EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

1 April 2016
EMA/CHMP/292464/2014
Committee for Medicinal Products for Human Use (CHMP)

# Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man

## The use of patient-reported outcome (PRO) measures in oncology studies

| | |
|---|---|
| Draft agreed by Oncology Working Party | December 2013 |
| Adopted by CHMP for release for consultation | 22 May 2014 |
| Start of public consultation | 17 June 2014 |
| End of consultation (deadline for comments) | 30 November2014 |
| Agreed by Oncology Working Party | November 2015 |
| Adopted by CHMP | 1 April 2016 |
| Date for coming into effect | 1 November 2016 |

| | |
|---|---|
| **Keywords** | ***Patient-reported outcome (PRO), health-related quality of life (HRQL)*** |

30 Churchill Place ● Canary Wharf ● London E14 5EU ● United Kingdom
**Telephone** +44 (0)20 3660 6000 **Facsimile** +44 (0)20 3660 5555
**Send a question via our website** www.ema.europa.eu/contact

# Appendix 2 to the Guideline on the evaluation of anticancer medicinal products in man

The use of patient-reported outcome (PRO) measures in oncology studies

## Table of contents

# Executive summary

The importance of the patient's point of view on their health status is fully acknowledged and such information may be used in drawing regulatory conclusions regarding treatment effects, in the benefit risk balance assessment or as specific therapeutic claims in Section 5.1 of the SmPC (Secord et al., 2015, CHMP Reflection paper on the regulatory guidance for the use of Health-related quality of life (HRQL) measures in the evaluation of medicinal products, 2005). For specific therapeutic claims, valid measures should be selected based on their 'fit' with a hypothesis led strategy, resulting in unbiased outcomes.

This appendix on the use of patient-reported outcome (PRO) measures in patients with cancer focuses on the value of these data from a regulatory perspective. The possible add-on value from a licensure perspective of such data to conventional efficacy and safety data in benefit risk assessment is therefore emphasised.

The oncology working party held a workshop on health-related quality of life (HRQL) in 2012, bringing together relevant experts to help inform on the content of this appendix. (http://www.ema.europa.eu/ema/index.jsp?curl=pages/news_and_events/events/2012/04/event_detail_000558.jsp&mid=WC0b01ac058004d5c3). This document has been developed to outline broad principles of scientific best practice and to provide guidance on the value of PRO data in the development of medicinal products for the treatment of cancer, whilst recognizing that PRO methodology is developing and evolving (Calvert et al., 2014, Kyte et al., 2014, Secord et al., 2015).

# 1. Background

A PRO includes any outcome evaluated directly by the patient himself or herself and is based on patient's perception of a disease and its treatment(s). PRO is an umbrella term covering both single dimension and multi-dimension measures of symptoms, HRQL, health status, adherence to treatment and satisfaction with treatment. PRO measures (PROMs) are the tools and/or instruments that have been developed to ensure both a valid and reliable measurement of these PROs. Like any other clinical outcome assessments such as a rating of a symptom, sign or performance by an observer or trained medical care provider, it is recognised that such data have inherent variability related to the assessor. Health-related quality of life is a specific type of PRO and is a broad concept which can be defined as the patient's subjective perception of the impact of his/her disease and its treatment(s) on his/her daily life, physical, psychological and social functioning and well-being. The notion of multidimensionality is a key component of the definition of HRQL. In clinical research, PROs provide a unique means of capturing the personal and social context of the disease and treatment experience, as OS (Overall Survival), PFS (Progression Free Survival), biomarker measures or adverse events may not necessarily capture the full impact of a treatment on how a patient feels or functions.

Over the last decades, PRO objectives have frequently been incorporated in confirmatory oncology studies. Historically, longitudinal HRQL data have rarely been informative from a licensure perspective. To what extent this is related to absence of a "true" difference between treatment arms, poorly defined objectives, poor validity, reliability, and responsiveness of the instruments, absence of off-study therapy data, high attrition rates and informative missing data, or simply reflects the resilience and dynamics of the individual's adaptive perception of HRQL during the course of disease, remains unknown (Sprangers & Schwartz et al., 1999). In addition, poorly defined PRO objectives and lack of a priori specification of the expected effect (e.g., improvement, maintenance) have further hampered the usefulness of PROs in licensure decisions. However, it is acknowledged that the lack of a HRQL difference between treatment arms should not be seen, per se, as a factor limiting the use of high

quality HRQL data, provided the selected methodology is sensitive enough to demonstrate differences if they exist.

More recently, time to deterioration in tumour related symptoms, as measured by PRO instruments, have been introduced and here differences have been demonstrated e.g. time to pain progression (Gravanis et al., 2013). Time to symptom deterioration places an emphasis on the patient's perspective, which is complementary to PFS, and hence may be of value for the demonstration of treatment benefit. However, these PRO relate to disease specific measures and therefore do not provide estimates of longitudinal HRQL, which may also be relevant, i.e. do not provide the patient's general perception of the effect of illness and treatment on physical, psychological, and social aspects of life. Therefore it is of importance to select the most appropriate instruments, in line with the study's objectives and the characteristics of the patient population.

In most cases, for a particular tumour type and disease stage, there is no reason to assume that the potential benefit of a delay in tumour progression, if of similar magnitude, is product specific. However, the tolerability and toxicity profiles may differ considerably between medicinal products. The differential impact on patient well-being is harder to estimate from conventional adverse event reporting, even though withdrawal rates prior to tumour progression may provide some insights. In relation to active compound comparative trials and from a licensure perspective, PRO data derived from instruments capturing the impact of adverse reactions on patient well-being, in an unbiased way and in relation to the study drugs, are welcomed.

There are a number of validated PRO instruments, including EORTC and FACT measures that aim to capture the consequences of adverse reactions on patient wellbeing. At the time of this appendix, there is no EMA/CHMP experience from the use of the NCI's PRO version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE) and more experience is needed before these tools can be used reliably. However, it is envisioned that the PRO-CTCAE could enhance the precision and patient centeredness of adverse event reporting in cancer clinical research and ultimately provide a more representative account of patients' treatment experiences (Basch et la., 2014, Dueck et al., 2015).

In summary, PROs can provide important patient perspectives on the disease and the treatment received; an evaluation that provides clinically important information that is not captured by conventional anti-tumour efficacy data and adverse event reporting (Basch, 2010). There are, however, methodological obstacles that historically have reduced the impact of PRO data on regulatory decisions e.g. bias, missing data, quality of data, timing of assessments, only single-dimensional PRO measure reporting, and lack of post-progression data. Key is careful planning and an in depth analysis of whether the inclusion of PRO measures is likely to make a potential difference to the study conclusions. In this regard, it is important to prioritise endpoints meaningful to patients, to select outcome strategies and to elucidate the relationships between PROs and other endpoints (Basch, 2013).

# 2. Scope

This appendix covers general aspects of the use of PRO endpoints in oncology studies such as the designing and carrying out of clinical studies, selecting instruments and the added value. By outlining broad principles of scientific best practice rather than prescribing a particular approach to PRO selection and application, the appendix aims to encourage developments in the methods and application of PROs in the oncology regulatory setting. This appendix does not cover the validation of instruments nor does it make specific recommendations regarding the instrument to select.

# 3. Legal basis and relevant guidelines

This document should be read in conjunction with Directive 2001/83/EC, as amended and Regulation 726/2004. In addition, relevant CHMP guidelines should be taken into account. These include but are not limited to:

- Guideline on the evaluation of anticancer medicinal products in man - EMA/CHMP/205/95/Rev.4
- Statistical principles for clinical trials – CPMP/ICH/363/96 (ICH E9)
- Reflection paper on the regulatory guidance for the use of HRQL measures in the evaluation of medicinal products - EMEA/CHMP/EWP/139391/2004
- Guideline on missing data in confirmatory clinical trials - EMA/CPMP/EWP/1776/99 Rev. 1
- Points to consider on multiplicity issues in clinical trials - CPMP/EWP/908/99

# 4. Patient-reported outcomes

A PRO includes any outcome evaluated directly by the patient himself or herself. A PRO can be measured by self-report, generally in the form of a questionnaire, or by interview, provided that the interviewer records only the patient's response. PRO measures must have acceptable responsiveness, reliability and validity, and may include reference to symptoms, functional status, treatment adherence or satisfaction with care. In clinical research, the use of a PRO measure is advised when measuring a concept best known to the patient or best measured from the patient perspective.

Oncology clinical studies to support regulatory submissions may include PRO measures as secondary or exploratory outcomes and rarely as primary outcomes. Examples as described in the CHMP 2012 Guideline on the evaluation of anticancer medicinal products in man (EMA/CHMP/205/95/rev.4) include:

- Section 7.1.5: In patients with tumour-related symptoms at base line, symptom control, if related to anti-tumour effects, is a valid measure of therapeutic activity and may serve as primary endpoint in late line therapy studies, provided that sources of possible bias can be minimised. In certain cases, time to symptomatic tumour progression may also be an adequate primary measure of patient benefit;
- Section 7.1.5.1: Especially in the palliative setting, HRQL/PRO using generally accepted instruments might be informative;
- Section 7.4. In a study conducted with BSC as reference therapy, the objective should be to demonstrate prolonged OS and/or globally improved symptom control or HRQL. The latter requires that all efforts are undertaken to reduce possible bias.

Reasons to include PRO assessment in the clinical development programme for oncology medicinal products may encompass:

- Provide a patient focused assessment of the burden and impact of disease, by understanding how a treatment impacts on patient functioning and well-being;
- Add information on the clinical benefit of a therapy by complementing efficacy and safety data with patient-reported evaluation;
- Assess the relationship/ agreement between clinical reported endpoints and patient-reported endpoints, e.g. to better understand the impact of an objective clinical response from a patient perspective;

- Attempt to differentiate two treatments in the non-inferiority trial setting, where the primary endpoint is an objective measure;
- Provide information to facilitate more accurate future patient-physician communication in terms of the quality of the survival time remaining for the patient and the burden of treatment-related morbidities and disease-related patient impacts by detailing a more complete evaluation of cancer treatment.

The general recommendations for the incorporation of PRO measures in clinical development include:

- An assessment or rationale (e.g. supported by a conceptual model) for the extent to which the inclusion of PRO measures can provide added value in the clinical trial setting;
- Consider whether the collection of PRO data can detect meaningful effects and make a difference to the study conclusions and regulatory benefit risk balance assessment;
- While it is expected that in the majority of cases, the collection of PRO data will be most informative in the confirmatory trial setting, on a case by case basis, the collection of PRO data within the context of exploratory studies could be informative;
- PROM should be considered early in the development programme, particularly if there is a need to develop a dedicated instrument;
- Consideration should be given to patient involvement in the study design process and in the evaluation of study feasibility;
- When used, PRO endpoints should be incorporated into the protocol development at the earliest stage and should be explicitly stated as a specific clinical trial objective or hypothesis in the study protocol and statistical analysis plan;
- PRO data reporting should be adequately performed, e.g. informed by the CONSORT PRO Extension;
- For specific therapeutic claims in Section 5.1 of the SmPC, a clear hypothesis lead strategy unbiased in relation to reference therapy is strongly recommended and measures should be selected based on this scientific rationale;
- PRO measures should be administered to study subjects at time points when there is a clear and hypothesis driven rationale for their use and when it is feasible to expect high levels of completion by the individual patient;
- Where feasible, PRO measures should be administered at the beginning of a clinic visit prior to medical interviews or procedures, in the event that medical information or immediate toxicities from chemotherapy could bias retrospective evaluation;
- PRO data should be treated with the same importance as other data in monitoring clinical site performance and collection methods, whilst acknowledging that more rapid follow up for missing data may be required due to the recall period (retrospective PRO data capture, unlike some clinical outcomes e.g. overall survival, may not be possible);
- PRO collection may give rise to PRO Alerts, medically concerning levels of psychological distress or physical symptoms that may require an immediate response. Where applicable, an a priori plan for the management of such alerts could be included in the protocol and communicated to trial staff.

### 4.1. Symptom PRO measures

Patients provide the primary perspective of treatment effectiveness, particularly as it has been reported in the literature that clinicians miss or underestimate a large proportion of the symptomatic adverse events experienced by patients (Basch, 2014). Measuring symptoms is important in understanding the burden of cancer, particularly as uncontrolled symptoms have been linked to increased mortality in patients with cancer (Khan SA et al., 2014). It is important to include patient-reported symptoms that are appropriate to the study population, intervention, objectives and setting (Basch et al., 2012). If symptom PRO measures are used to evaluate the impact on specific symptoms it should be ensured that a benefit in respect to these symptoms is not accompanied by a negative effect on other symptoms. One way of doing this could be to include a sensitive multidimensional HRQL instrument.

Symptom response rates and symptom control are particularly significant in the palliative setting. Assessment of palliation can be assessed by changes in symptom scores in general or change in symptom scores considering only certain prespecified symptoms. Symptoms (related to the disease, toxicity or multi-factorial) that are commonly found in the advanced cancers include anorexia, anxiety, constipation, depression, diarrhoea, dyspnoea, fatigue, insomnia, nausea, pain, neuropathy and vomiting (Basch et al., 2012). However determining which patient-reported symptoms should be investigated must be evidence based and derived from feedback from patients and carers, clinicians and other experts, as well as the literature (Basch et al., 2012).

As important as it is to select an instrument that properly captures disease-related symptoms, it is also important to use an instrument that captures side effects of therapy in an unbiased way.

### 4.2. Health Related Quality of Life (HRQL)

The impact of treatment and disease on patients can be measured using self-reported questionnaires. HRQL instruments attempt to measure complex aspects of life which are modified by the disease and therapeutic interventions. HRQL, and its resilience during disease and treatment, is a personal perspective and may vary with gender, experience, age, education, disease stage, and cultural background.

As with all PROs, the inclusion of HRQL assessment in clinical trials should have a strong scientific rationale and researchers should utilise existing validated instruments where available. HRQL complements the range of traditional indicators and the data can provide information regarding both positive and negative patient experiences (see also CHMP Reflection paper on the regulatory guidance for the use of HRQL measures in the evaluation of medicinal products, EMEA/CHMP/EWP/139391/2004).

## 5. Clinical trial design

### 5.1. General principles

There is no standard approach to collecting, analysing or interpreting PRO data in clinical trials. In common to other aspects of clinical trial design, good science applies and objectives need to be justified alongside realistic expectations. Careful thought must go into designing and implementing PRO measures in the oncology clinical trial setting in order to investigate a well-formulated predefined hypothesis. As with other endpoints of the trial, the design of the trial, choice and analysis of PROs should be justified from the question of interest. In order to ascertain meaningful information about the patient experience, a dedicated a priori power calculation for the PROs of greatest interest is

recommended where possible (Basch et al., 2012, see section 'Data collection and preventing avoidable missing data'). In the majority of circumstances, the patient is the best informant and the most appropriate way to measure PRO is self-reporting direct from the patient (see also section on carer/ proxy input and Observer-Reported Outcomes).

Importantly, measurements should not constitute an undue burden to the patient and in this context PRO experts suggest limiting estimated completion time of baseline PRO assessments to 20 minutes and 10-15 minutes completion for subsequent assessments (Basch et al., 2012). Patient burden may be reduced by the use of the right instrument at the right time; the concept measured by the instrument should be relevant to the patient, i.e. reflecting their experience at the time of completion. In ensuring patient well-being, consideration should be given to a plan for consistent and standardised management of 'PRO alerts' (e.g. medical response to the recording of extreme scores on questionnaires) and the subsequent opportunity to collect data on related *ad hoc* co-intervention (Kyte et al., 2013, Kyte et al., 2016).

Blinding patients to the treatment they have received in a controlled trial is a crucial method for reducing bias in randomised controlled trials. Non-blinded patients may report symptoms differently to blinded patients and mechanisms of unblinding involve perceptible differences in the compared treatments (Hrobjartsson, 2011). However, it is not always possible to ensure blinding of patients and treatment providers, and the priority should be capturing the patient experience through PRO data in all types of clinical trial designs. Each trial design will present different challenges to the validity and generalisability of the results which requires *a priori* planning of the study objectives and analysis of the PRO, and its context with the other clinical endpoints. Generalisability should be framed within both the efficacy and the associated toxicities and statements around potential bias, with justification of the criteria for reporting a PRO treatment benefit (see section 6).

Whilst the concern in relation to bias in open label studies remains, it might well be that data of clinical interest *a priori* can be produced only under open label randomised controlled trial conditions. In these circumstances, PRO data should also be supported by objective measures. One example being an experimental compound assumed to be more efficacious, but also more toxic or less well tolerated. Under these circumstances extensive planning in advance is required to increase the credibility of study data such as multidimensional reporting. For example, effects of neuropathy on functionality could be supported by conventional clinical measures of neuropathy.

It is of major importance to discuss in detail in the study protocol why certain timings of assessments were selected and why the selected instrument is unbiased in relation to the toxicity/tolerability profiles of study drugs (the instrument needs to be sensitive enough to capture a range of potential effects, including possible unanticipated effects).

## Frequency and duration of assessments

Timing and frequency of PRO assessment are key issues and frequency can greatly influence the scores received. The overall frequency of assessment depends on;

- The natural history of the disease;
- The hypothesis being tested;
- The method of data analysis;
- The nature of the investigative treatment including dosing frequency and how long after treatment the drug is expected to show changes in symptoms, impacts or side effects;
- The recall period for the instrument;

- The acceptable level of patient completion burden.

It is generally recommended to determine when expected changes in symptoms and or side effects are likely to occur over time and data collection should cover the clinically most important periods. If the timing of potential effects is poorly understood, PRO should be considered in earlier stage exploratory trials. The duration of assessment depends on the research questions being asked, but it is important to ensure that the duration of the clinical study and follow up is of adequate length to robustly support any planned analysis, including reversibility of adverse reactions.

In order to be able to accurately assess the PRO results on study therapy, continued assessment post-progression and during next-line therapy may also be informative. Such next-line PRO data allows contextualisation of the results observed on study treatment, which can be of particular importance in the palliative or maintenance setting, and could be supportive when therapeutic claims (section 5.1 of the SmPC) are intended.

Apart from the need for contextualisation, there is also a methodological rationale for collecting next-line or post-progression data when PROs are studied. Patients in the comparator arm are normally expected (as a group) to experience progression earlier than the patients in the experimental arm. Thus, if PRO assessments are stopped at progression, patients in the comparator arm will automatically have a shorter observation period compared with those in the experimental arm. This can be regarded as a form of informative missing data, affecting the possibilities to draw conclusions from the PRO data. Therefore, it is important to continue to collect PRO data after progression has been reached for comparisons of the trends over time (see also section on data collection and preventing avoidable missing data).

In general, the duration of assessment should be limited to a time period that is both feasible and interpretable. The assessment schedule should be terminated at a point when the results would no longer be interpretable either due to low compliance or because the next line of therapy is highly heterogeneous.

## Data collection and preventing avoidable missing data

High compliance has been attributed among other factors to comprehensive educational programmes prior to and during the trial for both research staff and study participants (Hansen LK et al., 2014). In general, assessments should be performed on schedule irrespective of whether study treatment has been given. Collecting PRO data from patients with advanced and progressive disease may be more difficult because of failing health and / or cognitive challenges. PRO data can be collected by administering PRO instruments through different modes – interviewing, telephone, mailing or self-administration. Variations in the mode of administration (how the self-report questionnaire is provided and when) can be a source of bias. The potential measurement bias introduced by mixing different modes could be justified if mixing modes can mitigate the more important issue of bias due to informative missing data and evidence of equivalence across different modes is available.

Electronic data capture methods may offer more convenience to some patients and may increase data quality, reduce missing data (allowing automatic reminders to be sent) and potentially reduce data entry errors (Basch et al., 2012, Coons et al., 2009). Adaptation of case report forms to electronic forms, including electronic modes of PRO administration, must ensure that the data collected via the different methods are equivalent or take account of any identified differences (Coons et al 2009). Where substantive changes have occurred, it is necessary to confirm that the adoption of electronic PRO format did not introduce significant response bias and that the different modes of administration produce essentially equivalent results (Coons et al., 2009). Of importance, whilst use of electronic data recording might be of benefit in some patient groups, alternatives should be made available, e.g. for elderly patients so that differential loss of data is minimised.

Incorporating PRO instruments as clinical trial endpoint measures introduces challenges in the analysis of clinical trial data, particularly because of their multi-dimensional nature and missing values. High attrition rates may occur if the instrument is burdensome to patients, especially as ability declines during the late stages of disease. As described above, informative missing data may also occur if PRO measurement is stopped upon clinical disease progression. The study protocol should describe the principal data analysis features in the statistical section, with a detailed elaboration of the analysis in the Statistical Analysis Plan (SAP), including how to control for multiplicity (for more guidance see CPMP/EWP/908/99). The clinical trial protocol should also describe how missing assessments will be handled in the analysis (e.g. use of imputation techniques, sensitivity analysis), clarify which question is answered by the strategy chosen to handle the missing data and justify why this question is relevant and which assumptions are made (e.g. an analysis under the missing-at-random assumption will typically address the question what the effect would have been if all patients had adhered, under the assumption that those that did adhere in the trial are representative for those that could adhere). The SAP should include a comprehensive specification of the scoring algorithms of PRO instruments used, including management of missing items in multi-item scores. Missing data should be put into the context of underlying reason, when known, and the reasons for missingness could inform which missing data strategies are (more) reasonable. Missing completely at random or missing-at-random is hardly ever a justified assumption and prevention of avoidable missing data is preferable to post-hoc strategies to handle the missing data (Bernhard et al., 1998, for more guidance, see EMA/CPMP/EWP/1776/99 Rev. 1). The SAP should also include a comprehensive specification of the scoring algorithms of PRO instruments used, including management of missing items in multi-item scores.

A high compliance rate is expected in order to provide a substantial amount of interpretable longitudinal data. It is therefore essential to minimise data loss and to employ strategies to increase patient compliance, such as;

- Completion of baseline PRO assessment included as part of the eligibility criteria checklist when ethically justifiable and practically feasible, and prior to randomization and administration of treatment;
- Appoint a PRO trained and qualified person responsible for PRO data collection in each study site;
- Comprehensive education and training of research staff, including investigators and the whole clinical research team, in order to ensure they understand the importance of PRO assessment and will be able to motivate their patients to complete the PRO instruments;
- Education and training of patients before completion of the questionnaire, including that there is no incorrect answer and explaining the purpose of the assessment. If such education and training is given, it should be carried out by means that do not influence the patient in their answers to the questionnaires themselves and it should be equally available to all patients participating in the project;
- Explore the use of automated electronic data collection and employing a reminder system;
- Checking for completeness of forms for omissions at the point of assessment or very soon after to minimise recall bias, and sensitively clarifying and documenting reasons for non-completion.

## *5.2. Instruments*

An instrument can be described as a means to capture data, such as a questionnaire, plus all the information and documentation that supports its use. Generally, this includes clearly defined methods and instructions for administration or responding, a standard format for data collection, and well-

documented methods for scoring, analysis and interpretation of results in the target patient population. Disease specific measures may be more relevant to patients, providing a more in-depth relevant analysis. However, they may fail to capture unexpected changes due to novel treatment toxicities. Generic measures are useful for comparisons across treatments but may be less sensitive to change and the relative importance of the different PRO domains needs to be determined a priori. It is therefore recommended to carefully select the most appropriate instrument (generic or specific or a combination of both), in line with the study's objective(s) and the characteristics of the patient population. PRO instruments should be relevant, reliable, have demonstrated content validity and responsiveness to change. For responsiveness, it is necessary to demonstrate that the PRO scores are sensitive to actual changes in health status (Revicki et al., 2008). A PROM previous 'track record' may also offer a practical indication of its properties (Luckett T et al., 2010). PROMs should be acceptable to the population in which they will be administered, both in terms of the questions they ask (e.g. are they appropriately worded?) and the overall burden to the patient (e.g. is the completion time for the PRO measure agreeable?). PRO instruments must also be easily interpretable, i.e. the meaning of differences in PRO score should be clearly understood (Mokkink L et al., 2010).

Depending on the chosen instruments, lack of linguistic and cultural validation of instruments may be problematic in multinational and global studies and consideration should be given to processes for selecting the number of languages required to adequately cover global trial populations and translation procedures (Wild et al., 2009, Wild et al., 2005, Dewolf et al. 2009). Investigation of PRO endpoints in a predefined sufficiently large subgroup, including patient representative for the EU target population, may be considered as feasible and appropriate to avoid issues relative to the availability of language versions of the instrument. However, proper linguistic validation should have been performed to optimize cross-cultural validity of the different language versions of the instruments used and the impact of culture on the assessment of PROs in the study may be explored to confirm homogeneity of measurement in the study.

## Selection of an instrument

It is beyond the scope of this appendix to make specific recommendations regarding valid instrument selection, but in general, the instrument should be evidence based, shown to measure the concept it is intended to measure, be appropriate for the research objective, the disease and patient population characteristics and the practical considerations (respondent burden, feasibility). In addition, the most appropriate and valid PRO measures have involved patients in their development (Staniszewska et al., 2012). Online resources can be informative regarding the range of available PRO measures (Luckett, 2010). Guidance on the principles for selecting PROMs for cancer research can be found in the paper by Luckett and King, 2010.

## Carer / proxy input and Observer-Reported Outcomes (ObsROs)

A proxy is a person who reports an outcome as if she/he was the patient him/herself. There can be discordance between patient-reported outcomes and outcomes that are reported by a proxy such as by a carer (e.g., rating of pain). However, it has been reported that substantial differences between patient and proxy ratings may occur in a minority of cases and proxy ratings tend to be in greater agreement with those of patients of physical HRQL domains compared to psychosocial domains (Sneeuw et al., 2002).

The evaluation of PROM by carers or other non-clinical caregivers may be utilised where it is clear that the patient themselves cannot contribute (e.g. very young children, patients with cognitive impairment, severe ill health). However, in general proxy reporting should be avoided, unless the use of such proxy raters may be the only effective means of obtaining information that might otherwise be

lost. The clinical study protocol must define clear rules with regards to whether and when patient-reported data might be replaced by proxy reported data and provide appropriate justifications.

An observer-reported outcome (ObsRO) is a measurement based on an observation by someone other than the patient or a health professional. This may be a parent, spouse, or other non-clinical caregiver who is in a position to regularly observe and report on a specific aspect of the patient's health. An ObsRO measure does not include medical judgment or interpretation, and observer reports include only events or behaviours that can be observed. ObsRO should only be used, where it is considered that the patient themselves cannot contribute.

## *5.3. Special patient populations*

### Paediatric

Children's daily activities and experiences differ substantially from those of adults and adult PRO measures may not be appropriate for use in paediatric populations, either due to content validity or differences in the measurement process itself. A successful paediatric instrument must adjust for age and take into account the rate and pattern of change that children experience over time (Connolly et al., 1999).

Recommendations for paediatric PRO instruments in research have been published and are considered to be a useful basis for the approach in children and adolescence (Matza et al., 2013). Specific issues to consider are development stage (maturation may also differ because of disease and or experiences) and meaning of self. As with adult patients, the best informants are the patients themselves and it is important to collect as much information directly from the child wherever possible, using creative and age related approaches e.g. the use of pictures instead of words can be used for children too young to read (Connolly et al., 1999). However it is acknowledged that some children will be too young or too sick to contribute to the data collection and parents or caregivers may be able to provide data in situations where the child is unable to provide it directly. These circumstances need to be carefully considered and the differences acknowledged (Eiser & Morse, 2001).

### Adolescents and young adults

Evidence suggests that adolescents and young adults affected by, or recovering from cancer, are at increased risk of poorer psychological functioning compared to their peers, or to children and or/adults with cancer. Adolescents and young adults affected by cancer are developmentally unique (less likely to comply with treatment, more likely to engage in risk taking behaviour, place a higher importance on their peer relationships), meaning that their needs and experiences may not be fully captured by existing instruments developed for children or adults. Adherence to medication in adolescents with cancer is particularly challenging because of the nature of the adolescent cancer and its treatment (Hullmann et al., 2015). Consideration should be given to using age appropriate instruments that are relevant and appealing to young people, and are sensitive to change as the cancer progresses through treatment into survivorship or palliative care (Wakefield et al., 2013).

### Elderly

Elderly patients present particular characteristics and instruments should be relevant to the special requirements of older patients wherever possible. Appropriate data collection methods, such as Interactive Voice Record System (IVRS), provision of grip-pens or interviewer administered PRO measures should be considered for this group of patients to reduce their burden.

In elderly patients, concomitant diseases are more frequent, affecting psychological status and general performance. It is important to consider that HRQL is affected by comorbidities, multiple medications

(polypharmacy), functional status, ability to carry out activities of daily living, mental status (depression, cognitive functioning) and social support.

### Palliative setting

Palliative care mainly refers to last line settings where the prognosis for survival is poor. Patients place a high priority on having good symptom control (Khan SA et al., 2014). Successful patient palliation has been described as disappearance or improvement of symptoms, improvement of a specific symptom from baseline, change in the severity of a specific target symptom, for example pain or composite outcomes of pain and analgesic requirements, a symptom difference perceived as beneficial by the patient, HRQL score changes or increased duration of survival (see also Section 7.4, Main Guideline).

In patients with advanced cancer where the aims are palliative, the focus of care is promoting and preserving the remaining quality of life and in confirmatory studies the primary outcome measure could be globally improved symptom control or HRQL, if related to anti-tumour activity. This aspect should be carefully considered in the clinical study design, including choice of appropriate PRO measures, collection of longitudinal HRQL data and selection of primary endpoint, in particular as complicated multidimensional changes can occur relatively quickly and patient survival time is relatively short. Given that patients might be severely ill or might have cognitive impairment in certain types of cancers, the use of proxy reporting or ObsRO might be considered.

### Patients with rare cancers

Assessment of the patient experience in the rare disease setting can be achieved when careful planning and rigorous methods are employed. However, 'off-the-shelf' questionnaires may not be appropriate or specific enough to measure important outcomes in some very rare diseases, and there may be a need to develop additional PRO tools (Basch and Bennett, 2014).

## 6. Clinical importance and added-value

PRO information can enhance decision making by providing a better understanding of the potential impact of both the disease and treatment on a patient (Au et al., 2010). PRO instruments and assessments should be capable of detecting clinically meaningful effects and can provide added value when they are actionable, by being instrumental in interpreting a study's findings and would be expected to influence clinical recommendations (Au et al., 2010). Added value may also be derived if patients and clinicians have a more complete picture of the expected impact of a treatment on the patients' perception of adverse reactions and on disease related symptoms, aiding decision making (Au et al., 2010).

Treatment benefit can be demonstrated by either an effectiveness or safety advantage. For example, the treatment effect may be measured as an improvement or delay in the development of symptoms or as a reduction or delay in treatment-related toxicity. Because responsiveness and change depend on the patient population and contextual characteristics, there is not necessarily a single value of change of relevance for a PRO instrument across all applications and patient samples (Revicki et al., 2008). However, treatment benefit and risk, and magnitude of relevance of change should be based primarily on relevant patient-based and clinical anchors (Revicki et al., 2008, Wyrwich et al, 2013). For outcome measures other than "time to", mean difference may be used for statistical testing, but justified definitions of "response" and "deterioration" in the individual patient and in relation to predefined measures should be used to facilitate the interpretation of PRO data.

Situations where PRO measures, including HRQL, could potentially be of added value in terms of affecting the benefit risk assessment in the context of a marketing authorisation application, include

late line palliative care, maintenance therapy, and in studies comparing agents with similar efficacy but different safety profiles. In some disease settings i.e. in patients with tumour-related symptoms at base line, symptom response or delay in symptom progression, if related to anti-tumour effects, is a valid measure of therapeutic activity and may serve as a valuable endpoint in late line therapy studies. Furthermore, time to symptomatic tumour progression (or time to relevant deterioration) may also be an adequate measure of patient benefit and should be supported by ORR and/or PFS.

Criteria used to assess the potential added value of PRO data include:

- The relevance, reliability, validity and responsiveness of the instrument/ assessment used in the study;
- The appropriateness of the frequency and duration of data collection and analysis methods, in light of the patient population, disease setting and treatment regimen;
- The adequacy of the study design including the hypothesis and methods for appropriate handling of multiple outcomes and missing data in the statistical analysis, and appropriate standards of reporting PRO data;
- The rationale for the anticipated magnitude of effect based on clinically relevant anchors – the proposed magnitude of relevant treatment benefit should be identified a priori in the statistical analysis plan;
- Statistical significance *per se* is not sufficient for specific claims without demonstration of a clinically relevant treatment benefit;
- Considerations of alternative explanations that may account for the observed changes or lack of change.

# 7. Glossary and abbreviations

**Health-related Quality of Life (HRQL):** In the context of drug approval, HRQL is considered to represent a specific type/subset of PROs, distinguished by its multi-dimensionality. Indeed, HRQL is a broad concept which can be defined as the patient's subjective perception of the impact of his disease and its treatment(s) on his daily life, physical, psychological and social functioning and well-being (Source: CHMP Reflection paper on the regulatory guidance for the use of HRQL measures in the evaluation of medicinal products - EMEA/CHMP/EWP/139391/2004 )

**Instrument** — A means to capture data (i.e., a questionnaire) plus all the information and documentation that supports its use. Generally, that includes clearly defined methods and instructions for administration or responding, a standard format for data collection, and well-documented methods for scoring, analysis, and interpretation of results in the target patient population (Source: FDA Guidance for industry; December 2009: Patient-Reported Outcome measures: use in medical product development to support labelling claims).

**Observer-reported outcome (ObsRO)** — An ObsRO is a measurement based on an observation by someone other than the patient or a health professional. This may be a parent, spouse, or other non-clinical caregiver who is in a position to regularly observe and report on a specific aspect of the patient's health. An ObsRO measure does not include medical judgment or interpretation. Generally, ObsROs are reported by a parent, caregiver, or someone who observes the patient in daily life. For patients who cannot respond for themselves (e.g., infants or cognitively impaired), we encourage observer reports that include only those events or behaviours that can be observed. As an example, observers cannot validly report an infant's pain intensity (a symptom) but can report infant behaviour thought to be caused by pain (e.g., crying)

(Source: http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/ucm370262.htm)

**Patient-reported outcome (PRO):** Any outcome evaluated directly by the patient himself and based on patient's perception of a disease and its treatment(s) is called patient-reported outcome (PRO). The term PRO is proposed as an umbrella term to cover both single dimension and multi-dimension measures of symptoms, health-related quality of life (HRQL), health status, adherence to treatment, satisfaction with treatment, etc. (Source: CHMP Reflection paper on the regulatory guidance for the use of HRQL measures in the evaluation of medicinal products - EMEA/CHMP/EWP/139391/2004)

**Proxy-reported outcome:** A measurement based on a report by someone other than the patient reporting as if he or she is the patient. A proxy-reported outcome is not a PRO. A proxy report also is different from an observer report where the observer (e.g., clinician or caregiver), in addition to reporting his or her observation, may interpret or give an opinion based on the observation. (Source: FDA Guidance for industry; December 2009: Patient-Reported Outcome measures: use in medical product development to support labelling claims).

**Treatment benefit:** The effect of treatment on how a patient survives, feels, or functions. Treatment benefit can be demonstrated by either an effectiveness or safety advantage. For example, the treatment effect may be measured as an improvement or delay in the development of symptoms or as a reduction or delay in treatment-related toxicity (Source: FDA Guidance for industry; December 2009: Patient-Reported Outcome measures: use in medical product development to support labelling claims).

# 8. References


Au, HJ, Ringash, J, Brundage, M, et al. 2010. Added value of healthrelated quality of life measurement in cancer clinical trials; the experience of the NCIC CTG. Expert Reviews in Pharmacoeconomics and Outcomes Research. 2010 Apr;10(2): 119-28

Basch E et al. 2014. Development of the National Cancer Institute's patient-reported outcomes version of the common terminology criteria for adverse events (PROCTCAE). J Natl Cancer Inst. 2014 Sep 29;106(9). pii: dju244. doi: 10.1093/jnci/dju244. Print 2014 Sep

Basch E, et al. 2012. Recommendations for Incorporating Patient-Reported Outcomes Into Clinical Comparative Effectiveness Research in Adult Oncology. Journal of Clinical Oncology. 2012;30(34):4249-55

Basch E., 2013. Towards patient-centred drug development in oncology. NEJM. Vol 369. No5; 397-400

Basch E. 2014. New frontiers in patient-reported outcomes: adverse event reporting, comparative effectiveness, and quality assessment. Annual Reviews of Medicine. 2014;65:307-17. doi: 10.1146/annurev-med-010713-141500.

Basch E. 2010. The Missing Voice of Patients in Drug-Safety Reporting, New England journal of Medicine (2010) p865-869

Basch and Bennett., 2014. Patient-Reported Outcomes in Clinical Trials of Rare Diseases. J Gen Intern Med 29(Suppl 3):S801–3

Bernhard J, Cella DF, Coates AS, Fallowfield L, Ganz PA, Moinpour CM, et al. 1998. Missing quality of life data in cancer clinical trials: serious problems and challenges. Statistics in Medicine. 1998;17(5-7):517-32.

Brundage M, Bass B, Davidson J, Queenan J, Bezjak A, Ringash J, Wilkinson A, Feldman-Stewart D. 2011. Patterns of reporting health-related quality of life outcomes in randomized clinical trials: implications for clinicians and quality of life researchers. Qual Life Res. 2011; 20:653-64.

Brundage M, Blazeby J, Revicki D, Bass B, de Vet H, Duffy H, Efficace F, King M, Lam CL, Moher D, Scott J, Sloan J, Snyder C, Yount S, Calvert M. 2013. Patient-reported outcomes in randomized clinical trials: development of ISOQOL reporting standards. Quality of Life Research. 2013 Aug;22(6):1161-75. doi: 10.1007/s11136-012-0252-1. Epub 2012 Sep 18

Brundage et al., 2008. NCIC Clinical Trials Group experience of employing patient-reported outcomes in clinical trials: an illustrative study in a palliative setting. *Exert Rev. Pharmacoeconomics Outcomes Res.* 8(3):243-253

Calvert M et al. 2014. Patient-Reported Outcome (PRO) Assessment in Clinical Trials: A Systematic Review of Guidance for Trial Protocol Writers. PLoS One. 2014 Oct 15;9(10):e110216

Calvert M, Blazeby J, Altman DG, Revicki DA, Moher D,Brundage MD; 2013. CONSORT PRO Group. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. JAMA. 2013 Feb 27;309(8):814-22. doi: 10.1001/jama.2013.879

Cocks K, King MT, Velikova G, et al, 2011. On behalf of the EBIG collaborative group. Evidence-based guidelines for determination of sample size and interpretation of the European Organisation for the Research and Treatment of Cancer quality of life questionnaire core 30 (EORTC QLQ-C30). Journal of Clinical Oncology. 2011; 29 (1): 89-96

Connolly MA, Johnson JA. 1999. Measuring quality of life in paediatric patients. Pharmacoeconomics. 1999;16:605–25

Coons , Gwaltney , Hays , et al; 2009. ISPOR ePRO Task Force. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. Value in Health. 2009 Jun;12(4):419-29.doi: 10.1111/j.1524-4733.2008.00470.x. Epub 2008 Nov 11

CMTPNET, Recommendations for Incorporating Patient-Reported Outcomes into the Design of Clinical Trials in Adult Oncology. http://www.cmtpnet.org/docs/resources/PRO_EGD.pdf

Dewolf, L, et al. 2009. EORTC Quality of Life Group Translation Procedure, 3rd Ed. EORTC, Brussels. http://groups.eortc.be/qol/sites/default/files/archives/translation_manual_2009.pdf

Dueck et al., 2015. Validity and Reliability of the US National Cancer Institute's Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). JAMA Oncol. 2015 Nov 1;1(8):1051-9.

Eiser C, Morse R. 2001. Can parents rate their child's healthrelated quality of life? Results of a systematic review. Quality of Life Research. 2001;10:347–57

Eunethta, Guideline Health-related quality of life and utility measures

http://www.eunethta.eu/outputs/methodological-guideline-rea-pharmaceuticals-health-related-quality-life

Ganz and Gotay., 2007. Use of patient-reported outcomes in phase III cancer treatment trials: lessons learned and future directions. J. Clin Oncol. 25(32): 5063-9

Gravanis I et al., 2013. The European Medicines Agency Review of Abiraterone for the Treatment of Metastatic Castration-Resistant Prostate Cancer in Adult Men After Docetaxel Chemotherapy and in

Chemotherapy-Naïve Disease: Summary of the Scientific Assessment of the Committee for Medicinal Products for Human Use. The Oncologist2013;18:1032–1042

Fairclough D. 2010. Design and Analysis of Quality of Life Studies in Clinical trials, Chapman and Hall 2nd edition. Examples: Value of PRO Data

Fayers P & Machin D. 2007. Quality of Life: The Assessment, Analysis and Interpretation of Patient-reported Outcomes, Wiley 2nd edition.

FDA Guidance for industry; December 2009: Patient-Reported Outcome measures: use in medical product development to support labelling claims

Hansen LK, Moinpour CM, Ermete RB. 2014. Enhancing Nurse Contributions To SWOG Clinical Trials Objectives. Seminars in Oncology Nursing. 2014; 30(1): 26-31

Hróbjartsson & Boutron. 2011. Clin Pharmacol Ther. 2011 Nov;90(5):732-6

Hullman et al., 2015. Medical and Psychosocial Associates of Nonadherence in Adolescents With Cancer. J Pediatr Oncol Nurs; 32(2): 103–113.

Jaeschke et al., 1989. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials. 1989 Dec;10(4):407-15.

Joly et al., 2007. Quality of life and/or symptom control in randomised clinical trials for patients with advanced cancer. Annals of Oncology. 18: 1935-1942

Khan SA et al., 2014, End-of-life care- what do cancer patients want? Nat. Rev. Clin Oncol. 11;100-108

King MT., 2011. A point of minimal important difference (MID): A critique of terminology and methods. Expert Review of Pharmacoeconomics & Outcomes Research. 2011; 11 (2): 171-184

Kyte D, Draper H, Calvert M., 2013. Patient-Reported Outcome Alerts: Ethical and Logistical Considerations in Clinical Trials. JAMA. 2013; 310(12): 1229-1230

Kyte D et al., 2014. Systematic Evaluation of the Patient-Reported Outcome (PRO) Content of Clinical Trial Protocols. PLoS One. 2014 Oct 15;9(10)

Kyte D et al., 2016. Management of Patient-Reported Outcome (PRO) Alerts in Clinical Trials: A Cross Sectional Survey. PLoS ONE 11(1): 1-14

Luckett T, King MT. 2010. Choosing patient-reported outcome measures for cancer clinical research – Practical principles and an algorithm to assist non-specialist researchers. European Journal of Cancer. 2010; 46(18):3149-57

Matza LS, Patrick DL, Riley AW, et al. 2013. Pediatric Patient-Reported Outcome Instruments for Research to Support Medical Product Labeling: Report of the ISPOR PRO Good Research Practices for the Assessment of Children and Adolescents Task Force. Value in Health. 2013;16(4):461-79

Mokkink L, Terwee C, Patrick D, et al. 2010. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. Journal of Clinical Epidemiology. 2010; 63: 737-745

Pasetto et al., 2007. Quality of life in elderly cancer patients. European Journal of Cancer. 43: 1508-1513.

Revicki D., et al 2008. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. Journal of Clinical Epidemiology 61;102-109

Secord AA et al., 2015. Patient-reported outcomes as end points and outcome indicators in solid tumours. Nat Rev Clin Oncol. 2015 Jun;12(6):358-370.

Sloan JA., et al 2006. The clinical significance of quality of life assessments in oncology: a summary for clinicians. Support Care Cancer. 14:988-998.

Sprangers M & Schwartz C. 1999. The challenge of response shift for quality-of-life-based clinical oncology research. *Annals of Oncology* 10: 747-749, 1999

Sneeuw KC, Sprangers MA, Aaronson NK. 2002. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease. Journal of Clinical Epidemiology. 55(11):1130-43, 2002

Staniszewska, S, Haywood KL, Brett J, Tutton E. 2012. Patient and public involvement in patient-reported outcome measures: evolution not revolution. Patient. 2012; 5(2): 79-87

University of Oxford, A structured review of patient-reported outcome measures for patients with lung cancer; Patient-reported outcome measurement group, Department of Public Health, University of Oxford, 2010

University of Oxford, A structured review of patient-reported outcome measures for patients with prostate cancer; Patient-reported outcome measurement group, Department of Public Health, University of Oxford, 2009

Wakefield C, Patterson P, McDonald FEJ, et al. 2013. Assessment of psychosocial outcomes in adolescents and young adults with cancer: a systematic review of available instruments. Clinical Oncology in Adolescents and Young Adults. 2013 February; 3: 13-27

Weldring and Smith., 2013. Patient-Reported Outcomes (Pros) and Patient-Reported Outcome Measures (Proms). Health Services Insights 2013:6 61–68

Wild D, Eremenco S, Mear I, et al. 2009. Multinational trials – Recommendations on the Translations Required, Approaches to Using the Same Language in Different Countries, and the Approaches to Support Pooling the Data: the ISPOR Patient Reported Outcomes Translation & Linguistic Validation Good Research Practices Task Force Report. Value in Health 2009;12:430-440

Wild, D, et al 2005. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. Value in Health, Vol. 8, Issue 2, 2005 pp95104http://www.ispor.org/workpaper/research_practices/PROTranslation_Adaptation.pdf

Wyrwich, K, Norquist JM, Lenderking, W, and Acaster S. 2013. Methods for interpreting change over time in patient reported outcome measures. Qual Life Res. 2013 Apr;22(3):475-83