

HEALTH TECHNOLOGY ASSESSMENT IN SURGERY

D. Bergqvist¹, M. Rosén²

¹ Department of Surgical Sciences, University Hospital, Uppsala, Sweden

² The Swedish Council on Health Technology Assessment and Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, Stockholm, Sweden

ABSTRACT

This review focuses on how surgical methods should be assessed from a health technology perspective. The use of randomized controlled trials, population based registries, systematic literature research and the recently published IDEAL method are briefly discussed.

Key words: Health technology assessment; HTA; randomized controlled trial; population based registry; surgery; learning curve; EVAR; GRADE

INTRODUCTION

The European organization EuNetHTA defines Health Technology as “the application of scientific knowledge in health care and prevention” (1). The Swedish Council on Health Technology Assessment (SBU, Statens Beredning för Medicinsk utvärdering) has adopted the term health technology as the use of systematic knowledge in techniques and methods to prevent, diagnose, treat and monitor diseases as well as rehabilitation. Health technology assessment (HTA) is truly multidisciplinary and summarises medical (including safety, efficacy and effectiveness), social, economic, organizational and ethical evaluation of properties, effects and/or impacts of health care technology, and this must be made in a systematic, transparent and reproducible manner, an important aim being to support decision makers. When the process has gone thus far political priorities come into work. Since 1985 there is a journal for scientific articles within this field (International Journal of Technology Assessment in Health Care). There are several categories of actions taken such as medical (where pharmacological substances are included), surgical with various types of interventions sometimes with the use of implantation products or de-

vices, screening programmes, psychological/behavioural interventions, diagnostic modalities etc.

Full SBU reports assess or evaluate one disease or a complex of symptoms. It is an extensive process and usually takes 2–4 years to complete, and it ends up with a scientific text of several hundred pages. SBU Alert is a system with rapid evaluation of the evidence for new methods and technologies, a similar function as horizon scanning or early warning system. The strengths with this system are the evaluation of one method at a time and the temporal effectiveness, usually finished in about 6–9 months.

Table 1 shows all SBU projects with a surgical interest from the start in 1987. This paper focuses on diagnostic and treatment procedures for surgical interventions.

PROBLEMS WITH NON-PHARMACOLOGICAL METHODS

When it comes to evaluation of pharmacological substances the regulatory pattern and requirements of authorities are well defined and there are strict rules how to perform a randomized study. How to evaluate other technologies is less well defined, and there have been less strict rules and often up to individual interventionists to test new surgical procedures and implant prostheses, grafts etc. A CE mark on a product does not guarantee testing in properly designed trials. The role of innovative, enthusiastic and brave colleagues has been important, perhaps in spite of initial poor clinical outcome. This can be exemplified by

Correspondence:

David Bergqvist, Professor Emeritus
Vascular Surgery Department of Surgical Sciences
Uppsala University
SE - 751 85 Uppsala, Sweden
Email: david.bergqvist@surgsci.uu.se

TABLE 1
Published SBU reports concerning surgical interventions.

Interventions	Publication year
Preoperative routines (nr 101)	1989
Bone marrow transplantation (nr 109)	1991
Surgery for epilepsy (nr 110)	1991
Coronary artery bypass graft and percutaneous transluminal coronary angioplasty (nr 120)	1994
Hysterectomy – ratings of appropriateness (nr 125)	1995
Surgical treatment of rheumatic diseases (nr 136)	1998
Chest pain: surgery, balloon dilation, drugs (nr 140)	1998
Laparoscopic surgery for treatment of colon cancer (Alert)	1998
Endovascular surgery for abdominal aortic aneurysm (Alert)	1999
Pallidotomy in Parkinson's disease (Alert)	1999
Autologous chondrocyte transplantation in treating cartilage damage in the knee (Alert)	1999
Xenotransplantation – renal transplantation example (Alert)	1999
Back and neck pain (nr 145)	2000
Transplantation of stem cells from umbilical cord blood (Alert)	2001
Stem cell transplantation for metastasized kidney cancer (Alert)	2001
Cell transplantation in Parkinson's disease (Alert)	2001
Stereotactic radiosurgery in treating arteriovenous malformations of the brain (Alert)	2001
Venous thromboembolism (nr 158)	2002
Obesity – problems and interventions (nr 160)	2002
Surgery for adult acquired flatfoot deformity (Alert)	2002
Drug eluting stents in coronary arteries (Alert)	2004
Gastric pacing for the treatment of obesity (Alert)	2004
Implantable defibrillator (Alert)	2006
Bilateral cochlear implantation (CI) in children (Alert)	2006
Cardiac Resynchronization therapy (CRT) in chronic heart failure (Alert)	2007
Obstructive sleep apnoea syndrome (nr 184)	2007
Dyspepsia and gastro-oesophageal reflux (nr 185)	2007
Peripheral arterial disease – diagnosis and treatment (nr 187)	2007
Laser eye surgery for the correction of refractive errors (Alert)	2007
Screening for abdominal aortic aneurysm (Alert)	2008
Open angle glaucoma – diagnosis, follow-up and treatment (nr 190)	2008
Antibiotic prophylaxis for surgical procedures (nr 200)	2010
Methods of diagnosis and treatment in endodontics (nr 204)	2010
Percutaneous vertebroplasty and balloon kyphoplasty in treating painful osteoporotic vertebral compression fractures (Alert)	2011
Benign prostatic hyperplasia (nr 209)	2011

early development of neurosurgery, open thoracic/ cardiac surgery and organ transplantation. Obviously there has been a methodological development without randomized studies and modern evaluation technology, and it is important that we define what cannot be questioned in today's treatment armamentarium. This is the basis against which new technologies

must be evaluated. Every speciality should define this level of basic knowledge. In this process common sense must not be forgotten but it is difficult to quantify. The lack of RCTs on the benefit of parachutes in air accidents is a drastic example (2). To exemplify, surgical treatment of femoral neck fractures is hardly to be questioned but when new methods, i.e. new

osteosynthetic devices, are developed, they must be properly tested. Regarding drainage of abscesses and suture of perforations the experience over time has given the answer. The same can be said about surgical treatment of ruptured abdominal aortic aneurysms, where no treatment goes with 100% mortality and with open surgery it is less than 50%, in the recent Swedvasc report 35% (3). Intuitively endovascular treatment should be even better because of the less surgical trauma, but the correct way to test this hypothesis is through a randomized trial comparing open and endovascular surgery.

This paper will deal with the use of health technology assessment methodology primarily exemplified by vascular surgery. An important issue, which will not be further discussed, is how ethical perspectives should be integrated (4–5).

RANDOMIZED CONTROLLED TRIAL IN SURGERY

The randomized controlled trial (RCT) has its strength in being truly experimental, ideally studying one fac-

tor, having all other factors constant and being evenly distributed between groups by the random allocation. Through the randomization process the risk of systematic bias or errors is minimized. The more strict the inclusion and exclusion criteria, however, the poorer will be the external validity and generalizability. One problem supporting this point is the often slow inclusion rate, much slower than the number of potential patients at risk, especially in large multicenter trials (6). This indicates the risk of selection bias. Often surrogate endpoints are used to keep the sample size small, the problem then being how relevant the surrogate is for the clinical outcome. One concern, true also for evaluation of pharmacological substances, is that there must be a genuine uncertainty which alternative is best for the patient. When the surgeon is not convinced which method is superior, it should be possible and also ethical to randomize – the principle of the “grey area of uncertainty” (7–8). The difficulty to perform blind studies in surgery makes fully independent monitoring boards vital – both concerning outcome and safety.

The pathophysiological background and physiological rationale motivating a specific treatment must be clear, a phenomenon illustrated by the external to internal carotid artery bypass study to prevent stroke (9). In that study the bypass, aimed at increasing blood flow to the brain, opened a free channel for embolization as the thrombotic ulceration in the carotid bifurcation was not removed. As could be expected the surgical group came out with significantly higher rate of stroke and death. We also have to be open for totally new etiologic explanations for diseases, exemplified in the change from large surgical resections to antibiotic eradication of *Helicobacter pylori* in patients with gastric ulcer.

Table 2 illustrates some problems which are met in surgical RCTs.

One great concern when a new technology is to be evaluated is the learning curve, which involves the surgical handcraft but also development of devices and instruments as well as continuous refinement of patient selection and team performance and communication regarding peri- and postoperative care. During the learning process randomization can start too early, when the methodology is still under development, the results rapidly becoming obsolete, and it is not clear how to deal with acquisition of skill and experience during the development phase of a surgical procedure. An example can be illustrated in Table 3 concerning endovascular aneurysm repair (EVAR). The question is when during the technical development a potential RCT against open repair is optimal to start. One study design, which has aimed at taking such problems into account, is the so called tracker trial (10). The tracker trial combines advantages with RCTs and registries and is designed to track the study progress over time by having a flexible protocol and sophisticated interim analysis. The sample size is open for modification and is not absolutely fixed from the study start.

Another type of RCT is where the patients are randomized to the surgeon or to the surgical team, with the best knowledge of the technologies, which are to

TABLE 2

Problems with RCT in surgery

-
- Reluctance to participate (from the patient and from the surgeon)
 - Ethical problems (i.e. sham operations, the placebo effect of surgery)
 - Methodological problems (observation bias)
 - Impossibility to perform blind studies making blinded assessment, blinded adjudication and participation of data and safety monitoring boards necessary
 - Problems with standardization (learning curve, improvement and modifications in technology over time, different surgeons)
 - Problems with external validity
 - Logistic problems
 - Lack of official control bodies
 - Difficulties to get economic support (cf. sponsored pharmacological studies)
-

TABLE 3

Development of endovascular abdominal aortic aneurysm repair (EVAR)

-
- Only proximal fixation of the graft leading to distal leakage
 - Proximal and distal fixation with smooth stents with the risk of graft migration
 - Use of stent hooks anchoring the graft in the aorta but requirement of a rather long aneurysmal neck
 - Stent support of the entire graft to avoid kinking
 - Stent coverage of the renal artery orifices making it possible to treat aneurysm without or with a short neck
 - Branched stent graft with grafts extending into the renal arteries making the fixation even more stable
-

be compared. This means that the patients have to be geographically transported, perhaps between hospitals at different locations, and this has logistic as well as economic problems which must be solved (11). This study type was suggested more than 30 years ago but is underused.

It is easier to obtain economic support for pharmacological than non-pharmacological studies through industry sponsorship. However, there seems to be a systematic bias overestimating effects and favouring products which are made by companies funding the research (12). On the other hand, it is more difficult to obtain funding for evaluation of non-pharmacological methods. Having observed the association between industry sponsorship and proindustry conclusions, it is, however, also fair to remember that the quality of industry-supported studies is often high (13). One problem with industry-sponsored trials is that there could be a company induced delay in publication, especially if results are negative, and selection of data. Especially if negative trials are not published, this could distort conclusions in meta-analyses (14). Small and underpowered RCTs, which is often the case in surgery, may contribute to unreliable results of meta-analyses (15). A further problem is the phenomenon of multiple publications of the same study, which is not always easily detected (16).

POPULATION BASED REGISTRIES

There are, however, problems with RCTs in surgery. One is the generalizability of results from the RCTs to the population at risk. As already stated the inclusion rate sometimes casts doubt on this issue. Another problem is what happens with the results obtained in an RCT, when the technique or innovation is disseminated outside the well defined study situation. A third problem is to detect rare complications and side effects which are not usually seen in an RCT, simply because the sample size is too small, and the majority of RCTs are not primarily designed and dimensioned to study side effects.

A recent systematic review showed no difference on average in the risk estimate of adverse effects of an intervention derived from meta-analyses of RCTs and observational studies (17). This suggests that systematic reviews of adverse effects should not be restricted to specific study type.

There is also a substantial risk of conducting meta-analyses of small RCTs where the purpose of studying the specified outcome is not defined in advance or if the follow-up method or time are not adequately described. This is evident from the case of aprotinin to reduce blood loss during coronary artery bypass surgery, where a meta-analysis of 52 RCTs in a Cochrane review could not display any excess mortality in spite of the fact that several observational studies showed this (18–19). Later, a large RCT (20) also showed excess mortality and aprotinin was therefore withdrawn from the market. The aprotinin saga illustrates overconfidence in small RCTs of inferior quality compared to well-conducted observational studies.

Therefore it is important to follow what happens outside the RCT situation using a prospective population based registry with a protocol with well defined variables. Then it is possible to compare regions within one country (as in the Swedish Vascular Registry, Swedvasc) and also to make comparisons between countries (as in the Vascunet cooperation, where a recent publication has focused on abdominal aortic aneurysm (21)).

Differences in such registries can of course be due to differences in prevalence and incidence of diseases but are more often due to differences in practice and indications for treatment (NB that the background concerning risk factors etc is comparable). This can lead to further in depth analysis to find explanations and can be used to generate new hypotheses and new studies.

When a method has been introduced on the basis of good evidence after RCTs or systematic literature research, it is important to analyze what happens outside the well controlled trials, that is in clinical reality, and here population based quality registries have an important task, they are increasingly used and also demanded by health care authorities (22). When establishing such registries it is of utmost importance that the profession is involved so that registries are developed with the highest scientific scrutiny.

Although registries have been around for more than 25 years there are several general problems, which have to be solved. Registries should be automatically linked to patient records, which, however, have logistic as well economic implications. The coordination on a national level is at the moment sub-optimal concerning IT support, medicolegal problems, data linkage between various registries etc. There is an urgent need to develop quality indicators and common definitions to make comparisons between hospitals, regions etc possible and meaningful. Existing registries are a rich source for research activities, but are underused. There should be an automatic linkage between RCTs and population based registries. RCTs and population based registries both have their advantages and disadvantages and should be looked upon as complementary research strategies (23).

SYSTEMATIC LITERATURE EVALUATION

With the increasing number of publications and the almost impossible task for individual doctors to synthesize the information and to summarize the evidence of new health technologies, the systematic literature search as originally described by Archie Cochrane has been of great help, since 1992 established in the International Cochrane Collaboration. When doing such a critical evaluation it is important to have very well defined and explicit methods for the strategy of searching the literature and for grading the scientific quality of the individual papers, where well performed RCTs are of highest value and case reports of lowest. The process of evaluating published evidence is quite rigorous. In the next step there must be a system where the results in the collected literature

TABLE 4

Projects on venous thromboembolism (VTE) where well performed studies are needed

-
- The effect of extended prophylaxis on clinically manifest thromboembolic morbidity and mortality after surgery for cancer
 - The prophylactic effect of low molecular weight heparins, oral inhibition of factors II and X on pulmonary embolism detected with spiral CT
 - The effect of thrombolysis in patients with right ventricular dysfunction and stable haemodynamics
 - Comparison between low molecular weight heparin and vitamin K antagonist to treat patients with VTE and malignancy
 - The optimal duration of treatment for VTE with regard to various risk factors
 - Prospective follow up of patients with negative thoracic CT with regard to pulmonary embolism
 - Do all females with previous VTE need thromboprophylaxis during pregnancy
 - What is the risk for caval thrombosis during pregnancy and can MR be used to define the proximal extension.
 - Risk and prevention of VTE during spinal surgery
 - The relation between venous and arterial thrombosis postoperatively
-

TABLE 5

Projects on lower limb ischaemia where well performed studies are needed

-
- The relation between vascular surgical activity and amputations on a population basis
 - RCT: treatment of intermittent claudication, intervention or training in addition to best medical treatment
 - RCT: treatment of critical limb ischaemia, open/endovascular intervention or best medical treatment
 - A management study on the value of duplex ultrasonography as the sole diagnostic tool before intervention
 - The effect of anticoagulation on the healing of diabetic foot ulcers
 - The effect of aspirin alone or combined with clopidogrel to prevent postreconstructive thrombosis (in below knee reconstruction)
 - Thrombolysis or open surgery in acute ischaemia
 - Optimal pharmacological prophylaxis for longterm postreconstructive patency
 - Quality of life after revascularization
 - RCT: subintimal angioplasty or surgical bypass in patients with critical limb ischaemia
-

are used to give an evidence mark, where the highest degree normally can be transferred into a clinical recommendation and low degrees should be a stimulation for further research. Today there are several systems for marking the evidence. The Swedish Council on Health Technology Assessment (SBU) has recently adopted the GRADE system (24).

When performing systematic literature overviews a depressing fact is that the majority of what is published is of low quality, often less than 10% for a given problem is of acceptable quality.

When a systematic review is performed, by necessity a number of questions and unresolved problems will be identified. As an example two SBU projects with the same chairman were chosen. The questions identified as well as low evidence statements were listed and all project members prioritized listed items. In Tables 4 and 5 the ten items judged to be of highest priority are illustrated. SBU is at present working on already finished projects to identify clinically relevant problems, where the evidence is low or insufficient. By identifying knowledge gaps research could hopefully be focused on questions where we lack answers and stimulate well designed studies, bringing the evidence to a higher level. When a systematic review has identified gaps in knowledge and uncertainties about effects, controlled trials should be the logical step forward. SBU has started to develop a database for interventions where we lack knowledge or where uncertainties exist (25). The database is primarily based on a similar database "DUET" in UK (26).

IDEAL

The IDEAL method has recently been suggested when evaluating surgery and other invasive therapies (27–29). IDEAL is the acronym for Idea, Development, Exploration, Assessment and Long-term study and resulted from conferences by the Balliol Collaboration, supported by the Nuffield Department of Surgery in Oxford, England, and the Department of Surgery, McGill University in Montreal, Canada, the initiative taken by Jonathan Meakins, professor of surgery in Oxford. During the idea phase there is a proof of concept studies and reports of small series, often as structured case reports. It is important to report new methods, whether successful or not. If there seems to be a benefit, studies are focused on technical development and there may be modifications and the safety aspects are also brought into focus. This is the phase where there still may be a learning curve with alterations over time, also because of increasing experience with the method. During the exploration phase the number of included patients has increased and ideally an RCT should be performed. One controversial issue is when there is appropriate timing to start an RCT. During the assessment phase the intervention has become stable in clinical outcome on an intermediate basis which is important to report. The effectiveness of new technology should be tested against current standards. During the long-term study phase the method is diffused into the population of interest and during this surveillance phase all consecutive patients should be prospectively included into a registry, which has been discussed above. During this phase rare side effects may be identified. It is important that standardized protocols are used, and that there is adherence to reporting standards, such as has been advocated in CONSORT and STROBE (30–32).

CONCLUDING REMARKS

The motivation for proper health technology assessment is to make it possible for decision makers to choose the best options for diagnosis and treatment. In this way HTA is a bridge between evidence based research and the process of decision-making. The methodology has developed considerably during the 30–40 years HTA has been around. Its importance does not only lie in evaluating technology for use in the health sector but also to stimulate research, where there is no evidence that a method works or to stop methods where there is evidence that they do not work. One problem still is the lack of awareness how HTA functions and to overcome this problem is a major challenge.

REFERENCES

1. EuNetHTA: EuNetHTA 2009 <http://www.eunethanet/Public/HTA/>
2. Smith GC, Pell JP: Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. *BMJ* 2003;327(7429):1459–1461. PMID: 300808
3. Troëng T: Nationella kvalitetsregister för kärllkirurgi. Swedvasc's Årsrapport för 2011 för 2010 års verksamhet. *Swedvasc*. 2011
4. ten Have H: Ethical perspectives on health technology assessment. *Int J Technol Assess Health Care* 2004;20(1):71–76
5. Hofmann B: Toward a procedure for integrating moral issues in health technology assessment. *Int J Technol Assess Health Care* 2005;21(3):312–318
6. Bergqvist D: Introduction of new technology: the surgical point of view. *Scand J Surg* 2009;98(1):3–7
7. Djulbegovic B: Acknowledgment of uncertainty: a fundamental means to ensure scientific and ethical validity in clinical research. *Curr Oncol Rep* 2001;3(5):389–395
8. European Carotid Surgery Trialists' Collaborative Group. MRC European Carotid Surgery Trial: interim results for symptomatic patients with severe (70–99%) or with mild (0–29%) carotid stenosis. *Lancet* 1991;337(8752):1235–1243
9. The EC/IC Bypass Study Group. Failure of extracranial-intracranial arterial bypass to reduce the risk of ischemic stroke: Results of an international randomized trial. *N Engl J Med* 1985;313(19):1191–1200
10. Lilford RJ, Braunholtz DA, Greenhalgh R, Edwards SJ: Trials and fast changing technologies: the case for tracker studies. *Bmj* 2000;320(7226):43–46
11. van der Linden W: Pitfalls in randomized surgical trials. *Surgery* 1980;87(3):258–262
12. Lexchin J, Bero LA, Djulbegovic B, Clark O: Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ* 2003;326(7400):1167–1170. PMID: 156458.
13. Bekelman JE, Li Y, Gross CP: Scope and impact of financial conflicts of interest in biomedical research: a systematic review. *JAMA* 2003;289(4):454–465
14. Egger M, Smith GD: Bias in location and selection of studies. *BMJ* 1998;316(7124):61–66. PMID: 2665334
15. Rerkasem K, Rothwell PM: Meta-analysis of small randomized controlled trials in surgery may be unreliable. *Br J Surg* 2010; 97(4):466–469
16. Melander H, Ahlqvist-Rastad J, Meijer G, Beermann B: Evidence b(i)ased medicine-selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. *BMJ* 2003;326(7400):1171–1173. PMID: 156459
17. Golder S, Loke YK, Bland M: Meta-analyses of adverse effects data derived from randomised controlled trials as compared to observational studies: methodological overview. *PLoS Med* 2011;8(5):e1001026. PMID: 3086872
18. Henry DA, Carless PA, Moxey AJ, O'Connell D, Stokes BJ, McClelland B, et al: Anti-fibrinolytic use for minimising peri-operative allogeneic blood transfusion. *Cochrane Database Syst Rev* 2007(4):CD001886
19. Rosen M: The aprotinin saga and the risks of conducting meta-analyses on small randomised controlled trials – a critique of a Cochrane review. *BMC Health Serv Res*. 2009;9:34. PMID: 2657782
20. Fergusson DA, Hebert PC, Mazer CD, et al: A comparison of aprotinin and lysine analogues in high-risk cardiac surgery. *N Engl J Med* 2008; 358(22):2319–2331
21. Mani K, Lees T, Beiles B, et al: Treatment of abdominal aortic aneurysm in nine countries 2005–2009: a Vascunet report. *Eur J Vasc Endovasc Surg* 2011; 42(5):598–607
22. Rosén M: Guldgruvan i hälso- och sjukvården. Översyn av de nationella kvalitetsregistren. Förslag till gemensam satsning 2011–2015: SKL Stockholm; 2010
23. Bergqvist D, Bjorck M, Sawe J, Troeng T: Randomized trials or population-based registries. *Eur J Vasc Endovasc Surg* 2007; 34(3):253–256
24. Guyatt GH, Oxman AD, Vist GE, et al: GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336(7650):924–926. PMID: 2335261
25. www.sbu.se
26. <http://www.library.nhs.uk/duets/>
27. Ergina PL, Cook JA, Blazeby JM, et al: Challenges in evaluating surgical innovation. *Lancet* 2009;374(9695):1097–1104. PMID: 2855679
28. McCulloch P, Altman DG, Campbell WB, et al: No surgical innovation without evaluation: the IDEAL recommendations. *Lancet*. 2009;374(9695):1105–1112
29. Barkun JS, Aronson JK, Feldman LS, et al: Evaluation and stages of surgical innovations. *Lancet* 2009;374(9695):1089–1096
30. Altman DG: Better reporting of randomised controlled trials: the CONSORT statement. *BMJ* 1996;313(7057):570–577. PMID: 2352018
31. von Elm E, Altman DG, Egger M, Pocock SJ, Gotsche PC, Vandenbroucke JP: The Strengthening of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007;147(8):573–577
32. Moher D, Schulz KF, Altman D: The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001;285(15):1987–1991.

Received: January 27, 2012