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

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

<table>
<thead>
<tr>
<th>Recommendation</th>
<th>The recommendation is based on the following elements</th>
</tr>
</thead>
</table>
| **Recommendation 1**  
In relative safety assessment of health technologies main objectives of HTA assessors are summarised as follows:  
- 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 |
| **Recommendation 2**  
HTA assessors may focus their investigation on the following areas:  
- 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 |
| **Recommendation 3**  
The HTA assessors should use consistent and precise terminology to avoid misleading results. They should use the MedDRA Dictionary for describing adverse reactions. | 2.2 Terminology |
| **Recommendation 4**  
Main sources of information of HTA assessors are:  
- 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 |
**Recommendation 5**
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.
To assess the data on adverse reactions, how the adverse effects were collected and reported should be evaluated.

Useful questions to assess how the adverse reactions are collected:
- 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?

Useful questions to assess how the adverse effects are reported:
- 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?

**Recommendation 6**
Characteristics of selected studies should be summarised in a table. Useful information on studies characteristics are the following:
- 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.

Different tables should be elaborated for RCTs and observational studies.

<table>
<thead>
<tr>
<th>2.4 Evaluation of sources of information</th>
<th>2.5.1 Description of included sources of information</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>2.4 Evaluation of sources of information</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>2.5.1 Description of included sources of information</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>2.5.1 Description of included sources of information</th>
</tr>
</thead>
</table>
### Recommendation 7
Results from individual studies should be presented by group in tabular form, using the following measures:
- 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

Different tables should be elaborated for RCTs and observational studies.

Adverse effects should be grouped according to the System Organ Class (SOC).

Adverse effects which are common and serious should be reported separately.

If possible, adverse effects should also be provided by severity grade.

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.

### Recommendation 8
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.

A table is preferable for the comparison of the safety profile of the new health technology and the comparator(s).

HTA assessors should describe if there is a clinically significant difference in adverse reactions between the interventions compared.

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.

### Recommendation 9
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.

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

### 2.5.3 The comparison of the safety profile of the health technology to the comparator(s)

### 2.5.4 Balanced discussion of benefits and adverse effects
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.


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.


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.


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.


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.


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.

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”.


**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.


**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.


**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.

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

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.

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).

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”.

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.

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


**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**
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.

Serious Adverse Device Effect (SADE)
Adverse device effect that has resulted in any of the consequences characteristic of a serious adverse event.

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.


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.

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

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

**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.

**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.

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.

### 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. 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.

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)

- EUnetHTA guideline on Endpoints used in REA: Composite endpoints
- EUnetHTA guideline on levels of evidence: applicability of evidence in the context of relative effectiveness assessment
- EUnetHTA guideline on levels of evidence: internal validity (of randomized controlled trials)
- EUnetHTA guideline on comparators and comparisons: direct and indirect comparisons
- EUnetHTA guideline on therapeutic medical devices (JA2)
- EUnetHTA Guideline on internal validity of non-randomised studies (NRS) on interventions (JA2)
- EUnetHTA Guideline on process of information retrieval for systematic reviews and health technology assessment on clinical effectiveness (JA2)
- 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 patients.

2.2. Terminology

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

For this purpose the Medical Dictionary for Regulatory Activities (MedDRA), developed by the International Conference on Harmonisation, could be a useful instrument. 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 (HTs); high level group terms (HLGTs) and it is organised in 26 system organ classes (SOCs). 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.

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.
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.

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 medicines 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.

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. 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.
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). 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 Devices@FDA: www.accessdata.fda.gov/scripts/cdrh/devicesatfda/index.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. 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. 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. 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. 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. 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. 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% pharmaceuticals 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 webpage. 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. 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. 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. As a result, studies using different methods of monitoring effects are difficult to be compared. 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 Collaboration, 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.\textsuperscript{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.\textsuperscript{36}

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

- 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
- 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:

- 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). 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.

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.\textsuperscript{44} Different approaches are available, including PRISMA Statement (Preferred reporting for Systematic Reviews and Meta-Analyses) and Cochrane Collaboration’s approach.\textsuperscript{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

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

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

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

Source: Adapted from NICE. Single Technology Appraisal. Specification for Manufacturer /sponsor submission of evidence.
Table 3b. Adverse effects by seriousness

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

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

<table>
<thead>
<tr>
<th>Adverse Event</th>
<th>nNSAIDs HR (95% CI)</th>
<th>COXibs HR (95% CI)</th>
<th>Opioids HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composite cardiovascular&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1 [Reference]</td>
<td>1 [Reference]</td>
<td>1 [Reference]</td>
</tr>
<tr>
<td>Upper or lower GI tract bleeding</td>
<td>1.28 (1.01-1.62)</td>
<td>0.60 (0.35-1.00)</td>
<td>0.60 (0.35-1.00)</td>
</tr>
<tr>
<td>Composite fracture&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.00 (0.00-1.49)</td>
<td>0.47 (0.12-1.64)</td>
<td>4.47 (1.32-1.41)</td>
</tr>
<tr>
<td>Hospitalized adverse event</td>
<td>1.12 (0.91-1.38)</td>
<td>1.68 (1.37-2.07)</td>
<td>1.68 (1.37-2.07)</td>
</tr>
<tr>
<td>Death related to adverse event</td>
<td>1.02 (0.62-2.02)</td>
<td>1.98 (1.65-2.37)</td>
<td>1.98 (1.65-2.37)</td>
</tr>
<tr>
<td>All-cause mortality</td>
<td>1.16 (0.85-1.57)</td>
<td>1.87 (1.39-2.53)</td>
<td>1.87 (1.39-2.53)</td>
</tr>
</tbody>
</table>

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 (≥1/10), Common (≥1/100 and <1/10), Uncommon (≥1/1000 and <1/100), Rare (≥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

<table>
<thead>
<tr>
<th>System organ class</th>
<th>Frequency</th>
<th>Event*</th>
<th>Pemetrexed/cisplatin</th>
<th>Cisplatin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>All grades toxicity</td>
<td>Grade 3 - 4 toxicity</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(%)</td>
<td>(%)</td>
</tr>
<tr>
<td>Blood and lymphatic system disorders</td>
<td>Very common</td>
<td>Neutrophils/Granulocytes decreased</td>
<td>56.0</td>
<td>23.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Leukocytes decreased</td>
<td>53.0</td>
<td>14.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Haemoglobin decreased</td>
<td>26.2</td>
<td>4.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Platelets decreased</td>
<td>23.2</td>
<td>5.4</td>
</tr>
<tr>
<td>Metabolism and nutrition disorders</td>
<td>Common</td>
<td>Dehydration</td>
<td>6.5</td>
<td>4.2</td>
</tr>
<tr>
<td>Nervous system disorders</td>
<td>Very common</td>
<td>Neuropathy-Sensory</td>
<td>10.1</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>Common</td>
<td>Taste disturbance</td>
<td>7.7</td>
<td>0.0***</td>
</tr>
</tbody>
</table>

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.  

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.\textsuperscript{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.\textsuperscript{f}

Clinical trials are usually powered to detect statistical significance of possible benefits of a technology and only secondly designed to study safety.\textsuperscript{16} As a result the evidence on adverse reactions generated by RCTs may be inconclusive.\textsuperscript{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.\textsuperscript{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

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

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

*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)
Hauté Autorité de Santé (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:


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

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of studies addressing the outcomes</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Outcome</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Quality of the evidence</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Frequency (very common and common)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Severity</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>System Organ class</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>No. patients in the group</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>No. of patients with the event</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Absolute risk</td>
<td>✓</td>
<td>✓</td>
<td>✓ (per 1000 people) (CI 95%)</td>
<td>✓</td>
<td>✓ (CI 95%)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Hazard ratio</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Absolute risk reduction</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Relative risk</td>
<td>✓ (CI 95%)</td>
<td>✓</td>
<td>✓</td>
<td>✓ (CI 95%)</td>
<td>✓</td>
<td>✓ (CI 95%)</td>
<td>✓</td>
</tr>
<tr>
<td>Relative risk reduction</td>
<td>✓</td>
<td>✓</td>
<td>✓ (CI 95%)</td>
<td>✓</td>
<td>✓ (CI 95%)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Risk difference</td>
<td>✓</td>
<td>✓</td>
<td>✓ (CI 95%)</td>
<td>✓</td>
<td>✓ (CI 95%)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>NNH</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Odds ratio</td>
<td>✓ (CI 95%)</td>
<td>✓</td>
<td>✓ (CI 95%)</td>
<td>✓</td>
<td>✓ (CI 95%)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Mean difference</td>
<td>✓ (CI 95%)</td>
<td>✓</td>
<td>✓ (CI 95%)</td>
<td>✓</td>
<td>✓ (CI 95%)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Standardised mean difference</td>
<td>✓ (CI 95%)</td>
<td>✓</td>
<td>✓ (CI 95%)</td>
<td>✓</td>
<td>✓ (CI 95%)</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
**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**

- **Research in medline:**
  - Adverse drug reaction or (adverse event and drug therapy) or (adverse effect and drug therapy) or comparative effectiveness or comparative effectiveness research or drug safety or comparative safety.
  - Titles and abstracts identified and screened n = 4772

- Not available and excluded because clearly not relevant for the scope of the guideline n=4535

- Excluded because
  - Discussing only relative effectiveness
  - Not discussing methological issues concerning relative safety of interest of HTA assessors n = 211

- Full copies retrieved and assessed for eligibility n = 237

- Documents identified from reference list and web sites n=29

- Publications meeting inclusion criteria n=55
Annexe 4. Bibliography


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.


Golder S, Lohe 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


Vandenbroucke JP. Case reports of suspected adverse drug reactions. Case reports were dismissed too quickly. BMJ February 2006;332: 488.


Golder S, McIntosh HM, Duffy S, Glasnave J: Developing efficient search strategies to identify reports of adverse effects in MEDLINE and EMBASE. Health Info Libr J 2006; 23(1):3-12.


43 Aronson JK. Anecdotes as evidence. BMJ 2003; 326: 1346


