



Wie werden Nebenwirkungen im HTA (Health Technology Assessment) bewertet?

Diskussionsforum 2010 der SEG 6 "Arzneimittelversorgung"
Dortmund, 14. September 2010

Dr. med. Dagmar Lühmann
Institut für Sozialmedizin, Universität zu Lübeck



Definition von HTA (Health Technology Assessment)

HTA ist eine Form der Politikfeldanalyse die systematisch kurz- und langfristige Konsequenzen der Anwendung einer medizinischen Technologie, einer Gruppe verwandter Technologien oder eines technologiebezogenen Sachverhalts untersucht

Ziel von HTA ist die Unterstützung von Entscheidungen in Politik und Praxis

Grundlegend für HTA ist die Ausrichtung auf Entscheidungsfindung sowie der multidisziplinäre und umfassende Ansatz



Was ist HTA?

(modifiziert nach: EUnetHTA 2007)

Health technology assessment (HTA) ist ein **multidisziplinärer Prozess**, im Laufe dessen Informationen zu medizinischen, gesellschaftlichen und ökonomischen Auswirkungen sowie ethischen Aspekten von gesundheitsrelevanten Technologien **systematisch, transparent, unverzerrt und methodisch belastbar** zusammengestellt werden.

Sein Ziel ist die **Bereitstellung einer Informationsbasis** als Grundlage für die Gestaltung einer sicheren, wirkungsvollen, kostenbewussten und patientenzentrierten Gesundheitsversorgung.

Trotz seiner strategischen Zielsetzung arbeitet HTA ausschließlich mit den jeweils relevanten wissenschaftlichen Methoden.



Anlass, Akteure und Adressaten von HTA

Policy Question: Anstehende, systemrelevante Entscheidung -
meistens Kostenübernahmeentscheidung – häufig Zeitdruck

Akteure: (unabhängige) HTA-Institutionen, wissenschaftliche
Forschungsabteilungen, gesundheitspolitische
Entscheidungsinstanzen, Kostenträger, Leistungserbringer,
Hersteller

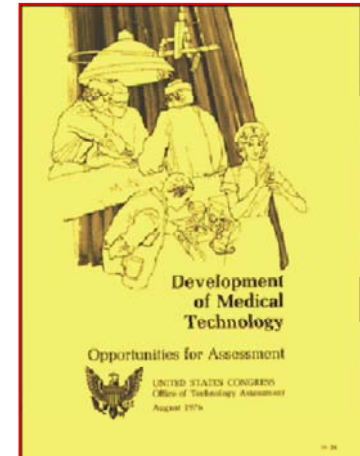
Adressaten: Entscheidungsträger in den relevanten Institutionen

HTA \neq Entscheidung!

Meilensteine in der Entwicklung von HTA

(modifiziert und ergänzt nach Perleth, 2000)

1968	Der Begriff "Technology Assessment" zum ersten Mal benutzt
1972	U.S. Kongreß "Office of Technology Assessment" (OTA) etabliert
1975	OTA etabliert Gesundheitsprogramm
1976	Schwedische Studie zur Computertomographie (CT)
1978	OTA-Studie zur Computertomographie
1980	Dänischer Bericht zu HTA
80er	Private Organisationen entstehen in den USA
1982	Australisches "National Health Technology Advisory Panel"
1984	Katalonisches Komitee für die Evaluation von Hochtechnologien
1985	Niederländisches "White Paper" zu HTA
1985	Gründung der International Society for Technology Assessment in Health Care (ISTAHC)
1987	Swedish Council on Health Care Technology Assessment (SBU) gegründet
1988	Conceil d'Évaluation des Technologies Sanitaires (CÉTS) in Quebec, Kanada etabliert
1990	Nationale Programme in Frankreich und Kanada entstehen
1990	Neue Regulation für die U.S. Agency for Health Care Policy and Research (AHCPR) – Outcome-Forschung
1992	Neues nationales Programm in Großbritannien
1993	Gründung: International Network of Agencies in Health Technology Assessment (INAHTA)
1994 -97	Europäische Projekte EUR-ASSESS und HTA-Europe
1994	Neue nationale Programme in Spanien und Italien
1995	Neues nationales Programm in Finnland
ab 1995	deutsches HTA-Forschungsprojekt
1997	Neue nationale Programme in Dänemark und Norwegen
2000	Start des Netzwerks European Collaboration in HTA (ECAH/ECHTA)
2000	Deutschland: Gesundheitsreformgesetz mit HTA Auftrag an DIMDI
2004 ff	Institut für für Qualität und Wirtschaftlichkeit im Gesundheitswesen
2006/08	EUNetHTA-Project
2009 ff	EUNetHTA Joint Action





Dimensionen von HTA

- Gesundheitsproblem und Technologiegebrauch
- Technische Beschreibung, technische Charakteristika
- Sicherheit
- Erwünschte Effekte (clinical effectiveness)
- Kosten, ökonomische Evaluation
- Ethische Aspekte
- Organisatorische Aspekte
- Soziale Aspekte
- Rechtliche Aspekte

(EUnetHTA 2007, Core Model for medical and surgical interventions™)

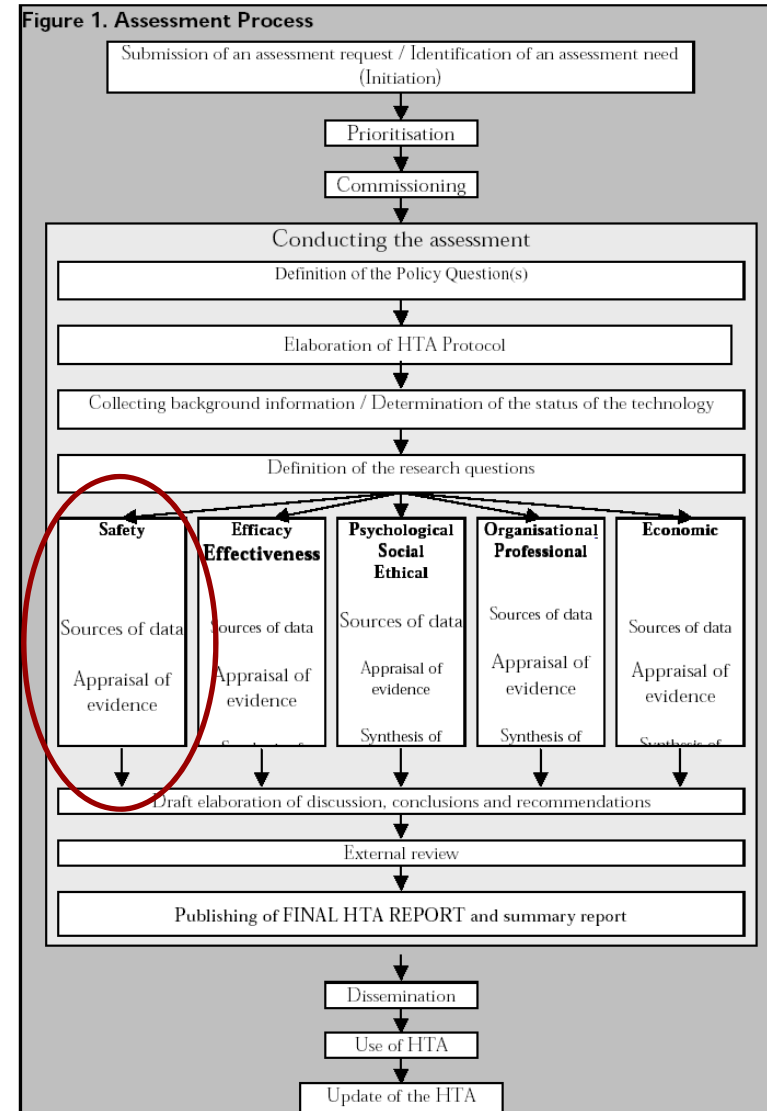
Working Group 4 Best Practice in Undertaking and Reporting HTA

To develop and disseminate best practice in undertaking and reporting assessments. To identify needs for methodological development

The ECHTA/ECAHI Project

Grant Agreement No. SI2.122594 (99CVF3-508)

- "Policy Question"
- Evaluationsrahmen ("Scoping")
- Protokoll
- Recherchen
- Kritische Bewertung (Validität, Relevanz)
- Synthese
- Schlussfolgerungen
- Peer Review
- Dissemination
- Aktualisierung





NICE Technology Appraisal

Beteiligte an Bewertung (assessment), Kommentierung, Analyse

- Unabhängige Bewertungsgruppe
(Assessment Group, Evidence Review Group)
- Technologiehersteller, -sponsoren
- (organisierte) Patientenvertreter
- Gesundheitsprofessionen/ Leistungsanbieter
- klinische Experten/ Patientenexperten



Inhalte der Sicherheitsbewertung

Informationen zu unerwünschten Effekten: Art, Dauer/ Schweregrad, Inzidenz, Zeitabfolge, vulnerable Gruppen

Auswirkungen unerwünschter Effekte: Morbidität, Mortalität, Beeinträchtigung, Lebensqualität der Betroffenen

Bewertung der Auswirkungen: Utilities

Auswirkungen/Anforderungen auf Versorgungssystem
Vergleich mit alternativen Technologien
Umweltverträglichkeit, Arbeitssicherheit
Möglichkeiten der Schadenminimierung

→ **Methodik: Systematischer Review**

(Nach: EUnetHTA 2007, Core Model for medical and surgical interventions™)

(EUnetHTA, 2007)



Systematischer Review zu unerwünschten Effekten I

Fragestellung: Nahe am entscheidungsrelevanten Kontext!

Stellenwert der Sicherheitsbewertung?

- unsichere Wirksamkeit/ unerwünschte Effekte wahrscheinlich
- vielversprechende Wirksamkeit/ schwere Sicherheitsbedenken
- keine unmittelbaren positiven Effekte, unerwünschte Effekte möglich

- alternative Technologien mit unterschiedlichen Sicherheitsprofilen
- alternative Technologien mit unterschiedlichen Wirksamkeits- und Sicherheitsprofilen



Systematischer Review zu unerwünschten Effekten II

Fragestellung: Nahe am entscheidungsrelevanten Kontext!

Auswahl der relevanten Zielgrößen:

- Problem: möglicherweise nicht alle relevanten Zielgrößen bekannt
- Fokussierter Ansatz: bewertet eine spezifische Zielgröße – einfacher zu bearbeiten, aber liefert kein "Sicherheitsprofil"
- Breiter Ansatz: gibt Überblick über potentielle Probleme. Wegen Datenheterogenität schwierig (interpretierbar), ressourcenintensiv – ggf. als Überblick sinnvoll im Scoping-Prozess
Cave: Heterogenität von Terminologie und Klassifikation!



Systematischer Review zu unerwünschten Effekten IIIa

Methoden: Auswahl der Datenquellen

Zielsetzung: Hohe Ergebnissicherheit – Einfluss von Bias minimal

Potentielle Datenquellen:

Literatur (RCT, andere Studiendesigns)

Herstellerdaten, Zulassungsunterlagen

Monitoring Systeme

Register (Prozeduren, Patienten)

Surveillance Berichte

Routinedaten

....

(Nach: Jefferson and Demicheli: Balancing benefits and harms in health care:observational data on harm are already included in systematic reviews. BMJ 327(7417):750, 2003

(LUNENIA, 2007)



Hypothese: Unterschiedliche Datenquellen berichten unterschiedliche Raten unerwünschter Effekte

- Amiodaron zur Behandlung von Herzrhythmusstörungen
- Meta-Analysen von RCT; publizierte Fallberichte; WHO International Drug Monitoring Programme
- Fokus auf 8 "Typen" von unerwünschten Effekten: thyreoidal, respiratorisch, hepatisch, dermal, neurologisch, kardial, gastrointestinal, visuell

Loke et al.: A comparison of three different sources of data in assessing the frequencies of adverse reactions to amiodarone. British J of Clinical Pharmacology 57(5): 616, 2004

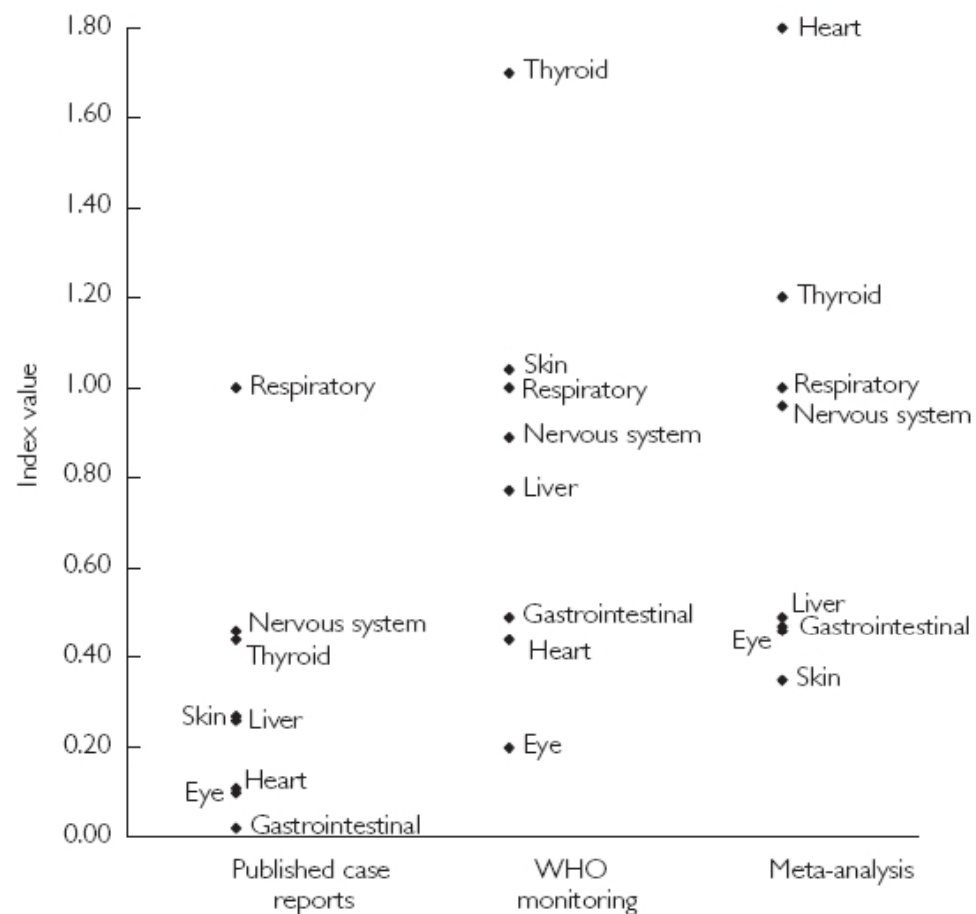


Figure 1

Rank order of adverse drug reactions, according to data source

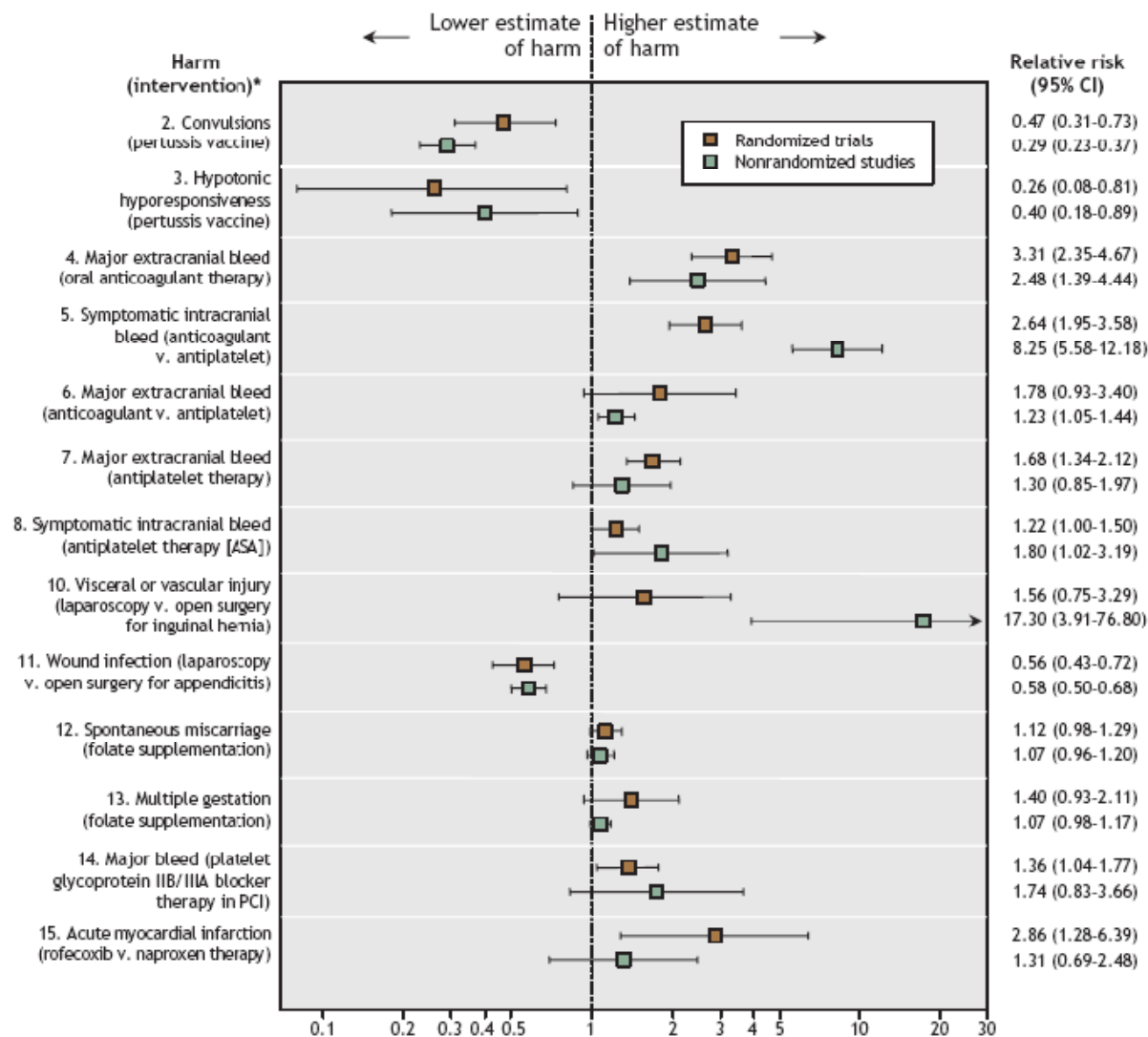


Frage: Kommen RCT und Non-RCTs zu unterschiedlichen Risikoeinschätzungen für unerwünschte Effekte?

- Systematische Reviews von RCT, ≥ 4000 Patienten, Bericht zu klar definierten unerwünschten Wirkungen (Quelle: Cochrane Library)
- Nicht-randomisierte Studien zur gleichen Indikation und mit Berichten zu den gleichen unerwünschten Effekten, ≥ 4000 Patienten (Quelle: PubMed)
- Absolute Risiken: Vergleich zu 8 unerwünschten Effekten
→ in 6/8 Vergleichen höhere Raten in RCT
- Relative Risiken: Vergleiche zu 13 unerwünschten Effekten

(Papanikolaou et al.: Comparison of evidence on harms of medical interventions in randomized and nonrandomized studies, CMAJ, 2006)

Fig. 1: Comparison of relative risk estimates for specific harms of medical interventions from randomized and nonrandomized studies.



Papanikolaou et al.: Comparison of evidence on harms of medical interventions in randomized and nonrandomized studies, CMAJ, 2006

Systematischer Review zu unerwünschten Effekten IIIb

Methoden: Auswahl der Datenquellen

Table 1

Strengths and weaknesses of studies included in systematic reviews of harmful effects

Method/study design	Strengths	Weaknesses
Case report	Early warning	Bias, differing case definitions, lack of comparators
Passive and active surveillance	Early warning or detection of rare events	Bias, differing case definitions, lack of comparators
Ecological study	Powerful, cheap	Difficulty in interpretation, confounding, bias, differing case definitions
Case crossover and case based studies	No need for independent controls	Lack of wide acceptance, bias, differing case definitions
Multiple time series	Flexible, powerful	Credibility, bias, differing case definitions
Case-control study	Can test hypotheses, especially rare events	Confounding, bias, differing case definitions
Cohort study	Powerful, cheap (if retrospective)	Confounding, bias (especially attrition), differing case definitions
Historical control study	Powerful, cheap	Bias, differing case definitions, difficulty in interpretation, differing case definitions
Randomised and clinically controlled studies	Powerful, minimisation of all biases	Short follow up, limited power, differing case definitions

(Nach: Jefferson and Demicheli: Balancing benefits and harms in health care: observational data on harm are already included in systematic reviews. *BMJ* 327(7417):750, 2003)



Systematischer Review zu unerwünschten Effekten IV

Methoden: Bewertung der Anfälligkeit für systematisch verzerrte Ergebnisse

Beurteilung:

Methodik der Detektion von unerwünschten Effekten
(Definitionen, aktive vs. passive Detektionsverfahren?)

Berichtsqualität: Vollständigkeit (Patienten, Gruppierungen, Kategorien von unerwünschten Effekten, Daten)

Checklisten zur Bewertung der methodischen Qualität:

Chou R, Helfand M. Challenges in systematic reviews that assess treatment harms. *Ann Intern Med.* Jun 21;142(12 Pt 2):1090-9. 2005 (8 Items)

Santaguida PL, Raina P. The development of the McHarm quality assessment scale for adverse events: Delphi consensus on important criteria for evaluating harms. <http://hiru.mcmaster.ca/epc/mcharm.pdf>. (acc. 12.9.2010) (12 Items)

aber:



Hypothese: Studien mit eingeschränkter methodischer Qualität unterschätzen Nebenwirkungsraten

- Karotis-Endarteriektomie (CEA) bei symptomatischer Stenose
- 111 RCT, Kohortenstudien, Fallserien
- 8 methodische Qualitätskriterien (Selektionsbias vermieden, Patientenpopulation beschrieben, wenige Patientenverluste, unerwünschte Effekte vordefiniert, Erhebungstechnik beschrieben, unverzerrte Erhebung durchgeführt, statistische Kontrolle für Confounding, Adäquate Nachbeobachtungszeit)
- 4-Kriterieninstrument prädiziert Komplikationsrate (Schlaganfall, Tod)
- Ergebnisse nicht replizierbar für 18 Studien zu asymptomatischer Stenose und 16 RCT zu Rofecoxib bei Arthrose (MI-Rate)
 - im letzten Fall "unabhängige Erhebung" prädiktiv

(Chou et al.: Empirical Evaluation of the Association Between Methodological Shortcomings and Estimates of Adverse Effects, 2006)

Systematischer Review zu unerwünschten Effekten V

Methoden: Datenextraktion

Inhaltliche Probleme:

- Heterogene Begrifflichkeiten
- Grenzwerte?
- "No adverse effects reported....."
- Studienabbruch wegen unerwünschter Effekte als Surrogat?

Probleme der Berichtsqualität, z. B.:

- unspezifische "Feststellungen"
- zusammengefasste Daten
- nur häufige oder signifikant unterschiedliche unerwünschte Effekte berichtet

Table 3. Common Poor Reporting Practices for Harms-Related Data

1. Using generic or vague statements, such as "the drug was generally well tolerated" or "the comparator drug was relatively poorly tolerated."
2. Failing to provide separate data for each study arm.
3. Providing summed numbers for all adverse events for each study arm, without separate data for each type of adverse event.
4. Providing summed numbers for a specific type of adverse event, regardless of severity or seriousness.
5. Reporting only the adverse events observed at a certain frequency or rate threshold (for example, $>3\%$ or $>10\%$ of participants).
6. Reporting only the adverse events that reach a P value threshold in the comparison of the randomized arms (for example, $P < 0.05$).
7. Reporting measures of central tendency (for example, means or medians) for continuous variables without any information on extreme values.
8. Improperly handling or disregarding the relative timing of the events, when timing is an important determinant of the adverse event in question.
9. Not distinguishing between patients with 1 adverse event and participants with multiple adverse events.
10. Providing statements about whether data were statistically significant without giving the exact counts of events.
11. Not providing data on harms for all randomly assigned participants.



Systematischer Review zu unerwünschten Effekten Va

Methoden: Synthese

Synthese ≠ Meta-Analyse!

"In many circumstances adverse effect information is best summarised in a qualitative or descriptive manner. For instance, data derived from divergent sources cannot be combined because of different study design, different populations or, different data collection methods.

Analysis and presentation of results categorized by study design can potentially provide useful insights into a particular adverse effect...."



Beispiel:

Drug Class Review on Cyclo-oxygenase (COX)-2 Inhibitors and Non-steroidal Anti-inflammatory Drugs (NSAIDs)

Final Report Update 3

November 2006



The purpose of this report is to make available information regarding the comparative effectiveness and safety profiles of different drugs within pharmaceutical classes. Reports are not usage guidelines, nor should they be read as an endorsement of, or recommendation for, any particular drug, use or approach. Oregon Health & Science University does not recommend or endorse any guideline or recommendation developed by users of these reports.

Roger Chou, MD
Mark Helfand, MD, MPH
Kim Peterson, MS
Tracy Dana, MLS
Carol Roberts, BS

Produced by
Oregon Evidence-based Practice Center
Oregon Health & Science University
Mark Helfand, Director



TABLE OF CONTENTS

INTRODUCTION	4
Scope and Key Questions	5
METHODS	7
Literature Search	7
Study Selection	7
Data Abstraction	7
Validity Assessment.....	8
RESULTS	8
Overview	8
Key Question 1. Are there differences in effectiveness between coxibs and other NSAIDs?.....	9
Effectiveness	9
Efficacy	9
Celecoxib vs NSAIDs.....	9
NSAID vs NSAID	10
Key Questions 2 and 3. Are there clinically important differences in short-term safety or adverse effects between celecoxib, other NSAIDs, or the combination of a nonselective NSAID plus antiulcer medication? Are there clinically important differences in long-term safety or adverse effects between celecoxib, other NSAIDs, or the combination of a nonselective NSAID plus antiulcer medication?	10
Celecoxib	11
Partially selective NSAIDs	12
Meloxicam	12
Nabumetone and etodolac.....	12
Nonselective NSAIDs (with and without antiulcer medications).....	13
Salsalate	14
Tenoxicam and tiaprofenic acid.....	14
Key Question 4. Are there subgroups of patients based on demographics, other medications (e.g., aspirin), or co-morbidities for which one medication is more effective or associated with fewer adverse effects?.....	15
Demographic subgroups.....	15
Concomitant anticoagulant or aspirin use.	15
Concomitant anticoagulants.....	15
Concomitant aspirin.....	15
Co-morbidities	16
SUMMARY	16
REFERENCES	19

SUMMARY

Table 3. Strength of evidence by key question

Key Question	Level of Evidence	Conclusion
1. Are there differences in efficacy between celecoxib, partially selective NSAIDs, nonselective NSAIDs, the combination of a nonselective NSAID plus antiulcer medication or salsalate?		
Celecoxib	Good. Evidence is available from many published trials.	No clear differences in pain reduction
Meloxicam	Good. Consistent evidence from many published trials	No consistent differences
Nabumetone	Fair. Fewer RCTs/systematic review	No consistent differences
Etodolac	Good. Consistent evidence from many published trials	No consistent differences
Nonselectives	Good. Consistent evidence from many published trials and several good-quality systematic reviews	No consistent differences
Salsalate	Fair. Limited evidence from few RCTs	No consistent differences
Tenoxicam	Good. Many published RCTs, meta-analysis	No consistent differences
Tiaprofenic acid	Good. Several RCTs and one fair-quality review	No consistent differences

SUMMARY

Table 3. Strength of evidence by key question

Key Question	Level of Evidence	Conclusion
<p>2. Are there clinically important differences in short-term safety or adverse effects between celecoxib, partially selective NSAIDs, nonselective NSAIDs, the combination of a nonselective NSAID plus antiulcer medication or salsalate?</p> <p>3. Are there clinically important differences in long-term safety or adverse effects between celecoxib, partially selective NSAIDs, nonselective NSAIDs, the combination of a nonselective NSAID plus antiulcer medication or salsalate?</p>		
Celecoxib	Good. Evidence from many published trials and systematic reviews	<p><i>Short-term GI safety:</i> Fewer GI complications for celecoxib</p> <p><i>Long-term GI/CV safety:</i> Evidence suggests a higher CV risk (primarily MI) for celecoxib at doses of 200 or 400 mg twice daily, or 400 mg once daily; evidence is inconclusive for GI risk</p> <p><i>Other serious adverse events:</i> No consistent differences</p>
Meloxicam	Fair for GI safety; poor for others	<p><i>Short-term and long-term GI safety:</i> No consistent differences</p> <p><i>Long-term CV safety:</i> no conclusive evidence of increased risk relative to nonselectives</p> <p><i>Hepatotoxicity:</i> no evidence of increased risk relative to placebo</p> <p><i>Other serious adverse events:</i> no evidence</p>



Systematischer Review zu unerwünschten Effekten Vb

Methoden: Synthese

Zusammenfassung der Auswirkung von erwünschten und unerwünschten Effekten in "Utilities" – insbesondere für Kosten-Nutzen-Abschätzungen

3.2 Evidence for relative treatment effects

Introduction

3.2.1 The treatment effect of a technology can be summarised as the difference between the duration and state of health or HRQL (including the impact of any adverse effects of treatment) that would be experienced on average by patients receiving the technology and that experienced by the same group were they to receive alternative care.

NICE-Standard:

Messung der Lebensqualität: EQ-5D (Patienten)

Bewertung der Lebensqualität: nationale EQ-5D Referenzwerte –

Nutzwert zur Berechnung von QALYs, DALYs

NICE: Guide to the methods of technology appraisal, June 2008

(EUNetHTA, 2007)



Schlussfolgerungen

- Die Bewertung von "Nebenwirkungen" ist obligates Kernelement eines Health Technology Assessment
- Empfehlungen zur Durchführung von HTA stecken den inhaltlichen Rahmen der Bewertung von unerwünschten Effekten ab – die Entwicklung der präzisen Reviewfragestellung hängt vom Thema und dem Entscheidungskontext ab.
- Die Bewertung erfolgt zumeist im Rahmen einer systematischen Literaturübersicht – wobei deren Methoden noch nicht als "etabliert" gelten können.
- Forschungs- und Entwicklungsbedarf besteht insbesondere hinsichtlich
 - der Vereinheitlichung der Nomenklaturen und Klassifikationen
 - der Bestimmung der Validität und Biasanfälligkeit unterschiedlicher Datenquellen
 - der (Weiter)entwicklung von Methoden zur Informationssynthese



Ich bedanke mich für Ihr Interesse!

Kontakt:

Dr. med. Dagmar Lühmann

Institut für Sozialmedizin

Ratzeburger Allee 160

23538 Lübeck

Email: dagmar.luehmann@uk-sh.de

