17 December 2015
EMA/CHMP/703715/2012 Rev. 2
Committee for Medicinal Products for Human Use (CHMP)

Appendix 4 to the guideline on the evaluation of anticancer medicinal products in man

Condition Specific Guidance

Agreed by Oncology Working Party	June 2014
Agreed by CHMP for release for consultation	23 October 2014
Start of public consultation	15 December 2014
End of consultation (deadline for comments)	30 June 2015
Agreed by Oncology Working Party	November 2015
Adoption by CHMP for publication	17 December 2015
Date for coming into effect	1 July 2016

Revision 2 of this appendix refers to the addition of point 7 – 'Minimal residual disease as an endpoint in chronic lymphocytic leukaemia studies' (EMA/CHMP/629967/2014). All other content remains unchanged. For background information please see 'Concept paper on the need to revise Condition – Specific guidance, Appendix 4 to the guideline on the evaluation of anticancer medicinal products in man'.

Keywords	*NSCLC, Prostate Cancer, CML, MDS, HSCT, Breast Cancer, pCR, neoadjuvant treatment, surrogate endpoint, Minimal residual disease (MRD), Chronic lymphocytic leukaemia (CLL)*

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

Appendix 4 to the guideline on the evaluation of anticancer medicinal products in man

Table of contents

1. Non-Small Cell Lung Cancer 3

2. Prostate Cancer 4

3. Chronic Myeloid Leukaemia 8

4. Myelodysplastic Syndromes 10

5. Haematopoietic Stem Cell Transplantation 12

6. The role of the pathological Complete Response as an endpoint in neoadjuvant breast cancer studies 13

7. Minimal residual disease as an endpoint in chronic lymphocytic leukaemia studies NEW 16

1. Non-Small Cell Lung Cancer

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 to determine the molecular features of the tumour is highly recommended, and should be carried out in line with current scientific knowledge (see section 5).

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, performance 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 (see section 7.5.2). 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 may be an acceptable endpoint for early evaluation of new medicinal products in NSCLC (see section 6), though modest response rates may in fact underestimate patient reported benefits. In light of this, endpoints which also 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 (see section 7.1.5).

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. Autopsy analyses show that almost every man will ultimately develop prostate cancer, the majority being slowly progressive, and only a minority 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, metastatic prostate cancer is frequently found in the bone only, and 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 procedures between active and placebo groups in order to avoid sampling bias. Additionally, 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 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 for localized prostate cancer are associated with side effects that interfere with HRQoL, a concept of minimal invasive treatment, i.e. focal therapy, 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 effective treatment options become available in the metastatic setting, more trials are expected also in the (neo) adjuvant treatment.

Adjuvant endocrine treatment 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 and need for treatment of recurrence 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 biochemical, local and systemic failure should be reported separately, as well as the rate of secondary treatment.

Therapy for high-risk localized disease and locally advanced disease

Clinical stage, Gleason score and PSA level should all be considered in the evaluation of risk of recurrence in patients with localized disease. High-risk localized prostate cancer, includes either locally advanced disease, i.e. 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 serum 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 the achievement and maintenance of castrate levels of testosterone in the absence of breakthroughs and micro-surges. If the aim is to achieve “surgical level” of castration, 20 ng/dL and below, clinical benefit should be demonstrated in a randomized trial vs. standard therapy (target 50 ng/dL and below) if the benefit of a lower serum testosterone target level cannot be demonstrated by other means.

For non-hormonal products to be used as add-on or instead of, it is expected that favourable effects on PFS (see below) and/or OS are demonstrated in-line with the main guideline text.

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 castration-resistant 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 compounds 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 always suitable to assessment according to RECIST. Therefore the occurrence of new bone lesions as a marker for progressive disease might be considered acceptable, provided that pre-specified criteria (an example being the PCWG2 criteria) adequately addressing the possibility of a flare reaction or trauma are defined in the protocol. Indeed, 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, however, can even decrease in progressive late castration refractory state due to a dedifferentiation of the cancer cells making them unable to produce PSA.

Concomitant radiation therapy to prevent fractures may confound the efficacy evaluation, but should not be considered as an event in the efficacy analysis.

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.

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 standardization 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 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, for example, those provided by the US National Comprehensive Cancer Network or the European Leukaemia Net 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	Loss of MMolR	Loss of CHR
	Mutations[a]	Loss of CCgR	Mutationsb

CHR, complete haematological response (WBC <10x109/l, differential with no immature granulocytes and <5% basophils, platelet <450x109/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).
aBCR-ABL KD mutations still sensitive to imatinib.
bBCR-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. The choice of alternative time-points for primary or secondary endpoints may also be acceptable if fully justified, for example if a response to treatment is expected to occur earlier during therapy 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.

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.

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 in this appendix.

a) Conditioning treatment

Whenever possible, conducting comparative studies against accepted standard conditioning treatment(s) will be expected and the choice of endpoints and time points will depend on the specific clinical condition. 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 applies. 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 is 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) Treatment prior to high dose therapy

The aim should be to improve overall outcome and the principles of ITT should be adhered to, i.e. also patients not undergoing (autologous) HSCT should be followed for PFS/EFS and OS, prioritized as regards primary outcome measure as recommended in the main guideline. Quality of response prior to high dose therapy may be reported as a secondary endpoint.

c) 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 mobilization, including tumour stem cells, and graft contamination should be addressed.

In cases where the PBSC mobilisation is intended for use in allogeneic transplant a safety assessment of the donor including short and long term data will be expected.

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 including details on specific prophylaxis and treatment measures and donor type (related or unrelated HLA matched transplant).

6. The role of the pathological Complete Response as an endpoint in neoadjuvant breast cancer studies

Neoadjuvant treatment is commonly used in early breast cancer to facilitate breast conserving surgery (*Romero et al. 2013*), but is also increasingly used in the development of new regimens to be used in the metastatic or adjuvant settings.

As new therapies have emerged, disease-free survival (DFS) and ultimately overall survival (OS) of patients with early breast cancer have improved and thereby the time needed to procure confirmatory data. A new endpoint that would allow the assessment of efficacy at an earlier point in time would therefore be valuable, as it could potentially bring novel and improved therapies faster to the market for the benefit of the patients with high-risk early breast cancer.

Pathological complete response

Based on data from several meta-analyses and clinical trials, pathologic complete response (pCR) has been shown to have an association with long term outcomes. pCR has therefore been proposed as a

surrogate endpoint for the evaluation of the efficacy of therapies for invasive breast cancer without distant metastasis.

There are several definitions of pCR in the scientific literature. However, for regulatory purposes the following definition is recommended: absence of any residual invasive cancer on haematoxylin and eosin evaluation of the resected breast specimen and all sampled ipsilateral lymph nodes following completion of the neoadjuvant systemic therapy (ypT0/is ypN0).

The relationship between pCR and OS/EFS

Several randomised trials, systematic reviews and meta-analyses have shown a consistent association between achievement of pCR following primary systemic therapy and better overall survival. Based on these trials, however, the importance of the treatment factor cannot be fully disentangled from the patient factor. This means that the true surrogacy of pCR, i.e. to what extent a certain difference in pCR rate can predict a certain difference in EFS, however, has not been established. (Fisher et al, 1998, Bear et al, 2006; Wapnir et al, 2006; Mieog et al, 2007; Mazouni et al, 2007; Kong et al, 2011; Berruti et al, 2011; Mamounas et al, 2012; Esserman et al, 2012; von Minckwitz et al, 2012; Cortazar et al, 2014; Bonnefoi et al, 2014).

The pCR rate differs according to subtype of breast cancer. A meta-analysis of neoadjuvant studies has shown that pCR rates were lower in patients with low-grade, hormone receptor-positive (HR+) tumours, and higher in the following tumour subtypes in increasing order: high-grade HR+, HR+/HER2+, triple negative, and hormone receptor-negative (HR–)/HER2+. In addition, patients with more aggressive tumour subtypes who achieved pCR seemed to have greater event-free survival (EFS) benefit compared to patients who did not achieve pCR.

In conclusion, there appears to be a stronger association between pCR and EFS in patients with aggressive tumour subtypes compared to patients with less aggressive tumours (Cortazar et al., 2014). Of note, in case of biomarker guided therapy, the value of pCR as outcome measure may be limited if the biomarker is not associated with aggressive tumours. Due to these limitations, this amendment to the anti-cancer guideline is focused on patients with high risk tumours.

pCR as endpoint in neoadjuvant breast cancer studies from a regulatory perspective

Eligibility criteria should define a patient population with high risk tumours expressing suitable biomarkers in relation to the selected background regimen and the experimental compound such as HER2 expression, hormone receptor status, BRCA status, markers for impaired DNA repair in general, etc.

It is expected that pCR assessment, whether locally or centrally conducted, is undertaken with experienced pathologists blinded for treatment assignment and in accordance with the principles of good clinical laboratory practise.

All efforts should be undertaken to provide best-evidence based surgical and adjuvant treatment to all patients. The surgical approach and management of the primary tumour and axilla should follow standard algorithms and should be clearly referenced or explained in the protocol. Eradication of the tumour from both breast and lymph nodes has been shown to be associated with better EFS and OS compared with eradication in only the breast (Cortazar et al. 2014; Von Minckwitz et al. 2012). Presence or absence of ductal carcinoma in situ does not appear to affect long-term outcomes (Cortazar et al. 2014).

As different strategies in terms of radiotherapy and adjuvant treatment may have an impact on the long term outcome, foreseeable difficulties in the interpretation of the results should be addressed by clearly defined and protocol pre-specified treatment strategies, including the neoadjuvant background

treatment regimen, and conditions for use of defined adjuvant regimens, depending on patients and disease characteristics.

Due to putative tumour heterogeneity and micro-metastatic spread, which leaves patients at continued risk of distant metastases and death, and the established efficacy of current (neo)adjuvant regimens, randomised trials in which the new agent is added to an established (neo)adjuvant treatment regimen are likely to be required. If, in the exceptional case, it is not possible to add the experimental compound to an existing regimen, the Sponsor is advised to request CHMP scientific advice.

The mechanism of action of the experimental compound should be well-known and there should be no reason to suspect any adverse interactions with the established treatment regimen based on PK and PD studies as needed, studies in the advanced stage and safety data.

For confirmatory neoadjuvant studies, EFS (the interval from randomisation to the earliest occurrence of disease progression resulting in inoperability, locoregional recurrence, distant metastases, or death from any cause) is accepted as primary outcome measure. It is, however, recommended that studies are also designed to capture distant metastasis-free survival. This means that patients should be followed after, e.g. locoregional recurrence, not only for survival, but also distant metastases. DFS is considered to be an appropriate primary endpoint in adjuvant studies (EMA/CHMP/205/95/Rev.4).

In principle, confirmation in terms of EFS/OS benefit may be achieved through the same study documenting the increase in pCR rate, but it is foreseen that possible adjuvant therapy will be at least partly guided by the response to the neoadjuvant therapy. This means that it will be harder to document a difference in EFS. Therefore a confirmatory, adjuvant DFS/OS trial with the same regimens that were used in the neoadjuvant trial may be undertaken instead, provided that the reference treatment is considered as an appropriate reference regimen for the adjuvant setting.

EFS/DFS assessment should follow the methodological consideration for using progression-free survival (PFS) or DFS in confirmatory trials as expressed in the Appendix 1 to the guideline on the evaluation of anticancer medicinal products in man.

Currently available data do not allow a precise prediction of the magnitude of the EFS/DFS/OS effect from a certain pCR effect. Therefore, a substantial increase in pCR shown in sufficiently large randomised trials is required for there to be a reasonable likelihood that this will translate into a clinically meaningful improvement in long-term outcomes.

As the magnitude of the effect in terms of EFS/DFS/OS cannot be well estimated, only minor add-on changes in toxicity are acceptable. In addition, agents which pose a theoretically increased risk of secondary tumours must include appropriate patient monitoring and follow up in the trials to assess and report such risks. The safety data base derived from the neoadjuvant trial should be sufficiently large to capture relevant increases in common adverse reactions compared with the underlying standard regimen and follow-up should be sufficiently long to assess reversibility of known side effects, such as neuropathy and cardiomyopathy and as appropriate secondary tumours.

Studies conducted with the regimen in the metastatic setting may provide important safety data and important supportive evidence of efficacy; however, this may not be necessary in all cases and should be determined on a case by case basis. The totality of the data available will be used in the assessment of benefit/risk.

Summary

Approval based on pCR may be acceptable for a medicinal product as add-on to an established (neo)adjuvant regimen for the treatment of patients with high-risk early stage breast cancer, provided that that mechanism of action is well-characterised and provided that the results show a major

increase in pCR with only minor changes in toxicity. In most cases supportive evidence of efficacy and safety of the experimental compound is expected from studies conducted in the metastatic setting.

These data may lead to an approval with agreed conditions for confirmatory study data in terms of EFS/DFS/OS to be submitted. Confirmation may in principle be achieved through prolonged follow-up of the neoadjuvant study if sufficiently large or through a separate adjuvant study.

7. Minimal residual disease as an endpoint in chronic lymphocytic leukaemia studies

Undetectable minimal residual disease (MRD) in patients with chronic lymphocytic leukaemia (CLL) in clinical complete remission (= MRD response rate) after induction therapy may be used as an intermediate endpoint for licensure in randomised well controlled studies designed to show superiority in terms of PFS. Regulatory recommendations with regards to laboratory aspects for MRD measurements, definition as a clinical intermediate efficacy endpoint and the inclusion in the statistical analysis plan should be followed.

Introduction

Chronic lymphocytic leukaemia (CLL) is the most common leukaemia in the Western world with an incidence of 4.2/100000/year that increases to >30/100000/year at an age >80 years.

Treatment is recommended only for those patients with active, symptomatic disease. With the introduction of new immune-chemotherapeutic combinations over the last decade the efficacy of treating patients with CLL has greatly improved and median PFS now ranges from 3.5 to 6.7 years after first line therapy whilst median OS for patients with advanced stages (Binet C or Rai IV) is approximately 6.5 years. Despite these significant advances, the disease remains incurable when treated with chemotherapy and monoclonal antibodies alone. Allogeneic stem cell transplant remains the only curative therapy and it is recommended for patients with very high risk and/or refractory disease.

Currently, PFS is considered an appropriate primary endpoint to demonstrate clinically meaningful patient benefit in randomised phase III CLL studies. However, with such an endpoint the timeframe to achieve meaningful statistical and clinical results from pivotal studies with new therapies in earlier treatment lines is well over 5 years. In the effort to develop efficacious treatment options to address the unmet medical need of CLL patients, there is a need to find alternatives to the currently used time-to-event variables so that the efficacy of novel therapies can be evaluated at an earlier time point.

Because patients achieving clinical complete remission (CR) according to international guidelines will eventually relapse, residual disease (RD) undetectable at clinical and morphological level must have been present at the time of CR. Therefore, the quality of response to treatment should be also assessed for the absence of detectable RD. The vast improvement in RD detection over the last two decades has now led to the concept that low RD levels are a desirable and achievable goal of CLL therapy.

Scope

The scope of this document is to describe the basis and regulatory requirements for the use of **minimal residual disease (**MRD) as an intermediate endpoint to predict clinical benefit in trials in CLL. At present, this guidance is not applicable to other clinical settings, such as other B-cell lymphomas. Of note, this document is not intended to discuss MRD-guided treatment of CLL.

General aspects of MRD

Definition & threshold

MRD is an objective measure of disease status defined by the number of leukaemic cells remaining in peripheral blood or bone marrow following treatment. According to current international definitions undetectable MRD (also known as MRD negativity) equals a quantitative detection of less than 1 CLL cell in 10000 leukocytes (MRD level < 10^{-4}).

There is no prospective data currently available to support that a further reduction of MRD level below the 10^{-4} threshold would provide added clinical benefit.

Laboratory assays

Although MRD evaluation is still not widely available there are currently two analytical methods capable of assessing MRD status at the required threshold, i.e. real-time quantitative PCR and four (or more)-colour flow cytometry. There is no specific recommendation on the method to be used as both are considered appropriate. Additional methods for which equivalent sensitivity, specificity and quantitative ranges have been demonstrated may be used in the future.

A quality management system that includes the laboratory(s) organisational structure, responsibilities, policies and standards needed to ensure accuracy and satisfactory quality of the MRD evaluation assay would be required. It is recommended that MRD should be evaluated under GLP, or an equivalent quality management system, and that the analytical method should be appropriately validated.

The use of central laboratories is not considered a regulatory requirement provided a robust quality system is in place and that the same protocol is used for that particular analytical method. All local laboratories within a clinical trial should undergo inter-laboratorial comparisons (round-robin tests) in order to normalize results and thereby render them comparable between different laboratories and may be between different trials.

- *Real-time quantitative PCR (RQ-PCR)*

Every leukaemic B-cell clone carries a unique IGHV-IGHD-IGHJ rearrangement that can be amplified by PCR using primers. Allele specific oligonucleotide immunoglobulin heavy chain real-time quantitative PCR (ASO IGH RQ-PCR) is labour intensive as it requires the sequencing of each clone-specific rearrangement prior to induction therapy but has sensitivity in the range of 10^{-4} to 10^{-5}.

Limitations of the method apply in case of changes in phenotype or genotype between baseline and follow up investigations. Since specific primers address a single rearranged IGH gene sequence, there is a certain risk of target gene loss due to ongoing rearrangements/somatic mutations in the IGH region which would result in reduced sensitivity. If biclonal disease is found at baseline, two IG PCR targets should be used to accurately quantify MRD. Patients with oligoclonal disease where accurate quantification of the CLL cell count of all clones is not possible should not be assessed for MRD by ASO-PCR.

A major advantage is that the samples do not need to be fresh and can be shipped to a single centre for analysis. Conserved samples could further enable retrospective analysis in clinical trials. In addition, ASO RQ-PCR offers a higher qualitative sensitivity below the threshold of 10^{-4} which might be relevant in clinical trials exploring complete eradication of the disease.

- *Four (or more) -colour flow cytometry*

Because CLL cells show a characteristic unique phenotype, low amount of leukaemic cells can be detected using flow cytometry to the required sensitivity level of 10^{-4}. The sensitivity of MRD flow primarily depends on the availability of sufficient numbers of leukocytes in a sample.

The main advantage of this method is that it is simpler and faster as it does not require the design of clone-specific primers. However, in the context of a clinical trial, sequence analysis of the IGH gene is expected to be available at baseline because mutational status is considered a prognostic factor.

It uses a widely available technology and is therefore a broadly applicable method. A disadvantage is that samples are required to be fresh (48h). Appropriate handling and transport to central laboratories may be difficult to establish in multi-centre, multi-national clinical trials. Implementation of regional laboratories may offer an acceptable solution as long as data handling and analysis are consistent across all sites.

Samples

MRD status can be assessed either from peripheral blood (PB) or bone marrow (BM).

It is recommended that for all medicinal products irrespective of drug class, patients are screened for CLL eradication in PB first. If MRD is not detectable in PB, it is mandatory to confirm MRD status in the BM.

Utility

It is accepted that response to therapy is the most important prognostic factor for survival. A profound reduction of tumour load, as evaluated with the MRD assay and not the treatment regimen by which this reduction is induced is the key factor for durable remission.

Available data has shown that undetectable MRD at the end of induction treatment is a strong predictor of PFS and OS irrespective of the following:

- Type and line of treatment

  Although patients are more likely to reach undetectable MRD with some therapies compared to others, for those patients that achieved undetectable MRD there appear to be no differences in terms of PFS or OS regardless of therapy received. Data are still limited, however.

- Known poor pre-treatment risk factors (e.g. deletion chromosomes 11q and 17p, mutated TP53, un-mutated IGHV status, ZAP70 expression)

The availability of MRD data shortly after treatment is important because with more effective treatment regimens PFS will be evaluable only after a long observation period.

Current evidence suggests that an MRD level $\geq 10^{-2}$ is associated to a median PFS of about 2 years, whereas a MRD level $< 10^{-4}$ predicts a median PFS of around 6 years.

The validation of MRD response rate (undetectable MRD + CR) as a surrogate endpoint requires that the treatment effect on this marker can explain quantitatively the treatment effect in terms of PFS. Qualitatively available data are sufficiently convincing for MRD response rate to be used as an intermediate endpoint in randomised controlled trials as long as the benefit in terms of long term efficacy can eventually be confirmed.

MRD as an endpoint for licensure

A difference in MRD response rates can be used as primary evidence of clinical benefit to obtain early licensure in randomised CLL trials designed to show superiority in terms of PFS but where mature PFS data will only become available at a later stage. Regulatory considerations (e.g. legal basis of the marketing authorisation application or other considerations, for example conditional approval) should be decided on a case by case basis.

The following conditions are required and any deviations should be fully justified.

Study design and results

- All patients with clinical response (CR or PR) should be assessed for MRD in PB first. Only patients with undetectable MRD in PB should have confirmation of MRD status in BM.
- The control regimen should be selected according to the criteria set out in the main anticancer guideline.
- The trial should be prospectively powered for PFS.
- The statistical analysis and methods for assessment of MRD and PFS should be pre-planned and clearly described in the statistical analysis plan.
- The difference in MRD response rate between study arms should be large enough to predict that a relevant PFS benefit will appear on mature data
- In case of early approval based on MRD response rate, an analysis of PFS would be required from the holder of the marketing authorisation in an agreed timeframe.
- All patients should be followed for OS
- The medicinal product is expected to show an improved or no significant major difference in terms of safety profile compared to the control.

MRD definitions as clinical endpoint and methods

- MRD status should be measured by a standardised method with a quantitative lower limit of at least $< 10^{-4}$ following guidelines that define specificity, sensitivity and reproducibility.
- A quality control scheme for each laboratory providing CLL MRD analysis in the clinical trial will be required.
- Measurement of MRD should be conducted at end-of-treatment response final staging assessment (around 3 months after end of treatment) to fully represent the effect of treatment. Deviations from the recommended time point for MRD assessment may be acceptable if justified by appropriate clinical data on the mechanism of action of the drug and prior knowledge on the kinetics of responses.
- MRD will be considered undetectable if the proportion of malignant cells is $< 10^{-4}$.
- MRD response rate is defined as the proportion of patients in the ITT population in whom a clinical complete response (CR) and undetectable MRD status in bone marrow is achieved following induction treatment in CLL.
- Patients who achieve clinical CR and undetectable MRD status at the end of treatment will be counted as MRD responders. Based on current available data it is not acceptable at present to include patients with a clinical partial response and undetectable MRD status as MRD responders.
- Patients with missing MRD assessment (any cause), patients with detectable MRD- status and patients in PR with undetectable MRD will be counted as MRD non-responders. A sensitivity analysis in patients with missing bone marrow samples may be conducted.

Additional recommendations and considerations

- Exploratory analyses are recommended using different cut-offs for “undetectable MRD” in patients with CR as well as PR. The prognostic value of different levels of MRD response may also be explored.

- For exploratory purposes, it is recommended that all patients responding to therapy (including PR) should have their MRD status assessed at least in peripheral blood.
- Correlation between MRD in PB and BM should be explored. BM MRD assessment is not informative for patients with $>10^{-2}$ MRD in the PB
- For patients that undergo allogeneic stem cell transplant (SCT), early detectable MRD is common probably due to the fact that the onset of graft-versus-leukaemia is not immediate. Undetectable MRD can be achieved several months after allogeneic SCT.

Additional areas of uncertainty

It has been suggested that the kinetics of MRD rather than a single MRD assessment may be more meaningful because it is the increase of MRD over time and not only its persistence that is eventually followed by clinical relapse. The kinetics of relapse is exponential even at the lowest evaluable levels of the disease.

The prognostic significance of MRD assessments during induction therapy is unknown, in particular, for tailoring treatment according to MRD response in order to reduce duration of treatment and subsequent reduction of toxicity.

At present it is not known whether long term outcome can be improved if MRD assessment is used to guide therapy, either to improve the quality of response through consolidation/maintenance therapy or to prevent relapse by therapies based on reappearance of MRD. MRD assessments in these areas of uncertainty may be very useful to explore.

References

1. Minimal residual disease quantification is an independent predictor of progression-free and overall survival in chronic lymphocytic leukaemia: a multivariate analysis from the randomized GCLLSG CLL8 trial. S. Böttcher *et al*; Journal of clinical oncology, vol 30 (9), 2012
2. A look into the future: can minimal residual disease guide therapy and predict prognosis in chronic lymphocytic leukaemia? Paolo Ghia, Haematology 2012
3. Chronic lymphocytic leukaemia: ESMO clinical practice guidelines for diagnosis, treatment and follow up, Annals of Oncology 2011
4. ASCO & FDA public workshop on MRD as a surrogate endpoint in CLL (Silver Spring, February 2013)
5. International standardized approach for flow cytometric residual disease monitoring in chronic lymphocytic leukaemia. A.C. Rawstron *et al*; Leukaemia (2007) 21, 956-964
6. Flow cytometry and polymerase chain reaction-based analyses of minimal residual disease in chronic lymphocytic leukaemia. S. Uhrmacher *et al*; Advances in haematology, vol 2010, article ID 272517
7. The role of minimal residual disease measurement in the therapy for CLL. Is it ready for prime time? S. Böttcher *et al;* hematology Oncology Cli N Am (2013) 267-288.
8. Hallek M *et al.* Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute Working Group 1996 guidelines. Blood. 2008; 111: 5446 - 5456.