PROTON PUMP INHIBITORS IN BLEEDING PEPTIC ULCER: A META-ANALYSIS

DUANE SHEPPARD

Proton Pump Inhibitors in Bleeding Peptic Ulcer: A Meta-Analysis

By

Duane Sheppard, MD, FRCPC

A Thesis submitted to the

School of Graduate Studies

in partial fulfillment of the

requirements for the degree of

Master of Science (Medicine)

Faculty of Medicine

Memorial University of Newfoundland

January 2006

St. John's Newfoundland

Your file Votre référence
ISBN: 978-0-494-31286-5
Our file Notre référence
ISBN: 978-0-494-31286-5

Abstract

Introduction: The purpose of this meta-analysis was to determine the efficacy of proton pump inhibitors (PPI) in preventing recurrent bleeding and death in bleeding peptic ulcer.

Methods: PubMed and hand searching of cited references was used to identify relevant studies comparing PPI to either placebo or histamine 2 receptor antagonists for the prevention of rebleeding or death. Non english studies were translated Quality assessment and data extraction was perfomed by 2 independent reviewers. The RevMan statistical software was used to combine studies with the peto method.

Results: 15 studies with a total of 3063 patients enrolled were included. The pooled Odds Ratio(pOR) for the use of PPI to prevent rebleeding was significant at 0.55 (95%CI 0.46-0.66). There was a reduction in rebleeding regardless of whether or not therapeutic endoscopy had been used. Infusions of PPI were most beneficial with a pOR of 0.40 (0.29-53) with a number needed to treat to prevent 1 rebleed of only5 while bolus injections failed to provide any benefit over control medication. There was no difference in mortality between PPI and control groups.

Discussion: PPI are effective for the prevention of rebleeding in peptic ulcer disease regardless of whether or not therapeutic endoscopy has been employed. PPI are most effective when given as an infusion. Intravenous PPI should only be used in high risk peptic ulcer bleeding. Further investigations are required to determine the role of oral PPI for the prevention of rebleeding as only 2 studies used oral PPI.

Acknowledgements

I would like to thank Drs. John Fardy, Patrick Parfrey and Brendan Barrett for their support during my work on this thesis.

I would also like to thank the physicians who were all involved in the curriculum while completing the course work for the Master of Science degree.

I would like to thank my secretary for providing secretarial support and assisting with editing this thesis.

Thank you Carla for your input, support and encouragement throughout this process. Also, my family and friends have been supportive all my endeavors.

List of Figures

List of Tables

Table of Contents

1.0 Introduction

1.1 Definition of meta-analysis

Meta-analysis is a tool which has become very popular in medical literature. It involves the process of critically reviewing all the available literature on a selected topic. Most meta-analyses involve assessing treatment effect, that is whether or not a certain treatment will or will not improve disease outcome. The quality of the study is assessed and relevant qualitative data are extracted. Quantitative data are then extracted and combined mathematically yielding an overall estimate of the treatment effect. Thus, meta-analysis involves both qualitative and quantitative synthesis.

The definition of meta-analysis has varied since it was first introduced. In 1959, Mantel and Haenzel wrote their landmark paper describing the statistical aspects of combining data from retrospective studies[1]. However, it was not until 1976 that the term meta-analysis was actually used for the first time[2]. Since then, it has become a very valuable tool to quantify the overall effect of many therapies and to help clinicians in the decision making process.

Meta-analyses are also referred to as systemic reviews as well as overviews. However, there are significant differences between a meta-analysis and a

traditional narrative review. Narrative reviews often reflect the opinion and potential bias of the authors. One such criticism of the narrative review compared it to an essay being written by a student; "if you found something that did not fit your theory, you left it out."[3] The systematic review of the literature decreases bias and meta-analysis by mathematically pools the data strengtheng the conclusions reached compared to that of traditional reviews.

Meta-analysis is now an accepted research tool. Journals such as Evidenced Based Medicine and ACP Journal Club review current medical literature and often include meta-analyses in their reviews[4]. Consensus groups often refer to meta-analyses to summarize the best available data[5]. Currently, conclusions reached in meta-analysis is considered level 1 evidence used to support or contradict medical therapeutics, which is on par with RCTs and cohort studies.

Completing a meta-analysis is analogous to performing a randomized controlled trial. There is a defined stepwise approach which begins with the formulation of a question or hypothesis. One then needs to determine which features are or are not needed in the experimental or control groups. Next, rather than performing an experiment or clinical trial to gather data, data are gathered by reviewing completed research projects. Analysis of the data involves applying the Mantel-Haenzel test or a variation of it to pool the data and to draw conclusions.

1.2 Formulating a Question

Most meta-analyses performed arise from questions encountered in every day clinical work. The question must be clinically relevant and generally arises from our own lack of knowledge. In fact, many of the queries are raised by our patients. It is important to formulate a question which has been well thought out. "Ask a poor question and you will get a poor review."[6] For example, there is a significant difference in the following questions:

1. Does acid suppression reduce poor outcomes in upper gastrointestinal bleeding (UGIB)?

2. Does the use of Proton Pump Inhibitors(PPI) in UGIB from peptic ulcer reduce the incidence of rebleeding or mortality?

The first question raises the possibility of multiple interventions such as oral antacids (eg. calcium and magnesium hydroxide), Histamine 2 Receptor Antagonists(H_2RA), PPI and even somatostatin. Question 1 also raises the possibility of many different sources of UGIB which include many different etiologies. Finally, question 1 does not specify which outcomes are important. Clinical studies on bleeding may utilize several different outcomes such as rebleeding, need for blood transfusion, need for surgery or death. In contrast, question 2 is quite specific in identifying the intervention to be examined, the

specific disease entity and the outcomes of interest.

Even with a well formulated question, there exists the probability that different studies will define endpoints differently. This is not a problem when it comes to an endpoint such as mortality but could very well be important in endpoints such as rebleeding. Therefore, it is important to precisely define what is considered to be relevant rebleeding prior to carrying on to the next step in performing meta-analysis. If studies use similar definitions for the endpoints of interest, heterogeniety may be reduced into the meta-analysis and the end result is likely to be more reliable.

1.3 Determining Study Criteria

Before setting out to find all of the information one can on the topic of interest, the study characteristics of the trials to be included in the meta-analysis need to be determined. When comparing this process to the randomized controlled trial, it is analogous to deciding on inclusion and exclusion criteria for the sample population. However, in a meta-analysis, we are determining the inclusion and exclusion criteria for the studies which will be included from those retrieved in the search.

The inclusion and exclusion criteria will depend on the question being asked. In the typical situation, the effect of a treatment is being examined and several variables need to be determined. These variables concern the study population, treatment allocation (ie: whether the study was a randomized control trial or an observational or retrospective study) and the outcome. Other factors inherently linked to the study design include issues such as randomization, blinding of treatment allocation and blinding of the outcome assessment. These issues will be discussed later in the section on quality assessment.

There are many different methods to examine the efficacy of a treatment. These include retrospective studies, historically controlled studies and prospective randomized controlled trials (RCT). Retrospective studies and studies using

historical controls have several potential sources of bias with the most important one being treatment allocation or allocation bias. The RCT, on the other hand, minimizes allocation bias by randomizing patients provided there is blinding of the allocation process. Therefore, the most important inclusion criteria for a study to be included in a meta-analysis is that it be randomized particularly when dealing with trials of therapy.

The study population should be well described. Comparing groups with similar characteristics may strengthen the meta-analysis (using a single disease process is less likely to introduce heterogeneity). In the example used previously, patients with PUD bleeding may respond differently to a potential therapy than those with either varices or a Mallory Weiss tear. Therefore, the characteristics of the study population from any article should match those set out before retrieval begins. That is, many articles are discovered in the retrieval process but they all need to be examined to determine whether or not they will be included or excluded from the meta-analysis. The study enrollment of various papers affects how generalizable the results are to other populations of interest (age, gender, race etc.)

There are many potential differences in treatment. There are different medication formulations such as oral and injectable forms or formulations from different manufacturers. One needs to determine a-priori whether studies using different methods of treatment should be combined or analyzed separately. The

timing of the administration of the treatment may also be important. In the case of peptic ulcer bleeding, one would prefer to see the medication given as soon as the diagnosis has been established. Other factors related to the treatment would include an adequate description of the dose used and how it was administered: orally, bolus intravenous infusion, continuous intravenous infusion, subcutaneous or intra-muscular injections.

The remaining issue for determining inclusion and exclusion criteria is the outcome of interest. Outcomes need clear definitions and should ideally be assessed by an observer who is unaware of the treatment allocation. For mortality, there is only one definition but some studies may differentiate between disease related mortality or overall mortality. However, rebleeding may be defined in several ways. First, it could be the appearance of bloody aspirate from a naso-gastric tube, observed bleeding at repeat endoscopy or it may have a more clinical definition such as recurrent hematemesis or melena associated with hemodynamic changes or a significant fall in hemoglobin. Therefore, since it is always possible that an endpoint may have different definitions, the endpoint should be clearly defined prior to retrieving any study.

Once the issues of patient population, treatment allocation and outcome assessment have been addressed, it is likely that the type of studies to be included in the analysis has already been determined. That is, if the purpose of the meta-

analysis is to examine the efficacy of a therapeutic intervention, then one would preferably include only prospective randomized studies. However, the decision to include other studies such as those with a historical control group may be tempered by the availability of data. Furthermore, not all randomized studies will be double blind, thus a similar decision will need to be made with how to deal with these studies. Ideally, the double blind RCT would provide the best source of data as it is the "gold standard" of evidence for clinical medicine.

Data may also be found in full peer reviewed articles or in abstract form. The fully published paper offers significant advantages over the abstract in that methodology can be clearly outlined and critically appraised. There are often gaps in both the methods and results from abstracts. There is no clear consensus on how to deal with data from abstracts. The data from abstracts may be included in the primary analysis or included later in a sensitivity analysis. By including the data from abstracts, one can determine if it would have altered any conclusions.

1.4 Heterogeneity

Heterogeneity is one of the most important issues which needs to be considered in a meta-analysis. Stedman's Medical Dictionary defines it as "heterogeneous state or quality."[7] Thus being "composed of parts having various and dissimilar characteristics or properties." Heterogeneity in a meta-analysis may exist on two levels, clinical heterogeneity and statistical heterogeneity.

1.4.1 Clinical Heterogeneity

Clinical heterogeneity results from the differences in and the variability between individual studies compiled together in a meta-analysis. Individual studies are usually designed to answer a single clinical question, which is based on the primary endpoint, but there may also be secondary endpoints. Likewise, meta-analysis attempts to answer specific questions using the data from those studies. Differences between the included studies can lead to clinical heterogeneity within a meta-analysis. Obvious sources of clinical heterogeneity are revealed when comparing patient characteristics, study centres, the intervention and the clinical endpoints used in the various studies.

The patients enrolled in each study may differ in gender, age, comorbid illnesses and there may be difference in the results obtained from patients of different racial origins. We cannot assume that different therapies are equally

effective in men and women. The gender issue has been quite relevant in many early cardiovascular studies interested in the effect of lipid lowering agents on decreasing coronary events in which the majority of patients were men[8]. Such studies limit the generalizability to the general population and combining studies with different gender mixes can also introduce clinical heterogeneity into the meta-analysis.

Age is important in reviewing the effect of any intervention. Again, using cholesterol lowering agents as a typical example, there is an age related decrease in the risk of cardiovascular events with increasing age. Law et. al. found that a decrease of 0.6 mmol/L of serum cholesterol in men resulted in the following decreased risks of Ischemic Heart Disease (IHD) for the respective age group[9]:

Table1: Risk reduction of IHD with 0.6 mmol/L decrease in serum cholesterol

Age	% reduced risk of IHD
40	54
50	39
60	27
70	20
80	19

Clearly, this is a single example, however it demonstrates that we may not be able to generalize the treatment effect observed on one age group to all age groups. Combining such groups can lead to clinical heterogeneity.

Finally, racial difference may be a source of clinical heterogeneity as the

effectiveness of therapies may vary between races. A very recent example of this was the difference in the efficacy of an HIV vaccination program on high-risk individuals. In that study, people of African origin were found to have reduced incidences of HIV acquisition after vaccination compared with those of Caucasian origin[10]. Clearly, combining the results obtained from studies conducted on single race with distinctly different subgroups will likely introduce heterogeneity when meta-analysis is performed.

Thompson et. al.[11]discussed underlying risk as a source of heterogeneity in meta-analysis in their paper. The authors explored the likelihood that there are inherent differences in patient populations for reaching the defined clinical endpoint. The likelihood of a treatment benefiting a higher risk individual may be much higher than in a lower risk individual. The authors used sclerotherapy for esophageal varices as an example and demonstrated heterogeneity is in fact introduced by combining patient populations with differing initial risk. Esophageal varices are graded according to size and high risk features such as a cherry red spot on a varix. Larger varices, those that occlude the esophageal lumen (Grade IV) and those with a "red spot" have the highest risk of bleeding[12]. Smaller varices, approximately 5mm in size (Grade II) are at lower risk of bleeding. Combining studies which have a greater proportion of low risk patients with studies having a a greater proportion of high risk patients will almost certainly introduce

heterogeneity.

The intervention is typically the next source of clinical heterogeneity as there may be differences in the trial medication for different therapeutic trials. There may be differences in the dose used, the route administered (ie orally vs intramuscularly vs intravenously), the timing of drug administration or the duration in which the medication is given. When discussing intervention, it is also important to remember that there may be differences between medications being used within a class of drugs. There are numerous examples of meta-analyses using and combining classes of drugs such as different angiotensin converting enzyme inhibitors and/or angiotensin II receptor blockers[13,14,15]. These two classes of medications which have numerous formulations that can be given one, two, three or even four times daily. It is not important to review the pharmacology of these medications at this time but it is important to acknowledge that there may or may not be a "class effect." Subtle differences betweeen medications may make one better than another. Therefore, combining these medications, which may potentially be given in different doses or by different routes, is an obvious potential source of clinical heterogeneity.

The "class effect" also needs to be explored with regard to potential biochemical or physiologic effects within a group of medications. Beta- blockers are a large class of commonly used medications with different physiologic activity (ie non-

selective and cardioselective). These differences may be important when evaluating different clinical endpoints. In cardiovascular studies, the physiologic differences between beta-blockers may not be as relevant as they are in Gastroenterology (eg. prevention of variceal bleeding). Both non and cardioselective beta-blockers affect the heart but cardioslective beta-blockers have little to no effect on portal pressures within the portal venous system of the liver.[16] Therefore, combining studies utilizing these medications, which may potentially be given in different doses or by different routes, or with different physiologic activity is an obvious potential source of clinical heterogeneity.

The final potential source of clinical heterogeneity is in the definition of the clinical endpoint. For example, there is no way to dispute an endpoint such as mortality, however, not all endpoints are as clearly defined. A clear example of this is found within the Crohn's disease literature where the definition of clinical response to therapy is often variable. Studies will often define clinical response and/or remission based on a scoring system referred to as the Crohn's disease activity index (CDAI). These studies often define remission as a CDAI of less than a score of less than150 and a response usually as a decrease of between 50 and 100 points[17]. Therefore, combining studies that are looking for a response to therapy as opposed to remission can lead to overestimation of the effect of an intervention and contribute to the heterogeneity between studies.

1.4.2 Statistical Heterogeneity

The degree of statistical heterogeneity is determined by a mathematical assessment of the variability between the studies combined in meta-analysis. Statistical heterogeneity may be caused by documented clinical differences or methodological differences between studies or possibly related to unknown or unrecorded trial characteristics[18]. The imprecision of individual trials is demonstrated by their respective confidence intervals about the odds ratio. Subsequently, when combining multiple studies, there needs to be an assessment as to whether the variation in estimates of treatment effect between studies is greater than can be explained by chance alone. This is accomplished by calculating the Chi Square statistic in meta-analysis which is the test of homogeneity and then comparing studies visually.

The Chi Square statistic for heterogeneity (lack of homogeneity) in meta-analysis consists of the mathematical assessment of within study and between study variation between outcomes of the studies included in the systemic review. Heterogeneity exists if the Chi Square statistic is greater than the degrees of freedom(df = number of studies minus one)[18,19,20]. This test is very non-specific and lacks statistical power. Concluding that study outcomes are homogeneous because there is no statistical heterogeneity is not necessarily true. It is therefore necessary to examine the graphical presentation of outcomes. The concept of

comparing studies visually is very simple. After plotting the odds ratios and their confidence intervals of each study on a standard meta-analysis graphical axis (Forest Plot), the upper and lower limits of the confidence intervals are compared. For homogeneous data it is expected that a single line can be drawn down through and intersecting the confidence interval of all studies (eg Fig.1)[18]. When the upper limit of a single study does not cross the lower limit of all other studies, this test of visual inspection will determine that heterogeneity is likely to exist (Fig.2).

Identifying whether or not heterogeneity exists and the potential explanations for that heterogeneity are obviously important to the overall concept of meta-analysis. Dealing with it and deciding what to do if heterogeneity is present in a meta-analysis is a more complex process. Purists argue that if heterogeneity exists, the data should not be combined in meta-analysis[1,3,11,18,19]. However, the purpose of meta-analysis is that of pooling available studies to find the best estimate of the overall treatment effect. Therefore, if heterogeneity is present, it should not preclude statistical pooling with the Mantel-Haenszel test. Performing sensitivity analysis within the framework of the meta-analysis may provide the answers as to why heterogeneity exists[17]. Subgroup analysis is the process of pooling subgroups within the overall analysis to determine if the heterogeneity is resolved in the smaller subgroups. Combining subsets of studies with more similar characteristics such as study population, treatment, and outcomes may be more clinically

homogeneous and allow for subgroup conclusions to be made.

Meier argues, in his commentary on heterogeneity, that we are mostly looking for a quantitative effect rather than a qualitative effect when dealing with heterogenous studies[20]. If, as he states, there is a relatively common effect in all studies rather than a less uniform effect, then studies should be able to be pooled in meta-analysis. If several studies with non-significant results all tend to favour a treatment but the Chi Square statistic suggests heterogeneity, Meier would continue with his analysis. Though he does not state it in his commentary, it would appear that he is indirectly referring to visually assessing Forest plots for heterogeneity. He acknowledges that experts such as Peto would argue against this statement. There is less support for the Meier argument and no other literature supporting his argument could be found. When the desired result is finding the best overall estimate of treatment effect, it seems to make less sense to not pool studies.

Fig.1 Schematic Illustration of homogeneity Visually

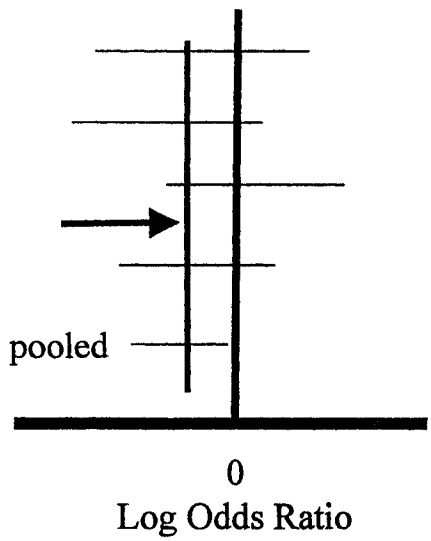
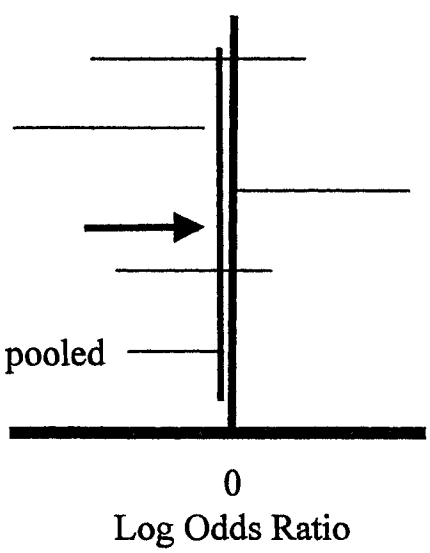


Fig.2 Schematic Illustration of heterogeneity Visually



1.5 Publication Bias in Meta-Analysis

Meta-analysis is a widely accepted research tool, however, there are limitations associated with the methodology. One of the most commonly cited problems with meta-analysis is publication bias.

Publication bias refers to the possibility that "research with statistically significant results is more likely to be submitted and published than work with null or non-significant results"[21]. However, others have expanded publication bias in meta-analysis to English language bias, multiple publication bias and paper selection bias. English Language bias refers to the meta-analysis reviewing only those papers published in English, thus eliminating many potential studies. Positive papers are also more likely to be presented in English journals. Papers may occassionally be submitted to more than one journal resulting in the small possibility that a given study is included more than once but multiple publication also increases the probabilty of discovering the study. Multiple publication bias should be avoided careful review of papers to rule out multiple publication. Paper selection bias refers to the likelihood that journal editors are more likely to publish papers which show a therapeutic advantage over papers which fail to display a benefit. The result of these potential biases is that inappropriate conclusions may be reached as a result of over-representation of published positive papers[22]. It is therefore possible that a potentially non-beneficial therapy could be assumed to be beneficial if only positive

studies were included.

Eggar and Smith have reviewed the potential sources of bias which are involved in meta-analysis[23]. In their paper, they compared two meta-analyses, Leizorovicz et. al.[24] and Nurmohamed et. al.,[25] that examined the role of low molecular weight heparin for prophylaxis of deep venous thrombosis in the post-operative setting. The literature searches both covered the same time period, 1984-91, but the authors selected somewhat different study inclusion criteria and the Nurmohamed et. al. study was limited to English only. Therefore, 39 and 23 studies were included respectively, yet only 18 studies were common to both meta-analyses. This simple example illustrates one of the concerns of those who critique meta-analysis: how can we be certain that all possible data has been included in the analysis?

Easterbrook et. al.[21] wished to examine factors contributing to statistical significance and the effect of statistical significance on publication. After identifying 720 studies that were approved by the ethics department at Oxford University, they contacted the principal investigator of 487 studies. They queried: was the work completed? What were the results? Was it presented or published? Their study showed that statistically significant results had a direct effect on whether or not a study was published or presented at a major meeting.

68% of published and presented studies had statistically

significant results compared with only 65% of those published, 55% of those presented and 29% of those neither published or presented. Conversely, only 15% of studies with statistically significant results remained unpublished or not presented compared with 44% of those with null results. The difference between being published or not based on statistical significance was significantly different using a Chi Square test. Moreover,"the mean citation index of the publishing journal was significantly higher for those studies with statistically significant results compared with non-significant results."

There may be other causes for non-publication of a study other than lack of a statistically significant result. In the Easterbrook et. al. study, the principal investigator was asked if there was any reason for non-publication[21]. Obtaining a null results was cited in over one and a half of instances as the reason for not writing a paper. However, pharmaceutical companies were blamed for non-publication of results in 11% of cases because data was managed at the pharmaceutical company. Editorial rejection accounted for only 9% of non-published studies.

Ioannidis examined the effect of statistical significance on the time to publication of randomized controlled trials sponsored by the National Institutes of Health in the field of the Human Immunodeficiency Virus[26]. He found that positive trials were submitted more rapidly than negative trials(1 versus 1.6 years) and were published

more rapidly after submission(0.8 versus 1.1 years). Therefore, at any point in time, a literature search is more likely to identify positive trials rather than negative trials thus contributing to the possibility of introducing publication bias. The only way that one could possibly retrieve more negative studies is if there had been many more negative than positive studies submitted.

The probability of a study being published in a recognized journal is quite often affected by the language in which the paper is written or alternatively the country in which the paper had been published. For example, Gregoire et. al. identified 36 meta-analyses published in leading English language medical journals and found that 26 of these had limited their search to English language only[27]. They then went on to search the German, Austrian and Swiss medical literature for articles published by lead authors included in any of the cited literature. Studies with significant results were more likely to be published in English. For example, 63% of studies with significant results were published in English versus 35% of studies published in German by the same authors. This resulted in the odds of English publication of 3.8 for studies with statistically significant results. Thus, by excluding Non-English studies, there is significant bias introduced in favour of any therapeutic trial.

Eggar and Smith also identify the impact of database bias as a subset of publication bias. We are unlikely to see the results of studies being done in

developing countries because the journals in developing countries are not cited in medline or similar citation indexes[23]. India, for example, has the highest research output of all developing countries and most journals are published in English. Yet, only 30 of its journals are indexed on Medline[28]. Therefore, this set of research is effectively hidden from any researcher planning a meta-analysis.

Thus far, the problems associated with adequately identifying research studies and how that causes publication bias have been discussed. However, it is also possible to introduce multiple publication bias. Studies with significant results are more likely to have multiple publications and presentations [21]. These studies are thus more likely to be identified in computer searches or cited by other researchers or reviewers. It is also possible that authors include data from the same group of patients in multiple studies. This effectively increases the likelihood of finding a statistically significant study compared to that of finding a negative study. Data which may be retrieved from studies with multiple publications could potentially be duplicated in the data presented in the meta analysis. These factors will usually lead to bias in favour of the therapeutic arm.

Publication bias plays a major role in the delivery of clinical information to the medical community and to those who wish to perform meta-analysis. Because of this, methods to assess publication bias have been devised, including the assessment of funnel plots or the statistical manipulation of the data using a linear regression

technique.

A funnel plot is a graphical representation of effect estimates (taken from each study) plotted against the sample size[29]. Plots of studies avoiding publication bias should resemble an inverted funnel and the plots of studies with publication bias will appear asymmetrical[30]. Duval et. al.[30] and Sutton et. al.[31] have used a trim and fill method to assess whether or not publication bias was present. The assessment of asymmetry in funnel plots is often subjective and observers do not always agree on the appearance[29]. The trim and fill method examines the structure of the funnel plot specifically looking for asymmetry. Sutton et. al. showed asymmetry in 54% of reviews on the Cochrane database and stated that publication bias may be present in up to 50% of meta-analyses but is strongly indicated in 20%[31].

However, the funnel plot can be modified using a linear regression technique. In this linear regression model, the standard normal deviate (SND), defined as the odds ratio divided by its standard error, is regressed against the estimate precision, the latter being defined as the inverse of the standard error. If there is symmetry within the funnel plot, the intercept of the SND will be zero.

Eggar et. al. used the linear regression model to compare the results of meta-analysis to subsequently published large randomized studies to test this model in prediction of whether or not publication bias existed[29]. They found that the 90% confidence interval around the SND intersected with zero in 4 of 4 concordant

studies. That is, the direction and magnitude of effect as shown in the meta-analysis was similar to that of the large randomized study. Furthermore, the 90% confidence interval did not include zero in 3 of 4 discordant pairs (ie: the results of the meta-analysis and the large randomized study differed in regard to the direction or size of the effect).

It is important to attempt to determine whether or not publication bias exists. However, it is not possible to exclude publication bias by using either of the above mentioned methods. Sutton et. al. [31] did not test their trim and fill method in a manner similar to that of Eggar et. al. However, Eggar et. al. only compared the results of 8 meta-analyses to subsequent large randomized studies. He does not explain to us why he used the 90% confidence interval for the SND. If we examine the upper limit of the 90% confidence interval in the 3 pairs of studies with discordant results, we see that it is close to zero. The implications of drawing conclusions based on the comparison of 8 meta-analyses to 8 large randomized studies are paramount. In no other research field could we rely on such a small sample size.

Eggar and his colleagues were criticized in a number of letters to the editor in BMJ. Stuck et. al. question whether the assymetry in funnel plots is due to publication bias, or, is it due to true heterogeneity[32]. Others point to the fact that the linear regression test had a 10% false positive rate which is higher than the standard

acceptable level of 5% for a type 1 error. Clearly, there is no simple answer and no absolute way to ensure that publication bias has been avoided.

1.6 Definition of Peptic Ulcer Disease

Peptic ulcers are mucosal lesions in the stomach (gastric ulcer) or the duodenum (duodenal ulcer). Peptic ulcers occur as a result of a breakdown in the complex interaction between acid and digestive enzymes and the mucosal defence barrier[35]. Clinically, they can present in several different ways with the most common being dyspepsia (upper abdominal pain, nausea or vomiting). Peptic ulcers can also present as a more complicated lesion such as bleeding, upper gastrointestinal obstruction or perforation.[35]

1.7 Epidemiology of Upper Gastrointestinal Bleeding

Bleeding from the upper gastrointestinal (UGI) tract is a common medical problem with significant morbidity and mortality. UGI bleeding accounts for approximately 100 admission per 100 000 population per year[33,34]. Bleeding is most commonly secondary to peptic ulcer disease(PUD) which accounts for approximately 50% of those patients who undergo diagnostic endoscopy[35]. The morbidity from UGI bleeding is due to rebleeding, blood transfusion and emergency surgery. Mortality rates have been estimated to be as high as 14% but in more recent studies, the estimated mortality from bleeding secondary to PUD is approximately 4.1%[35,49].

1.8 Risk of Rebleeding

Recurrent rebleeding from bleeding PUD is a significant problem and the rebleeding rate ranges from 5 to 65% depending on the clinical or endoscopic risk factors or both.

1.8.1 Clinical Data to Predict Rebleeding

Several investigators have examined the ability of clinical data to predict the probability of rebleeding[34,51-54]. These studies consistently conclude that the clinical factors associated with an increased probability of rebleeding include hemodynamic instability and melena. The odds ratio for these factors when present is between 3 and 6. As well, an elevated blood urea nitrogen was present in more patients who had rebleeding but it was not considered significant on logistic regression analysis[34]. Other demographic and laboratory data did not impact on the risk of rebleeding. The problem with using clinical factors to predict rebleeding is they have a low sensitivity. One study by Jaramillo concluded that shock, defined as a systolic blood pressure less than 100 mm Hg, was the strongest clinical predictor of rebleeding[54]. However, only 65 of 312 patients who had rebleeding had shock (sensitivity=21%) and 1161 of 1255 who did not rebleed did not have

shock(specificity=93%). Thus, with a sensitivity of only 21%, it is quite obvious that even the strongest clinical predictor of rebleeding would be quite unsatisfactory to stratify patients into high or low risk of rebleeding.

1.8.2 Endoscopic Signs to Predict Rebleeding

The role of endoscopic assessment in the prediction of rebleeding in PUD has been clearly established[55]. Various endoscopic features of ulcers have been examined, however, the features which consistently show high predictive value for recurrent haemorrhage are active bleeding (spurting or oozing), a non bleeding vessel (NBVV) and fresh clot in the ulcer base resistant to gentle washing. The odds ratio for rebleeding when either of these four endoscopic findings are present ranges between 3 and 17. Other features such as a black spot in the ulcer base and a clean ulcer base are of low predictive value. Laine has reviewed the studies which prospectively examined endoscopic findings as predictors of hemorrhage and these are summarized in Table 2[55]. Patients included in these studies did not have therapeutic endoscopy. Endoscopic predictors will identify approximately 85% of the patients likely to rebleed.

Table 2: The prevalence of endoscopic features in PUD and the probability of further hemorrhage[55]

Endoscopic Characteristic	Prevalence % (range)	Further Bleeding % (range)
Clean base	42 (19-52)	5 (0-10)
Flat spot	20 (0-42)	10 (0-13)
Adherent clot	17 (0-49)	22 (14-36)
NBVV	17 (4-35)	43 (0-81)
Active bleeding	18 (4-26)	55 (17-100)

1.8.3 Clinical Data and Endoscopic Signs to predict rebleeding

Two authors have developed scoring systems which combine clinical data and endoscopic signs in attempt to characterize patients who were at highest risk of rebleeding[33,51]. The first is referred to as the Baylor Bleeding Score[51]. It considers age, the number of comorbid illnesses and whether these illnesses are acute or chronic to calculate a pre-endoscopy score. An endoscopy score is based on the bleeding stigmata associated with the ulcer whether or not it is on the posterior wall bulb. By combining the two scores, a post-endoscopy score is generated. The relevant data needed to calculate this score are summarized in Table 3. This score is used to divide patients into groups with high or low risks of rebleeding after

therapeutic endoscopy has been performed based on the pre-endoscopy score or the post endoscopy score. A pre-endoscopy score of greater than 5 has a sensitivity of 100% and a specificity of 74% in predicting rebleeding while a post-endoscopy score greater than 10 increases the specificity to 79%. Saeed[53] validated his bleeding score in a second prospective study of 100 patients with ulcer hemorrhage. Eight patients who had rebleeding in his study were labelled high risk and no patients labelled as low risk had subsequent bleeding. Therefore, he concluded that the Baylor Bleeding Score accurately identifies those patients at highest risk of rebleeding following therapeutic endoscopy.

Table 3: Clinical scores in the Baylor Bleeding Score

Assign Score	Pre-Endoscopy Score			Endoscopy Score	
	Age	#Illnesses	Acuity		
0	<30	0			Clean base
1	31-49	1-2			Clot
2	50-59				
3	60-69				NBVV**
4		3-4	Chronic	PWB*	Ooze
5	>69	>5	Acute		Spurt

*PWB=posterior wall bulb, **NBVV=non-bleeding visible vessel

Consider a 48 year old man (1 point) presenting with melena and a history of chronic bronchitis (1+4=5 points). His pre-endoscopy score is 6 and thus has a high probability of rebleeding on clinical criteria. If his endoscopy shows a clean base (0 points), he could be safely discharged with therapy as his risk of further bleeding would be quite low (less than 5%) given a post endoscopy score of 6. However, active bleeding (4 or 5 points) would suggest that this man be admitted for monitoring as his post-endoscopy score would be 10 or 11 and his probabilty of rebleeding would be high. The post endoscopy score has more impact on younger patients as one can often change a patient from high risk based on clinical criteria as in the above case to a low risk as in this example. However, if the patient were 72 (5 points) with chronic congestive heart failure admitted with acute pancreatitis (3 diseases = 4 points and acuity = 5 points: total score is 14), the endoscopy score would not be able to prevent the post endoscopy score from exceeding the high risk threshold of 10 points.

The second scoring system which has been developed to assess the risk of rebleeding in PUD is called the Rockall Bleeding Score[34]. However, the Rockall Bleeding Score was not found to be useful in a prospective study by Vreeburg[56] and coworkers in the prediction of rebleeding but did help predict mortality.

1.9 Therapy of Bleeding Peptic Ulcer

The management of bleeding PUD has evolved considerably over the past 3 decades. The percentage of patients who require surgery and the mortality rate from bleeding have both significantly decreased. The major reason for the decline in these two rates has been the development of therapeutic endoscopic hemostasis[56].

Medical management of bleeding PUD has been directed primarily toward reducing gastric acidity. Two classes of pharmaceuticals broadly classified as histamine 2 receptor antagonists(H_2RA) and proton pump inhibitors (PPI) have the ability to increase gastric pH and have been used in attempts to decrease the rates of rebleeding and mortality. The rationale for this approach is based on the fact that neither platelets nor the clotting cascade function optimally in the acidic environment of the UGI tract. Several investigators have examined the function of platelets and clotting factors under the influence of gastric contents[59,60]. The results of these studies consistently show that the low gastric pH renders platelet aggregation and the coagulation cascade ineffective. Furthermore, pepsin enzymatically degrades all of the clotting factors including fibrin, the end product of the clotting cascade.

1.9.1 Therapeutic Endoscopy

Therapeutic endoscopy involves applying a therapeutic modality directly to the site of active bleeding or to an ulcer base which appears to be at high risk of rebleeding. This may involve several different therapeutic options such as injection of either a sclerosant, epinephrine, saline, bipolar coagulation, heater probe therapy or laser therapy. The National Institute of Health consensus guidelines on therapeutic endoscopy in bleeding PUD recommends that all patients with arterial spurting, oozing or those who have a NBVV receive an endoscopic therapeutic modality[61]. Therapeutic endoscopy is not necessary when a clean ulcer base or a pigmented spot is seen. Consideration should be given to removing a clot in the ulcer base that is resistant to gentle washing (this would allow the therapeutic endoscopist to apply therapeutics to the actual lesion). Several randomized controlled trials have investigated the different therapeutic modalities in bleeding PUD and their results have been systemically reviewed by two groups[56,57].

Two authors have analyzed the data on therapeutic endoscopy in bleeding PUD using meta-analysis and demonstrate similar results[56,57]. The meta-analysis by Cook and coworkers found a remarkable reduction in the rates of rebleeding, need for surgery and mortality regardless of the therapeutic method (laser, heater probe or injection) with an odds ratio of 0.57, 0.37 and 0.40 respectively[56]. Benefits of

therapeutic endoscopy were even more significant when lesions at higher risk of rebleeding were analysed. The odds ratios were 0.23 and 0.26 and 0.62 for the risk of recurrent bleeding, need for surgery and mortality respectively. In the subgroup analysis, no therapeutic method was found to give superior results to the other modalities.

1.9.2 Histamine 2 Receptor Antagonists

H_2RA's are a class of pharmaceuticals which have the ability to decrease acid production by chief cells in the gastric mucosa. No randomized controlled trial (RCT) has ever proven a benefit for H_2RA over placebo in decreasing the rate of rebleeding or to decreasing mortality[62,63]. In 1986, a meta-analysis of 26 RCT's failed to show any benefit of H_2RA over placebo[62]. However, there was a small benefit in the reduction of rebleeding in the gastric ulcer subgroup. Subsequently, a large RCT of involving 1005 patients who were at very high risk of rebleeding based on endoscopic criteria also failed to display any benefit of famotidine over placebo[63].

1.9.3 Proton Pump Inhibitors

Proton Pump Inhibitors (PPIs) are a second class of pharmaceuticals and include omeprazole, lansoprazole, esomeprazole, rabeprazole and pantoprazole. PPI's are more potent acid suppressors than H_2RA's. PPI's get absorbed into the chief cells at the basal surface of the gastric parietal cell and bind irreversibly to the "proton pump," an active hydrogen-potassium exchange pump. This pump uses ATP energy to actively exchange potassium cations for hydrogen cations on the basal surface of the parietal cell. The hydrogen cation is then secreted into the gastric lumen which acidifies the gastric lumen.

The first RCT using omeprazole in an attempt to reduce the rate of recurrent bleeding was in 1992[43]. It included UGI bleeding from all potential causes but failed to show any benefit in favour of omeprazole in the PUD subgroup. Since then several RCT's have been published with varying results[36-42]. There has been no consensus as to the benefit of PPI's to prevent rebleeding or death in bleeding peptic ulcer. One such study carried out on India and published by Khuroo[38] was heavily criticized. This study was a randomized double blind placebo controlled trial of high dose oral omeprazole which included patients with endoscopic criteria placing them at high risk of rebleeding. There was a very significant reduction in rebleeding and the need for surgery in the omeprazole group and no difference in

mortality. The major problem with this study is that therapeutic endoscopy was not performed on any patients in this study while it is the standard of care in North America[61].

2.0 Objective

The purpose of this study is to systemically review the results of published RCT's to better define the role of PPI's in bleeding PUD. The primary endpoints chosen were rebleeding and overall mortality. Subgroup analysis was planned to examine the effect of therapeutic endoscopy on both rebleeding and mortality. Other analyses were planned for infusional PPI and bolus PPI compared with placebo for both rebleeding and mortality.

3.0 Methods

3.1 Selection of Topic

The therapeutic use of proton pump inhibitors in bleeding peptic ulcer disease (PUD) lends itself well to meta-analysis. The selection of this topic is based on several factors. First, bleeding PUD is a common disorder causing significant morbidity and mortality. Prior to the introduction of proton pump inhibitors as acid suppressing medication, histamine receptor antagonists such as ranitidine were the most important acid suppressors, yet, their use did not reduce either morbidity or mortality in peptic ulcer bleeding. Since then, many studies examining the effect of proton pump inhibitors on bleeding from PUD have been published. It was initially a study from India[38] which made us consider performing a systematic review on the topic.

3.2 Literature Search

The literature review was begun by carrying out a computer search using the search engine PubMed. Pubmed is a recognized search engine provided free of charge on the internet by the National Library of Medicine in the USA. The goal of the search is to identify as many relevant studies as possible and therefore the following key words were utilized and cross referenced:

1. omeprazole or pantoprazole or lansoprazole or esomeprazole or rabeprazole or proton pump inhibitor. This selection of terms allows us to select the group of PPI's currently in use thus ensuring that all members of the class have been included. The linkage term "or" expands the search so that all of the terms will be included.

This was then linked to the following using the cross-linking Boolean "and". Using "and" for cross-linking selects studies which include all the terms of interest.

2. rebleeding or mortality. These are the main clinical endpoints of interest.
3. peptic ulcer or non-variceal. Many published studies on upper GI bleeding have differentiated between variceal and non-variceal rather than peptic ulcer bleeding. Therefore, by including non-variceal, studies which may have had a subgroup of peptic ulcer will not be missed.
4. randomized controlled trial or clinical trial. By using these we hope to help the screening process by eliminating observational type studies.

English Language was not pre-selected as an exclusionary tool. Foreign Language studies were to be retrieved and translated.

The references from relevant studies were reviewed to identify any further published studies. Occasionally, some authors may have had knowledge of other relevant studies and cited those. These were also included in the systematic review

Finally, Index Medicus continues to reference published medical studies and it was also reviewed for the appropriate time period using the terms included in the medline search. Paper searching continues to be the most time consuming part of meta-analysis, however, searching Index Medicus helps to ensure no relevant studies have been excluded. Meeting abstracts were not reviewed because sensitivity analysis with abstracts included was not considered in the original plan of this thesis.

3.3 Inclusion/Exclusion Criteria

To be included in the analysis, studies needed to be prospective and randomized with an adequately identified control group (either placebo or H_2RA). H_2RA were considered an adequate control group because they have not been shown to have an overall significant benefit when compared to placebo[62-63]. The treatment protocols had to identify the timing, route and dose of the specific medications given. Outcomes had to include the endpoints rebleeding and/or mortality. Only papers published in full were considered. Studies appearing as abstracts do not provide detailed methodology to fully determine whether or not they should be included in the analysis. Retrospective studies and historical control studies were excluded because they do not constitute an adequate control group and definition of clinical endpoints has to be determined retrospectively.

3.4 Qualitative Analysis

The data from papers meeting the inclusion criteria were assessed. The quality of published studies was assessed using a modification of the protocol developed by Chalmers et al.[64]. This assessment placed major importance on randomization and blinding. Other extracted information included:

1. Definition of the study population. Where were the patients recruited from, as in were they seen in the emergency room, or were they already in the hospital with prior illnesses, demographics of the study population.
2. Whether or not therapeutic endoscopy was performed. Studies not adhering to the National Institute of Health's guidelines on therapeutic endoscopy[61] were considered not to have performed therapeutic endoscopy. It was decided after discussion that if a study performed therapeutic intervention only on a subgroup of patients which technically should all have received therapeutic endoscopy, that study would then be sub-grouped with studies not performing therapeutic endoscopy.
3. Medication used, route given and the doses for the study. It was anticipated that several different PPIs would be utilized and they may be given by alternate routes. Therefore, it was planned to gather this data prior to carrying out any analysis so that if subgroups needed to be examined, the data would be prepared.

4. Single or multi-centered. Although the data collected from single or multi-centered studies is not treated any differently in meta-analysis, multi-centered studies have the advantage and may in fact supply better quality data as there are fewer opportunities to introduce investigator bias based on well developed protocols. However, there is also the possibility having more variability in the administration of the protocols as well.
5. Definition of rebleeding. The definition of rebleeding was considered appropriate if it included: recurrent hematemesis with fresh blood or fresh blood up a nasogastric tube, recurrent melena with a change in hemodynamic status defined by an increase in heart rate by 10 beats per minute or a drop in systolic blood pressure to 100 mm Hg (after the patient had been stabilized hemodynamically), or a drop of 20 g/L in hemoglobin over a 24 hour period. Rebleeding could also include active bleeding at the time of repeat endoscopy. Coffee ground emesis (sometimes defined as hematemesis) was not considered to be active rebleeding unless it was associated with one of the above.

3.5 Data Extraction

Data was extracted independently by 2 reviewers (DS and JF). This included the number of experimental and control subjects and the number of subjects in each group who had rebleeding or who died. These data were collected on standardized forms in 2X2 tables. There were no discrepancies.

3.6 Quantitative Analysis

Rebleeding and mortality were considered the two primary analyses in this meta-analysis. Data were collected for both endpoints as described above for both the treatment (PPI group) or control medication (placebo or H_2RA). It was planned to pool all studies initially and then to perform meta-analysis on selected subgroups. Subgroups were identified prior to retrieving data and included:

1. Therapeutic endoscopy. One of the major criticisms of the study which sparked interest in this meta-analysis was that therapeutic endoscopy had not been performed in what would be considered high risk patients. These concerns were initially raised at a journal club for Medical Residents at Memorial University by one of the staff Gastroenterologists. Therefore, it was felt that this would be an important subgroup of patients to see if those concerns and criticisms were founded.
2. Dose of the medication given and whether it was administered as a bolus or infusion. It was not known whether there was an optimal dose of PPI to be given or whether it should be given by bolus injection or by a continuous infusion.

The RevMan statistical software package (as supplied by the Cochrane Collaboration, http://hiru.mcmaster.ca/cochrane/) was used to perform the statistical analysis. Data was entered in simple 2x2 tables identifying the control and experimental groups. After identifying which studies were to be included in each

analysis, these were pooled and the output recorded. The statistical package calculated the odds ratios and 95% confidence intervals for each study and then pooled the data. The Peto-method with a fixed effects model was selected to calculate the pooled odds ratio's and 95% confidence interval.

Homogeneity was assessed both visually and with the Chi square test of homogeneity. Heterogeneity is discussed extensively in the introduction. The Chi Square statistic is calculated by the RevMan software.

Publication bias, which was discussed extensively in the introduction, was also assessed. This was assessed using the funnel plot method and also the linear regression of the funnel plot method. The Stata statistical software package was used to perform funnel plots and linear regression.

The number needed to treat (NNT) was calculated by taking the reciprocal of the difference between the bleeding rate in the control and treatment groups.

4.0 Results

4.1 Literature Search

The literature search identified 60 studies. 43 of these were excluded because they were letters, reviews or did not relate specifically to the topic. The remaining 17 studies were reviewed[35-49]. Two of these were excluded from the analysis. One because it used historical controls,[66] the other because it's endpoint was successful control of hemorrhage as shown by repeat endoscopy on day 4 [44].

4.2 Qualitative Analysis

The 15 studies included in the meta-analysis are summarized in Table 4. All studies used rebleeding as a major endpoint while 12 also included mortality data. Only 5 studies were double blind. Five studies[37,39,40,46,50] did therapeutic endoscopy in all patients according to the National Institute of Health guidelines[61] while another study performed therapeutic endoscopy on only patients with Forrest IA (defined as arterial spurting from the ulcer base) bleeding[42]. This last study was grouped with those not performing therapeutic endoscopy according to the NIH guidelines.

Table 4: Studies of Proton Pump Inhibitors in Bleeding Peptic Ulcer

Author	Study Population	Design	Medication/route
Daneshmend	All with UGI bleeding No specific criteria PUD subgroup	Randomized before Endo Single Centre - Blinded & no TE	IV-Om (80 &40 TID) vs Pl
Lin-1	Endoscopic high risk bleeding PUD Forrest Ia-b & IIa	Randomized after Endo Single Centre - Not Blinded & all TE	IV-Om (40&160inf) vs Cim (300 and 1600)
Lin- 2	Endoscopic high risk bleeding PUD Forrest IIa	Randomized after Endo Single Centre - Not Blinded & no TE	IV-Om (40od or bid) vs Cim (300 q6h)
Khuroo	Endoscopic high risk bleeding PUD Forrest Ia-b & IIa-b	Randomized after Endo Single Centre - Blinded & no TE	PO-Om (40 bid) vs Pl
Schaffalitzky	Endoscopic high risk bleeding PUD Forrest Ia-b & IIa-b & 2/3 of SBP<100, HR>100, or Hgb<7	Randomized after Endo Multi-Centre – Blinded & all TE	IV-Om (80 &8 mg/h) vs Pl
Lanas	Endoscopic high risk bleeding PUD Forrest Ib & IIa-b	Randomized after Endo Single Centre not Blinded & no TE	IV-Om (80 &40 bid vs Ran (50 q4h)
Villanueva	Endoscopic high risk bleeding PUD Forrest Ia-b	Randomized after Endo Single Centre not Blinded & all TE	IV-Om (80 &40 bid) vs Ran (50 q6h)
Hasselgren	Endoscopic high risk bleeding PUD - Forrest Ia-b & IIa-b All pts > 60 yrs	Randomized after Endo Multi-Centre – Blinded TE in F-Ia only	IV-Om (80 & 8mg/h) vs Pl
Orti	Endoscopic high risk bleeding PUD Forrest Ib & IIa-b	Randomized after Endo Single Centre not Blinded & no TE	IV-Om (80 &40 tid) vs Ran (50 q6h)
Uribarrena	Not all Endoscopic high risk and 207/282 PUD 77/207 Forrest Ia-b or IIa	Randomized after Endo Single Centre not Blinded & no TE	IV-Om (80 &40 bid) vs Cim (1200inf)
Michel	Endoscopic high risk bleeding PUD Forrest Ia-b & IIa-b	Randomized after Endo Multi-Centre not Blinded & no TE	PO-Lans (30 bid) vs Ran (300 bid)
Prassler	Endoscopic high risk bleeding PUD Forrest Ib & IIa-b	?Randomized after Endo Single Centre not blinded & all TE	IV-Om (80 &40 tid) vs Ran (50 &100q6h)
Flores	Endoscopic high risk bleeding PUD Forrest Ib & IIa-b	Randomized after Endo Single Centre not Blinded & no TE	IV-Om (80 &40 tid) vs Ran (50 &100q6h)

Brunner	Endoscopic high risk bleeding PUD Forrest Ib & IIa-b All patients in ICU & severely ill	Randomized after Endo Single Centre not Blinded & no TE	IV-Om (80 inf) vs Ran (400 inf)
Lau	Endoscopic high risk bleeding PUD Forrest Ia-b & IIa	Randomized after Endo Single Centre, Blinded & all TE	IV-Om (80 & 8mg/h) Vs Pl

PUD = peptic ulcer disease, TE = therapeutic endoscopy, IV = intravenous, PO = per oral, Om = omeprazole, Ran = ranitidine, Cim = cimetidine, Lans = lansoprazole and Pl = placebo, inf = infusion

4.3 Statistical Analysis

4.3.1 Rebleeding

Combining the 15 studies which examined rebleeding shows a statistically significant benefit of PPI versus either placebo or H_2RA with a pooled odds ratio of 0.55 (95% CI: 0.46 - 0.66)(Fig. 3). There was statistical heterogeneity (Chi Square = 42, df=15) between studies. Three studies had either upper or lower limits of their respective 95% CI which did not intersect with either the upper or lower limit of the remaining 12 studies[36,43,46]. Re-analyzing the data without two of these studies did not alter the pooled odds ratio[36,43] (Fig.4) The rationale to exclude these studies and to re-analyze the data was based on the fact that the entry criteria for these two studies were significantly different than the enrollment criteria for the others. The Brunner[36] study used what should be considered the highest risk patients (see description of Baylor Bleeding Score in the introduction) for enrollment as these included patients already admitted to an intensive care unit with other unrelated illnesses[36]. The Daneshmend study, however, used lower risk patients, in that patients were enrolled prior to endoscopy and many of these would have had a low risk of rebleeding according to the endoscopic appearance of the ulcer[42]. The remaining 13 studies were more homogeneous in the type of patient recruited into the study. No explanation could be given as to why the Prassler study seemed to be an outlier.

The pooled odds ratio when the experimental drug was given without therapeutic endoscopy was 0.54 (0.43 - 0.67)(Table 5) while it was 0.58 (0.42 - 0.79) when therapeutic endoscopy was performed (Table 5). The pooled odds ratio remained significant regardless of the dose of medication used (studies were divided based on whether they gave 80 mg of PPI per day or more than 80mg). However, there was a significant difference in favor of a reduction of rebleeding when the PPI was given as an infusion rather than as bolus injections (studies which used oral PPI[38,47] were excluded from this analysis because the comparison involved only intravenous formulations of PPI). The pooled odds ratio was 0.40 (0.29 – 0.53) for infusion therapy while it was 0.82 (0.63 – 1.07) for the group receiving bolus therapy.

The number of patients required to be treated with PPI to prevent one further episode of rebleeding ranged from 5.8 to 13.8 (Table 5) depending on the dose of PPI, the manner in which it was given and whether or not therapeutic endoscopy had been performed.

Fig. 3 : # Overall Rebleeding

	Citation	EffectName	Yea	Treated	Control	Effect	Lower	Upper	NTotal	PValue
	Brunner	Rebleeding		3 / 19	17 / 20	.03	.01	.19	39	.00
	Daneshmend	Rebleeding		58 / 246	70 / 257	.82	.55	1.23	503	.35
	Flores	Rebleeding		2 / 38	8 / 43	.24	.05	1.23	81	.07
	Hasselgren	Rebleeding		51 / 159	75 / 163	.55	.35	.87	322	.01
	Khuroo	Rebleeding		12 / 110	40 / 110	.21	.10	.44	220	.00
	Lanas	Rebleeding		6 / 28	9 / 23	.42	.12	1.45	51	.17
	Lau	Rebleeding		5 / 120	24 / 120	.17	.06	.47	240	.00
	Lin-1	Rebleeding		2 / 50	12 / 50	.13	.03	.63	100	.00
	Lin-2	Rebleeding		4 / 26	5 / 13	.29	.06	1.36	39	.11
	Michel	Rebleeding		8 / 38	11 / 37	.63	.22	1.80	75	.39
	Orti	Rebleeding		11 / 252	15 / 267	.77	.35	1.70	519	.51
	Prassler	Rebleeding		30 / 106	33 / 126	1.11	.62	1.99	232	.72
	Schaffalitzky	Rebleeding		20 / 134	37 / 140	.49	.27	.89	274	.02
	Uribarrena	Rebleeding		6 / 131	6 / 151	1.16	.36	3.69	282	.80
	Villanueva	Rebleeding		13 / 45	11 / 41	1.11	.43	2.85	86	.83
Fixed	Combined (15)			231 / 1502	373 / 1561	.57	.47	.70	3063	.00
Random	Combined (15)			231 / 1502	373 / 1561	.47	.32	.69	3063	.00

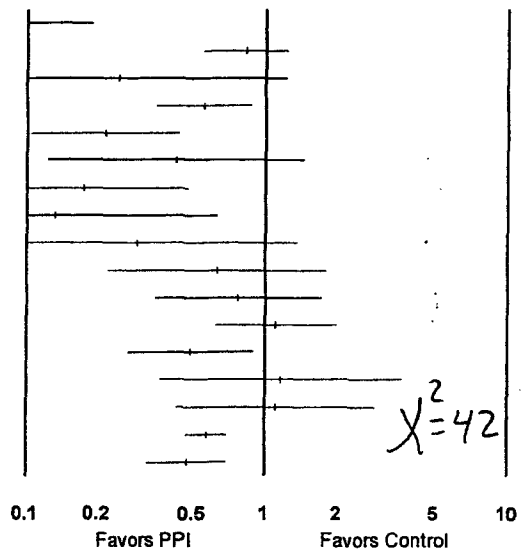


Fig. 4: # Rebleeding Excluding Outliers

	Citation	EffectName	Yea	Treated	Control	Effect	Lower	Upper	NTotal	PValue
	Flores	Rebleeding		2 / 38	8 / 43	.24	.05	1.23	81	.07
	Hasselgren	Rebleeding		51 / 159	75 / 163	.55	.35	.87	322	.01
	Khuroo	Rebleeding		12 / 110	40 / 110	.21	.10	.44	220	.00
	Lanas	Rebleeding		6 / 28	9 / 23	.42	.12	1.45	51	.17
	Lau	Rebleeding		5 / 120	24 / 120	.17	.06	.47	240	.00
	Lin-1	Rebleeding		2 / 50	12 / 50	.13	.03	.63	100	.00
	Lin-2	Rebleeding		4 / 26	5 / 13	.29	.06	1.36	39	.11
	Michel	Rebleeding		8 / 38	11 / 37	.63	.22	1.80	75	.39
	Orti	Rebleeding		11 / 252	15 / 267	.77	.35	1.70	519	.51
	Prassler	Rebleeding		30 / 106	33 / 126	1.11	.62	1.99	232	.72
	Schaffalitzky	Rebleeding		20 / 134	37 / 140	.49	.27	.89	274	.02
	Uribarrena	Rebleeding		6 / 131	6 / 151	1.16	.36	3.69	282	.80
	Villanueva	Rebleeding		13 / 45	11 / 41	1.11	.43	2.85	86	.83
Fixed	Combined (13)			170 / 1237	286 / 1284	.54	.43	.67	2521	.00
Random	Combined (13)			170 / 1237	286 / 1284	.49	.34	.72	2521	.00

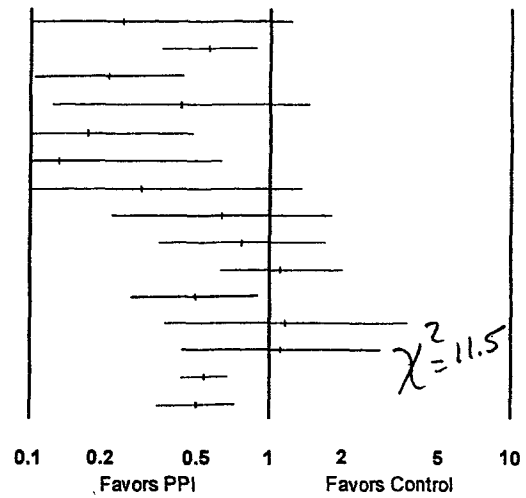


Fig. 5: Rebleeding Without Therapeutic Endoscopy

	Citation	EffectName	Yea	Treated	Control	Effect	Lower	Upper	NTotal	PValue
	Brunner	Rebleeding		3 / 19	17 / 20	.03	.01	.19	39	.00
	Daneshmend	Rebleeding		58 / 246	70 / 257	.82	.55	1.23	503	.35
	Flores	Rebleeding		2 / 38	8 / 43	.24	.05	1.23	81	.07
	Hasselgren	Rebleeding		51 / 159	75 / 163	.55	.35	.87	322	.01
	Khuroo	Rebleeding		12 / 110	40 / 110	.21	.10	.44	220	.00
	Lanas	Rebleeding		6 / 28	9 / 23	.42	.12	1.45	51	.17
	Lin-2	Rebleeding		4 / 26	5 / 13	.29	.06	1.36	39	.11
	Michel	Rebleeding		8 / 38	11 / 37	.63	.22	1.80	75	.39
	Orti	Rebleeding		11 / 252	15 / 267	.77	.35	1.70	519	.51
	Uribarrena	Rebleeding		6 / 131	6 / 151	1.16	.36	3.69	282	.80
Fixed	Combined (10)			161 / 1047	256 / 1084	.54	.43	.68	2131	.00
Random	Combined (10)			161 / 1047	256 / 1084	.46	.29	.72	2131	.00

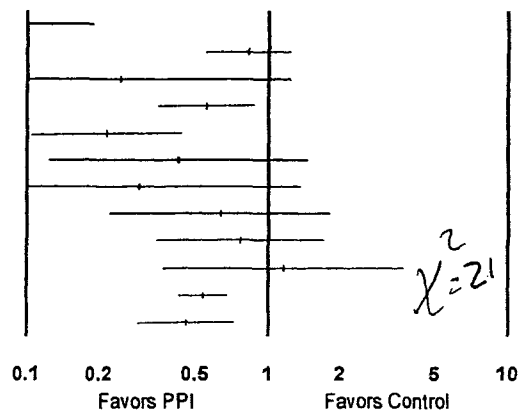
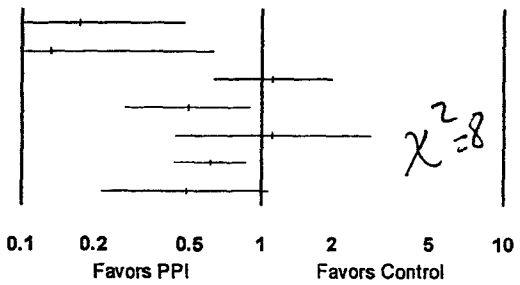


Fig 6: Rebleeding After Therapeutic Endoscopy

	Citation	EffectName	Yea	Treated	Control	Effect	Lower	Upper	NTotal	PValue
	Lau	Rebleeding		5 / 120	24 / 120	.17	.06	.47	240	.00
	Lin-1	Rebleeding		2 / 50	12 / 50	.13	.03	.63	100	.00
	Prassler	Rebleeding		30 / 106	33 / 126	1.11	.62	1.99	232	.72
	Schaffalitzky	Rebleeding		20 / 134	37 / 140	.49	.27	.89	274	.02
	Villanueva	Rebleeding		13 / 45	11 / 41	1.11	.43	2.85	86	.83
Fixed	Combined (5)			70 / 455	117 / 477	.61	.43	.86	932	.01
Random	Combined (5)			70 / 455	117 / 477	.48	.21	1.08	932	.08



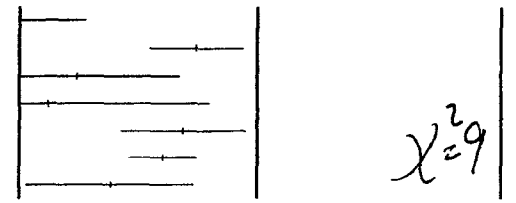
0.1 0.2 0.5 1 2 5 10

Favors PPI Favors Control

Fig 7:

Rebleeding Infusion PPI

	Citation	EffectName	Yea	Treated	Control	Effect	Lower	Upper	NTotal	PValue
	Brunner	Rebleeding		3 / 19	17 / 20	.03	.01	.19	39	.00
	Hasselgren	Rebleeding		51 / 159	75 / 163	.55	.35	.87	322	.01
	Lau	Rebleeding		5 / 120	24 / 120	.17	.06	.47	240	.00
	Lin-1	Rebleeding		2 / 50	12 / 50	.13	.03	.63	100	.00
	Schaffalitzky	Rebleeding		20 / 134	37 / 140	.49	.27	.89	274	.02
Fixed	Combined (5)			81 / 482	165 / 493	.40	.29	.56	975	.00
Random	Combined (5)			81 / 482	165 / 493	.24	.11	.55	975	.00



0.1 0.2 0.5 1 2 5 10

Favors PPI Favors Control

Fig. 8:

Rebleeding Bolus PPI

	Citation	EffectName	Yea	Treated	Control	Effect	Lower	Upper	NTotal	PValue
	Daneshmend	Rebleeding		58 / 246	70 / 257	.82	.55	1.23	503	.35
	Flores	Rebleeding		2 / 38	8 / 43	.24	.05	1.23	81	.07
	Lanas	Rebleeding		6 / 28	9 / 23	.42	.12	1.45	51	.17
	Lin-2	Rebleeding		4 / 26	5 / 13	.29	.06	1.36	39	.11
	Orti	Rebleeding		11 / 252	15 / 267	.77	.35	1.70	519	.51
	Prassler	Rebleeding		30 / 106	33 / 126	1.11	.62	1.99	232	.72
	Uribarrena	Rebleeding		6 / 131	6 / 151	1.16	.36	3.69	282	.80
	Villanueva	Rebleeding		13 / 45	11 / 41	1.11	.43	2.85	86	.83
Fixed	**Combined (8)**			**130 / 872**	**157 / 921**	**.83**	**.63**	**1.08**	**1793**	**.16**
Random	**Combined (8)**			**130 / 872**	**157 / 921**	**.83**	**.63**	**1.08**	**1793**	**.16**

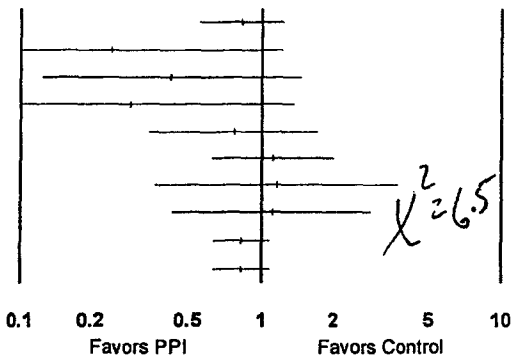


Table 5: Results of Subgroup analysis and the Number Needed to Treat

Endpoint	**Number of studies**	**Sample Size**	**Pooled OR**	**95% CI**	**NNT**
Rebleeding	15	3063	0.55	0.46 - 0.66	11.7
Rebleeding - no TE	10	2131	0.58	0.46 - 0.73	12.2
Rebleeding - TE	5	932	0.69	0.48 - 0.99	11.1
Rebleeding - Inf	5	975	0.40	0.29 - 0.53	5.8
Rebleeding - bolus	8	1793	0.82	0.63 – 1.07	N/A
Rebleeding - 80 mg	7	792	0.39	0.27 - 0.57	8.3
Rebleeding - >80 mg	8	2271	0.61	0.50 - 0.76	13.8
Mortality	12	2910	1.32	0.90 – 1.94	N/A
Mortality – no TE	7	1978	1.52	0.94 – 2.46	N/A
Mortality - TE	5	932	0.77	0.45 – 1.33	N/A

TE = Therapeutic Endoscopy

Inf = Infusion

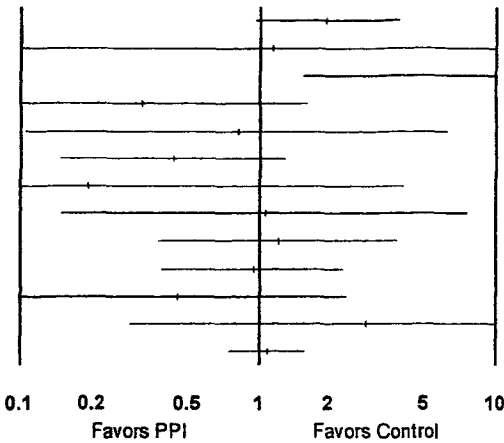
N/A = Not Applicable

4.3.2Mortality

Pooling the 12 studies which provided mortality data showed no statistical difference in mortality between the experimental and control groups with a pooled odds ratio (95% CI) of 1.13 (0.79 – 1.62)(Fig. 9). There was no statistical heterogeneity. Subgroup analysis revealed differing results based on whether or not therapeutic endoscopy was performed. There was a trend toward a reduction in mortality when PPI's were given after therapeutic endoscopy with a pooled odds ratio was 0.77 (0.45 – 1.33)(Table 5). In contrast there was a trend toward an increased mortality in those studies in which therapeutic endoscopy was not performed with a pooled odds ratio of 1.52 (0.94 - 2.46) (Table 5). There was no heterogeneity in the therapeutic endoscopy subgroup, however, there was significant heterogeneity in the subgroup of studies not performing therapeutic endoscopy. There were no significant differences in mortality in the other subgroups of patients.

Fig. 9: Overall Mortality

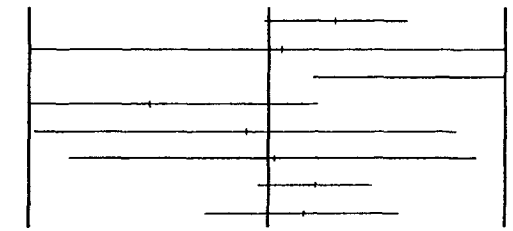
	Citation	EffectName	Yea	Treated	Control	Effect	Lower	Upper	NTotal	PValue
	Daneshmend	Mortality		23 / 246	13 / 257	1.94	.96	3.91	503	.06
	Flores	Mortality		0 / 38	0 / 43	1.13	.02	58.32	81	.95
	Hasselgren	Mortality		11 / 159	1 / 163	12.04	1.54	94.40	322	.00
	Khuroo	Mortality		2 / 110	6 / 110	.32	.06	1.63	220	.15
	Lanas	Mortality		2 / 28	2 / 23	.81	.10	6.23	51	.84
	Lau	Mortality		5 / 120	11 / 120	.43	.14	1.28	240	.12
	Lin-1	Mortality		0 / 50	2 / 50	.19	.01	4.10	100	.24
	Orti	Mortality		2 / 252	2 / 267	1.06	.15	7.58	519	.95
	Prassler	Mortality		6 / 106	6 / 126	1.20	.38	3.84	232	.76
	Schaffalitzky	Mortality		10 / 134	11 / 140	.95	.39	2.31	274	.90
	Uribarrena	Mortality		2 / 131	5 / 151	.45	.09	2.37	282	.34
	Villanueva	Mortality		3 / 45	1 / 41	2.86	.29	28.62	86	.35
Fixed	Combined (12)			66 / 1419	60 / 1491	1.09	.74	1.60	2910	.68



$\chi^2 = 18$

Fig 10 : Mortality Without Therapeutic Endoscopy

	Citation	EffectName	Yea	Treated	Control	Effect	Lower	Upper	NTotal	PValue
	Daneshmend	Mortality		23 / 246	13 / 257	1.94	.96	3.91	503	.06
	Flores	Mortality		0 / 38	0 / 43	1.13	.02	58.32	81	.95
	Hasselgren	Mortality		11 / 159	1 / 163	12.04	1.54	94.40	322	.00
	Khuroo	Mortality		2 / 110	6 / 110	.32	.06	1.63	220	.15
	Lanas	Mortality		2 / 28	2 / 23	.81	.10	6.23	51	.84
	Orti	Mortality		2 / 252	2 / 267	1.06	.15	7.58	519	.95
Fixed	**Combined (6)**			**40 / 833**	**24 / 863**	**1.58**	**.90**	**2.76**	**1696**	**.11**
Random	**Combined (6)**			**40 / 833**	**24 / 863**	**1.41**	**.55**	**3.61**	**1696**	**.48**



$\chi^2 = 6$

Fig.11 : Mortality After Therapeutic Endoscopy

	Citation	EffectName	Yea	Treated	Control	Effect	Lower	Upper	NTotal	PValue
	Lau	Rebleeding		5 / 120	24 / 120	.17	.06	.47	240	.00
	Lin-1	Mortality		0 / 50	2 / 50	.19	.01	4.10	100	.24
	Prassler	Mortality		6 / 106	6 / 126	1.20	.38	3.84	232	.76
	Schaffalitzky	Mortality		10 / 134	11 / 140	.95	.39	2.31	274	.90
	Villanueva	Mortality		3 / 45	1 / 41	2.86	.29	28.62	86	.35
Fixed	Combined (5)			24 / 455	44 / 477	.60	.35	1.05	932	.07
Random	Combined (5)			24 / 455	44 / 477	.65	.22	1.88	932	.42

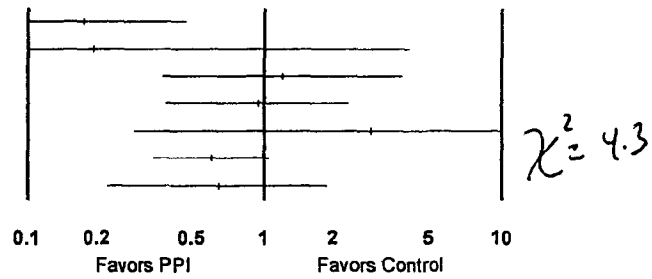


Fig. 12: Mortality - PPI Infusion

	Citation	EffectName	Yea	Treated	Control	Effec	Lowe	Uppe	NTotal	PValue
	Hasselgren	Mortality		11 / 159	1 / 163	12.04	1.54	94.40	322	.00
	Lau	Mortality		5 / 120	11 / 120	.43	.14	1.28	240	.12
	Lin-1	Mortality		0 / 50	2 / 50	.19	.01	4.10	100	.24
	Schaffalitzky	Mortality		10 / 134	11 / 140	.95	.39	2.31	274	.90
Fixed	Combined (4)			26 / 463	25 / 473	.86	.45	1.63	936	.64
Random	Combined (4)			26 / 463	25 / 473	1.02	.25	4.08	936	.98

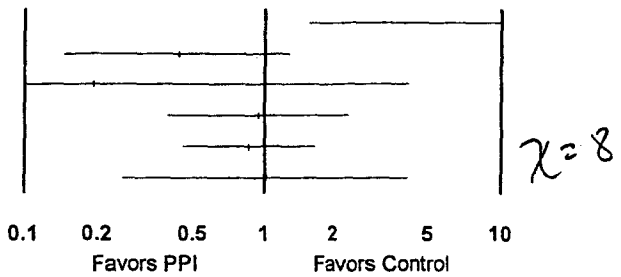


Fig 13: Mortality - PPI Bolus

	Citation	EffectName	Yea	Treated	Control	Effect	Lower	Upper	NTotal	PValue
	Daneshmend	Mortality		23 / 246	13 / 257	1.94	.96	3.91	503	.06
	Flores	Mortality		0 / 38	0 / 43	1.13	.02	58.32	81	.95
	Lanas	Mortality		2 / 28	2 / 23	.81	.10	6.23	51	.84
	Prassler	Mortality		6 / 106	6 / 126	1.20	.38	3.84	232	.76
	Uribarrena	Mortality		2 / 131	5 / 151	.45	.09	2.37	282	.34
	Villanueva	Mortality		3 / 45	1 / 41	2.86	.29	28.62	86	.35
Fixed	**Combined (6)**			**36 / 594**	**27 / 641**	**1.45**	**.85**	**2.45**	**1235**	.17
Random	**Combined (6)**			**36 / 594**	**27 / 641**	**1.45**	**.85**	**2.45**	**1235**	.17

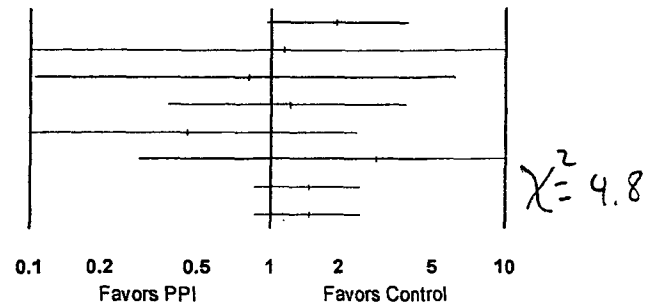
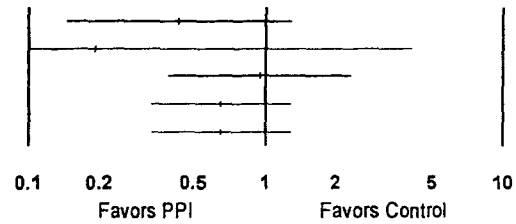


Fig. 14:

Mortality After Therapeutic Endoscopy with PPI Infusion

	Citation	EffectName	Yea	Treated	Control	Effect	Lower	Upper	NTotal	PValue
	Lau	Mortality		5 / 120	11 / 120	.43	.14	1.28	240	.12
	Lin-1	Mortality		0 / 50	2 / 50	.19	.01	4.10	100	.24
	Schaffalitzky	Mortality		10 / 134	11 / 140	.95	.39	2.31	274	.90
Fixed	Combined (3)			15 / 304	24 / 310	.65	.33	1.27	614	.21
Random	Combined (3)			15 / 304	24 / 310	.65	.33	1.27	614	.21



4.4 Publication Bias

Funnel plots were performed using two methods, both of which yielded identical results. Figures 15 and 16 plot the standard error and precision by effect size for rebleeding data respectively. The graphical presentation shows several studies lying outside the boundary of the triangle and the inverted funnel for their respective plots. Figure 15, which plots precision of the Log Odds ratio (defined by the reciprocal of the standard error of each study) against the Log Odds Ratio identified 8 of 15 studies outside the boundaries of the inverted funnel. Similarly, Figure 16 displayed nearly identical results as only 4 studies fully lie inside the triangle. Therefore, it is quite likely that publication bias exists in the publication of studies examining the effect of PPI therapy of upper GI bleeding. If there were as few as 4 small negative studies not published, the funnel plot for figure 15 would be quite symmetrical.

Funnel plots for the standard error and precision by effect size were also plotted for the mortality data. These are presented in Figures 17 and 18 respectively. Both figures identified 4 full studies lying outside the boundaries of the triangle and inverted funnel. This points to the existence of publication bias in the mortality data as well.

Fig 15: **Funnel Plot of Precision by Effect Size**

Rebleeding

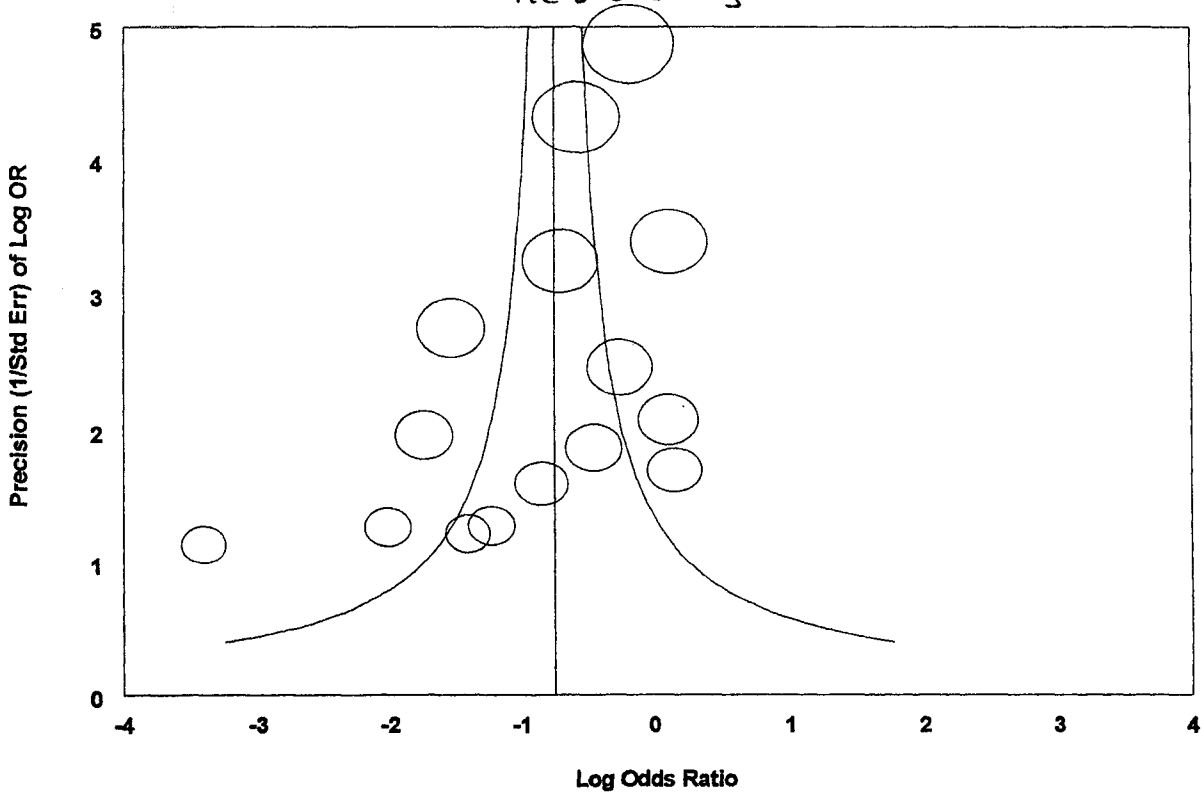


Fig 16: **Funnel Plot of Standard Error by Effect Size**

for Rebleeding

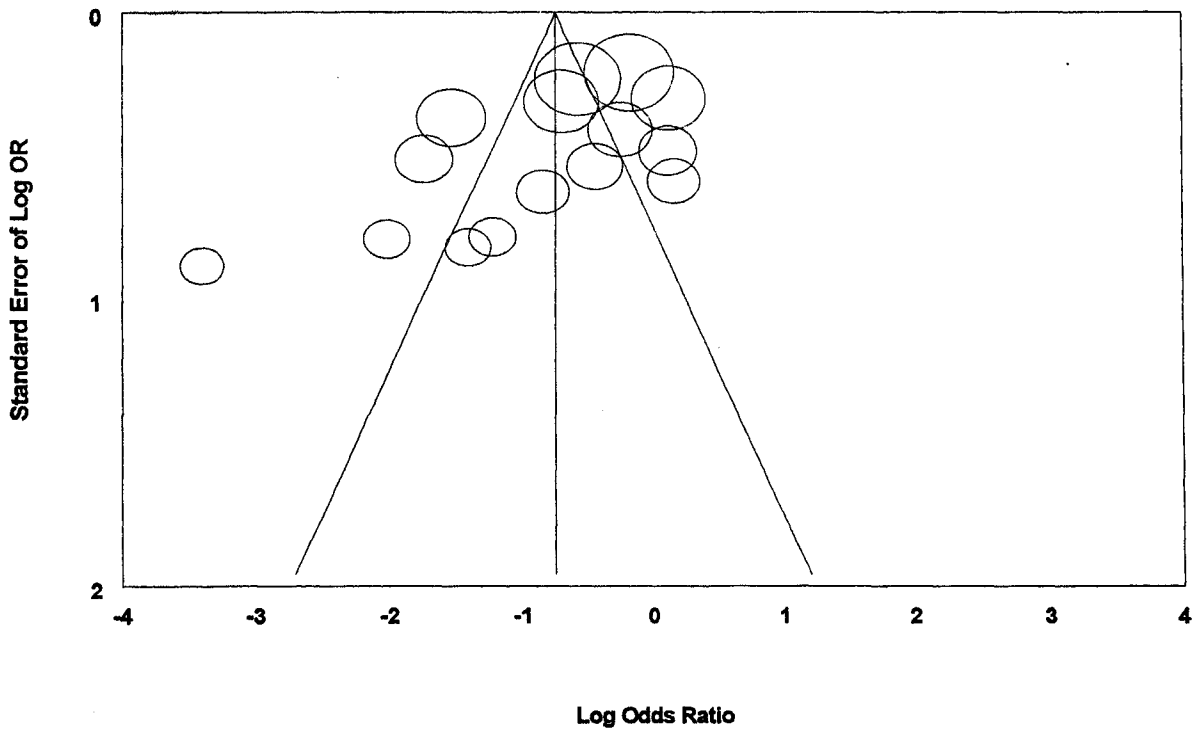


Fig 17: **Funnel Plot of Precision by Effect Size**

for Mortality

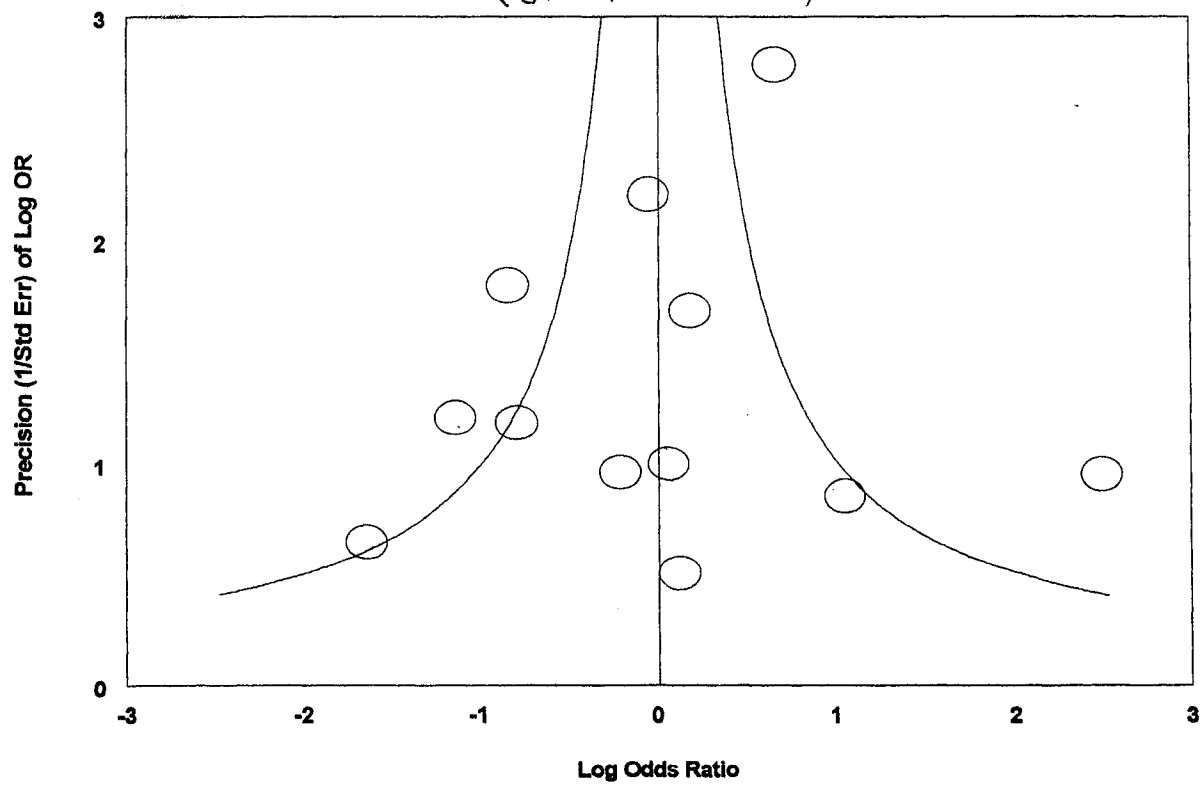
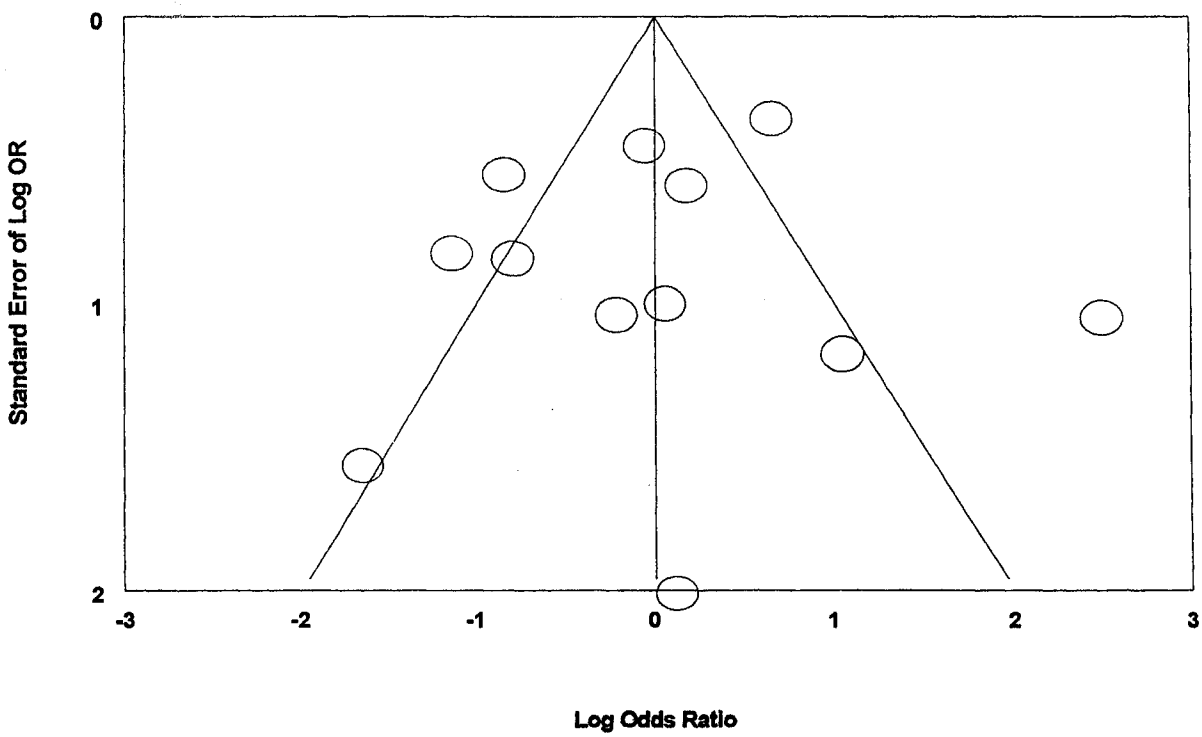


Fig 18: **Funnel Plot of Standard Error by Effect Size**

for mortality



5.0 Discussion

5.1 Interpreting the Results

The goal of treatment for patients with UGIB from PUD is the prevention of further bleeding. The results of this systematic review of 15 studies with a total sample size of approximately 3000 patients indicates that therapy with PPI in high risk patients with bleeding PUD (Forrest I - II) is beneficial with a significant reduction in the rebleeding rate. The benefit of PPI therapy in the prevention of rebleeding persisted regardless of whether or not therapeutic endoscopy had been performed. There is added benefit to giving intravenous PPI's as an infusion rather than as bolus injections, given the lack of a significant benefit in the bolus subgroup. Yet, two studies included in this systematic review used oral PPI, one demonstrating a significant benefit[38] in favor of PPI and the other[44] showing no benefit. The optimal dose of PPI for a high risk bleeding peptic ulcer is yet to be determined as there was no difference in the reduction of rebleeding when the 80 mg doses were compared to more than 80 mg.

It has long been speculated that control of acid output from the stomach would help to reduce the rate of recurrent bleeding. In-vitro studies clearly demonstrate that neither the fibrin plug nor the platelet clot is stable at the low pH found in the normal stomach[59,60]. Brunner has shown that an infusion of PPI after an initial bolus

provided more control of acid output and prevented decreases in pH relative to bolus injections in normal subjects[65]. It is not surprising then that the subgroup analysis of PPI infusion compared with the control group had the most significant reduction in rebleeding with a NNT of only 6. The PPI bolus subgroup analysis was the only subgroup which failed to demonstrate a benefit over the control medication for the prevention of rebleeding. This further supports the conclusion that PPI should be given as an infusion after an initial bolus to patients with high risk bleeding ulcer. No studies compared bolus injections to infusions or had three experimental arms with an accepted placebo group. Therefore, this conclusion is not fully supported by a good randomized controlled trial, however, there is strong evidence based on this meta-analysis to support infusion therapy over bolus.

There was statistical heterogeneity between studies which included rebleeding as an endpoint. Excluding the two of the three outlying studies resolved the heterogeneity[36,43]. Visually inspecting the studies in Figure 1 demonstrated that the Brunner study had an upper limit 95% confidence interval for it's Odds Ratio which failed to overlap with the lower limit of the 95% confidence interval of several other studies Odds Ratios[36,37,39,41,42,46-9]. However, the lower limit of the 95% confidence interval for the Odds Ratios for both the Prassler[49] and Daneshmend[43] study failed to overlap with the upper limit of the 95% confidence interval of the Khuroo[38] study Odds Ratio resulting in it being considered an outlier. Interestingly though, the

lower limit of the 95% confidence intervals for both the Daneshmend and Prassler studies did overlap with the pooled Odds Ratio and it's 95% confidence interval.

Two of the three studies which were most likely responsible for this statistical heterogeneity were in fact different methodologically from the others with respect to patient recruitment. All other studies had recruited patients from the Emergency rooms at the time of admission and included only those with high risk bleeding ulcers according to the Forrest classification. However, Daneshmend et. al.[43] included all sources of UGIB, with a peptic ulcer disease subgroup. Subjects were given medication prior to endoscopy and thus there was no stratification of patients thereby including many patients who were at low risk of recurrent bleeding. This approach would technically dilute the drug effect as many patients enrolled in their study would only have had a 5% chance of rebleeding[55].

Brunner et. al. included only patients who developed PUD bleeding while in an intensive care unit for other severe illnesses[36]. This study demonstrated a very significant therapeutic gain with a narrow confidence interval in favor of PPI, despite its small size. This study was not blinded, though not the only one, and therefore investigator bias could have been a factor. However, there may be potential for more therapeutic gain in seriously ill patients. The Baylor Bleeding Score[51,52] is a predictive model for the probability of rebleeding in bleeding peptic ulcer disease after therapeutic endoscopy. It utilizes a combination of clinical and endoscopic

criteria to predict the probability of rebleeding after effective endoscopic hemostasis. It would have predicted a high rate of recurrent bleeding in the Brunner study even with therapeutic endoscopy (none was employed during this study). It is thus possible that PPIs are more efficacious in seriously ill patients[36]. This study was not blinded and again this has the potential to introduce investigator bias.

PPI's provided no therapeutic gain over placebo/H_2RA with regard to mortality. The overall pooled Odds Ratio was nearly 1 indicating no difference between treatment and control. Surprisingly, the pooled odds ratio's were different when groups were compared based on whether or not therapeutic endoscopy had been performed. Studies were divided based on whether or not therapeutic endoscopy had been performed and the medication was given after endoscopy in both subgroups. The paper by Hasselgen et. al. was the only study to show a significant difference in mortality, however, it was in favor of the control group as 11 subjects died in the experimental group compared to only 1 in the placebo group[42]. We have no explanation for this finding. Most deaths in this study were due to co-morbid illnesses and unrelated to the UGIB in a period up to 3 months after presentation.

It is not surprising that there was no difference in overall mortality in this meta-analysis. There was an overall mortality rate of 3.8% in the control group. Calculating the size of a study needed to detect a 20% reduction in mortality would require nearly 100 000 individuals to be enrolled and it would require a sample size

of 22000 patients to detect a 50% reduction in mortality. It is extremely unlikely that any further large studies looking for such a PPIs benefit in the reduction of mortality will ever be performed. PPIs are currently widely used in clinical practice and given the costs and time associated with performing randomized clinical trials, it would not make economic sense.

Therapeutic endoscopy must continue to be the *gold standard* in the treatment of bleeding PUD. The meta-analysis by Cook and coworkers found a remarkable reduction in the rates of rebleeding and mortality regardless of the therapeutic method used (laser, heater probe or injection)[56]. The odds ratios were 0.57and 0.40 respectively. Benefits of therapeutic endoscopy were even more significant when ulcers with stigmata with a higher risk of rebleeding were analyzed: these odds ratios were 0.23 for rebleeding and 0.62 for mortality respectively. There is a further reduction in rebleeding when PPIs are given after therapeutic endoscopy in this meta-analysis, pOR of 0.53 while there is a trend toward a decrease in mortality, pOR of 0.65 when PPIs are given post therapeutic endoscopy. Though there have been no head to head comparisons, the Odds Ratio for rebleeding without therapeutic endoscopy was still 0.54, approaching the odds ratio of the overall odds ratio for therapeutic endoscopy (Cook meta-analysis). However, given the significant reduction in mortality seen with therapeutic endoscopy, it must remain the primary treatment for bleeding from peptic ulcers. PPIs should be considered for primary

treatment of bleeding PUD where the expertise for therapeutic endoscopy do not exist.

Somatostatin and octreotide have also been used in the past for actively bleeding PUD. A meta-analysis by Imperiale and Birgisson displayed similar reductions in the rates of rebleeding as the PPI's[67]. The pooled odds ratio in that meta-analysis for overall rebleeding was 0.53 (0.43 – 0.63). This raises the possibility of another therapeutic choice in the treatment of high risk PUD. There are no clinical trials comparing these two groups of medications and therefore direct comparisons are unavailable. There is no evidence available to suggest an added benefit of somatostatin or octreotide after therapeutic endoscopy or in conjunction with PPI's.

In conclusion, we recommend the use of PPI in high risk bleeding peptic ulcers following appropriate therapeutic endoscopy. The appropriate dose and choice of agent is yet to be resolved, however, it would appear that infusions of intravenous PPIs following an initial bolus provided more benefit than bolus PPIs alone. This conclusion is supported by the fact that infusions of PPIs provide greater in vivo acid suppression than bolus PPIs[65].

There have been no documented cost analysis for the use of PPI therapy in bleeding PUD. The cost involved with administration of PPI for a typical three day period (drug cost in Canada is approximately $60 per day) is less than the cost of repeated thereapeutic endoscopy, surgery and the risks involved with blood transfusion

should the patient need blood. A thorough cost benefit analysis with multiple decision trees would be quite useful.

5.2 The Debate between Heterogeneity and Publication Bias

The assessment of heterogeneity is performed by calculating the Chi-Square statistic and by comparing overlapping confidence intervals visually whereas publication bias is assessed by plotting an effect size against the precision (sample size). The net result is that the two plots can take on very similar appearances particularly when the standard meta-analysis plot is arranged by sample size. Therefore, it seems that heterogeneity and publication bias are assessing a similar element within meta-analysis, variance.

An extensive search of the literature was unable to discover an authoritative statistical review of heterogeneity and publication bias together. Figure 18 which plotted the width of the 95% confidence interval against sample size for rebleeding data demonstrated the existence of publication bias. This non-mathematically manipulated graphical presentation demonstrated that the confidence interval for many of the smaller studies is much narrower than that of some of the larger studies. These data could not be transposed over the standard meta-analysis to graphically compare the two.

Finally, in the introduction, it was stated that heterogeneous data should not be combined with meta-analysis. However, if the heterogeneity can be explained, then it is reasonable to combine similar studies. An argument can be made against not combining heterogeneous data on the whole premise that the purpose of the

systematic review is to provide an unbiased presentation of all available published data. Provided a thorough search had been performed, data should be attained from a multitude of sources which are distinct. That is so of this study which included data on patients from China, India, Arabic countries, Germany, Spain, the UK and the Netherlands. Data attained from such distinct groups and combined in a meta-analysis provides the best overall summary available on the use of proton pump inhibitors in the bleeding peptic ulcer.

This thesis summarized a large amount of data but there is still data which should be attained in further research. Two studies used oral PPI, one demonstrated a benefit in favour of the PPI[38] and one showed no difference[47]. The smaller study was obviously not powered to detect a significant difference between treatment and control and a post hoc combination of the two studies does show a benefit in favour of oral PPI[47]. Neither of these studies used therapeutic endoscopy and are therefore more limited in their generalizability to the standard of care of upper GI bleeding in Canada. The future direction of research for PPI therapy in UGI bleeding could consist of comparing oral to intravenous PPI. Currently, drug costs for three days of treatment with intravenous PPI costs approximately $210.00 (~13.50/40 mg vial of pantoprazole) while an equivalent dose of oral PPI would cost approximately $40.00 ($2.20/40 mg tablet – community phamaceutical price. This comparison suggests that an equivalent oral dose of PPI be given as the intravenous dose (approximately

172 mg/24 hours). The steps involved in this comparison would likely involve comparing intragastric pH in normal subjects given oral and intravenous PPI, then progressing to similar comparison with patients followed finally by a large equivalence study.

H_2RA's were considered to be equivalent to placebo for this meta-analysis as no large RCT or meta-analysis had been able to show a statistically significant benefit over placebo[63,64]. When this meta-analysis is compared to the data available on H_2RA's, one very significant difference is obvious. The normal oral dose of a PPI for example is 20-40 mg per day, yet, the intravenous protocols with omeprazole utilized up to 172 mg per 24 hours. The normal oral ranitidine dose is 300 mg per 24 hours and most bleeding trials included in the meta-analysis used 150 mg per 24 hours while the large famotidine study used the standard oral dose (40 mg/24 hours) given intravenously. Therefore, it would appear that the H_2RA studies were significantly flawed compared to the PPI studies. It is impossible to predict what the outcome of the H_2RA studies would have been if they had administered a dose equivalent to eight times the normal oral dose rather than half or equivalent iv to oral dose. In addition, the H_2RA's were almost exclusively delivered in a bolus fashion. Given that the PPI infusion showed a significant reduction in rebleeding and bolus PPI did not, perhaps the same would be true of the H_2RA's. Given the current standing of PPI in the treatment of the bleeding peptic ulcer, we are unlikely to see

any further H_2RA research.

6.0 References

1. Mantel,N & Haenszel,W. 1959. Statistical aspects of the analysis of data from retrospective studies of disease. Journal of the National Cancer Institute, 22, 719-748.

2. Glass, GV. 1976. Primary, secondary and meta-analysis of research. Educational Researcher 5: 3-8.

3. Wittes, RE. 1987. Problems in the Medical interpretation of overviews. Statistics in Medicine 6: 269-276.

4. Shojania, KG & Bero LA. 2001. Taking advantage of the explosion of systematic reviews: An efficient MEDLINE search strategy. Effective Clinical Practice 4(4): 157-62.

5. Barkun A, Bardou M & Marshall JK. 2003. Consensus recommendations for managing patients with nonvariceal upper gastrointestinal bleeding. Annals of Int Med. 139(10): 843-857.

6. Counsell C. 1997. Formulating Questions and locating primary studies for inclusion in systematic reviews. Annalsof Int Med. 127(5): 380-7.

7. Stedmans Medical Dictionary. Hensyl W (ed). 1990.

8. Pederson et al. 1994. Randomised ttrial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandanavian simvastatin survival study (4S). Lancet 344: 1383-9.

9. Law MR, Wald NJ & Rudnicka AR. 2003. Quantifying the effect of statins on

low density lipoprotein cholesterol, ischemic heart disease and sttroke: systematic review and meta-analysis. BMJ326: 1423-27.

10. Montefiori DC, Metch B, McElraith MJ, Self s, Weinhold, KJ, Corey L HIV Vaccine Trials Network. 2004. Demographic factors that influence the neutralizing antibody response in recipients of recombinant HIV-1 gp120 vaccines. J Inf Dis 190: 1962-69.

11. Thompson SG, Smith TC & Sharp S. 1997. Investigating underlying risk as a source of heterogeneity in meta-analysis. Satats in Med 16: 2741-58.

12. Fardy J & Laupacis A. 1994. A meta-analysis of prophylactic endoscopic sclerotherapy for esophageal varices. Am. J Gastroenterology 89: 1938-48.

13. Abuissa H, Jones PG, Marso SP & Craig JC. 2005. Angiotensin converting enzyme inhibitors or angiotensin receptor blockers for the prevention of type 2 diabetes: a mete-analysis of randomized controlled trials. J Am Coll Cardiol 46(5): 821-6.

14. Healey JS, Baranchuk A, Crystal E, Monillo CA, Garfincle M, Yusuf S and Connolly SJ. 2005. Prevention of atrial fibrillation with angiotensin converting enzyme inhibitors and angiotensin receptor blockers: a meta-analysis. J Am Coll Cardiol 45(11) 1832-9.

15. Doulton TW, He FG and MacGregor GA. 2005. Systematic review of combined angiotensin converting enzyme inhibition and angiotensin receptor blockade in hypertension. Hypertension 45(5): 880-6.

16. Hillon P, Leberc D, Munoz C, Jungers M, Goldfarb G & Benhamou JP. 1982. Comparison of the effects of cardioselective and non-selective beta blockers

on portal hypertension in patients with cirrhosis. Hepatology 2(5) 528-31.

17. Xiao LF & Sutherland LR. 2002. Assessing disease activity and disease activity indices for inflammatory bowel disease. Curr Gastrol Rep 4(6): 490-6.

18. Thompson SG. 1994. Systematic review: Why sources of heterogeneity should be investigated. BMJ 1009: 1351-1355.

19. Mulrow C, Langhorne P & Grimshaw J. 1997. Integrating heterogeneous pieces of evidence in systematic reviews. Ann Intern Med 127:989-95.

20. Meier P. 1987. Commentary. Stats in Med. 6: 329-31.

21. Easterbrook PJ, Berlin JA, Gopalan R &Mathews DR. 1991. Publication bias in clinical reearch. Lancet 337: 867-72.

22. Spector TD &Thompson SG. 1991. The potential and limitations of meta-analysis. J. Epidemiol Community Health. 45(2): 89-92.

23. Eggar M & Snith GD. 1998. Meta-analysis: Bias in location and selection of studies. BMJ 316: 61-6.

24. Leizorovicz A, Haugh MC, Chaupis FR, Samama MM, &Boissel JP. 1992. Low molecular weight heparin in prevention of perioperative thrombosis. BMJ 305: 913-20.

25. Nurmohamed MT, Rosendaal FR, Bueller HR, Decker E, Hommes DW et al. 1992. Low molecular weight heparin versus standard heparin in general and orthopedic surgery: a meta-analysis Lancet 340: 152-6.

26. Ioannidis JPA, Cappelleri JC and Lau J. 1998. Issues in comparisons between meta-analysis and large trials. JAMA 279(14): 1089-94.

27. Gregoire G, Derderian F & LeLorier J. 1995. Selecting the Language of the publications included in a meta-analysis: is there a Tower of Babel bias? J Clin Epidemiol 48: 159-63.

28. Singh R, Singh S. 1994. Research and doctors. Lancet 344: 546

29. Eggar M, Davey Smith G, Schneider M, Mindor CE. Bias in meta-analysis detected by a simple graphical test. BMJ 315: 629-34.

30. Duval, S5 and Tweedie RL. 2000. A simple funnel plot-based method of testing and adjusting for publication bias in meta-analysis Biometrics 56:276-84.

31. Sutton AJ, Duval SJ, Tweedie RL, Abrams KR and Jones DR. 2000. Empirical assessment of the effect of publication bias on meta-analysis. BMJ 320: 1574-7.

32. Stuck AE, Rubenstein L2 and Wieland. Letter: Bias in meta-analysis detected by a simple graphical test. BMJ 316:469.

33. Longstreth, GF. 1995. Epidemiology of hospitalization for acute upper gastrointestinal hemorrhage: A population based study. Am J Gastroenterol. 206: 210-15.

34. Rockall, TA, Logan, RFA, Devlin, HB and Northfield, TC. 1995. Incidence of and mortality from acute upper gastrointestinal hemorrhage in the United Kingdom. BMJ; 311(222): 331-45.

35. Sleisenger and Fordtran. Gastrointestinal and Liver Disease. Ed. Feldman, M. 1998. W.B. Saunders and Company, 6th Ed. Vol. 1.

36. Brunner, G and Chang, J. 1995. Intravenous therapy with high doses of Ranitidine and Omeprazole in critically ill patients with bleeding peptic ulcerations of the upper intestinal tract: An open randomized control trial. Digestion 45: 217-25.

37. Villaneuva, C, Balanzo, J, Torras, T, Dainz, S, Soriano, G, Gonzalex, D, Vilardell, F. 1995. Omeprazole versus Ranitidine as adjunct therapy to Endoscopic injection in actively bleeding ulcers: a prospective and randomized study. Endoscopy 27: 308-312.

38. Khuroo,MS, Yattoo, GN, Javid, G., Khan, BA, Gulzar, GM, Sodi, JS. 1997. A comparison of Omeprazole and placebo for bleeding peptic ulcer. NEJM 336(15): 1054-58.

39. Schaffalitzky, MOB, et al. 1997. Effect of Omeprazole on the outcome of endoscopically treated bleeding peptic ulcers. Randomized double-blind placebo-controlled multi-centre study. Scand J Gastro 32(4); 320-7.

40. Lin, HJ, Lo, WC, Lee, FY, Perng, CL, and Tseng, GY. 1998. A prospective randomized comparative trial showing that Omeprazole prevents rebleeding in patients with bleeding peptic ulcer after successful endoscopic therapy. Arch Intern Med 158(1): 54-8.

41. Lanas, A et al. 1995. Effect of parenteral Omeprazole and Ranitidine on gastric pH and the outcome of bleeding peptic ulcer. Clin Gastro 21(2): 103-6.

42. Hasselgren, G et al. 1997. Continuous intravenous infusion of Omeprazole in elderly patients with peptic ulcer bleeding. Results of a placebo-controlled multi-centre study. Sacnd J Gastro; 32(4): 328-33.

43. Daneshmend, TK et al. 1992. Omeprazole versus placebo for acute upper gastrointestinal bleeding: Randomized double-blind controlled trial. BMJ 302:143-7.

44. Cardi M el. 1997. Intravenous Omeprazole versus Ranitidine in the treatment of hemorrhagic duodenal ulcer: A prospective randomized study. Ann Chir 51(2): 136-9.

45. Flores, PR et al. 1994. The treatment of upper digestive hemorrhage of peptic origin: Intravenous Ranitidine versus Omeprazole. Rev Esp Enferm Dig 86(3): 637-41.

46. Prassler, R et al. 1995. Is medication with proton pump inhibitors more effective than H_2 rceptor antagonists after endoscopic hemostasis of ulcer bleeding. Z Gastro 33: 431-4.

47. Michel, P et al. 1994. Lansoprazole versus Ranitidine for the prevention of bleeding from peptic ulcer disease. Gastroenteral Clin Bio 18: 1102-5.

48. Uribarrena, R et al. 1994. Omeprazole versus Cimetidine in the treatment of upper GI hemorrhage. Rev Esp Enf Digest 86(6): 879-883.

49. Orti, E et al. 1995. Bleeding evolution influenced by the used antisecretory? Rev Esp Enf Digest 87: 427-30.

50. Lau JY, Sung JJ, Lee KK, et al. 2000. Effect of intravenous omeprazole on

recurrent bleeding after endoscopic treatment of bleeding peptic ulcers. NEJM. Aug3;343(5): 310-6.

51. Katschinski B, Logan R, Davies J, Faulkner G, Pearson J and Langman M. 1994. Prognostic factors in upper gastrointestinal bleeding. Dig Dis & Sciences 39(4): 706-12.

52. Saeed, ZA et. al. 1993. A scoring system to predict rebleeding after endoscopic therapy of nonvariceal upper gastrointestinal hemorrhage, with a comparison of heat probe and ethanol injection. Am J Gastro 88(11): 1842-49.

53. Saeed, ZA et. al. 1995. Prospective validation of the Baylor bleeding score for predicting the likelihood of rebleeding after endoscopic hemostasis of peptic ulcers. Gastrointestinal Endoscopy, 41(6): 561-5.

54. Jaramillo, JL, Galvez C, Carmona C et. al. 1994. Prectiction of further hemorrhage in bleeding peptic ulcer. AM J Gastr 89:2135-8.

55. Laine, L and Peterson, WL. Bleeding peptic ulcer. NEJM 1994; 331(11):717-26.

56. Cook, DJ, Guyatt GH, salena BJ et. al. 1992. Endoscopic therapy for acute nonvariceal upper gastrointestinal hemorrhage: A meta-analysis. Gastro102(1): 139-48.

57. Naveau S, Barotto E, Giraud V and Chaput JC. 1991. Endoscopic hemostasis for gastroduodenal ulcer. Meta-analysis of randomized clinical trials. Gastro Clin Biol 15: 580-7.

58. Vreeberg EM, Snel P, de Bruijne JW, Bartelsman JF, Rauws EA and Tystat GN. 1997. Acute upper gatrointestinal bleeding in the Amsterdam area: incidence, diagnosis and clinical outcome. AJ Gastro 92: 236-43.

59. Green FW, Kaplan MM, Curtis LE and Levine PH. 1978. Effect of acid and pepsin on blood coagulation and platelet aggregation. Gastro 74: 38-43.

60. Patchett SE, Enright H, Afdhal N et. al. 1989. Clot lysis by gastric juice: an in vitro study. Gut 30: 1704-7.

61. Consensus Development Panel. Consensus statement on endoscopy and bleeding ulcers. Gastrointestinal Endoscopy 1990; 36 suppl.4: S62-65.

62. Collins, R and Lansman M. 1985. Treatment with histamine H_2 antagonists in upper gastrointestinal hemorrhage: Meta-analysis. NEJM 313: 660-6.

63. Walt RP, Cottrell, Mann SO et. al. 1992. Continuous intravenous Famotidine for hemorrhage from peptic ulcer. Lancet 340: 1058-62.

64. Chalmers TC, Smith H, Blackburn B et. al. 1981. A method for assessing the quality of a randomized controlled trial. Controlled Clin Trials;2: 31-49.

65. Brunner G. Luna P and Thiesemann C. 1995. Drugs for pH control in upper gastrointestinal bleeding. Aliment Pharmacol Ther 9 Supp 1: 47-50.

66. Mohamed SA. Omeprazole versus histamine H_2 receptor antagonists in the treatment of acute upper non-variceal bleeding. Hepatogastroenterology 1996: Jul-Aug;43(10):863-5.

67. Imperiale, TF and Birgisson, S. 1997. Somatostatin or octreotide compared

with H_2 antagonist and placebo in the management of acute non variceal upper gastrointestinal hemorrhage: a meta-analysis. Ann Intern Med 127(12): 1062 - 71.

68. Netzer P; Gaia C; Sandoz M; Huluk T; Gut A; Halter F; Husler J; Inauen W. 1999. Effect of repeated injection and continuous infusion of omeprazole and ranitidine on intragastric pH over 72 hours. Am J Gastroenterol, 94(2):351-7.

Statistical Methods Programmed in MetaView

Version 4

Jon Deeks

on Behalf of the Statistical Methods Working Group

of the Cochrane Collaboration

December 1999

Data structure

Consider a meta-analysis of k studies. When the studies have a binary outcome the results of each study can be presented in a 2x2 table (Table 1) giving the numbers of subjects who do or do not experience the event in each of the two groups (here called intervention and control).

Table 1 Binary data

Study i	Event	No event	Total
Intervention	a_i	b_i	n_{1i}
Control	c_i	d_i	n_{2i}

If the outcome is a continuous measure, the number of subjects in each of the two groups, their mean response and the standard deviation of their responses are required to perform meta-analysis (Table 2).

Table 2 Continuous data

Study i	Group size	Mean response	Standard deviation
Intervention	n_{1i}	m_{1i}	sd_{1i}
Control	n_{2i}	m_{2i}	sd_{2i}

Formulae

Individual Study Responses: Binary outcomes

For study i denote the cell counts as in Table 1, and let $n_{1i} = a_i + b_i$, $n_{2i} = c_i + d_i$, and $N_i = n_{1i} + n_{2i}$. For the Peto method the individual odds ratios are given by

$$\hat{OR}_i = \exp\{(a_i - E[a_i])/v_i\}$$

with its logarithm having standard error

$$se\{\ln(\hat{OR}_i)\} = \sqrt{1/v_i}$$

where $E[a_i] = n_{1i}(a_i + c_i)/N_i$ (the expected number of events in the intervention group) and

$v_i = [n_{1i}n_{2i}(a_i + c_i)(b_i + d_i)]/[N_i^2(N_i - 1)]$ (the hypergeometric variance of a_i).

For other methods of combining trials, the odds ratio for each study is given by

$$\hat{OR}_i = a_i d_i / b_i c_i$$

the standard error of the log odds ratio being

$$se\{\ln(\hat{OR}_i)\} = \sqrt{1/a_i + 1/b_i + 1/c_i + 1/d_i}$$

The risk ratio for each study is given by

$$\hat{RR}_i = (a_i / n_{1i})/(c_i / n_{2i})$$

the standard error of the log risk ratio being

$$se\{\ln(\hat{RR}_i)\} = \sqrt{1/a_i + 1/c_i - 1/n_{1i} - 1/n_{2i}}$$

The risk difference for each study is given by

$\hat{RD}_i = (a_i / n_{1i}) - (c_i / n_{2i})$ with standard error $se(\hat{RD}_i) = \sqrt{a_i b_i / n_{1i}^3 + c_i d_i / n_{2i}^3}$

Where zero cells cause problems with computation of effects or standard errors, 0.5 is added to all cells (a_i, b_i, c_i, d_i) for that study, except when $a_i = c_i = 0$ or $b_i = d_i = 0$, when the relative effect measures $\hat{OR}_i$ and $\hat{RR}_i$ are undefined.

Individual Study Responses: Continuous outcomes

Denote the number of subjects, mean and standard deviation as in Table 2, and let

$$N_i = n_{1i} + n_{2i}$$

and

$$s_i = \sqrt{((n_{1i} - 1)sd_{1i}^2 + (n_{2i} - 1)sd_{2i}^2)/(N_i - 2)}$$

be the pooled standard deviation of the two groups. The weighted mean difference is given by

$\hat{WMD} = m_{1i} - m_{2i}$ with standard error $se(\hat{WMD}) = \sqrt{sd_{1i}^2 / n_{1i} + sd_{2i}^2 / n_{2i}}$

There are several popular formulations of the standardised mean difference. The one implemented in MetaView is Hedges adjusted g, which is very similar to Cohen's d, but includes an adjustment for small sample bias

$\hat{g}_i = ((m_{1i} - m_{2i})/s_i)(1 - 3/(4N_i - 9))$ with standard error $se(\hat{g}_i) = \sqrt{N_i/(n_{1i}n_{2i}) + \hat{g}_i^2/(2(N_i - 3.94))}$.

Pooling Methods

Mantel-Haenszel Methods for Combining Trials

For each study, the effect size from each trial $\hat{\Theta}_i$ is given weight w_i in the analysis. The overall estimate of the pooled effect, $\hat{\Theta}_{MH}$ is given by

$$\hat{\Theta}_{MH} = \frac{\sum w_i \hat{\Theta}_i}{\sum w_i}$$

For combining odds ratios, each study's OR is given weight

$w_i = b_i c_i / N_i$,

and the logarithm of $\hat{OR}_{MH}$ has standard error given by

$$se\{\ln(\hat{OR}_{MH})\} = \sqrt{((PR)/R^2 + ((PS+QR)/(R \times S)) + (QS)/S^2)/2}$$

where

$R = \sum a_i d_i / N_i$; $S = \sum b_i c_i / N_i$;

$PR = \sum (a_i + d_i) a_i d_i / N_i^2$; $PS = \sum (a_i + d_i) b_i c_i / N_i^2$;

$QR = \sum (b_i + c_i) a_i d_i / N_i^2$; $QS = \sum (b_i + c_i) b_i c_i / N_i^2$

For combining risk ratios, each study's RR is given weight

$w_i = [c_i(a_i + b_i)]/N_i$,

and the logarithm of $\hat{RR}_{MH}$ has standard error given by

$$se\{\ln(\hat{RR}_{MH})\} = \sqrt{P/(R \times S)}$$

where

$P = \sum (n_{1i} n_{2i} (a_i + c_i) - a_i c_i N_i)/N_i^2$; $R = \sum a_i n_{2i} / N_i$; $S = \sum c_i n_{1i} / N_i$

For risk differences, each study's RD has the weight

$w_i = n_{1i} n_{2i} / N_i$

and $\hat{RD}_{MH}$ has standard error given by

$$se\{\hat{RD}_{MH}\} = \sqrt{(P/Q^2)}$$

where

$P = \sum (a_i b_i n_{2i}^3 + c_i d_i n_{1i}^3)/(n_{1i} n_{2i} N_i^2)$; $Q = \sum n_{1i} n_{2i} / N_i$

The heterogeneity statistic is given by

$$Q = \sum w_i' (\hat{\Theta}_i - \hat{\Theta}_{MH})^2$$

where $\hat{\Theta}$ is the log odds ratio, log relative risk or risk difference and the w_i' are the weights calculated as $1/se(\hat{\Theta}_i)^2$. Under the null hypothesis that there are no differences in treatment effect between trials this follows a chi-squared distribution on $k-1$ degrees of freedom (where k is the number of studies contributing to the meta-analysis).

Inverse Variance Methods for Combining Trials

Inverse variance methods are used to pool both standardised mean differences, and weighted mean differences for continuous data. In the general formula the effect size is defined to be $\hat{\Theta}_i$ which is the trials SMD or WMD. The individual effect sizes are weighted according to the reciprocal of their variance (calculated as the square of the standard error given in the individual study section above) giving

$$w_i = 1/se(\hat{\Theta}_i)^2$$

These are combined to give a pooled estimate

$$\hat{\Theta}_{IV} = \frac{\sum w_i \hat{\Theta}_i}{\sum w_i}$$

with

$$se\{\hat{\Theta}_{IV}\} = 1/\sqrt{\sum w_i}$$

The heterogeneity statistic is given by a similar formula as for the Mantel-Haesznel method, using the inverse variance form of the weights, w_i

$$Q = \sum w_i(\hat{\Theta}_i - \hat{\Theta}_{IV})^2 .$$

Peto's Assumption Free Method for Combining Trials

Here, the overall odds ratio is given by

$$\hat{OR}_{Peto} = \exp\{\sum w_i \ln(\hat{OR}_i)/\sum w_i\},$$

where the odds ratio $\hat{OR}_i$ is calculated using the approximate method described in the individual trial section, and the weights, w_i are equal to the hypergeometric variances, v_i.

The logarithm of the odds ratio has standard error

$$se\{\ln(\hat{OR}_{Peto})\} = 1/\sqrt{\sum v_i}$$

The heterogeneity statistic is given by

$$Q = \sum v_i\{(\ln \hat{OR}_i)^2 - (\ln \hat{OR}_{Peto})^2\}.$$

DerSimonian and Laird Random Effects Models

Under the random effects model, the assumption of a common treatment effect is relaxed, and the effect sizes are assumed to have a distribution

$\Theta_i \approx N(\Theta, \tau^2)$.

The estimate of τ^2 is given by

$\hat{\tau}^2 = \max\{[Q-(k-1)]/[\sum w_i - (\sum(w_i^2))/\sum w_i], 0\}$, where the w_i are the inverse variance weights (calculated as $1/se(\hat{\Theta}_i)^2$) for log OR, log RR, RD, WMD and SMD, as appropriate.

The estimate of the combined effect for the heterogeneity may be taken as either the Mantel-Haenszel or the inverse variance estimate. Again, for odds ratios and risk ratios, the effect size is taken as the natural logarithm of the OR and RR. Each study's effect size is given weight

$w_i' = 1/(se(\hat{\Theta}_i)^2 + \hat{\tau}^2)$

The pooled effect size is given by

$\hat{\Theta}_{DL} = (\sum w_i' \hat{\Theta}_i)/(\sum w_i')$

and

$se\{\hat{\Theta}_{DL}\} = 1/\sqrt{\sum w_i'}$

Note that in the case where the heterogeneity statistic Q is less than or equal to its degrees of freedom $(k-1)$, the estimate of the between trial variation, $\hat{\tau}^2$, is zero, and the weights reduce to the those as given by the inverse variance method.

Confidence intervals

The $100(1-\alpha)\%$ confidence interval for $\hat{\Theta}$ is given by

$\hat{\Theta} - se(\hat{\Theta})\Phi(1-\alpha/2)$, to $\hat{\Theta} + se(\hat{\Theta})\Phi(1-\alpha/2)$

where $\hat{\Theta}$ is the log odds ratio, log relative risk, risk difference,mean difference or standardised mean difference, and Φ is the standard normal deviate.

Test statistics

In all cases, the test statistic is given by

$z = \hat{\Theta}/se(\hat{\Theta})$

where the odds ratio or risk ratio is again considered on the log scale.

Bibliography

Breslow NE, Day NE. Combination of results from a series of 2x2 tables; control of confounding. \ In: Statistical Methods in Cancer Research, Volume 1: The analysis of case-control data. IARC Scientific Publications No.32. Lyon: International Agency for Health Research on Cancer, 1980.

Deeks JJ, Altman DG, Bradburn MJ. Statistical methods for examining heterogeneity and combining results from several studies in a meta-analysis. In: Egger M, Davey Smith G, Altman DG. Systematic Reviewes and Healthcare: meta-analysis in context. BMJ Publications (in press).

DerSimonian R, Laird N. Meta-analysis in clinical trials. Controlled Clinical Trials 1986; 7: 177-188.

Greenland S, Robins J. Estimation of a common effect parameter from sparse follow-up data. Biometrics 1985;41: 55-68.

Greenland S, Salvan A. Bias in the one-step method for pooling study results. Statistics in Medicine 1990; 9:247-252.

Hedges LV, Olkin I. Statistical Methods for Meta-analysis. San Diego: Academic Press 1985. Chapter 5.

Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. Journal of the National Cancer Institute 1959;22: 719-748.

Robins J, Greenland S, Breslow NE. A general estimator for the variance of the Mantel-Haenszel odds ratio. American Journal of Epidemiolgy 1986; 124:719-723.

Rosenthal R. Parametric measures of effect size. In: Cooper H, Hedges LV (eds.). The Handbook of Research Synthesis. New York: Russell Sage Foundation, 1994.

Sinclair JC, Bracken MB. Effective Care of the Newborn infant.Oxford: Oxford University Press 1992.Chapter 2.

Yusuf S, Peto R, Lewis J, Collins R, Sleight P. Beta blockade during and after myocardial infarction: an overview of the randomized trials. Progress in Cardiovascular Diseases 1985;27:335-371.