World J. Surg. 23, 768–780, 1999

WORLD
Journal of
SURGERY

Second Step: Testing—Outcome Measurements

Wilfried Lorenz, M.D.,[1] Hans Troidl, M.D.,[2] Joseph S. Solomkin, Ph.D., M.D.,[3] Christoph Nies, M.D.,[4] Helmut Sitter, Ph.D.,[1] Michael Koller, Ph.D.,[1] Walter Krack, Ph.D.,[1] Michael F. Roizen, M.D.[5]

[1]Institute of Theoretical Surgery, Philipps-University of Marburg, Baldingerstrasse, 35033 Marburg, Germany
[2]Second Department of Surgery, University of Cologne, Ostmerheimerstrasse 200, 51109 Köln, Germany
[3]Department of Surgery, University of Cincinnati, College of Medicine, Cincinnati, Ohio 45267-0558, USA
[4]Department of General Surgery, Philipps-University of Marburg, Baldingerstrasse, 35033 Marburg, Germany
[5]Department of Anesthesia and Critical Care, University of Chicago, 5841 South Maryland Avenue, Chicago, Illinois 60637, USA

Abstract. Despite worldwide enthusiasm for endoscopic surgery, this new technology is now on the top of McKinlay's "product life circle curve." Critical questions are being asked about its benefits and burdens, but the concepts applied and the methodologies used for technology assessment are in a similar position as endoscopic surgery and need a critical evaluation. (1) There are incorrect and outdated concepts for the scientific basis of surgery (surgical theory) including the basic sciences involved; biomedicine still dominates, but assessment of outcome after operations is no longer possible without clinical epidemiology and social psychology. (2) Based on an outdated scientific theory for surgery, an outdated concept of disease is still propagated. It is denoted as mechanical and is based solely on biomedicine. Human subjects are reduced to biologic machines, and outcomes measurement excludes most dimensions of functioning and well-being. To achieve a valid result for outcome measures, a hermeneutic approach must be combined with the mechanical approach. (3) Based on an outdated model of disease, the outcomes used in endoscopic surgery rely too much on traditional measures, such as mortality rate, complication rate, hospital stay, and especially an endless list of biochemical mediators. Their alterations during the perioperative period have not yet been shown to be related to clinical or hermeneutic outcomes. A new method of assessment for clinical trials in endoscopic surgery and for other surgical problems is outlined, such as for surgical infections and for surgical oncology. It includes an index of recovery and objective health status assessed by the doctor, a quality-of-life self-report by the patient, and the true endpoint concept as a critical weighting of both types of outcome by patients and doctors.

Endoscopic surgery demonstrates the classic features of high technology in Western medicine [1]. First, it was born in the face of considerable economic and career-oriented pressure: Peptic ulcer surgery was replaced by drugs, and a similar loss of gallstone surgery was feared because of shock-wave lithotripsy and chemical lysis. Second, it offered fascinating challenges in technical performance and equipment when a video of the procedure debuted at the 1989 International Surgical Week in Toronto. It created the paradigm [2] of an important step forward in the goal of all surgery: to be without pain, stress, or risk. Third, its basic discipline was biomedicine, especially biochemistry, immunology, and cell biology (Table 1), to support the new technology with science [4]. The research concept of neuroendocrine–immune interactions [5, 6] was successfully applied to it. Finally, despite critical views [7] of such a procedure, it assessed benefits and risks predominantly by safety aspects [1] and traditional outcomes such as the rate of complications and hospital stay (Table 2), provoking extensive criticism of surgical research per se [9–11]. One prominent physician said "I actually find it a scandal that surgeons introduced endoscopic surgery in the absence of evidence from properly conducted trials" [12].

Correspondence to: W. Lorenz, M.D.

The problem, however, is much more general. The management and therapy of severe infections (sepsis) and malignant disease also reveal the classic features of high technology (molecular biology) and failure to assess its effectiveness [13, 14].

At the heart of the matter, a *causal chain* must be constructed that starts with theoretic items but ends with practical consequences. From an incorrect concept of basic science in medicine follows an incorrect concept of disease and outcomes and, derived from that, a flawed methodology for assessing outcomes of clinical trials.

The result is that no decision can be made reliably for the patient who seeks our help, certainly not our pathobiochemical explanations, which have never been validated. Surgeons who always think practically are motivated by Collingwood's statement to follow our arguments: "Thought is primarily practical, and only in the second place theoretical. Without theory, there would only be a few rudimentary types of practice, but without practice there could be no theory at all" [4].

Four Inadequate Concepts for Surgical Theory: Many Practical Consequences

Surgery, at least in significant parts, is a science; and as such [4] it must create and cultivate a scientific basis (surgical theory). Inability to achieve a valid assessment of endoscopic surgery and of other modes of surgical treatment has deep roots in the failures of traditional thinking.

Table 1. Postoperative changes of mediators following cholecystectomy.

Mediators	No.	Postoperative changes							
		Endoscopic				Conventional			
		0	↓	↑	> End	0	↓	↑	> Conv
Biogenic amines									
Epinephrine	5	2	0	3	0	1	0	4	4
Norepinephrine	6	2	0	4	2	3	0	3	2
Dopamine	2	1	0	1	1	1	0	1	1
Total catecholamines	1	1	0	0	0	0	0	1	1
Oligo- and polypeptides									
CRH	1	0	0	1	0	0	0	1	0
ACTH	3	1	0	2	0	0	0	3	2
Prolactin	3	1	1	1	0	1	0	2	1
Vasopressin	1	0	0	1	0	0	0	1	1
Growth hormone	3	0	0	3	1	0	0	3	0
Renin	1	0	0	1	0	0	0	1	0
Insulin	4	3	0	1	0	1	0	3	2
Glucagon	1	1	0	0	0	0	0	1	1
IL-1	2	2	0	0	0	2	0	0	0
IL-2	1	1	0	0	0	1	0	0	0
IL-6	14	3	0	11	0	0	0	14	9
IL-8	1	0	0	1	0	0	0	1	1
TNF-α	3	3	0	0	0	0	0	3	3
Proteins									
CRP	14	1	0	13	0	1	0	13	10
Elastase	1	0	0	1	0	0	0	1	0
Coeruloplasmin	1	1	0	0	0	1	0	0	0
Transferrin	1	1	0	0	0	1	0	0	0
α_1-Glycoprotein	1	1	0	0	0	1	0	0	0
α_1-Antitrypsin	1	1	0	0	0	0	0	1	1
α_2-Macroglobulin	1	1	0	0	0	1	0	0	0
Haptoglobulin	1	1	0	0	0	0	0	1	1
Fibrinogen	2	0	0	2	0	0	0	2	0
Fibrin dipeptide	1	0	0	1	0	0	0	1	0
vWFA	1	0	0	1	0	0	0	1	0
tPA	1	0	0	1	0	0	0	1	0
PAI-1	1	0	0	1	0	0	0	1	0
C3	1	1	0	0	0	1	0	0	0
Fatty acid derivates									
Not studied	—	—	—	—	—	—	—	—	—
Other									
Cortisol	18	3	1	14	0	0	1	17	9
$0_2^{\bullet}$ (monocytes)	1	1	0	0	0	0	0	1	1
$0_2^{\bullet}$ (neutrophils)	1	1	0	0	0	0	0	1	1

Results of a qualitative (systematic review) and quantitative meta-analysis [3] on the effectiveness of endoscopic versus conventional (including minilap) cholecystectomy. The study included 4,310,000 publications in several databases: MEDLINE, database of the Institute of Theoretical Surgery (n = 62,102 articles) and manual tracing as the gold standard. Thirty-three trials in patients were identified with the mechanical approach to systematic gallstone disease from the beginning of the study in 1989 up to 1996.

CRH: corticotropin-releasing hormone; ACTH: adrenocorticotropic hormone; 1L-1: interleukin 1, etc.; TNF-α: tumoral necrosis factor alpha; $0_2^{\bullet}$: superoxide (anion); CRP: C-reactive protein; vWFA: von Willebrand factor activator; tPA: tissue plasmin activator; PAI-1: plasmin activator inactivator 1; C3: complement factor 3; No.: number of studies dealing with the particular mediator; 0: no change; ↓: decrease; ↑: increase; > End: effect of endoscopic surgery greater than that of conventional surgery; > Conv: effect of conventional endoscopy greater than that of endoscopic surgery.

Misleading Concepts of Medical Philosophy and Basic Sciences for Medicine and Surgery

We still believe that modern philosophy is speculative, belongs to art and the humanities, and has little relevance for daily practical work and clinical research. Furthermore, we believe that the basic sciences for surgery are those of technology and biomedicine (today molecular biology) following the classic statement of Billroth: "The greatest upturn which surgery of the nineteenth century achieved in Germany was owed in the first place to the endeavour to unify all medical knowledge on the base of competent anatomical and physiological education" [15]. We have not recognized that modern physics, behavioral biology, clinical epidemiology, and social psychology have changed these two paradigms dramatically resulting in a revolution of medical care: the outcome movement [16–18].

The physicist H. Reichenbach made the distinction between the history of speculative philosophy, "the story of the errors of men who asked questions they were unable to answer," and the history of scientific philosophy, "the story of the development of problems" [19, 20]. Problems are solved not through vague generalities or picturesque descriptions of the relation between humans and the world but through *technical work.* Such work is done in the sciences, and in fact the development of problems must be traced through the history of the *individual sciences* [19].

Table 2. Outcomes of the mechanical and hermeneutic approach in endoscopic versus conventional cholecystectomy: results obtained in randomized trials and other types of clinical trials comparing the two groups.

Outcome variable	No. of studies choosing the endpoint: Randomized groups (n = 18)	Concomitant test and control group (n = 36)	Prospective test and historical control group (n = 14)
Mortality rate (%)	5	5	4
Complication rate (%)	11	17	12
Rate of reoperation (%)	0	1	0
Hospital stay (days)	14	22	13
Inflammatory response			
Fever	2	3	0
Duration of antibiotics	1	0	0
Pulmonary function	4	10	2
Intestinal atony	1	4	0
Weight loss	0	1	0
Diet	3	1	4
Muscle strength	0	2	0
Quality of life	1	0	1
No well-being			
Fatigue score	1	1	1
Complaint score	1	2	0
Pain			
Score	7	6	3
Use of analgetics	11	10	6
Patient's autonomy	0	1	0
Reconvalescence period	3	3	1
Inability to work	5	5	5

The study [3] was a qualitative (systematic review) and quantitative meta-analysis including a protocol before starting the conduct of the study, a review committee, control of publications bias with letters to leading authors in the field, and statistical analysis according to Fleiss [109].

n: number of studies.

Note: The line space after muscle strength divides mechanical from hermeneutic endpoints.

Outcome in medicine and surgery is such a problem and examples for its development are found in the concept of theoretic surgery [21], in principles and practice of research [22], in the philosophy of medicine [23], and in humane medicine [24]. The *value base,* which justifies medicine and surgery in normal life and end-stage of life conditions [25] is another problem of surgical theory, as it rationalizes assessments of quality of life and outcomes [12]. Handling *complexity* [12], "the art and science of *uncertainty*" in surgical indications [12, 26] and the ranking of *surgical intuition* as a mode of thought [12, 27, 28] are other important problems of surgical theory. Finally, the *concept,* not only the formal measurement of *effectiveness* in surgery is a problem of surgical theory. Hence, evidence-based surgery [29] and clinical practice guidelines are not independent goals but consequences of surgical theory.

Can all these questions be answered solely with the aid of two basic sciences: technology science and biomedicine? Certainly not. Two other basic sciences emerged during this century that are essential for, if not predominant in [12, 29, 30], clinical practice. Unfortunately, one, *clinical epidemiology* as defined by Feinstein, has only partly and reluctantly been recognized in medicine and surgery with great variations of research and training remaining even in Western countries [12, 29–33]. The other, *social psychology* as defined by Philipchalk [34], has not been mentioned by clinicians although they use its results daily. Both are essential for understanding modern outcome analysis.

Social psychology, which has emerged from sociology and psychology, is the study of the way people influence each others's thoughts, feelings, and actions [34]. It is an *experimental science* that applies strictly the criteria essential for biomedicine: reproducibility, validity, and refutability. It measures not only what people say or write in questionnaires but also nonverbal cues such as facial information, body positioning, movement, and touch [34]. Its major domains, some of its subdomains, and the specifications related to outcome are compiled in Table 3. Social psychology is not psychoanalysis; it is not one of the almost 100 speculative constructs and myths about the mind and the body; it is not ethics, or homoeopathy, or religion. It is as basic a science as molecular biology, and it includes sophisticated, multivariate statistics and refined information technology. The chess computer Deep Blue is a product of this fascinating new basic science; but quality of life concepts and assessment of outcomes will be products at least as important as these computers for clinical medicine.

Misleading Concepts of Disease

We still believe that the nature and extent of disease in human subjects to be treated with endoscopic or other forms of surgery can be described *sufficiently* with attributes obtained from measurement of physical and biochemical variables. We call the variables hard data. We categorize patient characteristics in clinical trials: age, disease name and severity, tumor TMN status; or we use the term "sepsis" or "SIRS" as a particular descriptor [36]. We classify concomitant perioperative risk (ASA) or intensive care unit (ICU) risk (APACHE score) [37]. Does a patient survive better or develop fewer complications if he *smiles* more often or more frequently enjoys the company of friends or develops a stronger will to survive? A good surgeon and anesthetist intuitively may consider such risk factors during the preoperative assessment. A systematic analysis of such items has already been developed that shows that such comprehensive considerations of health status are useful [38].

We need a paradigmatic, inspiring statement that allows us to understand how even disastrous biomedical science *alone* can be for medicine and surgery in general and for endoscopic surgery in particular. Such a statement can be constructed from Wulff's dialogue [23] between two surgical theory-minded clinicians, which describes the differences between the mechanical and the hermeneutic/critical approach to disease and outcome. (Hermeneutic is derived from ancient Greek: ἑρμηνευω = I describe, explain, exchange my complaints, views, judgments).

The Opponents. Doctor B (B = biologic paradigm) believes in the biologic model of disease, cell and molecular biology. He believes that cancer, coronary heart disease, sepsis, gallstone disease, and peptic ulcer can be reduced to abnormalities in biochemistry, physiology, or the structure or regulation of genes.

Doctor C (C = critical paradigm) believes in both models of disease, the mechanical *and* the hermeneutic one. He believes that diseases also have subjective dimensions in patients such as anxiety, lack of freedom and autonomy, and professional and

Table 3. Domains of social psychology related to outcome assessment.

Major domains	Subdomains	Outcome specifics	Ref. no.
Understanding ourselves and others	Social cognition: mental heuristics	Daily decision-making by doctors (experience, intuition)	[27]
	Social perceptions of ourselves and others	Expectations from treatments	[68]
	Attitudes: self-reports	Quality-of-life assessment	[47]
Understanding social influence	Group influence	Consensus procedures	[112]
Understanding social relations	Prejudice and discrimination	Social stigma (e.g., cancer patients)	[39]
	Prosocial behavior: helping others	Mechanical versus hermeneutic approach to outcome	[18]

The table was constructed by using the titles of chapters in Philipchalk's book [34] for domains and studies of our group for specifications.

social consequences that can be measured in monetary units [18]. In general, the patient suffers from a disturbed understanding or comprehension (social stigma) that can be measured in units lost in quality of life [39] and that has to be communicated and interpreted between the patient and the doctor, a hermeneutic circle [40].

The Dialogue. Doctor B proposes disease as a fault in the biologic machine. He defends the biologic concept of disease and therefore a biologic outcome *only.* That defense includes the outcomes in Table 1 with all the mediators and the outcomes in Table 2 that are listed in the first section: mortality rate and complication rate.

Doctor C criticizes this concept of outcome as *reductionistic:* Humans are reduced to biologic organisms.

Doctor B, explaining his model of disease, seems scientific and totally free from personal interests; but he has theorized from a concept called *species design* [23]. The patient should be cured by normalizing her or him according to the model that defines what an ordinary German, European, American, or Asian should be.

Doctor C protests. He has also a concept of his patients, but his is more complex. He wants to heal subjective disease, he wants to create well-being, he wants to restore autonomy (even in a wheel chair), *and* he wants to normalize the biologic variables.

The criterion of validity (evaluation) in these two views of Doctor B and Doctor C is clear. Is the outcome after an operation (e.g., endoscopic surgery) only the first part of Table 2 or the first *and* second parts of Table 2? The decision in favor of the first *and* second seems obvious, but there are deep roots for opposite opinions. Indeed, an etymologic analysis of the words for illness and ill in European languages [18] reveals striking differences that probably can explain to some extent why German doctors have such problems with outcomes research and understanding clinical research in other European countries. In contrast to most European languages, which operationalize subjective bad- and well-being (e.g., English *disease* = not free from discomfort; Russian *boljezin* = suffering from pain), the German word for ill, *krank,* means curved or bent, suggesting that the purpose of medicine is to straighten up the patient. This concept of disease is clearly *mechanical.* It has been developed, as the language shows, in German society over centuries.

The mechanical model of disease is responsible for the present evaluation of major life-threatening human diseases such as cancer and sepsis and also for the evaluation of functional diseases such as gallstone disease (Tables 1, 2). It has failed in many ways [13], but this failure has not yet influenced many biomedical scientists nor, as a consequence, the policies of major research grant-giving authorities or of drug administration. These policies reflect ordinary, nonrevolutionary thought and science [2].

There are historical reasons why the hermeneutic approach has not gained acceptance in the scientific domains of medicine and surgery, whereas it has always been operative in the patient–doctor relationship at an intuitive level [41]. Hermeneutics was first used in theology and philosophy to explain to individuals or communities what Jesus or a particular Greek philosopher meant in a given text. Hence hermeneutics was a part of the arts and humanities, not the life sciences. During the 1970s Gadamer defined it as the *art of understanding the views and judgment of other people* [42]. This view has dramatically changed during the last 20 years with the rise of social psychology as the new basic science for hermeneutics; it combines classic concepts and the methods of an exact natural science [43]. Thousands of articles on quality of life, but also on judgments about the different utilities of outcome (e.g., quality-adjusted life years) [44], during the last 20 years have analyzed patients' views, judgments, and behavior toward known risk and have already added experience to the new scientific theory.

Incorrect Concepts about Outcomes of Endoscopic Surgery

We still believe that *only* the mortality rate, complication rate, recurrence rate, hospital stay, and changes in biochemical mediators characterize a good outcome for endoscopic surgery. Have a look at Tables 1 and 2.

First, an attempt to retrieve *any* publication for a quantitative meta-analysis that compared endoscopic cholecystectomy with conventional cholecystectomy in a clinical trial and used either a clinical outcome or alterations of biochemical variables (mediators) as endpoints demonstrated that 33 of the 68 studies in Table 2 used mediators as endpoints. There were differences in favor of endoscopic surgery because of the stress response as indicated by measurement of epinephrine and cortisol (Table 1). Yet *clinical relevance* is not apparent, as the differences could not be related to any clinical outcome [45].

Quality of life, as measured by a valid construct to include somatic domains (symptoms), psychic domains (including pain as a separate group of items), and social domains [35] was measured only once in randomized trials (Table 2). This endpoint was assessed by Barkun et al. [46] using the Eypasch index [47]. In all other types of comparative trials it was also only measured *once* [48].

Third, quantitative meta-analysis of the mortality rate, compli-

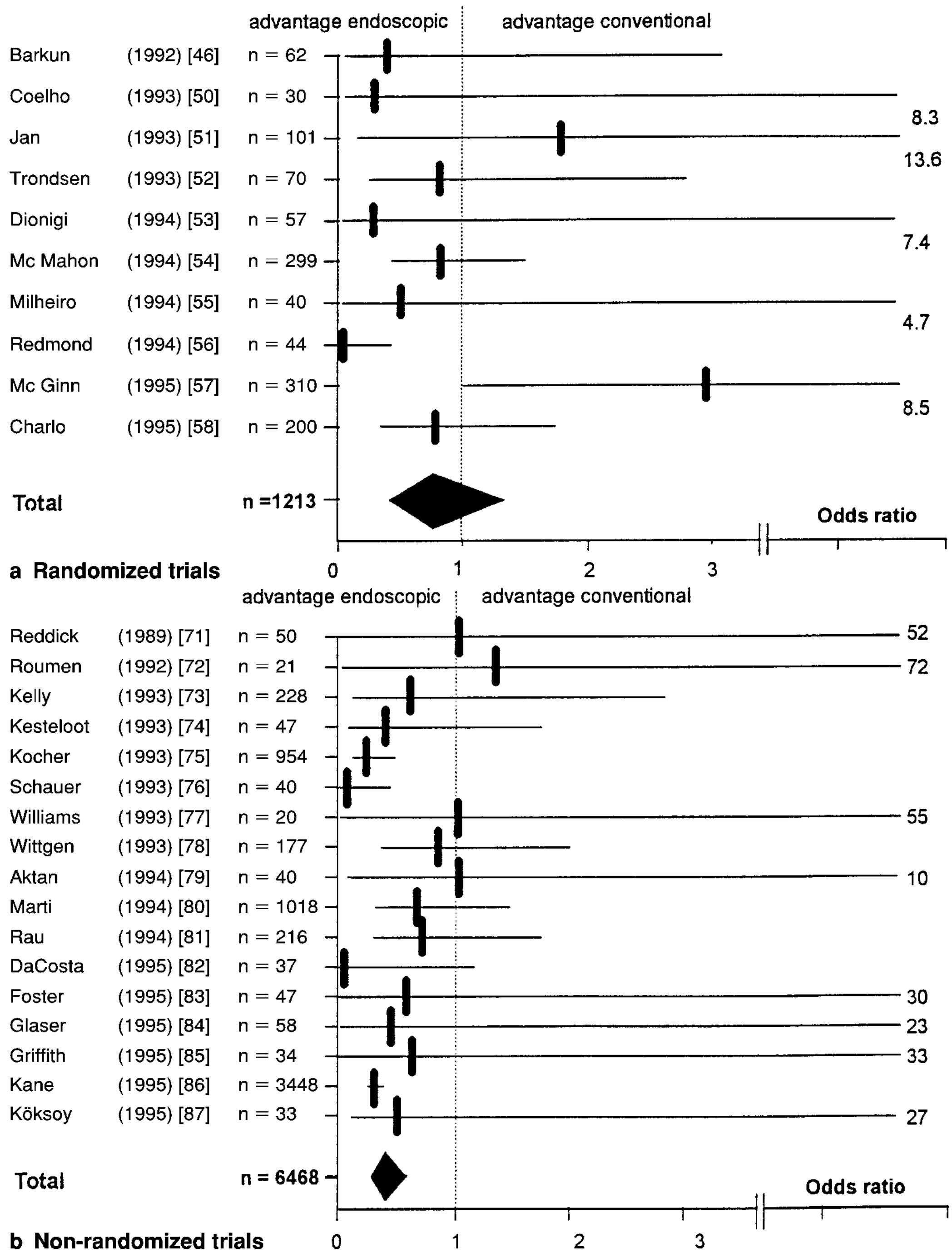

Fig. 1. Peto plots obtained from quantitative meta-analyses of endoscopic versus conventional cholecystectomy: comparison of randomized trials (**a**) and prospective cross-sectional trials (nonrandomized trials with concurrent control group, **b**). Criterion: postoperative complications. Individual trials were obtained from those listed in Table 2. Odds ratio (rhombus) for randomized trials: 0.77 (0.44–1.34 CI), which was statistically not of significant benefit for any of the treatments. For nonrandomized trials: an odds ratio (rhombus) of 0.41 (0.30–0.55 CI) was obtained. This is statistically significant for a benefit of endoscopic cholecystectomy because the confidence limits (CI) of the treatment effect exclude 1. The statistical methods involved in the calculations of the plots were described by Fleiss [109].

cation rate, and hospital stay (Table 2) showed no statistically significant difference for the first two endpoints [3] (upper part of Fig. 1) and a significant difference of 3 days for the third endpoint, in favor of endoscopic surgery. We know, however, from the randomized study of Majeed et al. [49] and the systematic review of Downs et al. [59] how questionable this difference is owing to the bias in unblinded studies even if they are randomized.

To distinguish between the mechanical concept of outcome alone and the critical concept including both the mechanical *and* the hermeneutic approach, we may use the following anecdote as a paradigmatic statement: Two well-dressed gentlemen were on their way home from a party, where they had obviously dined and wined too well. One was on his knees systematically examining the side wall underneath a street light. His friend volunteered helpfully "I am sure I heard your keys drop back here where it is dark!" The searcher on his knees replied "I know, but what is the use of looking back there where I can't see when it is so much *easier* here in the light" [60]. The biologic outcome in light, with the aid of simple animal and cell experiments, is so much *easier* to measure than health status or quality of life in the dark.

The critical concept or construct of outcomes is illustrated in Table 4. It synthesizes mechanical and hermeneutic variables. There are flaws in the concept of outcome that are found not only in the hands of traditional biomedical evaluators but also in the hands of social psychologists. A systematic analysis of incorrect concepts about outcomes is necessary that includes all the four basic sciences involved in surgical theory (Table 5):

There are still many studies about technologies in endoscopic surgery that have no concept of outcome at all. These studies contain one or a few traditional measures, and assessment of the operating time or blood loss are feasible but in this situation—after the learning time—not clinically relevant.

More difficult and at the moment the most important unsolved problem is handling the *complexity* of endpoints in a construct of outcome [12] (Table 5). The *true endpoint* [45] in endoscopic surgery for cholecystectomy or hernia repair may no longer be a single traditional endpoint, such as the recurrence rate, but a group of endpoints in an yet unknown hierarchic, time- and value-based order. Troidl et al. (Table 6) proposed such a scoring system for short-term outcome with new attributes: It is patient-friendly surgery and enables the surgeon to plan an operation with the particular outcome of the patient in mind. It is "neutral" in a legal sense, is easily understood by patients and doctors, and is prepared to weigh short-term outcomes. An example for reevaluation of long-term endpoints in a hierarchic construct of outcome is chronic pain in the groin following hernia repair. It is just now that in clinical trials [62, 63] this has become an undesirable endpoint: Patients are incapable of work and experience a poor quality of life.

With scoring systems on health status or quality of life [64] the single number of points obtained may not reflect the true severity of the patient's problems: The points are not interpretable [65]. In cancer research, therefore, we have worked hard to develop a *profile* of quality of life which in its various domains and subscales is readable like an x-ray film or computed tomography (CT) scan [66, 67].

The individual profile in Figure 2 (profile of quality of life) was reported to the general practitioner: "The 69-year-old patient came into our clinic with a very low quality of life. However, at the end of the in-hospital stay she demonstrated improvement in five of eight dimensions of quality of life. Especially impressive was the improvement in global quality of life (from 10 to 50) and pain (from 0 to 50). This patient came into our clinic in a physically weakened and discouraged state. Despite side effects of radiotherapy (WHO grade 3, diarrhea), physical and psychological stabilization of the patient was achieved during the hospital stay."

A similar profile as for cancer patients using the EORTC-C30 index can be developed for endoscopic surgery using the Gastrointestinal Quality of Life Index (GIQLI) of Eypasch et al. [47]. The items of the index are listed in Table 7. The five domains and a few combinations of items within the domains can certainly be used to construct the quality of life profile.

The concept of the critical approach to outcome assessment with the endpoints shown in Table 4 includes *essentially* to demonstrate that first the patient, then the doctor who treats the patient, and lastly the methodologist are partners who *separately* offer their judgments of health status and quality of life endpoints (Table 5). Studies in cancer patients showed that there was almost no or only a small and highly variable correlation between the judgment of the individual patient and doctor about the patient's quality of life [39]. Black et al. confirmed this disparity in a study on resection for benign prostatic hypertrophy [68].

Table 4. New directions in assessing outcomes: addition of hermeneutic variables.

Formerly
Mortality
Readmission
Complications
Other traditional measures of clinical outcome
Now—in addition
Functional status
Emotional health
Social interactions
Cognitive function
Degree of disability
Subjective well-being
Other valid indicators of health

Adapted from Epstein [17], Lorenz [18], and Diener [35].

There are publications on the hierarchy of outcomes [69], but they are dominated by methodologists, not by clinicians or patients, and are remarkably free of values [69]. Numerous surgical trials listing outcomes separately such as postoperative infections, thrombotic complications, recurrences, and prolonged hospital stay *without a scale of values* are equally disappointing (Table 5). A hierarchy of values as suggested by Little [24] depends on a concept of well-being [35] as it is affected by surgery, but this is not yet an objective of surgical research. On the other hand, methodologists have prepared utility scales with healthy test persons that may have little relevance in a real clinical scenario. Finally, although patients and doctors disagree on the hierarchy of outcomes (e.g., patients consider pain more important than do doctors) [39, 67], we have no concept to deal with the *disagreements.* Should we always believe the patient? That the patient can be influenced by social stigma demonstrates that he is not independent [39].

We discuss outcomes without analysis of what patients do after discharge from the hospital, which may strongly involve the health care insurers and employers (Table 5). In addition, cultural heterogeneity results in different outcomes [12]. One humorous example illustrates the differences in perspective regarding outcome [70]. A surgeon said to a bedouin woman with symptomatic gallstone disease: "You have two choices: one, you will get a long incision. You will have pain for a long time, and you will be forced to stay in the hospital for more than 1 week. Or, you will have no real incision, only a few small holes. You will have a little pain for 1 day; you will be discharged from the hospital the day after the operation. What do you choose?" The beduin woman responded, "Please, doctor, give me 2 weeks free from my 10 children!"

Incomplete and Incorrect Assessment of Outcome and Clinical Trials

We still believe (on the quiet) that randomized trials are *not* necessary for demonstrating the effectiveness of operations [7] and go through nigthmares with the "rise and fall of the random controlled trial in surgery" [88]. However, as in the case of aprotinin for acute pancreatitis, of shunt surgery for portal hypertension, and of surgery for prostatic cancer, we are again wrong in the case of endoscopic surgery: Randomized trials have demonstrated that the outcome results of nonrandomized trials were *overly*

Table 5. Some characteristics of incorrect concepts of outcome.

General defects	Specification in detail
No concept at all	Traditional measures only used as an alibi: only formal methodologic evaluation Methodologic feasibility driving the concept: clinical relevance of the effect not detectable
No concept for complexity	Single endpoints without clinical relevance: neglecting the concept of the true endpoint [45] Only a single number in scoring systems, without intention or possibility for interpretation
No concept for the critical approach: mechanical and hermeneutic outcomes	Aggregate endpoints, but only mechanical Aggregate endpoints, not including the patient's experience No distinction between health status and quality of life
No concept for a hierarchy of outcome values for patient and doctor	Listing of outcomes without a value system No concept for well-being Utility scales driven only by decision-making methodology Presentation of utilities that are not target-oriented to doctors and patients No concept for agreement/disagreement between patient and doctor
No concept for community's and cultural heterogeneity	No concept for the analysis of the health care insurers and employers No concept for the interaction of outcome with religion or ideologies

Table 6. Classification of positive and negative events in endoscopic surgery: modeling complexity in short-term outcome by a patient-related hierarchic order.

Class I, incident-free surgery, no incidents: no surgical technical problems and no negative outcome for the patient
Class II, inconsequential incident surgery: one or more surgical technical problems, but no negative outcome for the patient (intraabdominal stone loss, bleeding)
Class III, consequential nonincident surgery: no surgical technical problems but one or more negative outcomes for the patient (hematoma, wound infection)
Class IV, incidents: one or more surgical technical problem with corresponding negative outcomes for the patient (changing operative method intraoperatively, relaparotomy)
Class V, incidents: death in any relation to operation

From Troidl et al. [61].

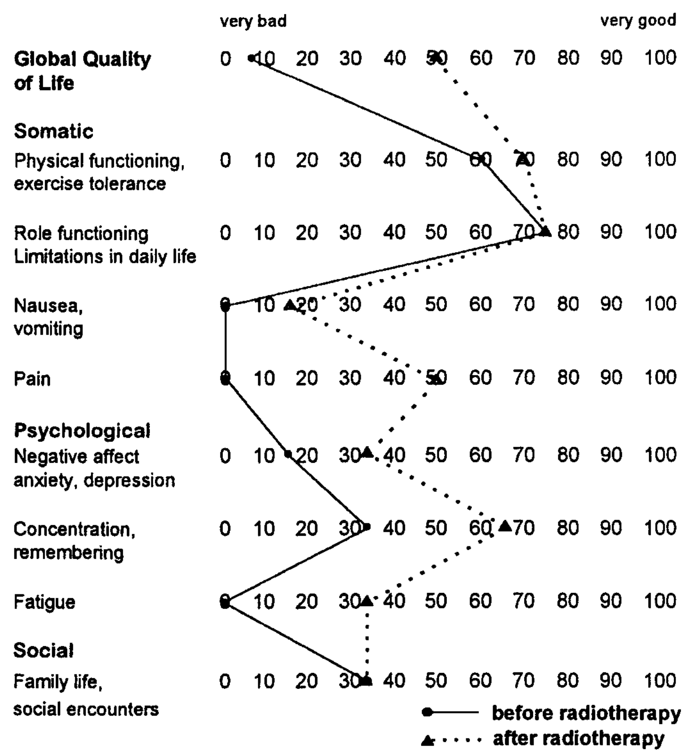

Fig. 2. Profile of quality of life obtained from a patient with proctectomy and adjuvant radiotherapy. Each of the items of this European Organization for Research and Treatment of Cancer (EORTC)-QLQ-C30 index including the colorectal cancer module is constructed from answers to several questions graded by expressing four levels of severity. By a mathematic process they are transformed to a scale of 0 (very bad) to 100 (very good). The clinicians obtained the answers from a questionnaire completed by patients in the absence of either health providers or members of the family (avoiding social stigma) [39]. The questions are so simple they can be answered by almost all patients. Drop-out rate was 10% (from Wagner et al. [68], with permission of Springer-Verlag.)

optimistic (Fig. 1). Complications were equal in number (not statistically different) in a meta-analysis of randomized trials but significantly fewer in nonrandomized trials for endoscopic surgery, pointing up the biases. Some characteristics of incomplete and incorrect assessment of outcome in clinical trials are listed in Table 8.

We still believe that we can go the easy way (Table 8). However, with assessment of outcome *only within the hospital,* not at other relevant points of time, or only within *one medical discipline* (surgery or anesthesiology), or *without formal criteria* of evaluation, reliability, responsiveness, and validity, assessment is inadequate (Table 8). Such inadequacy is an unacceptable basis on which to create clinical practice guidelines and evidence-based medicine [89].

A paradigmatic example is needed for the lower part of Table 8. It uses the analogy of a shooting gallery as inaugurated by a review of Büttner et al. [90]. Shooting failures are distinguished by four possibilities (Fig. 3): gross error, or not hitting the target at all; poor reliability due to large variation, but good validity with all the hits around the center; poor validity despite excellent reliability, possibly because the rifle has a warped barrel; and finally the inherent, unsurmountable error of even a perfect measurement. The third case is most important when measuring quality of life. The indices may have excellent Cronbach α values for reliability, but quality of life may not be measured at all. We may be precise but *precisely wrong.*

Such cases became a major matter of concern (Table 8). First, assessment of quality of life was strongly influenced by measurable *expectations* of patients [67] and doctors [67, 91] before treatment started. Furthermore, it was even more influenced by a *negative affect* of patients [39], a summary category that has been conceptualized as a general dimension of subjective distress, reflecting undifferentiated bad moods and a low self-concept and that can be

Table 7. Eypasch index: items of the gastrointestinal quality of life index (GIQLI).

Core symptoms	Psychological items
Pain	Sadness
Bloating	Nervousness
Epigastric fullness	Frustration
Flatus	Happiness
Belching	Bothered by treatment
Bowel frequency	Cope with stress
Abdominal noises	Social items
Restricted eating	Daily activities
Enjoyed eating	Leisure activities
Fatigue	Disease-specific items
Physical items	Regurgitation
Strength	Dysphagia
Feeling unwell	Eating speed
Feeling unfit	Nausea
Endurance	Diarrhea
Wake up at night	Bowel urgency
Appearance	Constipation
	Blood in stool

Developed by Eypasch et al. [47]. The score values for the items are obtained by 36 questions. Each is graded from least desirable option (0 points) to the most desirable option (4 points), with the GIQLI score the sum of all points (maximum $4 \times 36 = 144$ points).

measured again by sociopsychological scales [39]. Finally, experienced *social stigma* associated with the illness by the family and the friends was measured to be a powerful confounder in quality of life assessment with the EORTC-C30 index [39]. Hence we have to plan carefully when quality of life is measured in clinical trials. Researchers and practitioners must be aware that quality of life is not an entity that can be interpreted "as such" but is interwoven with numerous psychological, social, and health variables. Understanding these influences (or confounding variables) is essential for understanding what quality of life means and how it can be used and interpreted in clinical studies and everyday patient care. In general, health status assessment by doctors and quality of life assessment by patients are currently not separated precisely [92, 93]. The hermeneutic approach, assessment of quality of life by the *patient* and undisturbed by other influences should not be heavily biased by the so-called objective assessment of physicians or health administrators [39, 67].

From the many other defects in measuring outcome in clinical trials listed in the lower part of Table 8 two flaws deserve special mention. First, when methodologists use students of psychology for the development of outcome scales, the utility scales derived from such investigations (death = 0, perfect well-being = 1) may lack clinical relevance. One way to escape such fundamental failures is the critical approach: Patients and doctors must combine with the methodologists to bring the real world into the laboratory of professional decision makers and social psychologists [39, 65, 66].

Second, outcome is measured mainly *within* a particular medical discipline, although the patient is connected with several disciplines perioperatively but especially with surgery, anesthesiology, and intensive care. There are remarkable fences around the disciplines, and failures after operation are easily thrown into the neighbor's garden. For example, silent myocardial ischemias perioperatively have an influence on the rate of myocardial infarctions for *2 years* after operation [94]. Careful preoperative risk assessment [95] and refusal to oversimplify causality of bad outcomes (anesthetists are responsible for the first 2 days, surgeons are responsible thereafter) are necessary for convincing and honest outcome assessment.

Outline for Outcome Analysis

From our criticism of present outcome analyses it is apparent that a single strategy cannot be recommended. However, we can propose an outline that differs in specific items from present research and clinical practice in outcomes assessment.

1. Outcome analysis in surgery should be based on *three obligatory elements:* the true endpoint for the particular surgical problem [45], the doctor's assessment of recovery and health status through a corresponding index, and a quality of life index self-reported by the patient and assessed by him or her independent from foreign observation. This method excludes, for instance, sending questionnaires to patients at home.
2. The three outcome components should have a value base. Biomedical indices describing perioperative recovery are available from anesthesia [97, 98] (Table 9) or from surgery [24]. Some indices developed for other areas are also recommended for surgery, such as the Quality of Well-Being Index [100] or the Short Form 36 Health Survey in standard and acute versions [92, 101]. For quality of life assessment as the third component, the Eypasch index [47] may be used for benign diseases of the gastrointestinal tract. For malignant diseases and surgical infection trials, the European Organization for Research and Treatment of Cancer (EORTC) QLQ-C30 index with an organ-specific module [102, 103] can be utilized. Again, self-assessment by the patient must be considered [39]. The new concept of the three outcome components can and should replace mortality as the only primary endpoint in clinical trials—not only in those that deal with chronic disease but also in those that study emergency situations and include ICU treatment, such as for sepsis [14].
3. That scoring systems for health status (assessed by doctors) and quality of life (expressed by the patients in a self-report under well standardized conditions [39]) can be interpreted should be guaranteed. *Profiles* with the subscales should be constructed for daily use in clinical practice [66, 99] (Fig. 3) and for interpretation of the results in clinical trials.
4. Postoperative health status indices, such as the McPeek recovery index [98, 104] or quality of life indices [47], can be used for constructing *utility scales.* It is unlikely that classic methods for utility assessment, such as standard gamble, time trade-off (shorter life traded for better health), or visual analog scales, are applicable in real patients under various conditions [44]. Utility scales must be validated by doctors and patients.
5. Statistical evaluation of a combination of endpoints, including variables of the mechanical approach (survival curves) and of the hermeneutic approach (scales and subscales of quality of life indices), is possible [103, 105]. Also the problem of protecting the α error in case of three-times-repeated significance testing (three endpoints) can be solved without markedly increasing the sample sizes for clinical trials [106].
6. Surgical trials, especially those comparing endoscopic versus conventional operation techniques or trials of sepsis with various biologic effects of the pleiotropic cytokines or antibody treatments cannot be conducted in a double-blind fashion [11, 108]. New surgical procedures will continue to lack credibility

Table 8. Some characteristics of incomplete and incorrect assessment of outcome in clinical trials.

General defects	Specification in detail
Methodology driven by convenience	Measurement only within hospital Measurement perioperatively restricted to one medical discipline (surgery or anesthesiology)
Measurement without application of basic science	Questionnaires, lists, scores not validated by sociopsychological methodology [43]
Measurement without valid evaluation criteria	Tests and test constructs not formally evaluated by the criteria of reliability, responsiveness, validity Test constructs not suitable for the particular clinical scenario (malignant or benign disease)
Measurement without bias control	No experimental design No assessment of possible confounders (social stigma, expectancies, negative affect)
Measurement with only methodologists' utilities[a]	Utilities without clinical relevance Utilities without relation to real patients (no hermeneutic approach in constructing the value scales)
Measurement without relation to perioperative risk	Preoperative risk not considered at all Utilities not related to medical disciplines, such as surgery and anesthesiology

[a]Utilities correspond to the definition given by Weinstein and Fineberg [113]: they are a value between 0 (usually death) and 1.0 (usually perfect health).

Table 9. McPeek index for assessing postoperative recovery.

Patient characteristics	Score
Patients who died	
In the operating theater	1
Within 30 days after leaving the operating theater	2
Patients who survived 30 days, yet	
Required a great amount of care in the intensive care unit (ICU)	4
Required a moderate or minimal amount of care in the ICU	5
Patients who underwent routine recovery on a normal ward, yet	
Had a postoperative hospitalization > 7 days longer than standard	7
Had a postoperative hospitalization ≤ 7 days longer than standard	8
Had a standard length of hospitalization for the particular operation	9

Index published originally in 1986 [98] but modified in 1988 [105], in particular by "expected versus found" criteria.

unless assessed by properly randomized trials with objective outcome measures [11]. The solution to this problem, however, may not come from hiding the operative wounds by surgical dressings [49] but from measuring the expectations of doctors and patients in favor or against a new treatment and appropriate regression models [107].

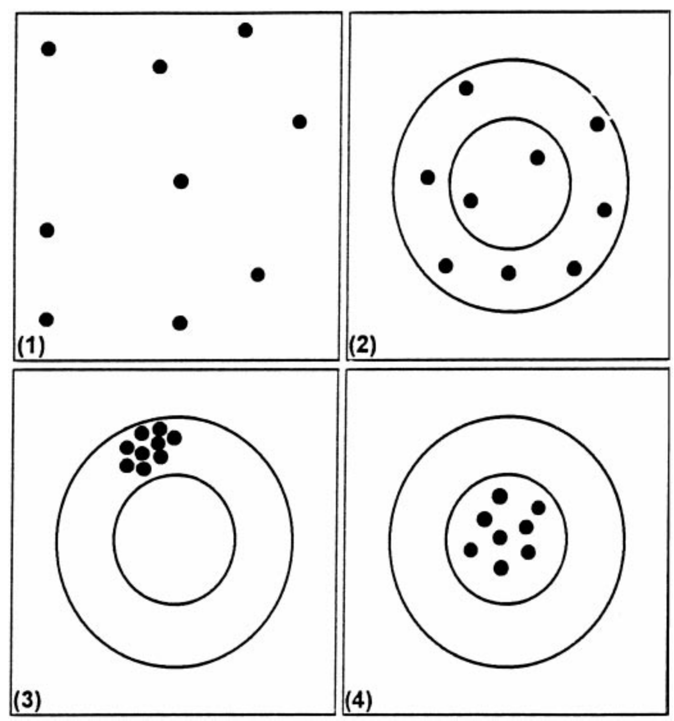

Fig. 3. Types of errors in methodology. Scheme developed by the German mathematician C.F. Gauss (1777–1855). Illustration by example of shots at a target with a rifle. **1:** gross error; **2:** error of reliability; **3:** error of validity; **4:** perfectly possible outcome. Modified (with different terminology) from [90], with permission of Wiley-VCH.

Conclusions

We need better clinical trials in endoscopic surgery and in surgery in general that are not dominated by methodology but by surgical concepts. We do not need thousands of patients with an increasing heterogeneity of treatment algorithms [8] and with decreasing opportunities to determine the effectiveness of new treatments. We need careful trials with relevant endpoints and thereafter quantitative meta-analyses which, although less precise than megatrials, may be more representative for nations and continents. The systematic overestimation of the mechanical approach to disease treatment and outcome (molecular biology) is disastrous for this new and certainly important basic science itself. Using thousands of patients in clinical trials with the survival rate as the only accepted endpoint prohibits the valuable use of products obtained by gene technology in clinical practice. The most successful company in this field, Amgen (Thousand Oaks, CA, USA), has introduced its product granulocyte colony-stimulating factor (G-CSF) for improvement of the quality of life of cancer patients [109]. This product was successful without megatrials.

Résumé

En dépit d'un enthousiasme mondial pour la chirurgie endoscopique, cette nouvelle technologie est maintenant au

sommet de la courbe circulaire de McKinlay appelée la "product life circle curve." Les questions critiques sont les bénéfices et les charges encombrées par cette technique. Cependant, les concepts appliqués et les méthodologies utilisées pour évaluer la technologie sont au même point que l'évaluation de la chirurgie laparsocopique et ont besoin d'une revision critique: (1) Les concepts pour établir les bases scientifiques de la chirurgie (la chirurgie «théorique») sont parfois inexacts et démodés y compris les sciences de base concernées. La biomédicine domine toujours, certes, mais l'évaluation des résultats postopératoires n'est plus possible sans inclure des notions d'épidémiologie clinique et de psychologie sociale. (2) Basé sur la théorie scientifique, démodée, on propose toujours le concept selon lequel la maladie est mécanique, et basée sur la biomédecine. Les sujets humains (patients) sont réduits à des machines biologiques et la plupart des critères de jugement excluent les critères fonctionnels et le bien-être. Pour arriver à un résultat valide en ce qui concerne les critères de jugement, il faut combiner une approche herméneutique à l'approche mécanique. La part relative des éléments de chaque approche doit être pesée par les patients et les médecins, et en dernier lieu, par les méthodolgistes. Ceci nous amènera à définir le «vrai critère de jugement». (3) Basé sur le modèle démodé de la maladie en tant qu'état pathologique, les critères de jugement utilisés en chirurgie laparoscopique sont trop souvent ceux de la chirurgie traditionnelle, c'est-à-dire la mortalité, le taux de complications, la durée d'hospitalisation, et surtout, sur une liste sans fin de perturbations ou conservation de médiateurs biochimiques. Cependant, il n'a jamais été démontré que les perturbations constatées dans la période postopératoire étaient en rapport avec l'évolution clinique ou herméneutique. Donc, surtout en ce qui concerne ces médiateurs, des nouveaux concepts en ce qui concerne les résultats attendus par la chirurgie laparoscopique pourraient offrir de nouveaux horizons pour la recherche clinique.

Resumen

A pesar del entusiasmo que se registra en todo el mundo por la cirugía endoscópica, esta nueva tecnología se encuentra actualmente en primer lugar en la denominada "curva del ciclo vital de un producto" de McKinley. Se plantean interrogantes sobre sus beneficios y sus aspectos negativos, pero los conceptos que se aplican para la evaluación de la tecnología se hallan en una posición similar en cuanto a la cirugía endoscópica y requieren evaluación crítica: 1) Existen conceptos incorrectos y obsoletos pertinentes a las bases científicas de la cirugía (teoría quirúrgica), incluyendo las ciencias básicas: la biomedicina es aún dominante, pero la valoración del resultado luego de la operación ya no es válida sin la epidemiología clínica y la sicología social. 2) Con base en una obsoleta teoría científica para la cirugía, persiste un concepto igualmente obsoleto de la enfermedad, de carácter mecánico y fundamentado en biomedicina. El ser humano es concebido como una máquina biológica y las valoraciones de resultado final excluyen dimensiones tales como funcionamiento y bienestar. Para lograr un resultado válido en la valoración de resultado final, se requiere la combinación de un enfoque hermenéutico con uno mecánico. El análisis crítico de los elementos de cada uno de estos enfoques debe ser hecho por pacientes y por médicos y—sólo como tercera instancia—por metodólogos. Este es el concepto de "resultado final verdadero." 3) Con fundamento en un concepto de enfermedad ya pasado de moda, los resultados que se utilizan en la cirugía endoscópica se centran excesivamente en mediciones tradicionales tales como tasa de mortalidad, tasa de complicaciones, duración de la hospitalización y, especialmente, en una innumerable lista de indicadores bioquímicos. Las alteraciones que tales mediciones exhiben en el periodo perioperatorio, todavía no han demostrado estar relacionadas con los resultados hermenéuticos clínicos. Es por ello que, especialmente en el campo de los mediadores, la nueva concepción de resultado final ofrece amplias oportunidades para la investigación clínica.

Acknowledgments

This conceptual study was inaugurated by the Lucerne Group for Consensus Assisted Development of the Study Protocol on Prevention of Abdominal Sepsis: Example G-CSF [111]. The manuscript was reviewed by experts to whom the authors are greatly indebted (listed by course of the article): P. Janich (methodological philosophy, Institute of Philosophy, University of Marburg); H. Wulff (medical philosophy, Department of Medical Philosophy and Clinical Theory, University of Copenhagen); M. Healy (medical statistics, Department of Medical Statistics and Epidemiology, London School of Hygiene and Tropical Medicine); H. Schäfer (medical statistics, Department of Medical Biometrics and Medical Informatics, University of Marburg); K. Voigt (neuroendocrinology, Institute of Normal and Pathological Physiology, University of Marburg); M. Rothmund and B. Stinner (surgery, Department of General Surgery, University of Marburg); A. Bauhofer and I. Celik (cell and molecular biology, Institute of Theoretical Surgery, University of Marburg); S. Kozlik (department editing, Department of Anesthesiology and Critical Care, University of Chicago).

The meta-analysis included in this article was peer-reviewed by H.K. Selbmann (medical statistics, University of Tübingen) and M. Büchler (surgery, University of Berne) and was accepted by the Medical Faculty, University of Marburg on February 4, 1998 [3]. It is available upon request. The study was supported by grants from Deutsche Forschungsgemeinschaft (Ba 1560/2-2 and Sonderforschungsbereich 297, A8) and Amgen Inc.

References

1. Jennett, B.: High Technology Medicine—Benefits and Burdens. Oxford, Oxford University Press, 1986, pp. 1–317
2. Kuhn, T.S.: The Structure of Scientific Revolutions (2nd ed.). Chicago, University Press, 1970, p. 23
3. Nies, C.: Non-optimum results in recovery from biliary surgery: a result of interaction between strain of the neuroendocrine system, mast cell activation and modulation of the immune system. Habilitation thesis, Philipps-University, Marburg, 1997, pp. 1–179
4. Healy, M.J.R.: Is statistics a science? J. R. Statist. Soc. A *141*:385, 1978
5. Besedovsky, H.O., del Rey, A.: Immune-neuro-endocrine interactions: facts and hypotheses. Endocr. Rev. *17*:64, 1996
6. Udelsman, R., Holbrook, N.J.: Endocrine and molecular responses to surgical stress. Curr. Probl. Surg. *31*:655, 1994
7. Neugebauer, E., Troidl, H., Spangenberger, W., Dietrich, A., Lefering, R.: Conventional versus laparoscopic cholecystectomy and the randomized controlled trial. Br. J. Surg. *78*:150, 1991
8. Encke, A., Hanisch, E., Sitter, H., Greger, B., Bauhofer, A., Margolis, C., Lorenz, W.: Models of evaluation in planning and standard-

ization of surgical treatment: sepsis as an example. Langenbecks Arch. Chir. Suppl. *2*:323, 1997
9. Horton, R.: Surgical research or comic opera: questions, but few answers. Lancet *347*:984, 1996
10. Johnson, A.: Laparoscopic surgery. Lancet *349*:631, 1997
11. Johnson, A.G., Dixon, J.M.: Removing bias in surgical trials. B.M.J. *314*:916, 1997
12. Lorenz, W., McKneally, M., Troidl, H., Banerjee, J.K., Benbassat, J., Dziri, C., Fagniez, P-L., Little, J.M., Wong, J., Wulff, H.R.: Surgical research around the world. In: Surgical Research: Basic Principles and Clinical Practice (3rd ed.), H. Troidl, W.O. Spitzer, B. McPeek, D.S. Mulder, M. McKneally, A. Wechsler, C.M. Balch, editors. New York, Springer, 1998, pp. 637–654
13. Bone, R.C.: Why sepsis trials fail. J.A.M.A. *276*:565, 1996
14. Petros, A.J., Marshall, J.C., van Saene, H.K.F.: Should morbidity replace mortality as an endpoint for clinical trials in intensive care? Lancet *345*:369, 1995
15. Billroth, T., von Winiwarter, A.: Die allgemeine chirurgische Pathologie und Therapie in fünfzig Vorlesungen. Vorlesung 1. Berlin, Reimer, 1885, pp. 15–17
16. Relman, A.S.: Assessment and accountability: the third revolution in medical care. N. Engl. J. Med. *319*:1220, 1988
17. Epstein, A.M.: Sounding board: the outcomes movement—will it get us where we want to go? N. Engl. J. Med. *323*:266, 1990
18. Lorenz, W.: Outcome: definition and methods of evaluation. In: Surgical Research: Basic Principles and Clinical Practice (3rd ed.), H. Troidl, W.O. Spitzer, B. McPeek, D.S. Mulder, M. McKneally, A. Wechsler, C.M. Balch, editors. New York, Springer, 1998, pp. 513–520
19. Reichenbach, H.: The Rise of Scientific Philosophy. Berkeley, University of California Press, 1951, pp. 117–118
20. Hartmann, D., Janich, P.: Methodischer Kulturalismus. In: Methodischer Kulturalismus. Zwischen Naturalismus und Postmoderne, D. Hartmann, P. Janich, editors. Frankfurt, Suhrkamp, 1996, pp. 1–69
21. Lorenz, W., Rothmund, M.: Theoretical surgery: a new specialty in operative medicine. World J. Surg. *13*:292, 1989
22. Troidl, H., Spitzer, W.O., Mulder, D.S., Wechsler, A.S., McPeek, B., McKneally, M.F., Balch, C.M.: Principles and Practice of Research: Strategies for Surgical Investigators (2nd ed.). New York, Springer, 1991, pp. 1–566
23. Wulff, H.R., Pedersen, S.A., Rosenberg, R.: Philosophy of Medicine: An Introduction. Oxford, Blackwell, 1986, pp. 1–222
24. Little, M.: Humane Medicine. Cambridge, Cambridge University Press, 1995, pp. 1–195
25. Patrick, D.L., Starks, H.E., Cain, K.C., Uhlmann, R.F., Pearlman, R.A.: Measuring preferences for health states worse than death. Med. Decis. Making *14*:9, 1994
26. Dowie, J., Elstein, A.: Professional Judgement: A Reader in Clinical Decision Making. Cambridge, Cambridge University Press, 1988, pp. 1–565
27. Gross, R., Lorenz, W.: Intuition in surgery as a strategy of medical decision making: its potency and limitations. Theor. Surg. *5*:54, 1990
28. Abernathy, C.M., Hamm, R.M.: Intuition as a mode of thought in surgery. In: Surgical Intuition. What It Is and How To Get It, C.M. Abernathy, R.M. Hamm, editors. Philadelphia, Hanley & Belfus, 1995, pp. 19–68
29. Sackett, D.L., Richardson, W.S., Rosenberg, W., Haynes, R.B.: Evidence-based Medicine, How To Practice and Teach EBM. Edinburgh, Churchill Livingstone, 1997, pp. 1–250
30. Sackett, D.L., Haynes, R.B., Tugwell, P.: Clinical Epidemiology: A Basic Science for Clinical Medicine (2nd ed.). Boston, Little Brown, 1991, pp. 1–441
31. Lorenz, W.: Worse, or just different? Lancet *348*:1638, 1996
32. Feinstein, A.R.: Clinical Epidemiology: The Architecture of Clinical Research. Philadelphia, Saunders, 1985, pp. 1–812
33. Feinstein, A.R.: Clinimetrics. New Haven, Yale University Press, 1987, pp. 1–272
34. Philipchalk, R.P.: Invitation to Social Psychology. Fort Worth, Harcourt Brace, 1995, pp. 1–566
35. Diener, E.: Subjective well-being. Psychol. Bull. *95*:542, 1984
36. Solomkin, J.: Sepsis as a descriptor, not a disease. Theor. Surg. *9*:12, 1994
37. Knaus, W.A., Zimmermann, J.E., Wagner, D.P., Draper, E.A., Lawrence, D.E.: APACHE: acute physiology and chronic health evaluation; a physiologically based classification system. Crit. Care Med. *9*:591, 1981
38. Roizen, M.F.: RealAge: Are You as Young as You Can Be? New York, Cliff Street Division of Harper Collins, 1999 (in press)
39. Koller, M., Kussmann, J., Lorenz, W., Jenkins, M., Voss, M., Arens, E., Richter, E., Rothmund, M.: Symptom reporting in cancer patients: the role of negative affect and experienced social stigma. Cancer *77*:983, 1996
40. Gadamer, H-G., Boehm, G.: Seminar: Die Hermeneutik und die Wissenschaften. Frankfurt, Suhrkamp Taschenbuch Wissenschaft, 1978, pp. 1–496
41. Gross, R., Löffler, M.: Hermeneutik. In: Prinzipien der Medizin: Eine Übersicht über Grundlagen und Methoden. Berlin, Springer, 1997, pp. 139–142
42. Gadamer, H-G.: Hermeneutik als praktische Philosophie. In: Rehabilitierung der praktischen Philosophie (vol. 1), M. Riedel, editor. 1973, Freiburg, Rombach, pp. 325–344
43. Smith, E.R.: Research design. In: Handbook of Research Methods in Social Psychology, H. Reis, C.M. Judd, editors. Cambridge, Cambridge University Press (in press)
44. Torrance, G.W., Feeny, D.: Utilities and quality-adjusted life years. Int. J. Technol. Assess. Health Care *5*:559, 1989
45. Troidl, H., Wechsler, A.S., McKneally, M.F.: How to choose a relevant endpoint. In: Surgical Research: Basic Principles and Clinical Practice (3rd ed.), H. Troidl, W.O. Spitzer, B. McPeek, D.S. Mulder, M. McKneally, A. Wechsler, C.M. Balch, editors. New York, Springer, 1998, pp. 303–320
46. Barkun, J.S., Barkun, A.N., Sampalis, J.S., Fried, G., Taylor, B., Wexler, M.J., Goresky, C.A., Meakins, J.L.: Randomised controlled trial of laparoscopic versus mini cholecystectomy. Lancet *340*:1116, 1992
47. Eypasch, E., Williams, I., Wood-Dauphinee, W., Ure, B.M., Schmülling, C., Neugebauer, E., Troidl, H.: Gastrointestinal quality of life index: development, validation and application of a new instrument. Br. J. Surg. *82*:216, 1995
48. Sanabria, J.R., Clavien, P.A., Cywes, R., Strasberg, S.M.: Laparoscopic versus open cholecystectomy: a matched study. Can. J. Surg. *36*:330, 1993
49. Majeed, A.W., Troy, G., Nicholl, J.P., Smythe, A., Reed, M.W.R., Stoddard, C.J., Peacock, J., Johnson, A.G.: Randomised, prospective, single-blind comparison of laparoscopic versus small-incision cholecystectomy. Lancet *347*:989, 1996
50. Coelho, J.C.U., de Araujo, R.P.M., Marchesini, J.B., Coelho, I.C.M.M., de Araujo, L.R.R.: Pulmonary function after cholecystectomy performed through Kocher's incision, a mini-incision, and laparoscopy. World J. Surg. *17*:544, 1993
51. Jan, Y.Y., Chen, M.F.: Laparoscopic versus open cholecystectomy: a prospective randomized study. J. Formosa Med. Assoc. *92*:243, 1993
52. Trondsen, E., Reiertsen, O., Andersen, O.K., Kjaersgaard, P.: Laparoscopic and open cholecystectomy, a prospective, randomized study. Eur. J. Surg. *159*:217, 1993
53. Dionigi, R., Dominioni, L., Benevento, A., Giudice, G., Cuffari, S., Bordone, N., Caravati, F., Carcano, G., Gennari, R.: Effects of surgical trauma of laparoscopic vs. open cholecystectomy. Hepatogastroenterology *41*:471, 1994
54. McMahon, A.J., Russell, I.T., Baxter, J.N., Ross, S., Anderson, J.R., Morran, C.G., Sunderland, G., Galloway, D., Ramsay, G., O'Dwyer, P.J.: Laparoscopic versus minilaparotomy: a randomised trial. Lancet *343*:135, 1994
55. Milheiro, A., Castro Sousa, F., Manso, E.C., Lettao, F.: Metabolic responses to cholecystectomy: open vs. laparoscopic approach. J. Laparoendosc. Surg. *4*:311, 1994
56. Redmond, H.P., Wastson, R.W.G., Houghton, T., Condron, C., Gordon, R., Watson, K., Bouchier-Hayes, D.: Immune function in patients undergoing open vs laparoscopic cholecystectomy. Arch. Surg. *129*:1240, 1994
57. McGinn, F.P., Miles, A.J.G., Uglow, M., Ozmen, M., Terzi, C., Humby, M.: Randomized trial of laparoscopic cholecystectomy versus mini-cholecystectomy. Br. J. Surg. *82*:1374, 1995
58. Charlo Dupont, T., Fernandez Martin, M., Tejido Sanchez, C.: Analisis de costes de la colecistectomia laparoscopica en comparacion con la abierta. Rev. Esp. Enferm. Dig. *87*:449, 1995

59. Downs, S.H., Black, N.A., Devlin, H.B., Royston, C.M.S., Russell, R.C.G.: A systematic review of the effectiveness and safety of laparoscopic cholecystectomy. Ann. R. Coll. Surg. Engl. *78*:241, 1996
60. Wood-Dauphinee, S., Troidl, H.: Endpoints for clinical studies: conventional and innovative variables. In: Principles and Practice of Research: Strategies for Surgical Investigators, H. Troidl, W.O. Spitzer, B. McPeek, D.S. Mulder, M.F., McKneally, A.S. Wechsler, C.M. Balch, editors. New York, Springer, 1991, pp. 151–168
61. Troidl, H., Spangenberger, W., Dietrich, A., Neugebauer, E.: Laparoskopische Cholecystektomie. Erste Erfahrungen und Ergebnisse bei 300 Operationen: eine prospektive Beobachtungsstudie. Chirurg *62*:257, 1991
62. Wright, D.M., Kennedy, A., Baxter, J.N., Fullarton, G.M., Fife, L.M., Sunderland, G.T., O'Dwyer, P.J.: Early outcome after open versus extraperitoneal endoscopic tension-free hernioplasty: a randomized clinical trial. Surgery *119*:552, 1996
63. Liem, M.S.L., van der Graaf, Y., van Steensel, C.J., Boelhouwer, R.U., Clevers, G-J., Meijer, W.S., Stassen, L.P.S., Vente, J.P., Weidema, W.F., Schrijvers, A.J.P., van Vroonhoven, J.M.V.: Comparison of conventional anterior surgery and laparoscopic surgery for inguinal-hernia repair. N. Engl. J. Med. *336*:1541, 1997
64. Ware, J.E., Jr.: The SF-36 health survey. In: Quality of Life and Pharmacoeconomics in Clinical Trials (2nd ed.), B. Spiker, editor. Philadelphia, Lippincott-Raven, 1996, pp. 337–345
65. Wennberg, J.E.: What is outcomes research? In: Modern Methods of Clinical Investigation, A.C. Gelijns, editor. Washington, National Academy Press, 1990, pp. 33–46
66. Koller, M., Kussmann, J., Lorenz, W., Rothmund, M.: Die Messung von Lebensqualität in der chirurgischen Tumornachsorge [Measurement of quality of life in surgical tumour follow-up]. Chirurg *65*:333, 1994
67. Wagner, K., Koller, M., Keill, A., Trott, D., Engenhart-Cabillic, R., Pfab, R., Lorenz, W.: Strahlentherapie beichirurgischen ünd nichtchirurgischen patienten: Therapieerwartüngen, lebensqualität ünd arzteinschätzügen. Chirurg *62*:252, 1998
68. Black, N., Petticrew, M., Ginzler, M., Flood, A., Smith, J., Williams, G., Davies, J., Doll, H., McPherson, K.: Do doctors and patients agree? Int. J. Technol. Assess. Health Care *7*:533, 1991
69. Fries, J.F., Singh, G.: The hierarchy of patient outcomes. In: Quality of Life and Pharmacoeconomics in Clinical Trials (2nd ed.), B. Spilker, editor. Philadelphia, Lippincott-Raven, 1996, pp. 33–40
70. Lorenz, W.: Standards klinisch-chirurgischer Studien [Standards of clinical-surgical trials]. In: Standards in der Chirurgie, O. Boeckl, H.W. Waclawiczek, editors. Munich, W. Zuckschwerdt Verlag, 1995, pp. 12–15
71. Reddick, E.J., Olsen, D.O.: Laparoscopic laser cholecystectomy: a comparison with mini-lap cholecystectomy. Surg. Endosc. *3*:131, 1989
72. Roumen, R.M.H., van Meurs, P.A., Kuypers, H.H.C., Kraak, W.A.G., Sauerwein, R.W.: Serum interleukin-6 and C-reactive protein responses in patients after laparoscopic or conventional cholecystectomy. Eur. J. Surg. *158*:541, 1992
73. Kelly, J.E., Burrus, R.G., Burns, R.P., Graham, L.D., Chandler, K.E.: Safety, efficacy, cost, and morbidity of laparoscopic versus open cholecystectomy: a prospective analysis of 228 consecutive patients. Am. Surg. *59*:23, 1993
74. Kesteloot, K., Penninckx, F.: The costs and effects of open versus laparoscopic cholecystectomies. Health Econ. *2*:303, 1993
75. Kocher, T., Herzog, U., Schuppisser, J.P., Tondelli, P.: Laparoskopische und offene Cholezystektomie bei 954 Patienten. Eine prospektive Evaluation. Helv. Chir. Acta *60*:761, 1993
76. Schauer, P.R., Luna, J., Ghiatas, A.A., Glen, M.E., Warren, J.M., Sirinek, K.R.: Pulmonary function after laparoscopic cholecystectomy. Surgery *114*:389, 1993
77. Williams, M.D., Sulentich, S.M., Murr, P.C.: Laparoscopic cholecystectomy produces less postoperative restriction of pulmonary function than open cholecystectomy. Surg. Endosc. *7*:489, 1993
78. Wittgen, C.M., Andrus, J.P., Andrus, C.H., Kaminski, D.L.: Cholecystectomy: which procedure is best for the high-risk patient? Surg. Endosc. *7*:395, 1993
79. Aktan, A.Ö., Büyükgebiz, O., Yegen, C., Yalin, R.: How minimally invasive is laparoscopic cholecystectomy? Surg. Laparosc. Endosc. *4*:18, 1994
80. Marti, M., Herzog, U., Kocher, T., Schuppisser, J.P., Ackermann, C., Tondelli, P.: Die Cholezystektomie heute: eine prospektive Analyse. Schweiz. Med. Wochenschr. *124*:763, 1994
81. Rau, H.G., Meyer, G., Maiwald, G., Schardey, M., Merkle, R., Lange, V., Schildberg, F.W.: Konventionelle oder laparoskopische Cholecystektomie zur Behandlung der akuten Cholecystitis? Chirurg *65*:1121, 1994
82. DaCosta, M.L., Burke, P.E., Qureshi, M.A., Grace, P.A., Brindley, N.M., Bouchier-Hayes, D.: Normal inspiratory muscle strength is restored more rapidly after laparoscopic cholecystectomy. Ann. R. Coll. Surg. Engl. *77*:252, 1995
83. Foster, D.: Prospective comparison of laparoscopic and open cholecystectomy in a community hospital. W. V. Med. J. *91*:270, 1995
84. Glaser, F., Sannwald, G.A., Buhr, H.J., Kuntz, C., Mayer, H., Klee, F., Herfarth, C.: General stress response to conventional and laparoscopic cholecystectomy. Ann. Surg. *221*:372, 1995
85. Griffith, J.P., Everitt, N.J., Lancaster, F., Boylston, A., Richards, S.J., Scott, C.S., Benson, E.A., Sue-Ling, H.M., McMahon, M.J.: Influence of laparoscopic and conventional cholecystectomy upon cell-mediated immunity. Br. J. Surg. *82*:677, 1995
86. Kane, R.L., Luri, N., Borbas, C., Morris, N., Flood, S., McLaughlin, B., Nemanich, G., Schultz, A.: The outcomes of elective laparoscopic and open cholecystectomies. J. Am. Coll. Surg. *180*:136, 1995
87. Köksoy, C., Kuzu, M.A., Kurt, I., Kurt, N., Yerdel, M.A., Tezcan, C., Aras, N.: Haemodynamic effects of pneumoperitoneum during laparoscopic cholecystectomy: a prospective comparative study using bioimpedance cardiography. Br. J. Surg. *82*:972, 1995
88. Pollock, A.V.: The rise and fall of the random controlled trial in surgery. Theor. Surg. *4*:163, 1989
89. Brook, R.H.: Practice guidelines: to be or not to be. Lancet *348*:1005, 1996
90. Büttner, H., Hansert, E., Stamm, D.: Statistical analysis, control and assessment of experimental results. In: Methods of Enzymatic Analysis (vol. 1), H.U. Bermeyer, editor. Weinheim, Verlag Chemie and San Diego, Academic Press, 1974, pp. 319–395
91. Koller, M., Lorenz, W.: Threats to double-blinding in sepsis trials: role of expectancies toward outcome. In: The Immune Consequences of Trauma, Shock and Sepsis: Mechanisms and Therapeutic Approaches (vol. 2), E. Faist, A.E. Baue, F.W. Schildberg, editors. Lengerich, Pabst, 1996, pp. 640–645
92. Ware, J.E., Jr.: The status of health assessment 1994. Annu. Rev. Public Health *16*:327, 1995
93. Leplège, A., Hunt, S.: The problem of quality of life in medicine. J.A.M.A. *278*:47, 1997
94. Eagle, K.A., Froehlich, J.B.: Reducing cardiovascular risk in patients undergoing noncardiac surgery. N. Engl. J. Med. *335*:1761, 1996
95. Eagle, K.A., Brundage, B.H., Chaitman, B.R., Ewy, G.A., Fleisher, L.A., Hertzer, N.R., Leppo, J.A., Ryan, T., Schlant, R.C., Spencer, W.H., III, Spittell, J.A., Jr., Twiss, R.D., Ritchie, J.L.: Guidelines for perioperative cardiovascular evaluation for noncardiac surgery. Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation *93*:1278, 1996
96. Roizen, M.F.: Preoperative evaluation. In: Anesthesia (4th ed.), R.D. Miller, editor. New York, Churchill Livingstone, 1994, pp. 827–882
97. Grogono, A.W., Woodgate, D.J.: Index for measuring health. Lancet *2*:1024, 1971
98. McPeek, B., Gasko, M., Mosteller, F.: Measuring outcome from anesthesia and operation. Theor. Surg. *1*:2, 1986
99. Stucki, G., Daltroy, L., Katz, J.N., Johannesson, M., Liang, M.H.: Interpretation of change scores in ordinal clinical scales and health status measures: the whole may not equal the sum of the parts. J. Clin. Epidemiol. *49*:711, 1996
100. Kaplan, R.M., Anderson, J.P.: The general health policy model: an integrated approach. In: Quality of Life and Pharmacoeconomics in Clinical Trials (2nd ed.), B. Spilker, editor. Philadelphia, Lippincott-Raven, 1996, pp. 309–322
101. Lohr, K.N., Aaronson, N.K., Alonso, J., Burnam, M.A., Patrick, D.L., Perrin, E.B., Roberts, J.S.: Evaluating quality-of-life and health status instruments: development of scientific review criteria. Clin. Ther. *18*:979, 1996
102. Aaronson, N.K., Cull, A.M., Kaasa, S., Sprangers, M.A.G.: The European Organization for Research and Treatment of Cancer (EORTC) modular approach to quality of life assessment in oncol-

ogy: an update. In: Quality of Life and Pharmacoeconomics in Clinical Trials (2nd ed.), B. Spilker, editor. Philadelphia, Lippincott-Raven, 1996, pp. 179–189
103. Olschewski, M., Schulgen, G., Schumacher, M., Altman, D.G.: Quality of life assessment in clinical cancer research. Br. J. Cancer *70*:1, 1994
104. Lorenz, W., Dick, W., Junginger, T., Ohmann, C., Ennis, M., Immich, H., McPeek, B., Dietz, W., Weber, D., Trial Group Mainz/Marburg: Induction of anaesthesia and perioperative risk: influence of antihistamine $H_1 + H_2$-prophylaxis and volume substitution with Haemaccel-35 on cardiovascular and respiratory disturbances and histamine release. Theor. Surg. *3*:55, 1988
105. Fairclough, D.L.: Summary measures and statistics for comparison of quality of life in a clinical trial of cancer therapy. Stat. Med. *16*:1197, 1997
106. Holm, S.: A simple sequentially rejective multiple test procedure. Scand. J. Statist. *6*:65, 1979
107. Koller, M., Kussmann, J., Elstein, A.S.: Double-blind clinical trials in sepsis research: problems connected with physicians' and patients' expectancies. Theor. Surg. *9*:52, 1994
108. Fleiss, J.L.: The statistical basis of metaanalysis. Stat. Methods Med. Res. *2*:121, 1993
109. Hartmann, L.C., Tschetter, L.K., Habermann, T.M., Ebbert, L.P., Johnson, P.S., Mailliard, J.A., Levitt, R., Suman, V.J., Witzig, T.E., Wieand, H.S., Miller, L.L., Moertel, C.G.: Granulocyte colony-stimulating factor in severe chemotherapy-induced afebrile neutropenia. N. Engl. J. Med. *336*:1776, 1997
110. Schein, M., Wittmann, D.H., Lorenz, W.: Duration of antibiotic treatment in surgical infections of the abdomen. Eur. J. Surg. *162*(Suppl. 576):1, 1996
111. Lorenz, W., Troidl, H., Fingerhut, A., Rothmund, M.: Introduction: the different ways to reach consensus. Eur. J. Surg. *162*:5, 1996
112. Weinstein, M.C., Fineberg, H.V.: Clinical Decision Analysis. Philadelphia, Saunders, 1980, pp. 1–351