EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

London, 15 December 20101
EMA/CHMP/205/95/rev.4
Oncology Working Party

Guideline on the evaluation of anticancer medicinal products in man

Draft Agreed by Working Party	**September 2011**
Adoption by CHMP for release for consultation	15 December 2011
End of consultation (deadline for comments)	31 May 2012

This guideline replaces guideline / NfG Reference.

Comments should be provided using this template. The completed comments form should be sent to **ONCWPsecretariat@ema.europa.eu**

Keywords	Cancer, malignancy, biomarker, targeted drugs, pharmacogenomics

7 Westferry Circus ● Canary Wharf ● London E14 4HB ● United Kingdom
Telephone +44 (0)20 7418 8400 **Facsimile** +44 (0)20 7418 8416
E-mail info@ema.europa.eu **Website** www.ema.europa.eu
An agency of the European Union

Table of contents

Executive summary

The purpose of this guideline is to provide guidance on all stages of clinical drug development for the treatment of malignancies, including drug resistance modifiers or normal tissue protective compounds. Supportive measures such as anti-emetics and haematopoietic growth factors, however, are covered by separate guidelines.

Alongside conventional aims such as defining the proper dose(s) and schedule(s), the importance of identifying a target population with optimised benefit risk is emphasised in Section 6: Exploratory Studies. Guidance is also provided on combination studies. Combinations of drugs with minimal activity as monotherapy, but synergistic effects when combined, as well as combinations of conventional cytotoxics, are also discussed.

Convincingly demonstrated favourable effects on overall survival (OS) are from both a clinical and methodological perspective the most persuasive outcome of a clinical trial. Prolonged progression-free or disease-free survival (PFS/DFS), however, are in most cases as such considered relevant measures of patients benefit, but the magnitude of the treatment effect should be sufficiently large to outbalance toxicity and tolerability problems. In order to capture possible negative effects on the activity of next-line therapies and also treatment related fatalities, informative data on overall survival compatible with a trend towards favourable outcome are normally expected at time of submission. This has consequences with respect to interim analyses and cross-over, which thus should be undertaken only when survival data will provide the information needed for a proper evaluation of benefit/risk.

An assessment of benefit/risk should encompass all relevant data on efficacy and safety, also taking into account uncertainties as well as external data of relevance in relation to the experimental compound and the disease to be treated. Therefore no precise definition of “trend towards favourable effects on survival” or “reasonably excluding negative effects on OS” is given in this document. If a major increase in toxicity is foreseeable (see section 8), it is recommended that confirmatory studies are undertaken with the aim to show an OS benefit. It is also acknowledged that improved safety without loss in efficacy may constitute tangible aims and the design of non-inferiority studies are discussed in 8.7.3.

In section 10, definitions and abbreviations used in this guideline are summarised. Appendix 1 provides methodological guidance on the use of PFS as endpoint in confirmatory studies. A planned appendix 2 will focus on the use of patient report outcome (PRO) measures and health-related quality of life (HRQoL) from a regulatory perspective. A revised paediatric guideline is also foreseen.

1. Introduction

The guideline on anticancer medicinal products adopted in 1996, and revised in 2001 and 2003, focused on conventional cytotoxic compounds. In 2005, a major revision was undertaken, aiming at covering non-cytotoxic compounds, to expand on the sections on exploratory trials and to provide more guidance with respect to methodological issues. Later, there followed an appendix on methodological issues related to use of PFS and in early 2010 an appendix on haematological malignancies followed. In this appendix disease specific guidance was introduced and the section on confirmatory studies based on aims of therapy and relative toxicity was restructured. These latter elements have now been incorporated in the revised main guideline. In this revision, the chapter on exploratory trials for cytotoxic compound has been shortened as it was considered too detailed and too prescriptive. The section on condition specific guidance has been expanded and now includes Non-Small Cell lung Cancer, Prostate Cancer, Chronic Myeloid Leukaemia, Myelodysplastic Syndrome and Haematopoietic Stem Cell Transplantation.

2. Scope

Whilst the thrust of a regulatory guideline should be on confirmatory studies, the aim of this guideline is also to underline the importance of exploratory studies in order to properly define the most appropriate target population in addition to the usual aims: to define dose, schedule, tumour type and line of therapy. The role of biomarkers to achieve these objectives is also further emphasised in this revised guideline.

There are numerous possible ways to classify anti-cancer drugs such as direct anti-tumoural *vs.* indirect anti-tumoural, or based on pharmacology or molecular target (*e.g.* hormones, immune

modulators, nuclear-targeting, signal-transduction targeting, etc.). As this document is meant to provide guidance on clinical drug development, the aim has been to classify compounds according to reasonable designs of exploratory studies, i.e. cytotoxic compounds where toxicity and ORR are considered suitable markers of activity in drug development vs. non-cytotoxic compounds where ORR and/or toxicity may not serve this purpose.

A very large number of anti-cancer compounds have been and currently are under development. Only a minority, however, have completed the clinical development and obtained a marketing authorisation, due to insufficient evidence of efficacy or evidence of a detrimental safety profile. Until non-clinical models with good predictive properties have been defined, this situation is likely to remain essentially unchanged and the absence of such models is considered to constitute the greatest hurdle for efficient drug development within the foreseeable future.

Since chemoprotective agents and drug resistance modifiers are used as part of anticancer regimens, some guidance on these agents will also be provided in appropriate sections of this guideline. Anti-emetics and haematopoietic growth factors, however, are covered in separate documents.

Additional recommendations of relevance for childhood malignancies and paediatric drug development are provided as a separate "*Addendum* on paediatric oncology".

3. Legal basis

This document should be read in conjunction with Directive 2001/83/EC, as amended. Applicants should also refer to other relevant European and ICH guidelines on the conduct of clinical trials, including those on:

- Nonclinical evaluation for anticancer pharmaceuticals EMEA/CHMP/ICH/646107/2008 (ICH S9)
- Reflection paper on methodological issues associated with pharmacogenomic biomarkers in relation to clinical development and patient selection EMA/CHMP446337/2011
- Clinical Investigation of the Pharmacokinetics of Therapeutic Proteins CHMP/EWP/89249/2004
- Evaluation of the Pharmacokinetics of Medicinal Products in Patients with Impaired Hepatic Function - CPMP/EWP/2339/02
- Investigation of drug interactions - CPMP/EWP/560/95
- Points to Consider on Adjustment for Baseline Covariates - CPMP/EWP/2863/99
- Points to Consider on Multiplicity Issues in Clinical Trials - CPMP/EWP/908/99
- Guideline on the choice of the choice of non-inferiority margin - CPMP/EWP/2158/99
- Guidance on Statistical Principles for Clinical Trials - CPMP/ICH/363/96
- Guideline on clinical trials in small populations- CPMP/EWP/83561/2005
- Choice of Control Group in Clinical Trials -CHMP/ICH/364/96 (ICH E10)
- Guideline on clinical evaluation of diagnostic agents - CPMP/EWP/1119/98
- Note for guidance on clinical safety data management: definitions and standards for expedited reporting - CPMP/ICH/377/95 (ICH E2A)
- Note for guidance on clinical safety data management: data elements for transmission of individual case safety reports - CPMP/ICH/287/95 (ICH E2B)
- Points to consider on application with 1. Meta-analyses 2. One pivotal study - CPMP/EWP/2330/99

4. Pharmacokinetics

In general, the same recommendations are valid for anticancer products as for other medicinal products and reference is made to the clinical pharmacology guidelines available including the conduct of food interaction studies prior to phase III. In the past, mass-balance studies have not been performed to the same extent for anticancer drugs as for other medicinal products. Due to the importance of the information gained in these studies for the understanding of the clinical pharmacology of the investigational drug, including the drug-drug interactions assessment, mass-balance studies are strongly recommended. (Investigation of drug interactions, CPMP/EWP/560/95)

Studies to be undertaken in patients with impaired organ function should mainly be selected based on prior information on the mode of elimination of the drug and formation/elimination of potential pharmacologically active metabolites. If a study in hepatic impairment is needed, as a first step a study in patients with liver metastases is warranted. Whether studies in more advanced liver disease are needed should be decided on a case by case basis.

It is recommended to evaluate the influence of intrinsic factors through population PK analyses of sparse phase III plasma concentration data. This evaluation could include factors such as age, weight, renal function, S-bilirubin, genotype etc. (Evaluation of the Pharmacokinetics of Medicinal Products in Patients with Impaired Hepatic Function, CPMP/EWP/2339/02)

It is recommended to collect sparse samples in pivotal trials. This information aids in understanding the exposure-response relationships for the drug, and may allow for a rational selection of treatment strategies in patients who are risk for excessive toxicity or ineffective therapy.

Exploratory studies, including PK, in patients with malignant ascites or other third space conditions are encouraged.

5. Biomarkers

In order to optimise benefit – risk, it is essential to identify the proper target population for therapy. This might be possible to accomplish through the judicious use of biomarkers in all phases of clinical drug development. A biomarker should be capable of measuring and evaluating a normal biological process, a pathological process or the pharmacological response to a therapeutic intervention, depending upon its purpose.

Irrespective of pharmacological class, it is assumed that entrance into clinical development of new molecule today is guided by translational research. This means that in most cases there are hypotheses to be tested and candidate biomarkers available. The utility of biomarkers is broad e.g. prospective stratification of clinical trial subjects according to biomarker status, determination of the biologically effective dose, early proof of mechanism or concept, assessment of toxicity and an indication of the natural course of a disease. However, although efforts to identify targets and explain variability in PK and PD are essential, the need to confirm the findings should not be overlooked in the planning of the drug development programme (technical and clinical validation).

It is acknowledged that biomarkers tested in early clinical trials often were exploratory in nature, but it is essential that technical/quantitative reliability is assured. While serum biomarkers or other sources of biological samples might be informative, tumour samples are expected to constitute an integral part of the biomarker exercise if not properly justified based on pharmacological properties. Normal tissues samples may also be used in early clinical studies, if non-clinical studies indicate that there is a correlation between the changes observed in normal tissues and changes in tumour tissue. The role of functional imaging in early drug development is not regarded as well established, but its use is encouraged.

The development of companion diagnostic methods should be considered early in clinical development, maximising the clinical application of the technology. Commercially available diagnostic assays should comply with the requirements laid down in IVD Directive (98/79/EC).

For the use in confirmatory studies and e.g. as measures of efficacy, biomarkers must be carefully and rigorously validated following systematic evaluation in well designed prospective clinical trials (Reflection paper on methodological issues associated with pharmacogenomic biomarkers in relation to clinical development and patient selection EMA/CHMP446337/2011). In order to assist in interpretation of results across studies and limit sources of variability when developing biomarkers, the use of available guidelines is encouraged, e.g. reporting recommendations for tumour marker prognostic studies (REMARK).

6. Exploratory studies

Exploratory studies are essential in rational drug development. The distinction between Phase I/II exploratory and Phase III confirmatory trials has been adhered to in this Guideline. However, this does not mean that exploratory aims should not form an important part of Phase III trials. Similarly, hypothesis generation, testing and confirmation may form parts of Phase II trials.

6.1. Cytotoxic compounds

This refers to conventional cytotoxic agents, i.e. compounds inducing irreversible lethal cellular lesions following short-term exposure through interference with DNA replication, mitosis, etc. For these compounds, toxicity and tumour response are considered suitable indicators of activity.

As for non-cytotoxic compounds, non-clinical and clinical studies encompassing aims to characterise prerequisites for activity/resistance and to identify markers of resistance are encouraged.

6.1.1. Phase I, single agent dose and schedule finding trials

The basic assumption governing the design of these trials is that, for dose finding purposes, toxicity is an acceptable endpoint. The main objective is thus to define dose-limiting toxicities and the dose to bring forward into further trials. While meeting this objective is generally straightforward, in spite of the fact that the inter-patient variability in PK might be large, it is often more complex to define reasonable dose schedules to study further.

It is accepted that the dose initially is calculated per body surface area (BSA), but the empirical support for the notion that this approach meaningfully reduces inter-patient variability in exposure is weak. Whether a flat dose or a dose calculated according to BSA or weight is used, it is recommended that the importance of BSA or weight for variability in exposure is explored through modelling.

Main Objectives

- Maximum Tolerated Dose (MTD), Dose Limiting Toxicity (DLT) and a recommended Phase II dose (RP2D) (usually one dose step below MTD) should be identified for defined schedules and modes of administration.
- Frequent side effects and target organs for toxicity should be characterised as regards relationship to dose and schedule. Extent, duration and reversibility should be determined.

Eligibility of patients

These trials should normally be undertaken in cancer patients without established therapeutic alternatives.

Routes of administration and schedules

In most cases, intravenous administration, when feasible, is advisable for first use in man studies since it eliminates variability related to bioavailability.

For schedule finding, experience related to class of compounds is helpful. Non-clinical data with respect to cycle dependency and the ratio tumour / normal tissue cytotoxicity *ex vivo* may be of some interest.

Dose escalation

In case of minimal toxicity, or occasionally in case of non-significant toxicity, within-patient dose escalation may be appropriate in order to reduce the number of patients exposed to non-active doses, but is acceptable only if non-clinical data provide no evidence of cumulative toxicity.

If toxicity is acceptable, the patient may be re-exposed upon recovery and preferably should receive at least 2 cycles at the same dose level.

Evaluation of toxicity

The minimal requirements for evaluation of adverse effects include assessment of symptoms, physical examination, ECG, blood and urine laboratory analyses and radiological assessment as appropriate. Preclinical data should be used to guide the need for further examinations. If there are no signals with respect to QTc in preclinical studies or related to class of products, no dedicated QTc studies are expected, but inclusion of ECG as part of routine monitoring is recommended. Local toxicity at the site of administration should be specifically recorded. The toxicity should be graded according to a generally recognised system (e.g. WHO toxicity criteria, Common Terminology Criteria for Adverse Events, CTCAE).

Factors influencing toxicity (organ dysfunction, concomitant therapy) should be explored as appropriate. These factors should be further elucidated in Phase II/III.

6.1.2. Phase II, single agent therapeutic exploratory studies

Phase II trials may investigate single-agent activity in a variety of tumour types, or in a selected tumour type, or investigate activity and feasibility of combination or multimodality regimens.

This section is focused on trials where the primary objective is to estimate single agent antitumor activity in patients with a defined tumour type in order to identify compounds to bring forward to confirmatory trial.

Objectives and design

Phase II trials may use a variety of study designs and early studies should provide initial evidence of treatment activity and tolerability. Inclusion of a randomised control arm is ~~to be~~ encouraged, particularly if only one confirmatory pivotal trial is foreseen

The studies are intended to:

- To assess the probability of response in the target tumour type and conclude on the need for further studies (investigate earlier stages of the disease, combinations, compare with standard therapy).
- To investigate pharmacogenomics, where appropriate
- Further characterise dose and schedule dependency, with respect to safety and activity
- Further characterise the side-effects of the medicinal product:
- When applicable, further characterise the optimum route of administration

Selection and number of patients

Exact definition of the target disease, previous therapy (if any) and stage should be given, in line with internationally agreed diagnostic criteria.

Provided safety and activity is reasonably established and there is a scientific rationale, it might be appropriate to conduct studies also in patients for whom alternative therapies are available. This includes the neo-adjuvant setting in treatment naïve patients scheduled for surgery, provided that delay in surgery cannot be detrimental to the patient. The safety and interests of the patient must always be guaranteed and a detailed justification should be provided in the study protocol. In these cases, the use of sensitive measures of anti-tumour activity is expected.

Dose and schedule

The dose and schedule should be clearly defined. Details on the administration of the medicinal product with special precautions (hydration of patients, protection against light and temperature, etc.) should be stated as well as other agents, which are contraindicated during the study period.

- Guidance should be supplied outlining dose modifications related to the severity of the observed toxicity.
- Rules for dose escalation in case of low toxicity should be considered.
- Consideration should be given to study high-risk patients (e.g. high risk with respect to target organ toxicity or compromised metabolic or excretory mechanisms for the experimental compound) separately.

Any evidence of cumulative toxicity should be recorded and estimated as a function of total dose. This should be specifically studied according to target organ or function.

Evaluation of activity

ORR should be documented according to international standards (e.g. RECIST, or WHO criteria). Modifications of these criteria may be appropriate in certain situations, but should be justified.

In evaluating ORR, data for all patients entered into the trial should be reported. Where ORR in the per-protocol analysis set is considered to be of primary interest, then data for all patients included into the trial should also be reported. External independent review of tumour response is encouraged, according to the objectives of the trial.

In haematological malignancies, disease specific response criteria are unavoidable in many cases and full harmonization has not yet been accomplished for some disease entities. Therefore it is of importance to follow the progress made by international working groups on these issues. Especially if

less conservative disease specific response criteria are introduced in new clinical guidelines, a justification with focus on aspects of drug development is expected from the sponsor.

Data on duration of response, TTP/PFS and available data on OS should normally be reported. The use of tumour biomarkers and other dynamic measures of activity is encouraged.

In patients with symptomatic disease at base line, the assessment of symptom control is encouraged, if a randomised phase II trial is undertaken.

6.2. Non-cytotoxic compounds

This refers to a very heterogeneous group of compounds ranging from antihormonal agents to antisense compounds, signal transduction, angiogenesis or cell cycle inhibitors, immune modulators, etc. The common element affecting the design of clinical trials is that toxicity may not be an appropriate endpoint in dose and schedule finding trials and ORR may not be an appropriate measure of anti-tumour activity.

For these reasons, the early stages of clinical drug development are more complex and have to be tailored according to the assumed pharmacology of the individual compound as defined in non-clinical studies. The rather strict delineation between Phase I and II trials, as for conventional cytotoxic compounds, may be less relevant as measures of anti-tumour activity, e.g. based on assessment of biomarkers might be needed early in order to define dose and schedule.

Otherwise, most of the elements discussed in relation to cytotoxic drugs are of relevance also here such as restrictions with respect to patient eligibility, recommendations as regards routes of administration, evaluation of toxicity and anti-tumour activity, etc. These issues will not be further discussed here.

6.2.1. Phase I, single agent dose and schedule finding trials

Based on preclinical tolerability and toxicology findings and the assumed pharmacology of the compound, early trials may sometimes be conducted in healthy volunteers. Tolerability, safety, PK and, if at all possible, PD measures of activity are appropriate objectives.

Non-clinical data and, when available, data from healthy volunteers should be used to design the studies to be conducted in patients, e.g. as regards eligibility criteria and starting dose. In accordance with the guidance for cytotoxic compounds, availability of established therapies should normally be regarded as an exclusion criterion. Refractoriness to conventional cytotoxic compounds, however, may confer resistance also to some clearly non-related compounds. This obviously affects the possibility to define a dose/concentration – effect relationship. All sensible and ethically acceptable measures undertaken to increase the assay sensitivity of these clinical trials, including the conduct of window of opportunity studies are encouraged. Whenever appropriate, this includes measuring the expression of the assumed target(s) for drug activity.

PD measures may include biochemical measures (receptor binding, enzyme inhibition, downstream events, etc. as defined in non-clinical studies), functional imaging, proteomics, immunological measures (antibody or T-cell response), etc. Population PK/PD studies are encouraged. For compounds shown to be cytostatic in non-clinical models, prolonged exposure may be needed to elicit tumour shrinkage in clinical studies. If in these cases unexpected, early tumour shrinkage is observed this constitutes a signal indicating that further studies exploring the underlying mechanisms behind early response are warranted.

While it is acknowledged that drug development for compounds with a single main target for activity, such as mutated BRAF, is more straight forward, it is still expected that the pharmacological rational behind poly-targeting compounds is reflected in the exploratory studies programme, e.g. in terms biomarkers selected in order to identify the proper target population for treatment.

Until now available experience indicates that tumour selectivity is not to be expected. Tolerability and toxicity thus remain important measures in dose and schedule finding studies. Even if, e.g. saturation of the target for drug activity, or a desired PD activity can be demonstrated without significant toxicity, it is still advisable to investigate higher dosages in order to better define the safety of the compound and possible irregularities in PK and PD. This may include defining MTD.

If not pharmacologically justified, proper analyses of biopsies from tumours (primaries and metastatic lesions), are expected to constitute a pivotal role in studies undertaken to identify the proper target

population for confirmatory studies. This might be crucial and has to be considered in the recruitment of investigators and patients.

As for conventional cytotoxic drugs, the use of tumour markers and sensitive imaging techniques, in combination with conventional methods, are recommended in order to delineate possible antitumor activity. Also in exploratory trials it is recommended that technical standardisation of, e.g. functional imaging techniques, is implemented in order to reduce inter-centre variability.

Eligibility criteria and the number of patients should be defined according to the objectives of the study, also taking into account variability in PK and PD at doses and schedules selected for further studies.

6.2.2. Phase II, single agent therapeutic exploratory studies

For the purpose of simplification, it is assumed that a dose/exposure range has been defined that shows pharmacological activity/target occupancy with or without dose limiting toxicity. If not otherwise justified, it is postulated that activities related to identification of the proper target population, as discussed above, continues in these studies.

Study designs and measures of activity

ORR, despite all its shortcomings related to patient-selection, etc, is a rather convincing measure of activity as for most tumours, spontaneous regression fulfilling criteria for at least partial response is an uncommon phenomenon. For exploratory purposes, studies without a randomised reference are therefore considered interpretable and guidance provided in the section about cytotoxic compounds is relevant. Irrespective of this, inclusion of a randomised reference arm is encouraged and might be of special interest in order to explore whether, e.g. a selected biomarker is prognostic and/or predictive.

Time to progression (TTP) and progression-free survival (PFS), however, are in principle a function of underlying tumour growth rate and the activity of the anti-tumour compound. Also, if documented progressive disease is an inclusion criterion, underlying growth rate is hard to define in most patients and historical data will be even harder to interpret. Therefore, the interpretation of TTP/PFS data without a randomised reference is problematic.

Exploratory trials with time-related endpoints

There is probably no ideal yet feasible design of exploratory studies for compounds assumed to mainly elicit tumour growth control. In the following some design alternatives are discussed, all with pros and cons, but in principle acceptable from a regulatory perspective.

- A randomised, dose comparative trial, e.g. comparing the lowest dose likely to be pharmacologically active with higher dose(s), if showing a difference in TTP/PFS, will obviously provide evidence of activity, but not in absolute terms.
- Randomised withdrawal of therapy in patients with non-progressive disease after a defined period of time on experimental therapy. The acceptability of this design to patients and investigators, however, may constitute an obstacle and carry-over effects may be a reality for some compounds.
- In previously treated patients, a within patient comparison of TTP/PFS might provide evidence of activity. Here TTP on last prior therapy is compared with TTP/PFS on the experimental therapy. It should be noted, however, that the underlying assumption of non-decreasing growth rate over time cannot always be substantiated. For exploratory purposes this constitutes no major concern. It is advisable to recruit patients with secondary as well as primary resistance on prior therapy. This ensures at least to some extent, that the study population is representative. It should also be noted that patients with early failure (primary resistance) on prior therapy may show some inversions in terms of TTP just due to fluctuations in tumour growth rate and variability related to imaging techniques.

  For certain indications, a within patient comparison may be justified also in treatment naïve patients.
- A randomised phase II study versus a compound known to be active in the selected population (or placebo/BSC if justified) provides another alternative. If such a study is regarded as exploratory, there is no need for, e.g. well-defined non-inferiority criteria. In a comparison in terms of TTP it should be noted, that a purely growth inhibitory compound is “favoured” compared with a compound inducing tumour shrinkage, as progression is defined in relation to best tumour response. At the time of tumour progression, the tumour burden in patients failing

a purely growth inhibitory compound will therefore be higher than in patients where tumour shrinkage was elicited. .

- If no more refined techniques are applicable, TTP/PFS without an internal reference has to be accepted as a measure of Phase II anti-tumour activity. A systematic literature review is advised in these cases. Fixed-time related endpoints such as percentage of patients without progression after a predefined period of therapy may be used in order to define whether the apparent anti-tumour activity is sufficiently high to justify the conduct of, e.g. Phase III confirmatory studies.

In principle, a statistical approach similar to that for Phase II trials with ORR as outcome measure is applicable. It is harder to set up criteria for early termination, however. The number of patients should be sufficient to obtain a reasonably precise estimate of the percentage of progression-free patients at a predefined time point. The underlying assumptions as regards progression rate without therapy are more problematic and “promising activity” is harder to define.

For these studies, the use of conventional criteria for ORR and tumour progression is recommended and independent review is encouraged. It is recognised, however, that, e.g. an apparent increase in tumour size due to inflammatory oedema might be a first sign of activity for certain compounds. If prior trials indicate that this is the case, it is accepted that this is accounted for in the study protocol. The use of ORR and TTP as key measures of activity should not be regarded as contradictory to the use of tumour/PD markers in parallel.

If a randomised design is considered appropriate, the use of generally accepted instrument to estimate HRQoL or symptom control may provide valuable information.

For window of opportunity studies and if sensitive measures of pharmacological activity are available, e.g. functional tumour imaging, and a target population has been identified with tumours likely to be sensitive, placebo-controlled trials with one or preferably more doses of the experimental compound might be feasible. Sensitive measures, even if not fully validated with respect to relationship to ORR, are from a regulatory perspective acceptable for exploratory purposes and allow not only for refined dose comparisons, but also early escape in case of absence of activity. It is advisable though to clearly define in the protocol criteria for progressive disease, whether a composite (e.g. biomarkers, or imaging, or symptoms) is used or not.

6.2.2.1. Monoclonal antibodies (MoAb)

Monoclonal antibodies may affect tumour cells directly, e.g. through ADCC and/or blocking of growth factor/anti-apoptotic receptor signalling, or through the targeting of growth factors for the tumour or tumour supportive structures.

As appropriate, tumour cells or plasma should be screened for (over-)expression of the target and the relationship between target expression and activity should be investigated.

Tumour specificity is frequently not attainable, but it is possible to screen for “unwanted” targets in vitro, facilitating the safety assessment.

Understanding PK provides some guidance for dose-finding as clearance may be related to target saturation.

If, e.g., a growth factor receptor is targeted and pending of the characteristics of the MoAb (Ig subclass, association with toxin, etc.) it is of relevance to try to elucidate whether blockade of the receptor or ADCC is of prime importance for antitumor activity. Studies conducted in the neo-adjuvant setting allowing for repeated microscopic examinations may provide means to investigate this.

The experience as regards immunogenicity of MoAbs in other fields of clinical medicine should be taken into account with respect to choice of assays, markers for loss of activity and possible safety problems.

6.2.3. Immune modulating compounds including tumour vaccines

Therapeutic cancer vaccines are aimed to induce specific anti-tumour immunity toward existing malignant disease. Such immune therapy is normally aimed to induce an adaptive T cell immune response in cancer patients. The nature of the drug substances used is highly variable, including synthetic peptides, recombinant proteins, virus-like particles, immune-modulating antibodies, gene therapy, and cell-based products. As it is difficult to break tolerance towards tumour antigens which

are normally derived from self-antigens, cancer vaccines are often combined with pharmacologically active adjuvants such as cytokines or toll-like receptor agonists. One other approach to break immune tolerance is to block T cell inhibitory signals. The resulting T-cell activation and proliferation leads to wanted and unwanted immune stimulatory effects: the desired anti-tumour effect as well as the appearance of immune related toxicities like colitis and endocrine insufficiency.

Non-clinical in vitro and in vivo proof-of-concept studies should be presented to justify the planned doses and schedule in early clinical studies. The dose level and the schedule in the non-clinical toxicology studies should be based on a dose that showed biological activity in proof-of-concept studies. It is acknowledged that for products relying on human-specific antigens which need to be presented on human MHC molecules, predictive animal models are often not available. Nevertheless, animal models using homologous antigens or animals being human MHC transgenic might be considered for non-clinical pharmacology and toxicology studies, if available.

The aim of early clinical trials is to determine the safety and the dose and schedule that induced a desired immune response. Monitoring the immune response, i.e. the induction of antigen-specific T cells or the presence of a humoral response is essential to determine appropriate dose and schedule. To achieve this goal multiple monitoring assays may be necessary and these should be carefully explored. The analytical methods should be described in detail in the clinical trial protocol.

Tumour biopsy data taken before and after treatment is expected to play a pivotal role in assessing the extent and type of immune activation in the target tissue and could serve as an early marker for possible anti-tumour activity.

The induction of tumour response in patients with high tumour burden might be a too high hurdle to overcome and may favour the inclusion of patients with minimal or low tumour burden. Examples are therapy of patients with NSCLC after complete tumour resection where cancer immunotherapy can be assessed in the adjuvant setting. Another example is patients suffering from non-resectable NSCLC who have responded to chemotherapy. The design of clinical studies using clearly experimental therapies in patients with limited and measurable disease, not heavily pretreated with cytotoxic regimens has to be carefully justified. As for other non-cytotoxic or cytotoxic agents evidence of anti-tumour activity is essential prior to the initiation of confirmatory studies.

Oncology patients are usually taken off treatment upon disease progression. Induction of an effective immune response and clinical response may need more time to develop (delayed effect) compared to classical cytotoxic compounds. Patients may thus experience disease progression prior to the onset of biological activities or clinical effects. Discontinuation of active cancer immunotherapy in case of slow progression may not be appropriate. In these situations a detailed definition of “slow progressive disease” is expected in the study protocol. In exploratory studies, revised criteria defining progression is accepted if properly justified, while in confirmatory studies OS should be prioritized.

Possible toxicities like induction of autoimmune reactivity (cellular and humoral) and induction of tolerance should be carefully monitored during the clinical development.

6.3. Combination therapy studies

Conventional cytotoxic compounds have for long been used in combination in order to increase the anti-tumour activity at acceptable levels of toxicity. This may be accomplished by combining compounds with at least partly non-overlapping toxicity and, perhaps, partly non-overlapping prerequisites for activity/resistance. Regulatory agencies, as well as learned societies, have accepted this approach, but it is acknowledged that it is frequently unknown whether combined use results in a better long-term outcome than consecutive use.

6.3.1. Combining conventional cytotoxic compounds

In the selection of patients with available alternative therapies, the documented activity of the individual components of the combination regimen should be taken into account.

The exploratory phase encompasses the determination of MTD and RP2D for the combination and a preliminary assessment of anti-tumour activity in terms of ORR and PFS/TTP. While the degree of anti-tumour activity for a new combination relies on assumptions, it is often possible to predict toxicity, based on the toxicities of the individual components. If relevant PK interactions can be excluded, and pending on the dose-response/toxicity profiles, dose-finding studies may be initiated at about 1/2 of the recommended mono-therapy dose for each compound. It might also be appropriate to start at the

full recommended mono-therapy dose for one of the compounds and reduced dose (<50%) for the other compound. As the sequence of administration may be of importance with respect to potential PK interactions and anti-tumour activity, this has to be accounted for in the design of the studies. More patients on each dose level are normally needed compared with single agent dose finding studies.

There is no uniform way to balance dose intensity between components of a combination regimen to optimise benefit – risk. It is thus accepted that, e.g. priority in terms of dose intensity is given to the compound with the highest monotherapy activity.

If one of the components is regarded as an acceptable treatment regimen in monotherapy, a randomised phase II study comparing the monotherapy regimen with the combination is informative. For confirmatory studies a comparison with the best available, evidence-based reference regimen is expected.

6.3.2. Combinations involving a non-cytotoxic drug.

If there are no strong biological/pharmacological arguments to the contrary, the selected chemotherapy regimen to be combined with the non-cytotoxic should normally be “best available”. If the dose intensity/systemic exposure of the chemotherapy regimen is unaltered it can be assumed that all patients will receive appropriate therapy. Therefore there is no need to restrict the eligibility of patients from this perspective.

Whenever previous non-clinical and clinical experience has suggested that PD markers, etc. might be informative with regard to anti-tumour activity, they should be part of the experimental plan. This may include investigations whether the expression of the target for the non-cytotoxic compound is affected by treatment with cytotoxic agents and if appropriate *vice versa*.

Given the current status with respect to predictability of add-on activity in non-clinical models, randomised phase II studies comparing the experimental regimen with the chemotherapy-alone regimen are considered essential. For these studies, it is recommended that conventional anti-tumour activity data (ORR and TTP) are supplemented with tumour markers and sensitive measures of, e.g. tumour metabolic activity as appropriate.

When add-on activity of the non-cytotoxic compound to a chemotherapy regimen has been demonstrated, the need for further randomised phase II studies when new indications are studied may be dispensable. This, however, should be justified as the importance of target expression and inhibition thereof might differ between malignancies.

If the expression of the target for the non-cytotoxic compound may be differently affected by different chemotherapy regimens, it is advisable to study target expression during treatment with a new chemotherapy regimen prior to the conduct of add-on studies.

Research aiming at understanding the mechanisms and prerequisites for the add-on effects is encouraged, as it may allow for an improved characterisation of target populations in future studies.

It is conceivable that for some non-cytotoxic compounds, combinations are needed not only to optimise anti-tumour activity, but actually are required in order to obtain activity. For such compounds, e.g. target saturation in monotherapy and, importantly, non-clinical toxicity for the combination may be used to define suitable starting doses and schedules. Otherwise dose/schedule exploratory and therapeutic exploratory studies may proceed essentially as for a monotherapy regimen.

If supported by strong biological and/or pharmacological non-clinical and early proof-of-principle clinical data, two new compounds may be combined in a co-development program.

Uni-enhancement refers to scenarios when one combination partner *B,* which has no or minimal anti-tumour activity per se, but enhances the anti-tumour activity of the other partner *A* (e.g. through prevention of resistance development). The contribution of *B* needs to be established by data from appropriate non-clinical models. In phase II the comparison to a reference treatment is encouraged, while Phase II monotherapy data for *B* may be considered dispensable. An appropriate phase II design would be a randomised three-arm study *AB vs. A vs. reference treatment.*

Co-enhancement is considered when both combination partners demonstrate (modest) anti-tumour activity per se and the anti-tumour activity of the combination is considerably increased. In phase II, the new combination should be compared to both combination partners as single agents at efficacious doses and preferably a reference treatment: *AB vs A vs B vs reference treatment*. Depending on the phase II results one or both monotherapy arms may be dispensable in phase III.

Synthetic lethality refers to a scenario when both combination partners have no or minimal anti-tumour activity per se but exhibit potent activity as a combination. If the contribution of both partners is established by data from appropriate models and if dose escalations studies investigating an extensive range of doses have excluded that inappropriately low doses led to the assumption of minimal anti-tumour activity in monotherapy, monotherapy treatment arms may be dispensable for phase 2 studies.

7. Phase III, confirmatory trials

Confirmatory trials should be designed with the aim to establish the benefit - risk profile of the experimental medicinal product, including supportive measures, in a well-characterised target population of relevance for clinical practice.

In the general part of this section (8.2 – 8.4), the aim of therapy, curative versus long term disease control vs. palliation and not the underlying disease has been used to structure the discussion. This is of relevance also for medicinal products developed for the treatment of conditions where there are no meaningfully active treatment options.

For some malignancies where treatment is administered without curative intent, there are alternative, in clinical practise still well established regimens, showing major differences in anti-tumour activity. This reflects that selection of therapy in the clinic is guided by efficacy and safety. It is therefore of relevance in the planning phase to take into account the expected tolerability/toxicity profile of the experimental regimen compared with the selected reference regimen. It is fully acknowledged that safety data may be rather limited prior to the conduct of the first confirmatory trial, but main toxicities should normally have been identified and this should be sufficient for a rough estimate of the expected relative toxicity of the experimental regimen compared with alternative reference regimens.

Three categories are used in this document: Reduced or similar toxicity, increased toxicity and major increase in toxicity. No precise definition is given here due to heterogeneity of the conditions. ”Major increase in toxicity”, however, in most cases refers to a fear that the experimental regimen might be associated with an increase in treatment related deaths, irreversible adverse events with an impact on QoL, or severe impairment to patient condition. Other issues to take into account include risk for secondary tumours. This categorisation is mainly meant for guidance in the planning of confirmatory studies and in order to provide advice on regulatory expectations with respect to study outcome measures in order to enable a proper benefit – risk assessment.

7.1. Design

7.1.1. Patient population

With respect to diagnosis, criteria for initiation of treatment, eligibility, response criteria and choice of reference therapy, a justification based on scientific evidence and/or generally acknowledged and updated treatment guidelines are expected. While this is true in general, it is also expected that the exploratory studies through the judicious use of biomarkers provide guidance with respect to selection of patients in order to optimise benefit – risk, whether in need for confirmation or not in the planned phase III trials.

There is a general wish to reduce heterogeneity of study populations in order to increase the ability of the study to detect differences between study arms. This has to be balanced against the availability of patients for inclusion and the wish to enrol a clinically representative selection of patients. Therefore investigators should normally be encouraged to include patient’s representative of those likely to be treated with the experimental compound in clinical practice. Restrictions as regards, e.g. performance status should be reflected in the SPC. With respect to studies with a non-inferiority efficacy objective, please refer to 8.7.3.

Patients are expected to be characterised by relevant tumour parameters, e.g. stage, grade, target expression, other biological markers of importance for prognosis and/or tumour sensitivity, prior therapy (responsive/ resistant/refractory as appropriate), as well as performance status, co-morbidity, organ dysfunction, etc. Stratification based on important and well established prognostic covariates should be considered. In case adjusted analyses are to be undertaken for covariates other than those used for stratification, these factors should be pre-specified in the protocol or the statistical analysis plan (Points to Consider on Adjustment for Baseline Covariates CPMP/EWP/2863/99).

If exploratory studies provide a basis for including/excluding certain patients based on tumour phenotype/genotype, this will be reflected in the labelling. As a corollary, if patients with tumours not expressing the target for activity are eligible, a restricted labelling may still be appropriate if it has not been demonstrated, e.g. by subgroup analyses, that target expression is irrelevant for anti-tumour activity.

As some of the conditions are rare, it is understood that the sponsor might wish to define the target population using alternative criteria to those commonly employed. For example, in studies investigating the activity of a compound targeting a specific, molecularly well-defined structure assumed to be pivotal for the condition(s), it might be possible to enrol patients with formally different histological diagnosis, but expressing this target. The driving role of the target in different histological diagnoses must be demonstrated. This should also be addressed in exploratory studies, but it is accepted that formal testing with adequate statistical power of such a hypothesis cannot always be done. Possible consequences with respect to selection of proper reference therapy(ies) must be considered and the study should be designed so that it is possible, based on all available evidence, including non-clinical and pharmacological data, to conclude on the benefit – risk in the different subgroups of patients for which a claim is to be made, taking into account multiplicity issues (Points to Consider on Multiplicity Issues in Clinical Trials CPMP/EWP/908/99). Prior to the initiation of confirmatory studies using non-conventional criteria for eligibility, EU scientific advice should be sought.

Some possible target indications comprise very small groups of patients, so small that “exceptional circumstances” might apply. Unless the target for activity is expressed only in these rare conditions, sponsors are in general advised to initiate confirmatory studies in these small patient groups when benefit – risk is established in indications allowing a more comprehensive evaluation, especially with respect to safety.

7.1.2. Reference therapy

The choice of reference regimen should be justified and normally this regimen should be selected from best available, evidence-based therapeutic options. In this context, “best available, evidence-based” should be read as a widely used, but not necessarily licensed regimen with a favourable benefit-risk convincingly documented through randomised trials and considered at least as good from a benefit/risk perspective as alternative, treatment options. It is acknowledged that there are different, region-preferred standards. For superiority studies (test vs. ref.) this should normally not constitute a problem as long as the reference is evidence-based. For add-on studies (ref. + test vs. ref.), it might be possible to use a few, region-preferred references. Here a convincing clinical/pharmacological justification is needed, and EU scientific advice is recommended.

If the aim is to demonstrate non-inferiority, the selected reference regimen must enable a proper definition of the non-inferiority margin. In most cases, this would require randomized well-controlled studies have showed the superiority of the selected reference versus control. This is of particular relevance if the reference regimen is non-licensed. Please also refer to 8.6.3.

Amongst best available references, regimens with similar cycle lengths should be prioritised as it facilitates the identical scheduling of tumour assessments. If the objective is not to improve tolerability and toxicity, a regimen with similar expected toxicity to the experimental regimen is also preferred

In some cases there is no well documented reference regimen, even though patients in clinical practice are treated with certain regimens. Even though BSC is acceptable in these cases, an active comparator, documented e.g. in terms of response rate, is often preferable. If a single reference regimen cannot be defined, investigator’s best choice is an option. In these cases reference regimens with low toxicity are favoured and superiority in terms of patient relevant endpoints should be demonstrated.

The absence of evidence-based therapies often refers to patients who have failed several lines of therapy. In this situation, it might be easier to obtain the data needed for marketing authorisation based on a properly conducted randomised study in less advanced patients, supported by “salvage” single arm studies, compared with conducting a last line, randomised BSC/investigator’s best choice comparative study.

7.1.2.1. Single agent and combination therapies

Whether the experimental agent is used as a single agent or in combination, the experimental regimen should be compared with the “best available” comparator again referring to benefit/risk, not only to efficacy.

If the experimental agent (A) is added to an established regimen (B), superiority of AB vs. B should be demonstrated and benefit-risk should be shown to be favourable. A discussion is expected based on available data as regards dose intensity of B and benefit risk. Traditionally, this type of studies does not include an A alone third arm, but this should be justified based on available exploratory study data.

In case of substitution studies, i.e. studies where a component (C) of an established regimen (BC) is replaced with an experimental agent (A) and if non-inferiority (BC vs. BA) is the aim, the contribution of C to the activity of BC has to be well defined (Guideline on the choice of the choice of non-inferiority margin CPMP/EWP/2158/99).

Uncommonly, an entirely new combination AB is tested against a reference regimen. In these cases, solid non-clinical and clinical phase I/II data should support the need for all components in the experimental regimen.

7.1.3. Cross-over

In order to enable a qualified benefit – risk assessment, cross-over at time of progression should be undertaken only when precise estimates of OS data have been established (see Appendix 1).

7.1.4. Randomisation and blinding

Randomisation and stratification should adhere to the general principles laid down in current guidelines (Guidance on Statistical Principles for Clinical Trials CPMP/ICH/363/96). In many cases, a double-blind design is no option due to obvious differences in toxicity between study regimens or due to safety concerns. If the study has to be conducted as an open label study, this has implications with respect to choice of study endpoints and conduct of sensitivity analyses and other measures to be undertaken to limit potential bias related to the open-label nature of the trial.

7.1.5. Endpoints

Confirmatory trials should demonstrate that the investigational product provides clinical benefit. There should thus be sufficient evidence available demonstrating that the chosen primary endpoint can provide a valid and reliable measure of clinical benefit in the patient population described by the inclusion criteria. In the following, superiority trials are the focus of the discussion.

Acceptable primary endpoints include cure rate, OS and PFS/DFS. Convincingly demonstrated favourable effects on survival are, from both a clinical and methodological perspective, the most persuasive outcome of a clinical trial. Prolonged PFS/DFS as such, however, is considered to be of benefit to the patient. The choice of primary endpoint should be guided by the relative toxicity of the experimental therapy, but e.g. expected survival after progression, available next-line therapies and the prevalence of the condition must also be taken into account. Irrespective of chosen primary endpoint, it is emphasised that it is the magnitude of the treatment effect on all relevant outcome measures that forms the basis in the benefit – risk assessment.

If PFS/DFS is the selected primary endpoint, OS should be reported as a secondary and *vice versa*.

When OS is reported as secondary endpoint, the estimated treatment effect on OS should be sufficiently precise, to ensure that there are no relevant negative effects on this endpoint, in most cases by showing trends towards superiority. In situations where there is a large effect on PFS, a long expected survival after progression, and/or a clearly favourable safety profile, precise estimates of OS may not be needed for approval.

When OS is reported as primary endpoint, consistency is expected as regards effects on PFS. If foreseen not to be the case, e.g. in case of certain immune modulating therapies, this should be made clear already in the study protocol.

For some conditions, events of progression will be observed at a slow rate making frequent assessments of events of progression a burden to the patients. Event rate at a pre-specified and justified fixed point in time might be used as primary measure in these cases. When event rate at a single point in time is selected for the primary analysis, it is in most cases recommended that all patients should have been on study for that period of time. PFS, in a time to event analysis, and as assessed by the investigator should be reported as a secondary endpoint when a fixed time-point assessment is used as primary outcome measure.

For further methodological guidance as regards PFS, please refer to appendix 1.

It should be noticed that it is expected that the tumour's drug resistance profile is affected by therapy. This might be of relevance for the activity of next-line therapies. This is most obvious if maintenance/prolonged therapy is compared with no treatment or placebo such as in areas where a fixed number of cycles is the standard, for example, first-line ovarian cancer, NSCLC and some haematological conditions. The consequences of progression **on** maintenance therapy, signifying resistance to at least the maintenance regimen, might thus differ from progression **off** therapy. In principle, this applies to all comparisons between different regimens, i.e. the degree of cross resistance as regards next-line therapy might differ between experimental and control regimens.

From a regulatory perspective, this concern has mainly been emphasised in settings where a new concept is introduced such as maintenance therapy or an increased number of "induction" cycles. If at all possible, these studies should therefore be designed with the aim to document patient benefit in terms of survival. If non-feasible, endpoints such as PFS on next-line therapy (PFS 2) should be determined. This should be done within the study so that agreed next line therapy is used after progression in the control and maintenance arms and so that PFS 1 and 2 in the maintenance arm can be compared with PFS 1 and 2 the control arm. In order to capture possible negative effects on next-line therapy and to outbalance tolerability and toxicity concerns related to maintenance therapy, it is expected that PFS2 in the experimental arm is sufficiently superior to PFS2 in the control arm. As the regulatory experience is limited and as methodological issues are foreseeable, EU scientific advice should be considered.

Alternative primary endpoints, such as TTP or time to treatment failure (TTF) might uncommonly be appropriate. This has to be fully justified.

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.

There are also examples where tumour response-related activities, e.g. limb-saving surgery may be reasonable primary measures of patient benefit. Analyses of location- or cause-specific events, however, should in general be avoided as the focus may be drawn away from the main objective, namely the overall success of the treatment strategy in question.

Biomarkers convincingly demonstrated to reflect tumour burden can be used, in combination with other measures of tumour burden, to define tumour response and progression, an example being multiple myeloma and the M-component. For new classes of compounds, however, it has to be demonstrated that the marker is a valid measure of tumour burden and that no bias in the assessment is introduced, e.g. through differential suppression of the tumour marker.

The first line of therapy administered after tumour progression on study drugs should be documented and when feasible further therapy. The putative effects of treatments after progression on study drugs on later events should be discussed in the study report.

7.1.5.1. Secondary endpoints and exploratory analyses

Irrespective of the choice of primary endpoint OS or PFS, ORR and rate of tumour stabilisation for, e.g. 3 or 6 months should be reported. Especially in the palliative setting, HRQoL/PRO using generally accepted instruments might provide valuable information (Appendix 2)

7.2. Treatment administered with curative intent

The ultimate aim of developing new therapies, e.g., in patients with high grade lymphoma, germ cell tumours or in the adjuvant setting, is to improve cure rate and survival or to relevantly decrease toxicity without loss of efficacy. Nevertheless, in some cases and due to the complexity of administered therapies, e.g. in AML, the impact of a relevantly active experimental compound on these endpoints may be hard to demonstrate.

It is foreseen that the experimental compound rarely will be used as single agent therapy, but will be used as add-on to an established, perhaps modified regimen, or as substitution for a compound being part of the established regimen. In this context, maintenance therapy may be regarded as add-on therapy if maintenance therapy is considered non-established.

In the treatment of acute leukemia, lack of achievement of CR, relapse and death without relapse are counted as events in an EFS analysis. Those patients who did not reach CR during the pre-specified induction phase will be considered as having an event at time 0.

In case EFS is found to be a justified primary endpoint, it is of importance that study data are analysed only when sufficiently mature, i.e. when it is foreseen that the EFS plateau is stable or when additional disease recurrence is rare.

In patients with high grade lymphoma or solid tumours, PFS may be used as outcome measure. Not achieving at least PR after a defined period/number of cycles may be regarded as treatment failure in some protocols and only those achieving at least PR continue on therapy. In the primary analysis it is recommended that patients not reaching PR are followed off or on next-line therapy until an event of progression or death is reached.

When improved cure rate is the objective of therapy, it is advised that disease-free survival at a pre-specified time point is used as outcome measure (see above with respect to timing).

7.2.1. Reduced or similar toxicity expected

In most cases, a substitution design is foreseen. From a regulatory perspective, a non-inferiority design is acceptable and in most cases EFS or PFS, as appropriate, are acceptable primary endpoints.

In cases where induction is followed by consolidation and/or maintenance therapy, confounding effects of therapies administered after the end of experimental therapy may make endpoints other than PFS or EFS more appropriate. This means that CR (and CR + PR, if specifically justified) after end of experimental therapy could be an acceptable primary endpoint when further therapy is scheduled. In these cases, the possible influence of the experimental compound on the activity of consolidation therapy should always be addressed and outcomes with respect to CR should be supported by EFS or PFS data.

It is recommended that CR is defined according to established clinical criteria, but supportive evidence in terms of Minimal Residual Disease (MRD) as defined by molecular criteria should be sought when applicable. MRD data, however, should only be used after proven intra- and inter-laboratory validation.

7.2.2. Increased toxicity expected

Substitution or add-on designs may apply. In most cases, superiority in terms of EFS, PFS, or OS as appropriate, should be demonstrated and the benefit in terms of prolonged time to event should be sufficiently large to balance increased toxicity.

A major increase in CR after induction therapy associated with trends in PFS or EFS, and survival, however, might be sufficient if scheduled treatments administered after the end of the experimental therapy are likely to confound overall outcome. This is of special relevance if the target population is small.

7.2.3. Major increase in toxicity expected

The aim should be to demonstrate increased cure rate or improved OS. In some cases, such as in small study populations, a major increase in EFS or PFS, as appropriate and supportive data compatible with a favourable trend on survival might be sufficient.

7.3. Treatment administered with the intent to achieve long-term disease control

Typical conditions include early lines of therapy in advanced breast cancer, colorectal cancer, low-grade lymphomas and the chronic leukaemias for which established reference therapies are available and next-line treatment options are likely to be meaningfully efficacious.

7.3.1. Reduced or similar toxicity expected

Substitution or single agent studies are foreseen. From a regulatory perspective, a non-inferiority design is acceptable and PFS is considered an appropriate primary endpoint. In case of relevantly reduced toxicity, mature survival data may be submitted post licensure if justified by study data.

7.3.2. Increased toxicity expected

The aim should be to demonstrate superiority at least in terms of PFS.

Survival data should be made available at the time of submission. It is acknowledged that mature survival data cannot be expected in all cases, though a justification explaining why this is the case should be provided. Post approval follow-up with respect to survival is expected in these cases. If absence of an increase in treatment-related mortality is not established with reasonable certainty, mature survival data should be available for the assessment of benefit – risk prior to licensure.

It is acknowledged that alternative endpoints may be more appropriate in certain situations, e.g. when maintenance therapy is investigated in areas where this has not established (Endpoints, 8.1.5). The aim may also be to enable a long treatment-free interval after intense induction therapy.

7.3.3. Major increase in toxicity expected

The principal objective should be to demonstrate improved survival.

In individual cases this might be non-achievable due to expected good prognosis with respect to survival and availability of several active next-line regimens, including experimental therapies, at the time of disease progression and a small target population. If PFS is the selected primary endpoint for the study, this requires a thorough justification. A careful discussion at the planning stage is also needed for the assessment of possibly therapy-related fatalities. Even though only a major benefit in terms of PFS prolongation would be acceptable, whenever possible the number of patients included should be sufficient to obtain an estimate on overall survival where a trend in a favourable direction is expected.

7.4. Palliative therapy

In the context of this appendix, this mainly refers to last line settings where the prognosis for survival is poor and where it might be problematic to identify sufficiently documented reference therapies. In other cases, patients are considered not suitable for intensive, potentially curative therapy as defined by clear and as far as possible unambiguous criteria.

In cases where there is no established reference therapy, investigator's best choice or BSC are acceptable.

In a study conducted with BSC as reference therapy, the objective should be to demonstrate prolonged OS and/or globally improved symptom control or quality of life (QoL). The latter requires that all efforts are undertaken to reduce possible bias (Appendix 2) and that the treatment is well tolerated. If the reference regimen is known to be active, but not established, superiority in terms of PFS might be acceptable. In these cases, the following will be taken into account in the benefit – risk assessment: the evidence showing activity of the reference therapy, the magnitude of the PFS benefit over the reference regimen, the tolerability/toxicity profiles and the prevalence of the condition.

It is acknowledged that patients may be considered suitable only for palliative therapy at baseline due to, e.g. poor performance status, but may respond so well that further therapy can be administered with curative intent, including, e.g. reduced intensity HSCT. How to handle these patients should be defined in the analysis plan.

7.5. Special considerations

7.5.1. Haematopoietic stem cell transplantation, methodological considerations

If allogeneic haematopoietic stem cell transplantation (HSCT) is a foreseeable treatment option, it is of importance to define how transplantation should be handled in the analysis plan. It is fully acknowledged that criteria for HSCT (e.g. patient eligibility, HLA matching, conditioning regimen, graft versus host disease prevention, etc) vary between institutions and regions. Nevertheless, these criteria should be defined as far as possible in the protocol and reasons for performing or not performing HSCT should be captured by the CRF.

Even though transplant related mortality is an issue and long-term benefit need prolonged follow-up, it is normally expected that patients undergoing HSCT are followed for OS and EFS as randomised. Patients may be censored at time of conditioning for HSCT as a sensitivity analysis.

As treatment administered prior to transplantation might affect outcome of HSCT, proportion of patients undergoing HSCT is not considered to be a suitable primary outcome measure even if all patients responding sufficiently well to treatment are scheduled for transplantation.

Autologous stem cell transplantation constitutes less of a concern from an assessment perspective and may be viewed as intensified consolidation therapy where the consequences on short-term mortality and possible long-term benefit are less pronounced than after HSCT. Nevertheless, heterogeneity in the conduct of autologous transplantation should be avoided as far as possible, but censoring should normally not be undertaken.

With respect to drug development specifically in relation to HSCT, please refer to 9.5.

7.5.2. (Neo)adjuvant therapy

In the adjuvant setting, the ultimate aim is to increase cure rate. While effects on DFS are considered relevant to the individual patient, it is of importance to consider in the planning of the study whether it is at all possible to demonstrate a favourable effect on cure rate, i.e. in analyses conducted when recurrence rates have reached an apparent plateau.

As the use of adjuvant therapy may limit therapeutic options at time of recurrence OS data should be reported. For established areas of adjuvant therapy, e.g. breast and colorectal cancer, and if benefit-risk is considered favourable for the experimental regimen based on DFS and available safety and survival data, mature survival data may be reported post-licensing. In some cases and due to toxicity concerns, favourable effects on OS have to be demonstrated.

The objectives of neoadjuvant therapy may include improved overall outcome and organ preservation (e.g. more conservative surgery). If organ preservation is the main objective, at least non-inferior DFS/PFS should be documented. As for adjuvant therapy, a defined number of cycles is frequently administered. Pending on the objectives of the study it is accepted that treatment is withdrawn if tumour shrinkage is not observed after a defined treatment period.

When pathological CR at time of surgery is reported as secondary endpoint, patients withdrawn should be considered as non-responders. Major increase in pathological CR over established therapy in high risk patient such as those with inflammatory breast cancer, might be indicative of patient benefit if not associated with major increase in toxicity.

7.5.3. Drug resistance modifiers, chemoprotective agents and radio/chemo sensitisers

In principle, the design of confirmatory studies for experimental drug resistance modifying agents and radio/chemo sensitizers (A) is straight forward; AB should be demonstrated to be more active than an established regimen (B) in terms of anti-tumour activity and the benefit – risk for the combination should be shown to be favourable. If there are PK interactions, or dynamic interactions not related to anti-tumour activity, dose adjustments of B in the combination arm might be needed in order to make the comparison AB vs. B at similar overall toxicity. If the full effects of the PK interaction is captured by changes in the plasma levels of B (e.g. no changes in distribution), however, dose adjustments of B in order to compare AB vs. B at similar exposure of B is preferred.

For a chemoprotective agent, it has to be shown that normal tissues are more protected from toxicity than tumour tissue. For most cytotoxic compounds, it is, however, easier to detect dose-related differences in toxicity than in efficacy. This means that in many cases very large studies are needed with tight confidence intervals around measures of anti-tumour activity in order to prove that normal tissue protection is achieved without loss of anti-tumour activity. Co-primary endpoints are thus needed, testing the hypotheses of improved safety and non-inferior anti-tumour activity. In some cases, it might actually be easier to convincingly demonstrate differential tissue protection by increasing the dose of the cytotoxic compound in the experimental arm aiming to show enhanced anti-tumour activity without increased toxicity.

However, if it can be shown conclusively that there is no PK interaction and that the chemoprotective compound cannot interact with the tumour, e.g. by absence of target in tumour cells, it might be acceptable only to show reduced toxicity without formal non-inferiority testing of tumour protection.

7.5.4. Tumour Prevention

Regulatory experience is limited, but conceptually the situation is rather similar to the adjuvant setting. Thus individuals at risk should be defined so that the observed risk reduction in tumour incidence outweighs the side effects of therapy. As tumour prevention may select for tumours with altered biological behaviour, comparative data on tumour pheno/genotype are expected and data on OS may be needed. In the planning of these studies, regulatory scientific advice is recommended.

7.6. Methodological considerations

Frequently, only one single study is foreseen for a specific indication. Licensing based on one pivotal study, however, requires demonstration of efficacy at levels beyond standard criteria for statistical significance (CPMP/EWP/2330/99). This is of special relevance in non-inferiority trials, in trials with PFS as primary endpoint and in a comparison with BSC/investigator's best choice. It is acknowledged that supportive evidence from confirmatory studies conducted in other indications should be taken into account in the assessment. The supportive value of these studies might vary and a discussion is expected as regards the relevance of these findings in relation to the application for the new indication.

7.6.1. Interim analyses

Interim analyses are frequently undertaken in Phase III trials, but early stopping whether for futility or superiority is a sensitive issue. Early stopping for superiority requires an assumption of proportional hazard, i.e. that the treatment effect in patients with rapidly progressing tumours is similar to that in less aggressive tumours in the absence of data demonstrating that the magnitude of effect is maintained.

If a clear majority of the total number of expected events in the long term has been observed and a difference has been documented, this is normally accepted as an indicator that the study is reasonably mature and that the study results will remain stable over prolonged follow-up. The interpretation of interim analyses conducted on a less mature data set is problematic.

In cases where the treatment effect has been underestimated in the planning of the study, this may create a dilemma if statistically convincing effects in terms of overall survival have been demonstrated before a representative and mature dataset is available. Other monitoring committee decisions might be investigated in this instance such as restricting the continuation of the trial to the under-represented subsets to which the observed effect cannot be extrapolated. Analyses according to stratification factors of major importance for prognosis might provide insights.

In general, interim analyses based on PFS data are not encouraged (Appendix 1).

7.6.2. Time to event analyses and assessment of response and progression

For studies with PFS/DFS as primary endpoint, symmetry with respect to imaging and study visits is pivotal and adherence to protocol-defined schedules is essential and deviations should be reported (Appendix 1).

As discussed above (Exploratory trials with time-related endpoints), a comparison in terms of PFS between a predominantly tumour shrinking compound and a predominantly growth inhibiting compound may "favour" the latter compound with respect to tumour burden at time of progression.

Until now, there is no regulatory experience with respect to comparisons with clearly discordant outcomes in terms of ORR and PFS and there are no established ways to adjust for this. If exploratory studies indicate that this might become the case, alternative endpoints such as OS should be considered.

Differences in mode of action between the experimental and reference therapy might generate problems in relation to measurements of tumour burden and anti-tumour activity, one example being early tumour swelling as discussed previously. Whenever such problems are foreseen, which may require deviation from standard approaches (RECIST, WHO), it is recommended that agreement is reached with regulatory agencies prior to the initiation of pivotal trials. Similarly, if tumour assessment techniques cannot be used that allow for independent adjudication, it is advisable to discuss available alternatives with regulatory agencies.

Pseudo-response should always be considered a possibility when tumour related oedema is an issue such as in high grade gliomas. Updated response and progression criteria taking this into account should be applied when available. If such criteria has not yet been established, scientific advice is recommended in order to discuss alternative ways forward.

7.6.3. Non-inferiority studies

Guidance of design, conduct and analysis of non-inferiority studies is given in other regulatory guidance documents (Choice of a Non-Inferiority Margin CPMP/EWP/2158/99), but some topics deserve particular attention in the oncology setting. For a PFS endpoint, which can be considered a composite endpoint, the discussion of a non-inferiority margin should consider the effect of the reference treatment overall but inference should also include a discussion on each type of progression (local, distant, clinical, radiological etc.) including description of the effect of the reference regimen on each component. If differences in the profiles of progressive disease in each treatment arm might be expected, this should be accounted for in the planning stage with a suitably conservative margin and appropriate sample size to obtain the required number of events for reliable inference.

Given the importance of study sensitivity (i.e. the ability of a trial to detect differences) for the assessment of non-inferiority trials, where similar activity is assumed for test and reference, it is of importance to plan in advance for a subgroup analysis excluding patients with poor prognostic factors at baseline such as poor PS, co-morbidities, etc. as in these patients it might be harder to detect a difference in activity between treatment regimens, if there were one. Similarly a per protocol analysis set should be defined so that protocol violations, compliance problems, etc. do not reduce the possibility to detect a difference. These analyses are expected to be undertaken with the aim to show consistency.

7.6.4. Analyses based on a grouping of patients on an outcome of treatment

Comparisons of time-to-event variables (like OS, or PFS) by grouping patients on a post-randomisation outcome of treatment are problematic. Since outcomes like tumour response, dose intensity, toxicity, or compliance represent an interaction between therapy, patient and tumour the contribution of therapy cannot be disentangled. Nevertheless, certain unexpected outcomes such as clearly improved survival despite dose-reduction due to toxicity, or absence of prolonged survival in responding patients might be informative. A search for unexpected findings constitutes a rationale for conducting these exploratory analyses.

Response duration comparing groups of patient on different therapies may be regarded as informative. Data should be reported with confidence intervals for the individual study arms, but significance testing comparing duration of response between study arms should not be undertaken as the comparison refers to groups that are not fully randomised. “Time in response” where patients without response are assigned a duration of zero enables a statistical comparison between study groups.

7.6.5. Studies in small study populations, very rare tumours

For some truly rare tumours or very narrow indications, whether due to tumour phenotype or restrictions related to target expression, it is simply not possible to recruit a sufficiently large number of patients to conduct reasonably powered, randomised studies in order to detect clearly relevant differences in anti-tumour activity. In some cases a small, randomised, reference controlled study is

the best option, in other cases a within-patient TTP/PFS analysis (or the combination) might be a better alternative. In the latter case, TTP on last prior therapy is compared with time to progression or death on the experimental therapy. This would require that the clinical appropriateness of the last administered therapy prior to study therapy and progression on prior therapy is independently adjudicated and that the study protocol clearly defines the proper conditions for the analysis. Superiority should be demonstrated.

Problems related to studies in small populations are further discussed in the Guideline on clinical trials in small populations (CPMP/EWP/83561/2005). In these small target populations all evidence with respect to efficacy and safety must be taken into account. This encompasses clinical response rate, duration of response as well as outcome measures such as HSCT rate, use of minimal residual disease (MRD) to define response rate and recurrence of disease, as appropriate. Mature time to event endpoints such as PFS and OS should be reported even though it is acknowledged that formal statistical significance cannot always be expected, even if the experimental compound is relevantly more efficacious.

As there is no general solution to the problem of how to document benefit – risk in these cases, scientific advice is recommended.

7.7. Special populations

7.7.1. Elderly and frail patients

In many indications elderly patients represent the majority of the patient population. In these cases it is expected that the study data base makes a benefit – risk assessment possible in the elderly.

Some compounds may be specifically suitable for the treatment of elderly, e.g. due to PK properties such as low sensitivity to impaired organ function. In these cases, dedicated studies in the elderly are encouraged. It is acknowledged that it may be hard to identify appropriate reference therapies in some of these cases and that other outcome measures than PFS/OS might become more relevant. In these cases it is advisable to seek regulatory agreement on the development program.

Frail patients, whether elderly or not, with clearly impaired performance status (PS) constitute a vulnerable group of patients rarely included in conventional confirmatory studies. Clinical studies in this group of patients are encouraged from a regulatory perspective.

7.7.2. Children

See *Addendum* (CPMP/EWP/569/02 under revision).

7.7.3. Gender

For some tumours and/or therapies, a difference in antitumor activity related to gender has been reported. Where a priori it is likely that there may be a treatment by gender interaction, this should be taken into account in the design of the study. Otherwise it is expected that the proportion of females and males reflects the prevalence of the disease and that the sponsor provides exploratory subgroup analyses (efficacy and safety) by gender.

7.7.4. Patients with impaired organ function

Please refer to Section 5, Pharmacokinetics.

7.8. Safety

In addition to standard reporting of adverse events, it is expected that effects of preventive measures, such as anti-emetics or use of growth factors are delineated. Acute, sub-acute, chronic and late toxicities should be described. Safety in special populations, as detailed above, should be summarised from the full studies programme.

For common events, safety in relation to treatment cycle, first, second, third etc., is of value. Similarly, timing and duration, including grade, of some events such as diarrhoea, mucositis, or cytopenias should be reported.

Monitoring of frequency and type (viral, bacterial, fungal) of possible, probable or proven infections should be undertaken in patients undergoing more intensive cytotoxic/immunosuppressive therapy. For compounds known or suspected to cause long term immunodeficiency, monitoring for opportunistic infections for up to one year after the end of therapy should be considered.

Cumulative toxicity should always be investigated.

If cure is the objective, long term follow up for toxicity is highly relevant. Late toxicity includes secondary malignancies and certain organ toxicities (e.g. CNS, cardiovascular). The number of patients suffering from late toxicities may increase over time and is therefore an objective for post licensure pharmacovigilance activities.

As radiation therapy is a standard treatment option in malignant tumours, it is foreseeable that patients will be receiving radiation therapy, e.g. for symptom palliation, concomitantly with or in a time frame close to administration of the medicinal agent. Safety information on concomitant or sequential use of the medicinal agent with radiotherapy should be collected throughout the entire study programme, including data on “radiation recall”.

In haematological malignancies, bone marrow failure is often a presenting symptom and is frequently aggravated by treatment. In contrast to the approach in solid tumours, dose reduction for this reason is often not indicated, in particular if the aim is curative.

If the aims of the study include demonstration of improved safety, the protocol should specify how this should be accomplished. It is not acceptable to focus on one toxic effect only. The outcome measure(s) should provide unbiased information on overall toxicity and tolerability, perhaps in addition to a specific item such as neuropathy where a clinically relevant improvement is expected. As there is limited experience with this type of studies, EU regulatory advice should be considered.

Where appropriate, pharmacogenomics may be used to identify patients at increased risk for severe toxicities.

8. Condition specific guidance

8.1. Non-small cell lung carcinoma

NSCLC is a leading cause of cancer morbidity and mortality. Most patients diagnosed with NSCLC present with advanced disease and many of the patients who do present early will go on to develop metastatic lung disease. Common disease related symptoms include pulmonary effects (cough, dyspnoea) and general symptoms of pain, anorexia and high degrees of psychological distress.

Recent developments in the knowledge of NSCLC biology have uncovered targets for therapeutic agents, creating new opportunities but also adding complexity to the interplay between potential biomarkers and drug candidates and consequently, to the assessment of their value in the management of this disease
These factors warrant a specific guidance for the assessment of medicinal agents directed at the management of NSCLC in the context of the present guideline. Namely, criteria, definitions, and other reflections are provided for the use of biomarkers, the systematization of therapeutic phases in the course of the disease, and the endpoints applicable to the assessment of clinical benefit.

Classification of NSCLC
NSCLC must be classified using pathological and molecular features. The importance of consistent, accurate and reproducible histological subtyping cannot be understated.

Pathological evaluation using internationally agreed criteria should determine the histological classification (WHO Classification) and the extent of the disease (UICC TNM Classification). Immunohistochemical analysis may improve pathological diagnosis, particularly for small biopsies. Pathological evaluation should also determine the molecular features of the tumour and this must include EGFR status, presence of k-ras mutation, level of expression of ERCC1, RRM1 and thymidylate synthase and the presence of ALK translocations.

Stratification according to disease and patients characteristics

Exploratory trials should clearly test hypotheses of activity in accordance with known or presumed biological roles of their intended molecular targets. For this purpose, trial subjects must be constituted by patients with disease that is well characterized according to relevant biomarkers. Subsequently, the same applies to confirmatory trials which must restrict inclusion to categories of patients with clinical and molecular characteristics that increase the likeliness of response and hence clinical benefit.

It is particularly important to perform specific trials, or at least to stratify patients based on baseline characteristics such as tumour histology and expression of predictive molecular biomarkers. Such markers help delineate distinct disease entities, enriching the patient population to those with the target of interest and defining subsets of patients most likely to benefit from therapy. However, the success of such an approach depends heavily on having an accurate diagnosis.

At least a third of lung cancer patients are 70 years or older, older patients should be actively recruited into clinical trials. Other variables such as smoking status and geographical origin should also be considered in the recruitment of patients.

Treatment definitions

Adjuvant or neoadjuvant therapy may improve survival in certain groups of patients by decreasing the risk of metastatic disease. For adjuvant therapy, patients should generally be relatively young without significant co-morbidities who have undergone complete resection by lobectomy. The tolerability of any adjuvant therapy must be considered. Neoadjuvant therapy may reduce tumour volume, control micrometastasis and if adequate tumour samples are obtained may provide valuable information regarding tumour response and tumour biology.

The concept of maintenance therapy should be considered for well tolerated medicinal products and a maintenance approach may represent an effective way of delivering second line therapy. Maintenance therapy is the prolongation of treatment at the end of a defined number of initial treatment cycles following tumour control (tumour response or stable disease). Continuation or true maintenance therapy refers to the continuous administration of at least one of the agents given in first line therapy (either at the same intensity or at a lower intensity). Switch maintenance or early second line therapy refers to the immediate administration of a different agent not included as part of the first line regimen following completion of therapy.

Efficacy endpoints

For exploratory studies, ORR is an acceptable endpoint for early evaluation of new medicinal products in NSCLC, though modest response rates may in fact underestimate patient reported benefits. In light of this, endpoints which capture clinical benefit and record palliative control (pain control, weight loss, performance status) may be included in the study design. Prognostic and predictive molecular markers and mechanisms of resistance should be actively investigated.

Improving survival remains the principal objective for patients with NSCLC and in many cases OS should be selected as the primary endpoint for confirmatory studies. If, however, the experimental regimen is likely to be well tolerated, PFS benefit might enable a proper benefit – risk assessment, especially if supported by data on HRQoL/PRO (Appendix 2).

For maintenance studies, if conducted versus placebo/BSC, the recommended endpoint is OS (8.1.5).

8.2. Prostate cancer

The proper design of prostate cancer studies is a challenge since there are several complicating issues.

Firstly there is a large variability in the biology of prostate cancer. Almost every man will ultimately develop prostate cancer, the majority being slowly progressive, but some are aggressive with fatal outcome. There is thus a risk related to the detection of indolent tumours and a challenge to identify clinical significant prostate cancer of importance to treat. Treatments with curative intent include surgery and/or radiotherapy but active surveillance is an alternative and reduces the risk for overtreatment and side effects related to radical therapy.

Secondly there are to date no method to properly quantify the tumour burden, making it difficult to interpret therapy outcome. Imaging techniques such as computed tomography (CT), magnetic resonance imaging (MRI), radionuclide imaging and positron emitting tomography (PET) with different traces are less suitable to estimate bone disease and soft tissue metastases are uncommon clinical presentation of prostate cancer.

Prostate specific antigen (PSA) is not cancer specific but changes in PSA levels during different therapies are used as a biomarker. Individuals' PSA values are not comparable to each other but changes and nadir are prognostic.

Prostate cancer is diagnosed on histopathology of core biopsies, but the likelihood to detect a cancer is dependent on number of biopsies, the prostate volume and the cancer location (anterior cancer and cancer located near the urethra is difficult to biopsy using transrectal technique).

Cancer prevention studies

The recommended primary outcome measure in prostate cancer prevention trials is disease free survival or the rate of diagnosed prostate cancer at a predefined point in time.

Baseline risk factors of likely prognostic importance include age, ethnicity, family history of prostate cancer, serum PSA, normal/abnormal digital rectal examination or transrectal ultrasonography.

It is crucial to have identical diagnostic procedure between active and placebo groups in order to avoid sampling bias and long observation periods are needed as both the induction period and the latency period to detect a prostate cancer are long. Even small differences in management between the treatment groups may harbour confounding factors of importance. It is also crucial to assess the clinical relevance of the diagnosed cancer, i.e. the diagnosed cancer should be clinically significant.

Stage, Gleason score and PSA level are regarded as the most appropriate prognostic factors of outcome of new diagnosed prostate cancers.

Minimally invasive treatment

Since available treatment options with curative potential are associated with side effects that interfere with health related quality of life, a concept of minimal invasive treatment has been introduced. The aim is to delay or avoid the need for, e.g. surgery using techniques and/or medicinal compounds that offer low risk of side effects.

As a first step, anti-tumour activity has to be proven. This may be achieved in trials using subjects planned for radical surgery where one lobe containing cancer is treated with the minimally invasive concept before radical surgery.

For confirmatory trials, an acceptable primary end point is time to need for radical therapy, or proportion of patients in need for such therapy at a predefined point in time. Until now, however, there is no consensus as regards criteria defining need for radical therapy. Clinical guidelines developed by European Urology Association (EAU), National Cancer Comprehensive Network (NCCN) and National Institute for health and clinical excellence (NICE) suggest several options. This unfortunate situation is acknowledged; nevertheless clear criteria defining need for radical therapy should be in place in study protocols, especially if the study cannot be conducted under double-blind conditions. Independent adjudication is recommended.

PROs and genitourinary function preservation should be reported as secondary endpoints.

Prognostic factors of relevance in the planning of the study include: age/life expectancy, disease stage, Gleason score and PSA.

Neoadjuvant and Adjuvant therapy

As more treatment options become available in the metastatic setting, more trials are expected also in the (neo)adjuvant treatment.

Adjuvant treatment using hormones has been proven effective in patients receiving radiotherapy or surgery in terms of improved progression free survival; however adjuvant androgen deprivation has improved overall survival only for patients receiving radiotherapy. Neoadjuvant hormonal treatment prior to radiotherapy improves progression free survival but prior to surgery hormonal treatment only reduces the number of positive surgical margins without any favourable outcome on progression free survival.

The definition of progression-free survival is usually based on PSA, and differs between radiotherapy and surgery groups. After successful surgery the PSA levels is immediately <0.2 ng/ml and a commonly used definition of relapsed disease is any measurable PSA levels above 0.2 ng/ml confirmed by two consecutive measures. But after successful radiotherapy a decrease in PSA is observed over several months not always reaching levels <0.2 ng/ml.

There have also been cases of demonstrated “PSA bounce” in patients proven relapse-free with long-term follow-up. This type of PSA kinetics after radiotherapy has urged for a consensus and a definition of relapse after radiotherapy is an increase from nadir of 2.0 ng/ml (RTOG-ASTRO criteria Phoenix).

It is acknowledged, however, that there is an ongoing debate on how to best define relapse. Irrespective of this, criteria defining progression should be clearly stated in the protocol. PSA measurement and any other clinical assessment should be done at the same pre-specified time-point in experimental and control groups. The rate of locally and systemic failure should be reported separately.

Therapy for locally advanced disease

No consensus regarding the definition of locally advanced disease has been reached, but the term often refers to either a bulky tumour with growth outside the prostate capsule (T-stage 3-4) based on per rectal assessment, or a tumour that express several high-risk factors indicating a more advanced tumour stage. Common is the absence of distant metastases; however this is a function of which diagnostics is performed.

The protocol should define methods to be used to exclude distant metastases. Digital rectal examination is still considered the most appropriate method to assess local progression. If studies cannot be conducted under proper double blind conditions, examination by two independent urologists is recommended. Response criteria are otherwise similar to those for metastatic disease presented below.

Distant metastases-free survival, PFS including local progression, genitourinary function and validated PRO questionnaires constitute relevant outcome measures.

Therapy for metastatic disease

Hormone naive

During more than 60 years the treatment of choice in metastatic prostate cancer has been androgen depletion therapy. More than 90% of the cancers are androgen dependent, but eventually the disease becomes castration refractory. Currently androgen depletion is often introduced in the adjuvant setting or at PSA relapse without detectable metastases. The first sign of castration refractory state is often detected as PSA increase despite S-testosterone at castration levels.

Several definitions have been discussed, but a consensus has been reached during the work of The Prostate Cancer Clinical Trials Working Group (PCWG2). The PCWG2 proposes that subjects should be categorised according to rising PSA state (non-castrate or castrate) and the occurrence of clinical detectable metastases (non-castrate or castrate) throughout the natural prostate cancer history.

It is foreseen that active medicinal agents in late castration refractory state of prostate cancer will challenge the use of androgen depletion therapy in order to avoid the symptoms associated with castration treatment.

The use of anti-androgens provides an additional treatment option in the hormone naive status. The anti-androgens treatment has both a direct effect and a withdrawal effect. This has to be taken into account when designing clinical trials and it is often stated that anti-testosterone treatment should have been removed at least 4-6 weeks before inclusion to avoid PSA decrease from withdrawal effect.

For medicinal products aiming at achieving medical castration, it is sufficient to convincingly demonstrate this while for non-hormonal products to be used as add-on or instead of, it is expected that favourable effects on PFS (see below) or OS are demonstrated.

Castration refractory

In the castration refractory state of the disease, there is still some hormonal treatment available including CYP-17 inhibitors, anti-androgens, oestrogens and corticosteroids before the disease is classified as androgen refractory. Androgen depletion should continue during the disease course as androgen sensitive clones are assumed to prevail.

It is important to emphasise that androgen-independent prostate cancer is a heterogeneous group of disease and today known prognostic factors include: Gleason score, PSA levels and kinetic, tumour stage at diagnose (including bone only, nodal visceral spread), primary treatment, time to relapse, duration of androgen depletion therapy, time to castration refractory disease, time with clinical detectable metastatic disease, use of cytotoxic and the response. Additionally, general performance status, age and co-morbidity are important prognostic factors. From this perspective, it is advisable to consider whether it is more informative to conduct separate studies in high and low risk patients.

The evaluation of response is performed according to RECIST criteria when soft-tissue metastases are detectable. However, prostate cancer is characterised by osteblastic bone metastases not suitable to assessment according to RECIST. Therefore the relevance of new bone lesions as a marker for progressive disease is emphasized. However, subclinical lytic bone lesion successfully treated may firstly responds with an osteoblastic reaction before restitution. Specifically for bone scan it is also of importance to consider uptake caused by trauma and other benign conditions such as osteoporotic fractures. Medicinal compounds acting as inhibitors of osteoblast activity may confound the assessment of disease activity by bone scans.

Progression in bone metastases is often accompanied by PSA increase. PSA increase may thus be taken into account in the definition of progressive disease based on imaging, although PSA increase alone cannot serve as primary end point in confirmatory studies. PSA can even decrease in progressive late castration refractory state due to a dedifferentiation of the cancer cells making them unable to produce PSA.

Currently a large number of new medicinal products are under late clinical development or have recently been marketed. Guidance is therefore not provided as regards suitable reference therapies in patients with castration resistant tumours.

Time to symptomatic progression, PFS and OS are considered appropriate outcome measures and the overall guidance provided in the general section apply.

8.3. Chronic Myeloid Leukaemia

CML is uniquely well characterised among human malignancies with respect to underlying molecular cause, course of disease, response to BCR-ABL tyrosine kinase inhibitors (TKI) and molecular events causing drug resistance. Due to the continuous scientific advance in this field it is of major importance to follow the progress with respect to standardisation of laboratory techniques used in the assessment of the disease. Generally acknowledged clinical diagnostic and treatment guidelines should also be followed and CHMP regulatory advice is recommended particularly when new diagnostic techniques or treatments emerge.

The diagnosis and stage of the disease should be well documented in any clinical study. Diagnosis of CML should be based on investigation of full blood count (FBC), bone marrow, cytogenetics and real time quantitative reverse transcriptase (RQ-PCR) for BCR-ABL transcripts.

When assessing the response to treatment there are three aspects that should be evaluated:

1. Haematological response
2. Cytogenetic response
3. Molecular response

The degree and timing of haematologic, cytogenetic and molecular responses provide very important prognostic information as time-dependent variables. Additionally, other prognostic scores such as age, spleen size and FBC should also be considered when defining high risk groups. The Sokal and Hasford scores are considered validated predictors of response in newly diagnosed patients.

Current international practice guidelines classify response to first line standard treatment (imatinib) into three categories and this approach including future updates should in general, be followed. An example as described by the ESMO is shown in Table 1. Other international practice guidelines such as those provided by the US National Comprehensive Cancer Network may also be acceptable. For newer drugs whose response may be faster, landmarks and standards of success and failure may need to be reassessed.

Table 1 Definition of response to imatinib

	Optimal	Suboptimal	Failure
3 months	CHR	<CHR	No HR
6 months	≥PCgR	<PCgR	No CgR
12 months	CCgR	<CCgR	<PCgR
18 months	≥MMolR	<MMolR	<CCgR
Any time	No response loss Mutations[a]	Loss of MMolR Loss of CCgR	Loss of CHR Mutations[b]

CHR, complete haematological response (WBC <$10x10^9$/l, differential with no immature granulocytes and <5% basophils, platelet <$450x10^9$/l, spleen non palpable);
PCgR, partial cytogenetic response (Ph+ metaphases 1%–35%); CCgR, complete cytogenetic response (Ph+ metaphases absent);
MMolR, major molecular response (BCR-ABL:ABL <0.10% by International Scale, on RT-Q-PCR).
[a]BCR-ABL KD mutations still sensitive to imatinib.
[b]BCR-ABL KD mutations still insensitive to imatinib.
[Chronic Myeloid Leukaemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow up; *Annals of Oncology* 21 (supplement 5); v 165-167, 2010]

Monitoring the therapeutic response and level of residual disease is essential and the following guide is recommended. However, if responses with a new therapeutic agent are more rapid testing at more frequent intervals may be required.

1. During the first 3 months clinical, biochemistry and haematological monitoring should be assessed every 2 weeks.
2. From the third month on:
    - cytogenetics (chromosome banding analysis of marrow cell metaphases) should be performed at least every 6 months until a complete cytogenetic response has been confirmed
    - RT-Q-PCR (BCR-ABL:ABL % on blood cells) should be performed every 3 months until a major molecular response is confirmed.
3. Once a complete cytogenetic response and major molecular response have been confirmed:
    - Cytogenetics every 12 months
    - RT-Q-PCR every 6 months

Screening for BCR-ABL KD mutations will be expected in cases of failure or suboptimal response.

Measuring drug concentration in blood may be required in some cases, such as failure, suboptimal response, dose-limiting toxicity and adverse events.

More frequent monitoring may be advisable in certain cases, for example when studies are conducted on a high risk population.

It is recommended that monitoring will take place in specialised central laboratories.

Whenever possible, it is expected that the mechanisms contributing to the lack or suboptimal response will be explored and may include the following:

- Mutations in the BCR-ABL kinase domain
- Clonal evolution, defined as the presence within CML cells of additional translocations that are thought to drive disease progression
- Pharmacokinetic variability (poor compliance, drug interactions, variability in metabolic enzymes etc)
- Amplification of the BCR-ABL fusion gene
- Overexpression of drug transporter genes and tyrosine kinases such as the SFKS
- Toxicity leading to dose interruptions or reductions

Chronic Phase (CP)

More than 90% of patients are diagnosed in CP.

As there are currently several medicinal products approved for the treatment of CML in CP a comparative trial should be undertaken against a licensed reference product.

If the aim is to show superiority versus a licensed comparator the recommended primary endpoint is major molecular response at 18 months. Appropriate secondary endpoints include complete cytogenetic response at 12 months, PFS and overall survival. Long term follow up of at least 8+ years is expected.

In the case of non-inferiority trials, a longer follow up will be required in order to evaluate the primary endpoint and major cytogenetic response after at least 2 years is recommended .

In patients failing a licensed TKI, studies may be undertaken in all patients fulfilling established criteria for non-response or secondary failure; alternatively patients may be enrolled also taking into account mutation patterns if properly justified.

When studies are conducted in special groups such as patients intolerant to prior TKI therapy, resistant to prior treatments (primary or secondary resistance), high risk patients or with new secondary mutations baseline characteristics should well defined before enrolment. Symptoms and signs defining intolerance to the prior TKI should be documented in detail (including grading) prior to inclusion in the study. As class related adverse reactions are common, it is of importance that “cross-intolerance” is excluded as objectively as possible due to the subjective nature of “intolerance” in many cases.

It is acknowledged that mutation analysis remains an essential assessment for patients in second line treatment and beyond. Enrolled patients should be well characterised with respect to secondary mutations and an important aim is to confirm activity in relation to relevant mutations. If justified by data, patients with certain mutations associated with low activity for the experimental compound may be excluded, but this will be reflected in the labelling.

If patients with increased risk of efficacy failure to TKIs are identifiable at baseline, it is foreseen that add-on studies with a non-TKI that is active in patients with CML may be undertaken. Superiority should be demonstrated comparing the combination regimen with a single TKI. In studies exploring the combination of two TKI the potential of additive toxicity should be fully addressed.

In cases where the target population may be small, for example patients who have no other available treatments, EU regulatory advice is recommended prior to the initiation of phase II/III trials.

Advanced disease (Accelerated Phase, Blast Crisis)

It is foreseen that the vast majority of these patients have been treated with a TKI.

For those patients that are on accelerated phase (AP) but had prior treatment for chronic phase a trial versus another TKI may be conducted if possible. In the case presentation at diagnosis is accelerated phase without prior chronic phase a trial versus a first line TKI will be expected. In general, as treatment on AP depends on type of prior therapy the comparator used will be defined by prior patient treatment history.

Patients on blast crisis receive conventional chemotherapy with or without allogeneic SCT. Due to the rarity of blast crisis and the foreseen complexity of the therapeutic situation, EU regulatory advice should be considered.

8.4. Myelodysplastic Syndromes

Myelodysplastic Syndromes (MDS) are a heterogeneous group of malignant clonal disorders which share two main features, i.e., progressive cytopenia and risk for transformation to AML. Until recently, supportive care, low dose Ara-C, intensive chemotherapy or HSCT were the only available treatment options. HSCT is potentially curative, but poses high mortality risk in the predominantly elderly MDS population. Supportive care options include blood transfusions, antibiotics, erythropoietin (EPO) and granulocyte colony-stimulating factor (G-CSF).

Diagnosis and Classification of MDS

Many patients with MDS are asymptomatic at the time of diagnosis, but eventually develop symptomatic anaemia, thrombocytopenia and neutropenia alone or in combination. The clinical course is highly variable and several classification systems have been developed, including FAB, WHO and the International Prognostic Scoring System (IPSS).

IPSS is based on the percentage of bone marrow blasts, cytogenetics and number and degree of peripheral cytopenias at diagnosis, enabling identification of four risks groups: low, intermediate-1, intermediate-2, and high risk. Recently, new clinical and laboratory variables were identified that might add prognostic information to the IPSS (red blood cell transfusion dependency, high levels of LDH). Sponsors are therefore advised to follow closely the expected refinement of prognostic scores to be used in the design of clinical trials when sufficiently validated.

The WHO classification of myeloid neoplasms encompasses disorders that show both dysplastic and proliferative features at the time of diagnosis. The following disorders belong to this category: chronic myelomonocytic leukaemia (CMML), atypical chronic myeloid leukaemia, juvenile myelomonocytic leukaemia, and myelodysplastic /myeloproliferative disease, unclassifiable (MDS/MPD, U).

Inclusion Criteria in Exploratory and Confirmatory Trials

Since evolution of bone marrow failure and survival depend on patients' baseline characteristics, any efficacy or safety conclusion may apply only to patients sharing similar prognostic features. It is, however, also acknowledged that pharmacological activity may vary in relation to, e.g. cytogenetic characteristics. There is thus a need for rather extensive exploratory studies in order to identify the proper target population for confirmatory studies.

Even though it is unwise in general to include patients with highly variable prognosis if left untreated, this might become necessary if exploratory studies indicate similar activity irrespective of prognostic score, e.g. due to common expression of a certain drug target. Stratification using a well established prognostic score such as IPSS is recommended in such cases.

Treatments Aiming at Symptom Improvement

Alleviation of symptoms related to cytopenia is an acceptable aim of treatment in patients with MDS. In most cases this means reduction of anaemia-related symptoms. Due to prevalent co-morbidities in this elderly population, symptom scales, even if properly validated, may be too insensitive to capture also relevant differences between treatment groups especially as transfusion of red blood cells must be individualised due to e.g. concomitant cardiovascular disorders. Loss of need for transfusion for a defined period of time (in combination with improved haemoglobin levels) is therefore considered an acceptable outcome measure.

These trials, however, must investigate the impact of treatments (test and reference) on safety and on more global outcome variables, including disease evolution. OS and disease evolution must therefore be prospectively assessed to exclude detrimental effects of the test drug that would outweigh documented benefits.

Placebo on top of best supportive care based on currently available treatment options is an acceptable comparator if no specific active drug is available to treat the targeted symptoms. It is acknowledged that EPO is not licensed within the EU for the treatment of anaemia in patients with MDS, but subgroups of patients are identifiable with an increased likelihood of meaningful response. For these patients EPO may serve as comparator. Alternatively, patients non-responsive to EPO may be enrolled.

Treatments aiming at reducing risk for disease progression

Since progression to more severe stages of MDS and to AML is common and signals poor prognosis, any treatment that could delay or avoid progression is expected to have a positive impact on clinical outcome. Concerning the respective merits of disease progression-related endpoints and OS, all recommendations expressed in the main text of this guideline apply. Haematological or cytogenetic responses cannot be accepted a priori to assess efficacy, and response rate is more suitable for exploratory trials (detecting activity and dose-effect relationships) than for efficacy purposes (and detection of a clinical benefit).

Confirmatory studies are expected to be randomised and well controlled using a licensed or evidence based medicinal product as reference. In principle, PFS is an acceptable primary endpoint, but survival data are needed in order to exclude with reasonable certainty detrimental effects on survival. In high risk MDS, however, survival is the preferred measure of patient benefit. In the case HSCT is a realistic treatment option in responding patients, please refer to the section "Treatment administered with curative intent". The definition of progression must be based on a combination of standardised clinical and biological data and centralised blinded review is needed in order to establish progression.

MDS is a condition that irrespective long-term prognosis severely can compromised patients QoL. With respect to the possible role of PRO/QoL outcome measure, please refer to appendix (X to be released for comments next year). The influence of treatments aiming at symptom improvement as part of background SOC on parameters relevant for the evaluation of safety and efficacy of the experimental drug, should be carefully addressed.

8.5. Haematopoietic Stem Cell Transplantation

Drug development in relation to HSCT can be conducted as part of conditioning treatment for HSCT and also for the mobilisation of peripheral blood (PB) stem cells that will be utilised in a peripheral blood stem cell transplant (PBSCT). Immune therapy in relation to HSCT, however, is not covered.

a) Conditioning treatment

The outcome measures will need to focus on two aspects, engraftment (short term outcome) and a long term outcome which depends on the indication and type of transplant. In addition

long term follow up will be required and its duration will depend on the clinical setting.

If autologous HSCT is established in a certain condition such as in multiple myeloma, a randomised comparison with an established conditioning regimen is expected. The guidance as regards long term endpoints provided in the general guideline document apply. If not established, a comparison with standard of care with survival as outcome measure is expected.

In allogeneic HSCT, standardisation as far as possible as regards immune suppressive therapy and post transplant infection prophylaxis are warranted.

In both cases it is advisable to restrict inclusion so that variability in prognosis is reduced, not least if the primary aim is to show improved tolerability and safety and non-inferiority in terms of efficacy.

b) PBSC mobilisation

This section reflects use of medicinal products for the mobilisation of autologous PBSC. The target population in terms of the condition to be treated, prior therapy etc. should be reflected in the eligibility criteria. Extrapolation to other patient populations will in general not be acceptable.

Endpoints should include short term and long term outcome. A target number of CD34 cells that translates into a successful engraftment together with long term data on the engraftment will be required for approval. Possible effects on the underlying condition should also be addressed.

Details on engraftment (time to engraft, outcome of engraft etc) will be expected. The potential for tumour stem cell mobilisation and graft contamination should be addressed.

Specific short and long term safety data in relation to the HSCT should be submitted. Data on early complications such as mucositis, infections, sinusoidal obstruction syndrome (also known as hepatic veno-occlusive disease) and transplant-related lung injury will be required. Delayed complications including fertility toxicity, secondary malignancies and impaired growth and development in children will also need to be collected.

In the case of allogeneic HSCT particular attention should be given to data on acute and chronic graft versus host disease (GVHD) including details on specific prophylaxis and treatment measures and donor type (related or unrelated HLA matched transplant).

9. Definitions and Abbreviations

Chemoprotectant: A compound which counteracts the activity of anti-tumour compounds on normal tissue without (or clearly less) affecting the anti-tumour activity.

Chemosensitizer (or drug resistance modifier): A compound without own anti-tumour activity which increases the activity through pharmacodynamic interaction with anti-tumour compound(s).

Cytostatic: Anticancer compound shown to inhibit cell division without direct effects on tumour cell viability in non-clinical studies.

Cytotoxic: Anticancer compounds inducing irreversible lethal lesions through interference with DNA replication, mitosis, etc. following short term exposure in non-clinical studies.

Data maturity: A clinical study is considered mature if the distribution of events over time (early – late) makes it feasible to estimate the treatment effect in the full study population. This refers to the assumption that there is a biological difference between e.g. tumours progressing early and late and that the treatment effect might differ. The number of late events should therefore be large enough for study data to be stable. In practice, if a treatment difference has been established and a clear majority of events expected over long term have occurred, the study may in most cases be regarded as “mature”.

Non-cytotoxic: Anticancer compounds not belonging to the class of cytotoxic compounds.

Primary (innate) resistance: Progression without prior objective response or growth inhibition.

Refractory: Progression on therapy or within a short period of time after last cycle of therapy.

Resistance: Progression within a defined timeframe after end of therapy.

Randomised phase II trial: Randomised exploratory study designed to provide data of importance for the design of Phase III confirmatory studies, e.g. with respect an estimate of the possible magnitude of the effect using a clinically relevant measure of activity and/or biomarkers.

Secondary resistance: Progression after documented objective response or period of growth inhibition.

Window of opportunity: Under certain well-defined conditions it is acceptable to conduct a clinical study with an experimental compound in settings (line of therapy, stage, etc.) where available data for this compound normally would be regarded as too limited. The conditions for conducting such a study must be set rigorously so that the interest of the patient is guaranteed. Circumstances to take into account include benefit-risk of available therapies, available safety/activity data for the experimental compound, tumour-related symptoms (in most cases absent), expected evolution of the disease if left untreated or treated with available therapies, ease of frequent monitoring of tumour evolution (including use of biomarkers), planned intervention post chemotherapy, etc.

ADCC: Antibody dependent cellular cytotoxicity

ANC: Absolute neutrophil count

BSA: Body surface area

BSC: Best supportive care – include antibiotics, nutritional support, correction of metabolic disorders, optimal symptom control and pain management (including radiotherapy), etc. but does not include tumour specific therapy

CR: Complete response

CRF: Case report form

DFS: Disease-free survival (time from randomisation to recurrence or death from any cause)

DLT: Dose limiting toxicities

EFS: Event-free survival in this guideline refers to lack of achievement of CR, relapse and death without relapse are counted as events in an EFS analysis. Those patients who did not reach CR during the pre-specified induction phase will be considered as having an event at time 0.

HRQoL: Health related quality of life

MoAb: Monoclonal antibody

MTD: Maximum tolerated dose, often defined by dose-limiting toxicity occurring in at least 2 of 6 patients so that further dose-escalation is not undertaken.

ORR: Objective response rate (the proportion of patients in whom a CR or PR was observed)

OS: Overall survival (time from randomisation to death from any cause)

RP2D: Recommended phase 2 dose

PD: Pharmacodynamics

PK: Pharmacokinetics

PR: Partial response

PRO: Patient reported outcome

PFS: Progression-free survival (time from randomisation to objective tumour progression or death from any cause)

TTF: Time to treatment failure (time from randomisation to discontinuation of therapy for any reason including death, progression, toxicity or add-on of new anti-cancer therapy)

TTP: Time to tumour progression (time from randomisation to observed tumour progression, censoring for death without progression)

Appendix