



eunethta

EUROPEAN NETWORK FOR HEALTH TECHNOLOGY ASSESSMENT

GUIDELINE

Endpoints used in Relative Effectiveness Assessment

SAFETY

Adapted version (2015)

based on

“Endpoints used for REA of pharmaceuticals – Safety” February 2013

The primary objective of the EUnetHTA methodology guidelines is to focus on methodological challenges that are encountered by HTA assessors while performing a relative effectiveness assessment. Originally the focus of the methodological guidelines in Joint Action 1 was on the assessment of pharmaceuticals. During Joint Action 2 this document has been revised by WP7 to extend the scope of text and recommendations to non-drug interventions with a special focus on medical devices. This process led to minor language adaptations, but also to supplements of the original text.

This guideline "Endpoints used for REA of pharmaceuticals - Safety" has been elaborated during Joint Action 1 by experts from AIFA, reviewed and validated by HAS and all members of WP5 of the EUnetHTA network; the whole process was coordinated by HAS.

The adaptation of the original guideline within JA2 has been made by experts from IQWiG, reviewed by KCE and AIFA, and the WP7 partners of the EUnetHTA network.

The guideline represents a consolidated view of non-binding recommendations of EUnetHTA network members and is in no case an official opinion of the participating institutions or individuals.

Disclaimer: EUnetHTA Joint Action 2 is supported by a grant from the European Commission. The sole responsibility for the content of this document lies with the authors and neither the European Commission nor EUnetHTA are responsible for any use that may be made of the information contained therein.

Table of contents

Acronyms – Abbreviations	4
Summary and recommendations	5
Summary	5
Recommendations	6
1. Introduction	9
1.1. Definitions	9
1.2. Context and problem statement	14
1.3. Scope/Objective(s) of the guideline	15
1.4. Related EUnetHTA documents	15
2. Analysis and synthesis of literature	16
2.1. Objectives of the HTA assessors	16
2.2. Terminology	16
2.3. Identification of adverse reactions: sources of information	17
2.3.1. Regulatory sources	17
2.3.1.1. Regulatory sources for pharmaceuticals	18
2.3.1.2. Regulatory sources for medical devices	20
2.3.2. Randomised Clinical Trials (RCTs)	21
2.3.3. Observational studies	22
2.3.4. Case Reports	22
2.3.5. How and where to find the information	22
2.4. Evaluation of sources of information	23
2.4.1. Regulatory sources	23
2.4.2. Randomised clinical trials (RCTs)	23
2.4.3. Observational studies	25
2.4.4. Case reports	25
2.5. Synthesis and reporting of results by HTA assessors	25
2.5.1. Description of included sources of information	26
2.5.2. Quantification of adverse effects in terms of frequency, incidence, severity and seriousness	27
2.5.3. Comparison of the safety profile of the technology to the comparator(s)	30
2.5.4. Balanced discussion of benefits and adverse effects	31
3. Discussion and Conclusions	32
Annexe 1. The STROBE statement	33
Annexe 2. Preferred contents of a case report	35
Annexe 3. Methods of documentation and selection criteria (applied during original guideline elaboration)	36
Annexe 4. Bibliography	39

Acronyms – Abbreviations

ADE: Adverse Device Effect

ADL: Activities of Daily Living

ADR: Adverse drug reaction

AE: Adverse Event

AusPAR: Australian Public Assessment Report

CER: Comparative effectiveness review

CHMP: Committee for medicinal product for Human Use

CMDh: Coordination Group for Mutual Recognition and Decentralized Procedures

EMA: European Medicines Agency

EPAR: European Public Assessment Report

EUDAMED: European Databank on Medical Devices

FDA: Food and Drug Administration

MAUDE-Database: Manufacturer and User Facility Device Experience Database

MD: Medical Device

MedDRA: Medical Dictionary for Regulatory Activities

MSAC: Medical services Advisory Committee

PASS: Post-authorisation safety studies

PMA: Premarket Approval

PRAC: Pharmacovigilance Risk Assessment Committee

PSUR: Periodic Safety Update Report

PT: Preferred Term

REA: Relative effectiveness Assessment

RCT: Randomised clinical trial

REMS: Risk evaluation and mitigation strategy

RMP: Risk Management Plan

SOC: System Organ Class

SPC: Summary of Product Characteristics

TGA: Therapeutic Goods Administration

Summary and recommendations

Summary

This guideline aims at providing a framework for the evaluation of relative safety performed by HTA assessors in the context of Relative Effectiveness Assessment (REA) .

When performing relative safety assessment the safety profile of e.g. a pharmaceutical is assessed in comparison to the comparator(s) of the same or different therapeutic class and to the safety profile of non-pharmaceutical alternatives (when available). In general, this is also true for non-drug interventions. For example the safety profile of a medical device (MD) can be assessed in comparison to the comparator(s) of the same MD category (e.g. drug-eluting stent A vs. drug-eluting stent B) or a different MD-category (e.g. drug-eluting stent vs. bare-metal stent) and to the safety profile of a pharmaceutical or another intervention (when available).

It is important to carry out balanced assessments of the interventions, taking into account both beneficial and adverse effects, in order to support clinicians, policy makers and patients in making informed decisions. For this reason beneficial and adverse reactions/effects should be assessed with similar methodological rigour and accuracy. Although the importance of a balanced assessment is well recognised, the assessment of adverse reactions/effects is still more troublesome than the assessment of benefits. In this guideline some important methodological issues concerning relative safety assessment have been addressed and recommendations were given.

Consistent and precise terminology should be used and for this purpose the Medical Dictionary for Regulatory Activities (MedDRA) should be used for describing adverse reactions/effects.

When conducting a rapid assessment, just after the marketing authorisation of a medicine, European Public Assessment Report (EPAR), Summary of Product Characteristics (SPC), Risk Management Plan (RMP) (when available), manufacturer's dossier and published and unpublished randomised clinical trials are generally used as primary sources of information. In the case of non-drug interventions primarily based on medical devices these are publicly available approval orders and safety reports from the FDA, manufacturer dossier and randomised clinical trials reports and/or publications (for each: when available). Documents elaborated for granting marketing authorisation (declaration of conformity) of medical devices in the EU are not publicly accessible and vigilance data are not systematically publicly accessible as well. It is important to assess, apart from risk of bias of studies, also the quality of data on adverse effects, taking into account how adverse effects data are collected and reported to decide on their inclusion and interpretation. Main characteristics of selected studies and their limitations should be described and reported preferably in tabular form.

Results from individual studies should be reported both for the technology and comparator(s), using summary tables for the different study designs.

Finally, the assessment of relative safety should be performed between the technology and its comparator(s) with special regard to the most frequent, serious and severe adverse reactions. This assessment together with the assessment of comparative benefits will contribute to establish a balanced assessment of the relative effectiveness of the health technology and to decide upon the possible consequences on coverage decision.

In the discussion of results limitations and external validity of results should be investigated and discussed, considering all factors (e.g. patient characteristics, co-morbidities, type and severity of disease) which may contribute to the occurrence of adverse reactions.

Recommendations

Recommendation	The recommendation is based on the following elements
<p>Recommendation 1</p> <p>In relative safety assessment of health technologies main objectives of HTA assessors are summarised as follows:</p> <ul style="list-style-type: none"> - To identify the adverse reactions - To quantify the adverse reactions in terms of frequency categories, incidence, severity and seriousness - To compare the safety profile of the health technology with its comparator(s)/best standard of care. 	2.1. Objectives of the HTA assessors
<p>Recommendation 2</p> <p>HTA assessors may focus their investigation on the following areas:</p> <ul style="list-style-type: none"> - The most serious adverse reactions. - The most frequent adverse reactions. - Other specific adverse reactions important to clinicians or patients. 	2.1 Objectives of the HTA assessors
<p>Recommendation 3</p> <p>The HTA assessors should use consistent and precise terminology to avoid misleading results. They should use the MedDRA Dictionary for describing adverse reactions.</p>	2.2 Terminology
<p>Recommendation 4</p> <p>Main sources of information of HTA assessors are:</p> <ul style="list-style-type: none"> - For pharmaceuticals: EPAR, SPC and RMP (when available) - For non-drug interventions: publicly available approval orders and safety reports from the FDA - Published and unpublished (where acceptable under the specific HTA system guidelines) randomised clinical trials - Manufacturer dossier - Unpublished full study reports (where acceptable under the specific HTA system guidelines) - Observational studies 	2.3 Identification of adverse reactions: sources of information

<p>Recommendation 5</p> <p>It is necessary to evaluate both the risk of bias of sources of information and the quality of data on adverse reactions. Methods used to assess the risk of bias should be clearly described and results should be reported. It should be clearly explained how the information on risk of bias will be used in the synthesis of data.</p> <p>To assess the data on adverse reactions, how the adverse effects were collected and reported should be evaluated.</p> <p>Useful questions to assess how the adverse reactions are collected:</p> <ul style="list-style-type: none"> - Were definitions given of reported adverse effects? - How were adverse effects data collected: prospective/routine monitoring, spontaneous reporting, patient checklist/ questionnaire/diary; systematic survey of patients? <p>Useful questions to assess how the adverse effects are reported:</p> <ul style="list-style-type: none"> - Were any patients excluded from the adverse effects analysis? - Did the report give numerical data by intervention group? - Which categories of adverse effects did the investigators report? - Did investigators report on all important or serious adverse effects, and how were these defined? - Were methods used for monitoring adverse effects reported? - Was an independent data safety monitoring board established? 	<p>2.4 Evaluation of sources of information</p>
<p>Recommendation 6</p> <p>Characteristics of selected studies should be summarised in a table. Useful information on studies characteristics are the following:</p> <ul style="list-style-type: none"> - methods (study design, follow-up period); - participants for both arms (e.g. setting, age, sex and country/geographic area), - intervention and comparators (e.g. for pharmaceuticals: the name, dose, frequency, way of administration and duration); - outcomes; - methods to collect adverse effects. <p>Different tables should be elaborated for RCTs and observational studies.</p>	<p>2.5.1 Description of included sources of information</p>

<p>Recommendation 7</p> <p>Results from individual studies should be presented by group in tabular form, using the following measures:</p> <ul style="list-style-type: none"> - Number of participants in both study arms - Number of patients excluded from the analysis dataset - Patient-years of exposure - Number of participants with the event - Number of events - Absolute risk; incidence rate (95% CI) - Relative risk (95% CI) - Quality of evidence <p>Different tables should be elaborated for RCTs and observational studies.</p> <p>Adverse effects should be grouped according to the System Organ Class (SOC).</p> <p>Adverse effects which are common and serious should be reported separately.</p> <p>If possible, adverse effects should also be provided by severity grade.</p> <p>When adverse effects are collected from different study designs and when the degree of heterogeneity is high the data cannot all be pooled together using standard meta-analysis principles. Therefore in these circumstances adverse effects data is best summarised in a qualitative or descriptive manner.</p>	<p>2.5.2 Quantification of adverse effects in terms of frequency, incidence, severity and seriousness</p>
<p>Recommendation 8</p> <p>The safety profile of the health technology is described in comparison to the comparator(s), with special regard to the most frequent, serious and severe adverse reactions.</p> <p>A table is preferable for the comparison of the safety profile of the new health technology and the comparator(s).</p> <p>HTA assessors should describe if there is a clinically significant difference in adverse reactions between the interventions compared.</p> <p>In the discussion of results, limitations and external validity should be investigated and discussed, considering all factors (e.g. patient characteristics, co-morbidities, type and severity of disease) which may contribute to the occurrence of adverse reactions.</p>	<p>2.5.3 The comparison of the safety profile of the health technology to the comparator(s)</p>
<p>Recommendation 9</p> <p>The assessment of relative safety together with relative benefits will contribute to establish a balanced assessment of the relative effectiveness of the intervention, and to decide upon possible consequences on coverage decision.</p>	<p>2.5.4 Balanced discussion of benefits and adverse effects</p>

1. Introduction

1.1. Definitions

The following list illustrates the range of definitions used in the context of regulation of pharmaceuticals and MD, respectively and the context of HTA with regard to safety.

Adverse effect and adverse reaction

The two terms refer to the same phenomenon, but an adverse effect is seen from the point of view of the pharmaceutical, whereas an adverse reaction is seen from the point of view of the patient. The pharmaceutical causes an effect, whereas the patient has a reaction.

Source: Aronson JK, Ferner RE. Clarification of Terminology in drug Safety. Drug Safety 2005; 28 (10): 851-870.

A general definition of **Adverse effect** to be found is: A harmful or undesirable outcome that occurs during or after the use of a drug or intervention for which there is at least a reasonable possibility of a causal relation.

Source: Chou R., Aronson N, Atkins D, Ismaila AS, Santaguida P, Smith DH, Whitlock E, Wilt TJ, Moher D. AHRQ paper 4: assessing harms when comparing medical interventions: AHRQ and the Effective Health Care Program. Journal of clinical epidemiology 2010; 63: 502-512.

A general definition of **Adverse reaction** to be found is: Events for which a causality link to the tested intervention is well established and strong enough (sensitive and specific) to warrant attribution of the event to the intervention. Attribution of causality in the setting of clinical trials may be misleading.

Source: Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, Moher D. Better reporting of harms in randomized trials: an extension of the CONSORT statement. Ann Intern Med 2004; 141(10):781-788.

Adverse Device Effect (ADE)

Adverse event related to the use of an investigational medical device.

This includes any adverse event resulting from insufficiencies or inadequacies in the instructions for use, the deployment, the implantation, the installation, the operation, or any malfunction of the investigational medical device as well as any event that is a result of a use error or intentional misuse.

Source: Guidelines on medical devices. Clinical investigations: serious adverse event reporting. MEDDEV 2.7/3 2010.

Adverse event

Any untoward medical occurrence in a patient administered a medicinal product and which does not necessarily have to have a causal relationship with this treatment.

An adverse event (AE) can therefore be any unfavourable and unintended sign for example, including an abnormal laboratory finding, symptom, or disease temporally associated with the use of a medicinal product, whether or not considered related to the medicinal product.

Source: ICH Harmonised Tripartite Guideline. Post approval safety data management: definitions and standards for expedited reporting (ICH E2D).

In the context of MD-regulation: Any untoward medical occurrence, unintended disease or injury or any untoward clinical signs (including an abnormal laboratory finding) in subjects, users or other persons whether or not related to the **investigational medical device**.

This includes events related to the investigational device or the comparator as well as events related to the procedures involved (any procedure in the clinical investigation plan). For users or other persons this is restricted to events related to the investigational medical device.

Source: Guidelines on medical devices. Clinical investigations: serious adverse event reporting. MEDDEV 2.7/3 2010.

A general definition of **Adverse event** to be found is: A harmful or undesirable outcome that occurs during or after the use of a drug or intervention but is not necessarily caused by it. When causality is uncertain or the purpose of the relative effectiveness assessment is to establish causality, "adverse event" should generally be the default term over "adverse effect" or "adverse reaction / adverse drug reaction".

Source: Chou R., Aronson N, Atkins D, Ismaila AS, Santaguida P, Smith DH, Whitlock E, Wilt TJ, Moher D. AHRQ paper 4: assessing harms when comparing medical interventions: AHRQ and the Effective Health Care Program. Journal of clinical epidemiology 2010; 63: 502-512.

Adverse reaction/adverse drug reaction (ADR)

Noxious and unintended effects resulting not only from the authorised use of a medicinal product at normal doses, but also from medication errors and uses outside the terms of the marketing authorisation, including the misuse and abuse of the medicinal product. The suspicion of an adverse drug reaction, meaning that there is at least a reasonable possibility of there being a causal relationship between a medicinal product and an adverse event, should be sufficient reason for reporting.

Source: Directive 2010/84/EU of the European Parliament and of the council of 15 December 2010.

Adverse reaction (Serious)

An adverse reaction which results in death, is life-threatening, requires in-patient hospitalisation or prolongation of existing hospitalisation, results in persistent or significant disability or incapacity, or in a congenital anomaly/birth defect, and is a medically important event or reaction.

For the terms "serious" and "severe," which are not synonymous, the following note of clarification is provided: The term "severe" is often used to describe the intensity (severity) of a specific event (as in mild, moderate, or severe myocardial infarction); the event itself, however, may be of relatively minor medical significance (such as severe headache). This is not the same as "serious," which is based on patient/event outcome or action criteria usually associated with events that pose a threat to a patient's life or functioning. Seriousness (not severity) serves as a guide for defining regulatory reporting obligations.

Sources: ICH Harmonised Tripartite Guideline. Clinical Safety Data Management: Definitions and Standard for Expedited Reporting E2A, (Article 1(12) of Directive 2001/83/EC). / ICH Harmonised Tripartite Guideline. Post approval safety data management: definitions and standards for expedited reporting (ICH E2D).

Adverse reaction (Severity Grade)

Grade 1

Mild; asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated.

Grade 2

Moderate; minimal, local or non-invasive intervention indicated; limiting age-appropriate instrumental Activities of Daily Living (ADL).

Grade 3

Severe or medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self care Activities of Daily Living.

Grade 4

Life-threatening consequences; urgent intervention indicated.

Grade 5

Death related to AE.

Source: Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0. Available and accessed in May 2011: http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf

Adverse reaction (Unexpected)

An adverse reaction, the nature or severity of which is not consistent with the applicable product information.

Source: ICH Harmonised Tripartite Guideline. Clinical Safety Data Management: Definitions and Standard for Expedited Reporting E2A.

An ADR whose nature, severity, specificity, or outcome is not consistent with the term or description used in the local/regional product labelling (e.g. Package Insert or Summary of Product Characteristics) should be considered unexpected. When a Marketing Authorisation Holder is uncertain whether an ADR is expected or unexpected, the ADR should be treated as unexpected.

Source: ICH Harmonised Tripartite Guideline. Post approval safety data management: definitions and standards for expedited reporting (ICH E2D).

Benefit - Risk Balance

In the regulatory context: an evaluation of the positive therapeutic effects of the medicinal product in relation to its risks (any risk relating to the quality, safety or efficacy of the medicinal product as regards patients' health or public health and any risk of undesirable effects on the environment).

Source: Directive 2001/83/EC of the European Parliament and of the Council of 6 November 2001 on the Community code relating to medicinal products for human use. (PHIS Glossary).

In the context of relative effectiveness assessment also see **Risk-benefit ratio**.

The most common expression for the comparison of harms and benefits. It is a technical term that assumes that a ratio can indeed be calculated. Because the benefits and harms of an intervention are often so different in character or are measured on different scales, the term "risk-benefit ratio" has no literal meaning. In addition, there may be several distinct benefits and harms. It is advocated to use "balance of benefits and harms" rather than "risk-benefit ratio".

Source: Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, Moher D. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004; 141(10):781-788.

Case by Case Causality assessment

The evaluation of the likelihood that a medicine was the causative agent of an observed adverse reaction. Causality assessment is usually made according to established algorithms.

Source: WHO. Glossary of terms used in Pharmacovigilance. Available and accessed on October 15 2010: <http://www.who-umc.org/graphics/15338.pdf>
<http://www.who-umc.org/Graphics/24729.pdf>

Classification of causality

- **Certain**
 - A Clinical event, including a laboratory test abnormality, that occurs in a plausible time relation to drug administration and which cannot be explained by concurrent disease or other drugs or chemicals
 - The response to withdrawal of the drug (dechallenge) should be clinically plausible
 - The event must be definitive pharmacologically or phenomenologically, using a satisfactory rechallenge procedure if necessary
- **Probable/likely**
 - A clinical event, including a laboratory test abnormality, with a reasonable time relation to administration of the drug, unlikely to be attributed to concurrent disease or other drugs or chemicals, and which follows a clinically reasonable response on withdrawal (dechallenge)
 - Rechallenge information is not required to fulfil this definition

- **Possible**
 - A clinical event, including a laboratory test abnormality, with a reasonable time relation to administration of the drug, but which could also be explained by concurrent disease or other drugs or chemicals
 - Information on drug withdrawal may be lacking or unclear
- **Unlikely**
 - A clinical event, including a laboratory test abnormality, with a temporal relation to administration of the drug, which makes a causal relation improbable, and in which other drugs, chemicals, or underlying disease provide plausible explanations
- **Conditional/unclassified**
 - A clinical event, including a laboratory test abnormality, reported as an adverse reaction, about which more data are essential for a proper assessment or the additional data are being examined
- **Unassessable/unclassifiable**
 - A report suggesting an adverse reaction that cannot be judged, because information is insufficient or contradictory and cannot be supplemented or verified

Source: Edwards IR, Aronson JK. Adverse drug reactions: definitions, diagnosis, and management. *Lancet*. 2000;356:1255-9. WHO-UMC Causality Categories

Causal relationship

A relationship between one phenomenon or event (A) and another (B) in which A precedes and causes B. Source: WHO. Glossary of terms used in Pharmacovigilance.

Harms

The totality of possible adverse consequences of an intervention or therapy; they are the direct opposite of benefits, against which they must be compared.

Source: Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, Moher D. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004; 141(10):781-788.

The nature and extent of actual damage that could be caused by a drug.

Source: WHO. Glossary of terms used in Pharmacovigilance.

In the context of MD regulation: Physical injury or damage to the health of people, or damage to property or the environment.

Source: Guidelines on medical devices. Vigilance System. MEDDEV 2.12-1 rev. 8 2013.

Incident

In the context of MD regulation: "Any malfunction or deterioration in the characteristics and/or performance of a device, as well as any inadequacy in the labelling or the instructions for use which, directly or indirectly, might lead to or might have led to the death of a patient, or user or of other persons or to a serious deterioration in their state of health." (Article 10 of the MDD).

There is a similar definition in Article 8 of the AIMDD and Article 11 IVD Directive with minor wording differences.

Source: Guidelines on medical devices. Vigilance System. MEDDEV 2.12-1 rev. 8 2013.

Pharmacovigilance

The science and activities relating to the detection, assessment, understanding and prevention of adverse effects or any other drug related problem.

Source: WHO. Glossary of terms used in Pharmacovigilance. Available and accessed on October 15 2010: <http://www.who-umc.org/graphics/15338.pdf>

Risk

The probability that an event will occur, e.g., that an individual will become ill or die within a stated period of time or by a certain age.

Also a nontechnical term encompassing a variety of measures of the probability of a (generally) unfavourable outcome.

Source: Last. A dictionary of epidemiology edited for the International Epidemiological Association. (PHIS Glossary).

Serious Adverse Device Effect (SADE)

Adverse device effect that has resulted in any of the consequences characteristic of a serious adverse event.

Source: Guidelines on medical devices. Clinical investigations: serious adverse event reporting. MEDDEV 2.7/3 2010.

Serious Adverse Event (SAE)

In the context of MD regulation: Adverse event that:

- Led to a death
- Led to a serious deterioration in health that either:
 - Resulted in a life-threatening illness or injury, or
 - Resulted in a permanent impairment of a body structure or a body function, or
 - Required in-patient hospitalization or prolongation of existing hospitalization, or
 - Resulted in medical or surgical intervention to prevent life threatening illness or injury or permanent impairment to a body structure or a body function.
- Led to fetal distress, fetal death or a congenital abnormality or birth defect

This includes device deficiencies that might have led to a SAE if suitable action had not been taken or intervention had not been made or if circumstances had been less fortunate. These are handled under the SAE reporting system.

A planned hospitalization for pre-existing condition, or a procedure required by the clinical investigation plan, without a serious deterioration in health, is not considered to be a serious adverse event.

Source: Guidelines on medical devices. Clinical investigations: serious adverse event reporting. MEDDEV 2.7/3 2010.

A general definition for **Serious adverse event** to be found is: Any adverse event with serious medical consequences, including death, hospital admission, prolonged hospitalization, and persistent or significant disability or incapacity.

Source: Chou R., Aronson N, Atkins D, Ismaila AS, Santaguida P, Smith DH, Whitlock E, Wilt TJ, Moher D. AHRQ paper 4: assessing harms when comparing medical interventions: AHRQ and the Effective Health Care Program. Journal of clinical epidemiology 2010; 63: 502-512.

Safety

Substantive evidence of an absence of harm. The term is often misused when there is simply absence of evidence of harm.

Source: Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, Moher D. Better reporting of harms in randomized trials: an extension of the CONSORT statement. Ann Intern Med 2004; 141(10):781-788.

Side effect

Unintended drug effects. The term, however, does not necessarily imply harm, as some side effects may be beneficial. Furthermore, it tends to understate the importance of harms because "side" may be perceived as denoting secondary importance.

Source: Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, Moher D. Better reporting of harms in randomized trials: an extension of the CONSORT statement. Ann Intern Med 2004; 141(10):781-788.

It is recommended that this term no longer be used and particularly should not be regarded as synonymous with adverse event or adverse reaction.

Source: ICH Harmonised Tripartite Guideline. Clinical Safety Data Management: Definitions and Standard for Expedited Reporting E2A.

Tolerability

A term that usually refers to medically less important (i.e. without serious or permanent sequelae) but unpleasant adverse effects of drugs. These include symptoms such as dry mouth, tiredness, etc, that can affect a person's quality of life and willingness to continue the treatment. As these adverse effects usually develop early and are relatively frequent, RCTs may yield reliable data on their incidence.

Source: Loke YK, Price D, Herxheimer A, Cochrane Adverse Effects Methods Group. Systematic reviews on adverse effects. Framework for a structured approach. BMC Med Res Method 2007; 7:32; 7:32.

Toxicity

Describes drug-related harms. The term may be most appropriate for laboratory-determined measurements, although it is also used in relation to clinical events. Abnormal laboratory values may be described as laboratory-determined toxicity. The disadvantage of the term "toxicity" is that it implies causality. If authors cannot prove causality, the terms "abnormal laboratory measurements" or "laboratory abnormalities" are more appropriate to use.

Source: Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, Moher D. Better reporting of harms in randomized trials: an extension of the CONSORT statement. Ann Intern Med 2004; 141(10):781-788.

As long as plausible, these considerations are also true for non-drug interventions primarily based on medical devices.

The term "toxicity" is used in pharmacology and microbiology to mean "the quality of being poisonous, especially the degree of virulence of a toxic microbe or of a poison." It is often measured in terms of the specific target affected (e.g. cytotoxicity, hepatotoxicity, and so on). In the context of relative effectiveness assessment, the term is often used to refer to laboratory-determined abnormalities, such as elevated liver-function tests. However, the terms "abnormal laboratory measurements" or "laboratory abnormalities" are more specific and appropriate.

Source: Chou R., Aronson N, Atkins D, Ismaila AS, Santaguida P, Smith DH, Whitlock E, Wilt TJ, Moher D. AHRQ paper 4: assessing harms when comparing medical interventions: AHRQ and the Effective Health Care Program. Journal of clinical epidemiology 2010; 63: 502-512.

1.2. Context and problem statement

The importance of assessing both benefits and adverse reactions with similar rigour in order to provide balanced assessments of alternative interventions is well recognised.^{1,2} Nevertheless, assessing adverse effects may be difficult because of the greater prominence given to the beneficial effects of interventions and the ongoing methodological issues with assessment of adverse effects.^{3,4,5,6,7}

In the context of safety two different evaluations are performed, on the one hand in the context of regulation and on the other hand in the context of HTA:

- In the context of pharmaceuticals: The benefit-risk assessment carried out by regulatory authorities, during the pre-approval phase to grant marketing authorisation and continuously during the post-approval phase taking into account new risks or changes in known risks.
- In the context of medical devices (in the EU): The evaluation of safety and performance carried out during conformity assessment to grant marketing authorisation (declaration of conformity) and continuously during the post-production phase in which manufacturers are required to implement a post-market surveillance procedure or program, which includes an obligation to report serious incidents to the relevant Competent Authorities.

- The relative safety assessment of a health technology, conducted by HTA assessors. In this case the safety profile of e.g. a pharmaceutical is compared to the safety profile of the comparator(s) belonging to the same or different therapeutic class and to the safety profile of non pharmaceutical alternatives (when available). Or for example the safety profile of a medical device (MD) can be assessed in comparison to the comparator(s) of the same MD category (e.g. drug-eluting stent A vs. drug-eluting stent B) or a different MD-category (e.g. drug-eluting stent vs. bare-metal stent) or to the safety profile of a pharmaceutical or another intervention (when available). In some cases it could be also necessary to consider how both the intervention and the comparator are administered and if different procedures may contribute to the occurrence of adverse reactions.

1.3. Scope/Objective(s) of the guideline

This guideline focuses on the relative safety assessment performed by the HTA assessors when conducting Relative Effectiveness Assessment (REA) and deals with the following methodological issues:

- objectives of HTA assessors
- terminology
- identification of adverse reactions: sources of information
- evaluation of sources of information
- synthesis and reporting of results compared to other interventions

These issues are addressed and discussed in the main chapters of the document.

Note: The diversity of types of health technologies means that there are many different types of safety issues. This guideline was originally developed in the context of the evaluation of pharmaceuticals. The scope and recommendations are now extended to non-drug interventions with a primary focus on medical devices. Specialities with regard to non-drug intervention like e.g. screening or diagnostics are out of the scope of this guideline. But many recommendations will also apply to these interventions.

1.4. Related EUnetHTA documents

This guideline should be read in conjunction with the following documents (www.eunethta.eu/eunethta-guidelines)

- o EUnetHTA guideline on Endpoints used in REA: Composite endpoints
- o EUnetHTA guideline on levels of evidence: applicability of evidence in the context of relative effectiveness assessment
- o EUnetHTA guideline on levels of evidence: internal validity (of randomized controlled trials)
- o EUnetHTA guideline on comparators and comparisons: direct and indirect comparisons
- o EUnetHTA guideline on therapeutic medical devices (JA2)
- o EUnetHTA Guideline on internal validity of non-randomised studies (NRS) on interventions (JA2)
- o EUnetHTA Guideline on process of information retrieval for systematic reviews and health technology assessment on clinical effectiveness (JA2)
- o EUnetHTA manufacturers' submission templates to support production of core HTA information and rapid assessments (JA2)

2. Analysis and synthesis of literature

2.1. Objectives of the HTA assessors

In relative safety assessment the main objectives of HTA assessors should be the following:

- To identify the adverse effects
- To quantify the adverse effects in terms of frequency, incidence, severity and seriousness
- To compare the safety profile of the health technology with its comparator(s)

Finally, in a context of Relative Effectiveness Assessment, HTA assessors will discuss in a balanced way the adverse effects and benefits related to the technology in comparison with alternatives.

Within the REA, the relative safety assessment can contribute to coverage decisions, because, on the basis of the safety profile of a technology in comparison with its alternatives, payers can decide:

- to limit the coverage to specific population subgroups and to specific therapeutic indications
- to partially reimburse or to not reimburse the technology
- to conditionally reimburse the technology and request that further information (e.g. on safety and patient-relevant outcomes) is gathered.

In order to carry out analyses in a systematic, manageable and useful way HTA assessors may focus their investigation on the following areas:

- the most serious adverse reactions
- the most frequent adverse reactions and
- other specific adverse reactions important to clinicians and patients⁸.

2.2. Terminology

It is important that the HTA assessors use consistent and precise terminology to avoid confusion and misleading conclusions.²

For this purpose the Medical Dictionary for Regulatory Activities (MedDRA), developed by the International Conference on Harmonisation, could be a useful instrument.⁹ MedDRA includes medical signs, symptoms, syndromes and diagnoses as well as social conditions, surgical and medical procedures and laboratory and clinical investigations. It comprises five levels: lowest level terms (LLTs); preferred terms (PTs); high level terms (HLT); high level group terms (HLGTs) and it is organised in 26 system organ classes (SOCs).¹⁰ It is important to note that due to the multiaxial structure of MedDRA, it may be necessary to combine several PTs in order to represent and to analyse one overlying medical concept (e.g. bleeding). In most cases it is not sufficient to rely upon one single PT. Standardised MedDRA Queries (SMQs), which are provided with the regular terminology updates, are preferred, especially when analysing medical conditions involving PTs across several SOCs.

MedDRA does not include a severity ranking. While the use of MedDRA is mandatory in the European Union for recording and reporting adverse effects/reactions data on marketed medicines, this is not the case for medical devices in the context of market access and vigilance systems in Europe. However, authors of HTA should always strive for consistent and precise usage of terminology when reporting data on harms. MedDRA is available free of charge to regulatory authorities and to certain non-profit-making organisations and on payment of an annual subscription to other users.¹¹

Since MedDRA contains neither severity descriptors nor descriptors of seriousness or intensity it may be useful to consider these characteristics of adverse effects in line with ICH E2D. The

definitions given in this guideline (chapter 1.1) provide an adequate and comprehensive foundation for the analysis of adverse effects regarding characteristics not covered by MedDRA.

Besides MedDRA-coded adverse events it may be useful to also analyse endpoints measuring harms not based on this terminology. These concern explicitly pre-planned endpoints which are recorded according to pre-planned definitions (e.g. different evaluations of bleeding events: major bleeding, minor bleeding etc). For observational studies in claims databases and electronic health records, other terminology may be used for AE assessment (ICD-9, ICD-10, READ).

2.3. Identification of adverse reactions: sources of information

A broad range of evidence sources may be considered to identify adverse effects relevant for the assessment. These sources may include regulatory sources (in the case of pharmaceuticals e.g. EPAR, SPC and RMP; in the case of medical devices e.g. publicly available approval orders and safety reports from the US Food and Drug Administration (FDA)), manufacturer dossier, randomised clinical trials, observational studies, country registries and case reports. Various sources can bring different and complementary information; randomised clinical trials may inform on common risks, whereas other data sources, although at higher risk of bias, (e.g. observational studies, country registries and case reports) can give insight on less frequent risks, long-term risks, and risks in populations not being part of randomised clinical trials. Singh et al used RCTs and observational studies to assess the risk of heart failure associated to thiazolidinediones and collected complementary information from case reports on specific characteristics of adverse effects such as dose, time and susceptibility factors.⁴⁹

In practice, for the identification of adverse effects in the first appraisal, the most important sources of data that are used by HTA assessors are:

- for pharmaceuticals: the EPAR, SPC, RMP (when available), manufacturer dossier and randomised clinical trials reports and/or publications.
- for non-drug interventions primarily based on medical devices: publicly available approval orders and safety reports from the FDA, manufacturer dossier, horizon scanning centres like NIHR Horizon Scanning Research & Intelligence Centre (UK), and randomised clinical trials reports and/or publications (for each: when available).

When possible, adverse effects relevant for the assessment should be identified in advance and should be listed in the protocol of the HTA report.

2.3.1. Regulatory sources

The regulatory authorities' documentation (mainly EPAR and SPC) elaborated during approval phase of pharmaceuticals is an important source of information and the most commonly available, especially when conducting a rapid assessment. Moreover regulatory agencies provide important data in the post approval phase as well. Documents elaborated for granting marketing authorisation (declaration of conformity) of medical devices in the EU are not publicly accessible and vigilance data are not systematically publicly accessible as well. However, if a medical device has already been assessed by the FDA, informative regulatory documents are freely accessible. There are also single cases of published FDA information on devices or manufacturers even if finally a device could not get an approval for the US market.

Data from different regulatory authorities, when available, should be used and compared for two main reasons:

- The safety assessments from different regulatory authorities may vary
- The technology may have different regulatory status across different jurisdictions: it could be in the approval phase in a jurisdiction and may have been marketed in a different jurisdiction.

2.3.1.1. Regulatory sources for pharmaceuticals

Data at approval phase

At approval phase regulatory authorities synthesise the available data in specific documents.

At European level (EMA):

- The European Public Assessment Report (EPAR) and the Summary of Product Characteristics (SPC) provide a useful summary of adverse effects of medicines¹² and may be useful to obtain data in order to evaluate adverse effects and to compare them between different products. They represent the main evidence for the HTA assessors.
- The Risk Management Plan (RMP), which is presented by applicants and/or marketing authorisation holders to describe the pharmacovigilance and risk minimisation activities summarises the important identified and potential risks of a medicinal product and important missing information on unidentified risks. A summary of the RMP will be publicly available according to the Regulation (EU) No 1235/2010 Directive 2010/84/EU (. Assessors should also determine whether risk minimisation activities (RMM) are required.
- HTA Assessors should verify if post-authorisation safety studies (PASS) were required. PASS can be required by regulatory authorities either as a commitment at the time of authorisation or in the post-authorisation phase to further assess a signal.¹³

At US level (FDA):

- Medical review
- Summary review
- Risk evaluation and mitigation strategy (REMS)
- Statistical review

At Australian level (TGA)

- Australian Public Assessment Report (AusPAR)

At Canadian level (Health Canada)

- Product Monograph

Data at post approval phase

After the marketing authorisation the safety profile is continuously monitored by pharmacovigilance systems of regulatory agencies. Spontaneous reports of suspected adverse drug reactions provide important early signals of safety concerns and include:

At European level:

Eudravigilance database, which collects reports received from EU regulatory agencies and from pharmaceutical companies. Data from EudraVigilance are published in the European database of suspected adverse drug reaction reports.

The periodic safety update report for marketed pharmaceuticals (PSUR) provides a critical evaluation of the benefit-risk balance of a medicinal product, in consideration of new or changing post-authorisation information and analyses all adverse reactions reported in the period since the last PSUR.¹⁴ According to the new legislation the following documents shall be made publicly available by means of the European medicines web-portal:

- List of EU reference dates and frequency of submission of PSURs, final assessment conclusions of the adopted assessment reports;
- Pharmacovigilance Risk Assessment Committee (PRAC) recommendations including relevant annexes;
- Coordination Group for Mutual Recognition and Decentralized Procedures (CMDh) position including relevant annexes and where applicable, detailed explanation on scientific grounds for any differences with the PRAC recommendations;

- Committee for medicinal products for Human Use (CHMP) opinion including relevant annexes and where applicable, detailed explanation on scientific grounds for any differences with the PRAC recommendations;
- European Commission decision

The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Resources Database is a public, fully searchable electronic index of the available EU research organisations, networks and data sources in the field of pharmacoepidemiology and pharmacovigilance. ENCePP is a collaborative scientific network coordinated by the European Medicines Agency and developed in collaboration with European experts in the fields of pharmacoepidemiology and pharmacovigilance. Its goal is to further strengthen the post-authorisation monitoring of medicinal products in Europe by facilitating the conduct of multi-centre, independent, post-authorisation studies focusing on safety and on benefit-risk, using available expertise and research experience across Europe.

At US level:

The Adverse Event Reporting System (AERS), which is the database which supports the FDA's post-marketing safety surveillance program for all approved pharmaceuticals and therapeutic biological products. In MedWatch website FDA collects information about adverse reactions and data are publicly available.

At international level:

Vigibase Services, which is an international collection of spontaneous reports of suspected adverse reactions, from countries participating in WHO Program for International Drug Monitoring and is maintained by Uppsala Monitoring Centre. National and regional centres in all official and associated member countries have access to the data.

Early signal detection is the task of regulatory authorities and not really of HTA assessors. Unless confirmed by regulatory authorities, they should not contribute to a re-appraisal of a pharmaceutical by an HTA agency. There are other possible safety triggers for re-appraisal of a pharmaceutical: relevant serious adverse events observed post-authorisation that may change benefit/harm balance or published literature data indicating an increased risk (e.g. increased incidence of cancer). However, this information is assessed by regulatory authorities and, if they change the benefit-risk balance, appropriate measures are taken so that the benefit-risk balance remains positive (a restriction of conditions of use or the need of periodic diagnostic procedures, among others). At re-appraisal, identified safety concern of a pharmaceutical (e.g. incidence of hepatotoxicity) is compared to the same safety concern (incidence of hepatotoxicity) of a comparator. However the whole overall safety picture of both medicines has to be considered, in order to assess other advantages/disadvantages.

The European Commission has amended the pharmacovigilance system in 2010 and proposed EU pharmacovigilance legislation (Regulation (EU) No 1235/2010 Directive 2010/84/EU) in order to continue to ensure greater patient safety. This pharmacovigilance legislation strengthened the system for safety monitoring of medicines on the European market in order to obtain major robustness and transparency. A scientific committee, the PRAC, was set up. The pharmacovigilance legislation covers all aspects of the risk management of the use of medicinal products for human use including the detection, assessment, communication relating to the risk of adverse reactions, taking into account the therapeutic effect of the medicinal product, the design and evaluation of post-authorisation safety studies and pharmacovigilance audit. It is useful to take into account the rules laid down by the pharmacovigilance legislation in consideration of the possible implications for the safety assessment in the field of REA.

Where to find the information

- EPAR: <http://www.ema.europa.eu/ema/>
- SPC: <http://www.ema.europa.eu/ema/>
- RMP: <http://www.ema.europa.eu/ema/>
- AERS: <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/ucm083765.htm>
- PASS: http://www.encepp.eu/encepp_studies/indexRegister.shtml
- ENCePP: <http://www.encepp.eu/index.shtml>
- Eudravigilance database: <http://www.adrreports.eu/EN/disclaimer.html>
- Vigibase Services, Uppsala Monitoring Centre: <http://www.who-umc.org>
- Medwatch: <http://www.fda.gov/medwatch/index.html>
- FDA summary review:; <http://www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm>
- Australian Public Assessment Report (AusPAR): <http://www.tga.gov.au/>
- Product Monograph: <http://www.hc-sc.gc.ca/index-eng.php>

2.3.1.2. Regulatory sources for medical devices

Data at approval phase

In the EU, there is no centralized authorization procedure for market access comparable with pharmaceuticals. Instead, medical devices have to run through a so-called conformity assessment. Herein the manufacturer or his authorised representative has to demonstrate through the use of appropriate conformity assessment procedures that the device complies with the relevant Essential Requirements covering safety and performance of the respective medical devices. Within this framework there is no need for demonstration of clinical effectiveness, and clinical trials are rarely required. Besides, all documents elaborated for granting marketing authorisation (declaration of conformity) of medical devices in the EU such as e.g. clinical evaluation report, clinical study report itself – if a clinical study has been performed – or the technical documentation are not publicly available.

In contrast to the EU, there is a centralized authorization procedure for medical devices in the US. Depending on the level of control to assure safety and effectiveness, the FDA provides two regulatory pathways for market access. The by far most common is the so called 510(k) premarket notification. A 510(k) is a premarket submission made to FDA to demonstrate that the device to be marketed is at least as safe and effective (that is substantially equivalent) to a legally marketed so-called “predicate device” that is not subject to premarket approval. While clinical trials are rarely required for this pathway, the so-called Premarket Approval (PMA) pathway is comparable to those requirements for a pharmaceutical approval. PMA is the process of scientific and regulatory review to evaluate the safety and effectiveness of so-called Class III medical devices. Under the US regulation Class III devices are those that support or sustain human life, are of substantial importance in preventing impairment of human health, or which present a potential, unreasonable risk of illness or injury. This definition in general includes those MD not allowed to go through 510(k) due to the non-existence of a predicate device. Please note, that the definition of Class III MD in the US and the EU is not identical.

Medical devices cleared in the context of a 510(k) procedure or approved under PMA and at least summaries of corresponding clinical data including adverse effects are made publicly accessible by the FDA:

- Cleared 510(k) are added to the 510(k) database
- PMA orders are added to the PMA database.

Data at post approval phase

In the EU the medical device directives require manufacturers to implement a post-market surveillance procedure or programme, which includes an obligation to report so called “incidents” (see section 1.1 Definitions) to the relevant competent authorities. Unlike the regulation of pharmaceuticals – which does not depend on the seriousness of an event – the MD vigilance

system focusses on circumstances "which might lead to or might have led to the death of a patient or user or to a serious deterioration in his state of health" (Medical Device Directive 93/42/EG Annex VII). Adverse events not corresponding to the definition of "incident" are therefore only recorded in the context of clinical investigations.

Irrespective of this limited spectrum of events to be reported by MD manufacturers, information held by national competent authorities in connection with the medical device vigilance system is to be held in confidence, as defined by the relevant articles of the directives (e.g. article 20 Medical Device Directive 93/42/EG)¹⁵. As such there is no publicly accessible database providing suitable vigilance data. The European Databank on Medical Devices (EUDAMED) which also contains data in accordance with the vigilance procedure only acts as a central repository for information exchange between national competent authorities and the Commission and is not publicly accessible.

In contrast to the EU, in the US there is a publicly accessible database which supports the FDA's post-marketing safety surveillance program for MDs. The US Manufacturer and User Facility Device Experience Database (MAUDE) represents reports of adverse events involving medical devices including mandatory reports from manufacturers and importers and voluntary reports from health care professionals, patients and consumers. The MAUDE data are presented in four logical record types (Master event data, device data, patient data, text data) – all four types of files should be downloaded. However, FDA itself states that MAUDE data is not intended to be used either to evaluate rates of adverse events or to compare adverse event occurrence rates across devices.

Where to find information

FDA 510(k) database: <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfPMN/pmn.cfm>

FDA Devices@FDA: www.accessdata.fda.gov/scripts/cdrh/devicesatfda/index.cfm

FDA PMA database: <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfPMA/pma.cfm>

FDA MAUDE database:

<http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfMAUDE/TextSearch.cfm>

NIHR Horizon Scanning Research & Intelligence Centre: <http://www.hsric.nihr.ac.uk>

2.3.2. Randomised Clinical Trials (RCTs)

From RCTs, data on well-recognised, frequent, easily detectable adverse reactions are normally identified.¹ Whether or not information on rare and long term effects that occur in clinical practice can be inferred from RCT, depends on the size, duration and other features of trials, but this also applies to all other types of studies. Clinical trials are usually powered to detect statistical significance of possible benefit(s) of a technology and only secondly designed to study safety.^{1,2,16}

As a result the evidence on adverse effects generated by RCTs may not be conclusive.

Patients included in clinical trials may not reflect the features of populations who will undergo the treatment in clinical practice affecting the external validity of the trial.¹ In fact RCTs may fail in identifying risks in populations not included in the trial and in identifying, depending on the design of the study, some categories of reactions such as reactions where young or old age is a risk factor, the effect of an intervention on other diseases, the effect on pregnancy, reactions related to sex, reactions due to genetic variations in different ethnic groups, those associated with other interventions not studied in the trial, withdrawal effects, and unfavourable changes in death rate because of low number of participants and short period of observation.¹⁷

Although head to head RCTs are the most direct evidence on comparative safety, placebo-controlled RCTs may be important to obtain information on absolute and relative risks and more precise estimates of adverse reactions.²

Studies sufficiently powered to assess (differences in) adverse effects, when possible, are important sources of information in the safety assessment.

The HTA assessors should also attempt to include results of completed but not published RCTs and unpublished results of published trials.² This kind of information may be collected searching abstracts presented during congresses and through the public assessment report performed by regulatory authorities.

2.3.3. *Observational studies*

RCTs are considered the best standard for evaluating benefits, but for safety assessment the use of additional sources of information could be necessary.²

In some cases, observational studies provide a wider number of participants observed for a longer period of time and may be more likely to capture rare and long term effects.¹¹ Whether or not data from observational studies have higher degree of external validity than RCTs, depends on the specific question and the design of specific studies. Nevertheless, because of the lack of randomisation the observational studies are more subject to risk of bias, which are normally taken into account in the design itself or in the data analysis.

There is debate concerning the capacity of the different study designs (randomised controlled trials and observational studies) to yield reliable quantitative estimate of adverse reactions.

Papanikolaou et al. found that non randomised studies are often conservative in estimating absolute risks. These differences were largely due to inconsistencies in study populations.¹⁸ On the other hand a recent review presented no difference on average between estimates of adverse effects from meta-analyses of RCTs and of observational studies. Therefore it suggested evaluating a broad range of studies to obtain an exhaustive assessment of adverse reactions with wider generalisability.¹⁹

Observational studies based on analyses of large administrative databases/registries are efficient in providing information on safety issues since they can be performed in shorter timeframes. They are probably more useful for evaluating serious adverse effects that are more reliably recorded than less serious adverse effects that may not generate a specific clinic visit or diagnostic code.²

2.3.4. *Case Reports*

Case reports are useful to collect data on uncommon, unexpected or long-term adverse reactions not normally identified in clinical trials.⁸ One of the most important advantages of case reports is to contribute to signal generation which may then be confirmed by further studies.^{20,21}

A study showed that 90% pharmaceutical withdrawals from French market between 1998 and 2004 were supported by spontaneous case reports.²² This study was in accordance with Arnaiz et al which investigated reasons for withdrawals of pharmaceuticals in Spain between 1990 and 1999.²³ However it should be pointed out that case reports present relevant limitations, being their role the generation of signals that need to be further confirmed. Limitations include the lack of important information, such as conclusive evidence on the estimate of the incidence.¹² In addition, even though they may be published in scientific journals they are seldom subjected to confirmatory investigations.²⁴

2.3.5. *How and where to find the information*

Because of poor indexing and inconsistent terminology the identification of studies reporting data on adverse reactions may be problematic.^{25,26,27}

There is no single approach of search strategy to collect evidence on adverse effects, but a combination of different approaches is required. Two main approaches can be followed: searching electronic database using index terms (such as MeSH in MEDLINE) and free text terms. The two approaches should be combined.²⁸

A search through clinical trials registers (i.e. www.clinicaltrials.gov, www.clinicaltrialsregister.eu or www.controlled-trials.com) to identify ongoing or completed but not published trials could provide additional information. In the EU, abstracts and summary results of post authorisation

safety studies requested by regulatory authorities to marketing authorisation holders will be public according to new legislation (Regulation (EU) No 1235/2010 Directive 2010/84/EU). EVIDENT database (access for EUnetHTA partners only) could be searched for identifying requests/recommendations by agencies for additional evidence collection concerning safety.

Further valuable advice concerning information retrieval in the area of the safety domain can be found in SuRE Info on the HTAi webpage²⁹. In addition, the EUnetHTA guideline "Process of information retrieval for systematic reviews and health technology assessments on clinical effectiveness" published in 2015 contains a lot of detailed recommendations on information retrieval and search strategies also applicable to safety issues.

2.4. Evaluation of sources of information

HTA assessors should adequately assess the risk of bias of studies and the quality of safety data to decide on the inclusion in the assessment and then on their interpretation.

2.4.1. Regulatory sources

The documentation supporting authorisation of pharmaceuticals usually includes clinical trials; therefore the quality issues relating RCTs described in paragraph 2.4.2 can be applied also to these regulatory sources. Depending on the availability of regulatory sources for medical devices and the type of clinical trial included, the respective information on evaluation of sources as documented below should be applied. Irrespective of the technology: the safety assessment from different regulatory authorities may vary, making a comparison useful. Moreover a technology may have a different regulatory status across jurisdictions; it could be in approval phase in a jurisdiction and may have been marketed in another jurisdiction and therefore different data may be available.

2.4.2. Randomised clinical trials (RCTs)

It is recommended to evaluate both the risk of bias of individual studies and the quality of data on adverse effects. For the assessment of risk of bias several methods are available and they may be in the form of scales, checklist or individual components. A tool was developed by the Cochrane Collaboration.³⁰ In any case methods used to assess the risk of bias should be clearly described and results should be reported, also in table format. Moreover it should be explained how the information on risk of bias will be used in the synthesis of data.⁴⁷ For general recommendations see EUnetHTA guideline on levels of evidence: internal validity (of randomized controlled trials).

In assessing quality of data HTA assessors should evaluate two main aspects:

- how adverse effects were identified and collected and
- how they were reported.

HTA assessors should bear in mind that methods used to monitor or detect adverse effects greatly influence their reported frequency.^{1,7,12} As a result, studies using different methods of monitoring effects are difficult to be compared.¹

For instance active methods (such as querying patients) and use of checklist are more likely to identify adverse effects than passive (relying on patient self report) or less focused methods.^{1,2} It is also important to consider how adverse effects are measured and if there are differences in measurements between studies (for example, for the adverse effect fever how and when the temperature is measured across trials should be taken into account).

Useful questions (also for observational studies), elaborated by the Cochrane Collaboration⁷, to evaluate how adverse effects are collected are the following:

- Were definitions given of reported adverse effects?

- How were adverse effects data collected: prospective/routine monitoring, spontaneous reporting, patient checklist/questionnaire/diary; systematic survey of patients?

The quality of adverse effects reporting in RCTs is often variable.^{31,32,33,34,35,36}

An evaluation of safety reporting in randomised trials across seven different medical areas demonstrated that safety reporting was often varying. Adverse effects were reported adequately only in 39% of the trials identified and many trials reported adverse effects without specifying severity and frequency.³⁶

The categorisation of adverse effects may differ across trials, making the synthesis of results difficult.⁶ Using a common, widely accepted scale would have the advantage that information can then be compared and synthesised across different studies.³⁷ Standardised scales are available for some conditions,² including National Cancer Institute (Common Terminology Criteria for Adverse Event)³⁸ and WHO scales.³⁹ The CTCAE is considered the standard for reporting the severity of adverse effects in oncology clinical trials.⁴⁰

For these reasons HTA assessors should adequately evaluate how adverse effects are reported. The Cochrane Collaboration identified questions to assess the quality of reporting (also valid for observational studies):

- Were any patients excluded from the adverse effects analysis?
- Did the report give numerical data by intervention group?
- Which categories of adverse effects did investigators report?
- Did investigators report on all important or serious adverse effects, and how were these defined?
- Were the methods used for monitoring adverse effects reported?
- Was an independent data safety monitoring board established?⁷

If adverse effects are independently adjudicated and/or confirmed by chart review, they should be evaluated.

The Consolidated Standards of Reporting Clinical Trials (CONSORT) presented the most frequent inadequate practices in reporting adverse effects in RCTs (see table 1)⁴¹ which should be considered when assessing data to include in the analysis.

Additional specific aspects to be considered in the context of non-drug interventions primarily based on medical devices (e.g. care providers' expertise, centres' volume, blinding of patients, care providers and outcome assessors) are dealt with in the EUnetHTA guideline on Therapeutic medical devices.

Table 1. Practices to avoid in reporting adverse effects in RCTs (Ioannidis JP et al. 2004)

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 randomised 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.

Although withdrawals are an important outcome, HTA assessors should be cautious when interpreting withdrawals as surrogates for safety or tolerability because of the potential for bias.

2.4.3. Observational studies

Because of the lack of randomisation the quality of observational studies should be adequately assessed.

Available instruments for evaluating observational studies vary in scope, number and types of items used and developmental rigor.²

The STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) is a guidance on how to properly report observational studies and provides a checklist of 22 items that should be addressed in papers reporting cohort, case-control, and cross-sectional studies (see Annex 1. The STROBE statement - checklist of items that should be addressed in reports of observational studies).⁴²

Special aspects are to be considered in assessing the quality of non-randomised studies, such as the selection of the study groups; the comparability of the groups; and the ascertainment of either the exposure or outcome of interest for case-control or cohort studies, respectively. Additional factors to be considered can be:

- a) any validation of the case definition and/or outcomes (or absence of outcome) in the data source - either as part of the ongoing study or from literature
- b) justification of the suitability of the proposed case/outcome definition for the safety study.

It is necessary to assess the quality of data on adverse reactions in observational studies. Most of the considerations presented for RCTs are valid and helpful also for observational studies.

For details on quality assessment of non-randomised studies (including registry-analysis) see EUnetHTA Guideline Internal validity of non-randomised studies (NRS) on interventions.

2.4.4. Case reports

Given the considerable limitations, HTA assessors should adequately assess published case reports when judging about their inclusion.

For this purpose HTA assessors should consider the following aspects:⁷

- Do reports have a good predictive value?
- Is there a causal relation between the intervention and the adverse event?
- Is there a plausible (biological) mechanism linking the intervention to the adverse events?
- Do reports provide enough information to allow detailed appraisal of the evidence?
- Are there any potential problems from using data from the reports, which might outweigh the perceived benefit of being comprehensive?

A set of desirable contents of a case report was proposed and included 14 items that are applied by the former Committee on Safety of Medicines' yellow card in the United Kingdom, and 14 from the MedWatch adverse event forms in the United States (see Annex 2. Preferred contents of a case report).⁴³ The respective information are desirable for case reports concerning MDs as well.

HTA assessors should take into account that published case reports and spontaneous reporting may provide different frequencies of adverse effects from those obtained from a meta-analysis of double-blind, randomised controlled trials.¹

2.5. Synthesis and reporting of results by HTA assessors

At this stage of the assessment, after collecting the evidence and judging on their inclusion, HTA assessors should clearly describe included sources of information, risk of bias and quality of

adverse reactions data, quantify adverse reactions in terms of frequency, incidence, severity and seriousness in comparison with its comparators.

2.5.1. Description of included sources of information

Characteristics of selected studies which may influence results and their external validity should be reported and summarised in tables.⁴⁴

Different approaches are available, including PRISMA Statement (Preferred reporting for Systematic Reviews and Meta-Analyses) and Cochrane Collaboration's approach.^{45,46}

PRISMA Statement recommends describing the characteristics of included studies (e.g. study size, follow-up period, PICOS). The Cochrane collaboration suggests to present a table entitled "Characteristics of included studies" including the following items: methods (study design, duration of the study); participants (setting, relevant details of health status of participants, age, sex and country); interventions (e.g. for pharmaceuticals: the name, dose, frequency, way of administration, duration); outcome and notes. Both approaches recommend describing the risk of bias of each study in a separate table.

According to these schemes useful information to be reported by HTA assessors are the following: methods (study design and follow-up period), characteristics of participants for both arms (e.g. setting, age, sex and country/geographic area, if appropriate race), intervention and comparator(s) (e.g. for pharmaceuticals: the name, dose, frequency, route of administration, duration), and outcomes. It is important to consider and to report the exposure of patients to the treatment. Methods used to collect adverse effects should be described as well (see table 2. as an example).

Table 2. Characteristics of included studies

Study (reference)	Methods (Study design and follow up)	Participants (e.g. setting, age, sex and country/geographic area)	Intervention and comparator(s) (e.g. for pharmaceuticals: name, dose, frequency, route of administration and duration)	Outcomes	Methods used to collect adverse effects

Tables describing characteristics of included studies can be presented separately for each study design as reported by Singh et al.⁴⁷

During JA2 subgroup 4 members of Work Package 7 developed manufacturers' submission templates to support production of core HTA information and rapid assessments. Some modules of the submission template are dedicated to safety issues. They include also tables to list, evaluate and synthesise the given evidence concerning the technology's safety. Special modules in regard to the safety of pharmaceuticals (e.g. safety risk management) and medical devices (e.g. manufacturer vigilance data) have been integrated. [Source to be mentioned when the templates will be published by EUnetHTA]

Additional specific aspects to be considered in the context of non-drug interventions primarily based on medical devices (e.g. precise description of the interventions or details on intervention standardisation) are dealt with in the EUnetHTA guideline on Therapeutic medical devices.

2.5.2. Quantification of adverse effects in terms of frequency, incidence, severity and seriousness

According to previously developed guidelines (see Annex 3. Methods for documentation and selection criteria), results of individual studies can be presented by study arm in tabular form, using at least the following measures:

- Number of participants per study arm
- Number of patients excluded from the analysis dataset
- Patient-years of exposure
- Number of participants with the event
- Number of events
- Absolute risk; incidence rate (95% CI)
- Relative risk (95% CI)
- The quality of the evidence (e.g. high, moderate, low and very low)

If the adverse effects data is from an observational study, the relative risk estimate should be adjusted for potential confounding/effect modifying factors.

Different tables can be elaborated for RCTs and observational studies.

The adverse effects should be grouped according to the System Organ Class (SOC). (The classification is available at: <http://www.meddrmsso.com>).

The adverse effects which are common and serious should be reported separately. An example of reporting adverse effects is given in tables 3 adapted from NICE.

The description of adverse effects in terms of duration and reversibility is advisable to understand their burden, taking also into account the exposure to treatment.

Table 3a. Adverse effects by frequency

System organ/ class/adverse effects	Frequency (very common, common, uncommon, rare, very rare, not known)	Intervention % of patients (n = x)	Comparator % of patients (n = x)	Relative risk (95% CI)
Class 1 (for example, nervous system disorders)				
Adverse event 1				
Adverse event 2				
Class 2 (for example, vascular disorders)				
Adverse event 3				
Adverse event 4				

Source: Adapted from NICE. Single Technology Appraisal. Specification for Manufacturer /sponsor submission of evidence.

Table 3b. Adverse effects by seriousness

System organ/class/adverse effects	Seriousness (death, life-threatening, requires hospitalisation or prolongation of existing hospitalisation, persistent or significant disability incapacity, or is a congenital anomaly/birth defect)	Intervention % of patients (n = x)	Comparator % of patients (n = x)	Relative risk (95% CI)
Class 1 (for example, nervous system disorders)				
Adverse event 1				
Adverse event 2				
Class 2 (for example, vascular disorders)				
Adverse event 3				
Adverse event 4				

Source: Adapted from NICE. Single Technology Appraisal. Specification for Manufacturer /sponsor submission of evidence.

Some authors proposed to use composite safety endpoints, merging in a single endpoint different types of adverse effects. As an example composite cardiovascular safety endpoint may include myocardial infarction and heart failure, stroke, coronary revascularisation and out-of-hospital cardiac death.⁴⁸ The data must be provided for composite and single effects (see EUnetHTA guideline on Composite endpoints). The data can be given as either number of effects or hazard ratio (see tables 4 as an example). The use of composite endpoints can have the advantage to facilitate understanding of comparative safety data and to increase the statistical power because of the larger number of participants.

Table 4a. Incidence Rates for Safety Effects per 1000 Person-years among Propensity Score–Matched Older Adults with Arthritis Initiating Prescription Analgesic Treatment (Solomon DH et al, 2010)

Adverse Event ^a	nsNSAIDs			Coxibs			Opioids		
	No. of Events	Person-years	Incidence Rate (95% CI)	No. of Events	Person-years	Incidence Rate (95% CI)	No. of Events	Person-years	Incidence Rate (95% CI)
Composite Safety Events									
Composite cardiovascular ^b	105	1369	77 (63-92)	208	2369	88 (76-100)	196	1604	122 (106-140)
Upper or lower GI tract bleeding	29	1385	21 (14-30)	28	2419	12 (8-16)	35	1668	21 (16-27)
Composite fracture ^c	36	1382	26 (19-36)	45	2419	19 (14-25)	168	1663	101 (87-117)
Hospitalized adverse event	143	1366	105 (89-123)	237	2375	100 (88-113)	247	1598	155 (136-175)
Death related to adverse event	18	1388	13 (8-20)	30	2428	12 (9-17)	20	1696	12 (7-18)
All-cause mortality	66	1388	48 (37-60)	115	2428	47 (39-57)	127	1696	75 (63-89)
Individual Safety Events									
Myocardial infarction	19	1386	14 (9-21)	49	2417	20 (12-27)	48	1675	29 (21-38)
Heart failure	42	1381	30 (22-41)	83	2407	34 (28-43)	75	1654	45 (36-57)
Stroke	31	1379	22 (16-31)	51	2416	21 (16-28)	31	1682	18 (13-26)
Coronary revascularization	7	1386	5 (2-10)	39	2405	16 (12-22)	43	1671	26 (19-34)
Out-of-hospital cardiac death	15	1388	11 (6-17)	19	2428	8 (5-12)	29	1696	17 (12-24)
Upper GI tract bleeding	19	1385	14 (9-21)	16	2422	7 (4-10)	19	1385	14 (9-21)
Lower GI tract bleeding	12	1387	9 (5-15)	12	2426	5 (3-8)	13	1683	8 (4-13)
Bowel obstruction	3	1388	2 (1-6)	8	2427	3 (2-6)	14	1683	8 (5-14)
Acute kidney injury	64	1377	5 (36-59)	94	2408	39 (32-48)	105	1660	63 (52-76)
Falls	36	1374	26 (19-36)	43	2412	18 (13-24)	68	1664	41 (32-51)
Fracture									
Hip	6	1388	4 (2-9)	13	2426	5 (3-9)	19	1694	11 (7-17)
Humerus	7	1387	5 (2-10)	9	2427	4 (2-7)	69	1683	41 (32-52)
Pelvic	8	1384	6 (3-11)	19	2424	8 (5-12)	27	1690	16 (11-23)
Radius	15	1387	11 (6-17)	9	2425	4 (2-7)	58	1684	34 (26-44)
Hepatotoxic effects	16	1386	12 (7-18)	26	2407	11 (7-16)	19	1682	11 (7-17)

Table 4b. Safety Effects among Propensity Score–Matched Older Adults with Arthritis Initiating Prescription Analgesic Treatment (Solomon DH et al, 2010)

Adverse Event ^a	HR (95% CI)		
	nsNSAIDs	Coxibs	Opioids
		Composite Safety Events	
Composite cardiovascular ^b	1 [Reference]	1.28 (1.01-1.62)	1.77 (1.39-2.24)
Upper or lower GI tract bleeding	1 [Reference]	0.60 (0.35-1.00)	1.07 (0.65-1.76)
Composite fracture ^c	1 [Reference]	0.96 (0.62-1.49)	4.47 (3.12-6.41)
Hospitalized adverse event	1 [Reference]	1.12 (0.91-1.38)	1.68 (1.37-2.07)
Death related to adverse event	1 [Reference]	1.12 (0.62-2.02)	1.11 (0.58-2.10)
All-cause mortality	1 [Reference]	1.16 (0.85-1.57)	1.87 (1.39-2.53)
		Individual Safety Events	
Myocardial infarction	1 [Reference]	1.63 (0.96-2.77)	2.25 (1.32-3.84)
Heart failure	1 [Reference]	1.26 (0.87-1.84)	1.63 (1.12-2.38)
Stroke	1 [Reference]	1.04 (0.66-1.63)	0.91 (0.55-1.50)
Coronary revascularization	1 [Reference]	3.30 (1.47-7.38)	5.34 (2.40-11.90)
Out-of-hospital cardiac death	1 [Reference]	0.86 (0.44-1.71)	1.96 (1.05-3.67)
Upper GI tract bleeding	1 [Reference]	0.52 (0.27-1.02)	1.03 (0.55-1.91)
Lower GI tract bleeding	1 [Reference]	0.61 (0.27-1.36)	0.95 (0.43-2.09)
Bowel obstruction	1 [Reference]	1.89 (0.50-7.13)	4.87 (1.40-17.02)
Acute kidney injury	1 [Reference]	0.98 (0.71-1.35)	1.53 (1.12-2.09)
Falls	1 [Reference]	0.73 (0.47-1.14)	1.64 (1.09-2.47)
Fracture			
Hip	1 [Reference]	1.35 (0.51-3.57)	3.02 (1.20-7.58)
Humerus	1 [Reference]	1.06 (0.39-2.85)	9.26 (4.25-20.18)
Pelvic	1 [Reference]	1.68 (0.73-3.86)	3.13 (1.42-6.91)
Radius	1 [Reference]	0.50 (0.22-1.14)	3.74 (2.12-6.60)
Hepatotoxic effects	1 [Reference]	1.08 (0.58-2.01)	1.14 (0.58-2.22)

Adverse effects should also be provided by severity grade, for e.g. anticancer medicines according to the Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0, which includes 5 grades of severity (mild, moderate, severe, life-threatening and death) differently from the previous versions including 4 grades (CAVE: The EPAR may report the data according to the former versions). An example of reporting adverse effects for anticancer medicines is given in table 5. The effects are reported by frequency: Very common ($\geq 1/10$), Common ($\geq 1/100$ and $< 1/10$), Uncommon ($\geq 1/1000$ and $< 1/100$), Rare ($\geq 1/10,000$ and $< 1/1000$), Very rare ($< 1/10,000$) and not known.

Table 5. Frequency and severity of adverse effects classified by System Organ Class (SOC) in trials

System organ class	Frequency	Event*	Pemetrexed/cisplatin (N = 168)		Cisplatin (N = 163)	
			All grades toxicity (%)	Grade 3 - 4 toxicity (%)	All grades toxicity (%)	Grade 3 - 4 toxicity (%)
			Blood and lymphatic system disorders	Very common	Neutrophils/Granulocytes decreased	56.0
Leukocytes decreased	53.0	14.9			16.6	0.6
Haemoglobin decreased	26.2	4.2			10.4	0.0
Platelets decreased	23.2	5.4			8.6	0.0
Metabolism and nutrition disorders	Common	Dehydration	6.5	4.2	0.6	0.6
Nervous system disorders	Very common	Neuropathy-Sensory	10.1	0.0	9.8	0.6
	Common	Taste disturbance	7.7	0.0***	6.1	0.0***

Source: EMA. EPAR (pemetrexed)

Differences in adverse effects among population subgroups (e.g. elderly, adults and children) and specific safety concerns should be addressed and discussed.

Adverse effects from different study designs cannot be pooled together using standard meta-analysis principles. Further, the data from non-randomised studies are more prone to bias and they are often heterogeneous; they should not be combined if there is important heterogeneity. Therefore in these circumstances adverse effects data is best summarised in a qualitative or in a descriptive manner.¹

2.5.3. Comparison of the safety profile of the technology to the comparator(s)

At this stage HTA assessors should describe the safety profile of the technology in comparison to the comparator(s), with special regard to the serious and most frequent adverse reactions.

They should evaluate if differences identified in adverse reactions between the interventions are clinically relevant. The evaluation of the clinical relevance should be performed taking into account the condition for which the treatment is used and the co-morbidities of the population. For instance, in chronic diseases also no serious adverse reactions may have important implications, as they may impair the adherence to the treatment.

HTA assessors should describe limitations of the evidence and analyse how these limitations may affect estimates of the adverse reactions.²

The heterogeneity of the studies should be explored; differences in the characteristics of the studies may lead to different results. Possible effects of individual study characteristics (e.g. follow up period, methods used to identify adverse effects, study design, study size, characteristics of populations, severity of disease and funding sources) and the external validity of results should be studied and discussed in the interpretation of findings.¹²

2.5.4. *Balanced discussion of benefits and adverse effects*

The assessment of relative safety together with relative benefits will contribute to establish a balanced assessment of the intervention compared to its comparator(s). Some frameworks reporting both benefits and adverse effects were proposed.⁵³

While performing REA, HTA assessors should describe possible consequences of safety appraisal on coverage decisions:

- coverage restriction: patients with high risk of developing a serious adverse effect may be excluded from coverage
- reimbursement may be lower, restricted or not acceptable for technologies with safety concerns

3. Discussion and Conclusions

Although there is no doubt on the importance of the relative safety assessment of health technologies, significant methodological issues still persist.

The identification of adverse reactions to include in the assessment can be challenging; unlike benefits which are well identified, some adverse reactions associated with an intervention may not be identified in advance.⁷

Moreover the identification of studies with data on adverse reactions is not necessarily straightforward for several reasons: some studies don't collect and provide data on the frequency of adverse reactions and, even when adverse reactions are evaluated, the information is not reported in the title and abstract; in some cases papers are not assigned indexing terms for adverse reactions, even though they contain data on adverse reactions frequency, making difficult the identification of the study.¹⁹

Clinical trials are usually powered to detect statistical significance of possible benefits of a technology and only secondly designed to study safety.¹⁶ As a result the evidence on adverse reactions generated by RCTs may be inconclusive.² Depending on the size and duration, clinical trials may fail in capturing long term and uncommon adverse reactions; in order to identify uncommon adverse reactions they should enrol a larger number of participants impacting negatively on time needed for development of the technology.¹⁶

In spite of these limitations the relative safety assessment plays an important role in the relative effectiveness assessment. The assessment of relative safety together with benefits contribute to establish a balanced assessment of the intervention and of its therapeutic value and to support the payers in making informed decisions on the coverage of the health technologies.

For these reasons it is important to assess adverse reactions and benefits with the same methodological rigour and accuracy. It will be important to consider the rules laid down by the pharmacovigilance legislation (Regulation (EU) No 1235/2010 Directive 2010/84/EU) and their implications for the safety assessment in the field of Relative Effectiveness Assessment. In this regard, HTA can take advantage of the input of the continuous benefit risk assessments to be performed by regulatory authorities. Currently it is unclear whether and when such developments come true for MD as well.

When conducting relative safety assessment the objective of HTA assessors should be first to identify adverse reactions, then to examine data in terms of frequency, incidence, severity and seriousness and finally to compare the safety profile of the technology with its comparator(s). To ensure the use of appropriate terminology the MedDRA dictionary could be a useful tool. In rapid assessment primary sources of information are EPAR, SPC, RMP (when available) in the case of pharmaceuticals and publicly available approval orders and safety reports from the FDA in the case of MD, respectively, manufacturer dossier and published and unpublished (where acceptable under the specific HTA system guidelines) clinical trials. Other sources of information such as observational studies, registries and confirmed relevant signals are also useful when available.

As the quality of primary data on adverse reactions may be heterogeneous, it is important to evaluate, apart from the risk of bias, how adverse effects were collected and reported in the studies.

Main characteristics of sources information should be reported and summarised in tabular form. Results on adverse reactions, categorised by SOC, should be reported in terms of number of participants per study arm, number of patients excluded from the analysis dataset, patient-years of exposure, number of events, number of participants with the event, absolute risk, incidence rate, relative risk and the quality of evidence for the technology and the comparator(s).

Finally a description of the safety profile of the technology in comparison with its comparator(s) with special regard to the most frequent, serious and severe adverse reactions should be given. External validity and heterogeneity of included sources of information, considering all factors which may influence the occurrence of adverse reactions (follow up period, methods used to identify adverse effects, study design, study size and characteristics of the population included), should be taken into account in the interpretation of findings.

Annexe 1. The STROBE statement

Checklist of items that should be addressed in reports of observational studies

	Item number	Recommendation
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract, an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow up, and data collection
Participants	6	(a) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow up <i>Case-control study</i> —Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants (b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed <i>Case-control study</i> —For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) <i>Cohort study</i> —If applicable, explain how loss to follow up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses
Results		
Participants	13*	(a) Report the numbers of individuals at each stage of the study—e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow up, and analyzed (b) Give reasons for nonparticipation at each stage (c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) <i>Cohort study</i> —Summarize follow-up time (e.g., average and total amount)
Outcome data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time <i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure <i>Cross-sectional study</i> —Report numbers of outcome events or summary measures
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarize key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalizability	21	Discuss the generalizability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

Source: von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche Peter C, Vandenbroucke JP for the STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting Observational studies. *J of Clin Epid* 2008; 61: 344-349

Annexe 2. Preferred contents of a case report

Section	Contents
The title	The suspected adverse event
	The suspected drug
	Age and sex of the patient (single case reports)
	Number of patients (multiple case reports)
	Important risk factors
Structured summary	Adverse event
	Drug implicated
	The patient(s)
	Evidence that links the drug to the event
	Management
	Mechanism, if known
	Implications for therapy
Hypotheses to be tested	
The introduction	The suspected drug and the adverse event with which it was associated*
	Previous similar reports
	The purpose of the report
The case report	
Demographic information	Age*†, sex*†, weight*†, and ethnic background†
Diagnoses	All diagnoses, especially those for which drug therapy was indicated*†; specify allergies (present or absent)*†
Drug therapy	All current drug therapy, including dosage, duration, and indication*†; other recent drug therapy, if relevant*†
Other relevant history (including relevant negatives)	Relevant family history
	Relevant social history†
The adverse event	Assessment of severity*
	Time-course in relation to the administration of the suspected drug*†
	The effect of withdrawal*†, including time-course
	The effect of rechallenge*†, including time-course
	Results of diagnostic tests (in vivo, in vitro, or ex vivo)*†
	Plasma concentrations (parent compound and main metabolites)
	Data from animal or in vitro studies
	The final outcome*†
Treatment	Measures that were taken to treat the adverse event*
The discussion	Assessment of the likelihood that the event was an adverse drug reaction
	Why the drug was implicated
	Why other drugs that the patient took were not responsible
	Elimination of other possible causes
	A review of previous cases, published and unpublished
	Methods of diagnosis
	Possible mechanisms
	Possible forms of management
	Implications of the report for clinical practice
Hypotheses generated by the report	

*These 14 details are solicited on the yellow card for reporting suspected adverse drug reactions to the Committee on Safety of Medicines in the United Kingdom. These 14 details are solicited on the MedWatch adverse event forms in the United States.
Source: Aronson JK. Anecdotes as evidence. BMJ 2003; 326: 1346

Annexe 3. Methods of documentation and selection criteria (applied during original guideline elaboration)

For the elaboration of the original document a literature search was carried out using the following keywords:

adverse drug reaction, adverse event, adverse effect, comparative effectiveness, comparative effectiveness research, drug safety and comparative safety.

Source of information

Database:

Ovid Medline was searched for the literature review.

Websites:

European Medicines Agency (EMA)
Food and Drug Administration (FDA)

Agency for Healthcare Research and Quality (AHRQ)
Agency for Health Technology Assessment and Tarrif System (AOTMiT)
Autoridade Nacional do Medicamento e Produtos de Saude (INFARMED)
Canadian Agency for Drugs and Technologies (CADTH/CEDAC)
College voor zorgverzekeringen (CVZ)
Hautè Autoritè de Santè (HAS)
Health Information and Quality Authority (HIQA)
Institute for Quality and Efficiency in Health Care (IQWiG)
Medical services Advisory Committee (MSAC)
National Institute for Health and Care Excellence (NICE)
Pharmaceutical Benefits Advisory Committee (PBAC)
Pharmaceutical Management Agency (PHARMAC)
Scottish Medicines Consortium (SMC)

Centre for Reviews and Dissemination, University of York (CRD)
Consolidated Standards of Reporting Trials (CONSORT)
Guidelines International Network (GIN)
Health Technology Assessment international (HTAi)
Institute of Medicine (IOM)
International Conference on Harmonization (ICH)
International Society for Pharmacoeconomics and Outcomes Research (ISPOR)
National Guideline Clearinghouse
Transparent Reporting of Systematic Review and Meta-Analyses (PRISMA)
The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)
World Health Organization (WHO)

Guidelines, reports, recommendations already available

A review of the guidelines used by HTA agencies addressing the assessment of safety and the methods to report data on adverse events was conducted and the following documents were selected:

- Agency for Health Technology Assessment (AHTAPol). Guidelines for conducting Health Technology Assessment (HTA) Version 2.1 April 2009.
- Pharmaceutical Benefits Advisory Committee (PBAC). Department of Health. Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee. (Version 4.3).December 2008.
- Medical Services Advisory Committee (MSAC). Funding for new medical technologies and procedures: application and assessment guidelines September 2005.

- National Institute for Health and Clinical Excellence (NICE). Single Technology Appraisal. Specification for manufacturer/sponsor submission of evidence. October 2009.
- Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1 [updated September 2008].

The most common practices used to report the safety data in a HTA report were identified. The results of the review are shown in table 6.

Table 6. Methods to report safety data by EMA, FDA, Cochrane and by the HTA agencies

	EPAR (EMA)	Summary Review (FDA)	Cochrane Handbook	AHTAPol (Poland 2009)	Australia (PBAC 2008)	Australia (MSAC 2005)	UK and Wales (NICE 2009)
			Cochrane Handbook for Systematic Reviews of Interventions Cochrane Book Series ⁴⁵	Guidelines for conducting Health Technology Assessment (HTA) ⁴⁹	Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee ⁵⁰	Funding for new medical technologies and procedures: application and assessment guidelines ⁵¹	Single Technology Appraisal. Specification for manufacturer / sponsor submission of evidence ⁵²
No. of studies addressing the outcomes			✓				
Outcome	✓	✓	✓	✓	✓	✓	✓
Quality of the evidence			(GRADE)	(GRADE)		✓	
Frequency (very common and common)	✓						
Severity	✓						
System Organ class	✓	✓					✓
No. patients in the group	✓	✓	✓			✓	✓
No. of patients with the event		✓				✓	✓
Absolute risk	✓	✓	✓ (per 1000 people) (CI 95%)		✓	✓ (CI 95%)	✓
Hazard ratio			✓				
Absolute risk reduction				✓			
Relative risk			✓ (CI 95%)	✓	✓ (CI 95%)	✓	✓ (CI 95%)
Relative risk reduction				✓			
Risk difference					✓ (CI 95%)		✓ (CI 95%)
NNH				✓		✓	
Odds ratio			✓ (CI 95%)	✓			
Mean difference			✓ (CI 95%)		✓ (CI 95%)		
Standardised mean difference			✓ (CI 95%)				

Other

Bibliography of selected documents

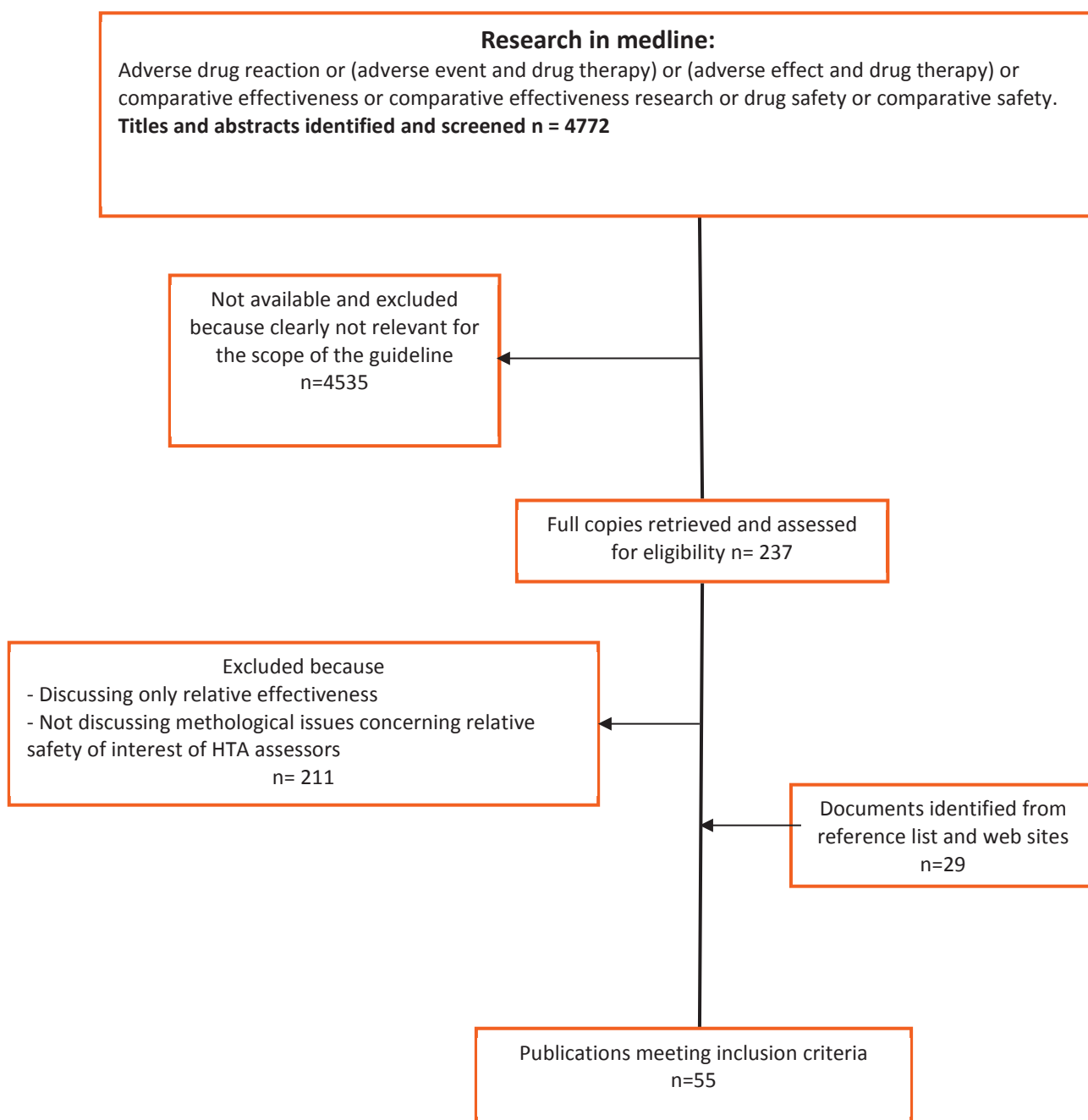
Bibliographic search strategy

The sources of information were searched for the period 2000-2010. The search was restricted to human subjects and to English language.

Selection criteria

Documents were selected for this review if they addressed the methodological issues related to the assessment of relative safety conducted by HTA assessors; documents related only to relative effectiveness and concerning issues of regulatory competence were excluded. The selection of the papers was carried out in two phases. A first screening was conducted in according to the title and abstract, afterwards the full texts of the papers selected relevant to the guidance were identified.

Figure 1. Flow chart diagram of study selection



Annexe 4. Bibliography

- ¹ Loke YK, Price D, Herxheimer A, Cochrane Adverse Effects Methods Group. Systematic reviews on adverse effects. framework for a structured approach. BMC Med Res Method 2007;7:32.
- ² Chou R., Aronson N, Atkins D, Ismaila AS, Santaguida P, Smith DH, Whitlock E, Wilt TJ, Moher D. AHRQ paper 4: assessing harms when comparing medical interventions: AHRQ and effective Health Care Program. Journal of clinical epidemiology 2010; 63: 502-512.
- ³ Golder, Loke Y, McIntosh HM. Room for improvement? A survey of the methods used in systematic reviews of adverse effects. Medical Research Methodology 2006, 6(3):1-6.
- ⁴ Cornelius VR, Perrio MJ, Shakir SAW et al. Systematic reviews of adverse effects of drug interventions: a survey of their conduct and reporting quality.
- ⁵ Etminan M, Carleton B, Rochon PA. Quantifying Adverse Drug Events. Are Systematic Reviews the Answer? Drug Safety 2004; 27 (11): 757-761.
- ⁶ Cuervo LG, Clarke M. Balancing benefits and harms in health care [Editorial]. BMJ. 2003;327:65-6.
- ⁷ Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1 [updated September 2008]. Chapter 14: Adverse effects. Authors: Yoon K Loke, Deirdre Price and Andrew Herxheimer on behalf of the Cochrane Adverse Effects Methods Group.
- ⁸ Chou R, Helfand M. Challenges in Systematic Reviews That Assess Treatment Harms. Ann Intern Med. 2005; 142(12):1090-1099.
- ⁹ MedDRA - the Medical Dictionary for Regulatory Activities. Available and accessed on October 15 2010: <http://www.meddramsso.com/>
- ¹⁰ Brown E G. Effects of Coding Dictionary on Signal Generation. A Consideration of Use of MedDRA Compared with WHO-ART. Drug Safety 2002; 25 (6): 445-452.
- ¹¹ Brown E G. Methods and Pitfalls in Searching Drug Safety Databases Utilising the Medical Dictionary for Regulatory Activities (MedDRA). Drug Safety 2003; 26 (3): 145-158.
- ¹² Centre of reviews and Dissemination (CRD). Systematic reviews of adverse effects. Systematic reviews. CRD's Guidance for undertaking reviews in health care. January 2009;p.177-197.
- ¹³ VOLUME 9A of The Rules Governing Medicinal Products in the European Union – Guidelines on Pharmacovigilance for Medicinal Products for Human Use. Accessed in January 2011 and available at: http://ec.europa.eu/health/files/eudralex/vol-9/pdf/vol9a_09-2008_en.pdf
- ¹⁴ Guideline on good pharmacovigilance practices (GVP). Module VII- Periodic safety update report. Data for coming into effect: 2 July 2012 . Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/06/WC500129136.pdf
- ¹⁵ MEDDEV 2.12-1 rev8, 2013. Accessed in June 2015 and available at: http://ec.europa.eu/growth/sectors/medical-devices_old/documents/guidelines/files/meddev/2_12_1_ol_en.pdf

- ¹⁶ Eichler HG, Pignatti , Flamion B et al. Balancing early market access to new drugs with the need for benefit/risk data: a mounting dilemma. *Nature Reviews. Drug Discovery* 2008;7:819-826.
- ¹⁷ Clark D WJ, Coulter DM, Besag F MC. Randomized controlled trial and Assessment of Drug Safety. *Drug Safety* 2008; 31 (12): 1057-1061.
- ¹⁸ Papanikolaou PN, Christidi GD, Ioannidis JPA (2006) Comparison of evidence on harms of medical interventions in randomized and nonrandomized studies. *CMAJ* 2006; 174 (5): 635–641.
- ¹⁹ Golder S, Loke YK, Bland M. Meta-analysis of adverse effects data derived from randomised controlled trials as compared to Observational studies: Methodological overview. *Plos Medicine* 2011; 8(5):1-13
- ²⁰ Vandenbroucke JP. In defense of case reports and case series. *Ann Intern Med* 2001;134:330-4.
- ²¹ Vandenbroucke JP. Case reports of suspected adverse drug reactions. Case reports were dismissed too quickly. *BMJ* February 2006;332: 488.
- ²² Olivier P, Montastruc JL. The nature of the scientific evidence leading to drug withdrawals for pharmacovigilance reasons in France. *Pharmacoepidemiology and drug safety* 2006; 15: 808–812.
- ²³ Arnaiz JA, Carne X, Riba N, Codina C, Ribas J, Trilla A. The use of evidence in pharmacovigilance. Case reports as the reference source for drug withdrawals. *Eur J Clin Pharmacol* 2001; 57:89-91.
- ²⁴ Loke YK, Price D, Derry S, et al. Case reports of suspected adverse drug reactions: systematic literature survey of follow up. *BMJ* 2006; 332: 335-9.
- ²⁵ Golder S, McIntosh HM, Loke Y. Identifying systematic reviews of the adverse effects of health care interventions. *BMC Medical Research Methodology* 2006; 6:22:1-11.
- ²⁶ Derry S, Kong Loke Y, Aronson JK. Incomplete evidence: the inadequacy of databases in tracing published adverse drug reactions in clinical trials. *BMC Med Res Methodol* 2001; 1:7.
- ²⁷ Golder S, Loke YK. Search strategies to identify information on adverse effects: a systematic review. *Med Libr Assoc* April 2009; 97(2):84-92.
- ²⁸ Golder S, McIntosh HM, Duffy S, Glanville J: Developing efficient search strategies to identify reports of adverse effects in MEDLINE and EMBASE. *Health Info Libr J* 2006; 23(1):3-12.
- ²⁹ <http://vortal.htai.org/?q=node/50>, accessed on 29th of June 2015
- ³⁰ Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.0.1 [updated September 2008]. Chapter 8: Higgins J PT, Douglas GA on behalf of the Cochrane Statistical Methods Group and the Cochrane Bias Methods Group. Assessing risk of bias in included studies.
- ³¹ Ioannidis JPA, Lau J. Completeness of safety reporting in randomized trials: an evaluation of 7 medical areas. *JAMA* 2001; 285: 437-43.
- ³² Loke YK, Derry S. Reporting of adverse drug reactions in randomised controlled trials – a systematic survey. *BMC Clin Pharmacol* 2001, 1:3.
- ³³ Chowes MY, Gottesman BS, Leibovici L, et al. Reporting of adverse events in randomized controlled trials of highly active antiretroviral therapy: systematic review. *Journal of Antimicrobial Chemotherapy* 2009; 64, 239–250.

- ³⁴ Nuovo J, Sather C. Reporting adverse events in randomized controlled trials. *Pharmacoepidemiol Drug Saf* 2007; 16:349-51.
- ³⁵ Pitrou I, Boutro I, Ahmad N et al. Reporting of Safety Results in Published Reports of Randomized Controlled Trials. *Arch Intern Med*. 2009; 169(19):1756-1761.
- ³⁶ Trotti A, Bentzen SM. The need for adverse effects reporting standards in oncology clinical trials. *J Clin Oncol* 2004; 22:19-22.
- ³⁷ Ioannidis JP, Lau J. Improving Safety Reporting from Randomised Trials. *Drug Safety* 2002; 25 (2): 77-84
- ³⁸ Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0. Available and accessed in May 2011: http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf
- ³⁹ World Health Organization (WHO). Cancer treatment: WHO recommendations for grading of acute and sub acute toxicity. *Cancer* 1981; 47: 207-14.
- ⁴⁰ Trotti A, Colevas A D, Setser A, Basch E. Patient-Reported Outcomes and the Evolution of Adverse Event Reporting in Oncology. *J Clin Oncol*. 2007 November; 25(32):5121-5127.
- ⁴¹ Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, Moher D. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004; 141(10):781-788.
- ⁴² von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche Peter C Vandembroucke JP for the STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J of Clin Epid* 2008; 61: 344-349.
- ⁴³ Aronson JK. Anecdotes as evidence. *BMJ* 2003; 326: 1346
- ⁴⁴ Busse R, Orvain J, Velasco M et al. Best Practice in undertaking and reporting health Technology Assessment. *Int J of Technology Assessment in Health care* 2002; 18(2):361-422.
- ⁴⁵ Liberati A, Altman DG, Tetzlaff J et al. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of studies that evaluate Health Care Interventions: Explanation and Elaboration. *Annals of internal Medicine* 2009; 151(4):w65-94.
- ⁴⁶ Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1 [updated September 2008]. Chapter 11: Presenting results and Summary of finding tables* Authors: Schunemann H G, Oxman A D , Higgins JPT, Vist GU, Glasziou P and Guyatt GH on behalf of the Cochrane Applicability and Recommendations Methods Group and the Cochrane Statistical Methods Group.
- ⁴⁷ Singh S, Loke YK, Furberg CD Thiazolidinediones and heart failure: a teleo-analysis. *Diabetes Care*. 2007 Aug; 30(8):2148-53.
- ⁴⁸ Solomon DH, Rassen JA, Glynn RJ et al. The comparative safety of analgesics in older adults with arthritis. *Arch Intern Med*.2010; 170(22):1968-1978.
- ⁴⁹ Agency for Health Technology Assessment. Guidelines for conducting Health technology assessment (HTA) version 2.1. April 2009. Available and accessed in May 2011: http://www.aotm.gov.pl/assets/files/wytyczne_hta/2009/09.06.29_wytyczne_HTA_eng_MS.pdf

⁵⁰ Pharmaceutical Benefit Advisory Committee. Department of Health. Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). December 2008. Available and accessed in May 2011:
[http://www.health.gov.au/internet/main/publishing.nsf/content/AECB791C29482920CA25724400188EDB/\\$File/PBAC4.3.2.pdf](http://www.health.gov.au/internet/main/publishing.nsf/content/AECB791C29482920CA25724400188EDB/$File/PBAC4.3.2.pdf)

⁵¹ Medical Services Advisory Committee. *Funding for new medical technologies* and procedures: application and assessment guidelines. September 2005. Available and accessed in May 2011:
[http://www.health.gov.au/internet/msac/publishing.nsf/Content/D81BE529B98B3DB6CA2575AD0082FD1B/\\$File/guidelines.pdf](http://www.health.gov.au/internet/msac/publishing.nsf/Content/D81BE529B98B3DB6CA2575AD0082FD1B/$File/guidelines.pdf)

⁵² National Institute for Health and Clinical Excellence. Single Technology Appraisal. Specification for Manufacturer /sponsor submission of evidence. October 2009. Available and accessed in May 2011:
<http://www.nice.org.uk/aboutnice/howwework/devnicetech/singletechnologyappraisalsubmissiontemplates.jsp?domedia=1&mid=4D9D8C83-19B9-E0B5-D4B0E148B3FE727F>