt Salt Etidronate, Tetrapotassium Tetrapotassium Salt Etidronate (1-hydroxyethylene)diphosphonic acid, Tetrapotassium Salt Xidifon Xydiphone Xidiphon Dicalcium Etidronate Etidronate, Dicalcium Dicalcium EHDP EHDP, Dicalcium Etidronate Disodium Sodium Etidronate Etidronate, Sodium HEDSPA Phosphonic acid, (1-hydroxyethylidene)bis-, disodium salt Disodium 1-Hydroxyethylene Diphosphonate 1-Hydroxyethylene Diphosphonate, Disodium Diphosphonate, Disodium 1-Hydroxyethylene Disodium 1 Hydroxyethylene Diphosphonate Disodium Etidronate Etidronate, Disodium
Alendronate	4-Amino-1-Hydroxybutylidene 1,1-Biphosphonate 4 Amino 1 Hydroxybutylidene 1,1 Biphosphonate Aminohydroxybutane Bisphosphonate Bisphosphonate, Aminohydroxybutane MK-217 MK 217 MK217 Fosamax Alendronate Sodium Sodium, Alendronate Alendronate Monosodium Salt, Trihydrate
Risedronate	risedronic acid risedronate 2-(3-pyridinyl)-1-hydroxyethylidenebisphosphonate 2-(3-pyridinyl)-1-hydroxyethylidene-bisphosphonate risedronic acid, monosodium salt risedronate sodium Actonel Procter and Gamble brand of risedronic acid, monosodium salt Procter and Gamble Pharmaceuticals brand of risedronic acid, monosodium salt Aventis brand of risedronic acid, monosodium salt
Tiludronate	iludronic acid (4-chlorophenyl)thiomethylene bisphosphonic acid Cl2SMBP tiludronate (chloro-4-phenyl)thiomethylene bisphosphonate (chloro-4-phenyl)thiomethylene biphosphonate tiludronate disodium Skelid
Zoledronate	zoledronic acid 2-(imidazol-1-yl)-1-hydroxyethylidene-1,1-bisphosphonic acid zoledronate Zometa Novartis brand of zoledronic acid CGP 42446A CGP-42'446 CGP-42446 CGP 42'446
Multiple Myeloma	Multiple Myelomas Myeloma, Multiple Myelomas, Multiple Myeloma, Plasma-Cell Myeloma, Plasma Cell Myelomas, Plasma-Cell Plasma-Cell Myeloma Plasma-Cell Myelomas

Table 13. The further identified MeSH / EMTREE searched key words.		
Key word	Osteonecrosis	Jaw Diseases
Entry Terms	Osteonecroses Necrosis, Avascular, of Bone Avascular Necrosis of Bone Bone Avascular Necrosis Kienbock Disease Kienboeck Disease Kienboeck's Disease Kienboecks Disease Kienbock's Disease Necrosis, Aseptic, of Bone Aseptic Necrosis of Bone Bone Aseptic Necrosis	Disease, Jaw Diseases, Jaw Jaw Disease
Previous Indexing	Bone Disease (1966-1976) Necrosis (1966-1976)	Bone Disease (1966) Jaw (1966) Mandible (1966) Maxilla (1966)

3.1.3. Selection criteria

For the evaluation of efficacy of bisphosphonates in multiple myeloma patients, randomised trials with a parallel design compared with placebo or no treatment as a control group were taken in consideration for inclusion (Table 14, 15).
For the evaluation of ONJ as a side-effect of bisphosphonates in multiple myeloma patients, observational studies reporting on frequency were included (Table 21). All case reports of multiple myeloma patients experiencing ONJ were gathered and listed (Table 19, 22, 23).

3.1.4. Included multiple myeloma trials

Study eligibility criteria were defined according to the review protocol developed beforehand (section 1.1.2.). Thirteen RCTs that satisfied the criteria were identified. Four of these were excluded due to a major publication bias (for more details, see Table 15 about exclusion reasons). The remaining nine studies (Table 14) also showed elements of publication bias, meaning that, out of thirteen, only six trial data could be used for meta-analysis of mortality and seven trial data for meta-analysis of skeletal related events reduction. This illustrates the gravity of the inaccessibility of data due to publication bias.

Table 14. Summary of included multiple myeloma trial.

First author, year	Study type	Treatment groups	No. of BP pts E (enrolled) /R (randomized)	No. of control pts E (enrolled) /R (randomized)	Route, dose, frequency	Treatment duration	Outcomes
Clodronate							
Heim 1995	Not double blind, not placebo controlled	Clodronate vs observation	R= 39	R=32	1600 mg/d po	12 mo	SRE (total), total fractures, pain, calcium, adverse events
Lahtinen 1992	Double blind	Clodronate vs placebo	E=R=168	E=R=168	400 mg capsules po tid	24 mo	Mortality, non-vertebral fractures, pain, calcium
McCloskey 2001	Double blind	Clodronate vs placebo	E=R=264	E=R=272	400 mg capsules po tid	24 mo	SRE (total), total fractures, vertebral fractures, non-vertebral fractures, total mortality, pain, calcium
Pamidronate							
Attal 2006	Not double blind, not placebo controlled	Pamidronate vs pamidronate and thalidomide vs placebo	R (pamidronate)= 196 R (pamidronate and thalidomide)=201	R=200	Pamidronate 90 mg IV, every 4 weeks and 400 mg pamidronate and thalidomide, po a minimum dose reduction of 50 mg for treatment related toxicity.	30 mo	SRE (total), response rates, event-free, relapse-free, overall survival
Berenson 1998	Double blind	Pamidronate vs placebo	E=205 R=198	E=187 R= 173	90 mg in 500 ml of 5% dextrose in water every 4 weeks.	21 mo	SRE (total), vertebral fractures, non-vertebral fractures, total mortality, pain, calcium, adverse events
Brincker 1998	Double blind	Pamidronate vs placebo	R=152	R=148	75 mg capsules po bid	24 mo	SRE (total), pain, calcium, adverse events
Kraj 2000	Not double blind, not placebo controlled	Pamidronate vs placebo	E=R=23	E=R=23	60 mg IV, every 4 weeks		Total mortality, vertebral fractures
Etidronate							
Belch 1991	Double blind	Etidronate vs placebo	E=98 R=92	E=78 R=74	5 mg/kg/d	Until death or discontinuation	SRE (total), pain, calcium, survival
Ibandronate							
Menssen 2002	Double blind	Ibandronate vs placebo	E=107 R=99	E=107R=99	2 mg IV every mo.	24 mo	SRE/year, vertebral fractures, non-vertebral fractures, total mortality, pain, hypercalcemia

3.1.5. Excluded multiple myeloma trials

Table 15 summarises the studies which contain evidence of interest but ultimately proved to be inaccessible due to publication bias. Table 16 includes studies and publications which did not satisfy the inclusion criteria.

As some exclusion decisions may be self-evident, since they were applied for study exclusion (Table 16), others may need further explanation, such as the case of clinical studies examining bisphosphonate therapy in asymptomatic patients [Musto 2003, Musto 2008, Barlogie 2008] which were excluded, because it's the benefit of this therapy in this population is highly controversial. The Terpos [2000] study was excluded as it reported purely in terms of pamidronate effect on markers of bone resorption. The outcomes of interest for this evaluation, skeletal related events, side-effects or survival were not reported.

Table 15. Summary of the excluded multiple myeloma trials	
First author, year	**Reason for exclusion**
Clodronate	
Delmas 1982	Confusing reporting
Pamidronate	
Terpos 2003a	No published results of interest
Etidronate	
Daragon 1993	Not sufficient reporting
Pamidronate/ Zoldronate	
Rosen 2004	Data not stratified by illness

Tabel 16. Summary of the excluded multiple myeloma RCTs trials	
First author, year	Reason for exclusion
Clodronate	
Adam 1996	Not randomized
Clemens 1993	More up-to-date data published in 1995 (Heim 1995)
Merlini 1990	Pseudorandomized study (treatment allocation was performed on alternate days, and not according to randomized allocation sequence)
McCloskey 1998	More up-to-date data published in 2001 (McCloskey 2001)
Pamidronate	
Abildgaard 1998	Subgroup analysis of a larger trial (Brinker 1998)
Ali 2001	Not randomized (see also zoldronate)
Berenson 1996	More up-to-date data published in 1998 (Berenson 1998)
Musto 2003	Asymptomatid patients
Caparrotti 2003	Not randomized. A combination therapy
Ciepłuch 2002	Not randomized. A combination therapy
Martin 2002	Asymptomatic patients
Morris 2001	Not randomized. A combination therapy
Barlogie 2008	Asymptomatic patients
Kraj 2000b	Duplicate publication (Kraj 2000)
Ibandronate	
Bergner 2007	Not randomized
Coleman 1999	Not extractable data for MM patients
Fontana 1998	More up-to-date published in 2002 (Menssen 2002)
Terpos 2003	No published data of interest
Zoldronate	
Ali 2001	Not randomized (see also pamidronate)
Musto 2008	Asymptomatic patients
Spencer 2008	Not randomized. A combination therapy
Tassinari 2007	An observational study
Tosi 2006a	A combination therapy
Vogel 2004	Not randomized

3.1.6. Included ONJ observational studies

ONJ has not been reported in RCTs except Attal [2006] study. Regarding the growing amount of new evidence being reported on this side-effect, it is necessary to re-evaluate risk/benefit of bisphosphonate therapy. This reevaluation could be based on observational studies reporting on the ONJ frequency. The eligible studies are listed in Table 17.

Table 17. Included ONJ studies.

First author, year	Study type
Badros 2006	Retrospective study
Calvo-Villas 2006	Not clear
Corso 2007	Retrospective study
Dimopoulos 2006	Retrospective study from 1997; Prospective from 2003 to 2005
Garcia-Gara 2006	Retrospective study
Tosi 2006b	Retrospective study
Zervas 2006	Retrospective study from 1991, prospective from 2001 to 2006

3.1.7. Excluded ONJ studies

Studies in which data on multiple myeloma were not extractable or available were excluded [Bujanda 2007, Hoff 2006]. Frequency of ONJ was used as and inclusion criterion only if it was reported in a reliable way. This was not the case with the Kut [2004] study, which was therefore excluded. The excluded studies are summarised in Table 18.

Table 18. Summary of excluded ONJ studies.

1st author, year	Reason for exclusion
Bujanda 2007	No multiple myeloma pts with ONJ
Hoff 2006	Not extractable data for MM pts (abstract)
Kut 2004	ASH 2004 (abstract No- 4933): Approximately 600 MM pts. The reported frequency: 7 pts. Exclusion due to not imprecise reporting (e.g. “approximately 600 MM pts”)

3.1.8. ONJ case reports

This thesis investigated the growing number of ONJ cases recently reported, with a view to assisting medical professionals, researchers, consumers and policy makers in their re-evaluation of the risk/benefits of bisphosphonate therapy. A list of fifty-three identified case reports is shown in Tables 19, 22 and 23.

Table 19. Summary of ONJ case reports

1st author, year	Comments
Abu-Id 2006	Abstract, not extractable data for MM pts
Agrillo 2006	Not extractable data for MM pts
Bagan 2006	
Battley 2006	
Braun 2006	
Broglia 2006	
Capalbo 2006	
Carneiro 2006	
Carter 2005	
Clarke 2007	
Curi 2007	
Dannemann 2007	
Diego 2007	
Dimitrakopoulos 2006	
Elad 2006	
Estilo 2004	Not extractable data for MM pts
Ficarra 2005	
Gibbs 2005	Abstract, not extractable data for MM pts
Hansen 2006	
Hay 2006	
Herbozo 2007	
Kademani 2006	
Katz 2005	
Khamaisi 2006	
Kumar 2007	
Lenz 2005	
Lugassy 2004	
Magopoulos 2007	
Marunick 2005	
Marx 2005	Not extractable data for MM pts
Mavrokokki 2007	
Melo 2005	
Merigo 2006	
Migliorati 2005	
Montazeri 2007	
Mortensen 2007	
Murad 2007	
Pires 2005	
Pozzi 2007	
Purcell and Boyd 2005	
Ruggiero 2004	
Pastor-Zuazaga 2006	
Phal 2007	
Polizzotto 2006	
Salesi 2006	
Senel 2007	
Sitters 2005	
Treister 2006	
Vannucchi 2005	
Walter 2007	
Wutzl 2006	
Yeo 2005	
Zarychanski 2006	

3.2. Meta-analyses of efficacy results

The meta-analysis was conducted as a measurement technique for the evaluation of efficacy of bisphosphonates regarding their ability to reduce the number of skeletal related events (SREs) and mortality.

This thesis avoids any quantification of findings which were not measured appropriately. For example, in contrast to the last Cochrane systematic review [Djulbegovic 2002], pain reduction by bisphosphonates was not assessed, because this should be tested against an appropriate palliative opiate treatment.

3.2.1. Meta-analysis of mortality reduction data

From the thirteen multiple myeloma trials that satisfied the selection criteria, six were included in the meta-analysis of the mortality data. Mortality data from seven other trials were not extractable for the purpose of this meta-analysis. Three of this group of excluded studies reported that mortality in comparison with the placebo group was not significantly reduced. One of these studies was recent and large with 597 patients [Attal 2006]. Data extracted from 6 randomised trials involved 1673 patients. There were 475 deaths among the 849 patients treated with bisphosphonates and 487 deaths among the 824 control patients. The identified heterogeneity (Figure 3) was investigated through a sensitivity analysis by excluding an outlier (Belch 1991 study). The resulting heterogeneity was not significant (Chi^2=3.35; degrees of freedom (df)=4; p=0.50) (Figure 4). The corresponding odds ratio of 0.75 (95% CI: 0.59-0.95; p=0.02) indicates an advantage regarding the mortality of patients treated by bisphosphonates.

In this thesis it was assumed that bisphosphonate effects on mortality can be generalised with data from studies using at least three different bisphosphonates. In order to get an analysis about mortality effects concerning three and not two bisphosphonates, the data for one of these, ibadronate, were obtained from the last Cochrane Collaboration Review [Djulbegovic 2002], since the single publication about this bisphosphonate [Menssen 2002] did not numerically report on mortality.

The sensitivity analysis obtained by excluding the Menssen study [2002] did not show significantly different results. The heterogeneity among trials tested by the Chi^2 test (Chi^2 =1.37; degrees of freedom (df)=3; p=0.71) was not significant. As above, the corresponding odds ratio of 0.69 (95% CI: 0.53-0.90; p=0.007) also indicates an advantage regarding mortality of patients treated by bisphosphonates (Figure 5).

3.2.2. Meta-analysis of SRE reduction data

From thirteen selected studies, data from seven trials were included in the meta-analysis of skeletal related event reduction. Among 880 patients treated with bisphosphonates, 298 experienced skeletal related events in comparison with 301 patients experiencing skeletal related events among the 973 control patients. The Chi^2 test (Chi^2 =8.55; degrees of freedom (df)=6; p=0.20) shows there was no significant heterogeneity among trials. The odds ratio of 0.89 (95% CI: 0.72-1.09; p=0.27) indicates no beneficial effect of bisphosphonates on the number of patients experiencing general skeletal events (Figure 6).

Figure 3: Efficacy of bisphosphonates measured as mortality reduction.

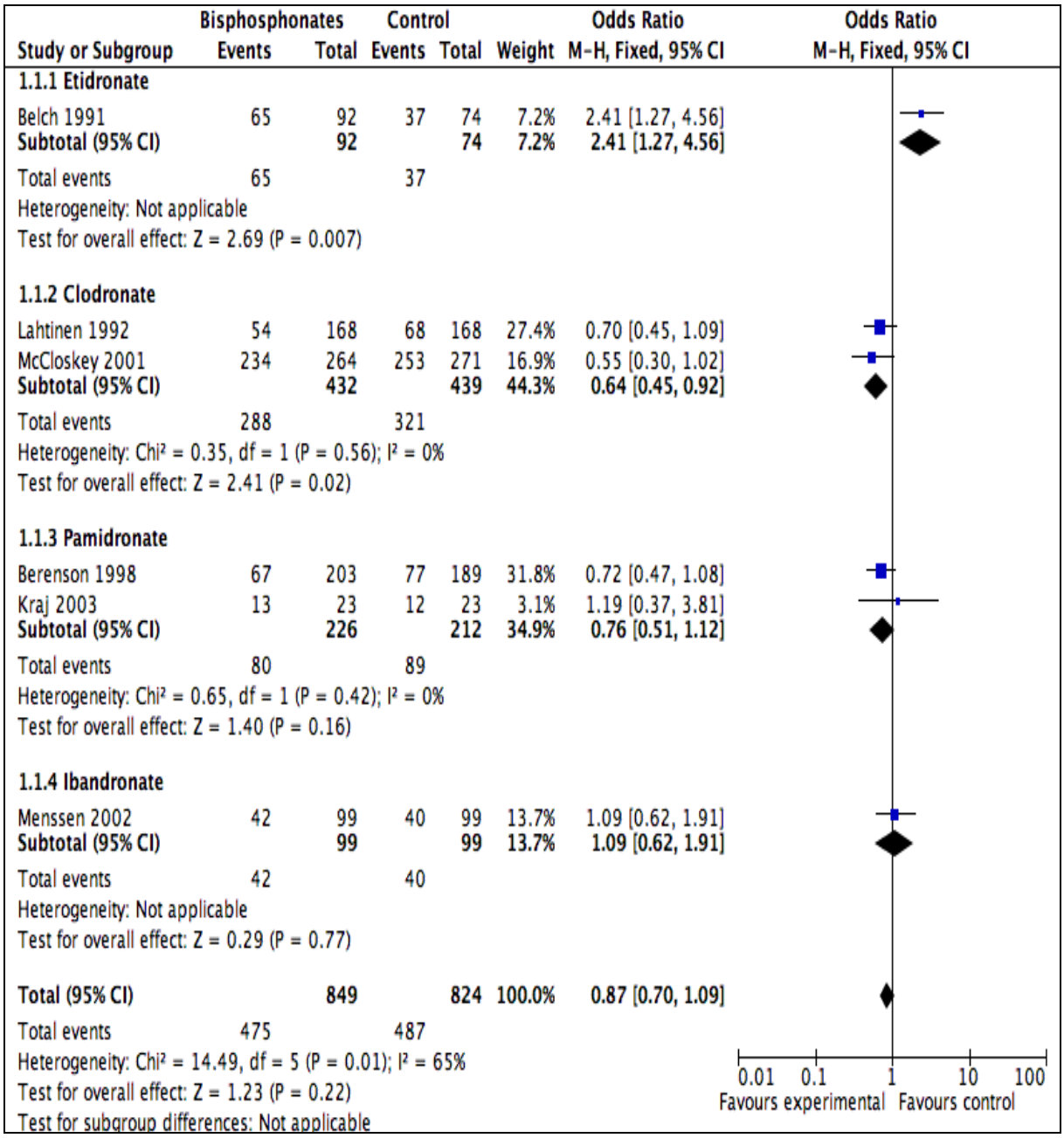


Figure 4. Sensitivity analysis (mortality assessment without the outlier, Belch 1991).

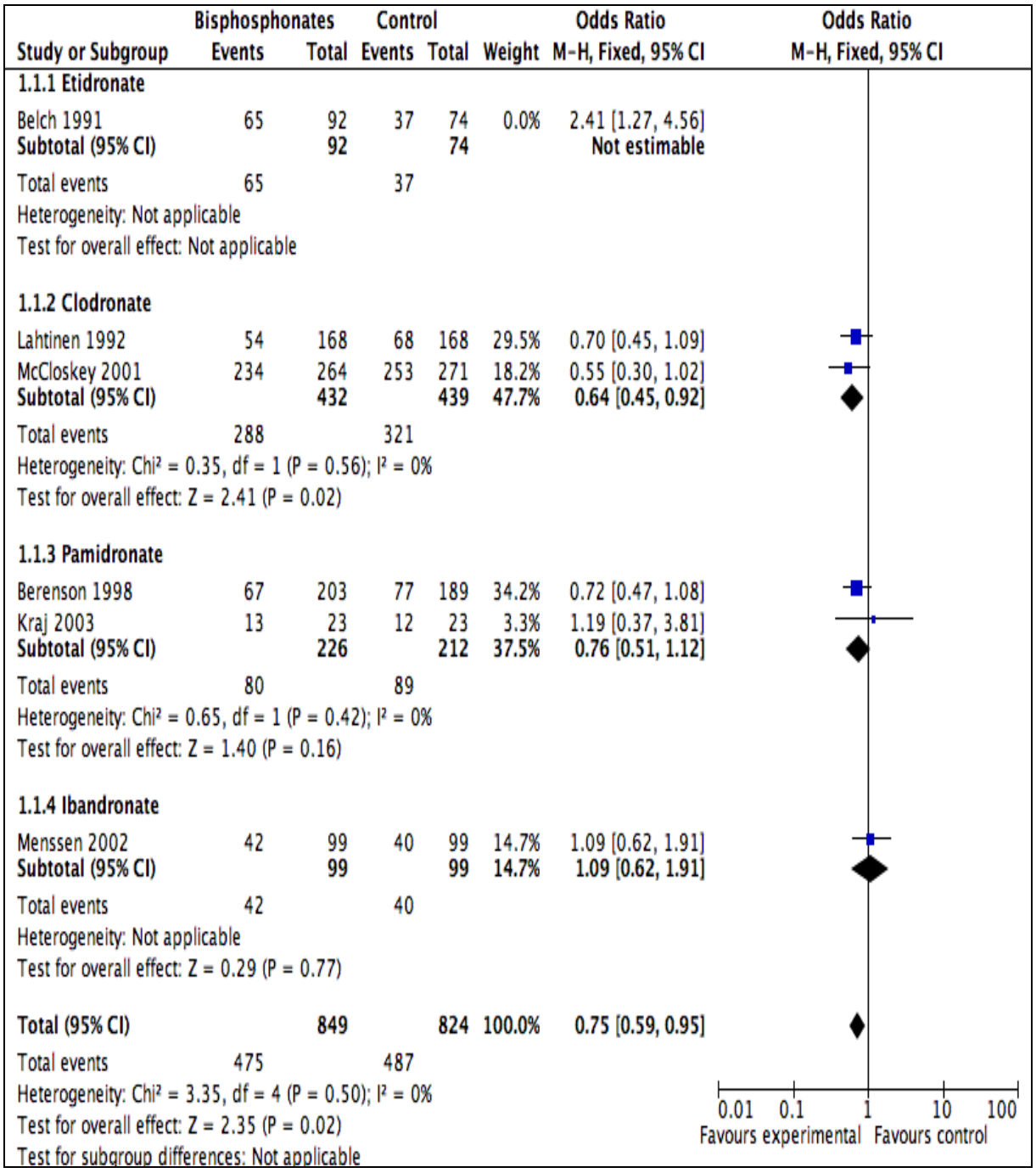


Figure 5. Sensitivity analysis (mortality assessment without Menssen 2002 study).

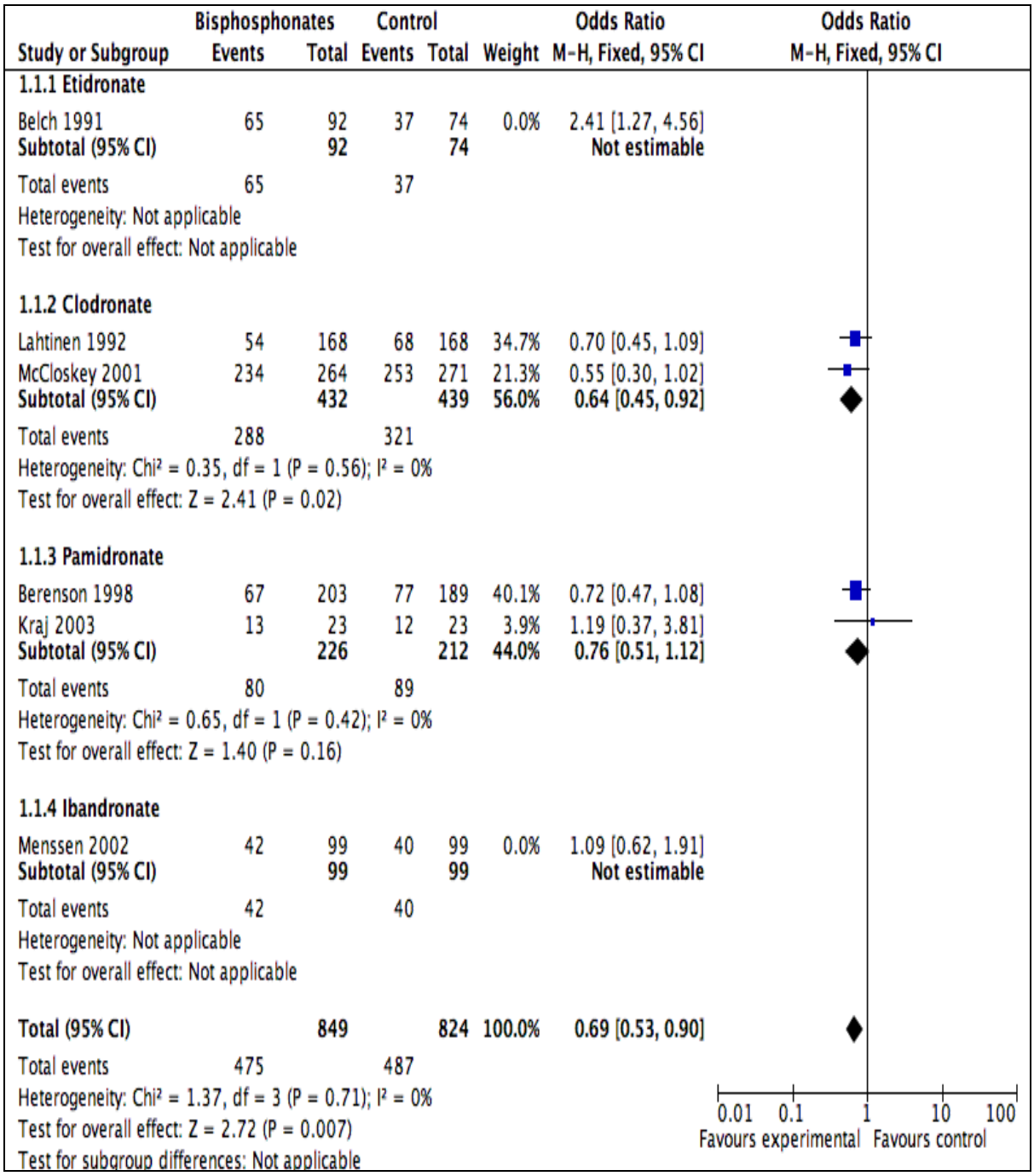
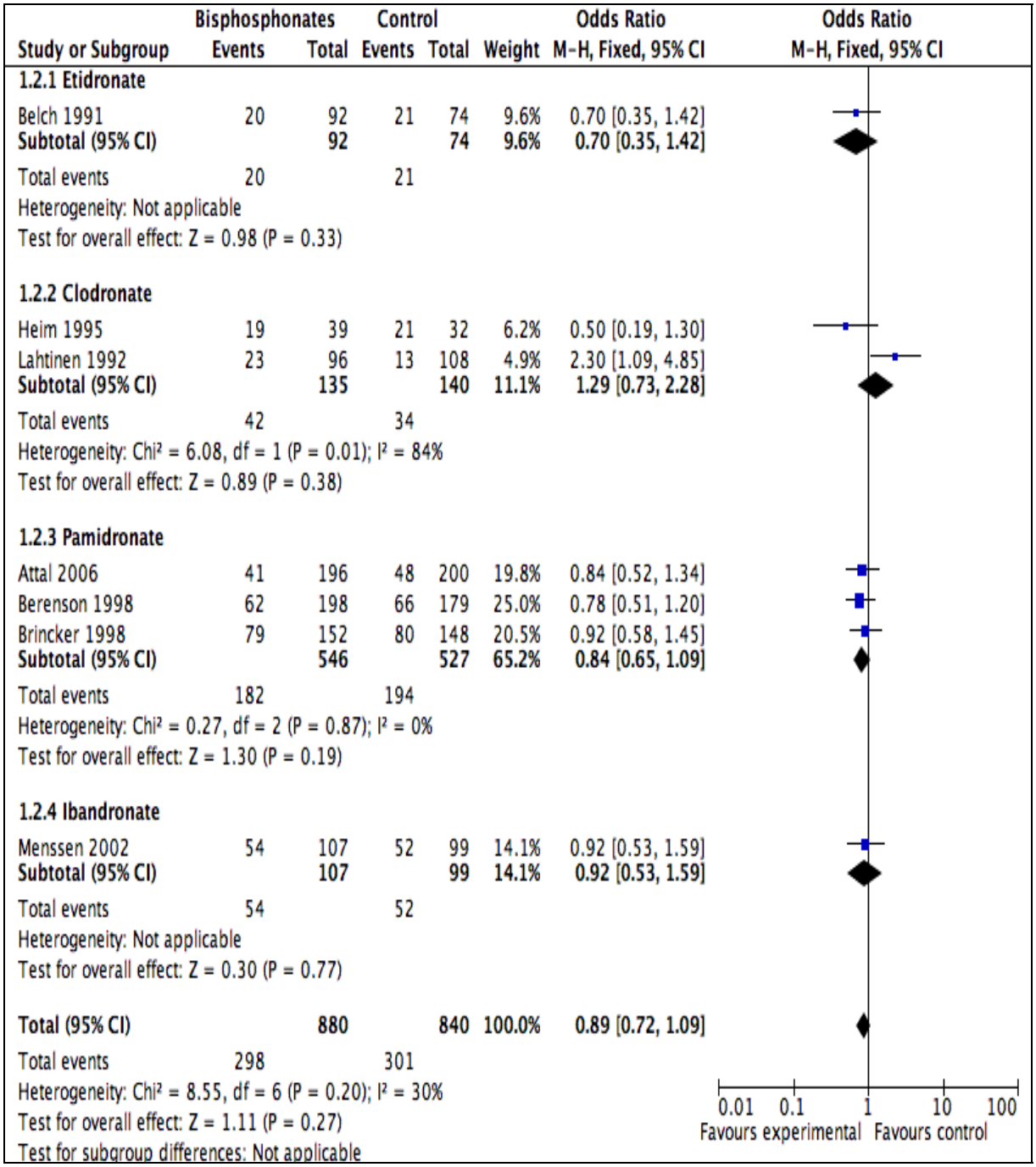


Figure 6. Efficacy of bisphosphonates measured as reduction of skeletal related events incidence.



3.3. Bisphosphonate side-effect analysis

Three out of the nine included multiple myeloma studies [Attal 2006, Lahtinen 1992, Berenson 1998] adequately reported side-effects, three [McCloskey 2001, Kraj 2000, Belch 1991] did not report side-effects at all and three [Brincker 1998, Heim 1995, Menssen 2002] did not adequately report side-effects. Four studies reported gastrointestinal side-effects as commonest toxicity. Only the study by Attal [2006] reported ONJ.

As the majority of side-effects were reported in a qualitative, narrative manner (Table 20) an analytical comparison among studies by type of identified side-effects and their frequencies would be quite misleading.

Seven observational trials with 1068 patients regarding ONJ were included in the present analysis (Table 21). The analysis of their findings resulted in very heterogeneous ONJ frequencies for various bisphosphonates (range: 0 to 51.51%).

Fifty-three case reports of multiple myeloma patients presenting this side-effect are listed in Table 22 and 23. In total, 46.68% of all ONJ cases were not stratified according to the illness. ONJ was more common in men than in women (50.42% vs. 49.58% of all patients and 55.29% vs. 44.71% in multiple myeloma patients) (Table 22). ONJ was present in the mandible site in 63.56% of all cases and 63.77% of multiple myeloma cases. Maxilla was less common in 25.15% of all cases and 28.26% of multiple myeloma cases. ONJ manifestation in both the mandible and maxilla was present in 6.54% of all patients and 6.52% of multiple myeloma patients.

ONJ occurred in 62.52% of all patient or 66.67% of multiple myeloma patients after a dental intervention (such dental extraction or dentoalveolar surgery).

It would be not be appropriate to draw conclusions about ONJ risk in various bisphosphonates based on case reports without control groups and frequencies (Table 23).

Table 20. Commonest side-effects reported in the multiple myeloma studies.			
Reference	Route, dose & frequency	Commonest toxicity	Comments
Clodronate			
Heim 1995	1600 mg/d po	Leukopenia, nausea, loss of appetite, vomiting, dyspnoea.	
Lahtinen 1992	400 mg capsules po tid	Nausea, diarrhoea, constipation, abdominal pain, alergic reations. No difference between clodronate and placebo group.	
McCloskey 2001	400 mg capsules po tid	Not reported.	
Pamidronate			
Attal 2006	Pamidronate 90 mg IV, every 4 weeks, and 400 mg pamidronate plus thalidomide po. a minimum dose reduction of 50 mg	Peripheral neuropathy, fatigue, constipation, neutropenia, thrombocytopenia, anemia, infection, osteonecrosis, nausea.	
Berenson 1998	90 mg po. every 4 weeks	Anemia, fever, nausea, upper respiratory tract infection, fatigue, constipation, diarrhea, coughing.	Two withdrawals due toxicities: an apparently allergic reaction and hypocalcemia (7.5 mg per deciliter)
Brincker 1998	75 mg capsules po. bid	Nausea, dysphagia/ dyspepsia and gastrointestinal ulcerations(56 vs. 43)	Only gastrointestinal events were reported.
Kraj 2000	60 mg IV, every 4 weeks	Not reported.	
Etidronate			
Belch 1991	5 mg/kg/d	Not reported.	
Ibandronate			
Menssen 2002	2 mg IV every month	Not spezified.	40 placebo and 42 ibandronate patients dropped out because of severe adverse advents.

Table 21. Included osteonecrosis of the jaw (ONJ) observational studies.

Reference	Study type	Type of bisphosphonates	Total No of pts	No. of ONJ pts	Route, dose & frequency	Time of treatment	ONJ inci-dence
Badros 2006	Retrospective study	Pamidronate	17	3	NR	NR.	17.65%
		Zoledronate	34	2			5.88%
		Pamidronate + zoledronate	33	17			51.51%
Calvo-Villas 2006	NC	Zoledronate	64	7	NR	NC	7(10.9%)
Corso 2007	Retrospective study	Pamidronate	20	0	NC	23 mo.	0%
		Zoledronate	37	5	NC	28 mo.	11.9%
		Pamidronate + zoledronate	42	2	NC	47 mo.	4.55%
Dimopoulos 2006	Retrospective study from 1997; Prospective from 2003 to 2005	Pamidronate	93	7	NR	39 mo ONJ pts (11-76) vs 28 without ONJ (4.5-123)	7.5%
		Zoledronate	33	1			3%
		Pamidronate+zoledronate	66	6			9.1%
		Ibandronate	1	0			0%
		Ibandronate +zoledronate	4	1			25%
		Clodronate+ zoledronate	1	0			0%
		Residronate+ zoledronate	1	0			0%
Garcia-Gara 2006	Retrospective study	Pamidronate	49	1	90 mg monthly	28 mo.	2%
		Zoledronate	64	6	4 mg monthly	12 mo (7-28)	9.3%
		Pamidronate+zoledronate	30	7		43.5 mo (24-59)	23.3%
Tosi 2006 b	Retrospective study	Zoledronate	225	6	NR	10 mo (4-35)	2.7%
Zervas 2006	Retrospective study from 1991, prospec-tive from 2001 to 2006	Pamidronate	78	1	90 mg IV	24 mo (4-120)	1.28%
		Zoledronate	91	6	4 mg IV 4-6 wks		6.59%
		Pamidronate+zoledronate	85	21			24.71%

Table 22. ONJ case reports: data stratified by patient sex, ONJ site and previous surgical / dental intervention

Reference	Total No. of MM pts	No. male	No female	Mandible	Maxilla	Both	Previous surgical/dental intervention
Abu-Id 2006	73*	24	49	57	12	4	38
Agrillo 2006	30*	10	20	18	7	5	20
Bagan 2006	9	4	5	7	0	2	5
Battley 2006	1	1	0	0	1	0	0
Braun 2006	1	1	0	1	0	0	1
Broglia 2006	1	1	0	0	1	0	NR
Capalbo 2006	9	3	6	6	3	0	9
Carneiro 2006	1	1	0	1	0	0	1
Carter 2005	2	2	0	0	0	1	2
Clarke 2007	21	14	7	16	5	0	9
Curi 2007	1	0	1	1	0	0	1
Dannemann 2007	7	NE	NE	NE	NE	NE	7
Diego 2007	3	3	0	1	1	1	3
Dimitrakopoulos 2006	5	4	1	3	2	0	5
Elad 2006	22	12	10	13	9	0	19
Estilo 2004	13	4	9	6	5	2	9
Ficarra 2005	2	0	2	1	0	1	2
Gibbs 2005	8	5	3	NR	NR	NR	7
Hansen 2006	5	3	2	3	1	1	3
Hay 2006	2	1	1	2	0	0	1
Herbozo 2007	1	0	1	0	1	0	0
Kademani 2006	1	1	0	1	0	0	1
Katz 2005	2	1	1	0	2	0	2
Khamaisi 2006	6	3	3	1	5	0	NR
Kumar 2007	2	2	0	1	1	0	2
Lenz 2005	1	0	1	0	0	1	NR
Lugassy 2004	3	2	1	3	0	0	1
Magopoulos2007	33	NR	NR	20	12	1	NR
Marunick 2005	2	1	1	2	0	0	NE
Marx 2005	119*	NR	NR	81	33	5	55
Mavrokokki 2007	114*	63	51	57	24	8	89
Melo 2005	7	6	1	5	1	1	NE
Merigo 2006	1	0	1	1	0	0	1
Migliorati 2005	3	2	1	3	0	0	2
Montazeri 2007	1	1	0	NR	NR	NR	1
Mortensen 2007	4	3	1	1	3	0	4
Murad 2007	2	2	0	1	1	0	2
Pires 2005	4	2	2	3	1	0	NE
Pozzi 2007	35	11	24	27	6	2	16
Purcell, Boyd2005	3	3	0	2	0	0	2
Ruggiero 2004	28	17	11	19	8	1	NE
Pastor-Zuazaga 2006	1	NE	NE	0	0	0	1
Phal 2007	3	1	2	0	2	0	NR
Polizzotto 2006	1	1	0	NR	NR	NR	1
Salesi 2006	2	1	1	NR	NR	NR	NR
Senel 2007	1	0	1	1	0	0	1
Sitters 2005	1	1	0	1	0	0	NR
Treister 2006	1	1	0	1	0	0	0
Vannucchi 2005	1	1	0	NR	NR	NR	NE
Walter 2007	9	4	5	6	0	3	1
Wutzl 2006	12	8	4	7	5	0	NR
Yeo 2005	2	1	1	0	1	1	2
Zarychanski 2006	10	6	4	9	1	0	6
Total	632	238 (472*) 50.42%	234 (472*) 49.58%	389 (612*) 63.56%	154 (612*) 25.16%	40 (612*) 6.54%	332 (531*) 62.52%
MM extractable pts	295 46.68%	141 (255**) 55.29%	114 (255**) 44.71%	176 (276**) 63.77%	78 (276**) 28.26%	18 (276**) 6.52%	130 (195**) 66.67%

NE not extractable NR not reported * all pts ** MM pts

Table 23. ONJ case reports: data stratified by bisphosphonate type.

	Total No. of MM pts	Clodronate	Pamidro-nate	Zoledronate	Pamidronate/ Zoledronate	iV not speci-fied	Others
Abu-Id 2006	73*				68		
Agrillo 2006	30*					30	
Bagan 2006	9		2	7			
Battley 2006	1			1			
Braun 2006	1			1			
Broglia 2006	1				1		
Capalbo 2006	9		2	4	3		
Carneiro 2006	1			1			
Carter 2005	1		2				
Clarke 2007	21		12	1	8		
Curi 2007	1			1			
Dannemann 2007	7			2	5		
Diego 2007	3			3			
Dimitrakopoulos 2006	5			2	3		
Elad 2006	22		17	4			1(A)
Estilo 2004	13*					13	
Ficarra 2005	2			1	1		
Gibbs 2005	8*		1	7			
Hansen 2006	5			1	4		
Hay 2006	2			2			
Herbozo 2007	1			1			
Kademani 2006	1			1			
Katz 2005	2			1	1		
Khamaisi 2006	6		6				
Kumar 2007	2		2				
Lenz 2005	1			1			
Lugassy 2004	3		1		2		
Magopoulos 2007	33		6	19	7		1(P,I,Z)
Marunick 2005	2		1	1			
Marx 2005	119*		3 2	48	36		3 (A)
Mavrokokki 2007	114*	2	20	43	13		30 (A) 2(R) 2(A/R) 1(P/A) 1(P/I)
Melo 2005	7		4	2	1		
Merigo 2006	1			1			
Migliorati 2005	3		1		2		
Montazeri 2007	1	1					
Mortensen 2007	4		2	2			
Murad 2007	2			2			
Pires 2005	4				4		
Pozzi 2007	35		3	14	18		
Purcell and Boyd 2005	3		2	1			
Ruggiero 2004	28		14	4	10		
Pastor-Zuazaga 2006	1				1		
Phal 2007	3			1	1		1(P/C)
Polizzotto 2006	1		1				
Salesi 2006	2			2			
Senel 2007	1	1					
Sitters 2005	1			1			
Treister 2006	1		1				
Vannucchi 2005	1			1			
Walter 2007	9		1	1	7		
Wutzl 2006	12		2	8	2		
Yeo 2005	2		2				
Zarychanski 2006	10		10				
Total	632	4	147	193	198	43	42
MM extractable pts	295	2	95	102	81	13	9
Pts (no/%)not stratified by illness	337 (53.32%)	2 (50%)	52 (35.37%)	91 (41.15%)	117 (59.09%)	30 (69.77 %)	33 (78.57%)

* MM pts not extractable A Alendronate C Clodronate I Ibandronate P Pamidronate Z Zoledronate

3.4. Discussion

The choice of therapy for a multiple myeloma patient ideally depends on evidence that the selected treatment leads to better outcomes and/or a lower risk of side-effects. However, the volume of data that need to be considered by medical professionals, researchers, consumers and policy makers is constantly expanding and it has become extremely difficult for the individual to keep up to date with current knowledge in his or her field of interest. Reviews are also required to identify new research questions to address in further studies, rather than simply giving a summary of the studies so far. Provided that the quality of research has been adequate, research synthesis has always an important impact on policy and practice [Chalmers 2001, Chelimsky 1995].

3.4.1. Method

A systematic review as a method for reviewing research evidence involves the careful and systematic collection of data from clinical trials, an assessment of each study and an unbiased synthesis and measurement of findings from individual studies. This is primarily done by means of a statistical tool called meta-analysis and requires due consideration of any flaws in the evidence.

A statistical synthesis of the "appropriate" results of separate but similar studies through meta-analysis has its major advantages. Firstly, many individual studies lack statistical power, as they are too small to detect modest but important effects. Statistical power is considerably improved by combining all the studies that have attempted to answer the same question. For example, one of the earliest and more important meta-analyses ever done, published in 1982 and concerning treatments in myocardial infarction, showed that thrombolysis was associated with a highly significant fall in mortality. This finding came about after synthesising eight smaller studies, each of which separately showed no significant result [Stampfer 1982]. The meta-analysis finding, however, was taken seriously and impacted medical practice only after the publication of two large clinical trials in the late 1980`s that confirmed its results [GISSI 1986, ISIS-2 1988].

Secondly, although critics of meta-analysis argue that combining data from different trials leads to the problem of "mixing apples and oranges", and is subject to the “garbage-in, garbage-out”, the method can have distinct advantages. In fact, by putting together all available data from separate but similar studies, meta-analyses generate results which are more generalisable to a wide variety of settings and clinical trial designs than those obtained by individual trials. The findings of a particular

study, on the other hand, may be valid only for a population of patients with the same characteristics as those investigated in the trial [Egger 1997a].

Although it is always appropriate and desirable to systematically review a body of data, this must be done with great prudence, as sometimes it may be inappropriate or even misleading to statistically pool results from separate studies [Egger 2001]. As often occurs in research, even an elegant statistical treatment of data, when performed on biased “rubble”, is incapable of generating unbiased precious stones [Chalmers 2001, Chelimsky 1995].

Confirmation of this problem is shown by the fact that the findings of some meta-analyses have later been contradicted by large randomised controlled trials [Egger 1995, 1997b]. Such discrepancies have brought discredit to a technique that has been controversial since the outset [Eysenck 1978]. As well as the problem of publication bias, there are many other sources of distortion that may be introduced in the process of locating, selecting, and combining studies that may lead a misleading meta-analysis [Easterbrook 1991, Gøtzsche 1987, Egger 1997a]. Therefore there is an urgent need not only to reduce biases in the data collection and reporting in individual studies that may contribute to reviews, but also to resist the temptation to combine biased or misleading findings in a questionable statistical synthesis.

One important example of this issue is the lack of clinical trials reporting side-effects. Observational studies may represent in this case the only source of evidence allowing a practitioner to find out the odds of exposing patients to a risk factor and to compare thiso to the odds of exposure among controls. Despite the strength of observational studies, they are relatively easy, cheap and quick to obtain from previously available patient records. However, taking into account the great level of reviewer subjectivity, their weaknesses are more serious than the advantages and make them only a fair source of general confirmatory data and not a source of unequivocal evidence.

More controversial is decision by the Cochrane Group to pool results on pain reduction by bisphosphonates [Djulbegovic 2002]. Contrarily, this thesis avoided doing so, since the main research on this issue was conducted inappropriately. To be able to estimate a true effect of bisphosphonates on pain, bisphosphonates should be compared against an appropriate treatment with opiate in clinical trials and avoid using the reduction of opiate consumption as a partial index of activity.

These examples not only illustrate some critical points but also further support the argument for implementation of systematic reviews. They also point out specific issues and give some direction for further practical steps. Therefore, if, as described above, any unreliability and invalidity of measurements in primary researches are identified, it becomes essential to develop new up-to-date evidence-based clinical trial bisphosphonate guidelines becomes to determine of standards

regarding how to assess therapy effects within future clinical trials. This would lead to valid results and consequently contribute to a valid body of cumulative evidence.

3.4.2. Multiple myeloma data evaluation and interpretation

Treatment of multiple myeloma, a B-cell cancer, is usually palliative. Bone disease affects 70% of multiple myeloma patients [Badros 2006] and causes complications, such as pathologic fractures, severe bone pain, impaired mobility, spinal cord compression and hypercalcaemia, all of them leading to greater morbidity and QOL.

To assist oncologists in management decisions, the following agencies have developed guidelines for the treatment of multiple myeloma: European Society for Medical Oncology (ESMO) [Harrouseau 2008], American Society of Clinical Oncology (ASCO) [Kyle 2007], United Kingdom Myeloma Forum (UK-MF) [Smith 2006], Italian Society of Hematology (SIE), Italian Society of Experimental Hematology (SIES) and Italian Group for Bone Marrow Transplantation (GITMO) [Barosi 2004], International Myeloma Foundation (IMF) [Durie 2003], and National Comprehensive Cancer Network (NCCN) [NCCN 2008]. The guidelines are based on a review of the available evidence from clinical studies and were followed by general consensus from an expert committee.

Some recommendations by guidelines regarding bisphosphonate therapy are summarized in Table 24.

Table 24. Currently available guidelines for bisphosphonate therapy

Guideline source	Oral	IV	Patients
ESMO 2008	Generally recommended	Generally recommended	With stage III or relapsed disease
ASCO 2007	Not	Pamidronate and zoldronate	Who have on plain radiogrph (s), lytic destruction of bone or osteopenia
UK-MF 2005	Clodronate	Pamidronte and zoldronate	All patients with or without bone lesions
SIE; SIES; GITIMO 2004	Clodronate	Pamidronte and zoldronate	With bone disease or severe osteopenia
IMF 2003	Clodronate	Pamidronate and zoldronate	With bone disease
NCCN 2008	Not	Pamidronate and Zoldronate	With documented bone disease including osteopenia

Generally, the problem with expert panels concerns a lack of established baseline qualifications of what an "expert" is, lack of transparency and the risk of conflicts of interests or domineering personalities influencing recommendations.

In clinical practiceh, most guideline recommendation are based on the agreement of a panel of experts (consensus expert committees), not on a systematic review process, and are considered the established truth and the gold standard about how to treat patients. The American College of Chest Physicians states that the phrases “evidence-based”, “guideline” and “we recommend” should not be used in the context of a consensus-based statement. Findings of a consensus panel should be rather stated as "opinions" or "suggestions" [Guyatt 2006]. The Australian National Health and Medical Research Council similarly excluded expert opinions and expert committee consensus from the classification of the level of evidence, as they do not arise directly from scientific investigation [NHMRC 2000].

A comparison between meta-analyses of study data on the effectiveness of treatments for myocardial infarction with "opinions" or "suggestions" made by clinical experts in textbooks and review articles, found that often the expert opinions were not in line with the evidence. They either failed to recommend treatments that were effective or recommended routine use of ineffective or potentially harmful treatments [Antman 1992]. Only if evidence is lacking should these, "opinions" or "suggestions" be, at most, taken into consideration and reviewed until new evidence becomes available [NHMRC 2000].

In comparison to the clinical guidelines listed above, this thesis developed a systematic review of bisphosphonate effects in multiple myeloma patients by means of a meta-analysis of study data. In this thesis, two searching strategies were applied due to findings of different methodological quality which had to be analysed separately.

Firstly, a primary analysis must adress an original examination of research data as reported from RCT. Moreover, a carefully defined search strategy must be used to detect and prevent publication bias [Pirozzo/Mayer 2004a].
The second search strategy was aimed at identifying and classifying either observational studies or case reports of ONJ (3.1.2. search strategy).

To prevent biases and *post-hoc* adaptation of data, the process of article selection was preceded by a review protocol written in advance. The research question was clearly defined as the benefits and harms of bisphosphonate therapy in multiple myeloma patients. Outcomes to be looked for were also set down as skeletal related events, mortality and side-effects.

Finally, this thesis identified 13 myeloma trials containing valuable information on bisphosphonate therapies. However, a large publication bias was identified and data from three studies could not be utilised at all: the Terpos [2000] study data have not been published at all (the last Cochrane review obtained the data from the manufacturer), the data from Rosen 2004 study [Rosen 2001, 2003, 2004, Berenson

2001] have not been stratified by illness and the Delmas study [1982] and Daragon study [1993] outcomes were ambiguously described meaning extraction was not possible.

Further analyses of the individual studies were performed according to their research designs. The studies were tested for heterogeneity to determine if they were qualitatively similar enough to be combined in the meta-analysis.

3.4.2.1. Efficacy of bisphosphonates concerning measurement of outcomes

The formal ranking system of levels of evidence assists health care decision makers to judge the strength of evidence associated with the reported clinical findings.

The system of assigning levels of evidence incorporates study quality ranking according to the strength of their research design, ranging from the gold standard, i.e. RCTs, to case series and case reports. The levels of evidence should also be ranked according to the evidential strength of the measured end points.

There are a variety of end points which may be measured and reported in oncology: mortality (or survival), cause-specific mortality, QOL or indirect surrogates of these three outcomes, such as disease-free survival, progression-free survival and tumour response rates.

Since more than a decade ago the Outcomes Working Group of the American Society of Clinical Oncology (ASCO) highlighted the priority of patient-outcomes (mainly survival and QOL), giving a secondary relevance to cancer outcomes and pharmacoeconomic evaluations [ASCO 1996]. The Food and Drugs Administration (FDA) and the European Medicines Agency (EMEA) also gave great relevance to QOL as a patient-outcome [Apolone 2003, Schilsky 2002]. Despite this, an appropriate QOL assessment in clinical trials is not deemed necessary. Rather, symptoms control, toxicity or performance statuses are assessed as surrogate end points for QOL.

The myeloma trials identified in this thesis do also not report (or at least not adequately) on QOL data. Therapy efficacy represented as reduction of skeletal related events and pain reduction are used as surrogate end points for QOL instead of the actual QOL assessment.

A reason for this omission is that in clinical practice the assessment of QOL is problematic and often inaccurate. The frequent obstacles of QOL assessment are represented by patients compliance, missing data, accuracy of the assessment, and complexity [Tassinari 2003]. Furthermore, there are many generic and disease-specific QOL and health survey measures. Comparison of their usefulness is often difficult [Contopoulos-Ioannidis 2009, Garratt 2009]. Varying measurements across

clinical trials make comparisons and syntheses of finding difficult. Therefore, future efforts should be made to appropriately use and standardise QOL measures in cancer RCTs [Garratt 2009].

A Cochrane systematic review examined relief of pain secondary to bone metastases by using bisphosphonates in thirty identified RCTs [Wong/Wiffen 2002]. The review concluded that the evidence is insufficient to recommend bisphosphonates for immediate effect. This finding is in contradiction with a Cochrane review examining the role of bisphosphonates in myeloma patients [Djulbegovic 2002] that showed a clear beneficial effect of bisphosphonate on pain reduction.

The discrepancy between the results of these two Cochrane reviews lays in the decision about which studies should be included. The review by Djulbegovic at al. included eight studies, all of which measured pain reduction in different ways.In contrast, the Wong and Wiffen`s study included only studies that reported the proportion of patients with pain relief within 12 weeks of bisphosphonate treatment. Using this inclusion criterion, just one trial of multiple myeloma was identified [Berenson 1996].

This thesis, as mentioned above, in contrast to the reviews of Djulbegovic et al. [Djulbegovic 2002] and Wong and Wiffen [Wong/Wiffen 2002], did not address pain reduction by bisphosphonates. In the protocol of this research, written in advance, two most important efficacy patient outcomes were set to be looked for: mortality and skeletal related event reduction (see 1.1.2)

3.4.2.2. Efficacy of bisphosphonates concerning mortality reduction

Survival as an appropriate measure of outcome is rarely achieved in treatments of common metastatic solid tumours [Chlebowski1994]. In view of this, a single trial claiming improved survival must be viewed sceptically as it is quite likely to be a false-positive result, especially if the p-value is “borderline” [Vardy 2004].

The factors that contribute to false-positive trials include publication bias in favour of positive trials, the use of multiple significance tests in the analysis of the data (at least one may be positive by chance), and a low probability that a new treatment will be superior [Vardy 2004]. Parmar et al. [1996] have shown that if the true prevalence of clinical trials comparing therapeutic strategies with a meaningful difference in survival is 10% (an arbitrary but not unreasonable estimate), and one designs an RCT with p=0.05 and with 80% power to detect a positive result, then about one trial in every three reported as positive will actually be a false-positive. For these reasons, any improvement in duration of survival needs to be verified in a second trial [Vardy 2004]. From ten myeloma studies reporting on mortality, there was not a single trial reporting positive results.

A relevant publication bias has been detected: from ten studies, one study did not investigate mortality [Heim 1995], three reported no significantly difference in mortality but without giving quantitative information [Attal 2006, Brincker 1998, Menssen 2002], and one study did not report on the number of deaths [Daragon 1993].
Mortality reduction was assessed by the meta-analysis of data from six studies that reported data adequately and from a study [Menssen 2002] based on the data that Cochrane Collaboration obtained from the manufacturer [Djulbegovic 2002].

The meta-analysis results show a significant difference between the treatment and the placebo group. There are, however, important reasons to believe this result is biased. The first is the publication bias mentioned above, which meant that non significant results were not adequately reported and could not be utilised in the meta-analysis. Secondly, since this finding is contradictory to practical findings and to the clinical studies (no one has found a mortality reduction benefit), it can be assumed that the number of non significant results (from studies big enough to detect a difference if present [McCloskey 2001, Berenson 1998]), when combined in a big "meta-study" led to significance because of the higher statistical power. Furthermore, the result of this meta-analysis is contradictory to the most recent clinical study by Attal et al. [2006] involving 597 myeloma patients. This study reported a non significant mortality reduction between the treatment and observation groups, without providing extractable data for meta-analysis purposes.

3.4.2.3. Efficacy of bisphosphonates concerning skeletal related event reduction

Skeletal related event (SRE) reduction can be used as an intermediate study end point to represent a preliminary index of the final outcome, such as mortality or QOL. This means that this surrogate end point *presupposes* that skeletal related morbidity is reduced through the reduction of skeletal related events, but without a direct measurement. At the same time, this end point is a surrogate of QOL instead of a real QOL measurement, as addressed above. The risk of confounding a surrogate end point with a final end point is a real danger for health care decision makers [Tassinari 2003]. An understanding of the difference between surrogate and final end points is therefore of great importance for the process of health care decision making.

The effect on reduction of skeletal related events obtained from the meta-analysis of six trials with 1673 patients was not significantly different from the effects in the placebo group. This result contradicts those obtained from Cochrane Collaboration Group [Djulbegovic 2002]. A possible explanation could be that the new results from the Attal [2006] study, having a large number of patients (396 included in

the meta-analysis), influenced the outcome in the favour of placebo. A further explanation is that different meta-analysis results were obtained by a different choice of outcomes to be analysed and therefore by the inclusion of different studies reporting on the outcomes. In other words, Djulbegovic et al. [2002] conducted a meta-analysis of the data of skeletal related events distinguishing in vertebral and non vertebral fractures, with the result that only vertebral fractures were shown to be significantly reduced by bisphosphonates. In contrast, this thesis analysed the total number of SREs, including two additional studies [Heim 1995 and Brincker 1998] ignored by the Cochrane review meta-analysis, because they do not distinguish between vertebral and non vertebral fractures. Additionally, the Cochrane review included the Terpos study [2000], which does not provide published data (Cochrane obtained the data directly from the manufacturer), the Delmas study [1982], whose published results were not adequately described and the McCloskey study [2001], with data obtained from event-free curves by the Cochrane group.

3.4.2.4. Sensitivity analysis

The meta-analysis of mortality reduction showed a significant heterogeneity between analysed trials. Through the examination of the funnel plot (Figure 4), the study by Belch [1991] was identified as an outlier. This was also expected, being the only study with mortality results in favor of placebo (after this study etidronate, which the study tested, ceased to be recommended for the use in myeloma patients). A sensitivity analysis conducted by excluding this study showed no heterogeneity between trials, meaning the meta-analysis can be conducted.

In this thesis it was assumed that bisphosphonate effects on mortality are generalisable with data from studies using bisphosphonates. The data for one of these, ibadronate, were obtained from the last Cochrane Collaboration Review [Djulbegovic 2002], as the single publication about this bisphosphonate [Menssen 2002] did not numerically reported on mortality. However, the sensitivity analysis obtained by excluding the Menssen study [2002] did not show significantly different results (Figure 5).

3.4.2.5. Harms of bisphosphonate therapy

A therapy evaluation is very often bound to be misleading, as safety data are assessed less rigorously in comparison to efficacy data [Lassere 2005] and clinical trials evaluate efficacy and toxicity asymmetrically.

Bisphosphonate RCTs do report almost only gastro-intestinal (GI) side-effects. Four identified multiple myeloma RCTs reported gastrointestinal side-effects as commonest toxicity and indicate that bisphosphonate therapy is well tolerated.

Getting evidence of bisphosphonate side-effects requires looking for case reports or case series in either descriptive studies or observational studies without a control group. Since no comparison is made to any controls, contributory cause cannot be provided. Descriptive studies are subjective and are at risk of being abused or misinterpreted. At best, they can suggest future directions for research on the treatment or test being reported on, using more stringent study design. A good general rule for this kind of evidence is “take them seriously and then ignore them” [Pirozzo/Mayer 2004].

Moreover, their credibility depends on the number of cases reported. This means, for example, that if ONJ is a rare side-effect and is very unlikely to occur on a regular basis, then descriptive studies maybe be considered sufficient evidence. On the other hand, if the figure exceeds a critical amount, then it can no longer be presumed to be a rarity and a true incidence should be investigated using more stringent study design. Should this be the case for ONJ, regulatory agencies should require bisphosphonate manufacturers to perform additional investigations about its incidence. This thesis sought to conduct a comprehensive research of all descriptive studies and also aimed to reach a critical number that could be potentially considered a persuasive argument for authorities to require more serious investigation into ONJ.

Only a randomised study by Attal et al. [2006] reported ONJ. Seven observational trials with a total of 1068 examined patients regarding ONJ were analysed. The finding resulted in very heterogeneous ONJ frequencies relating to various bisphosphonates (range: 0 to 51.51%).

The ONJ frequencies reported through observational studies included in the analysis were too heterogeneous to enable drawing a conclusion regarding frequency. However, 1068 patients from observational studies and 900 cases of ONJ reports (all published after 2003), is a large number that contradicts the apparently small number of cases presented in a retrospective chart review by the manufacturer of pamidronate and zoldronate (Novartis). This result was shown in 2005 at the public hearing of the FDA`s Oncology Drugs Advisory Committee (ODAC), which aimed to re-evaluate the risk/benefit ratio of intravenous bisphosphonates. This chart showed that, of 2500 patients who were treated with intravenous bisphosphonates at MD Anderson clinical centre over the last ten years, only 11 ONJ cases occurred in 631 breast cancer patients and 6 ONJ cases occurred in 148 multiple myeloma patients.

In this work all available evidence on ONJ was gathered and stratified into observational studies and case reports, according to their likelihood of being biased. The results showed that the number of cases is much bigger than that considered in the FDA risk-benefit re-evaluation. Therefore, it would be desirable for the national

authorities to demand an investigation from the manufacturers into the true incidence of side-effects in clinical studies, which are less likely to produce biased results than retrospective data collections.

From a general point of view, higher safety requirements should be in the agenda of future trials. Investigating and publishing side-effects adequately should be set as a standard. This would result not only in better systematic reviews but also in an improvement of decision making by policy makers and physicians, let alone benefiting patient care.

In summary, this thesis conducted a critical analysis on ONJ by means of a thorough review of all descriptive studies, aiming to call attention on some concerns regarding the actual standards for clinical studies. Other bisphosphonate side-effects were not investigated, since this would go beyond the scope of this work, but a general requirement of higher examination standards for other relevant side-effects should also be considered legitimate.

3.4.2.6. Clinical significance versus statistical significance

Clinical significance can be taken into consideration only after assessing the statistical significance of primary and secondary trial end points.

Generally, a widely spread belief is that a statistically significant result automatically means the same as *clinically significant* or true result. This is improper.

If the starting question is: “what is the probability (likelihood) that the difference we found between groups was obtained purely by chance?”, then the procedure used to give an answer would be a null hypothesis (H_o) testing. The H_0 assumes that there is no difference between treatment and placebo group and that the observation occurred due to chance alone. The customary scientific approach is to accept or to reject the null hypothesis within a defined margin of acceptable error. Rejecting the null hypothesis is a confirmation of the opposite statement “There is a real difference between the groups" based on a testing of the probability that the null hypothesis was falsely rejected. It is generally accepted that the probability that the null hypothesis is rejected when in fact it is true should be less than 5% ($p< 0.05$). This only means that the difference found between the groups is significant because it is unlikely for them to have occurred by chance alone. This does not automatically mean, that the difference is significant because of a clinically relevant effect of a therapy or drug. However, the higher the difference, the higher is the likelihood that this could be the case. On the other side, if the difference in effect (the effect size) is small, significance can still be reached providing higher statistical power by enlarging the sample size. By doing this, of course, the risk of intervening external factors determining the

effect found become greater. In this case the inference from statistically significant to clinically significant gets more problematic.

For example, consider a placebo with which 50% experience a clinical benefit and 50% do not. In testing a drug against this placebo with groups of 10 patients, the result will be significant ($P < 0.05$) if at least 80% of patients show a benefit. With 1000 patients, a 58% benefit ratio leads already to significance, with 10000, 50,8%.

In big studies small differences between groups are enough for reaching significance, whereas the same difference ratio in small studies could appear as due to chance alone. If the effect size we are looking for is not previously set, as in most cases, then a sample size could be arbitrarily enlarged until we get significant results also for marginal effects. This is a controversial point of discussion on clinical trials, but also a potential argument against the use of meta-analysis.

In the case of the meta-analysis on mortality reduction discussed here, a number of studies with non-significant results combined together in a broad "meta-study" showing the contrary. This does not mean that the meta-analysis discovered something that the individual studies did not see, but simply that effect difference previously seen as insignificant became significant thanks to the bigger sample size.

In our particular case, moreover, the significant result in the meta-analysis is contradicted by a recent big randomised trial by Attal et al. [2006].

In order to evaluate a therapy and its significance, the results from subgroup analyses should be taken into account. However their results could also be misleading.

If in the example above you have a not significant result of 55% of 100 patients with a clinical benefit (not significant being larger than 50%) and than you make a subgroup analysis with ten groups with ten patients each, then it is possible that in some of the groups a significant result is reached. Let's say that in a specific group nine out of ten patients experienced a beneficial effect of the treatment, which appears as highly significant (in a case of a chance concentration in a group such as in this example, then we should expect also that other ten-patiens subgroups show non-significant results or even a lower benefit ratio lower than the placebo). This result is obviously biased. Since the patients were not randomly assigned but those with the best outcomes concentrated in one of the subgroups, the probability that the beneficial effect occurred due to chance is ten times bigger:

$1-0{,}9893^{10} = 0{,}1020 = 10{,}2\%$

This means that the result occurred probably by chance and is not significant. It would be a severe case of "selection bias" if we were only to concentrate and communicate the results of the subgroup with the 90% beneficial effects, ignoring or keeping in the background how we got to that result.

This point should have been taken account when the meta-analysis of subgroups data on vertebral and non vertebral fractures were conducted and interpreted. This was also the reason behind the choice in this thesis to conduct a meta-analysis based on SRE *total* data in contrast to the last Cochrane Review [Djulbegovic 2002].

Although it is clear that a clinically significant result also has to be statistically significant, the statistical significance should not be taken at face value as a definitive proof of clinical efficacy without a critical appraisal.

3.4.3. Limitations

This thesis has taken a step in the direction of identifying the benefits and risks of supportive bisphosphonate therapy in multiple myeloma patients.
As mentioned above, publication bias was relevant to this issue, meaning that from 13 identified myeloma studies only a small number of information could be utilised. Moreover, the fact that the studies not published in English, German, Italian language journals were not reviewed may also have introduced bias into the research.

Further bias may have been introduced by taking into account only published literature. The little availability of published data also imposes a clear limit.
Generally, since “positive” studies are more likely to be published than “negative” studies, then any review, including this one, must be biased towards a “positive” result. This is an implicit problem concerning the whole domain of health care publications.

As it makes intuitive sense to take into account information on the quality of clinical trials when doing systematic reviews [Hayward 1995, Girling 2003, Altman 2001], the approach to selecting studies for this systematic review of multiple myeloma trials was to exclude trials that fail to meet some standard of quality of study design. In order to avoid the risk of excluding studies that might contribute valid information, only trials with gross deficiencies in design were excluded, as for example the case of the non randomised Merlini [1990] trial or the Musto [2003] study in asymptomatic patients.

In addition, it is important to emphasize that even large, well-designed randomised trials have limitations. Patients in clinical trials are often very much selected, with a focus on those with good performance status and near normal blood parameters and as such are frequently not representative of the general cancer patient population. Therefore, benefits seen in patients recruited to clinical trials are not necessarily generalisable to a less carefully selected sample of patients, even if they present the same tumour type and stage [Vardy 2004].

Another approach aimed at dealing with studies of different quality would be to directly incorporate information on study quality as weighting factors in the analy-

sis. Study weights can be multiplied by quality scores, thus increasing the weight of trials deemed to be of high quality and decreasing the weight of those of low quality [Jüni 2001]. As the poor quality of studies should not modify the precision of estimate, poor quality study inclusion would be advantageous because that information could be utilised as well.

However, weighting by quality scores is problematic for several reasons. As the choice of the scale influences the weight of individual studies in the analysis, the combined effect estimate and its confidence interval, reviewers' subjectivity would play a role. Additionally, bias associated with poor methodology is only reduced, not removed: dubious data continues to stay dubious also after giving them less importance. Including both good and poor studies may moreover increase heterogeneity of estimated effects across trials and may reduce the credibility of meta-analysis conduction.

3.5. Conclusion

Multiple studies (reviewed in Hillner et al. [2000]) have shown that patient outcomes depend on how frequently a practitioner or centre treats a particular cancer site [Hillner 2000, Glasgow 1996, Lieberman1995, Begg 1998, Hodgson 2003, Davis 1987, Feuer 1994] and several studies have suggested that patients treated in clinical trials have a better outcome than patients who receive similar treatment but who are not in a clinical trial [Davis 1985, Karjalainen 1989, Mayers 2001]. Therefore, in order to improve patient outcomes, oncologists should further improve how the therapy is delivered and also reflect on how patients are recruited for clinical trials. Additionally to the physician's efforts, the decision to treat a patient should be of course based on the premise that the treatment will do more good than harm, based on the best evidence available. Further improvements can only be made when a true risk/benefit assessment is possible. Since an overly positive risk/benefit assessment can result in inappropriate health care decisions (e.g. new and expensive drugs may be used instead of older, cheaper and thoroughly investigated products or inefficient or unsafe experimental therapies may be used by other investigators, unaware of the outcome of previous trials) the failure of such unbiased evidence to reach the medical community is a serious hazard [Bardy 1998].

The current evidence on bisphosphonate therapy shows there is no practical survival advantage. Future studies should therefore investigate bisphosphonate treatments as a palliative treatment by measuring its influence on true QOL end points. The QOL assessment using solely surrogate end points like SRE reduction is inappropriate from a methodological point of view and can be misleading since such QOL outcomes are just surrogate responses [Tassinari 2003].

Furthermore, there is no sufficient published evidence for the risk/benefit assessment of bisphosphonate treatment. Some of reasons for this is the lack of promotion and legal requirements requiring all significant clinical study outcomes to be accurately measured, reported and published, with exceptions. New published study data that compares different bisphosphonates are essential in order to offer to patients the best possible treatment. Efforts should be made to establish multinational databases that include all existing data on all clinical trials, including those notified, ongoing, suspended and completed [Chalmers 1992, Delamothe1996, Bardy 1998]. Such databases should be accessible to regulatory authorities and preferably to the entire medical community [Bardy 1998]. They could be helpful in overcoming biases of medical information based on intuitive interpretation, one-sided interests or authority-based opinions.

APPENDIX. QUALITY EVALUATION OF THE INCLUDED MULTIPLE MYELOMA TRIALS

The quality evaluation was performed using of the check-lists obtained from the Guideline of the Australian NSW Health Department [Liddle 1996].

Table A1.1. Descriptive information about the Attal 2006 study.		
Study Identification	Include author, title, reference and year of publication (if available) and the study timeframe.	Attal 2006
How is the study type described?	Randomized Controlled Trials (RCT), Non-Randomized Control Trials (N-RCS), Cohorts, Before and After Studies (BAS) with/without controls, Case Control Studies (C-CS) - define whether population or hospital based case control study.	RCT
What interventions are considered and how are they implemented?		Intensive therapy aimed at maintaining the duration of response and prolongation of survival after high dose therapy.
What outcomes are considered?	ie benefits and harms.	*Benefits:* Event free survival, overall survival and survival without skeletal related event. *Harms:* Peripheral neuropathy, fatigue, constipation, neutropenia, cardiac, thrombosis, thrombocytopenia, anemia, infection, mood change, renal, osteonecrosis of the jaw, nausea.
What factors other than the intervention could affect the outcome?	Include potential confounding factors, differences in baseline characteristics between intervention and control groups.	Deletion of chromosome 13, response rate at time of randomization.
What are the characteristics of the population and study setting?	Population characteristics eg age, sex, disease characteristics of the population, disease prevalence. Study Setting eg rural, urban, hospital inpatient or outpatient, general practice, community.	Patients less than 65 years without or with only one adverse prognostic factor (beta-2 microglobulin > 3 mg/l and deletion of chromosome 13 by FISH analysis) were enrolled. The criteria for exclusion were prior treatment for myeloma, another malignancy, abnormal cardiac function (systolic ejection fraction < 50%), chronic respiratory disease (vital capacity or carbon monoxide diffusion < 50% of normal), abnormal liver function (serum bilirubin > 35 µmol per liter or ALAT, ASAT > four times normal), psychiatric disease.
How many groups/sites in the study?		Multicenter (74 centers) study.

Table A1.2. The Evaluation Criteria for Attal 2006.							
EVALUATION CRITERIA FOR THE STUDY					***COMMENTS***	***CODE OPTIONS a, b1, b2, c, ?, n/a***
What is the study type?					RCT	
RCT	N-RCS	Cohort	BAS	C-CS		
Are study participants well-defined in terms of time, place and person?						n/a
	N-RCS	Cohort	BAS	C-CS		
Is the method of allocation to intervention and control groups/sites independent of the decision to enter the individual or group in the study (adequate allocation concealment)?						a
RCT						
What percentage (%) of individuals or clusters refused to participate?						n/a
	N-RCS	Cohort	BAS	C-CS		
Are individuals within groups/clusters blind to which intervention group they belong AND are those delivering the intervention (health professional's careers) blind to the intervention group?					No.	c
RCT	N-RCS					
Is exposure to interventions measured in a standard, valid and reliable way (avoidance of recall bias)?						n/a
				C-CS		
Is exposure to interventions measured in the same way for both case and control groups? (NB: Objective measures would meet these criteria).						n/a
				C-CS		
Are outcomes measured in a standard, valid and reliable way?						a
RCT	N-RCS	Cohort	BAS	C-CS		

Table A1.3. The Study Evaluation Criteria for Attal 2006.							
EVALUATION CRITERIA FOR THE STUDY						**COMMENTS**	**CODE OPTIONS a, b1, b2, c, ?, n/a**
Are outcomes measured in the same way for both intervention and control groups? (NB: Blinding or objective measures would meet these criteria).						Yes.	a
RCT	N-RCS	Cohort	BAS	C-CS			
Are factors other than the intervention eg confounding factors, comparable between intervention and control groups and if not comparable, are they adjusted for in the analysis?						Yes.	a
RCT	N-RCS	Cohort	BAS	C-CS			
What percentage (%) of individuals or clusters recruited into the study is not included in the analysis? (Loss to follow up).							
RCT	N-RCS	Cohort	BAS	C-CS			
Is the analysis by intention to intervene (treat)?						Criterion not described adequately to classify as a, b1, b2 or c.	?
RCT	N-RCS		BAS				
Are results homogeneous between sites? (Multicenter/multisite studies only).						Yes.	a
RCT	N_RCS	Cohort	BAS	C-CS			
OVERALL ASSESSMENT OF THE STUDY						**COMMENTS**	**CODE OPTIONS A, B1, B2, C**
How well (code A, B1, B2, C - see Table 3) was the study done to minimize bias? If coded as B1, B2 or C, what is the likely direction in which bias might affect the study results?						Some evaluation criteria from the checklist are fulfilled. Where evaluation criteria are not fulfilled or are not adequately described, the conclusions of the study or review are thought unlikely to alter.	B1
Is the overall effect of the study due to the study intervention?						Yes	
If the study type is not an RCT, explain if there is any practical or ethical reason why an RCT cannot be done.						Not applicable.	
Include other comments concerning areas for further research, applicability of evidence to target population, importance of study to policy development.						A skeletal event was defined as a bone lesion requiring a specific therapy (chemotherapy, irradiation or surgery).	

Table A2.1. Descriptive information about the Belch 1991 study.		
Study identification	Include author, title, reference and year of publication (if available) and the study time-frame.	Belch 1991
How is the study type described?	Randomized Controlled Trials (RCT), Non-Randomized Control Trials (N-RCS), Cohorts, Before and After Studies (BAS) with/without controls, Case Control Studies (C-CS) - define whether population or hospital based case control study.	RCT
What interventions are considered and how are they implemented?		Etidronate capsules (20 mg/kg x 28 days, then 5 mg/kg) until death or discontinuation; placebo: identical
What outcomes are considered?	ie benefits and harms.	patient height vertebral index; pathological fractures overall survival; pain hypercalcemia
What factors other than the intervention could affect the outcome?	Include potential confounding factors, differences in baseline characteristics between intervention and control groups.	Baseline characteristics included age, sex, performance status, bone lesions, hypercalcemia.
What are the characteristics of the population and study setting?	Population characteristics eg age, sex, disease characteristics of the population, disease prevalence. Study Setting eg rural, urban, hospital inpatient or outpatient, general practice, community.	166 eligible multiple myeloma patients with a majority of males, over 60 years in age. The median follow-up time was 3.7 years with a minimum of 1.5 years..
How many groups/sites in the study?		Multicenter.

Table A2.2. The Study Evaluation Criteria for Belch 1991 study.							
EVALUATION CRITERIA FOR THE STUDY						**COMMENTS**	**CODE OPTIONS a, b1, b2, c, ?, n/a**
What is the study type?						RCT	
RCT	N-RCS	Cohort	BAS	C-CS			
Are study participants well-defined in terms of time, place and person?							n/a
	N-RCS	Cohort	BAS	C-CS			
Is the method of allocation to intervention and control groups/sites independent of the decision to enter the individual or group in the study (adequate allocation concealment)?						Yes.	a
RCT							
What percentage (%) of individuals or clusters refused to participate?							n/a
	N-RCS	Cohort	BAS	C-CS			
Are individuals within groups/clusters blind to which intervention group they belong AND are those delivering the intervention (health professionals carers) blind to the intervention group?							a
RCT	N-RCS						
Is exposure to interventions measured in a standard, valid and reliable way (avoidance of recall bias)?							n/a
				C-CS			
Is exposure to interventions measured in the same way for both case and control groups? (NB: Objective measures would meet this criteria).							n/a
				C-CS			
Are outcomes measured in a standard, valid and reliable way?							a
RCT	N-RCS	Cohort	BAS	C-CS			

Table A2.3. The Study Evaluation Criteria Belch 1991 study.							
EVALUATION CRITERIA FOR THE STUDY						**COMMENTS**	**CODE OPTIONS a, b1, b2, c, ?, n/a**
Are outcomes measured in the same way for both intervention and control groups? (NB: Blinding or objective measures would meet these criteria).						Yes.	a
RCT	N-RCS	Cohort	BAS	C-CS			
Are factors other than the intervention eg confounding factors, comparable between intervention and control groups and if not comparable, are they adjusted for in the analysis?						Yes.	a
RCT	N-RCS	Cohort	BAS	C-CS			
What percentage (%) of individuals or clusters recruited into the study is not included in the analysis? (Loss to follow up).						6.12% (6) of etidronate group (98 patients) and 5.13% (4) of placebo group (74 patients)	a
RCT	N-RCS	Cohort	BAS	C-CS			
Is the analysis by intention to intervene (treat)?						Yes.	a
RCT	N-RCS		BAS				
Are results homogeneous between sites? (Multicenter/multisite studies only).						Yes.	a
RCT	N_RCS	Cohort	BAS	C-CS			
OVERALL ASSESSMENT OF THE STUDY						**COMMENTS**	**CODE OPTIONS A, B1, B2, C**
How well (code A, B1, B2, C - see Table 3) was the study done to minimize bias? If coded as B1, B2 or C, what is the likely direction in which bias might affect the study results?						All or most evaluation criteria from the checklist are fulfilled. Where evaluation criteria are not fulfilled, the conclusions of the study or review are thought very unlikely to alter.	A
Is the overall effect of the study due to the study intervention?						Yes.	
If the study type is not an RCT, explain if there is any practical or ethical reason why an RCT cannot be done.						Not applicable.	
Include other comments concerning areas for further research, applicability of evidence to target population, importance of study to policy development.						SRE=pathologic fractures.	

Table A3.1. Descriptive information about the Berenson 1998 study.		
Study Identification	***Include author, title, reference and year of publication (if available) and the study timeframe.***	Berenson 1996, 1998
How is the study type described?	Randomized Controlled Trials (RCT), Non Randomized Control Trials (N-RCS), Cohorts, Before and After Studies (BAS) with/without controls, Case Control Studies (C-CS) - define whether population or hospital based case control study.	RCT
What interventions are considered and how are they implemented?		Bisphosphonate therapy aimed at reduction of skeletal events in multiple myeloma patients.
What outcomes are considered?	ie benefits and harms.	SRE (total); vertebral fractures; non vertebral fractures; survival; hypercalcemia; pain; Quality of life adverse events
What factors other than the intervention could affect the outcome?	Include potential confounding factors, differences in baseline characteristics between intervention and control groups.	Assessments like physical examination, the evaluation of bone pain, scores for Eastern Oncology Group (ECOG) performance status and scores scores for quality of life.
What are the characteristics of the population and study setting?	Population characteristics eg age, sex, disease characteristics of the population, disease prevalence. Study Setting eg rural, urban, hospital inpatient or outpatient, general practice, community.	A total of 392 patients were enrolled (203 patients received pamidronate and 189 received placebo. Data of 196 patients receiving pamidronate and 181 receiving placebo were evaluated .Adult patients with Durie-Salmon stage multiple myeloma with an estimated life expectancy of at least nine months
How many groups/sites in the study?		Multicenter (88 centers in USA, Canada, Australia and New Zealand) study.

Table A3.2. The Evaluation Criteria for Berenson 1998 study.							
EVALUATION CRITERIA FOR THE STUDY						**COMMENTS**	**CODE OPTIONS a, b1, b2, c, ?, n/a**
What is the study type?						RCT	
RCT	N-RCS	Cohort	BAS	C-CS			
Are study participants well-defined in terms of time, place and person?							n/a
	N-RCS	Cohort	BAS	C-CS			
Is the method of allocation to intervention and control groups/sites independent of the decision to enter the individual or group in the study (adequate allocation concealment)?							a
RCT							
What percentage (%) of individuals or clusters refused to participate?							n/a
	N-RCS	Cohort	BAS	C-CS			
Are individuals within groups/clusters blind to which intervention group they belong AND are those delivering the intervention (health professional's careers) blind to the intervention group?						Yes	a
RCT	N-RCS						
Is exposure to interventions measured in a standard, valid and reliable way (avoidance of recall bias)?							n/a
				C-CS			
Is exposure to interventions measured in the same way for both case and control groups? (NB: Objective measures would meet these criteria).							n/a
				C-CS			
Are outcomes measured in a standard, valid and reliable way?							a
RCT	N-RCS	Cohort	BAS	C-CS			

Table A3.3. The Study Evaluation Criteria for Berenson 1998 study.							
EVALUATION CRITERIA FOR THE STUDY						**COMMENTS**	**CODE OPTIONS a, b1, b2, c, ?, n/a**
Are outcomes measured in the same way for both intervention and control groups? (NB: Blinding or objective measures would meet these criteria).						Yes.	a
RCT	N-RCS	Cohort		BAS	C-CS		
Are factors other than the intervention eg confounding factors, comparable between intervention and control groups and if not comparable, are they adjusted for in the analysis?						Yes.	a
RCT	N-RCS	Cohort		BAS	C-CS		
What percentage (%) of individuals or clusters recruited into the study is not included in the analysis? (Loss to follow up).						Loss to follow up 61% with pamidronate vs 58% with placebo.	
RCT	N-RCS	Cohort		BAS	C-CS		
Is the analysis by intention to intervene (treat)?						Criterion not described adequately to classify as a, b1, b2 or c.	?
RCT	N-RCS			BAS			
Are results homogeneous between sites? (Multicentre/multisite studies only).						Yes.	a
RCT	N_RCS	Cohort		BAS	C-CS		
OVERALL ASSESSMENT OF THE STUDY						**COMMENTS**	**CODE OPTIONS A, B1, B2, C**
How well (code A, B1, B2, C - see Table 3) was the study done to minimize bias? If coded as B1, B2 or C, what is the likely direction in which bias might affect the study results?						All or most evaluation criteria from the checklist are fulfilled. Where evaluation criteria are not fulfilled, the conclusions of the study or review are thought very unlikely to alter	A
Is the overall effect of the study due to the study intervention?						Yes	
If the study type is not an RCT, explain if there is any practical or ethical reason why an RCT cannot be done.						Not applicable.	
Include other comments concerning areas for further research, applicability of evidence to target population, importance of study to policy development.						SRE(total)=any pathologic fracture. Total number of deads reported in Berenson 1996	

Table A4.1. Descriptive information about the Brincker 1998 study.		
Study Identification	Include author, title, reference and year of publication (if available) and the study timeframe.	Abildgaard 1998; Brincker 1998
How is the study type described?	Randomized Controlled Trials (RCT), Non-Randomized Control Trials (N-RCS), Cohorts, Before and After Studies (BAS) with/without controls, Case Control Studies (C-CS) - define whether population or hospital based case control study.	RCT
What interventions are considered and how are they implemented?		Pamidronate 75 mg capsules po bid; identical placebo; duration at least 2 years Oral bisphosphonate therapy aimed at prevention of skeletal-related morbidity in newly diagnosed multiple myeloma patients.
What outcomes are considered?	ie benefits and harms.	SRE (total); pain; hypercalcemia survival significant gastrointestinal events
What factors other than the intervention could affect the outcome?	Include potential confounding factors, differences in baseline characteristics between intervention and control groups.	Non. The two groups characteristics were well balanced without any significant differences.
What are the characteristics of the population and study setting?	Population characteristics eg age, sex, disease characteristics of the population, disease prevalence. Study Setting eg rural, urban, hospital inpatient or outpatient, general practice, community.	Total enrolled: 304; Bisphos. analyzed: 152; Placebo: analyzed: 148. Median age of patients was 69
How many groups/sites in the study?		Multicenter(21 centers in Denmark and Sweden).

Table A4.2. The Study Evaluation Criteria for Brincker 1998 study.							
EVALUATION CRITERIA FOR THE STUDY						**COMMENTS**	**CODE OPTIONS a, b1, b2, c, ?, n/a**
What is the study type?						RCT	
RCT	N-RCS	Cohort	BAS	C-CS			
Are study participants well-defined in terms of time, place and person?							n/a
	N-RCS	Cohort	BAS	C-CS			
Is the method of allocation to intervention and control groups/sites independent of the decision to enter the individual or group in the study (adequate allocation concealment)?						Yes.	a
RCT							
What percentage (%) of individuals or clusters refused to participate?							n/a
	N-RCS	Cohort	BAS	C-CS			
Are individuals within groups/clusters blind to which intervention group they belong AND are those delivering the intervention (health professionals careers) blind to the intervention group?						Yes.	a
RCT	N-RCS						
Is exposure to interventions measured in a standard, valid and reliable way (avoidance of recall bias)?							n/a
				C-CS			
Is exposure to interventions measured in the same way for both case and control groups? (NB: Objective measures would meet these criteria).							n/a
				C-CS			
Are outcomes measured in a standard, valid and reliable way?							a
RCT	N-RCS	Cohort	BAS	C-CS			

Table A4.3. The Study Evaluation Criteria for Brincker 1998 study.							
EVALUATION CRITERIA FOR THE STUDY						**COMMENTS**	**CODE OPTIONS a, b1, b2, c, ?, n/a**
Are outcomes measured in the same way for both intervention and control groups? (NB: Blinding or objective measures would meet these criteria).						Yes.	a
RCT	N-RCS	Cohort	BAS	C-CS			
Are factors other than the intervention eg confounding factors, comparable between intervention and control groups and if not comparable, are they adjusted for in the analysis?						Yes.	a
RCT	N-RCS	Cohort	BAS	C-CS			
What percentage (%) of individuals or clusters recruited into the study is not included in the analysis? (Loss to follow up).						73.03% with pamidronate and 74.32% with placebo.	a
RCT	N-RCS	Cohort	BAS	C-CS			
Is the analysis by intention to intervene (treat)?						Yes.	a
RCT	N-RCS		BAS				
Are results homogeneous between sites? (Multicenter/multisite studies only).						Yes.	a
RCT	N_RCS	Cohort	BAS	C-CS			
OVERALL ASSESSMENT OF THE STUDY						**COMMENTS**	**CODE OPTIONS A, B1, B2, C**
How well (code A, B1, B2, C - see Table 3) was the study done to minimize bias? If coded as B1, B2 or C, what is the likely direction in which bias might affect the study results?						All or most evaluation criteria from the checklist are fulfilled. Where evaluation criteria are not fulfilled, the conclusions of the study or review are thought very unlikely to alter.	A
Is the overall effect of the study due to the study intervention?						Yes.	
If the study type is not an RCT, explain if there is any practical or ethical reason why an RCT cannot be done.						Not applicable.	
Include other comments concerning areas for further research, applicability of evidence to target population, importance of study to policy development.						It is a negative study and does not recommend oral pamidronate.	

Table A5.1. Descriptive information about the Heim 1995 study.		
Study Identification	Include author, title, reference and year of publication (if available) and the study timeframe.	Clemens 1993; Heim 1995
How is the study type described?	Randomised Controlled Trials (RCT), Non-Randomised Control Trials (N-RCS), Cohorts, Before and After Studies (BAS) with/without controls, Case Control Studies (C-CS) - define whether population or hospital based case control study.	RCT Not double-blind, Not placebo-controlled;
What interventions are considered and how are they implemented?		Clodronate 1600 mg/d po.; control: no treatment; duration 12 months
What outcomes are considered?	ie benefits and harms.	SRE (total); pain; calcium; adverse events
What factors other than the intervention could affect the outcome?	Include potential confounding factors, differences in baseline characteristics between intervention and control groups.	The distribution of drop outs.
What are the characteristics of the population and study setting?	Population characteristics eg age, sex, disease characteristics of the population, disease prevalence. Study Setting eg rural, urban, hospital inpatient or outpatient, general practice, community.	Total: 170; 13 withdrawn after Rx. premature termination in add. 75; Bisphos.: analyzed 39; Placebo analyzed: 32
How many groups/sites in the study?		Multicenter study.

Table A5.2. The Evaluation Criteria for Heim 1995 study.							
EVALUATION CRITERIA FOR THE STUDY						**COMMENTS**	**CODE OPTIONS a, b1, b2, c, ?, n/a**
What is the study type?						RCT	
RCT	N-RCS	Cohort	BAS	C-CS			
Are study participants well-defined in terms of time, place and person?							n/a
	N-RCS	Cohort	BAS	C-CS			
Is the method of allocation to intervention and control groups/sites independent of the decision to enter the individual or group in the study (adequate allocation concealment)?						No.	c
RCT							
What percentage (%) of individuals or clusters refused to participate?							n/a
	N-RCS	Cohort	BAS	C-CS			
Are individuals within groups/clusters blind to which intervention group they belong AND are those delivering the intervention (health professional's careers) blind to the intervention group?						No.	c
RCT	N-RCS						
Is exposure to interventions measured in a standard, valid and reliable way (avoidance of recall bias)?							n/a
				C-CS			
Is exposure to interventions measured in the same way for both case and control groups? (NB: Objective measures would meet this criteria).							n/a
				C-CS			
Are outcomes measured in a standard, valid and reliable way?						.	a
RCT	N-RCS	Cohort	BAS	C-CS			

Table A5.3. The Study Evaluation Criteria for Heim 1995 study.							
EVALUATION CRITERIA FOR THE STUDY						**COMMENTS**	**CODE OPTIONS a, b1, b2, c, ?, n/a**
Are outcomes measured in the same way for both intervention and control groups? (NB: Blinding or objective measures would meet these criteria).							c
RCT	N-RCS	Cohort		BAS	C-CS		
Are factors other than the intervention eg confounding factors, comparable between intervention and control groups and if not comparable, are they adjusted for in the analysis?							a
RCT	N-RCS	Cohort		BAS	C-CS		
What percentage (%) of individuals or clusters recruited into the study is not included in the analysis? (Loss to follow up).						51.76%	
RCT	N-RCS	Cohort		BAS	C-CS		
Is the analysis by intention to intervene (treat)?						No.	c
RCT	N-RCS			BAS			
Are results homogeneous between sites? (Multicenter/multisite studies only).						Criterion not described adequately to classify as a, b1, b2 or c	?
RCT	N_RCS	Cohort		BAS	C-CS		
OVERALL ASSESSMENT OF THE STUDY						**COMMENTS**	**CODE OPTIONS A, B1, B2, C**
How well (code A, B1, B2, C - see Table 3) was the study done to minimize bias? If coded as B1, B2 or C, what is the likely direction in which bias might affect the study results?						Some evaluation criteria from the checklist are fulfilled. Where evaluation criteria are not fulfilled or are not adequately described, the conclusions of the study or review are thought likely to alter.	B2
Is the overall effect of the study due to the study intervention?							a
If the study type is not an RCT, explain if there is any practical or ethical reason why an RCT cannot be done.						Not applicable.	
Other comments						SRE=bone progression. This study did not reported on the number of patients with new SREs.	

Table A6.1. Descriptive information about the Kraj 2000 study.		
Study Identification	Include author, title, reference and year of publication (if available) and the study timeframe.	Kraj 2000 a,b
How is the study type described?	Randomized Controlled Trials (RCT), Non-Randomized Control Trials (N-RCS), Cohorts, Before and After Studies (BAS) with/without controls, Case Control Studies (C-CS) - define whether population or hospital based case control study.	RCT Not Double-blind, Not placebo-controlled;
What interventions are considered and how are they implemented?		Pamidronate 60 mg iv, every 4 weeks; control: no treatment
What outcomes are considered?	ie benefits and harms.	SREs (total) Vertebral fractures Survival
What factors other than the intervention could affect the outcome?	Include potential confounding factors, differences in baseline characteristics between intervention and control groups.	?
What are the characteristics of the population and study setting?	Population characteristics eg age, sex, disease characteristics of the population, disease prevalence. Study Setting eg rural, urban, hospital inpatient or outpatient, general practice, community.	Bisphos. enrolled / analyzed 23; Placebo enrolled / analyzed 23
How many groups/sites in the study?		One.

Table A6.2. The Evaluation Criteria for Kraj 2000 study.							
EVALUATION CRITERIA FOR THE STUDY						**COMMENTS**	**CODE OPTIONS a, b1, b2, c, ?, n/a**
What is the study type?						RCT	
RCT	N-RCS	Cohort	BAS	C-CS			
Are study participants well-defined in terms of time, place and person?							n/a
	N-RCS	Cohort	BAS	C-CS			
Is the method of allocation to intervention and control groups/sites independent of the decision to enter the individual or group in the study (adequate allocation concealment)?						.	?
RCT							
What percentage (%) of individuals or clusters refused to participate?							n/a
	N-RCS	Cohort	BAS	C-CS			
Are individuals within groups/clusters blind to which intervention group they belong AND are those delivering the intervention (health professional's careers) blind to the intervention group?						No.	c
RCT	N-RCS						
Is exposure to interventions measured in a standard, valid and reliable way (avoidance of recall bias)?							n/a
				C-CS			
Is exposure to interventions measured in the same way for both case and control groups? (NB: Objective measures would meet this criteria).							n/a
				C-CS			
Are outcomes measured in a standard, valid and reliable way?						.	?
RCT	N-RCS	Cohort	BAS	C-CS			

Table A6.3. The Study Evaluation Criteria for Kraj 2000 study.							
EVALUATION CRITERIA FOR THE STUDY						**COMMENTS**	**CODE OPTIONS a, b1, b2, c, ?, n/a**
Are outcomes measured in the same way for both intervention and control groups? (NB: Blinding or objective measures would meet these criteria).						Yes.	a
RCT	N-RCS	Cohort		BAS	C-CS		
Are factors other than the intervention eg confounding factors, comparable between intervention and control groups and if not comparable, are they adjusted for in the analysis?						Criterion not described adequately to classify as a, b1, b2 or c	?
RCT	N-RCS	Cohort		BAS	C-CS		
What percentage (%) of individuals or clusters recruited into the study is not included in the analysis? (Loss to follow up).						0	
RCT	N-RCS	Cohort		BAS	C-CS		
Is the analysis by intention to intervene (treat)?						Criterion not described adequately to classify as a, b1, b2 or c.	?
RCT	N-RCS			BAS			
Are results homogeneous between sites? (Multicenter/multisite studies only).							n/a
RCT	N_RCS	Cohort		BAS	C-CS		
OVERALL ASSESSMENT OF THE STUDY						**COMMENTS**	**CODE OPTIONS A, B1, B2, C**
How well (code A, B1, B2, C - see Table 3) was the study done to minimize bias? If coded as B1, B2 or C, what is the likely direction in which bias might affect the study results?						Some evaluation criteria from the checklist are fulfilled. Where evaluation criteria are not fulfilled or are not adequately described, the conclusions of the study or review are thought likely to alter.	B2
Is the overall effect of the study due to the study intervention?						Yes	
If the study type is not an RCT, explain if there is any practical or ethical reason why an RCT cannot be done.						Not applicable.	
Include other comments concerning areas for further research, applicability of evidence to target population, importance of study to policy development.							

Table A7.1. Descriptive information about the Lahtinen 1992 study.		
Study Identification	Include author, title, reference and year of publication (if available) and the study timeframe.	Lahtinen 1992.
How is the study type described?	Randomized Controlled Trials (RCT), Non-Randomized Control Trials (N-RCS), Cohorts, Before and After Studies (BAS) with/without controls, Case Control Studies (C-CS) - define whether population or hospital based case control study.	RCT Double-blind, placebo-controlled;;
What interventions are considered and how are they implemented?		Clodronate 400 mg capsules po tid; identical placebo; duration 24 months.
What outcomes are considered?	ie benefits and harms.	bone lesions, vertebral fractures; non vertebral fractures; Total mortality; calcium pain side-effects
What factors other than the intervention could affect the outcome?	Include potential confounding factors, differences in baseline characteristics between intervention and control groups.	Patients treated with clodronate were younger.
What are the characteristics of the population and study setting?	Population characteristics eg age, sex, disease characteristics of the population, disease prevalence. Study Setting eg rural, urban, hospital inpatient or outpatient, general practice, community.	Bisphos. enrolled / analyzed 168; Placebo enrolled / analyzed 168
How many groups/sites in the study?		Twenty-tree hospitals in Finland.

Table A7.2. The Evaluation Criteria for Lahtinen 1992 study.							
EVALUATION CRITERIA FOR THE STUDY						**COMMENTS**	**CODE OPTIONS a, b1, b2, c, ?, n/a**
What is the study type?						RCT	
RCT	N-RCS	Cohort	BAS	C-CS			
Are study participants well-defined in terms of time, place and person?							n/a
	N-RCS	Cohort	BAS	C-CS			
Is the method of allocation to intervention and control groups/sites independent of the decision to enter the individual or group in the study (adequate allocation concealment)?						Yes.	a
RCT							
What percentage (%) of individuals or clusters refused to participate?							n/a
	N-RCS	Cohort	BAS	C-CS			
Are individuals within groups/clusters blind to which intervention group they belong AND are those delivering the intervention (health professional's careers) blind to the intervention group?						Yes.	a
RCT	N-RCS						
Is exposure to interventions measured in a standard, valid and reliable way (avoidance of recall bias)?							n/a
				C-CS			
Is exposure to interventions measured in the same way for both case and control groups? (NB: Objective measures would meet these criteria).							n/a
				C-CS			
Are outcomes measured in a standard, valid and reliable way?						Yes.	a
RCT	N-RCS	Cohort	BAS	C-CS			

Table A7.3. The Study Evaluation Criteria for Lahtinen 1992 study.							
EVALUATION CRITERIA FOR THE STUDY						**COMMENTS**	**CODE OPTIONS a, b1, b2, c, ?, n/a**
Are outcomes measured in the same way for both intervention and control groups? (NB: Blinding or objective measures would meet these criteria).						Yes.	a
RCT	N-RCS	Cohort	BAS	C-CS			
Are factors other than the intervention eg confounding factors, comparable between intervention and control groups and if not comparable, are they adjusted for in the analysis?						Criterion mostly fulfilled except the younger age of the patients in the clodronate group.	b1
RCT	N-RCS	Cohort	BAS	C-CS			
What percentage (%) of individuals or clusters recruited into the study is not included in the analysis? (Loss to follow up).						12.5 % of the clodronate group and 15.48% of the placebo group.	
RCT	N-RCS	Cohort	BAS	C-CS			
Is the analysis by intention to intervene (treat)?						Yes.	a
RCT	N-RCS		BAS				
Are results homogeneous between sites? (Multicenter/multisite studies only).						Yes.	a
RCT	N_RCS	Cohort	BAS	C-CS			
OVERALL ASSESSMENT OF THE STUDY						**COMMENTS**	**CODE OPTIONS A, B1, B2, C**
How well (code A, B1, B2, C - see Table 3) was the study done to minimize bias? If coded as B1, B2 or C, what is the likely direction in which bias might affect the study results?							A
All or most evaluation criteria from the checklist are fulfilled. Where evaluation criteria are not fulfilled, the conclusions of the study or review are thought very unlikely to alter.						Yes	
If the study type is not an RCT, explain if there is any practical or ethical reason why an RCT cannot be done.						Not applicable.	
Include other comments concerning areas for further research, applicability of evidence to target population, importance of study to policy development.						SRE= Progression of osteolytic lesions or vertebral fractures or non-vertebral fractures.	

Table A8.1. Descriptive information about the McCloskey 2001 study.		
Study Identification	Include author, title, reference and year of publication (if available) and the study timeframe.	McCloskey1998, 2001
How is the study type described?	Randomized Controlled Trials (RCT), Non-Randomized Control Trials (N-RCS), Cohorts, Before and After Studies (BAS) with/without controls, Case Control Studies (C-CS) - define whether population or hospital based case control study.	RCT Double-blind, placebo-controlled
What interventions are considered and how are they implemented?		Clodronate 400 mg capsules po qid; identical placebo; duration 24 months.
What outcomes are considered?	ie benefits and harms.	SRE (total); total fractures; vertebral fractures; non-vertebral fracture; pain; calcium***
What factors other than the intervention could affect the outcome?	Include potential confounding factors, differences in baseline characteristics between intervention and control groups.	An advantage in survival was shown in a subgroup analysis.
What are the characteristics of the population and study setting?	Population characteristics eg age, sex, disease characteristics of the population, disease prevalence. Study Setting eg rural, urban, hospital inpatient or outpatient, general practice, community.	Bisphos. enrolled / analyzed 264; Placebo enrolled / analyzed 272
How many groups/sites in the study?		Multicenter study (85 centers) in the U.K.

Table A8.2. The Evaluation Criteria for McCloskey 2001study.							
EVALUATION CRITERIA FOR THE STUDY						**COMMENTS**	**CODE OPTIONS a, b1, b2, c, ?, n/a**
What is the study type?						RCT	
RCT	N-RCS	Cohort	BAS	C-CS			
Are study participants well-defined in terms of time, place and person?							n/a
	N-RCS	Cohort	BAS	C-CS			
Is the method of allocation to intervention and control groups/sites independent of the decision to enter the individual or group in the study (adequate allocation concealment)?						Yes.	a
RCT							
What percentage (%) of individuals or clusters refused to participate?							?
	N-RCS	Cohort	BAS	C-CS			
Are individuals within groups/clusters blind to which intervention group they belong AND are those delivering the intervention (health professionals carers) blind to the intervention group?						Yes.	a
RCT	N-RCS						
Is exposure to interventions measured in a standard, valid and reliable way (avoidance of recall bias)?							n/a
				C-CS			
Is exposure to interventions measured in the same way for both caseand control groups? (NB: Objective measures would meet this criteria).							n/a
				C-CS			
Are outcomes measured in a standard, valid and reliable way?						Yes	a
RCT	N-RCS	Cohort	BAS	C-CS			

Table A8.3. The Study Evaluation Criteria for McCloskey 2001 study.							
EVALUATION CRITERIA FOR THE STUDY						**COMMENTS**	**CODE OPTIONS a, b1, b2, c, ?, n/a**
Are outcomes measured in the same way for both intervention and control groups? (NB: Blinding or objective measures would meet these criteria).						Yes.	a
RCT	N-RCS	Cohort		BAS	C-CS		
Are factors other than the intervention eg confounding factors, comparable between intervention and control groups and if not comparable, are they adjusted for in the analysis?							?
RCT	N-RCS	Cohort		BAS	C-CS		
What percentage (%) of individuals or clusters recruited into the study is not included in the analysis? (Loss to follow up).						17.80% with clodronate and 13.60% with placebo.	
RCT	N-RCS	Cohort		BAS	C-CS		
Is the analysis by intention to intervene (treat)?						Yes.	a
RCT	N-RCS			BAS			
Are results homogeneous between sites? (Multicenter/multisite studies only).						Yes.	a
RCT	N_RCS	Cohort		BAS	C-CS		
OVERALL ASSESSMENT OF THE STUDY						**COMMENTS**	**CODE OPTIONS A, B1, B2, C**
How well (code A, B1, B2, C - see Table 3) was the study done to minimize bias? If coded as B1, B2 or C, what is the likely direction in which bias might affect the study results?						All or most evaluation criteria from the checklist are fulfilled. Where evaluation criteria are not fulfilled, the conclusions of the study or review are thought very unlikely to alter.	Yes.
Is the overall effect of the study due to the study intervention?						Yes	
If the study type is not an RCT, explain if there is any practical or ethical reason why an RCT cannot be done.						Not applicable.	
Include other comments concerning areas for further research, applicability of evidence to target population, importance of study to policy development.						.	

Table A9.1. Descriptive information about the Menssen 2002 study.		
Study Identification	Include author, title, reference and year of publication (if available) and the study timeframe.	Menssen 2002
How is the study type described?	Randomized Controlled Trials (RCT), Non-Randomized Control Trials (N-RCS), Cohorts, Before and After Studies (BAS) with/without controls, Case Control Studies (C-CS) - define whether population or hospital based case control study.	RCT
What interventions are considered and how are they implemented?		Ibandronate 2 mg IV every month identical placebo, duration 24 months. The therapy aimed at prolongation of multiple myeloma patient survival and prevention of skeletal related events.
What outcomes are considered?	ie benefits and harms.	SREs total, median survival time, side-effects total.
What factors other than the intervention could affect the outcome?	Include potential confounding factors, differences in baseline characteristics between intervention and control groups.	An advantage in survival was shown in a subgroup analysis.
What are the characteristics of the population and study setting?	Population characteristics eg age, sex, disease characteristics of the population, disease prevalence. Study Setting eg rural, urban, hospital inpatient or outpatient, general practice, community.	Bisphos. enrolled 107 analyzed 99; Placebo: enrolled: 107 analyzed: 99 The patients with stag II and III multiple myeloma were enrolled.
How many groups/sites in the study?		Multicenter study.

Table A9.2. The Evaluation Criteria for Menssen 2002 study.							
EVALUATION CRITERIA FOR THE STUDY						**COMMENTS**	**CODE OPTIONS a, b1, b2, c, ?, n/a**
What is the study type?						RCT	
RCT	N-RCS	Cohort	BAS	C-CS			
Are study participants well-defined in terms of time, place and person?							n/a
	N-RCS	Cohort	BAS	C-CS			
Is the method of allocation to intervention and control groups/sites independent of the decision to enter the individual or group in the study (adequate allocation concealment)?						Yes.	a
RCT							
What percentage (%) of individuals or clusters refused to participate?							n/a
	N-RCS	Cohort	BAS	C-CS			
Are individuals within groups/clusters blind to which intervention group they belong AND are those delivering the intervention (health professional's careers) blind to the intervention group?						Yes.	a
RCT	N-RCS						
Is exposure to interventions measured in a standard, valid and reliable way (avoidance of recall bias)?							n/a
				C-CS			
Is exposure to interventions measured in the same way for both case and control groups? (NB: Objective measures would meet this criteria).							n/a
				C-CS			
Are outcomes measured in a standard, valid and reliable way?						Yes.	a
RCT	N-RCS	Cohort	BAS	C-CS			

Table A9.3. The Study Evaluation Criteria for Menssen 2002 study.							
EVALUATION CRITERIA FOR THE STUDY						**COMMENTS**	**CODE OPTIONS a, b1, b2, c, ?, n/a**
Are outcomes measured in the same way for both intervention and control groups? (NB: Blinding or objective measures would meet these criteria).						Yes.	a
RCT	N-RCS	Cohort		BAS	C-CS		
Are factors other than the intervention eg confounding factors, comparable between intervention and control groups and if not comparable, are they adjusted for in the analysis?						Yes.	a
RCT	N-RCS	Cohort		BAS	C-CS		
What percentage (%) of individuals or clusters recruited into the study is not included in the analysis? (Loss to follow up).						7.48 % of the each group.	
RCT	N-RCS	Cohort		BAS	C-CS		
Is the analysis by intention to intervene (treat)?							a
RCT	N-RCS		BAS				
Are results homogeneous between sites? (Multicenter/multisite studies only).						Yes.	a
RCT	N_RCS	Cohort		BAS	C-CS		
OVERALL ASSESSMENT OF THE STUDY						**COMMENTS**	**CODE OPTIONS A, B1, B2, C**
How well (code A, B1, B2, C - see Table 3) was the study done to minimize bias? If coded as B1, B2 or C, what is the likely direction in which bias might affect the study results?						All or most evaluation criteria from the checklist are fulfilled. Where evaluation criteria are not fulfilled, the conclusions of the study or review are thought very unlikely to alter.	A
Is the overall effect of the study due to the study intervention?						Yes.	
If the study type is not an RCT, explain if there is any practical or ethical reason why an RCT cannot be done.						Not applicable.	
Include other comments concerning areas for further research, applicability of evidence to target population, importance of study to policy development.							

ABSTRACT

This research develops in detail a systematic review of therapy evidence on bisphosphonate effects in multiple myeloma patients. The objectives of this work are threefold: to introduce and discuss the advantages and shortcomings of systematic reviews, providing insights into a still not very diffused methodology in the healthcare decision-making process, to give an up-to-date overview of multiple myeloma therapy and management, focusing on bone disease management with bisphosphonates and to develop a systematic review, by means of a meta-analysis of study data on bisphosphonate effects in multiple myeloma patients.

The meta-analysis of mortality reduction showed a significance advantage of the patients treated with bisphosphonates. There are, however, some reasons to believe that these results are biased due to the poor quality of the input data. This is due partly to publication bias, since non significant results were not adequately reported. Also, the findings are contradicted by empirical results of all identified randomised trials. This does not invalidate the analysis, but shows the need for further examination.

For the evaluation of a bisphosphonate effect on the reduction of skeletal related events (SRE), seven trials were included in the meta-analysis. The included trials were not significantly heterogeneous and the meta-analysis of their results indicates no benefit from bisphosphonates on the number of patients experiencing SREs. However, the quality of life (QOL) assessment using solely surrogate end points such SRE reduction is inappropriate by a methodological point of view and can be misleading since such QOL outcomes are simply surrogate responses.

This thesis conducted a critical analysis on ONJ evidence and clinical relevance by means of a thorough review of all descriptive studies. Finally, some concerns regarding the actual standards for clinical studies and the need for higher examination standards for long-term post-marketing safety are expressed. To overcome biases in health care decisions, efforts should be made to establish multinational databases including all relevant data covering the whole domain of existing clinical trials.

Key words: systematic review, meta-analysis, bisphosphonates, therapy management, clodronate, pamidronate, ibandronate, tiludronate, zoledronate, skeletal related events, side-effects, osteonecrosis of the jaw, mortality

ZUSAMMENFASSUNG

Diese Doktorarbeit beschäftigt sich mit der Bewertung des medizinischen Wissens in Bezug auf die Therapie von Myelompatienten mit Bisphosphonaten.

Dabei wird besonderes Augenmerk auf drei Schwerpunkte gelegt. Die Vor- und Nachteile von systematischen Übersichtsarbeiten und Meta-Analysen werden beschrieben und diskutiert, wodurch ein Einblick in eine noch nicht verbreitete methodische Grundlage von Entscheidungsprozessen im Gesundheitswesen ermöglicht wird. Die Aktualisierung des jetzigen Kennisstandes in der Myelomtherapie mit dem Schwerpunkt auf der Biphosphonattherapie von Knochenerkrankungen wurde durch diese systematische Übersichtsarbeit und die Meta-Analysen vorgenommen.

Die Bewertung der Effektivität von Biphosphonaten in Bezug auf Mortalität, die durch eine Meta-Analyse durchgeführt wurde, zeigt eine signifikante Reduzierung der Mortalität in der Patientengruppe, die mit Biphosphonaten behandelt worden sind. Es gibt jedoch einige Gründe von einer Verzerrung des Ergebnisses auszugehen, vor allem weil empirischen Ergebnissen widersprochen wird. Mit Wahrscheinlichkeit handelt es sich um eine Folge mangelhafter Berichterstattung und des Nichtveröffentlichens nicht signifikanter Ergebnisse (*publication bias*). Dennoch wird die Meta-Analyse dadurch nicht ungültig, obwohl zu hinterfragen ist, ob es sich bei diesem Ergebnis um eine therapie-rechtfertigende Größenordnung handelt.

Die Bewertung der Effektivität von Biphosphonaten in Bezug auf die Senkung von skelettalen Komplikationen basiert auf der Meta-Analyse von sieben randomisierten Studien, die untereinander keinen signifikanten Unterschied zeigen. Die anschließende Meta-Analyse zeigt keinen statistisch signifikanten Unterschied zwischen den Therapiearmen. Allerdings basiert die Messung klinischer Endpunkte, die Lebensqualitätsdaten (QOL) beschreiben, nur auf der Messung eines Surrogat-Endpunktes wie der Senkung von skelettalen Komplikationen. Das ist methodologisch nicht richtig und irreführend, da dies nur zu vorläufigen Antworten klinischer Fragen führt.

Die Evidenz und klinische Relevanz der Ostenekrose im Kiefer (ONJ) wurde durch eine umfassende Beobachtungsstudien-Analyse im Kontext einer systematischen Übersichtsarbeit untersucht. Schließlich drückt diese Doktorarbeit Bedenken über die jetzigen Anforderungen an den klinische Studien und die Notwendikeit aus, höhere Anforderungen an die Langzeit-post-marketing-Sicherheit der Arzneimittel zu stellen. Die Abschätzung des Nutzen-Risiko-Verhältnisses im Gesundheitwesen auf Basis nicht verzerrter Daten wird nur dann möglich, wenn

multinationale Datenbanken mit relevanten Daten aus allen exsistierenden klinischen Studien als Grundlage etalbliert werden.

Schlüsselwörter: Systematische Übersichtsarbeiten, Meta-Analyse, Biphosphonate, Clodronat, Pamidronate, Ibandronat, Zoledronat, skelettale Komplikationen , Nebenwirkungen, Kiefernekrose, Mortalität.

GLOSSARY

Definitions are quoted from Last J (ed), *A Dictionary of Epidemiology* (3rd edition), Oxford University Press, Oxford 1995 and Liddle J, Williamson M, Irwig L. Method for evaluating research and guideline evidence. NSW Health Department, Sydney, December 1996.

.

Adequate allocation concealment: No alternation reference to case record numbers, dates of birth, day of the week, or any other such approach. An allocation procedure that is entirely intransparent before assignment and there is no open lists of random numbers or assignments.

Analysis by intention to treat: the analysis compares study and control groups based on the original random allocation regardless of whether individuals in either group received the intervention.

Applicability: extent to which the results of a study or review can be applied to a population or patient group different to that in the original study or review.

Before-and-after study: study carried out before and after the introduction of an intervention where a group is usually the unit of observation. Where groups or individuals in a before-and-after study are allocated to an intervention or control group, then the study is classified as a non-randomized controlled study.

Benefit(s): an outcome of an intervention which is advantageous for an individual or a population.

Bias: systematic errors in the design and execution of a study which may lead to an over- or underestimation of the “true” effect of an intervention.

Blinded study: a study in which observer(s) and/or subjects are kept ignorant of the group to which the subjects are assigned or of the population from which the subjects come. When both the observer and subjects are kept ignorant, we refer to a double-blind study. The intent of keeping subjects and/or investigators blinded, ie unaware of knowledge that might introduce a bias, is to eliminate the effects of such biases.

Case: a person in the population or study group identified as having the particular disease, health disorder or condition under investigation.

Case-control study: a study that starts with the identification of persons with the disease (or other outcome variable) of interest and a suitable control group of persons without the disease. Case control studies are used to estimate relative risk. Case-control studies are useful where the study factor (disease) is rare.

Case report: detailed report on one case usually covering the course of a disease and the response to treatment.

Case series: description of several cases of a given disease (usually covering the course of a disease and the response to treatment).

Cohort study: a study in which subjects are grouped by the risk factor, and those with and without the risk are followed to see who develops the disease and who doesn't. The occurrence of the outcome of interested is compared in the two groups. The alternative terms for a cohort study ie follow-up, longitudinal and prospective study, describe an essential feature of the method.

Confidence interval: the computed interval with a given probability e.g. 95%, that the true value of a variable such as a mean, proportion or rate is contained within the interval.

Confounding factor: a variable that can cause or prevent the outcome of interest, is not an intermediate variable and is associated with the factor under investigation.

Descriptive study: a study concerned with and designed only to describe the existing distribution of variables, without regard to causal or other hypotheses. An example is a community health survey used to determine the health status of people in a community.

Effectiveness: measure of the extent to which a specific intervention, procedure, regimen or service, when deployed in the field, does what it is intended to do for a specified population.

Evaluation criteria: specific features of a study or guideline/recommendation relating to quality. Coded as a, b1, b2, c, ? or n/a (Table 5).

Experiment: a study in which the investigator intentionally alters one or more factors under controlled conditions in order to study the effects of so doing.

Guideline: systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances.

Harm(s): an outcome of an intervention which is disadvantageous for an individual or population. Also referred to as risks.

Heterogeneity: occurs when the results of a set of independently performed studies on which a meta-analysis is based, are not enough similar to make statistical pooling valid. It is usually assessed by a Chi^2 test for heterogeneity.

Homogeneous: implies there is no significant heterogeneity.

Intervention: public health/health promotion policy or program or clinical treatment regimen aimed at improving health, preventing or minimizing disease or changing some other health related characteristic or behavior.

Meta-analysis: a technique which uses special adaptations of conventional statistical methods to combine results of the selected set of studies in order to investigate, compare and interpret pattern of findings. This allows making the best use of all the information gathered in the first part of the research using a systematic reviewing process.

Multicentre RCT: randomized controlled trial performed in several different settings, e.g. in different hospitals over a broad geographic area.

Non-randomized controlled study: a study or clinical trial where the allocation to the intervention or control groups has not been randomized.

Non-systematic reviews: an explicit and systematic approach has not been used to identify evidence relating to a particular topic. An adequate literature searching procedure has not been used and dimensions of study quality of the primary studies have not been considered.

Null Hypothesis (Ho): The null hypothesis states that the findings from the study are the result of chance or random factors. Therefore the overall purpose of a typical trial is to reject the null hypothesis.

Observational study: analytic methods such as case control and cohort study designs are called observational studies because the investigator is observing without intervention other than to record, classify, count and statistically analyze results.

Outcomes: all the possible results that may stem from exposure to a causal factor or from preventive or therapeutic interventions; all identified changes in health status arising as a consequence of the handling of a health problem.

Overall assessment: an overall rating on quality of a study, guideline or recommendation using the evaluation criteria. Coded as A, B1, B2, C for study checklists (Table 6).

Quality of evidence: degree to which bias has been prevented through the design and conduct of research from which evidence is derived.

Randomization: a procedure is used so that study units have an equal chance of being allocated to an intervention or control group.

Randomized controlled trial or study (RCT): an experiment in which subjects are randomly allocated into groups, usually called "study" and "control" groups, to receive or not to receive an experimental preventive or therapeutic procedure or intervention.

Recommendation: advised course of action.

Representativeness: extent to which the population or patient group in a study or review is comparable to other populations or patient groups.

Retrospective study: any study in which the outcomes have already occurred before the study and collection of data has begun.

Case control studies are also referred to as retrospective studies.

Selection bias: error due to systematic differences in characteristics between those who are selected for study and those who are not.

Sensitivity analysis: a method to determine the robustness of an assessment by examining the extent to which results are affected by changes in methods, values or variables or assumptions.

Single centre RCT: randomized controlled trial performed in one setting eg in one hospital.

Statistical significant result: a statistical significant result means that it is highly unlikely that the difference found between groups could have occurred by chance alone. In a clinical research context, it is common to interpret a result as statistically significant if the difference between groups could have occurred by chance alone in less than 1 time in 20 (5% of the times). This is expressed as a p value lower than 0.05 ($p< 0.05$).

Strategy: clinical treatment regimen or public health program (including program aimed at preventing disease or some health-related characteristic).

Strength of Association: extent to which the intervention is associated with the outcome(s) of interest.

Study checklist: one of five checklists used to evaluate the quality of research depending on study type or study purpose.

Study group: in a randomized controlled trial, the group which receives an experimental preventive or therapeutic procedure or intervention. More generally, the group participating in a study.

Study quality: an assessment of the degree to which bias has been prevented through the design and conduct of the study.

Study type: includes randomized controlled trial, cohort, non-randomized controlled trial, population based case-control, hospital-based case-control, cross-sectional analytic, ecological, descriptive. Randomized controlled trials are the study type of highest quality.

Systematic review: a method to synthesize and analyze the results of different researches on a specific topic using a careful handling of data, mostly by means of statistical tools, called *meta-analysis*.

Target population: population receiving an intervention or for whom an intervention is planned.

Variability: extent to which the results of different studies differ from each other. Variability may occur because of random error or differences in study design, study setting, participants, interventions, exposure(s) or outcome(s) or in the way these are measured.

ACRONYMS AND ABBREVIATIONS

A alendronate
ABMTR Autologous Blood and Marrow Transplant Registry
AERS Adverse Event Reporting System of FDA, known as MedWatch
BPs bisphosphonates
C clodronate
CCTR Cochrane Controlled Trials Register
Chi^2 chi-square
CI confidence interval
CR complete remission
dDV vincristine, dexamethasone and liposomal doxorubicin
DTPACE dexamethasone, thalidomide, cisplatin
E etidronate
EBMT European Group for Blood and Bone Marrow Transplantation
EMEA The European Medicines Agency
FDA Food and Drug Administration is a regulatory agency of the United States of America
GI gastro-intestinal
IBMTR International Bone Marrow Transplant Registry
Ho the null hypothesis
HDT high-dose chemotherapy
HR hazard ratio
I ibandronate
ISS international staging system
IV (intravenous): Within a vein. IV is the abbreviation for "intravenous"
MeSH medical subject headings
MM multiple myeloma
MR minimal response
nCR near complete remission
NE not extractable
NHMRC Australian National Health and Medical Research Council
NNH number needed to harm
NNT number needed to treat
NR not reported
ODAC FDA's Oncologic Drugs Advisory Committee
ω Inverse Variance Weights
ONJ osteonecrosis of jaw
OR objective remission=CR+PR

OR odds ratio
P pamidronate
PD progressive disease
PR partial response
pts patients
QOL quality of life
R risedronate
RCT randomised controlled trial
RD risk difference
RR relative risk/risk ratio
SCT stem cell transplantation
SD stable disease
SE standard error
VAD vincristine, doxorubicin and dexamethasone
VBCMP vincristine, carmustine, cyclophosphamide, melphalan and prednisone
VGPR very good partial response
Z zoledronate

REFERENCES

AAE. (2008). American Association of Endodontists Position Statement: Endodontic implications of bisphosphonate-associated osteonecrosis of the jaws. Available at www.aae.org/ManagedFiles/pub/0/Pulp/bisphosonatesstatement.pdf. Accessed November 20.

Abildgaard N, Rungby J, Glerup H, et al. (1998). Long-term oral pamidronate treatment inhibits osteoclastic bone resorption and bone turnover without affecting osteoblastic function in multiple myeloma. Eur J Hematol. 61:128–134.

Abu-Id MH, Açil Y, Gottschalk J, Kreusch T. (2006). Bisphosphonate-associated osteonecrosis of the jaw. Mund Kiefer Gesichtschir. 10(2):73-81.

Adam Z, Prokes B, Znojil V, et al. (1996). Effect of clodronate on bone density of patients in multiple myeloma-2 year study [Vliv clodronatu na kostin denzitu pacientu s mnohocenym myelomem-dvoulete sledeovani lecebneho ucinku]. Vnitr Lek. 42:379–385.

Agrillo A, Petrucci MT, Tedaldi M, et al. (2006). New therapeutic protocol in the treatment of avascular necrosis of the jaws. J 3. 3. 3. Craniofac Surg. 17(6):1080-1083.

Alexanian R, Barlogie B, Dixon D. (1990). Renal failure in multiple myeloma. Pathogenesis and prognostic implication. Arch Int Medicine. 150:1693-1695.

Alexanian R, Dimopulos M. (1994). The treatment of multiple myeloma. N Eng J Med. 330:484-489.

Ali SM, Esteva FJ, Hortobagyi G, et al. (2001). Safety and efficacy of bisphosphonates beyond 24 months in cancer patients. J Clin Oncol. 19(14):3434-3437.

Altman DG, Schulz KF, Moher D et al. (2001). The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med. 134: 663–694.

Antman EM, Lau J, Kupelnick B et al. (1992). A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. J Am Med Assoc.1992; 268:240–248.

Aparicio A, Gardner A, Tu Y, et al. (1998). In vitro cytoreductive effects on multiple myeloma cells induced by bisphosphonates. Leukemia. 12:220-229.

Apolone G; De Carli G; Brunetti M, Garattini S. (2001). Health-related quality of life (HR-QOL) and regulatory issues. An assessment of the European Agency for the Evaluation of Medical Products (EMEA) recommendation on the use of HR-QOL measures in drug approval. Pharmacoeconomics.2001; 19:187–195.

Apolone, G. (2003). Clinical and outcome research in oncology. The need for integration. Health Qual Life Outcomes. 1:3.

ASCO Special Article. (1996). Outcomes of cancer treatment for technology assessment and cancer treatment guidelines. J Clin Oncol. 14:671–679.

Attal M, Harousseau JL, Leyvraz S, et al. (2006). Maintenance therapy with thalidomide improves survival in multiple myeloma patients. Blood. 108(10):3289-3294.

Atula S, Powles T, Paterson A, et al. (2003). Extended safety profile of oral clodronate after long term nuse in primary breast cancer patients. Drug Safety. 26: 661 – 671.

Avilés A, Nambo MJ, Neri N, et al. (2007). Antitumor effect of zoledronic acid in previously untreated patients with multiple myeloma. Med Oncol. 24(2):227-230.

Badros A, Weikel D, Salama A, et al. (2006). Osteonecrosis of the Jaw in Multiple Myeloma Patients: Clinical Features and Risk Factors. J Clin Oncol. 24(6):945-952.

Bagan JV, Jimenez Y, Murillo J, et al. (2006). Jaw osteonecrosis associated with bisphosphonates: multiple exposed areas and its relationship to teeth extractions. Study of 20 cases [Letter]. Oral Oncol. 42: 327-329.

Bamias A, Kastritis E, Bamia C, et al. (2005). Osteonecrosis of the jaw in cancer after treatment with bisphosphonates: incidence and risk factors. J Clin Oncol. 23:8580-8587.

Bardy AH. (1998). Bias in reporting clinical trials. Br J Clin Pharmacol. 46(2):h147–150.

Baris D, Brawn LM, Silverman DT, et al. (2000). Socioeconomic status and multiple myeloma among U.S. blacks and whites. Am J Public Health. 90:1277-1281.

Barlogie B, van Rhee F, Shaughnessy JD Jr, et al. (2008). Seven-year median time to progression with thalidomide for smoldering myeloma: partial response identifies subset requiring earlier salvage therapy for symptomatic disease. Blood. 112(8):3122-3125.

Barosi G, Boccadoro M, Cavo M, et al. (2004). Management of multiple myeloma and related-disorders: guidelines from the Italian Society of Hematology (SIE), Italian Society of Experimental Hematology (SIES) and Italian Group for Bone Marrow Transplantation (GITMO). Haematologica. 89:717 – 741.

Battley J, Jayathissa S, Seneviratne E. (2006). Jaw osteonecrosis associated with bisphosphonates. N Z Med J. 119(1246):U2341.

Begg CB, Cramer LD, Hoskins WJ, Brennan MF. (1998). Impact of hospital volume on operative mortality for major cancer surgery. J Am Med Assoc.1998; 280:1747–1751.

Belch AR, Bergsagel DE, Wilson K. (1991). Effect of daily etidronate on the osteolysis of multiple myeloma. J Clin Oncol. 9:1397-1402.

Berenson JR, Lichtenstein A, Porter L, et al. (1996). Efficacy of pamidronate in reducing skeletal events in patients with advanced multiple myeloma. Myeloma Aredia Study Group. N Engl J Med. 334:488–493.

Berenson JR, Lichtenstein A, Porter L. (1998). Long-term pamidronate treatment of advanced multiple myeloma patients reduces skeletal events. J Clin Oncol. 16:593–602.

Berenson JR, Rosen LS, Howell A, et al. (2001). Zoledronic acid reduces skeletal-related events in patients with osteolytic metastases. Cancer. 91(7):1191-200.

Berenson J, Hillner B, Kyle R, et al. (2002). American Society of Clinical Oncology Clinical Practice Guidelines: the role of bisphosphonates in multiple myeloma. J Clin Oncol. 20:3719–3736.

Bergner R, Henrich DM, Hoffmann M, et al. (2007). Renal safety and pharmacokinetics of ibandronate in multiple myeloma patients with or without impaired renal function. J Clin Pharmacol. 47(8):942-950.

Bladé J, Samson D, Reece D, et al. (1998). Criteria for evaluating disease response and progression in patients with multiple myeloma treated by high-dose therapy and haemopoietic stem cell transplantation. Myeloma Subcommittee of the EBMT. European Group for Blood and Marrow Transplant. Br J Haematol. 102:1115–1123.

Braun E, Iacono VJ. (2006). Bisphosphonates: case report of nonsurgical periodontal therapy and osteochemonecrosis. Int J Periodontics Restorative Dent. 26(4):315-319.

Brincker JW, Abildgaard N. (1998). Failure of oral pamidronate to reduce skeletal morbidity in multiple myeloma : a double-blind placebo-controlled trial. Br J Haematol. 101:280-286.

Broglia C, Merlati G, Valentino F, et al. (2006). Avascular jaw osteonecrosis associated with bisphophonate therapy. Recenti Prog Med. 97(3):140-144.

Brown JE, Neville-Webbe H, Coleman RE. (2004). The role of bisphosphonates in breast and prostate cancers. Endocr Relat Cancer. 11**:**207 – 224.

Bujanda AD, Sarmiento BU, Suárez CMA, Morales AJ. (2007). Assessment of renal toxicity and osteonecrosis of the jaws in patients receiving zoledronic acid for bone metastasis. Ann Oncol. 18(3):556-560.

Calle EE, Rodriguez C, Walker-Thurmond K, et al. (2003). Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. N Engl J Med. 348(17):1625-1638.

Calvo-Villas JM, Tapia Torres M, Govantes Rodriguez J, et al. (2006). Osteonecrosis of the jaw in patients with multiple myeloma during and after treatment with zoledronic acid. Med Clin (Barc). 127(15):576-579.

Campbell DT. (1957). Factors relevant to the validity of experiments in social settings. Psychol Bull. 54:297-312.

Campbell DT, Stanley JC. (1963). Experimental and quasi-experimental designs for research on teaching. In: Gage NL, ed. Handbook of research on teaching. Chicago: Rand McNally. 171-246.

Canadian Task Force on the Periodic Health Examination. (1979). The Periodic Health Examination. Can Med Assoc J. 121:1193–1254.

Capalbo S, Delia M, Diomede D, et al. (2006). Jaw osteonecrosis associated with use of bisphosphonates and chemotherapy: paradoxical complication of treatment of bone lesions in multiple myeloma patients. Int J Hematol. 83(5):439-442.

Caparrotti G, Catalano L, Feo C, et al. (2003). Perspective study on pamidronate in stage I multiple myeloma. Hematol J. 4(6):459-460.

Caraglia M, Santini D, Marra M et al. (2006). Emerging anti-cancer molecular mechanisms of aminobisphosphonates. Endocr Relat Cancer. 13:7–26.

Carneiro E, Vibhute P, Montazem A, Som PM. (2006). Bisphosphonate-associated mandibular osteonecrosis. AJNR Am J Neuroradiol. 27(5):1096-1097.

Carter G, Goss AN, Doecke C. (2005). Bisphosphonates and avascular necrosis of the jaw: a possible association. Med J Aust. 182:413-415.

Chalmers I. (1990). Underreporting is scientific misconduct. J Am Med Assoc. 263:1405–1408.

Chalmers I, Dickersin K, Chalmers TC. (1992). Getting to grips with Archie Cochrane's agenda. All randomised controlled trials should be registered and reported. Br Med J. 305:786–787.

Chalmers I. Foreword. In Egger M, Smith GD and Altman DG(Eds.). (2001). Systematic reviews in health care: meta-analysis in context. BMJ Publishing Group. London. xiii-xvii.

Chang JT, Green L, Beitz J. (2003). Renal failure with the use of zoledronic acid. N Engl J Med. 349:1676 – 1679.

Chelimsky E. (1995). Politics, Policy and Research Synthesis. Evaluation.1995; 1(1):97-104.

Ciepłuch H, Baran W, Hellmann A. (2002). Combination of pamidronate and thalidomide in the therapy of treatment-resistant multiple myeloma. Med Sci Monit. 8(4):PI31-PI36.

Clarke M, Oxman AD, eds. (1999). Cochrane reviewers' handbook 4.0. In: Cochrane Collaboration. Cochrane Library. Oxford: Update Software.

Clarke BM, Boyette J, Vural E, et al. (2007). Bisphosphonates and jaw osteonecrosis: the UAMS experience. Otolaryngol Head Neck Surg. 136(3):396-400.

Clemens MR, Fessele K, Heim ME. (1993). Multiple myeloma: effect of daily dichloromethylene bisphosphonate on skeletal complications. Ann Hematol. 66:141–146.

Coleman RE, Purohit OP, Black C, et al. (1999). Double-blind, randomised, placebo-controlled, dose-finding study of oral ibandronate in patients with metastatic bone disease. Ann Oncol. 10(3):311-316.

Contopoulos-Ioannidis DG, Karvouni A, Kouri I, Ioannidis JPA. (2009). Reporting and interpretation of SF-36 outcomes in randomised trials: systematic review. Br Med J. 339:a3006.

Cook DJ, Sackett DL, Spitzer WO. (1995). Methodologic guidelines for systematic reviews of randomized control trials in health care from the Potsdam consultation on meta-analysis. J Clin Epidemiol. 48:167-171.

Corso A, Varettoni M, Zappasodi P, et al. (2007). A different schedule of zoledronic acid can reduce the risk of the osteonecrosis of the jaw in patients with multiple myeloma. Leukemia. 21(7):1545-8.

Cox DR, Fitzpatrick R, Fletcher AE, Gore SM, Jones DR, Spiegel-halter DJ. (1992). Quality-of-life assessment: can we keep it simple? J R Stat Soc.155:353-393.

Curi MM, Cossolin GS, Koga DH, et al. (2007). Treatment of avascular osteonecrosis of the mandible in cancer patients with a history of bisphosphonate therapy by combining bone resection and autologous platelet-rich plasma: Report of 3 cases. J Oral Maxillofac Surg. 65(2):349-355.

Dannemann C, Grätz KW, Riener MO, Zwahlen RA. (2007). Jaw osteonecrosis related to bisphosphonate therapy: a severe secondary disorder. Bone. 40(4):828-834.

Daragon A, Humez C, Michot CXLL. (1993). Treatment of multiple myeloma with etidronate results of a multicentre double-blind study. Eur J Med. 2:449-452.

Davis S, Wright PW, Schulman SF et al. (1985). Participants in prospective, randomized clinical trials for resected non-small cell lung cancer have improved survival compared with nonparticipants in such trials. Cancer. 56:1710–1718.

Davis S, Dahlberg S, Myers MH et al. (1987). Hodgkin's disease in the United States: a comparison of patient characteristics and survival in the Centralized Cancer Patient Data System and the Surveillance. Epidemiology, and End Results Program. J Natl Cancer Inst. 78: 471–478.

Delamothe T. (1996). Whose data are they anyway? Br Med J. 312:1241–1242.

Delmas PD, Charhon S, Chapuy MC, et al. (1982). Long-term effects of dichloromethylene diphosphonate (CI2MDP) on skeletal lesions in multiple myeloma. Metab Bone Dis Relat Res. 4(3):163-168.

Dhodapkar MV, Singh J, Mehta J, et al. (1998). Anti-myeloma activity of pamidronate in vivo. Br J Haematol. 103:530-532.

Diego R, D'Orto O, Pagani D, et al. (2007). Bisphosphonate-associated osteonecrosis of the jaws: a therapeutic dilemma. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 103(3):e1-5.

Diel IJ, Solomayer EF, Costa SD et al. (1998). Reduction in new metastases in breast cancer with adjuvant clodronate treatment.Engl J Med. 339(6):357-363.

Dimitrakopoulos I, Magopoulos C, Karakasis D. (2006). Bisphosphonate-induced avascular osteonecrosis of the jaws: a clinical report of 11 cases. Int J Oral Maxillofac Surg. 35(7):588-593.

Dimopoulos MA, Kastritis E, Bamias A, et al. (2006). Osteonecrosis of the jaw in patients with multiple myeloma treated with bisphosphonates: evidence of increased risk after treatment with zoledronic acid. Haematologica. 91:968-971.

Djulbegovic B, Wheatley K, Ross J, et al. (2002). Bisphosphonates in multiple myeloma. The Cochrane Database of Systematic Reviews. 4. John Wiley & Sons, Ltd Chichester, UK.

Durie BG, Kyle RA, Belch A, et al. (2003). Myeloma management guidelines: a consensus report from the Scientific Advisors of the International Myeloma Foundation. Hematology J. 4:379 – 398.

Egger M, Smith GD. (1995). Misleading meta-analysis. Br Med J. 310:752-754.

Egger M, Smith GD. (1997a). Meta-analysis. Potentials and promise. Br Med J. 315:1371-1374.

Egger M, Davey SG, Schneider M, et al. (1997b). Bias in meta-analysis detected by a simple, graphical test. Br Med J. 315:629-634.

Egger M, Smith GD and O`Rourke K. (2001). Rationale, potentials, and promise of systematic reviews. In Egger M, Smith GD and Altman DG(Eds.). Systematic reviews in health care: meta-analysis in context. BMJ Publishing Group. London. 3-22.

Egger M and Smith GD. (2001). Principles of and procedures for systematic reviews. In Egger M, Smith GD and Altman DG(Eds.). Systematic reviews in health care: meta-analysis in context. BMJ Publishing Group. London. 23-42.

Elad S, Yarom N, Hamed W, et al. (2006). Osteomylelitis and necrosis of the jaw in patients treated with bisphosphonates: a comparative study focused on multiple myeloma. Clin Lab Haematol. 28(6):393-8.

Elomaa I, Blomqvist L, Porkka L, et al. (1987). Treatment of skeletal disease in breast cancer: A controlled clodronate trial. Bone. 8(1):53-56.

Estilo CS, Van Poznak CH, Williams T, et al. (2004). Osteonecrosis of the maxilla and mandible in patients treated with bisphosphonates: a retrospective study [Abstract]. Proc Am Soc Clin Oncol. 22:750.

Eysenck HJ. (1978). An exercise in mega-silliness. Am Psychol. 33:517.

Feinstein AR. (1995). Meta-analysis: statistical alchemy for the 21st century. J Clin Epidemiol. 48:71-79.

Feuer EJ, Frey CM, Brawley OW et al. (1994). After a treatment breakthrough: a comparison of trial and population-based data for advanced testicular cancer. J Clin Oncol. 12:368–377.

Ficarra G, Beninati F, Rubino I, et al. (2005). Osteonecrosis of the jaws in periodontal patients with a history of bisphosphonates treatment. J Clin Periodontol. 32:1123-1128.

Fleisch H. (1997). Bisphosphonates in Bone Disease-From the Labratory to the Patient (ed 3). New York, NY, The Pathenon Publishing Group.

Fontana A, Herrmann Z, Menssen HD, et al. (1998). Effects of intravenous ibandronate therapy on skeletal related events (SRE) and survival in patients with advanced multiple myeloma. Blood. 92(1):106a.

Fromer MJ. (2006). Report from FDA Advisory Committee: FDA Drug Safety Problems Many & Serious, Sweeping Changes Needed (FDA InSight). Oncology Times. 28(22)25:8-11.

Garcia-Gara C, Gonzalez-Garcia C, Juliana Majado M, et al. (2006). Osteonecrosis of the Jaw in Multiple Myeloma Patients. Experience of Two Hospitals. Blood. 108 (11):5086.

Garratt A. (2009). Patient reported outcome measures in trials. Br Med J. 338:a2597.

Gibbs SD, O'Grady J, Seymour JF, Prince HM. (2005). Bisphosphonate-induced osteonecrosis of the jaw requires early detection and intervention [Letter]. Med J Aust. 183:549-50.

Girling DJ. (2003). Clinical Trials in Cancer: Principles and Practice. New York: Oxford University Press.

Glasgow RE, Mulvihill SJ. (1996). Hospital volume influences outcome in patients undergoing pancreatic resection for cancer. West J Med. 165:294–300.

Greenstein G. (2003). Clinical versus statistical significance as they relate to the efficacy of periodontal therapy. J Am Dent Assoc. 134 (5):583-591.

Greipp PR, San Miguel J, Durie BMG et al. (2005). International staging system for multiple myeloma. Journal of Clinical Oncology. 20:3412-3420.

Gruppo Italiano per lo Studio della Streptochinasi nell'Infarto Miocardico (GISSI). (1986). Effectiveness of intravenous thrombolytic treatment in acute myocardial infarction. Lancet. 397-402.

Guyatt G., Gutterman D, Baumann M, et al. (2006). Grading strength of recommendations and quality of evidence in clinical guidelines: report from an American College of Chest Physicians Task Force, CHEST. 129:174-181.

Gwede C, Johnson D, Daniels S, et al. (2002). Assessment of toxicity in cooperative oncology trials: The long and short of it. J Oncol Manag. 11:15-21.

Gyorkos T, Tannenbaum TN, Abrahamowitcz M, et al. (1994). An approach to the development of practice guidelines for community health interventions. Can J Public Health. 85(1):8–13.

Hansen T, Kunkel M, Weber A, Kirkpatrick CJ. (2006). Osteonecrosis of the jaws in patients treated with bisphosphonates - histomorphologic analysis in comparison with infected osteoradionecrosis. J Oral Pathol Med. 35(3):155-160.

Harrouseau JL, Dreyling M. (2008). Multiple myeloma: ESMO Clinical Recommendations for diagnosis, treatment and follow-up Ann Oncol.19(2): ii55-ii57.

Hay KD, Bishop PA. (2006). Association of osteonecrosis of the jaws and bisphosphonate pharmacotherapy: dental implications. N Z Dent J. 102(1):4-9.

Hayward RS, Wilson MC, Tunis SR et al. (1995). Users' guides to the medical literature VIII. How to use clinical practice guidelines. A. Are the recommendations valid? The Evidence-Based Medicine Working Group. J Am Med Assoc. 274:570–574.

Hedges LV, Olkin I. (1985). Statistical methods for meta-analysis. Orlando: Academic Press.

Heim ME, Clemens MR, Queißer W, et al. (1995). Prospective randomized trial of dichloromethilene bisphosphonate (clodronate) in patients with multiple myeloma. Onkologie. 18: 439-448.

Herbozo PJ, Briones DL, Ferres AJ, Torrealba RL. (2007). Severe spontaneous cases of bisphosphonate-related osteonecrosis of the jaws. J Oral Maxillofac Surg. 65(8):1650-1654.

Hillner BE, Smith TJ, Desch CE. (2000). Hospital and physician volume or specialization and outcomes in cancer treatment: importance in quality of cancer care. J Clin Oncol. 18:2327–2340.

Hjorth M, Hellquist L, Holmberg E, et al. (1993). Initial versus deferred melphalan-prednisone therapy for asymptomatic multiple myeloma stage I – a randomized study. Myeloma Group of Western Sweden. European Journal of Haematology. 50: 95–102.

Hodgson DC, Zhang W, Zaslavsky AM et al. (2003). Relation of hospital volume to colostomy rates and survival for patients with rectal cancer. J Natl Cancer Inst. 95:708–716.

Hoff AO, Toth BB, Altunday K et al. (2006). Osteonecrosis of the jaw in patients receiving intravenous bisphosphonate therapy. J Clin Oncol. 24:8528a.

Hortobagyi GN, Theriault RL, Porter L, et al. (1996). Efficacy of pamidronate in reducing skeletal complications in patients with breast cancer and lytic bone metastases. N Engl J Med. 335(24):1785-1791.

Horton R. (1993). Data-proof practice. Lancet. 342:1499.

Imrie K, Stevens A, Makarski J, et al. (2007). The role of bisphosphonates in the management of skeletal complications for patients with multiple myeloma: A clinical practice guideline. Published online. http://www.cancercare.on.ca/pdf/pebc6-4f.pdf

International Myeloma Working Group. (2003). Criteria for the classification of monoclonal gammopathies, multiple myeloma and related disorders: a report of the International Myeloma Working Group. Br J Haematol. 121:749-757.

ISIS-2 Collaborative Group. (1988). Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17,187 cases of suspected acute myocardial infarction: ISIS-2. Lancet. ii:349-360.

ISIS-4 Collaborative Group. (1995). ISIS-4: a randomised factorial trial assessing early oral captopril, oral mononitrate, and intravenous magnesium sulphate in 58050 patients with suspected acute myocardial infarction. Lancet.345:669-85.

Janeway C. (2002). Immunologie. 5 Aufl.- Heidelberg; Berlin: Spektrum, Akad. Verl. 304.

Johnson JR; Williams G; Pazdur R. (2003). End points and United States Food and Drug Administration approval of oncology drugs. J Clin Oncol. 21:1404–11.

Jüni P, Altman DG, Egger M. (2001). Assessing the quality of controlled clinical trials. In: Egger M, Davey Smith G, Altman DG, eds. Systematic reviews in health care: meta-analysis in context 2nd ed. London: BMJ Books.

Kademani D, Koka S, Lacy MQ, Rajkumar SV. (2006). Primary surgical therapy for osteonecrosis of the jaw secondary to bisphosphonate therapy. Mayo Clin Proc. 81(8):1100-1103.

Kamangar F, Dores GM, Anderson WF. (2006). Patterns of Cancer Incidence, Mortality, and Prevalence. J Clin Oncol. 24(14):2137-2150.

Karjalainen S, Palva I. (1989). Do treatment protocols improve end results? A study of survival of patients with multiple myeloma in Finland. Br Med J. 299:1069–1072.

Katz H. (2005). Endodontic implications of bisphosphonate-associated osteonecrosis of the jaws: a report of three cases. J Endod. 31:831-834.

Kay NE, Leong T, Bone N, et al. (1998). T-helper phenotypes in the blood of myeloma patients on ECOG phase III trial E9486/E3A93. Br J Haematol. 100:459-463.

Khamaisi M, Regev E, Yarom N, et al. (2007). Possible association between diabetes and bisphosphonate-related jaw osteonecrosis. J Clin Endocrinol Metab. 92(3):1172-1175.

Kraj M, Poglód R, Pawlikowsky J, Maj S. (2000a). The effect of long-term pamidronate treatment on skeletal morbidity in advanced multiple myeloma. Acta Haematologica Polonica. 31:379-389.

Kraj M, Póglod R, Pawlikowski J, Maj S, Nasiloska B. (2000b). Effect of pamidronate on skeletal morbidity in myelomatosis. Part 1: The results of the 12 months of pamidronate therapy. Acta Poloniae Pharmaceutica. 57 (1):113–116.

Kumar V, Pass B, Guttenberg SA, Ludlow J, et al. (2007). Bisphosphonate-related osteonecrosis of the jaws: a report of three cases demonstrating variability in outcomes and morbidity. J Am Dent Assoc. 138(5):602-609.

Kyle RA. (1983). Long-term survival in multiple myeloma. N Engl J Med. 308: 314-316.

Kyle RA. Disperienzi A. (2004). Neurological aspects of monoclonal gammopathy of undermined significance, multiple myeloma, and related disorders. In multiple myeloma and related disorders (ed by G. Gahrton, B.G.M. Durie, D. Samson). Arnold, London. 350-360.

Kyle RA, Yee GC, Somerfield MR, et al. (2007). American Society of Clinical Oncology 2007 clinical practice guideline update on the role of bisphosphonates in multiple myeloma. J Clin Oncol. 25(17):2464-2472.

Lahtinen R, Laakso M, Palva I, et al. (1992). Randomized, placebo-controlled multicentre trial of clodronate in multiple myeloma. Lancet. 340:1049-1052.

Lassere MND, Johnson KR, Woodworth TG, et al. (2005). Challenges and progress in adverse event ascertainment and reporting in clinical trials. Journal of Rheumatology. 32(10): 2030-2032.

LeLorier J; Grégoire G, Benhaddad A et al. (1997). Discrepancies between meta-analyses and subsequent large randomized, controlled trials. N Engl J Med. 337:536–542.

Lenz JH, Steiner-Krammer B, Schmidt W, et al. (2005). Does avascular necrosis of the jaws in cancer patients only occur following treatment with bisphosphonates? J Craniomaxillofac Surg. 33:395-403.

Liberati A, Himel HN, Chalmers TC. (1986). A quality assessment of randomized control trials of primary treatment of breast cancer. J Clin Oncol. 4:942-951.

Liddle J, Williamson M, Irwing L. (1996). Method for Evaluating Research Guideline Evidence. NSW Department of Health, Sydney.

Lieberman MD, Kilburn H, Lindsey M, Brennan MF. (1995). Relation of perioperative deaths to hospital volume among patients undergoing pancreatic resection for malignancy. Ann Surg. 222:638–645.

Lipsey MW, Wilson DB. (2001). Practical meta-analysis (Applied social research methods series; v 49). Thousand Oaks, CA: Sage.

Lugassy G, Shaham R , Nemets A , Ben-Dor D, Nahlieli O. (2004). Severe osteomyelitis of the jaw in long-term survivors of multiple myeloma: a new clinical entity, Am. J. Med. 117:440–441.

Maerevoet M, Martin C, Duck L. (2005). Osteonecrosis of the jaw and bisphosphonates [Letter]. N Engl J Med. 353:99-102.

Magopoulos C, Karakinaris G, Telioudis Z, et al. (2007). Osteonecrosis of the jaws due to bisphosphonate use. A review of 60 cases and treatment proposals. Am J Otolaryngol. 28(3):158-163.

Martín A, García-Sanz R, Hernández J, et al. (2002). Pamidronate induces bone formation in patients with smouldering or indolent myeloma, with no significant antitumour effect. Br J Haematol. 118(1):239-242.

Martoni A, Guaraldi M, Camera P, et al. (1991). Controlled clinical study on the use of dichloromethylene diphosphonate in patients with breast carcinoma metastasizing to the skeleton. Oncology. 48:97-101.

Marunick M, Miller R, Gordon S. (2005). Adverse oral sequelae to bisphosphonate administration. J Mich Dent Assoc. 87:44-9.

Marx RE. (2003). Pamidronate (Aredia) and zoledronate (Zometa) induced avascular necrosis of the jaws: a growing epidemic, J. Oral Maxillofac. Surg. 61:1115–1118.

Marx RE, Sawatari Y, Fortin M, Broumand V. (2005). Bisphosphonate-induced exposed bone (osteonecrosis/osteopetrosis) of the jaws: risk factors, recognition, prevention, and treatment. J Oral Maxillofac Surg. 63:1567-75.

Mavrokokki T, Cheng A, Stein B, Goss A. (2007). Nature and frequency of bisphosphonate-associated osteonecrosis of the jaws in Australia. J Oral Maxillofac Surg. 65(3):415-423.

Mayer D. (2004). Essential Evidence-based Medicine. WB 102 M468e: 22-29.

Mayers C, Panzarella T, Tannock IF. (2001). Analysis of the prognostic effects of inclusion in a clinical trial and of myelosuppression on survival after adjuvant chemotherapy for breast carcinoma. Cancer. 91:2246–2257.

McCloskey EV, MacLennan IC, Drayson MT, et al. (1998). A randomized trial of the effect of clodronate on skeletal morbidity in multiple myeloma. MRC Working Party on Leukaemia in Adults. Br J Hematol. 100:317–325.

McCloskey EV, Dunn JA, Kanis JA, et al. (2001). Long-term follow-up of a prospective, double-blind, placebocontrolled randomised trial of clodronate in multiple myeloma. Br J Haematol. 113:1035-1043.

Melo MD, Obeid G. (2005). Osteonecrosis of the jaws in patients with a history of receiving bisphosphonate therapy: strategies for prevention and early recognition. J Am Dent Assoc. 136(12):1675-1681.

Menssen HD, Sakalová A, Fontana A, et al. (2002). Effects of long-term intravenous ibandronate therapy on skeletal-related events, survival, and bone resorption markers in patients with advanced multiple myeloma. J Clin Oncol. 20(9):2353-2359.

Merigo E, Manfredi M, Meleti M, et al. (2006). Bone necrosis of the jaws associated with bisphosphonate treatment: a report of twenty-nine cases. Acta Biomed. 77(2):109-117.

Merlini G, Parrinello GA, Piccinini L, et al. (1990). Long-term effects of parenteral dichloromethylene bisphosphonate (CL2MBP) on bone disease of myeloma patients treated with chemotherapy. Hematol Oncol. 8:23–30.

Migliorati CA, Schubert MM, Peterson DE, Seneda LM. (2005). Bisphosphonate-associated osteonecrosis of mandibular and maxillary bone: an emerging oral complication of supportive cancer therapy. Cancer.104:83-93.

Moher D, Jadad AR, Nichol G, et al. (1995). Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. Controlled Clin Trials. 16:62-73.

Moher D, Jones A, Lepage L. (2001). Use of the CONSORT statement and quality of reports of randomized trials. J Am Med Assoc. 285:1987-1991.

Montazeri AH, Erskine JG, McQuaker IG. (2007). Oral sodium clodronate induced osteonecrosis of the jaw in a patient with myeloma. Eur J Haematol. 79(1):69-71.

Morris TC, Ranaghan L, Morrison J. (2001). Phase II trial of clarithromycin and pamidronate therapy in myeloma. Med Oncol.18(1):79-84.

Mortensen M, Lawson W, Montazem A. (2007). Osteonecrosis of the jaw associated with bisphosphonate use: Presentation of seven cases and literature review. Laryngoscope. 117(1):30-34.

Mulrow CD. (1987). The medical review article: state of the science. Ann Intern Med.1987; 106:485-488.

Murad OM, Arora S, Farag AF, Guber HA. (2007). Bisphosphonates and osteonecrosis of the jaw: a retrospective study. Endocr Pract. 13(3):232-238.

Musto P, Falcone A, Sanpaolo G, et al. (2003). Pamidronate Reduces Skeletal Events but does not Improve Progression-free Survival in Early-stage Untreated Myeloma: Results of a Randomized Trial.Leukemia & Lymphoma. 44 (9):1545–1548.

Musto P, Petrucci MT, Bringhen S, et al. (2008). A multicenter, randomized clinical trial comparing zoledronic acid versus observation in patients with asymptomatic myeloma. Cancer. 113(7):1588-1595.

Myeloma Trialists' Collaborative Group. (1998). Combination chemotherapy versus melphalan plus prednisone as treatment of multiple myeloma: an overview of 6,633 patients from 27 randomized trials. J Clin Oncol. 16:3832-3842.

National Cancer Institute. (1999). Cancer Therapy Evaluation Program. Common toxicity criteria version 2.0 manual. Bethesda: National Cancer Institute.

Naveau A, Naveau B. (2006). Osteonecrosis of the jaw in patients taking bisphosphonates. Joint Bone Spine. 73(1):7-9.

NCCN. National Comprehensive Cancer Network, Inc. The Multiple Myeloma Clinical Practice Guidelines in Oncology (Version 2.2008). Available at www.nccn.org/professionals/physician_gls/PDF/myeloma.pdf. Accessed November 20, 2008.

NHMR. (2000). National Health and Medical Research Council, NHMR. How to use the evidence: assessment and application of scientific evidence. Canberra: Commonwealth of Australia, AusInfo.

Oxford handbook of oncology. (2006). Oxford University press. New York. 504-505.

Oxman A (Editor). (1994). The Cochrane Collaboration Handbook. Oxford: The Cochrane Collaboration.

Oxman AD, Guyatt GH. (1988). Guidelines for reading literature reviews. Can Med Assoc. 138:697-703.

Palumbo A, Bringhen S, Bertola A, et al. (2004a). Multiple myeloma: comparison of two dose-intesive melphalan regimens (100 vs 200 mg/m^2). Leukemia. 18: 133-138.

Palumbo A, Bringhen S, Petrucci MT, et al. (2004b). Intermediate-dose melphalan improves survival of myeloma patients aged 50-70: results of a randomized controlled trial. Blood. 104:3052-3057.

Parkin DM, Bray F, Ferlay J, Pisani P. (2005). Global cancer statistics, 2002. CA Cancer J Clin. 55:74–108.

Pastor-Zuazaga D, Garatea-Crelgo J, Martino-Gorbea R, et al. (2006). Osteonecrosis of the jaws and bisphosphonates. Report of three cases. Med Oral Patol Oral Cir Bucal. 11:76-79.

Pearn J. (1995). Publication: an ethical imperative. Br Med J. 310:1313–1315.

Perrone F, Maione P, De Maio E, et al. (2002). Survey of modalities of assessing and reporting toxicity in non-comparative prospective studies of chemotherapy in breast cancer. J Clin Oncol. 20:52-57.

Phal PM, Myall RW, Assael LA, Weissman JL. (2007). Imaging findings of bisphosphonate-associated osteonecrosis of the jaws. AJNR Am J Neuroradiol. 28(6):1139-1145.

Pires FR, Miranda A, Cardoso ES, et al. (2005). Oral avascular bone necrosis associated with chemotherapy and biphosphonate therapy. Oral Dis. 11:365-369.

Pirozzo S. (2004a). Searching the medical literature. In Mayer D editor. Essential Evidence-based Medicine. WB 102 M468e: 30-51.

Pirozzo S. (2004b). Study design and strength of evidence. In Mayer D editor. Essential Evidence-based Medicine. WB 102 M468e: 52-61.

Pogue J, Yusuf S. (1998). Overcoming the limitations of current meta-analysis of randomised controlled trials. The Lancet. 351:47-52.

Polizzotto MN, Cousins V, Schwarer AP. (2006). Bisphosphonate-associated osteonecrosis of the auditory canal. Br J Haematol. 132(1):114.

Powles T, Paterson S, Kanis JA, et al. (2002). Randomized, placebo-controlled trial of clodronate in patients with primary operable breast cancer [comment]. J Clin Oncol. 20:3219–3224.

Pozzi S, Marcheselli R, Sacchi S, et al. (2007). Bisphosphonate-associated osteonecrosis of the jaw: a review of 35 cases and an evaluation of its frequency in multiple myeloma patients. Leuk Lymphoma. 48(1):56-64.

Purcell PM, Boyd IW. (2005). Bisphosphonates and osteonecrosis of the jaw. Med J Aust. 2005; 182:417-418.

Riccardi A, Mora O, Tinelli C, et al. (2000). Long-term survival of stage I multiple myeloma given chemotherapy just after diagnosis or at progression of the disease: a multicentre randomized study. Cooperative Group of Study and Treatment of Multiple Myeloma. British Journal of Cancer. 82:1254–1260.

Rodan GA, Fleisch HA. (1996). Bisphosphonates: mechanisms of action.J Clin Invest. 97(12):2692-2696.

Rogers MJ, Gordon S, Benford HL, et al. (2000). Cellular and molecular mechanisms of action of bisphosphonates. Cancer. 88:2961 – 2978.

Rosen LS, Gordon D, Kaminski M, et al. (2001). Zoledronic acid versus pamidronate in the treatment of skeletal metastases in patients with breast cancer or osteolytic lesions of multiple myeloma: a phase III, double-blind, comparative trial. Cancer J. 7(5):377-387.

Rosen L, Gordon D, Kaminski M, et al. (2003). Long term efficacy and safety of zolendronic acid compared with pamidronate disodium in the treatment of skeletal complications in patients with advanced multiple myeloma or breast carcinoma. Cancer. 98(8):1735-1744.

Rosen LS, Gordon DH, Dugan W Jr, et al. (2004). Zoledronic acid is superior to pamidronate for the treatment of bone metastases in breast carcinoma patients with at least one osteolytic lesion. Cancer. 100:36–43.

Ruggiero SL, Meherotra B, Rosenberg TJ, Engroff SL. (2004). Osteonecrosis of the jaws associated with the use of bisphosphonates: a review of 63 cases, J. Oral Maxillofac. Surg. 62:527–534.

Ruggiero S, Gralow J, Marx RE, et al. (2006). Practical guidelines for the prevention, diagnosis, and treatment of osteonecrosis of the jaw in patients with cancer. J Oncol Pract. 2:7-14.

Rychetnik I, Frommer M, Hawe P, Shiell A. (2002). Criteria for evaluating evidence on public health interventions. Journal of Epidemiology and Community Health. 56:119-127.

Salesi N, Pistilli R, Marcelli V, et al. (2006). Bisphosphonates and oral cavity avascular bone necrosis: a review of twelve cases. Anticancer Res. 26(4B):3111-3115.

Schilsky, RL. (2002). End points in cancer clinical trials and the drugs approval process. Clin Cancer Res. 8:935–938.

Schottenfeld D, Fraumeni-JF J, eds. (1996). Cancer epidemiology and prevention. 2nd ed. Oxford, U.K.Oxford University Press.

Schulz KF, Chalmers I, Hayes RJ, Altman DG. (1995). Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. J Am Med Assoc. 273:408-412.

Senel FC, Saracoglu Tekin U, et al. (2007). Severe osteomyelitis of the mandible associated with the use of non-nitrogen-containing bisphosphonate (disodium clodronate): report of a case. J Oral Maxillofac Surg. 65(3):562-565.

Shipman CM, Rogers MJ, Apperley JF. (1997). Bisphosphonates induce apoptosis in human myeloma cells: a novel anti-tumour activity. Br J Haematol. 98:665-672.

Sim I, Hlatky MA. (1996). Growing pains of meta-analysis. Advances in methodology will not remove the need for well designed trials. Br Med J. 313:702–703.

Sitters MA, Caldwell CS. (2005). Bisphosphonates, dental care and osteonecrosis of the jaws. Tex Dent J. 122:968-972.

Smith A, Finn W, Samson D. (2005). Guidelines on the diagnosis and management of multiple myeloma. Br J Haematol. 132:410-451.

Sonnenveld P, van der Holt B, Vellenga E, et al. (2005). Intesive versus double intesive therapy in untreated multiple myeloma: final analysis of the HOVON 24 Trial (abstract). Blood. 106:2545.

Spencer A, Roberts A, Kennedy N, et al. (2008). Renal safety of zoledronic acid with thalidomide in patients with myeloma: a pharmacokinetic and safety sub-study. BMC Clin Pharmacol. 8:2.

Stampfer MJ, Goldhaber SZ, Yusuf S, et al. (1982). Effect of intravenous streptokinase on acute myocardial infarction. Pooled results from randomized trials. New Engl J Med. 307:1180-1182.

Sterne JAC, Egger M, Smith GD. (2001). Investigating and dealing with publication and other biases. In: Egger M, Smith SG, Altman DG editor(s). Systematic reviews in health care. Meta-analysis in context. London: BMJ Books. 189-208.

Tannock IF. (1987). Treating the patient, not just the cancer. N Engl J Med. 317: 1534–1535.

Tanvetyanon T, Stiff PJ. (2006). Management of the adverse effects associated with intravenous bisphosphonates. Ann Oncol. 17:897 – 907.

Tassinari D. (2003). Surrogate end points of quality of life assessment: have we really found what we are looking for? Health Qual Life Outcomes. 1:71.

Tassinari D, Poggi B, Nicoletti S, et al. (2007). Zoledronic acid treatment at home: safety data from an observational prospective trial.J Palliat Med. 10(2):352-358.

Terpos E, Palermos J, Tsionos K, et al. (2000). Effect of pamidronate administration on markers of bone turnover and disease activity in multiple myeloma. Eur J Haematol. 65:331-336.

Terpos E, Viniou N, de la Fuente J, et al. (2003). Pamidronate is superior to ibandronate in decreasing bone resorption, interleukin-6 and beta 2-microglobulin in multiple myeloma. Eur J Haematol. 70(1):34-42.

Tosi P, Zamagni E, Cellini C, et al. (2006a). First-line therapy with thalidomide, dexamethasone and zoledronic acid decreases bone resorption markers in patients with multiple myeloma. Eur J Haematol. 76(5):399-404.

Tosi P, Zamagni E, Cangini D, et al. (2006b). Osteonecrosis of the jaws in newly diagnosed multiple myeloma patients treated with zoledronic acid and thalidomide-dexamethasone. Blood. 108(12):3951-3952.

Treister N, Woo SB. (2006). Images in clinical medicine. Bisphosphonate-associated osteonecrosis of the jaw. N Engl J Med. 355(22):2348.

Tricot G, Vesole DH, Jagannath S, et al. (1996). Graft-versus-myeloma effect: proof of principle. Blood. 87(3):1196-1198.

Trotti A, Bentzen SM. (2004). The need for adverse effects reporting standards in oncology clinical trials. J Clin Oncol. 22:19-22.

Van Holten-Verzantvoort AT, Kroon HM, Bijvoet OL, et al. (1993). Palliative pamidronate treatment in patients with bone metastases from breast cancer. J Clin Oncol. 11:491 – 498.

Vannucchi AM, Ficarra G, Antonioli E, Bosi A. (2005). Osteonecrosis of the jaw associated with zoledronate therapy in a patient with multiple myeloma. Br J Haematol. 28:738.

Vardy J, Tannock IF. (2004). Quality of cancer care. Ann Oncol. 15:1001–1006.

Vogel CL, Yanagihara RH, Wood AJ, et al. (2004). Safety and pain palliation of zoledronic acid in patients with breast cancer, prostate cancer, or multiple myeloma who previously received bisphosphonate therapy. Oncologist. 9(6):687-695.

Walter C, Grötz KA, Kunkel M, Al-Nawas B. (2007). Prevalence of bisphosphonate associated osteonecrosis of the jaw within the field of osteonecrosis. Support Care Cancer. 15(2):197-202.

Wang T, Song ST, Jiang ZF, et al. (2004). Clinical trial on ibandronate in patients with tumor-associated hypercalcemia. Zhonghua Zhong Liu Za Zhi. 26(12):739-741.

Woloshin GWS and Schwarz LM. (2007). How Two Studies on Cancer Screening Led to Two Results. NY Times. 13:18.

Wong R, Wiffen PJ. (2002). Bisphosphonates for the relief of pain secondary to bone metastases. Cochrane Database Syst Rev.2:CD002068.

Woo SB, Hellstein JW, Kalmar JR. (2006). Narrative [corrected] review: bisphosphonates and osteonecrosis of the jaws. Ann Intern Med. 144(10):753-761.

World Health Organization. (1979). WHO handbook for reporting results of cancer treatment. Geneva: World Health Organization.

Wutzl A, Eisenmenger G, Hoffmann M, et al. (2006). Osteonecrosis of the jaws and bisphosphonate treatment in cancer patients. Wien Klin Wochenschr. 118:15-16:473-478.

Yeo AC, Lye KW, Poon CY. (2005). Bisphosphonate-related osteonecrosis of the jaws. Singapore Dent J. 27:36-40.

Yusuf S, Flather M. (1995). Magnesium in acute myocardial infarction. Br Med J. 310:751-752.

Zarychanski R, Elphee E, Walton P, Johnston J. (2006). Osteonecrosis of the jaw associated with pamidronate therapy. Am J Hematol. 81:73-75.

Zervas K, Verrou E, Teleioudis Z, et al. (2006). Incidence, risk factors and management of osteonecrosis of the jaw in patients with multiple myeloma: a single-centre experience in 303 patients. Br J Haematol. 134(6):620-623.