INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Université d'Ottawa · University of Ottawa

DOES ACUTE NORMOVOLEMIC HEMODILUTION REDUCE PERIOPERATIVE ALLOGENEIC BLOOD EXPOSURE? A SYSTEMATIC REVIEW

BY

© GREGORY L. BRYSON M.D., FRCP(C), B.Sc. (Medicine)

A thesis submitted to the

School of Graduate Studies and Research

in partial fulfillment of the requirements for the

M.Sc. degree in Epidemiology

University of Ottawa

December 1998

ABSTRACT

The objective of this study was to systematically review the literature and to statistically summarize the evidence evaluating the efficacy of acute normovolemic hemodilution (ANH). Review of 1575 citations identified 24 prospective, randomized, controlled trials of ANH containing a total of 1216 patients. When all trials were pooled ANH reduced the likelihood of exposure to allogeneic blood and the total units of allogeneic blood transfused. Marked heterogeneity of the results was present and adverse events were incompletely reported. Subgroup analyses could not account for all heterogeneity noted in this overview. Under reporting of small trials with negative results was suspected. In trials using a protocol to guide perioperative transfusion, ANH failed to reduce either the likelihood of transfusion or the units administered. It is possible that experimental design bias in favor of reducing transfusion in the ANH group is in part responsible for the reported efficacy of this technique.

ACKNOWLEDGEMENTS

My post-graduate education has been made possible by a network of concerned and helpful individuals whose support I would like to acknowledge.

I would first like to acknowledge the endurance and understanding of my wife, Pilar Bryson. Pilar has been a constant source of encouragement and support throughout medical school, internship, residency, and most recently, my M.Sc. training. I doubt I'd have completed any of these programs without her.

I would next like to thank my supervisors Drs. Andreas Laupacis and George Wells for their guidance throughout the design and conduct of this project. They have patiently tolerated my slow progress on the thesis while I returned to my "day job" in the Department of Anaesthesia at the Ottawa Civic Hospital.

I would also like to acknowledge a number of individuals whose contributions of time and expertise made this work possible. The comprehensive literature searches were designed and executed with the technical support of Ms. J. McGowan and Ms. J. MacLeod. Non-English language literature would have remained Greek to me without the efforts of Ms. B. Malanka and Drs. O. Rosaeg, G. Pagliarello, J. Cooper, and P. Hebert. Thanks also to Mr. Ba Pham and Ms. Liz Yetisir for their statistical assistance.

This work was conducted in collaboration with the International Study of Perioperative Transfusion (ISPOT). ISPOT members have graciously provided assistance at all points during the development and completion of this project. The International Study of Peri-operative Transfusion (ISPOT) Investigators are: Coordinating Centre and Canadian Investigators: Phil Wells, MD; George Wells, PhD; Fraser Rubens, MD; Annette O'Connor, PhD; Laura McAuley BSc; Andreas Laupacis,

MD (Chair); Jack Kitts, MD; Paul Hebert, MD; Ian Graham, PhD; Dean Fergusson, MHA (Coordinator); Greg Bryson, MD. Clinical Epidemiology Unit, Loeb Medical Research Institute, Ottawa Civic Hospital, University of Ottawa, Ottawa, Ontario, Canada. United States: Chris Hillyer, MD; Jeff Etchason, MD; Angela Blair. Kerr L. White Institute for Health Services Research, Decator, Georgia. United Kingdom: Brian McClelland, MD; Patricia Phillips. SES Scotland Transfusion Service, Edinburgh. Spain: Jordi Gol-Freixa, MD; Antonio Garcia. Agencia de evaluacion de Tecnologias Sanitarias, Madrid. Netherlands: Ankie Koopman-van Gemert, MD, PhD. Laren (NH). Japan: Akinori Hisashige, MD, PhD. University of Tokushima, Tokushima. Israel: Nachum Egoz MD, MPH; Luis Gaitini MD; Somri Mustafa MD; Edna Katz MD. Bnai Zion Medical Center, Haifa. France: Luc Noel, MD; Centre Hospitalier de Versailles, Versailles. Rachid Salmi, PhD; Charlotte Huet, MD; Université Victor Segalen Bordeaux II, Bordeaux. Denmark: Helga Sigmund, MSc. Danish Hospital Institute, Copenhagen. Australia: Katherine McGrath, MD; Kim Henderson. Hunter Area Pathology Service, Newcastle. David Henry, MD. Newcastle.

Lastly, ISPOT and its individual investigators have received funding from a number of sources that I'd like to acknowledge. Dr. Laupacis received the First Fellowship from the International Society of Technology Assessment in Health Care, funded by the PPP Medical Trust, United Kingdom in order to do this study. The Ottawa Coordinating Centre has received funding from Janssen Ortho Inc, Canada, the Australian Group from the National Health and Medical Research Council and the Hunter Area Pathology Services, the French Group from Haemonetics France, Ortho Diagnostics France and Université Victor Segalen Bordeaux II, and the group from the United States

from the Baxter Healthcare Corporation Biotech Group and the Emory Center for Clinical Evaluation Sciences.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1 Background

1.1.1. Risks Associated with Allogeneic Transfusion

Blundell first described transfusion of blood from non-genetically identical individuals in 1898 (1). In the nearly 100 years that followed, the use of allogeneic blood and blood products has become a mainstay of modern medical practice. The emergence of the human immunodeficiency virus (HIV) and the subsequent discovery of its transmission through blood transfusion focused the attention of patients, physicians and governments on the safe use of donor blood.

A brief summary of the complications associated with allogeneic transfusion is found in Table 1.

Table 1. Complications Associated with Blood Transfusion.

Complication	Etiology
Viral contamination	HIV, hepatitis B,C, and G, HTLV, CMV
Bacterial contamination	Y. enterocolitica
Parasitic contamination	Malaria, Chagas disease
Creutzfeldt-Jakob disease	Prion
Acute hemolytic reaction	ABO incompatibility
Delayed hemolytic reaction	Minor blood group incompatibility
Febrile, non-hemolytic, reaction	Recipient antibodies to donor WBC or platelets
Anaphylaxis	Complement activation
Urticaria	Recipient antibodies to donor plasma
Acute lung injury	Complement-mediated pulmonary edema
Immunosupression	Donor lymphocytes depress recipient immune function
Graft-versus-host disease	Immunocompetent donor lymphocytes engraft in recipient
Post-transfusion purpura	Recipient develops antibodies against donor and recipient platelets
Passive alloimmune thrombocytopenia	Donor blood contains platelet-specific alloantibodies to recipient platelets

Adapted from reference (2)

A detailed description of all complications associated with allogeneic transfusion is beyond the scope of this document. However, elaboration on several of these topics will place the efforts to reduce the use of allogeneic blood in context.

Of the complications associated with allogeneic transfusion, viral contamination has received the majority of public attention. In 1996 the Retrovirus Epidemiology Donor Study (REDS) provided the following estimates of the likelihood of an infected unit of allogeneic blood avoiding detection: HIV 1 in 676,000; Human T-lymphotrophic viruses (I and II) 1 in 641,000; hepatitis B virus 1 in 63,000; and hepatitis C virus 1 in 103,000 (3). Despite public concern the risks of infection following allogeneic transfusion are remarkably low. By means of comparison the likelihood of death during childbirth in Canada in 1994 was 1 in 12 463 (4). Unfortunately, the emergence of new infectious entities, such as hepatitis G virus and Creutzfeldt-Jacob disease, will ensure that the blood supply is never entirely safe.

Acute hemolytic transfusion reactions are estimated to occur with 1 in 25,000 transfusions and are fatal in 1 in 600,000 transfusions. The less severe delayed reactions follow 1 in 2,500 transfusions (2). The vast majority of these transfusion reactions result from clerical errors. Human errors of this kind complicate 1 in every 12,500 transfusions (5). Human error therefore represents an equally serious, and more prevalent, threat to the transfusion recipient as the transfusion-related viral infections discussed earlier.

The influence of white blood cells administered with allogeneic blood transfusion on the function of the recipient's immune system has recently been the focus of considerable attention. Donor white blood cells have been shown to decrease natural killer cell activity and increase T-suppressor cell activity in those patients receiving

allogeneic blood products (6). The importance of the resultant decrease in recipient immune function has been of some debate. A recent systematic overview and meta-analysis suggested that perioperative allogeneic transfusion does not increase the risk of cancer recurrence following surgery (7). On the other hand, use of white blood cell depleted allogeneic blood has been shown to reduce 60-day mortality following aortocoronary bypass (8) and postoperative infectious complications following colorectal surgery (9).

While transfusion-related HIV and hepatitis are now relatively rare events, emerging or unknown infections, transfusion errors, and immunomodulation remain compelling reasons to minimize the use of allogeneic transfusion.

1.1.2. Alternatives to Allogeneic Transfusion

In 1993 Justice Horace Krever began a judicial inquiry into the management and safety of Canada's blood supply. In 1997 the final findings and recommendations of this inquiry were released. Among Krever's recommendations was "...that the operator of the blood supply system promote appropriate use of, and alternatives to, blood components and blood products." Krever further noted that "the need for transfusion of blood components is being reduced in some cases by the use of medical techniques that allow the patient's blood to be conserved or reused during surgery (10)."

The preoperative donation of one's own (autologous) blood is the most commonly used means of reducing allogeneic blood transfusion in Canada. In 1995-96 Canadians donated 20, 495 units of autologous blood in the weeks preceding their planned surgery (11). Preoperative autologous donation (PAD) reduced the likelihood of allogeneic blood exposure in a systematic review (12) of 6 randomized controlled trials (Odds Ratio (OR)

0.17, 95% Confidence Interval (CI) 0.08 to 0.32) and 9 cohort studies (OR 0.19, 95% CI 0.14 to 0.26). An interesting finding of this overview was that the anemia resulting from preoperative phlebotomy led to a marked increase in the number of transfusions of either autologous or allogeneic blood in randomized trials (OR 3.03, 95% CI 1.70 to 5.39) and cohort studies (OR 12.32, 95% CI 5.90 to 25.40). This increase in transfusion likely resulted from lower preoperative hematocrits in the PAD group and a more liberal transfusion policy regarding autologous blood.

While undeniably effective, and popular with both surgeons and their patients, PAD programs have recently come under scrutiny. Many patients participating in PAD programs do not lose sufficient blood during their surgery to require any transfusion. As a result many autologous units are discarded at the end of the patient's hospital stay. Between 25% and 45% of autologous units donated in Forgie's review were eventually discarded (12). The frequent waste of units donated in PAD programs leads to autologous blood being more costly than allogeneic blood. Tretiak (13) estimated that the cost of a unit of autologous blood (CDN $ 328.00) donated in Canada was approximately 50% greater than a unit of allogeneic blood (CDN $ 210.00).

The increased cost and the relative rarity of serious adverse outcomes associated with allogeneic transfusion render PAD a relatively cost-ineffective technology. Etchason estimated that for a patient undergoing a total hip arthroplasty, participation in a PAD program cost US$ 235,000.00 per quality adjusted year of life (QALY) saved through avoidance of transfusion-related viral disease cost (14). Considering the US$ 8000.00 per QALY cost associated with performing the same hip replacement it is clear that PAD is a relatively inefficient use of health care dollars (15). As health resources

shrink, the utility of the widespread use of PAD programs must be questioned and less expensive alternatives explored.

Blood lost during surgery may be salvaged from the operative field and returned to the patient. A variety of cell salvage techniques exist and vary in complexity and the volume of blood that can be captured and reinfused. Cell salvage offers several advantages over PAD "in that it is logistically easier to organize, it is not affected by cancellation of operation, and it is applicable to emergency cases. (16)". Unfortunately, cell salvage may not be used in surgical fields contaminated with bacteria or malignant cells and may be associated with destruction of salvaged cells and subsequent coagulopathy. Cell salvage has yet to undergo a formal cost-effectiveness study but is associated with significant fixed costs. By way of example, the cost of the disposable components of Haemonetics 5 cell salvage system used in operating rooms of the Ottawa Civic Hospital is in excess of CDN $185.00 per patient. The device itself costs approximately CDN $35,000.00. If a hospital regularly performs surgical procedures associated with significant blood loss the rather substantial costs of this technology may be justified.

The use of a variety of pharmacological agents to reduce perioperative blood loss and resultant allogeneic transfusion has also been evaluated. Aprotinin is a protease inhibitor that reduces the degradation of blood clot and preserves platelet function. Epsilon-aminocaproic acid and tranexamic acid directly inhibit the enzyme plasmin from binding to and dissolving fibrin, the protein responsible for blood clotting. Desmopressin acetate reduces bleeding by releasing factor VIII:C and von Willebrand factor encouraging the adherence of platelets, the cellular element of blood clots, to the wall of

injured blood vessels. Recent systematic reviews of these agents suggest that aprotinin and tranexamic acid are effective in the reduction of allogeneic transfusion but that there is insufficient evidence to support the use of desmopressin and epsilon-aminocaproic acid (17). The use of human recombinant erythropoietin, a hormone that stimulates the production of red blood cells, has also been advocated. Erythropoietin was shown to reduce the likelihood of allogeneic blood in those patients undergoing cardiac (OR 0.09, 95% CI 0.03 to 0.28) and orthopedic surgery (OR 0.38, 95% CI 0.24 to 0.63) in a recent systematic overview (18).

Clinical practice guidelines to assist the clinician in the safe use of allogeneic blood and technologies to reduce or eliminate its use have recently appeared in general medical, surgical, and anesthesia literature (2,19,20). Common to these guidelines is the recommendation that acute normovolemic hemodilution (ANH), another technology for collecting autologous blood, be considered as a means of reducing allogeneic transfusion for patients undergoing surgery.

1.1.3. Acute Normovolemic Hemodilution (ANH)

Hemodilution is defined as "dilution of all of the blood constituents resulting from the exchange of the patient's whole blood for cell-free plasma like fluid (21)." Interest in hemodilution began in the early 1960s in response to the pressures placed on a limited donor blood supply by cardiac surgical patients. Initial experience with open-heart surgery relied heavily upon donor blood and blood products both to prime the cardiopulmonary bypass machine and to treat the often substantial blood loss associated with cardiac surgical procedures. Based on their experiences with the passive hemodilution associated with bloodless cardiopulmonary bypass primes, Buckley,

Hallowell, and colleagues at the Massachusetts General Hospital began to experiment with the "temporary removal of a patients' blood ... before extracorporeal circulation (22)."

Buckley and Hallowell's technique, now commonly referred to as ANH, is performed in the operating room shortly before or during surgery. A portion of the patient's blood volume is withdrawn through an intravenous or intraarterial cannula and replaced with an appropriate volume of intravenous fluid to maintain normal circulating volume. The blood is collected in bags containing an anticoagulant and stored in the operating room for later use. The blood reserved in this process contains oxygen-carrying red blood cells as well as platelets and proteins essential for normal blood clotting. During the hemodilution process the patient's hematocrit, a measure of red blood cell volume, is lowered from a normal value of 0.45 to 0.30 or less. As surgery proceeds, blood lost in the surgical field contains fewer red blood cells and results in a reduced loss of oxygen carrying ability. At the end of surgery, or when hematocrit falls to a point where oxygen delivery is compromised, ANH units are returned to the patient, thus restoring oxygen carrying capacity.

In 1972 Hallowell and Buckley reported the first prospective, randomized, controlled trial of acute normovolemic hemodilution (ANH) (22) followed a year later by Laks's report of the first use of the technique in non-cardiac surgery (23). ANH-related reductions in allogeneic blood exposure ranging from 18 to 90% have since been reported in a variety of surgical procedures (24). ANH offers several practical advantages over other technologies for the collection of a patient's own (autologous) blood. Unlike preoperative autologous donation (PAD), which requires that patients donate blood on

several preoperative visits in the weeks preceding surgery, ANH may be performed on the day of surgery in the operating theater. The minimal preoperative preparation required minimizes patient inconvenience and makes ANH suitable for both emergency and elective procedures. Unlike perioperative cell salvage, ANH requires no special equipment and may be employed in circumstances where the surgical field may be contaminated with infected or malignant tissue. ANH units are collected and stored at the patient's bedside reducing the administrative costs associated with the collection, storage, and testing of PAD units. ANH units also contain fresh clotting factors that may be of theoretical advantage in cardiac surgery where disorders of blood clotting are relatively common. Lastly, all units of autologous blood collected through ANH are returned to the patient at the end of the surgical procedure eliminating the risk of previously donated units spoiling should elective surgery be delayed or canceled.

Given the potential advantages and lower costs associated with ANH, why then is it not used as a source of autologous blood more frequently? Only 30% of Canadian hospitals surveyed in 1995 reported employing ANH in their operating rooms to reduce allogeneic transfusion (25). Proposed factors inhibiting wider use of this technology include: ANH is time consuming and requires intensive monitoring (26); the research supporting ANH is flawed (27); ANH is incapable of yielding significant blood savings (28); and the anemia inherent to the technique places patients at undue risk of cardiovascular morbidity (29).

1.1.4. Rationale for a Systematic Review of ANH Research

It could be argued that larger, perhaps multi-center, trials will be required to clearly define the role of ANH in clinical practice. Unfortunately some advocates of

ANH suggest that "ethical and moral considerations make it unlikely that they will ever be carried out. Any decrease in the use of homologous blood would make such studies extremely difficult to justify (26)."

Clinicians, patients, and health care policy-makers need research that will "separate the known from the unknown and save them from the position of knowing less than has been proved (30)." A systematic review provides an efficient means of identifying and integrating existing research on ANH. By critically examining and pooling evidence from a number of sources one can determine if ANH yields a consistent treatment effect across different studies and patient populations. Similarly one can identify variation in treatment effect in selected subgroups. While systematic review will not provide a definitive answer, one can use the results to identify and refine hypotheses, recognize and avoid pitfalls of previous work, and identify important covariates or adverse events that warrant consideration in future studies (30). A quantitative summary of past research offers an increase in the statistical power and may allow the identification of clinically relevant savings in allogeneic blood resulting from the use of ANH. This increase in statistical power is of particular benefit when the event of interest occurs infrequently, as might be the case with an adverse event.

Given the expense and ethical concerns regarding the mounting of a definitive trial of ANH our existing knowledge should be critically evaluated and summarized before proceeding with further prospective research. Systematic review and statistical summary of previous research using meta-analytic techniques will be a timely and useful addition to the debate regarding the use of ANH.

1.2. Aims and Objectives

The aim of this study was to conduct a systematic overview of all randomized controlled trials evaluating acute normovolemic hemodilution (ANH) as a means of reducing allogeneic blood transfusion in the perioperative period. Pooling the results of previously published research using meta-analytic techniques was used to create a summary estimate of the efficacy of ANH.

The primary objective of this study was:

1. to determine the effect of ANH on the likelihood of perioperative allogeneic blood exposure.

Secondary objectives included:

2. to determine the effect of ANH on the volume of allogeneic blood transfused and surgical blood loss in the perioperative period;
3. to evaluate the efficacy of ANH in different surgical procedures;
4. to evaluate the influence of the volume of preoperative phlebotomy (V_{ANH}) and the volume of surgical blood loss (SBL) on the efficacy of ANH;
5. to determine the effect of trial methods on the efficacy of ANH;
6. to assess for the presence of publication bias in the literature evaluating ANH; and
7. to determine if language of publication influences the conclusions of the overview.

2. METHODS

2.1. Study Selection

Systematic searches of Medline (1966-96), Current Contents (1995-6), and Embase (1978-1996) were undertaken in August 1996. A broad search using MeSH headings and textwords for both hemodilution and blood transfusion was conducted with the assistance of a medical librarian. Details of the search strategies employed in this overview can be found in Appendix A.

All titles and abstracts were reviewed to determine eligibility for analysis. Any prospective, randomized trial comparing intraoperative acute normovolemic hemodilution to a concurrent control group, regardless of language or medium of publication, was retrieved for further analysis. Only those trials in which whole blood was withdrawn on the day of surgery and replaced with a crystalloid or colloid solution were considered to represent ANH. If insufficient detail was provided in the abstract to determine eligibility, the article was retrieved for further review. The bibliographies of review articles and all retrieved trials were hand searched for additional publications. Duplicate publications and trials enrolling participants less than 18 years of age or parturients were excluded.

Publications meeting the inclusion and exclusion criteria that reported either the number of subjects exposed to at least one unit of allogeneic blood or the volume of allogeneic blood transfused were considered eligible for meta-analysis.

In addition to the basic search, two "filters" designed to identify randomized controlled trials in Medline were employed to evaluate their performance in the

anaesthesia and transfusion literature (31,32). Search terms used in the McMaster and Cochrane RCT filters are also provided in Appendix A.

2.2. Data Extraction

Trials meeting all eligibility criteria were photocopied in such a way that the authors, the date of publication, and the journal of publication were obscured. Two independent reviewers using standardized data collection forms abstracted data. Disagreements were resolved by consensus of the two reviewers. Non-English language trials were abstracted by one of the reviewers with the assistance of a translator. When necessary an attempt was made to contact the primary authors by mail to clarify results or provide missing data.

The primary outcome measure was the proportion of subjects exposed to at least one unit of allogeneic blood in the perioperative period. Other data abstracted included: volume of allogeneic blood transfused in the perioperative period, surgical blood loss in the perioperative period (SBL), type of surgery, the volume of blood withdrawn during hemodilution (V_{ANH}), the reporting of a perioperative transfusion protocol, and discharge hematocrit. If the volume of blood transfused was reported in milliliters, the number of units transfused was calculated by dividing the volume by 300.

The quality of the reporting of the trial methods was assessed using Jadad's 3 item scale (33). Trials received a single point for each of the following characteristics: randomization; blinding; and accounting for all enrolled subjects. A point was added if the mechanism of randomization was described and was appropriate (e.g. random numbers table) or subtracted if inappropriate (e.g. alternate or date of birth allocation). Similarly a point was added or subtracted for appropriate or inappropriate blinding

procedures respectively. The total score ranged from 0 (worst quality) to 5 (best quality). The form used to score trial methods, including definitions of the characteristics scored, can be found in Appendix B.

2.3. Analysis

The proportion of patients transfused was analyzed using both the Mantel-Haenszel fixed effects model (34)and the DerSimonian and Laird random effects model (35) with results expressed using relative risks (RR) and 95% confidence intervals (95% CI). Continuous variables such as the total units of allogeneic blood transfused and SBL were also analyzed using both fixed and random effects models (34). Results are expressed as weighted mean differences (WMD) with 95% CI. Heterogeneity of results was evaluated using the χ^2 statistic appropriate for the given method. All of the preceding analyses were performed using RevMan 3.0.1 (36). A detailed description of the statistical methods used for meta-analysis of dichotomous and continuous outcomes is included in Appendix C.

Volume of allogeneic blood transfused in a subset of trials was also analyzed as an ordinal variable using the method described by Whitehead and Jones (37). Subjects enrolled in the ANH and control groups were placed into five ordered categories according to the units of allogeneic blood they received. Patients were classified as having received no units of allogeneic blood, one unit, two units, three units, and four or more units. Transfusion outcome was then reduced to a binary response by defining a category cut-point to reflect treatment success or failure. For example, if treatment success were defined as receiving no units of allogeneic blood a patient receiving no transfusion would be a success and one or more units a failure. The cut-point was then

adjusted to one unit, two units, and three units creating four separate binary classifications of treatment success. Logistic regression (PROC LOGISTIC in SAS) was then used to determine the odds of treatment success in any given classification. Assuming a common log odds ratio across all classifications a proportional odds model was used to determine overall odds of treatment success regardless of how the categories were divided into success or failure. The summary odds of success and failure for each study were then pooled using standard meta-analytic techniques. The methodological details of the ordinal variable analysis used are discussed in greater detail in Appendix C and the relevant SAS files are included in Appendix D.

Subgroup analyses based upon type of procedure, V_{ANH}, SBL, language of publication, Jadad score, and the use of a transfusion protocol were proposed *a priori*. To be considered to have employed a transfusion protocol a trial must have clearly stated the indications for transfusion of allogeneic blood in the Methods section and have employed equivalent triggers for transfusion in both ANH and control groups. Subgroup analyses could not be performed using the weighted chi square because of an insufficient number of trials.

Evidence of publication bias was sought by using a "funnel plot", a graphical means of comparing effect sizes and sample sizes of trials in a meta-analysis (38). The log odds ratio is presented on the x-axis and sample size is plotted on the y-axis. Sampling theory suggests that estimates of effect from small trials should scatter widely on both sides of the estimate of true effect. As sample size increases variability in the estimates should decrease along with the scatter of the estimates to form a cone or funnel shaped distribution. Publication bias was suspected if the funnel plot was asymmetrical

or lacking data points in the portion of the funnel representing negative or equivocal studies.

To further characterize publication bias the natural log of standardized normal deviate (SND = effect size ÷ standard error) was regressed against precision (precision = standard error^{-1}). This regression technique provided an objective statistical test of funnel plot asymmetry (39). Estimates of effect from small trials should be associated with a large standard error and therefore small values for both the standardized normal deviate and precision. As standard error decreases with increasing sample size values for both SND and precision should both increase. When this information is plotted with lnSND on the y-axis and precision on the x-axis a line joining the data points should run through the origin in the absence of publication bias. If only small trials with positive results are reported the regression line passes through the x-axis and contacts the y-axis below the origin. When regressed using the formula ($\ln SND = \alpha + \beta precision$) the magnitude of α represents the degree of asymmetry and its p value the statistical significance of the estimate. P values <0.1 were considered to represent statistically significant asymmetry.

3. RESULTS

3.1. Search Strategies

Numbers of citations identified from each computerized database were as follows: Medline 880; Embase 758; and Current Contents 91. Following removal of references duplicated in the different searches, a total of 1568 unique references were identified and reviewed. 164 references were retrieved for detailed assessment, of which 21 were described as prospective randomized controlled trials (RCTs). Of the 21 RCTs identified from the computerized literature searches Medline provided 18 citations (40-57), Current Contents two (58,59) and Embase one (60).

Manual search of the bibliographies of retrieved publications yielded an additional 7 RCTs. Citations identified by manual review were sought in Medline by author name search. All seven references identified by manual searches were indexed in Medline, but under terms excluded by our search strategy. Five were indexed under the MeSH heading "blood transfusion, autologous" and did not contain the textword hemodilution (22,61-64). The screen for Mesh headings and textwords referring to blood transfusion eliminated the other two publications missed by the Medline search (65,66). The single reference identified using Embase (60) was also identified in Medline under "blood transfusion, autologous". In total, 1575 citations were identified from computerized and manual searches yielding 28 publications described as RCTs. The publication characteristics of the 28 identified trials are summarized in Table 1.

The RCT filters were successful in reducing the volume of references gathered by Medline searches. Of the 880 citations identified by Medline the use of the Cochrane and McMaster filters reduced the number of trials requiring review to 208 and 223

respectively. Reliance on the Cochrane filter would have resulted in 3 of the 18 publications identified through Medline being missed (46,49,52). Similarly the McMaster filter would have overlooked 2 publications (46,56).

Table 2. Publication Characteristics of Identified Trials

Author	Year	Source	Format	Language	Cochrane	McMaster
Ahlberg (40)	1977	MDL	Full	English	+	+
Atallah (41)	1993	MDL	Full	English	+	+
Bennett (42)	1994	MDL	Full	English	+	+
Boldt (43)	1991	MDL	Full	English	+	+
Bonnet (44)	1986	MDL	Full	French	+	+
Dietrich (45)	1989	MDL	Full	English	+	+
Lawson (46)	1974	MDL	Full	English		
Lilleaasen (47)	1977	MDL	Full	English	+	+
Lorentz (48)	1991	MDL	Full	German	+	+
Malinovsky (49)	1989	MDL	Abst	French		+
Moyes (50)	1985	MDL	Full	English	+	+
Rose (52)	1981	MDL	Full	English		+
Rosberg (51)	1981	MDL	Full	English	+	+
Triluzi (53)	1995	MDL	Full	English	+	+
van der Linden (54)	1994	MDL	Full	English	+	+
Vara Thorbeck (55)	1990	MDL	Full	French	+	+
von Bormann (56)	1986	MDL	Full	German	+	
Welch (57)	1993	MDL	Full	English	+	+
Herregods (58)	1995	CC	Full	English		
Kochamba (59)	1996	CC	Full	English		
Ferraris (60)	1993	EMBASE	Full	English		
Catoire (65)	1992	Manual	Full	English		
Dale (66)	1987	Manual	Full	English		
Hallowell (22)	1972	Manual	Full	English		
Kaplan (61)	1977	Manual	Full	English		
Ochsner (62)	1973	Manual	Full	English		
Pliam (63)	1975	Manual	Full	English		
Vedrinne (64)	1992	Manual	Full	English		

Author (reference number)
MDL = MEDLINE
CC = Current Contents
Full = Full Manuscript
Abst = Abstract

3.2. Data Extraction

Of the 28 publications described as RCTs, 3 used a "pseudo-random" date of birth allocation scheme and were excluded from further analysis (40,51,63). Two additional

trials were excluded: one enrolled patients less than 18 years of age (62) and another provided inadequate description of data (60). One paper separately reported the results of two parallel trials of ANH involving two different maxillofacial procedures and was considered as two trials, subsequently referred to as Bonnet_1 and Bonnet_2 (44). In the end, twenty-four RCTs enrolling a total of 1216 patients (630 ANH : 586 control) were available for analysis.

Twelve authors were sent written requests to supply supplemental data or clarify data presented in their publications, but only two responded (41,55). Sixteen trials reported the proportion of patients exposed to at least one unit of allogeneic blood, twelve documented the total units or volume of allogeneic blood transfused and thirteen reported total SBL in the perioperative period in a format suitable for analysis as a continuous variable (mean ± sd). Ten additional trials contributed results to the ordinal variable analysis of the volume of allogeneic blood transfused (22,43,44,46,48,49,52,55,56,66)

3.3. Trial Characteristics

The median sample size of the trials, including both ANH and control groups, was 30 (range 15-300). The mean preoperative hematocrits were similar in ANH and control groups (0.411 ± 0.017 vs 0.410 ± 0.020, respectively p=0.856). The mean volume of blood reserved preoperatively (V_{ANH}) in the ANH group was 989 ± 443 ml. Mean of the lowest intraoperative hemoglobin reported in the ANH groups was 0.270 ± 0.054 and was statistically lower than that reported in the control groups 0.326 ± 0.068 (p=.009). Surgical blood loss (SBL) in the perioperative period averaged 1211 ± 519 ml in those patients undergoing ANH and 1279 ± 538 ml in the control groups (p=0.692). The median score on the Jadad scale was one out of a possible five. Only five trials, all

scoring two, exceeded the median. Reviewers agreed on the Jadad score in 18 of 24 trials. Assessment of agreement beyond chance yielded a Cohen's kappa of 0.69. During the review process it became evident that the lack of agreement resulted from a difference in the interpretation of the third item of Jadad's scale that allots a point for accounting for all subjects enrolled. Consensus on the scoring of withdrawals was reached following discussion and clarification with Dr. Jadad. Fourteen trials did not specify a transfusion protocol or employed a protocol that set different transfusion thresholds in the ANH and control groups. Characteristics of the included RCTs are summarized in Table 3.

Table 3. Characteristics of Included Trials

Author	Surgery	Sample Size ANH/Control	Mean V_{ANH} (ml)	Mean SBL (ml) ANH Group	Mean SBL (ml) Control Group	Jadad Score	Transfusion Protocol
Hallowell (22)	Cardiac	25/25	1252	961	1110	1	no
Lawson (46)	Cardiac	150/150	900	700	900	1	no
Kaplan (61)	Cardiac	60/20	700	-	-	1	no
Lilleaasen (47)	Cardiac	15/15	855	812	1517	1	no
Dale (66)	Cardiac	7/8	850	1350	2850	1	no
Dietrich (45)	Cardiac	25/25	750	890	664	1	yes
Boldt (43)	Cardiac	15/15	890	1000	1080	1	yes
Vedrinne (64)	Cardiac	30/30	400	775	834	1	no
Herregods (58)	Cardiac	15/15	785	-	-	1	no
Triluzi (53)	Cardiac	18/28	924	475	735	2	yes
Kochamba (59)	Cardiac	50/50	866	696	916	2	yes
Vara Thorbeck (55)	Ortho	52/48	1000	-	-	1	no
Lorentz (48)	Ortho	16/15	700	1753	1720	1	yes
van der Linden(54)	Ortho	10/10	880	1453	1170	2	no
Bennett (42)	Ortho	20/20	900	1750	1750	1	yes
Rose (52)	GI	15/6	2700	-	-	1	no
Moyes (50)	Thoracic	10/10	1500	2338	1277	2	no
von Bormann (56)	Hepatic	22/22	-	1924	2094	2	yes
Bonnet 1 (44)	ENT	10/10	1015	535	801	1	no
Bonnet 2 (44)	ENT	10/10	870	1100	1340	1	no
Malinovsky (49)	Urology	15/15	740	-	-	1	yes
Atallah (41)	Urology	10/10	900	-	-	1	no
Catoire (41)	Vascular	10/10	875	1560	910	1	no
Welch (57)	Vascular	20/19	1500	1640	1340	1	yes
Summary		630/586	989 ± 443 mean ± SD	1211 ± 519 mean ± SD	1279 ± 538 mean ± SD		

Author (reference number)

3.4. Effects of ANH on Allogeneic Transfusion

The influence of ANH on the likelihood of exposure to allogeneic blood transfusion is shown in Figures 1 a and b. For ease of presentation all Figures are grouped together on pages 30-44 at the end of the results section. When all eligible trials were considered, ANH significantly reduced the likelihood of exposure to at least one unit of allogeneic blood in both the fixed effects (Relative Risk (RR) 0.65, 95% Confidence Interval (CI) 0.56 to 0.75) and random effects models (OR 0.67, 95% CI 0.50 to 0.92). Marked heterogeneity of results as assessed by the chi square test for heterogeneity was present in both fixed and random effects models ($p < 0.001$).

Reductions in the likelihood of exposure to at least one unit of allogeneic blood were echoed by reductions in the total number of units transfused (U) as shown in Figures 2 a and b. Patients in the ANH group received fewer units of allogeneic blood than those in the control group in both the fixed (Weighted Mean Difference (WMD) -1.31 U, 95% CI -1.60 to -1.02) and random effects models (WMD -2.37 U; 95% CI -3.74 to -1.01). Marked and statistically significant heterogeneity was again noted ($p <0.001$).

As shown in Figures 3a and b, ANH had a small and clinically insignificant effect on the volume of blood lost in the perioperative period regardless of the model chosen (WMD fixed effects -173 ml; 95% CI -236 to -110) or random effects -166 ml; 95% CI -321 to -10).

Marked heterogeneity of results was found in all of the analyses shown above. Indeed, heterogeneity was common in most of the analyses performed in this overview. In the presence of heterogeneity the fixed effects model yields estimates of treatment effect that are more precise than the data may support. Consequently, all further analyses

will report only analyses based on random effects models. Methods of dealing with heterogeneity will be reviewed in detail in the discussion.

3.5. Results Unsuitable for Analysis as a Continuous Variable

Meta-analysis of the volume of allogeneic blood transfused as a continuous variable requires that a summary mean and standard deviation for this outcome be reported in both the ANH and control groups. In ten trials data describing the volume of allogeneic blood transfused was reported without a valid standard deviation which could be used for summary analysis.

Table 4. Study Group and Volume of Allogeneic Transfusion

	Allogeneic Blood Transfused					
Group	0	1	2	3	4 +	Total
Included ANH		6	2	2	2	12
Included Control		1	4	1	6	12
Excluded ANH	4	0	1	1	4	10
Excluded Control	0	2	2	1	5	10

Included = those trials providing results for the analysis of the volume of allogeneic blood transfused.
Excluded = those trials which provided some measure of the volume of allogeneic blood transfused but could not be used for continuous variable meta-analysis.
Cell counts represent the number of trials reporting the given mean units of allogeneic blood transfused.

In four trials no patient in the ANH group received an allogeneic transfusion (43,44,49,52) yielding a zero value for standard deviation. Three trials reported the volume of allogeneic blood transfused in the intraoperative and postoperative periods separately (22,56,66) and failed to provide a summary mean and standard deviation. Two trials reported mean volumes of allogeneic blood transfused but used a range to describe variability (46,48) and a single trial provided no measure of variability (55). Information

from all trials that reported the volume of allogeneic blood transfused in some fashion is summarized in Table 4.

ANH arms of both included and excluded trials were more likely to report mean allogeneic blood exposure of zero or one unit than control arms. Similarly control groups were more likely to report 4 or more units transfused than ANH groups. In very crude terms, it can be seen that the included and excluded trials had a similar "pattern" of allogeneic blood exposure.

Three trials (42,43,50) reported the volumes of blood transfused to individual patients in ANH and control groups which permitted analysis of this outcome as an ordinal variable. The details of this analysis are discussed in Appendix C and results summarized in Table 5.

Table 5. Volume of Allogeneic Blood Transfused - Ordinal Variable Analysis.

Study	Odds Ratio	θ_i	ω_i
Moyes	3.676	1.302	1.332
Boldt	3.428	1.232	2.161
Bennett	1.086	0.083	2.406
Total	2.18 (1.44 to 3.29)	0.42 (0.37 to 1.19)	

Values in brackets represent 95% confidence intervals.
θ_i = logistic regression β coefficient representing treatment effect
ω_i = (variance of θ_i)$^{-1}$.

3.6. Adverse Events

The incidence of complications was poorly reported but there was no obvious increase in adverse events with ANH. Two deaths, one in each of the ANH and control groups, were reported among 314 patients in the eight trials that reported mortality (45,50,52,54,56,57,59,65). A single myocardial infarction was reported, in the control group, in the four trials enrolling 116 patients that reported this outcome (48,50,54,56,65). In the two trials that specifically assessed 120 patients for

thromboembolic complications (54,55) the likelihood of deep venous thrombosis (RR 0.61, 95% CI 0.14 to 2.69) and pulmonary embolism (RR 1.00, 95% CI 0.20 to 5.09) were not influenced by ANH.

3.7. Publication Bias

The "funnel plot" evaluating the presence of publication bias in trials evaluating the influence of ANH on the likelihood of allogeneic transfusion is displayed in Figure 4a. Negative or equivocal trials enrolling small numbers of subjects, found in the upper right corner of the plot, are underrepresented suggesting a systematic underreporting of trials not showing a reduction of allogeneic transfusion with ANH.

A graphical representation of a linear regression technique for evaluating funnel plot asymmetry is shown in Figure 4b. The natural log of the standardized normal deviate is plotted on the x axis and precision shown on the y axis. In the absence of trials with equivocal or negative results the small trials reporting incredibly beneficial effects of ANH pull the regression line well below the origin. Statistical testing of the α value in the regression ($\ln SND = \alpha + \beta precision$), which represents the y intercept, shows markedly significant asymmetry ($\alpha = -7.95$, $p < 0.001$). These findings again suggest the presence of publication bias.

3.8. Subgroup Analyses

3.8.1. Surgical Procedure

When surgical procedures were considered separately ANH was effective in reducing the likelihood of exposure to at least one unit of allogeneic blood in miscellaneous procedures (RR 0.25, 95% CI 0.08 to 0.76) but did not yield statistically

significant reductions in trials of cardiac surgery (RR 0.87, 95% CI 0.72 to 1.06) or orthopedic surgery (RR 0.88, 95% CI 0.39 to 2.02) (Figures 5 a,b,and c).

Reductions in the numbers of units of allogeneic blood transfused in different surgical procedures are presented in (Figures 6 a,b,and c). ANH groups undergoing cardiac surgery (WMD –2.83, 95%CI –5.34 to –0.31), and miscellaneous procedures (WMD –1.8, 95% CI –3.35 to –0.30) showed clinically and statistically significant reductions in the volume of allogeneic blood transfused. Reduction in allogeneic transfusion in ANH groups undergoing orthopedic surgery did not reach statistical significance (WMD –1.54, 95% CI –4.01 to 1.317).

As with the pooled analyses, statistically significant heterogeneity was present in most of the subgroup analyses presented above. The notable exception was in the cardiac surgery subgroup where heterogeneity in the analysis of the RR of allogeneic blood exposure was absent. Values of the c2 test for heterogeneity for the summary meta-analyses and the subgroups based on surgical procedure are found in Table 6.

Table 6. Heterogeneity in ANH Research

Measure	Group	χ^2	df	p
RR	Cardiac	4.5	4	0.343
RR	Ortho	11.1	2	0.004
RR	Misc	42.7	6	<0.001
RR	All Trials	53.5	14	<0.001
WMD	Cardiac	161.4	5	<0.001
WMD	Ortho	21.0	1	<0.001
WMD	Misc	12.3	3	0.006
WMD	All Trials	216.3	11	<0.001

χ^2 = chi square test for heterogeneity
df = degrees of freedom

3.8.2. Volume of Preoperative Phlebotomy

The influence of the volume of preoperative phlebotomy (V_{ANH}) on the efficacy of ANH is seen in Figures 7 a,b,c and d. In trials in which the V_{ANH} was less than 1000 ml (Figure 7a) the reduction in the likelihood of transfusion failed to reach statistical significance (RR 0.78, 95% CI 0.55 to 1.0). On the other hand, studies in which V_{ANH} was 1000 ml or more (Figure 7c) there was a large and statistically significant reduction in the likelihood of transfusion (RR 0.33, 95% CI 0.13 to 0.83). Remarkably, the significant reduction in the likelihood of transfusion associated with larger phlebotomy did not correlate with a greater reduction in the mean volume of allogeneic blood transfused (Figures 7 b and d). Similar reductions in the mean units of allogeneic blood transfused resulted if V_{ANH} was less than 1000 ml (WMD -2.30 U, 95% CI -3.79 to -0.81) or exceeded 1000 ml (WMD -1.69 U, 95% CI -3.42 to 0.03).

3.8.3. Volume of Surgical Blood Loss

The volume of perioperative surgical blood loss (SBL) did not appear to influence the likelihood that ANH reduced exposure to allogeneic blood. Relative risks were similar and showed a trend toward a reduction in transfusion whether SBL was less than 1000 ml (RR 0.84, 95% CI 0.58 to 1.22) or greater than 1000 ml (RR 0.79, 95% CI 0.53 to 1.19), (Figures 8 a and c respectively). Neither trend reached statistical significance. Surprisingly, large differences in the volume of blood transfused were noted despite the similar relative risks noted above. When SBL was less than 1000 ml (Figure 8b) the reduction in the mean number of units transfused failed to reach clinical or statistical significance (WMD -0.30 U, 95% CI -0.81 to 0.20). Conversely, when SBL was greater

than 1000 ml (Figure 8d) the use of ANH led to a large reduction in donor blood consumption (WMD -4.01 U, 95% CI -6.67 to-1.36).

3.8.4. Language of Publication

The language of publication did not influence the reported efficacy of ANH (Figures 9 a and b). Trials published in languages other than English showed similar reductions in the likelihood of exposure to one or more units of allogeneic blood (RR 0.61, 95% CI 0.34 to 1.10) as those reported in English language journals (RR 0.68, 95% CI 0.44 to 1.04). A single non-English language publication reported the number of units of allogeneic blood transfused as a continuous variable precluding subgroup analysis based on this outcome (44).

3.8.5. Methodological Quality

The quality of the reporting of trial methods, as assessed by the Jadad scale, did not greatly influence the efficacy of ANH. Trials scoring a single point on the Jadad scale (Figures 10 a and b) demonstrated both statistically significant reductions in the likelihood of transfusion (RR 0.60, 95% CI 0.38 to 0.94) and the units of allogeneic blood transfused (WMD –2.89, 95% CI -4.95 to –0.84). Reductions in allogeneic blood exposure were similar in those trials scoring greater than one on the Jadad scale (Figures 10 c and d) (RR 0.76, 95% CI 0.61 to 0.94 and WMD –1.06, 95% CI -2.95 to 0.43).

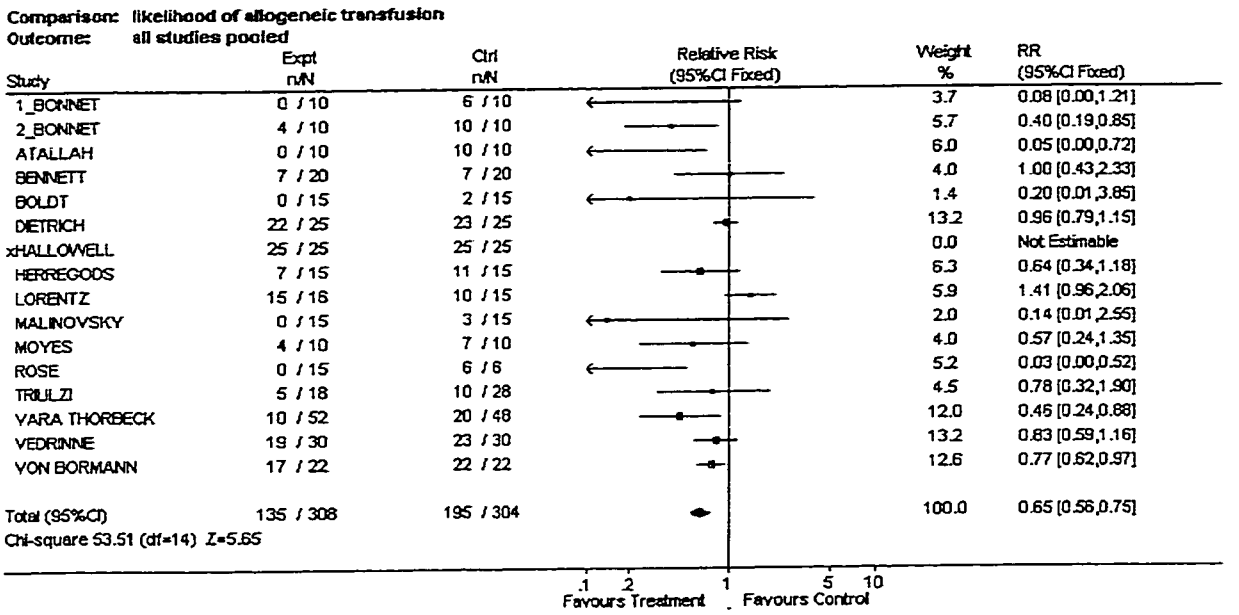
The influence of a transfusion protocol on the likelihood of allogeneic blood exposure is shown in Figures 11 a,b,c,and d. Trials without transfusion protocols showed marked reductions in both the likelihood of exposure to allogeneic blood (RR 0.39, 95% CI 0.22 to 0.69) and the units of allogeneic blood transfused (WMD –3.33. –5.54 to –1.12), (Figures 11 a and b respectively). In contrast, trials that carefully controlled

transfusion decisions with a transfusion protocol (Figures 11 c and d) were unable to demonstrate a reduction in allogeneic exposure (RR 0.92, 95% CI 0.76 to 1.10) or the units of allogeneic blood transfused (WMD -0.32 units; 95% CI -0.94 to 0.30). It is interesting to note that statistically significant heterogeneity is absent in trials employing a transfusion protocol despite significant clinical heterogeneity.

3.9. Figures

Figure 1. Effect Of ANH on the Likelihood of Allogeneic Transfusion (All Studies)

a. Fixed Effects Model



b. Random Effects Model

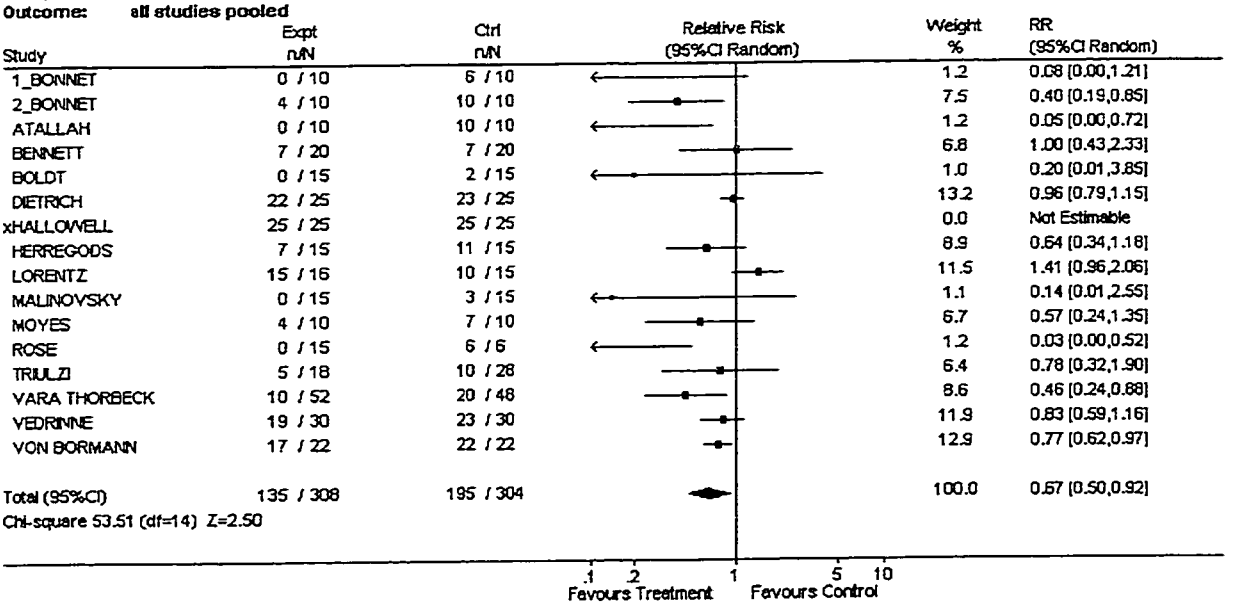


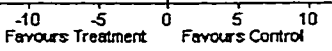
Figure 2. Effect of ANH on the Units of Allogeneic Blood Transfused (All Studies)

a. Fixed Effects Model

Comparison: units of allogeneic blood transfused
Outcome: all studies pooled

Study	Expt n	Expt mean(sd)	Ctrl n	Ctrl mean(sd)	WMD (95%CI Fixed)	Weight %	WMD (95%CI Fixed)
2_BONNET	10	0.40 (0.51)	10	3.30 (0.94)		19.1	-2.900 [-3.563,-2.237]
BENNETT	20	0.70 (1.20)	20	0.80 (1.30)		13.9	-0.100 [-0.875,0.675]
CATOIRE	10	0.90 (1.20)	10	1.40 (1.60)		5.5	-0.500 [-1.740,0.740]
DIETRICH	25	2.80 (2.00)	25	4.60 (3.20)		3.8	-1.800 [-3.279,-0.321]
KAPLAN	60	7.82 (5.99)	20	9.49 (3.61)		1.7	-1.670 [-3.861,0.521]
KOCHAMBA	50	0.96 (1.87)	50	1.20 (1.54)		18.6	-0.240 [-0.911,0.431]
LILLEAASEN	15	5.80 (2.00)	15	20.10 (3.80)		1.8	-14.300 [-16.473,-12.127
MOYES	10	1.30 (1.90)	10	2.83 (2.36)		2.4	-1.530 [-3.408,0.348]
TRIULZI	18	0.90 (1.80)	28	0.60 (0.90)		10.4	0.300 [-0.596,1.196]
VAN DER LINDEN	10	0.60 (0.90)	10	3.60 (1.30)		8.7	-3.000 [-3.980,-2.020]
VEDRINNE	30	1.70 (1.70)	30	1.90 (1.40)		13.5	-0.200 [-0.988,0.588]
WELCH	20	2.70 (3.10)	19	8.30 (8.70)		0.5	-5.600 [-9.741,-1.459]
Total (95%CI)	278		247			100.0	-1.313 [-1.602,-1.023]

Chi-square 216.27 (df=11) Z=8.89



b. Random Effects Model

Comparison: units of allogeneic blood transfused
Outcome: all studies pooled

Study	Expt n	Expt mean(sd)	Ctrl n	Ctrl mean(sd)	WMD (95%CI Random)	Weight %	WMD (95%CI Random)
2_BONNET	10	0.40 (0.51)	10	3.30 (0.94)		9.2	-2.900 [-3.563,-2.237]
BENNETT	20	0.70 (1.20)	20	0.80 (1.30)		9.1	-0.100 [-0.875,0.675]
CATOIRE	10	0.90 (1.20)	10	1.40 (1.60)		8.7	-0.500 [-1.740,0.740]
DIETRICH	25	2.80 (2.00)	25	4.60 (3.20)		8.5	-1.800 [-3.279,-0.321]
KAPLAN	60	7.82 (5.99)	20	9.49 (3.61)		7.6	-1.670 [-3.861,0.521]
KOCHAMBA	50	0.96 (1.87)	50	1.20 (1.54)		9.2	-0.240 [-0.911,0.431]
LILLEAASEN	15	5.80 (2.00)	15	20.10 (3.80)		7.6	-14.300 [-16.473,-12.127
MOYES	10	1.30 (1.90)	10	2.83 (2.36)		8.0	-1.530 [-3.408,0.348]
TRIULZI	18	0.90 (1.80)	28	0.60 (0.90)		9.0	0.300 [-0.596,1.196]
VAN DER LINDEN	10	0.60 (0.90)	10	3.60 (1.30)		9.0	-3.000 [-3.980,-2.020]
VEDRINNE	30	1.70 (1.70)	30	1.90 (1.40)		9.1	-0.200 [-0.988,0.588]
WELCH	20	2.70 (3.10)	19	8.30 (8.70)		5.0	-5.600 [-9.741,-1.459]
Total (95%CI)	278		247			100.0	-2.371 [-3.737,-1.005]

Chi-square 216.27 (df=11) Z=3.40

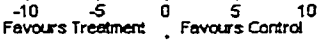
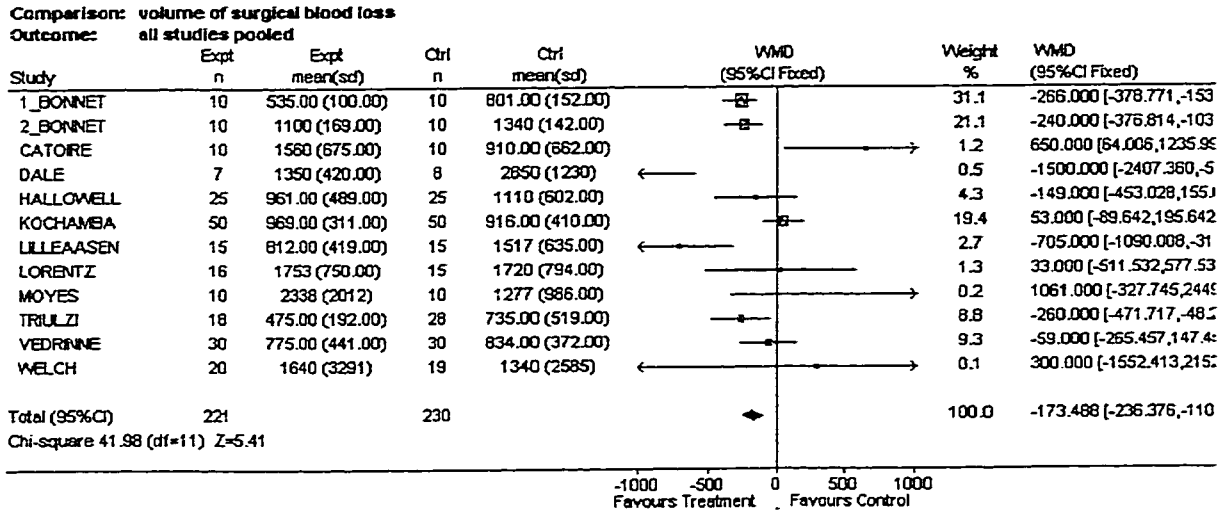


Figure 3. Effect of ANH on the Volume of Surgical Blood Loss (All Studies)

a. Fixed Effects Model



b. Random Effects Model

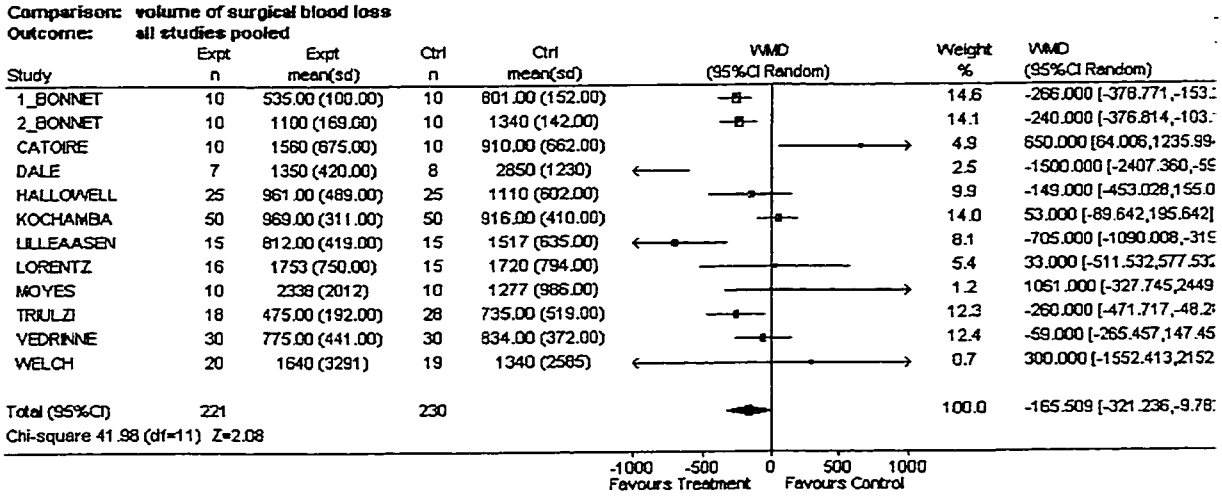
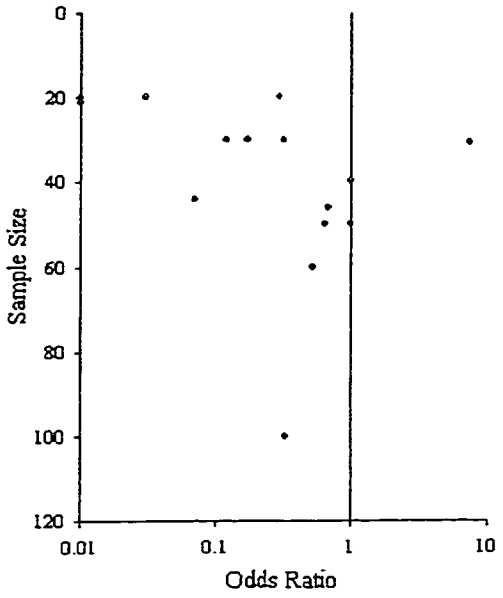


Figure 4. Publication Bias in ANH Research

a. Funnel Plot – Odds Ratio vs Sample Size



b. Egger Regression Model for Funnel Plot Asymmetry

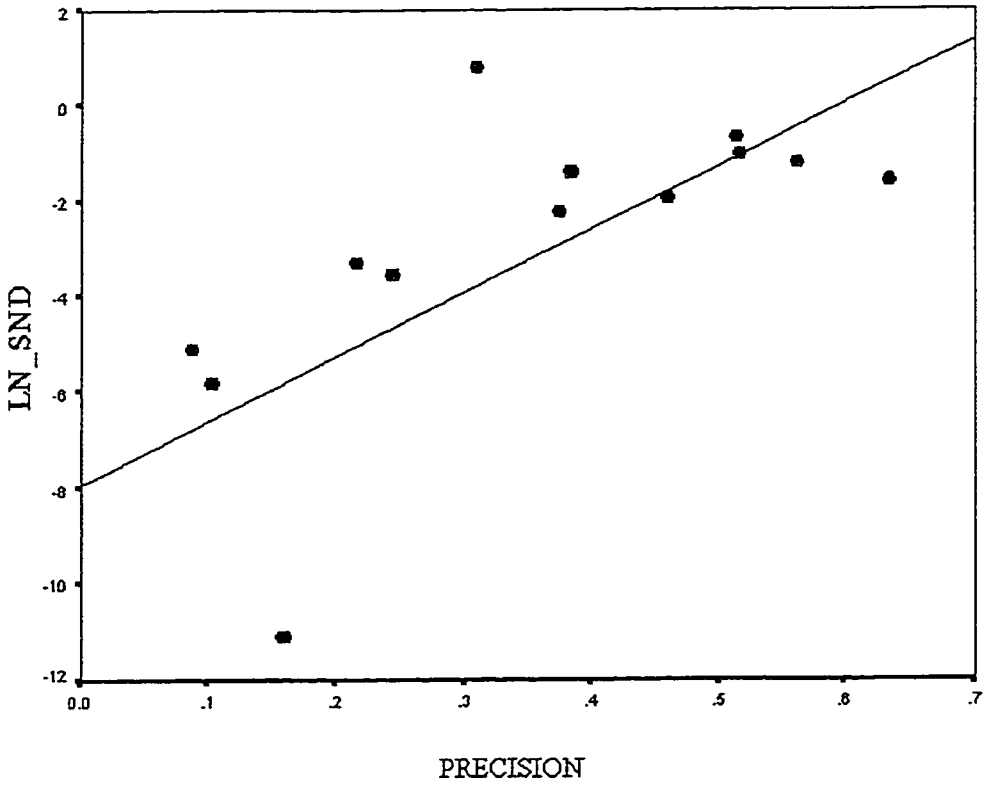
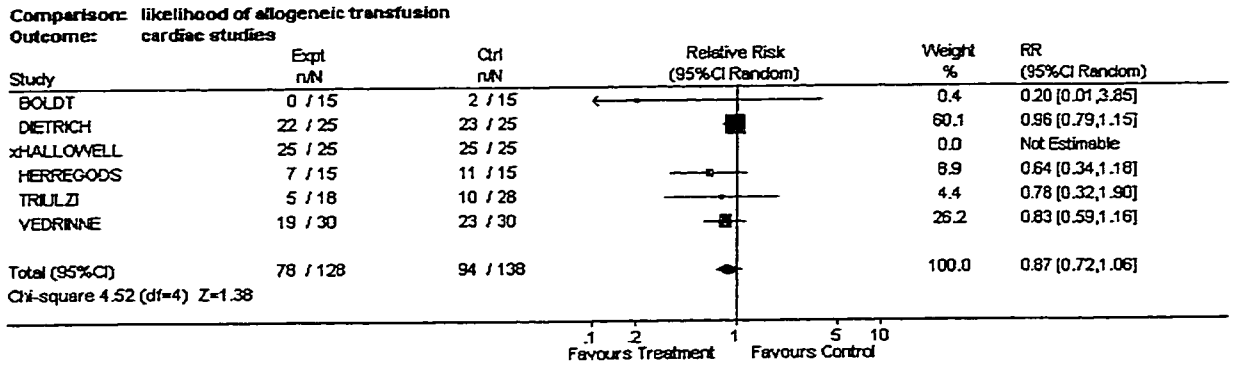
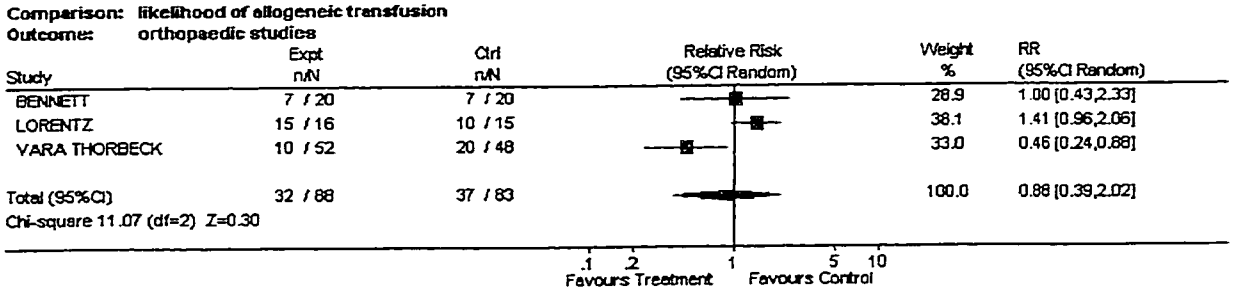


Figure 5. Effect of ANH on the Likelihood of Allogeneic Transfusion (By Surgical Procedure)

a. Cardiac Surgery



b. Orthopedic Surgery



c. Miscellaneous Surgical Procedures

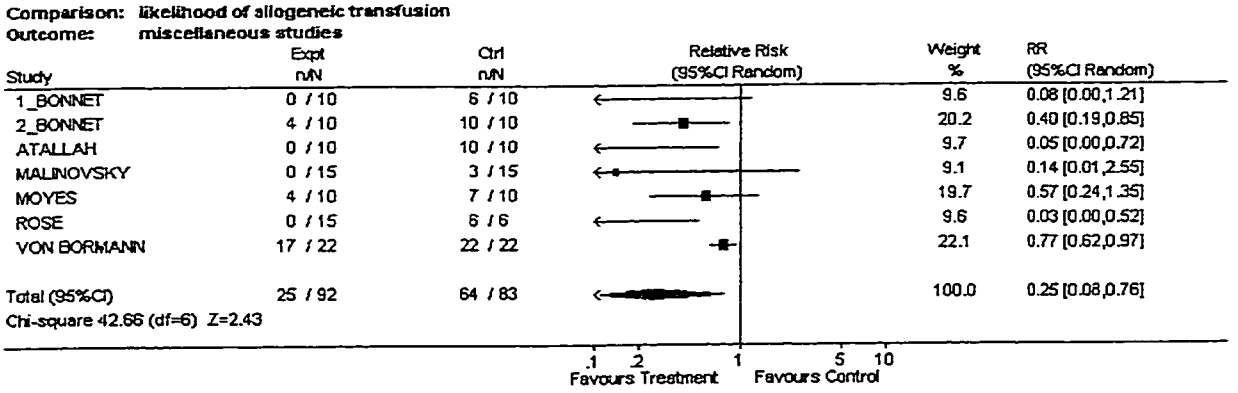


Figure 6. Effect of ANH on the Units of Allogeneic Blood Transfused (By Surgical Procedure)

a. Cardiac Surgery

Comparison: units of allogeneic blood transfused
Outcome: cardiac studies

Study	Expt n	Expt mean(sd)	Ctrl n	Ctrl mean(sd)	WMD (95%CI Random)	Weight %	WMD (95%CI Random)
DIETRICH	25	2.80 (2.00)	25	4.64 (3.21)		16.7	-1.840 [-3.323,-0.357]
KAPLAN	60	7.82 (5.99)	20	9.49 (3.61)		15.6	-1.670 [-3.861,0.521]
KOCHAMBA	50	0.96 (1.87)	50	1.20 (1.54)		17.5	-0.240 [-0.911,0.431]
LILLEAASEN	15	5.80 (2.00)	15	20.10 (3.80)		15.6	-14.300 [-16.473,-12.127
TRIULZI	18	0.90 (1.80)	28	0.60 (0.90)		17.3	0.300 [-0.596,1.196]
VEDRINNE	30	1.70 (1.70)	30	1.90 (1.40)		17.4	-0.200 [-0.988,0.588]
Total (95%CI)	198		168			100.0	-2.825 [-5.339,-0.311]
Chi-square 161.43 (df=5) Z=2.20							

-10 -5 0 5 10
Favours Treatment Favours Control

b. Orthopedic Surgery

Comparison: units of allogeneic blood transfused
Outcome: orthopaedic studies

Study	Expt n	Expt mean(sd)	Ctrl n	Ctrl mean(sd)	WMD (95%CI Random)	Weight %	WMD (95%CI Random)
BENNETT	20	0.70 (1.20)	20	0.80 (1.30)		50.5	-0.100 [-0.875,0.675]
VAN DER LINDEN	10	0.60 (0.90)	10	3.62 (1.30)		49.5	-3.020 [-4.000,-2.040]
Total (95%CI)	30		30			100.0	-1.544 [-4.405,1.317]
Chi-square 20.98 (df=1) Z=1.06							

-10 -5 0 5 10
Favours Treatment Favours Control

c. Miscellaneous Surgical Procedures

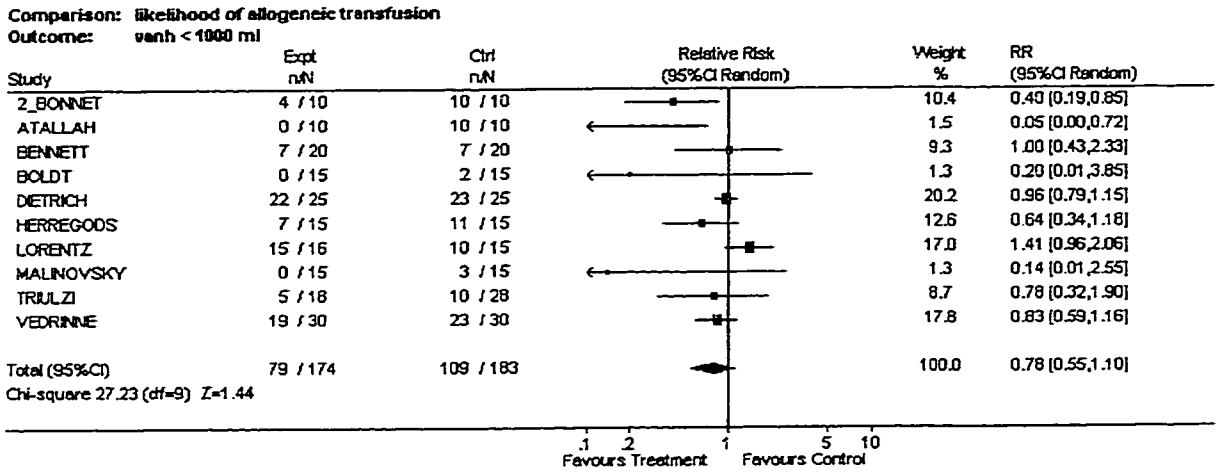
Comparison: units of allogeneic blood transfused
Outcome: miscellaneous studies

Study	Expt n	Expt mean(sd)	Ctrl n	Ctrl mean(sd)	WMD (95%CI Random)	Weight %	WMD (95%CI Random)
2_BONNET	10	0.40 (0.51)	10	3.33 (0.94)		35.6	-2.930 [-3.593,-2.267]
CATOIRE	10	0.90 (1.20)	10	1.40 (1.60)		30.4	-0.500 [-1.740,0.740]
MOYES	10	1.30 (1.90)	10	2.80 (2.40)		24.0	-1.500 [-3.397,0.397]
WELCH	20	2.67 (2.98)	19	5.33 (8.72)		10.0	-2.660 [-6.793,1.473]
Total (95%CI)	50		49			100.0	-1.820 [-3.345,-0.296]
Chi-square 12.28 (df=3) Z=2.34							

-10 -5 0 5 10
Favours Treatment Favours Control

Figure 7. Influence of the V_{ANH} of the Efficacy of ANH

a. V_{ANH} Less than 1000 ml (Likelihood of Allogeneic Transfusion)



b. V_{ANH} Less than 1000 ml (Units of Allogeneic Blood Transfused)

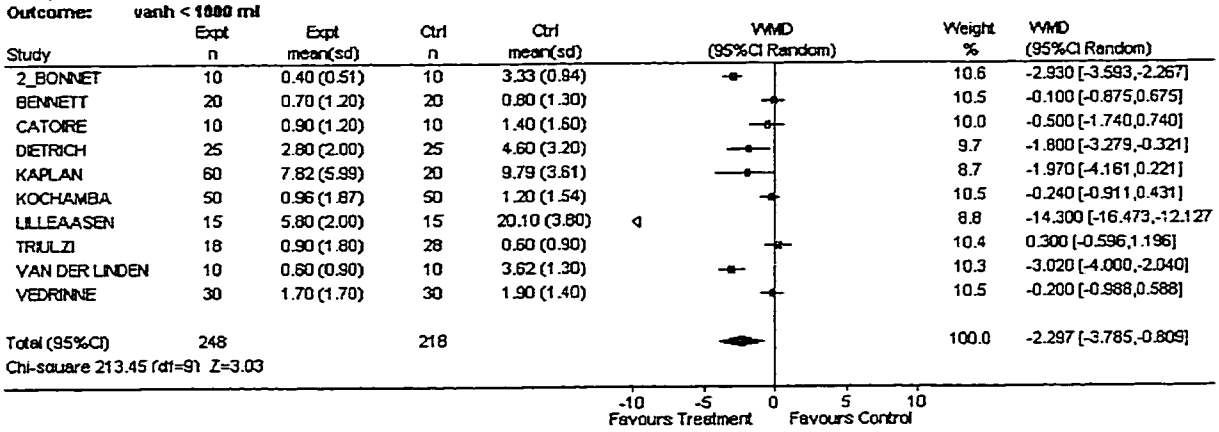
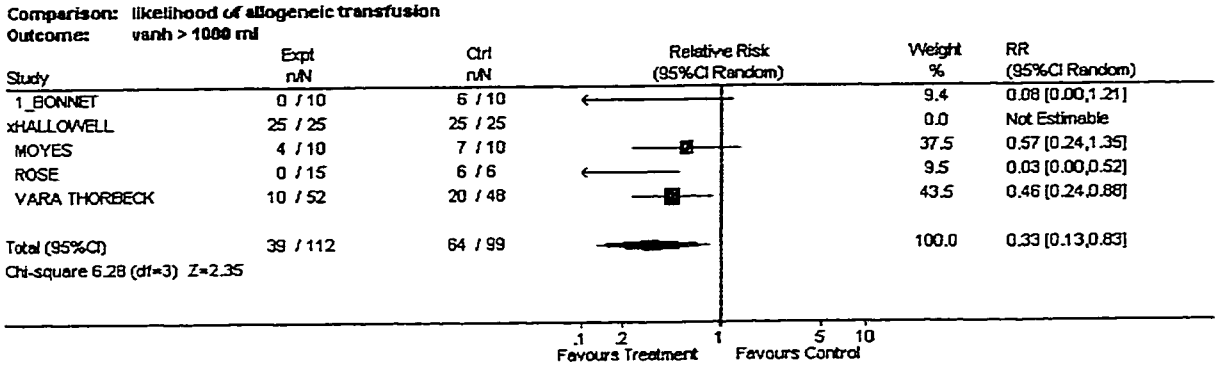


Figure 7. Influence of the V_{ANH} on the Efficacy of ANH

c. V_{ANH} Greater than 1000 ml (Likelihood of Allogeneic Transfusion)



d. V_{ANH} Greater than 1000 ml (Units of Allogeneic Blood Transfused)

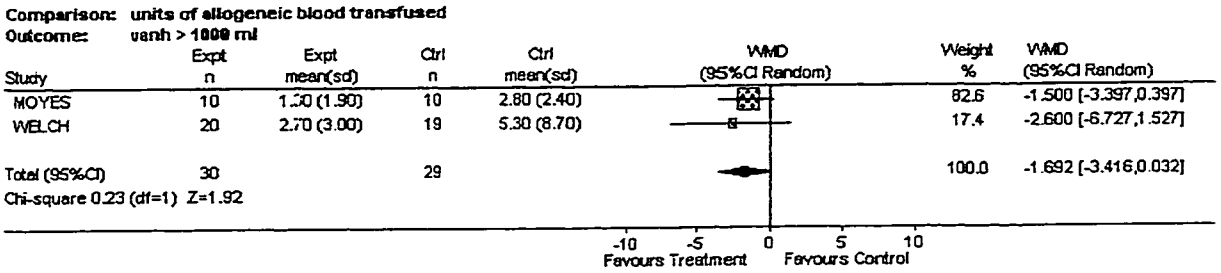
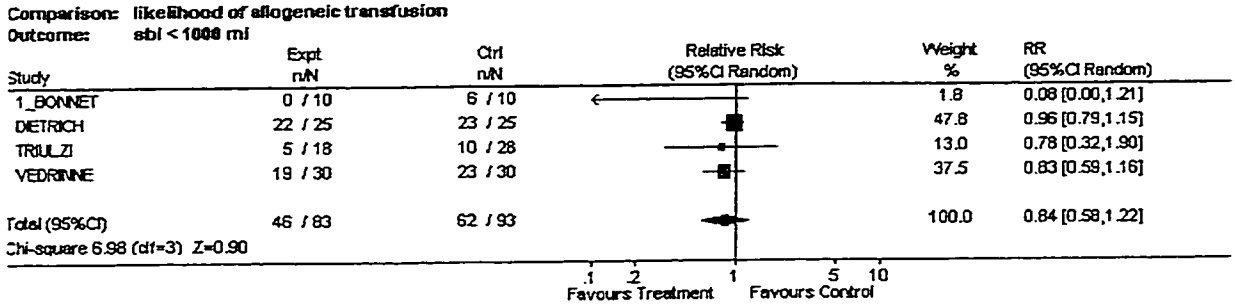


Figure 8. Influence of SBL on the Efficacy of ANH

a. SBL Less Than 1000 ml (Likelihood of Transfusion)



b. Blood Loss Less Than 1000 ml (Units of Allogeneic Blood Transfused)

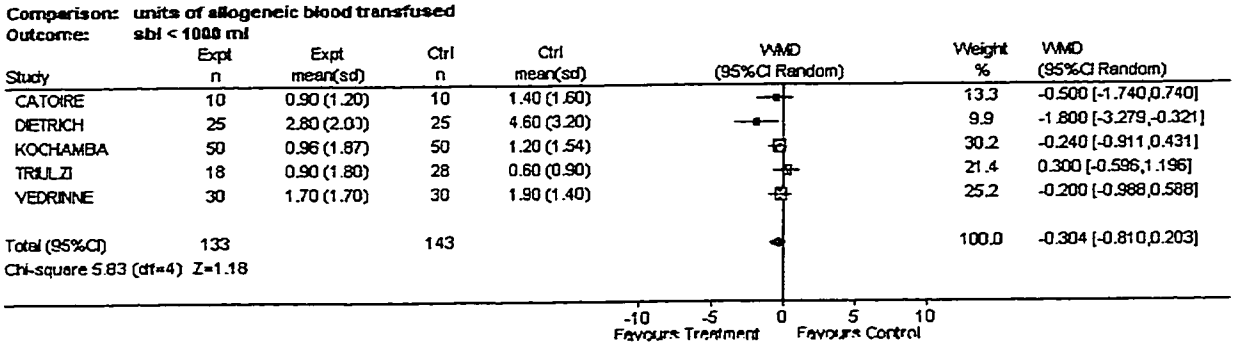
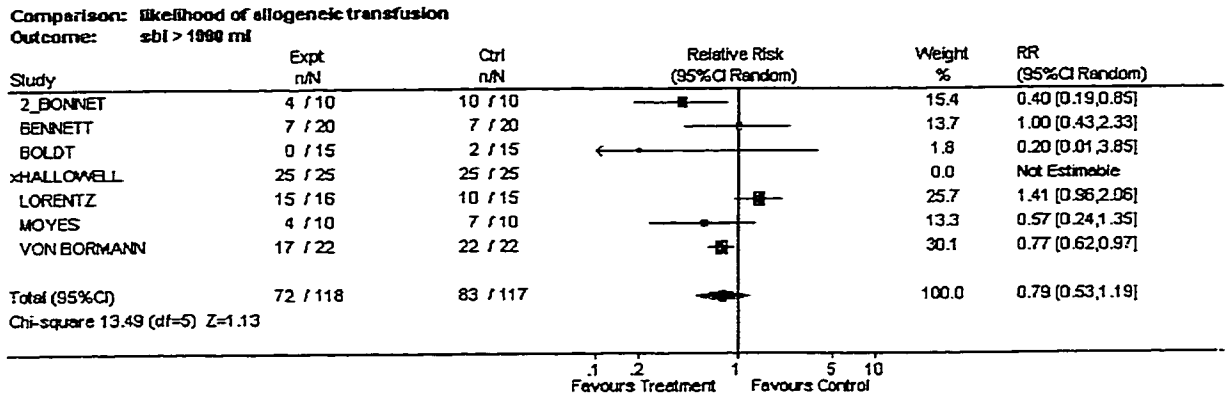


Figure 8. Influence of SBL on the Efficacy of ANH

c. SBL Greater Than 1000 ml (Likelihood of Allogeneic Transfusion)



d. SBL Greater Than 1000 ml (Units of Allogeneic Blood Transfused)

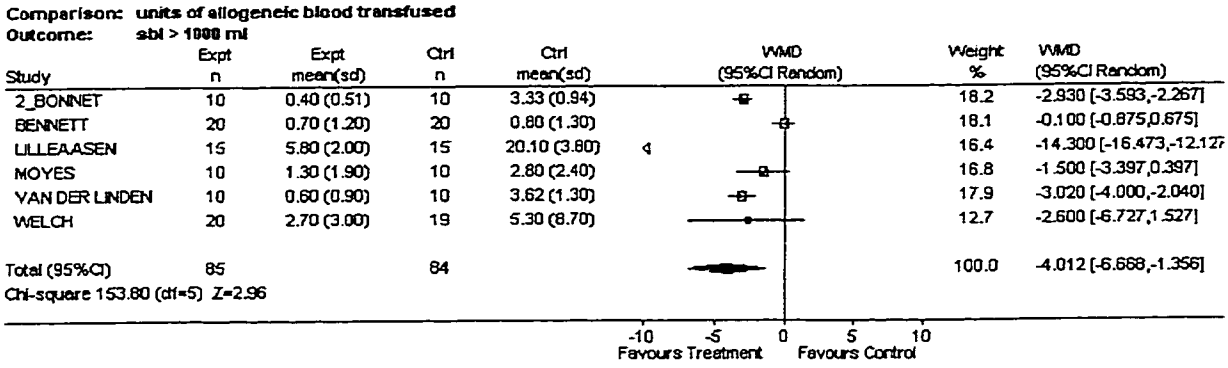
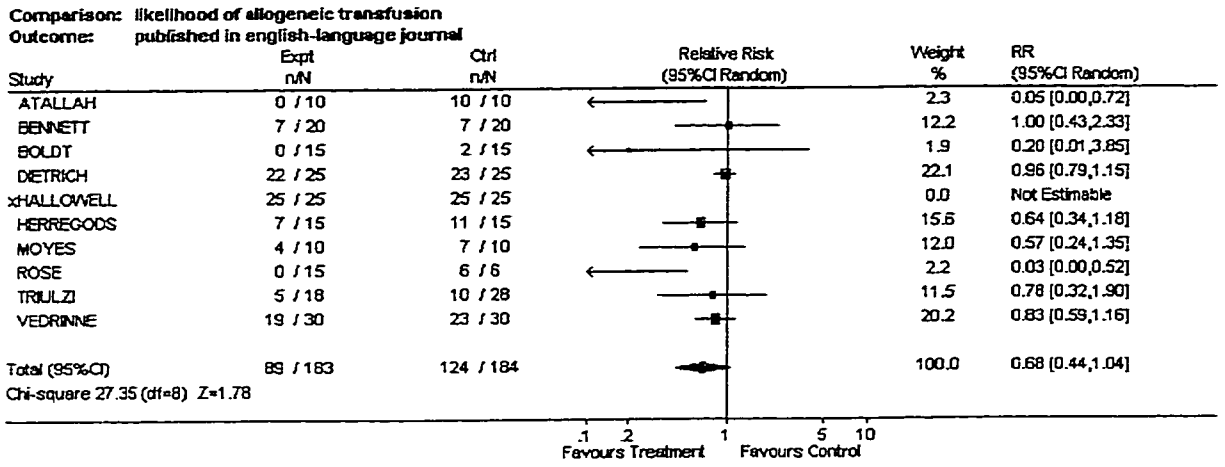


Figure 9. The Influence of Language of Publication on the Efficacy of ANH

a. Trials Published in English (Likelihood of Allogeneic Transfusion)



b. Trials Not Published in English (Likelihood of Allogeneic Transfusion)

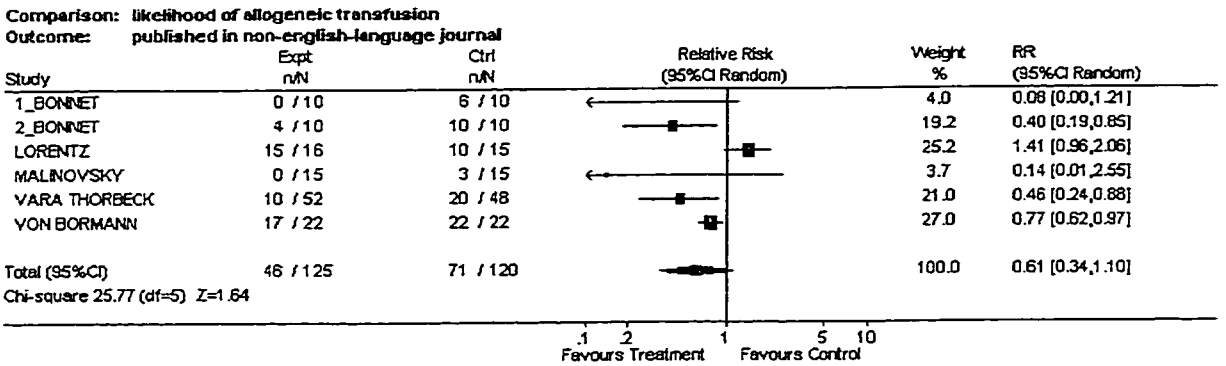
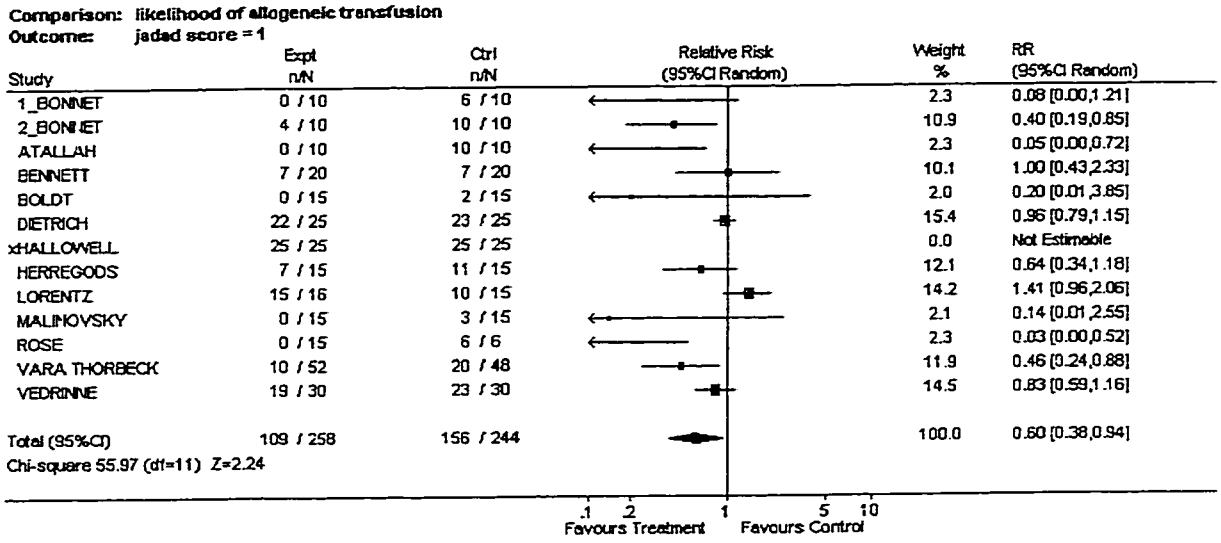


Figure 10. Influence of the Reporting of Trial Methods on the Efficacy of ANH

a. Jadad Score Equal to 1 (Likelihood of Allogeneic Transfusion)



b. Jadad Score Equal to 1 (Units of Allogeneic Blood Transfused)

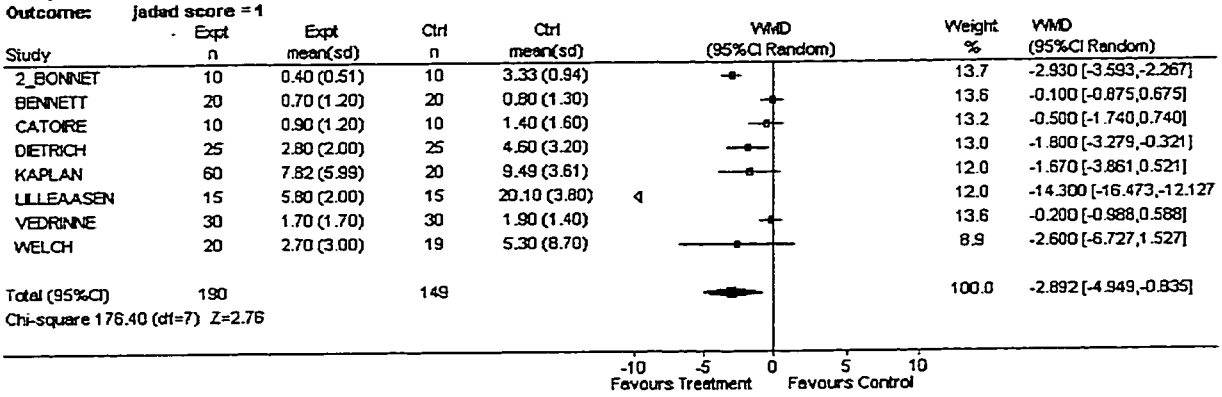
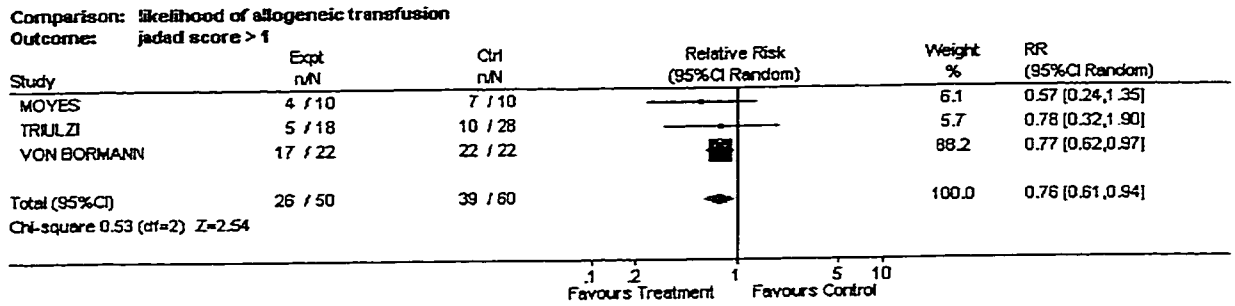


Figure 10. Influence of Reporting of Trial Methods on the Efficacy of ANH

c. Jadad Score Greater Than 1 (Likelihood of Allogeneic Transfusion)



d. Jadad Score Greater Than 1 (Units of Allogeneic Blood Transfused)

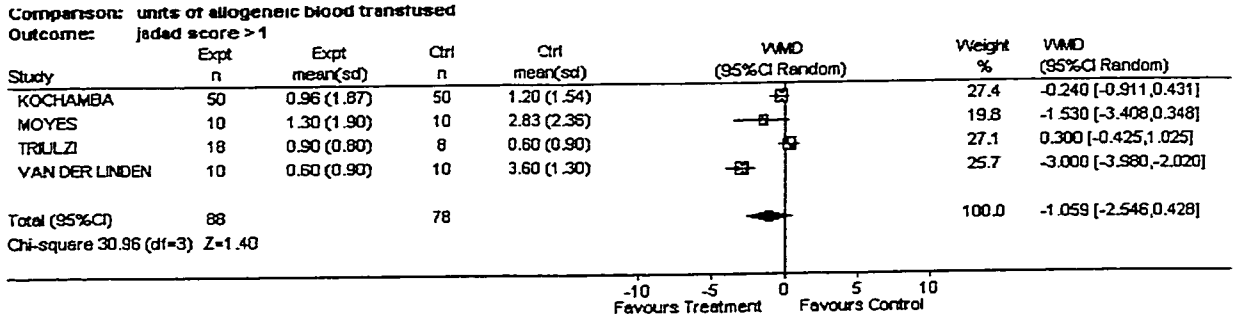
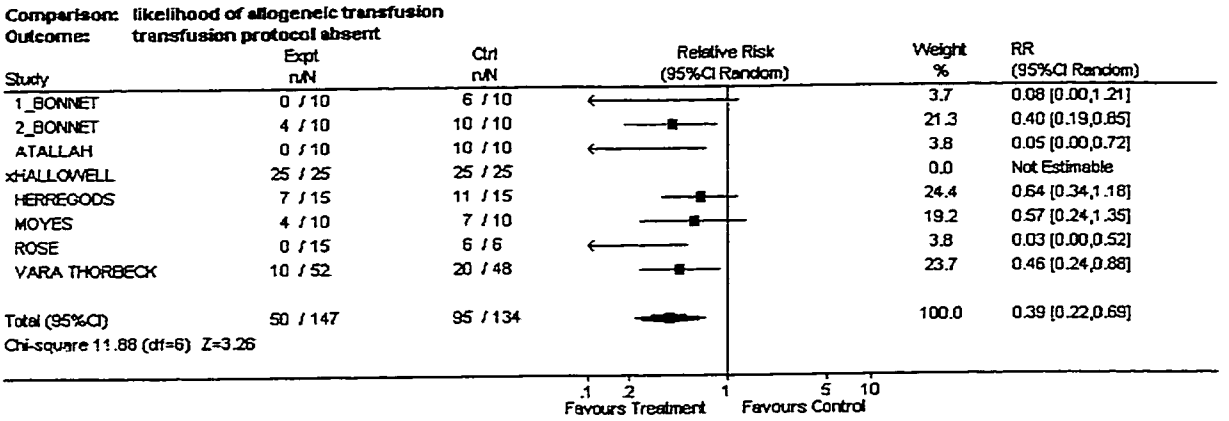


Figure 11. Influence of the Use of A Transfusion Protocol on the Efficacy of ANH

a. No Protocol Used (Likelihood of Allogeneic Transfusion)



b. No Protocol Used (Units of Allogeneic Blood Transfused)

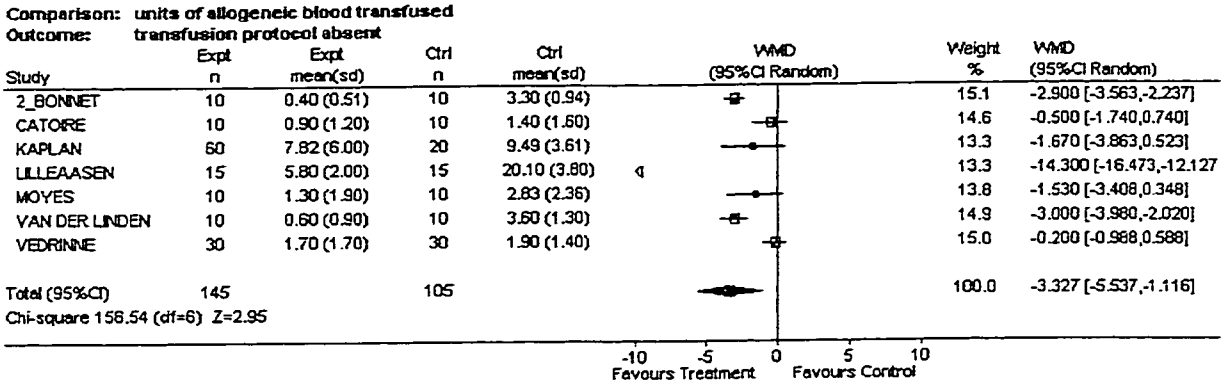
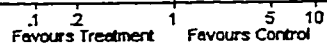


Figure 11. Influence of the Use of a Transfusion Protocol on the Efficacy of ANH

c. Protocol Used (Likelihood of Allogeneic Transfusion)

Comparison: likelihood of allogeneic transfusion
Outcome: transfusion protocol present

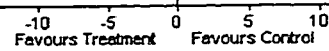
Study	Expt n/N	Ctrl n/N	Relative Risk (95%CI Random)	Weight %	RR (95%CI Random)
BENNETT	7 / 20	7 / 20		4.4	1.00 [0.43,2.33]
BOLDT	0 / 15	2 / 15		0.4	0.20 [0.01,3.85]
DIETRICH	22 / 25	23 / 25		30.5	0.96 [0.79,1.15]
LORENTZ	15 / 16	10 / 15		15.6	1.41 [0.96,2.06]
MALINOVSKY	0 / 15	3 / 15		0.4	0.14 [0.01,2.55]
TRIULZI	5 / 18	10 / 28		4.0	0.78 [0.32,1.90]
VEDRINNE	19 / 30	23 / 30		18.1	0.83 [0.59,1.16]
VON BORMANN	17 / 22	22 / 22		26.5	0.77 [0.62,0.97]
Total (95%CI)	85 / 161	100 / 170		100.0	0.92 [0.76,1.10]
Chi-square 10.84 (df=7) Z=0.92					



d. Protocol Used (Unit of Allogeneic Blood Transfused)

Comparison: units of allogeneic blood transfused
Outcome: transfusion protocol present

Study	Expt n	Expt mean(sd)	Ctrl n	Ctrl mean(sd)	WMD (95%CI Random)	Weight %	WMD (95%CI Random)
BENNETT	20	0.70 (1.20)	20	0.80 (1.30)		28.3	-0.100 [-0.875,0.675]
DIETRICH	25	2.80 (2.00)	25	4.60 (3.20)		13.1	-1.800 [-3.279,-0.321]
KOCHAMBA	50	0.96 (1.87)	50	1.20 (1.54)		31.8	-0.240 [-0.911,0.431]
TRIULZI	18	0.90 (1.80)	28	0.60 (0.90)		24.7	0.300 [-0.596,1.196]
WELCH	20	2.70 (3.10)	19	5.30 (8.70)		2.2	-2.600 [-6.741,1.541]
Total (95%CI)	133		142			100.0	-0.322 [-0.944,0.300]
Chi-square 7.04 (df=4) Z=1.02							



4. DISCUSSION

4.1. Literature Search

"Of primary importance to the results obtained in a systematic review or meta-analysis are the data collection methods used (32)". Use of the computerized Medline database, compiled by the National Library of Medicine, is the best-described and researched means of identifying randomized controlled trials (RCTs). Unfortunately, Medline does not provided complete coverage of all medical literature. Dickersin found that the sensitivity of Medline in identifying all known trials in vision research was only 51%. This study also noted that it is not possible to identify 25% of all published RCTs in journals indexed by Medline using Medline (32). If Dickersin's results can be generalized to other fields of research it is possible that 50% of eligible citations may be overlooked if Medline is used as the sole method of identifying trials. Use of alternate means of identifying RCTs is therefore essential to provide complete coverage of the topic at hand.

To increase the number of RCTs of ANH for review, this systematic overview utilized two additional computerized databases. Embase (Excerpta Medica Online) differs from Medline in that it emphasizes pharmacology, toxicology, Asian and European literature. Overlap of journals indexed in Embase and Medline is approximately 34% (67). Actual overlap of citations depends upon the topic reviewed and has been reported to range from 15 to 25% (68). Current Contents abstracts titles and bibliographic information on a weekly basis from 7000 journals. The prompt appearance of new research in the Current Contents database may be an important consideration in new or rapidly evolving technology.

Manual review of the bibliographies of review articles and retrieved trials identified seven trials not detected in the Medline search and was the most productive addition to Medline. Five citations identified by manual reviewed were present in Medline but were catalogued under the MeSH heading "blood transfusion, autologous." Unfortunately, a recent systematic overview of preoperative autologous donation identified 2275 citations using the MeSH heading "blood transfusion, autologous" (12). While inclusion of this term in future systematic overviews of ANH may improve capture of identification of trials within Medline, it will create an incredible amount of extra work. Two additional trials of ANH identified by manual review were excluded by terms in the search strategy intended to limit the Medline search to articles referring to blood transfusion practice. The limiting terms reduced the number of citations identified by Medline from approximately 3000 to the more manageable 880 reported here. Given the numbers of references identified in this overview, the use of a manual review in addition to a focussed literature search appears to strike a reasonable balance between comprehensiveness and identification of irrelevant articles.

The addition of an Embase search did not contribute significantly to this systematic overview. The single reference identified using Embase (60) was referenced in Medline and was also identified by manual review. Given the costs and time involved in performing an Embase search ($90.00 per hour of on-line time and $1.30 for each citation downloaded) future systematic overviews of ANH should carefully reconsider the use of this database.

It is evident that identification of all published RCTs is a time consuming and labour-intensive task. Medline identified 880 citations regarding ANH of which 18 were

RCTs. Use of the McMaster and Cochrane filters reduced the number of references to approximately 200 at the expense of missing 2 or 3 trials, respectively. While the filters offer some reduction in unnecessary publications their lack of sensitivity makes them unsuitable for future reviews of ANH. In the end there seems to be no reliable alternative to hard work when sorting through the considerable volume of medical research available for systematic overview.

4.2. Publication Bias

The most rigorous search of the literature will be unable to identify results of research that has not been published. Failure of completed, prospective, research to reach publication has a number of causes. Researchers report lack of interest and negative results as the most common reasons why they failed to publish the results of their research (69). Editorial practices of leading journals have also been identified as barriers to publication. Dickersin cites statements from several journal editors encouraging the submission of positive results and a quote stating "negative results have never made riveting reading" as evidence of editorial influence in the failure of completed research to reach the medical literature (70).

From the preceding comments it would appear that there is a bias against the publication of negative research findings. This impression has been confirmed by a number of researchers. Simes noted that published trials of combination chemotherapy for ovarian cancer suggested significant efficacy whereas results from prospectively registered trials, both published and unpublished, could demonstrate no improvement over single agent therapy (71). Results of Dickersin's survey of research publication practice suggested that unpublished trials, representing 23% of completed research,

favored the test therapy only 14% of the time. In contrast published research supported the test intervention in 55% of published research (69). Similarly Easterbrook found that among studies registered with Oxford's Research Ethics Committee (72), those with statistically significant results were more likely to be published than those with equivocal results (OR 2.32, 95% CI 1.25 to 4.28).

It appears that there is a bias against the publication of a subset of completed research and that the results of these studies differ in a systematic way from the findings of published research. Publication bias defined by Dickersin as "the selective submission and acceptance of positive over negative studies (69)" has the potential to influence the results of meta-analysis. It is intuitive that the results of a meta-analysis will be biased toward the study intervention if publication bias has limited access to trials with negative or equivocal results. Villar compared the results of meta-analyses with randomized controlled trials that enrolled more than 1000 subjects to determine if publication bias influenced the estimates of efficacy derived from meta-analyses. Among the 15 meta-analyses in which publication bias was suspected, 13 were found to have low or very low similarity to the results of the large trial compared with only 2 of 8 overviews lacking publication bias (73).

Several techniques to identify publication bias in the results of a meta-analysis have been published. Rosenthal suggests a technique that calculates the number of negative trials required to bring the summary positive result of a meta-analysis to the boundary of statistical significance and estimates the resistance of the results to publication bias (74). Begg and Mazumdar recommend the use of an adjusted rank correlation between effect size and its measure of variability to statistically test for

publication bias (75). Finally, Light and Pillemer describe a simple graphical means of detecting in which the estimate of treatment effect is plotted on the x-axis while sample size is plotted on the y-axis (38). Given these different approaches to the assessment of publication bias, which should one choose?

Villar compared the performance of these tests for publication bias in a subset of meta-analyses in the Pregnancy and Childbirth Module of the Cochrane Database of Systematic Reviews. Funnel plots were created for 23 meta-analyses and assessed blindly by two reviewers for their degree of asymmetry. These 23 meta-analyses were then assessed for publication bias using Rosenthal's measure of resistance to publication bias and Begg and Mazumdar's adjusted rank correlation. Agreement on the presence or absence of funnel plot asymmetry could be reached on 22 of 23 meta-analyses. Following consensus of the reviewers, funnel plot asymmetry was noted in 15 of 23 trials. Conversely, publication bias as assessed by the rank correlation method was absent in all 23 trials. There was no association between the funnel plot patterns and Rosenthal's measure. Similarly, there was no association between Rosenthal's number of trials predicted to make the results of the meta-analysis non-significant and the rank correlation coefficient.

The failure of the Begg and Mazumdar method to detect statistically significant publication bias is not surprising. The initial publication describing the performance of the method suggested that statistical power of the test was poor when the number of included trials was 75 or less and no meta-analysis in Villar's study had more than 16 trials. The lack of association between either the funnel plot or rank correlation and Rosenthal's estimates raises concern regarding the Rosenthal method. The funnel plot,

despite its subjective nature, had good inter-individual reliability and was more sensitive to the presence of publication bias than the rank correlation.

A funnel plot that assesses the presence of publication bias in ANH research is shown in Figure 4 a. The upper right corner of the plot representing trials showing no difference or an increase in the likelihood of allogeneic transfusion is absent. This funnel plot asymmetry indicates that there is a systematic underreporting of small negative trials of ANH and suggests publication bias.

Future systematic overview would benefit from improved, objective, statistical techniques to assess for the presence of publication bias. A test with power where the number of included trials is small would be a significant advance. Egger et al recently described a regression technique to measure funnel plot asymmetry (39). A graphical representation of Eggers's technique is shown in Figure 4b. Statistical testing reveals markedly significant asymmetry (α = -7.95, $p < 0.001$). Use of Egger's technique provides objective confirmation of the subjective results of the funnel plot and again suggests the presence of publication bias.

If publication bias is suspected unpublished research may be sought from industry, trial registries, and other non-peer-reviewed sources. Unfortunately, a new set of problems must be addressed when one incorporates the results of unpublished research into a systematic overview. First, how should one evaluate the unpublished results, often available as raw data, to avoid the introduction of bias? Second, can one be certain that unpublished data is of the same methodological quality as data subjected to peer-review? Last, should unpublished data be accorded the same value as published data in the meta-analysis?

Funnel plot asymmetry suggesting publication bias was evident in this systematic review of ANH research. Given the controversy surrounding the identification and incorporation of unpublished research in systematic overview no attempt was made to seek out unpublished literature. As trials reporting negative or equivocal results appear to be systematically under-represented it is likely that the effectiveness of ANH is overestimated in this meta-analysis. Furthermore, the reader cannot assume that the results of this meta-analysis can be used as a surrogate for a large well-conducted prospective trial.

4.3. Data Extraction

As is the case with other forms of retrospective research, there is ample opportunity for observer bias to influence the interpretation of the published results of the identified trials. Obscuring publication information to reduce bias has been suggested by a number of authors and has considerable face value (76,77). Jadad found that the scoring of the reporting of trial methods was lower and more consistent when trials were blinded (33).

In this overview, all trials were photocopied so that the title, authors, year and journal of publication were obscured. It is interesting to note that empirical evidence that blinding influences the final outcome of meta-analyses is lacking. Berlin compared open assessment to electronic scanning of the source publications and removal of any reference to author, institution, journal of publication or treatment group in five previously published meta-analyses. The considerable time and effort expended in blinding the reports failed to yield clinically or statistically significant differences in the results of the

subsequent meta-analyses (78). The importance of blinding of reviewers in systematic overviews is therefore uncertain.

To minimize observer bias, two independent investigators abstracted data from the trials. There was little disagreement about the objective evidence reported in the results (i.e. numbers of patients in the study groups, quantities of blood transfused, etc). Disagreement was noted on the more subjective scoring of trial quality as assessed by the Jadad score. Jadad scores, as assessed by two independent investigators, agreed in 17 of 23 trials (Cohen's $\kappa = 0.69$). The Jadad scale assigns points for the completeness of reporting three components of trial methods: randomization, blinding, and reporting of withdrawals. Disagreement arose between our reviewers over what constituted a complete report of withdrawals. Personal communication with Jadad revealed that this is a frequently encountered problem. Once settled, consensus could be reached on all items of Jadad's scale.

4.4. Continuous Variables and Measures of Variance

The primary outcome variable in this overview was the proportion of subjects exposed to at least one unit of allogeneic blood. In any surgical procedure associated with a substantial risk of transfusion a technology that is able to reduce the volume of allogeneic blood transfused, if not eliminate it all together, remains a valuable asset. Therefore a secondary, yet important, outcome in this overview was the quantity of allogeneic blood transfused.

Meta-analysis of continuous variables requires that a mean and standard deviation for the outcome of interest for both treated and control groups be entered for each trial. Please refer to Appendix C for a detailed description of the statistical techniques

employed for the meta-analysis of continuous variables. Unfortunately ten trials enrolling 644 subjects reported the volume of allogeneic blood transfused without a standard deviation which could be used for continuous variable analysis.

The primary goal of ANH is to reduce allogeneic blood exposure. If ANH worked perfectly no subject undergoing ANH would be exposed to allogeneic blood and the mean volume of allogeneic blood transfused to subjects in the ANH groups would be 0 ± 0. This was the case in four trials of ANH identified in this overview(43,49,52). Informative zero cells present a problem in the meta-analysis of both dichotomous and continuous data extracted from randomized controlled trials. For dichotomous data, addition of 0.5 of an event to each cell is an acceptable compromise and is the method used by Revman 3.0.1 to accommodate zero cells (79). Accommodation of zero cells in the meta-analysis of continuous data presents a much more difficult problem.

Revman 3.0.1 software, used to calculate the WMD for continuous outcomes reported in this overview, treats zero cells as missing values. Revman will not calculate a WMD or include the trial in the summary analysis if zero is entered as a value in its data tables. In order to include information from the 4 trials in which no subject received an allogeneic transfusion two attempts to impute a value for the standard deviation were assessed. First, the smallest non-zero value (0.01) was entered instead of zero for standard deviation of the given trial. This compromise proved unsatisfactory as the trial with the imputed data dominated the summary meta-analysis. Because the standard deviation is an important determinant of the weight applied to a given trial, an artificially small value for this term leads to an artificially large weight. As a result, this trial is

assumed, for purposes of analysis, as being an incredibly precise estimate of treatment effect and the trial assumes a disproportionate role in the summary analysis.

A second compromise evaluated was the use of the standard deviation in the control arm of the same trial. This solution weighted the trial in keeping with its sample size but removed the independence of the observations. Unfortunately, the calculation of confidence limits assumes independence of observation and this compromise was abandoned. A third compromise suggested by Follman (80), assumption of common variance for all trials, was rejected in view of the substantial qualitative and quantitative heterogeneity of the included trials.

Three additional trials presented data as mean ± standard deviation for allogeneic blood transfused in both the intraoperative and postoperative periods but no single value describing total allogeneic exposure. Mean values could be easily added to provide a mean total exposure; however, an accurate standard deviation for total exposure must be estimated.

A pooled variance was calculated using the following formula where sd_1 and sd_2 represent standard deviations from the intraoperative and postoperative periods and ρ represents an estimate of the correlation between variance in these periods.

$$\text{pooled} \quad sd = \left[sd_1^2 + sd_2^2 + \left(2\rho \sqrt{sd_1^2 \bullet sd_2^2} \right) \right]^{\frac{1}{2}}$$

Unfortunately no reliable estimate of the correlation between intraoperative and postoperative blood transfusion could be drawn. Intraoperative, postoperative, and summary values were reported in only two trials (45,53) precluding a reasonable estimate of ρ.

Two trials reported total allogeneic blood transfused as a mean and range. The standard deviation could be crudely estimated assuming that two to three standard deviations on either side of the mean encompass 95% and 99% of the range of values reported. Lastly, one trial provided a mean blood loss without a measure of variance following a request to the authors for unpublished data.

It became readily apparent that the number of assumptions made to incorporate the data from these 10 trials was becoming a threat to the validity of the results. Regrettably, exclusion of these trials resulted in the loss of information from 644 subjects enrolled in trials meeting the inclusion criteria for this analysis and raised concerns regarding the validity of the results of the analyses of the volume of allogeneic blood transfused.

Table 5 presents a descriptive approach to this problem. All trials reporting a summary measure of the quantity of allogeneic blood transfused to treatment and control have been arbitrarily divided into 6 discrete categories. Regrettably the hypothesis of lower transfusion levels in the ANH cannot be tested statistically. However, from this table a trend toward lower levels of transfusion in the ANH groups can be inferred. It would thus seem unlikely that inclusion of the other trials would substantially alter the direction of the findings of our summary analysis; this suggests a reduction in the volume of allogeneic blood transfused in the ANH groups. Beyond this generalization, trials excluded from analysis because of their measure of variance can provide no further information to the meta-analysis of this outcome. Another unified means of incorporating the results of all trials reporting the volume of allogeneic blood was therefore required.

Unfortunately, the volume of allogeneic blood transfused is of great importance in the evaluation of technologies that reduce perioperative transfusion. It is unrealistic to expect that any drug, technique or device will be able to entirely eliminate the requirement for allogeneic blood in some surgical procedures. However, a technology which reduces the number of units and donors that the patient is exposed is still of significant value. As this systematic review highlights, the statistical methods for the analysis of continuous variables are subject to a number of limitations. Of particular concern is the inability of these methods to incorporate meaningful zero values and standard deviations reported from different time periods. Further research is required to develop new means of incorporating continuous outcomes in the systematic overview of transfusion technology.

4.5. Meta-Analysis of Blood Transfusion as an Ordinal Variable

In an attempt to deal with the problems presented in the meta-analysis of the volume of allogeneic blood transfused, analysis of this outcome as an ordinal variable model was conducted. An ordinal approach better reflects the biological consequences of the administration of allogeneic blood than standard continuous-variable techniques. Allogeneic blood is usually administered in units containing approximately 300 ml of packed red blood cells. In most circumstances patients receive the full volume of the unit transfused. Once transfusion of a unit of allogeneic blood has begun the recipient is exposed to all the potential adverse events associated with that unit regardless of the total proportion of the unit administered. Measurement of the volume of allogeneic blood exposure in ANH research can therefore be thought of as discrete ordered categories.

Whitehead and Jones described a method of summarizing the results of randomized controlled trials with ordinal outcomes using a proportional odds model (37). A detailed description of Whitehead's model appears in Appendix C. Whitehead's ordinal approach obviates many of the problems with measures of variability encountered in the continuous variable analysis in this systematic overview. Treatment arms receiving no allogeneic blood or reporting variability in measures other than a single standard deviation may be easily accommodated.

Regrettably, Whitehead's model required that data regarding the volume of allogeneic blood transfused was available for each individual patient, making application of the technique in the meta-analysis of published data somewhat impractical. Individual subject data was available in only three trials of ANH identified in this overview (50). All three trials (42,43) were already represented in the continuous variable analysis.

Whitehead's model divides the volume of allogeneic blood transfused into binary outcomes – treatment success or treatment failure – based on an arbitrary threshold. For example, transfusion of a single unit of donor blood, or less, may be considered success while transfusion of two or more units will be considered failure. By adjusting the level of outcome representing a treatment success, the odds of a favorable outcome in the treated group relative to the control group for any level of allogeneic blood exposure may be calculated.

In the trials available for ordinal variable analysis the likelihood of lower allogeneic blood exposure was greater in those patients undergoing ANH (Odds Ratio 2.18, 95% CI 1.44 to 3.29). These results suggest that ANH is effective in reducing the volume of allogeneic blood transfused and agree with the findings of the continuous

variable analysis. This result should not be surprising given that the three analyzed trials represent a subset of the continuous variable analysis. While providing no new information, this analysis represents an interesting alternative to traditional continuous variable meta-analysis in transfusion trials.

4.6. Trials Excluded from Analysis as a Continuous Variable

Despite all attempts, data from ten trials reporting the volume of allogeneic blood transfused were excluded from analysis for "technical reasons." Table 4 represents an attempt to assess the impact of the removal of these trials in a qualitative sense. ANH arms of both included and excluded trials were more likely to report mean allogeneic blood exposure of zero or one unit than control arms. Similarly control groups were more likely to report 4 or more units transfused than ANH groups. Table 4 crudely demonstrates that the included and excluded trials had a similar "pattern" of allogeneic blood exposure. It is therefore unlikely that exclusion of these trials would markedly affect the direction of the findings of the summary analysis favoring lower volumes of blood transfused to ANH patients.

4.7. Heterogeneity

"The combining of heterogeneous material is a commonly accepted threat to the validity of meta-analysis (77)." Heterogeneity may arise from both clinical and statistical sources (81). Clinical heterogeneity results from differences in characteristics of the included trials such as subject selection, baseline risk of outcome, intervention technique, duration of observation, and study design. Statistical heterogeneity is found where there is greater variation in the results of the selected trials than would be expected by chance. Statistical heterogeneity may arise from clinical sources or simply from quantitatively

different results from qualitatively similar trials. Unfortunately statistical tests for heterogeneity lack power and may fail to detect substantially heterogeneous results (82,83). Through its attempt to gain objectivity, generalizability, and precision by combining all available evidence, meta-analysis would therefore seem to be doomed to some measure of heterogeneity. The question must then be how should one best deal with heterogeneity in systematic overview and meta-analysis?

4.8. Fixed and Random Effects Models

Two statistical models for the combining of data from previously published research, fixed and random effects, have been advanced. Considerable discussion has focussed on the appropriate choice of these meta-analytic models when faced with heterogeneity. Fixed and random effects models differ in one substantial respect, the choice of weights for the individual studies. The Mantel Haenszel (MH) method, a fixed effects model, weights studies by the inverse of the within-study variance. Since within-study variance is usually inversely correlated with sample size, the MH method may be considered to preferentially weight larger studies. The DerSimonian and Laird (D&L) method, a random effects model, weights studies by the inverse of a combination of both within- and between-study variation. Incorporation of between-study variance in the weighting procedure used in the D&L method has several implications (84). First, by adding another source of variation the D&L method should lead to an estimate of treatment effect with wider confidence intervals than the MH method. Second, the contribution of large trials to overall treatment effect in the D&L method is reduced as between-study variance increases and assumes a larger role in the weighting procedure. Lastly, the MH method may only yield valid inferences about the studies in hand,

whereas the D&L method may extend its scope to a hypothetical population of trials. The funnel plot asymmetry identified in this systematic overview suggests that there may indeed be a larger population of ANH research not identified in this overview. Berlin has stated that, "if there is heterogeneity that cannot be explained as a function of patient populations, protocols, etc., then the random effects model is more appealing."

If heterogeneity were present the contribution of between-study variance would be expected to increase. Berlin found that when the results of 22 previously published meta-analyses were reanalyzed using both fixed and random effects models the D&L method yielded a more conservative and less precise estimate of treatment effect where significant heterogeneity was present (84). Interestingly, three studies that supported treatment effect in the fixed effects analysis were equivocal in the random effects analysis. The use of random effects models in this systematic overview yielded results compatible with Berlin's observations. The influence of the chosen model was similar in the analysis of both dichotomous variables, upon which the discussion has thus far focused, and continuous variables for which both fixed and random effects models may be employed. In all cases statistical tests identified substantial heterogeneity. As a result, confidence limits surrounding the point estimates of treatment effect were wider and the weighting of larger studies was reduced under the random effects model (Figures 1 and 2).

Given the substantial heterogeneity present in the above analyses and the possibility of a larger hypothetical population of trials excluded from analysis by publication bias, the random effects model would seem the best choice for estimation of

the treatment effect of ANH. However, reliance on the statistical method to account for and manage heterogeneity is an overly simplistic approach to this problem.

4.9. Investigation of Clinical Heterogeneity

Clinical heterogeneity is clearly evident in the trials of ANH identified in this systematic overview. Included trials spanned a diverse range of surgical procedures and ANH techniques. An evaluation of the influence of these clinical factors should be undertaken "to investigate the influences of the specific clinical differences between studies rather than rely on a single overall test of heterogeneity. This then focuses attention on particular contrasts between the trials included, which will be more powerful at detecting genuine differences – and clinically and scientifically more relevant conclusions (81)." Subgroup analysis attempts to identify the sources of heterogeneity through analysis of subsets of trials based on these known clinical factors.

Several subgroup analyses were proposed *a priori*. Trials were grouped into three subgroups – cardiac, orthopedic, and miscellaneous - based on surgical procedure. Based on the preceding discussion of the influence of language of publication, trials were regrouped into English and non-English publications. Mathematical modeling of ANH, to be discussed in greater detail in the following section, suggests that the success of ANH is dependent upon the volume of blood withdrawn preoperatively (V_{ANH}) and the volume of blood lost during surgery (SBL). Subgroup analyses based on the V_{ANH} and SBL were undertaken to evaluate the role of these technical factors on efficacy. Lastly, the influence of trial methods was evaluated by comparing trials with a Jadad score greater than 1 with those scoring 1, the minimum inclusion criterion for this overview. During the review process it became apparent that a number of trials failed to employ a

protocol to guide transfusion decisions. A second subgroup analysis based on the reporting of a transfusion protocol was thus proposed.

4.9.1. Surgical Procedure

An obvious source of clinical heterogeneity is the use of ANH in different surgical procedures. Use of ANH in cardiac and orthopedic patients represented 11 and 4, respectively, of the 24 trials included in this overview. The remaining 9 trials were evenly distributed among a variety of other surgical procedures (Table 2) and were grouped together for simplicity. The reduction in the likelihood of transfusion associated with ANH was similar in cardiac and orthopedic trials while a marked and statistically significant reduction in transfusion was noted in the miscellaneous group. ANH was associated with similar reductions in the volume of allogeneic blood transfused regardless of the surgical procedure evaluated. Statistical heterogeneity persisted in many of these subgroup analyses. These results suggest that differences in surgical procedure are not solely responsible for the statistical heterogeneity identified in the pooled analysis and that other characteristics may contribute.

4.9.2. Language of Publication

Research published in languages other than English is often excluded from meta-analyses. Exclusion of non-English-language publications results in a decrease in the power of a systematic review and ignores valid research findings. Moher found that trials reported in non-English-language journals share similar methodological rigor and completeness of reporting as those trials that were published in English (85). Furthermore, it has been suggested that the exclusion of non-English language trials may bias the result of the meta-analysis in favor of the technology being evaluated. The

"Tower of Babel" bias occurs if non-English speaking investigators publish positive findings in English language journals and preferentially publish negative or equivocal research in local journals. Gregoire et al noted that the results of a systematic review favoring the use tk to tk would have been made non-significant by the inclusion of non-English language publications (86). Similarly Egger recently compared the results of matched pairs of publications period published in English and German from the same author in the same time period. Logistic regression identified that the only factor predicting publication in English was the presence of a p value <0.05 (87).

In this meta-analysis all publications regardless of language of publication were included, abstracted, and analyzed. It should be noted that a sizable reduction in the power of this meta-analysis would result from the exclusion of 20% (245/1216) of subjects reported in non-English journals. ANH produced statistically and clinically significant reductions in the likelihood of allogeneic transfusion in trials published in English and non-English language journals. Heterogeneity persisted in both analyses (Figures 9 a & b). It is therefore apparent that a "Tower of Babel" bias does not represent a source of statistical heterogeneity in this overview.

4.9.3. Technical Performance of ANH

The optimal method for performing ANH remains controversial and is reflected in the variety of methods employed in existing studies. In an attempt to better understand ANH a number of authors have used simple spreadsheet-based simulations of the technique to determine how many units of autologous blood need to be removed in order to save allogeneic blood. Most models assumed ANH was to be performed in an average 70-kilogram adult with an estimated blood volume of 5000 ml. Key variables used in the

models included the hematocrit (a measure of the amount of hemoglobin in the subject's blood) before dilution (H_i), the minimum hematocrit to which the subject could be safely diluted (H_m), the volume of autologous blood withdrawn during ANH (V_{ANH}), and the surgical blood loss (SBL). While the assumptions used in the models differed somewhat, all reached similar conclusions. Brecher first estimated that in an individual with a normal H_i of 0.400 the maximum possible saving in allogeneic blood was 0.6 units (28). To achieve this small saving in blood exposure the H_m would have to be reduced to 0.250 requiring a V_{ANH} of 2500 ml - half of the subject's blood volume. As a reference, the National Institutes of Health suggests that the lowest H_m tolerated in most healthy surgical patients is 0.21. Feldman found that increasing V_{ANH} and tolerating much lower Hms could increase savings of allogeneic blood. By removing 4500 ml of autologous blood and tolerating a H_m as low as 0.15 Feldman estimated savings of up to 3 units of allogeneic blood (27). Weiskopf further suggested benefit of ANH if aggressive reductions of H_m were tolerated. The volume of SBL tolerated before allogeneic transfusion was needed could be increased from 4900 to 8300 ml if H_m as low as 0.150 could be achieved (88). All of the preceding models concluded H_i must be in excess of 0.300 and that a V_{ANH} of one or two units (450-900 ml) is of little or no benefit. Kick found that if a H_m of 0.220 were tolerated, a smaller V_{ANH} (900 ml) would yield savings of a single unit of allogeneic blood only if SBL was greater than 3000 ml (89).

Subgroup analyses based on the mean volumes of V_{ANH} and SBL reported in trials included in this overview, 989 ± 443 ml and 1211 ± 519 ml respectively, were undertaken to determine the influence of these factors on the efficacy of ANH. Cut-off values for these subgroups were set at 1000 ml for ease of interpretation. V_{ANH} did not

appear to influence the reported efficacy of ANH (Figures 7 a-d). A trend toward a reduced likelihood of transfusion in trials in which V_{ANH} exceeded 1000 ml did reach statistical significance when compared to trials reporting smaller V_{ANH}. SBL had a similarly weak influence on the efficacy of ANH, although the reduction in the volume of allogeneic blood transfused was greater in those trials with greater than 1000 ml blood loss. Heterogeneity persisted in most of these subgroup analyses.

Given the importance of V_{ANH} and SBL on the modeled efficacy of ANH why did the subgroup analyses based on these variables not influence outcome? The lowest H_m that was reported in trials identified in this overview was 0.250. Mathematical models suggest that hemodilution to this level is associated with a single unit reduction in allogeneic transfusion at best. Similarly, the mean V_{ANH} and SBL reported (989 ± 443 ml and 1211 ± 519 ml, respectively) were substantially lower than the values predicted to be of benefit by mathematical modeling. The use of mean values to subdivide groups was somewhat arbitrary and reflected the data available. In light of the larger volumes of V_{ANH} and SBL that are suggested by mathematical models these cutpoints might have been inadequate to define differences in efficacy. However use of these larger values for subgroup analysis was impractical as only one trial reported a mean V_{ANH} greater than 2000 ml(52) while two reported SBL in excess of 2000 ml(56,66). It is evident that published ANH research falls well short of the conditions predicted by mathematical models that are required to yield significant allogeneic blood savings. Given these findings it is surprising indeed that when all trials of ANH were pooled, significant reductions in both the likelihood of transfusion and the volume of blood transfused were

seen. What then was responsible for the apparent efficacy of ANH reported in the literature?

4.9.4. Trial Methods in ANH Research

Several concerns have been raised regarding the validity of the literature supporting the use of ANH. Stehling, in a review advocating the use of ANH, admits, "carefully controlled trials of ANH are few and far between (26). Uncertainty regarding the quality of the original research may undermine the most rigorously conducted meta-analysis. Evaluation of the methodological quality of the included trials is therefore an important component of a well-conducted review.

Quality of the reporting of trial methods was assessed using Jadad's score. A number of scales have been devised for the scoring of the methodological quality of randomized controlled trials. Jadad's three-item scale was chosen for several reasons. First, four of five possible points in Jadad's score are allotted for randomization and blinding, factors shown to be of greatest importance in determining the validity of a trial's results (90). Second, Jadad's score was developed and validated in the anesthesia literature where it was anticipated the bulk of published research on ANH would be identified. And lastly, the simplicity of scoring three items was thought to minimize the likelihood of inter-observer disagreement.

A single point (for randomization) was a basic inclusion criterion for this overview and only five of twenty-four included trials exceeded this value. Scoring of reporting of trial methods was compromised by the consistent absence of blinding in randomized controlled trials of ANH. The removal of several liters of a subject's blood is indeed difficult to conceal and was understandably not performed by investigators

evaluating ANH. As a result most trials did poorly on the Jadad score. Jadad has suggested that only trials scoring 3 or more should be considered to be of good methodological quality. In this systematic overview no trial was given a score associated with good quality. Indeed, only five of twenty-four trials were assigned a Jadad score greater than one.

As may be seen in Figures 10 a-d, the likelihoods of allogeneic blood exposure and reductions in the number of units transfused were similar whether the Jadad score was less than or greater than one. The ability of this score to identify differences in outcome was reduced by the consistent lack of blinding and the failure of any trial in this overview to reach a Jadad score consistent with "good" trial methods.

Another element of trial design, the use of a protocol to guide transfusion decisions, was an important determinant of the reported efficacy of ANH. As mentioned earlier, blinding was not employed in any of the trials identified in this overview. Frequency and volume of allogeneic transfusion were outcome measures used in this meta-analysis and the trials it included. As such, the indications for transfusion should be carefully controlled if valid conclusions regarding the efficacy of ANH are to be drawn. Transfusion decisions left to the discretion of the investigators may be subject to their biases.

Trials in this overview that employed a protocol to define the indications for transfusion failed to yield a clinically or statistically significant reduction in allogeneic transfusion (Figures 11 c and d). Statistical heterogeneity was absent in analyses of both the likelihood of transfusion and the number of units transfused in this subgroup despite the variety of surgical procedures, languages of publication, V_{ANH}, SBL, and Jadad scores

included. The lack of both clinical effect and heterogeneity in these trials stands in marked contrast to those trials without transfusion protocols that showed dramatic and heterogeneous reductions in both the likelihood of transfusion and the number of units transfused (Figures 11 a and b).

These results suggest that statistical heterogeneity identified in this overview may be secondary to differences in the transfusion practices employed in the included trials. While the proportion of the heterogeneity attributable to transfusion practice cannot be determined it is apparent that a diverse subgroup of trials, similar only in their use of a transfusion protocol, yielded a homogeneous estimate of the efficacy of ANH. Similarly, the failure of trials reporting a protocol to guide transfusion practice to demonstrate clinically or statistically significant reductions in allogeneic blood exposure suggests that the reported efficacy of ANH may be secondary to biased trial design or transfusion decision making. The clinical importance of these findings is clear and their influence on future research will be discussed in subsequent sections.

5. IMPLICATIONS OF STUDY RESULTS

5.1 Meta-Analysis and Evidence Based Medicine

A frequent question raised regarding systematic review and meta-analysis is how the clinician should apply the findings. Critics of meta-analysis suggest that "pooled results incorporate the biases of individual studies and embody new sources of bias, mostly because of the selection of studies and the inevitable heterogeneity among them (91)." Others would "argue that the results of multiple small trials are more applicable to the practice of medicine in which the encountered patients are rarely homogeneous (77)."

In preparing clinical practice guidelines for anti-thrombotic agents Cook and colleagues developed a hierarchy of evidence to weigh the relative strength of evidence supporting a given medical technology (92). It is suggested that results from meta-analyses be given the same or greater value than a single randomized controlled trial if their estimates of treatment effect are consistent and beyond the minimal clinically important benefit (MCIB). Using Cook's approach, what recommendation can be made regarding the use of ANH?

There could be little doubt that the results of individual studies evaluating ANH are widely disparate, as evidenced by the significant heterogeneity identified throughout this systematic overview. Defining the MCIB is somewhat more difficult as no consensus has emerged from the literature to suggest the MCIB for a technology reducing allogeneic blood transfusion. A review of the summary results of this meta-analysis would suggest that ANH may result in a one third reduction in the relative risk of transfusion (RR = 0.67, 95% CI 0.50 to 0.92) and the savings of two units of allogeneic blood (WMD = -2.37, 95% CI –3.74 to –1.01). If we restrict our attention to only

properly designed trials employing a protocol to guide transfusion decisions, neither a statistically nor clinically significant benefit of ANH was identified (RR 0.92, 95% CI 0.76 to 1.10 and WMD = -0.32, 95% CI –0.94 to 0.30).

If the goal is the reduction of allogeneic blood exposure at all costs, as is the case with Jehovah's Witnesses who refuse allogeneic blood transfusion on religious grounds, any benefit demonstrated by ANH may be acceptable. Similarly, for patients wishing to do everything to avoid transfusion and who are able to withstand the anemia and fluid loads associated with ANH, the small benefits of the technique may be of value. However, based on the best quality evidence, it is difficult to support the widespread use of ANH in the perioperative period.

5.2. Meta-Analysis vs Large Randomized Controlled Trials

Large RCTs are widely considered the gold standard in the evaluation of a clinical intervention. It can be argued that a large RCT of hemodilution is required to determine the roll of ANH in clinical practice. Can the results of this systematic overview be considered representative of a large, well conducted RCT?

Comparison of meta-analyses and the results of large co-operative RCTs have yielded conflicting results. Chalmers compared the results of several multi-center trials with meta-analyses including multiple undersized previously published RCTs (93). Meta-analysis of twelve studies evaluating the use of β-blockers following acute myocardial infarction yielded results clinically and statistically similar to the MIAMI and ISIS-1 studies. Similarly, systematic review of eleven thrombolytic trials yielded mortality rates similar to the GISSI trial but showed both quantitative and qualitative differences in the estimate of a secondary outcome, reinfarction.

Several investigators have attempted to systematically compare the results of meta-analyses to large RCTs. Villar compared the results of 30 systematic reviews selected from the Pregnancy and Childbirth Module of the Cochrane database that contained at least one trial enrolling 1000 or more subjects (94). The large trial was then removed from the meta-analysis the summary odds ratios recalculated. The result of the meta-analysis agreed in direction of treatment effect reported in the large RCT in 24 of 30 comparisons (80%) but agreed in both direction and statistical significance in only 18 (60%). Cappelleri also identified systematic reviews from the Cochrane database and added reviews found using Medline (95). All eligible reviews contained a trial enrolling greater than 1000 subjects. The large trial was again withdrawn and the summary odds ratio recalculated and compared to the results of the large RCT. The results of the meta-analysis agreed with those of the large trial in 71 of 79 comparisons (90%). Finally Le Lorier identified trials enrolling 1000 or more subjects published in four influential medical journals between 1991 and 1994 (91). Systematic reviews evaluating the same topic that were published before the large RCT were then identified from Medline. Agreement between the results of the systematic review and the subsequent large trial was reached in only 27 of 40 comparisons (68%).

A recent review attempts to reconcile the disparate findings of the previous three studies (96). Ioannidis suggests that "comparisons of large trials with meta-analyses may reach different conclusions depending on how trials and meta-analyses are selected and how end points and agreement are defined." A relatively consistent finding is that disagreement beyond chance of between 10% and 23% occurs when meta-analyses and large trials are compared. Sources of disagreement include different baseline risk of the

primary outcome, differences in protocol, methodological quality, and publication bias. Indeed, these factors had a large influence on the results of this systematic review. If ANH is to gain acceptance, new, well-conducted RCTs confirming its benefits are required.

5.3. Anemia and Safety

Individuals able to tolerate the relatively extreme conditions required to yield allogeneic blood savings may still benefit from ANH. However, these patients must balance the anticipated benefits of ANH against the risks associated with the anemia and fluid loads associated with the technique. For these selected patients the appropriate question may then be - what is known about the risks of anemia and the safety of ANH? A detailed description of the physiologic response to acute anemia is beyond the scope of this document; for an in depth discussion of this topic the reader is encouraged to consult Hébert's recent review (97). Evidence from both laboratory and clinical research suggests that anemia in healthy individuals is remarkably well tolerated. Tissue oxygenation is maintained with hematocrits of 0.18 - 0.25 and the myocardium does not begin anaerobic metabolism until the hematocrit is further reduced to 0.15 - 0.20 (19). A selective review of 54 publications involving Jehovah's Witnesses revealed that no patient with a hematocrit of 0.15 - 0.24 died as a result of their anemia (98). A recent study of critically ill patients showed no difference in morbidity or mortality when were they randomly assigned to transfusion thresholds of hematocrits 0.21 - 0.27 or 0.30 - 0.36 (99). Similarly a recent randomized controlled trial in patients suffering hip fracture compared a transfusion policy designed to maintain a hematocrit >0.30 with transfusion decisions based on the presence of symptoms (100). Patients in the symptom-based

transfusion group received significant less allogeneic blood but showed no differences in morbidity or functional status 60 days following surgery. A trend towards increased 60 day mortality in the symptom-based transfusion group did not reach statistical significance (RR = 2.5; 95% CI, 0.5-12.2).

Despite research supporting the relative safety of moderate anemia, concern remains that anemia is detrimental to a number of patients and will increase their risk of perioperative cardiac morbidity. Weisel observed that patients undergoing coronary artery bypass grafting with hematocrits of less than 0.30 experienced a delay in myocardial metabolic recovery (101). Analysis of a retrospective cohort of 1958 adult Jehovah's Witnesses revealed that in patients with cardiovascular disease mortality increased exponentially with declining hemoglobin. Thirty-day mortality in patients with preoperative hematocrits of 0.36 or more was 1.3% compared to 33.3% in those patients with hematocrits of 0.18 or less. Mortality in those patients with pre-existing cardiovascular disease was 4.3 times higher than in their healthy counterparts (102). Identification of patients at risk for cardiovascular disease is notoriously difficult. Fifteen percent of an unselected group of surgical patients experienced silent myocardial ischemia preoperatively; thirty-five percent of these had no risk factors for cardiovascular disease identified on preoperative assessment (103).

The anemia essential for the successful application of ANH may therefore place a number of patients at risk for perioperative cardiovascular morbidity and mortality. Given that the risks of contracting HIV and HCV from an allogeneic transfusion are estimated at 1 : 676,000 and 1 : 103,000 respectively some authors are questioning the lengths that patients should be pushed in order to reduce allogeneic blood exposure (3).

Gillon summarized these concerns stating that "any reduction in the use of allogeneic blood must be assumed to be intrinsically beneficial...so long as the technique itself or the avoidance of transfusion does not introduce a greater risk (29)."

Systematic overview and meta-analysis has been described, perhaps a bit melodramatically, as "a tower of statistical power that allows researchers to rise above the body of evidence, survey the landscape, and map out future directions (104)." The increased statistical power is "particularly relevant to conditions of relatively low event rates or when small effects are being assessed (30)," as would be the case with adverse events. Regrettably all such advantages of meta-analysis are negated if adverse events are not reported.

This review identified only six trials enrolling a total of 170 subjects that specifically reported mortality. Cardiovascular and thromboembolic complications were identified in even fewer trials – two each. Although no increase in adverse events was noted in patients undergoing ANH, the poor reporting of these outcomes suggests that the results of the previously published trials cannot be used to establish the safety of this technique.

5.4. ANH and Other Technologies Reducing Allogeneic Transfusion

Based on current evidence the benefits of ANH are somewhat questionable and the risks are poorly understood. What role should ANH play, if any, among other technologies currently available to reduce allogeneic transfusion?

In Canada, preoperative autologous donation (PAD) is the most commonly used means of providing autologous blood for patients undergoing elective surgery. To date there have been few head-to-head comparisons of PAD and ANH. Ness compared ANH

and PAD in prospective randomized controlled trial enrolling 50 patients undergoing retropubic prostatectomy. None of the 25 patients in the hemodilution group required allogeneic transfusion compared with 2 in the predonation group (105). In a retrospective case-control study Monk found that the likelihood of exposure to allogeneic blood and volume of allogeneic blood transfused were similar in those patients undergoing either ANH or PAD (106). Unfortunately a number of the patients in the ANH group also predonated in the weeks prior to surgery raising the possibility of bias introduced through cointervention.

Available evidence does suggest that ANH is, at the very least, a less costly means of providing autologous blood than PAD. In Monk's study the total costs associated with the collection and transfusion of all blood products was 40% lower in the ANH group (106). Similarly Roberts reported that provision of 2 units of autologous blood through ANH was 55% less costly than 2 units collected through their PAD program (107).

Given the paucity of evidence supporting the safety and efficacy of ANH, it remains to be seen if ANH offers a useful, let alone cost-effective, alternative to PAD in the provision of autologous blood for patients undergoing surgery. Comparisons of ANH to other technologies to minimize perioperative blood exposure such as perioperative cell salvage, recombinant human erythropoietin, and antifibrinolytics have not yet been performed. Until prospective and properly conducted comparisons of ANH and other competing technologies have been performed the role of ANH will remain uncertain.

5.5. Implications for Future Research

The results of this systematic overview offer valuable insights into the proper design and conduct of future research. Future trials must address the significant concerns regarding bias in ANH research. Where possible, every effort should be made to blind both subjects and investigators. At a minimum, the indications for transfusion of both autologous and allogeneic blood products must be clearly defined and adherence to the protocol demonstrated. While blinding of ANH is difficult, it should be possible to have transfusion decisions made by a blinded third party not directly involved in the bedside care of the subject. Adverse events must be carefully sought and reliably reported.

Mathematical models indicate that there may be a limited role for ANH in reducing perioperative allogeneic blood transfusion. Future randomized controlled trials must incorporate the findings of these models in their design if ANH is to be of proven value. Subjects must present for surgery with hematocrits greater than 0.30 and have the physiologic reserve to tolerate the anemia and fluid shifts associated with the technique. V_{ANH} must be substantial, in excess of 2000 ml, for the technique to be of significant benefit. Anticipated SBL must also be substantial and in excess of the planned phlebotomy. Lesser volumes of SBL would be easily tolerated by passive dilutional anemia.

ANH should not receive wider clinical use until properly designed and conducted trials are able to demonstrate both the efficacy and the safety of the technique. Only if evidence supporting the use of ANH is available should its value relative to other technologies be investigated.

REFERNCES

1. Wick MR. Blood transfusion: an historic and current perspective. Am J Clin Pathol 1989;91(Suppl 1):S2-S6

2. Expert Working Group. Guidelines for red blood cell and plasma transfusion for adults and children. Can Med Assoc J 1997;156(11(Suppl)):S1-S24

3. Schreiber GB, Busch MP, Kleinman SH, Korelitz JJ. The risk of transfusion-transmitted viral infections. N Engl J Med 1996;334(26):1685-90.

4. Causes of death, 1994. Ottawa. Statistics Canada. 1996; Cat No 84-208-XPB.

5. Linden JV, Paul B, Dressler KP. A report of 104 transfusion errors in New York State. Transfusion 1992;32:601-6.

6. Bordin JO, Heddle NM, Blajchman MA. Biologic effects of leukocytes present in transfused cellular blood products. Blood 1994;84(6):1703-21.

7. McAlister FA, Clark HD, Wells PS, Laupacis A. Perioperative allogeneic blood transfusion does not cause adverse sequelae in patients with cancer: a meta-analysis of unconfounded studies. Br J Surg 1998;85:171-8.

8. van de Watering LMG, Hermans J, Houbiers JGA, van den Broek PJ, Bouter H, Boer F, Harvey MS, Huysmans HA, Brand A. Beneficial effects of leukocyte depletion of transfused blood on postoperative complications in patients undergoing cardiac surgery. A randomized clinical trial. Circulation 1998;97:562-8.

9. Jensen LS, Hokland M, Nielsen HJ. A randomized controlled study of the effect of bedside leucocyte depletion on the immunosuppressive effect of whole blood transfusion in patients undergoing elective colorectal surgery. Br J Surg 1996;83(7):973-7.

10. Krever H. Commission of Inquiry on the Blood System in Canada: Final Report. Ottawa: Canadian Government Publishing; 1997; 40, The blood system of the future. p. 1052-3.

11. Canadian Red Cross Society. National blood program. Summary of Activities. 1996; Ottawa: Canadian Red Cross Society.

12. Forgie MA, Wells PS, Laupacis A, Fergusson D. Preoperative autologous donation decreases allogeneic transfusion but increases exposure to all red blood cell transfusion. Arch Intern Med 1998;158:610-6.

13. Tretiak R, Laupacis A, Rivière M, McKerracher K, Souêtre E, Canadian Cost of Transfusion Study Group. Cost of allogeneic and autologous blood transfusion in Canada. Can Med Assoc J 1996;154(10):1510-08.

14. Etchason J, Petz L, Keeler E, Calhoun L, Kleinman S, Snider C, Fink A, and Brook R. The cost effectiveness of preoperative autologous blood donations. N Engl J Med 1995;332(11):719-24.

15. Drummond M, Torrance G, and Mason J. Cost-effectiveness league tables: more harm than good? Soc Sci Med 1993;37(1):33-40.

16. Desmond MJ, Thomas MJG, Gillon J, Fox MA. Consensus conference on autologous transfusion. Perioperative red cell salvage. Transfusion 1996;36(7):644-51.

17. Laupacis A, Fergusson D, International Study of Perioperative Transfusion. Drugs to minimize perioperative blood loss in cardiac surgery: meta-analysis using perioperative blood transfusion as the outcome. Anesth Analg 1997;85(6):1258-67.

18. Laupacis A, Fergusson D, International Study of Perioperative Transfusion. Erythropoietin to minimize peri-operative blood transfusion: a systematic review of randomized. Transfusion Medicine 1998;In Press.

19. Anonymous. Practice guidelines for blood component therapy: A report by the American Society of Anesthesiologists Task Force on Blood Component Therapy. Anesthesiology 1996;84(3):732-47.

20. Spence RK. Surgical red blood cell transfusion practice policies. Blood Management Practice Guidelines Conference. Am J Surg 1995;170(6A Suppl):3S-15S.

21. Messmer K. Hemodilution - possibilities and safety aspects. Acta Anaesthesiol Scand 1988;32(Supp 89):49-53.

22. Hallowell P. Transfusion of fresh autologous blood in open-heart surgery. A method for reducing blood bank requirements. J Thorac Cardiovasc Surg 1972;64:941-8.

23. Laks H, Pilon RN, Klovekorn WP, Anderson W, MacCallum JR, and O'Connor NE. Acute hemodilution: its effects on hemodynamics and oxygen transport in anesthetized man. Ann Surg 1974;180(1):103-9.

24. Martin E, Hansen E, Peter K. Acute limited normovolemic hemodilution: A method for avoiding homologous blood transfusion. World J Surg 1987;11(1):53-9.

25. The use of technologies to decrease perioperative allogeneic blood transfusion: an international survey of practice. 1997 Oct 08; Groningen, Netherlands: 1997; 22 p.

26. Stehling L, Zauder HL. Controversies in transfusion medicine. Perioperative hemodilution: Pro. Transfusion 1994;34(3):265-8.

27. Feldman JM, Roth JV, Bjoraker DG. Maximum blood savings by acute normovolemic hemodilution. Anesth Analg 1995;80(1):108-13.

28. Brecher ME, Rosenfeld M. Mathematical and computer modeling of acute normovolemic hemodilution. Transfusion 1994;34(2):176-9.

29. Gillon J. Controversies in transfusion medicine. Acute normovolemic hemodilution in elective major surgery: Con. Transfusion 1994;34(3):269-71.

30. Mulrow CD. Rationale for systematic overviews. BMJ 1994;309(6954):597-9.

31. Haynes RB, Wilczynski N, McKibbon KA, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. J Am Med Inform Assoc 1994;1(6):447-58.

32. Dickersin K, Scherer R, and Lefebvre C. Identifying relevant studies for systematic reviews. BMJ 1994;3091286-91.

33. Jadad AR, Moore A, Carroll D, Jenkinson C, Reynolds JM, Gavaghan DJ, McQuay DM. Assessing the quality of reports of randomized controlled trials: Is blinding necessary? Control Clin Trials 1996;17(1):1-12.

34. Bracken MB. Sinclair JC, Bracken MB, editors.Effective care of the newborn infant. Oxford: Oxford University Press; 1992; 2, Statistical methods for analysis of effects of treatment in overviews of randomized trials. p. 13-8.

35. DerSimonian R and Laird NM. Meta-analysis in clinical trials. Control Clin Trials 1986;7177-88.

36. The Cochrane Review Manager [computer program]. Cochrane Collaboration. 3.0.1. 1997;

37. Whitehead A, Jones NMB. A meta-analysis of clinical trials involving different classifications of response into ordered categories. Stat Med 1994;13(23/4):2503-15.

38. Light RJ, Pillemer DB. Quantitative procedures. Summing up: The science of reviewing research. Cambridge, MA: Harvard University Press; 1984;.

39. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315(7109):629-34.

40. Ahlberg A, Nillius A, Rosberg B, Wulff K. Preoperative normovolemic hemodilution in total hip arthroplasty. A clinical study. Acta Chir Scand 1977;143(7-8):407-11.

41. Atallah MM, Abdelbaky SM, Saied MM. Does timing of hemodilution influence the stress response and overall outcome? Anesth Analg 1993;76(1):113-7.

42. Bennett SR. Perioperative autologous blood transfusion in elective total hip prosthesis operations. Ann R Coll Surg Engl 1994;76(2):95-8.

43. Boldt J, Kling D, Weidler B, Zickmann B, Herold C, Dapper F, Hempelmann G. Acute preoperative hemodilution in cardiac surgery: volume replacement with a hypertonic saline-hydroxyethyl starch solution. J Cardiothorac Vasc Anesth 1991;5(1):23-8.

44. Bonnet MC, Julia JM, Mathieu-Daude JC, du Cailar J. Value of hemodilution in maxillofacial surgery for postoperative traumatic edema and graft viability. Ann Fr Anaesth Reanim 1986;5(3):243-8.

45. Dietrich W, Barankay A, Dilthey G, Mitto HP, Richter JA. Reduction of blood utilization during myocardial revascularization. J Thorac Cardiovasc Surg 1989;97(2):213-9.

46. Lawson NW, Ochsner JL, Mills NL, Leonard GL. The use of hemodilution and fresh autologous blood in open-heart surgery. Anesth Analg 1974;53(5):672-83.

47. Lilleaasen P. Moderate and extreme haemodilution in open-heart surgery. Blood requirements, bleeding and platelet counts. Scand J Thorac Cardiovasc Surg 1977;11(2):97-103.

48. Lorentz A, Osswald PM, Schilling M, Jani L. A comparison of autologous transfusion procedures in hip surgery. Anaesthesist 1991;40(4):205-13.

49. Malinovsky JM, Bouyer L, Rusterholtz T, Lepage JY, de Dieuleveult C, Cozian A, Pinaud M, Souron R. Saving of blood products by normovolemic hemodilution in prostate resection. Ann Fr Anaesth Reanim 1989;8 Suppl:R120

50. Moyes DG, Mistry BD, Conlan AA. Normovolaemic haemodilution using dextran 70 in thoracic surgery. S Afr Med J 1985;67(19):762-4.

51. Rosberg B. Blood coagulation during and after normovolemic hemodilution in elective surgery. Ann Clin Res 1981;13 Suppl 33:84-8.

52. Rose D, Coutsoftides T. Intraoperative normovolemic hemodilution. J Surg Res 1981;31(5):375-81.

53. Triulzi DJ, Gilmor GD, Ness PM, Baumgartner WA, Schultheis LW. Efficacy of autologous fresh whole blood or platelet-rich plasma in adult cardiac surgery. Transfusion 1995;35(8):627-34.

54. Van der Linden P, Wathieu M, Gilbart E, Engelman E, Wautrecht JC, Lenaers A, Vincent JL. Cardiovascular effects of moderate normovolaemic haemodilution during enflurane-nitrous oxide anaesthesia in man. Acta Anaesthesiol Scand 1994;38(5):490-8.

55. Vara Thorbeck R, Rosell Pradas J, Mekinassi KL, Prados Olleta N, Guerrero Fernandez-Marcote JA. Prevention of thromboembolic disease and post-transfusional complications using normovolemic hemodilution in arthroplasty surgery of the hip. Rev Chir Orthop Reparatrice Appar Mot 1990;76(4):267-71.

56. von Bormann B, Weidler B, Boldt J, Jooss D, Aigner K, Peil J, Hempelmann G. Acute normovolemic hemodilution in extensive surgical interventions. Chirurg 1986;57(7):457-64.

57. Welch M, Knight DG, Carr HM, Smyth JV, Walker MG. The preservation of renal function by isovolemic hemodilution during aortic operations. J Vasc Surg 1993;18(5):858-66.

58. Herregods L, Foubert L, Moerman A, Francois K, Rolly G. Comparative study of limited normovolaemic haemodilution in patients with left main coronary artery stenosis. Anaesthesia 1995;50(11):950-3.

59. Kochamba GS, Pfeffer TA, Sintek CF, Khonsari S. Intraoperative autotransfusion reduces blood loss after cardiopulmonary bypass. Ann Thorac Surg 1996;61:900-3.

60. Ferraris VA, Berry WR, Klingman RR. Comparison of blood reinfusion techniques used during coronary artery bypass grafting. Ann Thorac Surg 1993;56:433-40.

61. Kaplan JA, Cannarella C, Jones EL, Kutner EL, Hatcher CR, Dunbar RW. Autologous blood transfusion during cardiac surgery. A reevaluation of three methods. J Thorac Cardiovasc Surg 1977;74(1):4-10.

62. Ochsner JL. Fresh autologous blood transfusion with extracorporeal circulation. Ann Surg 1973;117(6):811-.

63. Pliam MB. Failure of transfusion of autologous whole blood to reduce banked-blood requirements in open-heart surgical patients. J Thorac Cardiovasc Surg 1975;70(2):338-43.

64. Vedrinne C, Girard C, Jegaden O, Blanc P, Bouvier H, Ffrench P, Mikaeloff P, Estanove S. Reduction in blood loss and blood use after cardiopulmonary bypass with high-dose aprotinin versus autologous fresh whole blood transfusion. J Cardiothorac Vasc Anesth 1992;6(3):319-23.

65. Catoire P, Saada M, Liu N, Delaunay L, Rauss A, Bonnet F. Effect of peroperative normovolemic hemodilution on left ventricular segmental wall motion during abdominal aortic surgery. Anesth Analg 1992;75:654-9.

66. Dale J, Lilleaasen P, Erikssen J. Hemostasis after open heart surgery with extreme or moderate hemodilution. Eur Surg Res 1987;19:339-47.

67. Smith B.J., Darzins P.J., Quinn M., and Heller R.F. Modern methods of searching the medical literature. Med J Aust 1992;157(9):603-11.

68. Ramos-Remus C, Suarez-Almazor M, Dorgan M, Gomez-Vargas A, and Russell AS. Performance of online biomedical databases in rheumatology. J Rheumatol 1994;21(10):1912-21.

69. Dickersin K, Chan S, Chalmers TC, Sacks HS, Smith Jr H. Publication bias and controlled clinical trials. Control Clin Trials 1987;8:343-53.

70. Dickersin K. The existence of publication bias and risk factors for its occurence. JAMA 1990;263(10):1385-9.

71. Simes RJ. Publication bias: the case for an international registry of clinical trials. J Clin Oncol 1986;4:1529-41.

72. Easterbrook PJ, Berlin JA, Gopalan T. Publication bias in clinical research. Lancet 1991;337(8746):867-72.

73. Villar J, Piaggio G, Carroli G, Donner A. Factors affecting the comparability of meta-analyses and largest trials results in perinatology. J Clin Epidemiol 1997;50(9):997-1002.

74. Rosenthal R. The "file drawer problem" and tolerance for null results. Psych Bull 1979;86(638):641

75. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1998;50:1088-101.

76. Felson DT. Bias in meta-analytic research. J Clin Epidemiol 1992;45(8):885-92.

77. Chalmers TC. Problems induced by meta-analyses. Stat Med 1991;10971-80.

78. Berlin JA. Does blinding of readers affect the results of meta-analyses? Lancet 1997;350(9072):185-6.

79. Mulrow CD, Oxman AD, editors.The Cochrane Library (database on disk and CDROM). Oxford: The Cochrane Collaboration; 1997; 8, Analysing and presenting results. Cochrane Collaboration Handbook (updated 1 March 1997). p. 53-73.

80. Follman D, Elliott P, Suh I, Cutler J. Variance imputation for overviews of clinical trials with continuous response variables. J Clin Epidemiol 1992;45(7):769-33.

81. Thompson SG. Why sources of heterogeneity in meta-analysis should be investigated. BMJ 1994;309(6965):1351-5.

82. Whitehead A, Whitehead J. A general parametric approach to the meta-analysis of randomized controlled trials. Stat Med 1991;10(11):1665-77.

83. Thompson SG, Pocock SJ. Can meta-analyses be trusted? Lancet 1991;338(8775):1127-30.

84. Berlin JA, Laird NM, Sacks HS, and Chalmers TC. A comparison of statistical methods for combining event rates from clinical trials. Stat Med 1989;8141-51.

85. Moher D, Fortin P, Jadad AR, Juni P, Klassen T, Le Lorier J, Liberati A, Linde K, Penna A. Completeness of reporting of trials published in languages other than English: implications for conduct and reporting of systematic reviews. Lancet 1996;347(8998):363-6.

86. Grégoire G, Derderian F, Le Lorier J. Selecting the language of the publications included in a meta-analysis: Is there a Tower of Babel bias? J Clin Epidemiol 1995;48(1):159-63.

87. Egger M, Zellweger-Zahner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. Lancet 1997;350(9074):326-9.

88. Weiskopf RB. Mathematical analysis of isovolemic hemodilution indicates that it can decrease the need for allogeneic blood transfusion. Transfusion 1995;35(1):37-41.

89. Kick O, Daniel E. Mathematical considerations in the practice of acute normovolemic hemodilution. Transfusion 1997;37(2):141-3.

90. Schulz KF, Chalmers I, Hayes RJ, and Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 1995;273(5):408-12.

91. Le Lorier J, Grégoire G, Benhaddad A, Lapierre J, Derderian F. Discrepancies between meta-analyses and subsequent large, randomized, controlled trials. N Engl J Med 1997;337(8):536-42.

92. Cook DJ, Guyatt GH, Laupacis A, Sackett DL, Goldberg RJ. Clinical recommendations using levels of evidence for antithrombotic agents. Chest 1995;108(4 Suppl):227S-30S.

93. Chalmers TC, Levin H, Sacks HS, Reitman D, Berrier J, and Nagalingam R. Meta-analysis of clinical trials as a scientific discipline. I: Control of bias and comparison with large co-operative trials. Stat Med 1987;6315-25.

94. Villar J, Carroli G, Belizán JM. Predictive ability of meta-analyses of randomized controlled trials. Lancet 1995;345:772-6.

95. Cappelleri JC, Ioannidis JPA, Schmid CH, de Ferranti SD, Aubert M, Chalmers TC, Lau J. Large trials vs meta-analysis of smaller trials. How do their results compare? JAMA 1996;276(16):1332-8.

96. Ioannidis JPA, Cappelleri JC, Lau J. Issues in comparisons between meta-analyses and large trials. JAMA 1998;279(14):1089-93.

97. Hébert PC, Ling QH, Biro GP. Review of the physiologic mechanisms in response to anemia. Can Med Assoc J 1997;156(11(Suppl)):S27-S40

98. Viele MK, Weiskopf RB. What can we learn about the need for transfusion from patients who refuse blood? The experience with Jehovah's Witnesses. Transfusion 1994;34(5):396-401.

99. Hébert PC, Wells G, Tweeddale M, Martin C, Marshall J, Pham B, Blajchman M, Schweitzer I, Pagliarello G. Does transfusion practice affect mortality in critically ill patients? Transfusion Requirements in Critical Care (TRICC) Investigators and the Canadian Critical Care Trials Group. Am J Respir Crit Care Med 1997;155(5):1618-23.

100. Carson JL, Terrin ML, Barton FB, Aaron R, Greenburg AG, Heck DA, Magaziner J, Merlino FE, Bunce G, McClelland B, et al. A pilot randomized trial comparing symptomatic vs. hemoglobin-level- driven red blood cell transfusions following hip fracture. Transfusion 1998;38(6):522-9.

101. Weisel RD, Charlesworth DC, Mickleborough LL, Fremes SE, Ivanov J, Mickle DAG, Teasdale SJ, Glynn MFX, Scully HE, Goldman BS, et al. Limitations of blood conservation. J Thorac Cardiovasc Surg 1984;88(1):26-38.

102. Carson JL, Duff A, Poses RM, Berlin JA, Spence RK, Trout R, Noveck H, and Strom BL. Effect of anemia and cardiovascular disease on surgical mortality and morbidity. Lancet 1996;348(9034):1055-60.

103. Muir AD, Reeder MK, Föex P, Ormerod OJM, Sear JW, Johnston C. Preoperative silent myocardial ischaemia: incidence and predictors in a general surgical population. Br J Anaesth 1991;67(4):373-7.

104. Gelber RD, Goldhirsch A. Meta-analysis: the fashion of summing up. Ann Oncol 1991;2:461-8.

105. Ness PM, Bourke DL, Walsh PC. A randomized trial of perioperative hemodilution versus transfusion of preoperatively deposited autologous blood in elective surgery. Transfusion 1992;32(3):226-30.

106. Monk TG, Goonough JD, Birkmeyer ME, Brecher ME, Catalona WJ. Acute normovolemic hemodilution is a cost-effective alternative to preoperative autologous blood donation by patients undergoing radical prostatectomy. Transfusion 1995;35(7):559-65.

107. Roberts WA, Kirkley SA, and Newby M. A cost comparison of allogeneic and preoperatively or intraoperatively donated autologous blood. Anesth Analg 1996;83(1):129-33.

APPENDIX A

MEDLINE SEARCH STRATEGIES

Acute Normovolemic Hemodilution

1. hemodilution/ all sub-headings in MeSH
2. (hemodilution or haemodilution). tw.
3. 1 or 2
4. exp blood transfusion/ all sub-headings in MeSH
5. exp hemorrhage/ all sub-headings in MeSH
6. transfusion$. tw.
7. bleed$. tw.
8. blood loss$. tw.
9. hemorrhag$. tw.
10. or/ 4-9
11. animal/ not (human/ and animal/)
12. not 11
13. 3 and 12

McMaster RCT Search Filter (For MEDLINE 1991 - Present)

1. clinical trial. pt.
2. randomized controlled trial. pt
3. tu. fs.
4. dt. fs.
5. random$. tw.
6. (double adj blind$). tw.
7. placebo$.tw
8. or /1-7

McMaster RCT Search Filter (For MEDLINE 1966 - 1990)

1. clinical trial. pt.
2. randomized controlled trial. pt
3. tu. fs.
4. dt. fs.
5. random$. tw.
6. (double adj blind$). tw.
7. (controlled adj trial). tw
8. placebo$.tw.
9. random allocation/ all sub-headings in MeSH
10. comparative study/ all sub-headings in MeSH
11. or/ 1-10

Cochrane Collaboration RCT Search Filter

1. clinical trial. pt.
2. randomized controlled trial. pt.
3. random allocation/ all sub-headings in MeSH
4. exp clinical clinical trials/ all sub-headings in MeSH
5. exp longitudinal studies/ all sub-headings in MeSH
6. double-blind method/ all sub-headings in MeSH
7. single-blind method/ all sub-headings in MeSH
8. placebos/ all sub-headings in MeSH
9. random$. tw.
10. placebo$. tw.
11. (control or controls or controlled). tw.
12. (double adj blind$). tw.
13. or/ 1-12

EMBASE SEARCH STRATEGY

Acute Normovolemic Hemodilution

1. hemodilution@EX
2. hemodilution@(TI, AB, KWDS)
3. haemodilution@(TI, AB, KWDS)
4. 1,2,or 3
5. transfusion@EX
6. bleeding@EX
7. transfus*@(TI, AB, KWDS)
8. bleed*@(TI, AB, KWDS)
9. blood los*@(TI, AB, KWDS)
10. hemorrhag*@(TI, AB, KWDS)
11. haemorrhag*@(TI, AB, KWDS)
12. 5,6,7,8,9,10,or 11
13. 4 and 12

CURRENT CONTENTS SEARCH STRATEGY

Acute Normovolemic Hemodilution

1. (hemodilution or haemodilution).ab,ti,kw,kp
2. transfusion$.ab,ti,kw,kp
3. hemmorhage. ab,ti,kw,kp
4. bleed$. ab,ti,kw,kp
5. blood loss$. ab,ti,kw,kp
6. or/2-5
7. 1 and 6

APPENDIX B

JADAD SCORE FOR QUALITY OF REPORTING OF TRIAL METHODS

Trial Characteristics	Yes	No	Bonus
Was the study described as randomized? This includes the use of words such as random, randomly, and randomization.			
Was the study described as double-blind?			
Was there a description of withdrawals?			N/A

Total Score = _____

Scoring the items:

Give a score of 1 point for each "yes" and a score of 0 points for each "no". There are no in-between marks.

Give 1 additional point if:

For question 1, the method of randomization was described and was **appropriate** (random numbers table, computer generated schedule, coin toss)

For question 2, the method of double-blinding was described and was **appropriate** (identical placebo, active placebo, dummy, etc.)

Deduct 1 point if

For question 1, the method of randomization was described and was **inappropriate**. (alternating allocation, date of birth, chart number, etc.)

For question 2, The method of double-blinding was described and was **inappropriate** (comparison of injectable and oral preparations without a "double dummy")

APPENDIX B (CONTINUED)

JADAD SCORE FOR QUALITY OF REPORTING OF TRIAL METHODS

Guidelines for Assessment

Randomization

A method to generate the sequence of randomization will be regarded as appropriate if it allowed each study participant to have the same chance of receiving each intervention and the investigators could not predict which intervention was next. Methods of allocation using date of birth, date of admission, hospital numbers, or alternation should not be regarded as appropriate.

Double-blinding

A study must be regarded as double-blind if the word double-blind is used. The method will be regarded as appropriate if it is stated that neither the person doing the assessments nor the study participant could identify the intervention being assessed, or if in the absence of such a statement the use of active placebos, identical placebos, or dummies is mentioned.

Withdrawals and drop-outs

Participants who were included in the study but did not complete the observation period or who were not included in the analysis must be described. The number and the reasons for withdrawal in each group must be stated. If there were no withdrawals, it should be stated in the article. If there is no statement on withdrawals, this item must be given no points.

APPENDIX C. STATISTICAL METHODS

1. Introduction

Meta-analysis attempts to summarize previously published data to determine if there is a meaningful association between a treatment and its desired effect in a number of independent trials. Fleiss (1) suggests that three questions should be asked in such an exercise:

1. Is there evidence that the degree of association, whatever its magnitude, is consistent from one group to another?
2. Assuming that the degree of association is found to be consistent, is the common degree of association statistically significant?
3. Assuming that the common degree of association is significant, what is the best estimate of the common value for the measure of association? What is its standard error? How does one construct a confidence interval for the underlying measure?

All meta-analytic techniques share a similar framework. This discussion will begin with a general approach to meta-analysis and will then highlight the different techniques used in this study.

2. A General Approach to Meta-Analysis

A number of trials (K) comparing a treated group (t) with a control group (c) have been assembled. In the ith trial (i = 1,...K) some measure of association (y_i) between treatment and the outcome of interest has been calculated. At this point, it is unnecessary to state whether the outcome is nominal, continuous or ordinal in nature.

Let w_i represent the weight attached to yi.

$$w_i = \frac{1}{[s.e.(y_i)]^2} \qquad 2.1$$

The standard error ($s.e._i$) represents the precision of the estimate of yi. A weighting scheme based on the squared reciprocal of $s.e._i$ will give more weight to trials with precise estimates with correspondingly small s.e.s and less weight to trials with less precise estimates and larger s.e.s.

To test the hypothesis that a consistent treatment effect exists in trial i let the value of yi = 0 when no association between treatment and outcome exists. When the hypothesis of no association is true the quantity (x_i) has a standard normal distribution,

$$x_i = \frac{y_i}{s.e.(y_i)} = y_i\sqrt{w_i} \qquad 2.2$$

and the quantity x_i^2 has a chi square distribution with one degree of freedom.

$$x_i^2 = w_i y_i^2 \qquad 2.3$$

If the hypothesis of no association is false then the value of x_i^2 is large and the null hypothesis is rejected when the chi squared test is applied.

In meta-analyses the results from the ith trial are not as important as the results from all groups together. The summary analysis begins with calculation of a total chi square.

$$x^2_{total} = \sum_{i=1}^{K} w_i y_i^2 \qquad 2.4$$

The total chi square is partitioned into two components: one representing the association between treatment and outcome (x^2_{assoc}) and another representing the consistency or homogeneity of effect (x^2_{homog}) across K studies.

The overall measure of association (ý) is calculated as the weighted average of the K individual measures with weights defined above.

$$\bar{y} = \frac{\sum_{i=1}^{K} w_i y_i}{\sum_{i=1}^{K} w_i} \qquad 2.5$$

The standard error of the overall estimate of effect may be calculated by

$$s.e(\bar{y}) = \frac{1}{\sqrt{\sum_{i=1}^{K} w_i}} \qquad 2.6$$

The x^2 for association may be calculated by

$$x^2_{assoc} = \frac{\left(\sum_{i=1}^{K} w_i y_i\right)^2}{\sum_{i=1}^{K} w_i} \qquad 2.7$$

Confidence limits on our overall estimate of effect then be constructed by

$$95\%\ CI\ (\bar{y}) = \bar{y} \pm 1.96 \times s.e.(\bar{y}) \qquad 2.8$$

Lastly an estimate of the homogeneity of treatment effect may be calculated by simply subtracting x^2_{assoc} from x^2_{total} or more directly by

$$x^2_{hom\ og} = \sum_{i=1}^{K} w_i (y_i - \bar{y})^2 \qquad 2.9$$

We now have answered the three questions posed earlier. First, the value of x^2_{homog} may be used to determine if the degree of association is consistent from one group to another. Next, statistical significance of the association may be determined using x^2_{assoc}. Last, an overall estimate of the degree of association ($\hat{y}$) with its standard error and 95% confidence limits may be determined.

3. Fixed Versus Random Effects Models

The general approach to meta-analysis described above assumes that all variability in the estimate of treatment effect is present in the studies pooled. Models with this perspective are described as fixed-effects models and effectively limit inferences about the treatment effect to those studies in hand. If an unknown population of trials exists, then the variability within the included studies cannot account for all possible variability in the estimate of effect. The random effects approach assumes that the trials in hand represent a broader population of studies. By adding among-study variability to within-study variability the random effects model allows inferences to be made about the hypothetical population of studies from those in hand.

Practically speaking, the main difference between fixed and random effects models is in the calculation of the weights applied to y_i. An additional term Δ^2 representing among-study variability is added to w_i to calculate a new weight w^*_i. Using the x^2_{homog} calculated above, Δ^2 may be determined as follows:

$$\Delta^2 = \max\left[0, \frac{x^2_{homog} - (K-1)}{\sum_{i=1}^{K} w_i - \sum_{i=1}^{K} w_i^2 \Big/ \sum_{i=1}^{K} w_i}\right] \qquad 3.1$$

The value for w^*_i may now be calculated as follows:

$$w_i^* = \frac{1}{w_i^{-1} + \Delta^2} \qquad 3.2$$

This new weight replaces w_i in the formulae derived to answer the questions posed in the previous section.

From this basic framework the individual methods used in this systematic overview may be described. The remaining discussion will highlight differences in the nature and calculation of y_i and w_i for each of the methods employed.

4. Binary Outcomes - Relative Risk

Assume a number of trials (K) comparing a treated group (t) with a control group (c) have been assembled. A 2x2 table may be constructed for a dichotomous outcome occurring in the ith trial.

	Outcome	No Outcome	Total
Treated	A_i	B_i	N_{1i}
Control	C_i	D_i	N_{0i}
	M_{1i}	M_{0i}	T_i

The rate of an outcome occurring in the treated (p_t) and control (p_c) groups may be calculated by:

$$p_{ti} = \frac{A_i}{A_i + B_i} \quad \text{and} \quad p_c = \frac{C_i}{C_i + D_i} \qquad 4.1$$

To summarize this information, the relative risk (RR) for the ith trial is calculated as follows:

$$RR_i = \frac{p_{ti}}{p_{ci}} = \frac{\frac{A_i}{A_i + B_i}}{\frac{C_i}{C_i + D_i}} \qquad 4.2$$

The Mantel Haenszel fixed effects model for combining relative risks (RR_{MH}) differs from the general meta-analysis model in the application of weights to the individual trials (2). Rather than calculate a trial specific relative risk and weight it during the pooling procedure, the summary RR_{MH} is calculated and weighted ($wi = C_iN_{1i}/T_i$) in a single step.

$$RR_{MH} = \frac{\sum_{i=1}^{K} A_i N_{0i} / T_i}{\sum_{i=1}^{K} C_i N_{1i} / T_i} \qquad 4.3$$

Variance of the RR_{MH} is then calculated using the following formula:

$$\text{var}[\ln(RR_{MH})] = \frac{\sum_{i=1}^{K} (M_{1i} N_{1i} N_{0i} - A_i C_i T_i) / T_i^2}{\sum_{i=1}^{K} \frac{A_i N_{0i}}{T_i} \sum_{i=1}^{K} \frac{C_i N_{0i}}{T_i}} \qquad 4.4$$

Confidence intervals may then be calculated by:

$$\exp[\ln(RR_{MH}] \pm Z\sqrt{\text{var}[\ln(RR_{MH})]} \qquad 4.5$$

The DerSimonian & Laird random effects model more closely follows the general meta-analytic design described earlier (3). Trial-specific relative risks are calculated

using equation 4.2 and the standard error used in the weighting procedure is calculated using the following formula:

$$SE[\ln(RR)] = \sqrt{\left(\frac{B_i}{A_i N_{1i}} + \frac{D_i}{C_i N_{0i}}\right)} \qquad 4.6$$

The weight applied to the ith trial is now determined using formulae 2.1, 3.1 and 3.2. Recall the weighting term form 2.1 reflects within-study variance and the weight from 3.1 reflects between-study variance. The overall weight for the ith trial is calculated using 3.2. From this point forward the DerSimonian and Laird random effects model follows the general meta-anlytic frame work described in Section 2.

5. Continuous Variables – Weighted Mean Difference

In a hypothetical population of K (i = 1,...K) trials in which a continuous outcome is described with a mean and standard deviation in treated ($x_t \pm sd_t$) and control ($x_c \pm sd_c$) groups of sample size n_t and n_c respectively, let the estimate of treatment effect yi be calculated by:

$$y_i = x_{ti} - x_{ci} \qquad 5.1$$

Similarly let the weight of ith trial (w_i) be calculated as

$$w_i = \frac{1}{\left(\sqrt{\frac{sd_{ti}^2}{n_{ti}} + \frac{sd_{ci}^2}{n_{ci}}}\right)^2} \qquad 5.2$$

The summary estimate of effect, standard error, and χ^2 for heterogeneity may be calculated using the general formula for fixed effects models defined in Section 2 (2).

The random effects model for analysis of continuous variables is largely similar to the fixed effects model but also incorporates between study variance into the weighting of the ith trial. For continuous variables the w_i and Δ^2 are determined using Equations 5.2 and 3.1, respectively. The random effects weight of the ith trial, w^*_i, is then determined using Equation 3.2. From this point y_i and w^*_i are entered into the general meta-analytic model described in Section 2.

6. Ordinal Variables – Proportional Odds Model

The proportion odds model described by Whitehead (4) was used to summarize volume of allogeneic blood transfused as an ordinal variable. Let the patient responses fall into m(i) ordered categories $C_{1i},....C_{m(i)i}$. The categories are ordered in terms of desirability such that C_{1i} is the best outcome and $C_{m(i)i}$ the worst.. Now let p_{jti} represent the probability that a patient in the ith study given the experimental treatment has a response in category j, where j = 1,....m(i) and let p_{jci} be similarly defined for the control group. Data from these ordinal responses can be reduced to binary responses by combining categories. Responses in categories C_{1i} to C_{ji} could be termed successes while responses in categories $C_{(j+1)i}$ to $C_{m(i)i}$ would represent treatment failures.

Three studies identified in this systematic overview provided data describing the volume of allogeneic blood transfused to individual patients (5) (6) (7). Each patient was assigned to one of five ordered categories reflecting their transfusion exposure. The five categories were:

1. No allogeneic transfusion
2. One unit of allogeneic blood
3. Two units of allogeneic blood
4. Three units of allogeneic blood
5. Four or more units of allogeneic blood

Summary data from the three eligible trials is presented in Table 1 Cell counts represent the number of individual patients in the given group in each category.

Table 1. Volume of Allogeneic Blood Transfused Expressed as an Ordinal Variable

Study	Group		Number of Subjects					
		Units	0	1	2	3	4 +	
		Category	1	2	3	4	5	Total
Moyes	ANH		6	1	2	1	0	10
	Control		3	1	3	2	1	10
Boldt	ANH		15	0	0	0	0	15
	Control		13	2	0	0	0	15
Bennett	ANH		13	3	2	1	1	20
	Control		13	2	2	2	1	20

The threshold for treatment success , j_i, was varied with cut-points set at no allogeneic blood transfused, one unit transfused, two units transfused and three units transfused.

Treatment effect was assessed using a log odds ratio calculated by the following formula where $Q_{jti} = p_{1ti} + \ldots + p_{jti}$ and $Q_{jci} = p_{1ci} + \ldots p_{jci}$:

$$\theta_{ji} = \log \left\{ \frac{Q_{jti}\,(1 - Q_{jci})}{Q_{jci}\,(1 - Q_{jti})} \right\} \qquad 6.1$$

Assuming a common odds ratio across all values of j, such that $\theta_{1i} = \theta_{2i} = \ldots = \theta_i$, the parameter θ_i can be considered the log odds of treatment success relative to control for the ith study regardless of how the ordered categories were divided.

θ_i was estimated using the linear logistic regression model available in PROC LOGISTIC in SAS. Programming details are provided in Appendix D.

The model may be described in general terms by:

$$\log\left[Q_{jti}/(1-Q_{jti})\right] = \alpha_{ij} + \theta_i \qquad 6.2$$
$$\log\left[Q_{jci}/(1-Q_{jci})\right] = \alpha_{ij}$$

Once the value of θ_i is known for the ith trial it may be entered as y_i into the general framework for meta-analysis described in Section 2.

The value of w_i is the inverse of the asymptotic variance of θ_i calculated in the logistic regression procedure. The method of calculating variance of estimated coefficients in logistic regression is derived from maximum likelihood estimation and is described in detail in Hosmer and Lemeshow (8).

REFERENCES

1. Fleiss JL. Statistical methods for rates and proportions. 2 ed. New York: John Wiley & Sons; 1981; 10, Combining evidence from fourfold tables. p. 160-87.

2. Bracken MB. Sinclair JC, Bracken MB, editors.Effective care of the newborn infant. Oxford: Oxford University Press; 1992; 2, Statistical methods for analysis of effects of treatment in overviews of randomized trials. p. 13-8.

3. DerSimonian R and Laird NM. Meta-analysis in clinical trials. Control Clin Trials 1986;7177-88.

4. Whitehead A, Jones NMB. A meta-analysis of clinical trials involving different classifications of response into ordered categories. Stat Med 1994;13(23/4):2503-15.

5. Moyes DG, Mistry BD, Conlan AA. Normovolaemic haemodilution using dextran 70 in thoracic surgery. S Afr Med J 1985;67(19):762-4.

6. Boldt J, Kling D, Weidler B, Zickmann B, Herold C, Dapper F, Hempelmann G. Acute preoperative hemodilution in cardiac surgery: volume replacement with a hypertonic saline-hydroxyethyl starch solution. J Cardiothorac Vasc Anesth 1991;5(1):23-8.

7. Bennett SR. Perioperative autologous blood transfusion in elective total hip prosthesis operations. Ann R Coll Surg Engl 1994;76(2):95-8.

8. Hosmer DW, Lemeshow S. Applied logistic regression. New York: John Wiley & Sons; 1989; 3, Interpretation of the coeffeients of the logistic regression model. p. 38-81.

APPENDIX D – SAS FILES FOR ORDINAL VARIABLE ANALYSIS

```
data study1;
set dir.meta2;
if study_co = 1;
run;
proc logistic data=study1;
model units = group_co;
run;

data study2;
set dir.meta2;
if study_co = 2;
run;
proc logistic data=study2;
model units = group_co;
run;

data study3;
set dir.meta2;
if study_co = 3;
run;
proc logistic data=study3;
model units = group_co;
run;
```

The LOGISTIC Procedure

Data Set: WORK.STUDY1
Response Variable: UNITS
Response Levels: 5
Number of Observations: 20
Link Function: Logit

Response Profile

Ordered Value	Units	Count
1	0	9
2	1	2
3	2	5
4	3	3
5	4	1

Score Test for the Proportional Odds Assumption

Chi-Square = 0.3173 with 3 DF (p=0.9567)

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	62.821	62.447	
SC	66.804	67.426	
2 LOG L	54.821	52.447	2.373 with 1 DF (p=0.1234)
Score			2.311 with 1 DF (p=0.1285)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Sqaure	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	-0.8804	0.6440	1.8688	0.1716		
INTERCP2	1	-0.4386	0.6191	0.5019	0.4787		
INTERCP3	1	0.8427	0.6489	1.6865	0.1941		
INTERCP4	1	2.4613	1.0596	5.3954	0.0202		
GROUP_CO	1	1.3020	0.8665	2.2576	0.1330	0.368235	3.677

Association of Predicted Probabilities and Observed Responses

Concordant = 39.3%	Somers' D = 0.264
Discordant = 12.9%	Gamma = 0.507
Tied = 47.9%	Tau-a = 0.195
(140 pairs)	C = 0.632

The LOGISTIC Procedure

Data Set: WORK.STUDY2
Response Variable: UNITS
Response Levels: 2
Number of Observations: 154
Link Function: Logit

Response Profile

Ordered Value	Units	Count
1	0	142
2	1	12

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	86.289	84.439	
SC	89.326	90.512	
2 LOG L	84.289	80.439	3.850 with 1 DF (p=0.0497)
Score			3.600 with 1 DF (p=0.0578)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Sqaure	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	1.9895	0.3545	31.5055	0.0001		
GROUP_CO	1	1.2320	0.6803	3.2791	0.0702	0.357552	3.428

Association of Predicted Probabilities and Observed Responses

Concordant = 45.2%	Somers' D = 0.332%
Discordant = 12.1%	Gamma = 0.578
Tied = 42.7%	Tau-a = 0.048
(1704 pairs)	C = 0.666

The LOGISTIC Procedure

Data Set: WORK.STUDY3
Response Variable: UNITS
Response Levels: 5
Number of Observations: 40
Link Function: Logit

Response Profile

Ordered Value	Units	Count
1	0	26
2	1	5
3	2	4
4	3	3
5	4	2

Score Test for the Proportional Odds Assumption
Chi-Square = 0.5426 with 3 DF (p=0.9094)

Model Fitting Information and Testing Global Null Hypothesis BETA=0

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	97.140	99.124	
SC	103.896	107.568	
2 LOG L	89.140	89.124	0.017 with 1 DF (p=0.8974)
Score			0.017 with 1 DF (p=0.8974)

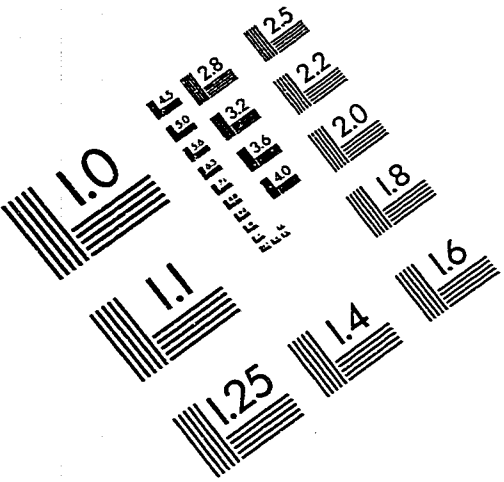
Analysis of Maximum Likelihood Estimates

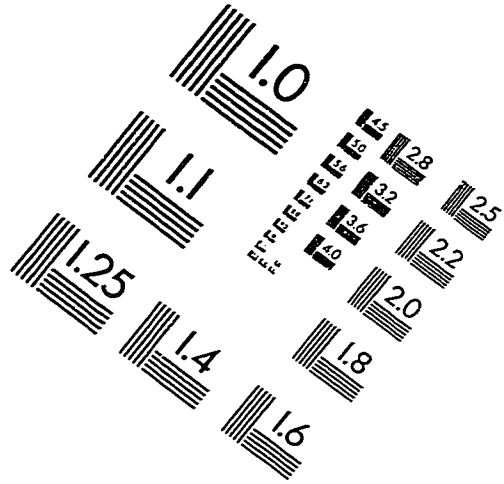
Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Sqaure	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCP1	1	0.5766	0.4593	1.5762	0.2093		
INTERCP2	1	1.1944	0.4930	5.8688	0.0154		
INTERCP3	1	1.9045	0.5721	11.0823	0.0009		
INTERCP4	1	2.9037	0.7902	13.5040	0.0002		
GROUP_CO	1	0.0832	0.6447	0.0167	0.8973	0.023225	1.087

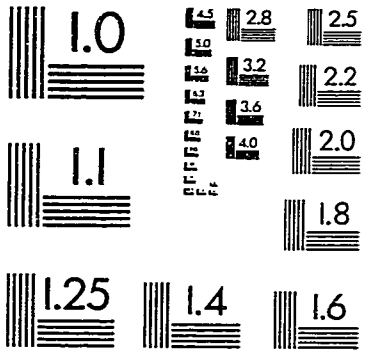
Association of Predicted Probabilities and Observed Responses

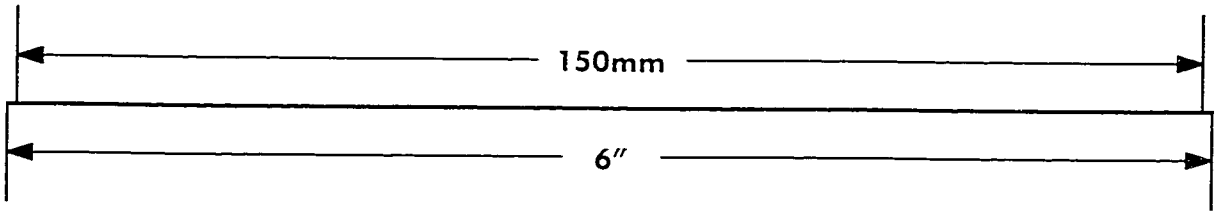
Concordant = 26.0% Somers' D = 0.018
Discordant = 24.1% Gamma = 0.037
Tied = 49.9% Tau-a = = 0.010
(435 pairs) C = 0.509

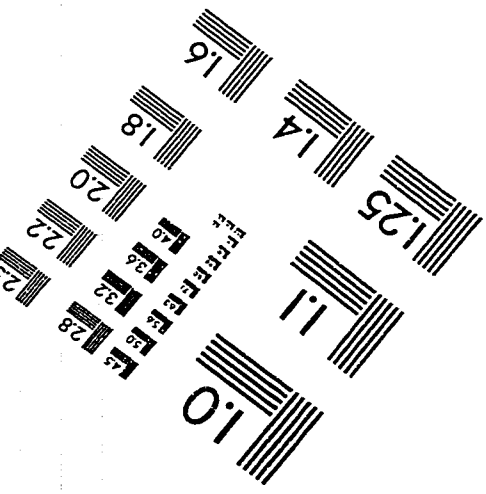
IMAGE EVALUATION TEST TARGET (QA-3)

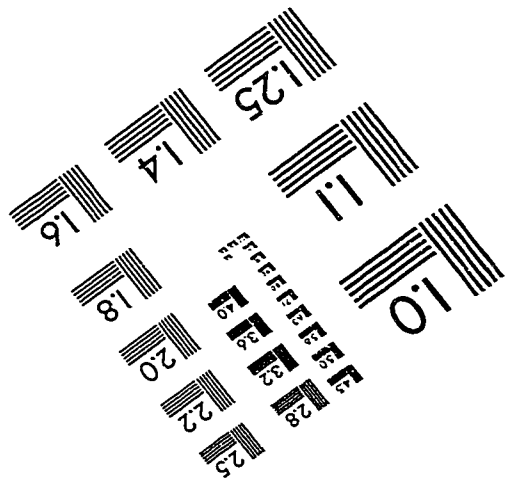












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989