Annals of Oncology 27: 2149–2160, 2016
doi:10.1093/annonc/mdw419

ESMO consensus conference on malignant lymphoma: general perspectives and recommendations for prognostic tools in mature B-cell lymphomas and chronic lymphocytic leukaemia

M. Ladetto[1]*, C. Buske[2], M. Hutchings[3], M. Dreyling[4], G. Gaidano[5], S. Le Gouill[6], S. Luminari[7,8], C. Pott[9], A. Zamò[10], E. Zucca[11]& the ESMO Lymphoma Consensus Conference Panel Members[†]

[1]Hematology Division, Azienda Ospedaliera Santi Antonio e Biagio e Cesare Arrigo, Alessandria, Italy; [2]Comprehensive Cancer Center Ulm and Department of Internal Medicine III, Institute of Experimental Cancer Research, University Hospital, Ulm, Germany; [3]Department of Hematology, Rigshospitalet, Copenhagen, Denmark; [4]Medizinische Klinik III, Klinikum der Universität München/LMU, Munich, Germany; [5]Division of Hematology, Department of Translational Medicine, Amedeo Avogadro University of Eastern Piedmont, Novara, Italy; [6]Clinical Hematology, Centre Hospitalo-Universitaire de Nantes, UMR892 Team 10, CIC Nantes, France; [7]Hematology, Arcispedale S. Maria Nuova, IRCCS Reggio Emilia; [8]Department of Diagnostic, Clinical and Public Health Medicine, University of Modena and Reggio Emilia, Modena, Italy; [9]Second Medical Department, University Hospital Schleswig-Holstein, Kiel, Germany; [10]Department of Diagnostics and Public Health, University of Verona, Verona, Italy; [11]Lymphoma Unit, Oncology Institute of Southern Switzerland, Ospedale San Giovanni, Bellinzona, Switzerland

Received 9 June 2016; revised 11 August 2016; accepted 23 August 2016

The European Society for Medical Oncology (ESMO) consensus conference on mature B-cell lymphomas and chronic lymphocytic leukaemia (CLL) was held on 20 June 2015 in Lugano, Switzerland, and included a multidisciplinary panel of 25 leading experts. The aim of the conference was to develop recommendations on critical subjects difficult to consider in detail in the ESMO Clinical Practice Guidelines. The following areas were identified: (i) the elderly patient, (ii) prognostic factors suitable for clinical use and (iii) the 'ultra-high-risk' group. Before the conference, the expert panel was divided into three working groups; each group focused on one of these areas in order to address four clinically relevant questions relating to that topic. All relevant scientific literature, as identified by the experts, was reviewed in advance. During the consensus conference, each working group developed recommendations to address each of the four questions assigned to their group. These recommendations were then presented to the entire panel and a consensus was reached. This manuscript presents recommendations dedicated to the second area of interest, i.e. prognostic factors suitable for clinical use. The four topics [i.e. interim positron emission tomography (PET), *TP53* mutations, cell of origin (COO) and minimal residual disease (MRD)] were primarily chosen because of the bulk of available data together with the lack of clear guidance regarding their use in clinical practice and within clinical trials. Results, including a summary of evidence supporting each recommendation, are detailed in this manuscript. The panel acknowledged that detection of *TP53* inactivation by deletion or mutation in CLL should be implemented in clinical practice (level of evidence I, strength of recommendation A). Due to their potentially high prognostic value, at least in some lymphoma entities, implementation of interim PET, COO and MRD was highly recommended in the context of clinical trials. All expert panel members approved this final article.

Key words: lymphoma, consensus, positron emission tomography, *TP53*, cell of origin, minimal residual disease

introduction

Laboratory-based and imaging tools are increasingly used in patients with lymphoid malignancies to better understand their prognosis and even to guide therapeutic decisions. Despite their documented predictive value in several specific settings, their use is often extended to conditions where there is little evidence of substantial therapeutic benefit. This could result in an increase in costs and inappropriate therapeutic decisions. As such, clear recommendations regarding the use of these tools are required.

In 2015, the European Society for Medical Oncology (ESMO) held a consensus conference on mature B-cell neoplasms and chronic lymphocytic leukaemia (CLL) in order to develop recommendations on critical subjects that were difficult to consider in detail in the ESMO Clinical Practice Guidelines (CPG). In this consensus conference, one of the working groups (Working Group 2) focused on prognostic factors suitable for clinical use.

*Correspondence to: Dr Marco Ladetto, ESMO Guidelines Committee, ESMO Head Office, Via L. Taddei 4, CH-6962 Viganello-Lugano, Switzerland. E-mail: clinicalguidelines@esmo.org

[†]See the appendix for members of the ESMO Lymphoma Consensus Conference.

As such, the objectives of this working group were: (i) to identify a restricted number of prognostic tools whose clinical use is established or under rapid technological development; (ii) to discuss the technical and clinical reliability of these prognosticators; (iii) to consider the prognostic value of these tools; (iv) to provide recommendations on the use of these prognosticators in the context of clinical research and routine practice. Here, we describe the recommendations developed by Working Group 2 and approved by the whole panel, and provide a summary of evidence supporting each recommendation.

methods

A consensus panel, comprising a multidisciplinary panel of 25 experts in the management of lymphoma, was convened by ESMO. Three consensus conference chairs (CB, ML, MH) were also appointed. The consensus panel was divided into three working groups, each of which was assigned a specific subject area and a working group chair as follows: Working Group 1: the elderly patient (Chair: CB); Working Group 2: prognostic factors suitable for clinical use (Chair: ML); Working Group 3: the 'ultra-high-risk' group (Chair: MH). The consensus conference was held on 20 June 2015 in Lugano, Switzerland. Before this consensus conference, four clinically relevant questions were identified for each subject area.

A literature review was conducted by each working group before the consensus conference, with each group responsible for compiling a summary of relevant information required to develop recommendations relating to each of their questions at the conference. No systematic literature search was undertaken. During the conference, in parallel sessions, the three working groups discussed and agreed on recommendations relating to each of their assigned questions. The level of evidence and strength of each recommendation were also noted, which were defined based on the 'Infectious Diseases Society of America-United States Public Health Service Grading System', as shown in Table 1 [1]. Recommendations from each group were then presented to the entire panel of experts, where they were discussed and modified, as required. Finally, a vote was conducted to determine the level of agreement among the expert panel for each of the recommendations. Discussion regarding each of the recommendations was completed after the consensus meeting, with additional supporting evidence published after the meeting also included in the final manuscript.

For Working Group 2, which is the focus of this report, four prognostic tools were identified for discussion in terms of their potential suitability as prognostic tools for clinical use. Discussions focused on B-cell lymphoma and CLL; plasma cell disorders and T-cell lymphoma were considered outside the scope of this consensus conference. In addition, working group members were asked to focus on disease entities in which the prognostic tools were most promising and where a greater need for clinical recommendations was required (front-runner entities, FRE). As such, the following prognostic tools and associated disease entities of specific interest were considered:

(i) interim positron emission tomography [PET; FRE: Hodgkin's lymphoma (HL), diffuse large B-cell lymphoma (DLBCL)],
(ii) *TP53* mutations and deletions (FRE: CLL),
(iii) cell of origin (COO) determination by gene expression profiling (GEP) or immunohistochemistry (IHC) (FRE: DLBCL),
(iv) molecular-based minimal residual disease (MRD) evaluation [FRE: mantle cell lymphoma (MCL); follicular lymphoma (FL), CLL]

Table 1. Levels of evidence and grades of recommendation (adapted from the Infectious Diseases Society of America-United States Public Health Service Grading System[a])

Levels of evidence	
I	Evidence from at least one large randomised, controlled trial of good methodological quality (low potential for bias) or meta-analyses of well-conducted randomised trials without heterogeneity
II	Small randomised trials or large randomised trials with a suspicion of bias (lower methodological quality) or meta-analyses of such trials or of trials with demonstrated heterogeneity
III	Prospective cohort studies
IV	Retrospective cohort studies or case–control studies
V	Studies without control group, case reports, experts' opinions
Grades of recommendation	
A	Strong evidence for efficacy with a substantial clinical benefit, strongly recommended
B	Strong or moderate evidence for efficacy but with a limited clinical benefit, generally recommended
C	Insufficient evidence for efficacy or benefit does not outweigh the risk or the disadvantages (adverse events, costs, . . .), optional
D	Moderate evidence against efficacy or for adverse outcome, generally not recommended
E	Strong evidence against efficacy or for adverse outcome, never recommended

[a]By permission of the Infectious Diseases Society of America [1].

Results from the section of the consensus conference dedicated to prognostic factors suitable for clinical use, together with a summary of evidence supporting each recommendation, are detailed in this article. A summary of these recommendations is shown in Table 2. Importantly, these additional recommendations should be read in conjunction with the already-published ESMO CPGs for the diagnosis, treatment and follow-up of malignant lymphomas and CLL [2–6].

results

1. Interim PET as a prognostic tool

18-F-fluorodeoxyglucose (FDG)-PET has recently been recommended as the standard tool for the evaluation, staging and response assessment for patients with FDG-avid lymphomas, including HL, DLBCL and FL [7]. With the use of FDG-PET, metabolic response has increasingly been acknowledged as one of the strongest available prognostic tools and has been identified as a surrogate test for chemosensitivity. For the purposes of this consensus manuscript, the definition of interim PET applies to any FDG-PET carried out during a planned systemic treatment, usually after 2–4 cycles in the case of a conventional chemotherapy programme, or after 2–4 cycles of reinduction chemotherapy

Table 2. Summary of recommendations[a]

Guidelines statement	LoE	GoR
1. The potential role of interim PET as a prognostic tool		
Recommendations		
1.1 The exploratory use of interim FDG-PET as a surrogate test of chemosensitivity and as a diagnostic tool to facilitate clinical decision-making is encouraged in clinical trials in HL, DLBCL and other aggressive FDG-avid lymphoma entities	III	B
1.2 There are little published data from randomised trials to support the use of an interim PET-driven therapeutic strategy in HL, DLBCL or other FDG-avid lymphomas. However, preliminary data strongly support the use of interim PET to tailor therapy in individual cases. On these grounds, results of interim PET may be applied in individual patients with early or advanced HL	II	C
1.3 Based on the lack of therapeutic consequences, the routine clinical use of interim PET is not recommended in patients with DLBCL	II	D
1.4 Based on the lack of data, the routine use of interim PET as a decision tool is discouraged in non-HL, non-DLBCL, FDG-avid lymphoma entities	V	E
2. The potential role of *TP53* mutations and deletions as a prognostic tool		
Recommendations		
2.1 Given the well-established, prognostic and predictive value of *TP53* disruption in CLL, the panel strongly recommends the inclusion of *TP53* analysis, both by FISH and DNA sequencing, in clinical trials of CLL for intervention and monitoring purposes. In particular, the availability of new drugs that overcome *TP53*-mediated chemorefractory disease mandates the acquisition of *TP53* status for all patients with CLL at the time of screening procedures in trials in which one or more arms may be based on drugs that are known to be ineffective in *TP53*-disrupted CLL. Therefore, the use of *TP53* screening for monitoring and intervention in clinical trials is encouraged in CLL	I	A
2.2 In other lymphoid neoplasms, *TP53* screening for investigational purposes is neither recommended nor discouraged. At present, the panel discourages clinical trials aimed at specific interventions based on *TP53* status unless prognostic markers are the major focus of the trial and the drug being evaluated has a strong biological rationale for overcoming *TP53*-mediated resistance	V	C (investigation); D (intervention)
2.3 In CLL, the panel supports analysis of *TP53* disruption at the time of treatment requirement, both in first-line and subsequent lines of therapy. Reassessing *TP53* status in previously *TP53* wild-type CLL at relapse requiring treatment is relevant since *TP53* disruption may develop, or become detectable only at relapse. In routine practice, characterising *TP53* status in a given patient with CLL is clinically relevant as this may affect treatment decisions. The use of *TP53* screening by FISH and mutational analysis for monitoring and intervention in clinical practice is therefore encouraged in CLL, provided there is availability of and access to therapies overcoming *TP53*-mediated resistance (e.g. inhibitors of the B-cell receptor and allo-SCT)	I	A
2.4 In other lymphoid neoplasms, the panel discourages the use of *TP53* outside of clinical trials as there is no general recommendation for treatment modification currently published. The results of currently recruiting trials might modify this attitude in the coming years	V	E
3. The potential role of COO determination by IHC or GEP as a prognostic tool		
Recommendations		
3.1 Given the limitations of IHC, the panel does not encourage its use in prospective clinical trials for prognostication	I	C
3.2 Given the limitations of IHC, the panel discourages its use in prospective clinical trials to guide intervention	I	D
3.3 The panel strongly encourages the use of GEP in prospective clinical trials for prognostication	I	A
3.4 Clinical trials of interventions based on GEP results are encouraged	I	B
3.5 Based on inadequate standardisation, and a lack of well-designed interventional studies, the use of COO determination by IHC or GEP in DLBCL is generally not recommended in routine clinical practice outside of clinical trials	V	D
4. The potential role of molecular-based MRD evaluation as a prognostic tool		
Recommendations		
4.1 The use of MRD evaluation for monitoring and intervention in clinical trials is encouraged in MCL, FL and CLL		
(i) MCL: for monitoring	I	B
(ii) MCL: for intervention	III	C
(iii) FL: for monitoring	I	B
(iv) FL: for intervention	IV	C
(v) CLL: for monitoring (depends on the drug used)	I	B
(vi) CLL: for intervention	IV	C

Continued

Table 2. *Continued*

Guidelines statement	LoE	GoR
4.2 The use of MRD evaluation for monitoring and intervention in clinical practice is not recommended in MCL, FL and CLL, with the exception of monitoring after allo-SCT	V	D

[a]There was 100% consensus from the panel of experts for all recommendations listed.
LOE, level of evidence; GOR, grade of recommendation; allo-SCT, allogeneic haematopoietic stem cell transplantation; CLL, chronic lymphocytic leukaemia; COO, cell of origin; DLBCL, diffuse large B-cell lymphoma; DNA, deoxyribonucleic acid; FDG, fluorodeoxyglucose; FISH, fluorescence *in situ* hybridisation; FL, follicular lymphoma; GEP, gene expression profiling; HL, Hodgkin's lymphoma; IHC, immunohistochemistry; MCL, mantle cell lymphoma; MRD, minimal residual disease; PET, positron emission tomography.

before the administration of a planned high-dose chemotherapy followed by stem cell support where intensified regimens are used.

In HL and DLBCL, the identification of metabolic response during treatment has been correlated with the individual risk of relapse, and of death in some cases, and has the potential to improve patient outcome through the early adaptation of treatment intensity [8–12]. There is general consensus that the achievement of an early metabolic response during treatment is predictive of favourable outcomes in terms of both progression and overall survival (OS). The high negative predictive value of interim PET, however, is counterbalanced by a variable rate of false-positive results that are usually more common in DLBCL than in HL [13].

methodological considerations

broad availability. Although FDG-PET is broadly available in high-income and in some middle-income countries, it remains inaccessible for many patients. As such, access to FDG-PET still needs to be improved worldwide.

reproducibility and standardisation. Reproducibility of FDG-PET has markedly improved with the application of standardised and recommended methods, particularly with the use of the Deauville 5-point scale (5PS) [14–17] [III, B]. However, quality assurance and training programmes are still needed. The application of semi-quantitative measurements of interim PET [i.e. delta standardised uptake value (SUV) max] is not recommended, although data suggest it may add prognostic detail in DLBCL [18].

clarity of reporting system. Currently, routine clinical reports are not well standardised. The panel recommends documenting the 5PS and SUV of the main lesions in the interim FDG-PET report of patients receiving front-line treatment [19].

prognostic value

The panel was confident of the high prognostic value of interim FDG-PET when used during induction therapy with doxorubicin/bleomycin/vinblastine/dacarbazine (ABVD) (after 1–3 cycles) in immunocompetent and human immunodeficiency virus (HIV)-negative patients with classical HL [II, A] [8, 12, 15, 20–25].

The panel was also confident of the prognostic value of interim FDG-PET when used during induction therapy with anthracycline-containing regimens (after 2–4 cycles) in immunocompetent and HIV-negative patients with DLBCL [III, A] [11, 13, 16, 18, 26–29].

Finally, the panel recognised that interim FDG-PET is prognostic when used after reinduction chemotherapy and before a preplanned high-dose therapy programme in relapsed or refractory HL and DLBCL [III, A] [9, 10, 30–36].

panel recommendations for the use of interim PET for monitoring and intervention in clinical trials

recommendation 1.1. The exploratory use of interim FDG-PET as a surrogate test of chemosensitivity and as a diagnostic tool to facilitate clinical decision-making is encouraged in clinical trials in HL, DLBCL and other aggressive FDG-avid lymphoma entities.

Level of evidence: III
Strength of recommendation: B
Consensus: 100% yes (23 voters)

panel recommendations for the use of interim PET for monitoring and intervention in routine clinical practice

recommendation 1.2. There are little published data from randomised trials to support the use of an interim PET-driven therapeutic strategy in HL, DLBCL or other FDG-avid lymphomas. However, preliminary data strongly support the use of interim PET to tailor therapy in individual cases. On these grounds, results of interim PET may be applied in individual patients with early or advanced HL [20–22, 25, 37–40].

Level of evidence: II
Strength of recommendation: C
Consensus: 100% yes (23 voters)

recommendation 1.3. Based on the lack of therapeutic consequences, the routine clinical use of interim PET is not recommended in patients with DLBCL [13].

Level of evidence: II
Strength of recommendation: D
Consensus: 100% yes (23 voters)

recommendation 1.4. Based on the lack of data, the routine use of interim PET as a decision tool is discouraged in non-HL, non-DLBCL, FDG-avid lymphoma entities.

Level of evidence: V
Strength of recommendation: E
Consensus: 100% yes (23 voters)

2. The potential role of *TP53* mutations and deletions as a prognostic tool

The tumour suppressor gene *TP53* maps at 17p13 and codes for a central regulator of the deoxyribonucleic acid (DNA) damage-response pathway; its activation leads to cell cycle arrest and DNA repair, apoptosis or senescence [41, 42]. In lymphoid malignancies, *TP53* may be disrupted by chromosomal deletions, mutations or a combination of both. Overall, 95% of mutations are localised within the central DNA binding domain of *TP53*, impairing DNA binding and transactivation of target genes [41–43]. Deletion of the *TP53* locus at 17p13 is detectable by fluorescence *in situ* hybridisation (FISH), while identification of *TP53* mutations requires DNA sequencing, either Sanger sequencing or next-generation sequencing. The frequency of *TP53* disruption at the time of diagnosis varies across different types of lymphoid malignancies, and may progressively increase at the time of relapse or development of chemorefractory disease, as clearly documented in the case of CLL [44]. The fact that *TP53* disruption may be acquired during the disease course is important from a diagnostic perspective, requiring, where clinically indicated, the sequential analysis of the locus at each time of treatment requirement [45–48]. The clinical importance of *TP53* abnormalities in lymphoid malignancies is best demonstrated in the case of CLL, where *TP53* disruption is tightly linked to the poor prognosis marked by this genetic lesion and its close association with chemorefractory disease, as documented by a number of observational studies and prospective trials conducted both in the chemotherapy and immuno-chemotherapy eras [49–55]. However, there is evidence that *TP53* disruption predicts an adverse outcome also in other mature B-cell neoplasms [56, 57].

methodological considerations

broad availability. A complete analysis of *TP53* disruption requires the availability of both FISH and DNA sequencing. Analysis of *TP53* deletion by FISH is widely available in many haematological referral centres as well as in diagnostic laboratories dedicated to genetic disorders. Conversely, analysis of *TP53* mutations by Sanger sequencing is currently restricted to highly specialised centres, and is not widely available. The panel agrees that, at least in the context of CLL, a complete analysis of *TP53* disruption, including analysis of *TP53* mutations, should be prioritised because *TP53* disruption is the only well-established genetic marker which requires adaptation of treatment in CLL [45–48].

reproducibility and standardisation. FISH analysis for del17p13 is considered a well-standardised and reproducible technique. Sanger sequencing analysis for *TP53* mutations is technically well standardised and adequately reproducible in experienced laboratories. Until recently, inter-laboratory reproducibility has not been systematically assessed. However, the European Research Initiative on CLL (ERIC) has now implemented a quality control initiative for *TP53* mutations in many centres in Europe [58].

clarity of reporting system. Currently, there is no standardised reporting system for *TP53* analysis across different centres. Data derived from randomised trials supporting these recommendations were obtained using a cut-off for FISH of 10%–20% of positive cells by Sanger sequencing. Regarding *TP53* mutation analysis, the cut-off for mutation detection by Sanger sequencing can be generally estimated at 15%–20% of positive cells, although it may vary according to the precise nucleotide position and sequence. Inter-observer variability in the interpretation of electropherograms may also affect the detection threshold of Sanger sequencing; the use of dedicated software for mutation detection may reduce, at least in part, such variability. The precise description of *TP53* mutations should be documented according to the well-codified Human Genome Variation Society (HGVS) nomenclature system (www.hgvs.org/mutnomen). Mutations also need to be validated through the International Agency for Research on Cancer (IARC) *TP53* database (p53.iarc.fr). The GenBank reference sequence used for mutation detection should also be clearly stated in diagnostic reports.

prognostic value

The panel is confident with the general prognostic and predictive value of *TP53* disruption in CLL [I, A]. The panel is also confident with the general prognostic value of *TP53* disruption in other diseases, namely MCL, DLBCL and FL [II, B].

Many studies, both prospective and retrospective, have demonstrated that *TP53* disruption is associated with a poor prognosis in CLL [48–55]. In particular, the CLL8 trial of the German CLL Study Group clearly documented that both del17p13 and *TP53* mutation identify a very high-risk category of patients with CLL who were treated with fludarabine/cyclophosphamide/rituximab (FCR), an immuno-chemotherapy regimen that is the gold standard first-line treatment for fit patients with CLL [51, 55]. Notably, the poor prognosis associated with *TP53* disruption in CLL appears to be independent of the chemotherapeutic agents utilised [48–55]. However, this might potentially change when non-genotoxic drugs, such as ibrutinib, idelalisib and venetoclax, become part of routine practice.

panel recommendations for molecular and cytogenetic analysis of TP53 disruption in CLL and other lymphoid neoplasms for monitoring and intervention in clinical trials

recommendation 2.1. Given the well-established, prognostic and predictive value of *TP53* disruption in CLL, the panel strongly recommends the inclusion of *TP53* analysis, both by FISH and DNA sequencing, in clinical trials of CLL for intervention and monitoring purposes. In particular, the availability of new drugs that overcome *TP53*-mediated chemorefractory disease mandates the acquisition of *TP53* status for all patients with CLL at the time of screening procedures in trials in which one or more arms may be based on drugs that are known to be ineffective in *TP53*-disrupted CLL [58–62]. Therefore, the use of *TP53* screening before the start of treatment is highly encouraged in CLL.

Level of evidence: I
Strength of recommendation: A
Consensus: 100% yes (23 voters)

recommendation 2.2. In other lymphoid neoplasms, *TP53* screening for investigational purposes is neither recommended nor discouraged. At present, the panel discourages clinical trials aimed at specific interventions based on *TP53* status unless prognostic markers are the major focus of the trial and the drug being

evaluated has a strong biological rationale for overcoming *TP53*-mediated resistance.

Level of evidence: V
Strength of recommendation for investigation: C
Strength of recommendation for intervention: D
Consensus: 100% yes (23 voters)

panel recommendations for molecular and cytogenetic analysis of TP53 disruption in CLL and other lymphoid neoplasms for monitoring and intervention in clinical practice outside of clinical trials

recommendation 2.3. In CLL, the panel supports analysis of *TP53* disruption at the time of treatment requirement, both in first-line and subsequent lines of therapy. Reassessing *TP53* status in previously *TP53* wild-type CLL at relapse requiring treatment is relevant since *TP53* disruption may develop, or become detectable only at relapse. In routine practice, characterising *TP53* status in a given patient with CLL is clinically relevant as this may affect treatment decisions. The use of *TP53* screening by FISH and mutational analysis for monitoring and intervention in clinical practice is therefore encouraged in CLL, provided there is availability of and access to therapies overcoming *TP53*-mediated resistance [e.g. inhibitors of the B-cell receptor and allogeneic haematopoietic stem cell transplantation (allo-SCT)].

Level of evidence: I
Strength of recommendation: A
Consensus: 100% yes (23 voters)

recommendation 2.4. In other lymphoid neoplasms, the panel discourages the use of *TP53* outside of clinical trials as there is no general recommendation for treatment modification currently published. The results of currently recruiting trials might modify this attitude in the coming years.

Level of evidence: V
Strength of recommendation: E
Consensus: 100% yes (23 voters)

3. The potential role of COO determination by IHC or GEP as a prognostic tool

DLBCL is the most common form of lymphoma in the Western world [63]. It shows a wide spectrum of morphology and is biologically heterogeneous [63]. To identify biological entities within DLBCL, GEP has been applied to tumour samples of DLBCL [64–66]. The seminal study by Alizadeh et al. [64] was the first to recognise that DLBCL contains at least two biological entities, one with a GEP similar to the normal purified germinal centre B-cell (the germinal centre B-cell profile, or GCB) and the other similar to the profile produced by a purified, *in vitro* immunoglobulin M (IgM)-stimulated B-cell (the activated B-cell profile, or ABC). Consequently, DLBCL was commonly divided into these two subtypes, which show different clinical and molecular features. The robustness of this profile based on GEP has been confirmed in other studies [67–69].

As the use of high-throughput GEP was considered unfeasible in routine laboratory practice, there have been several attempts to simplify the procedures for COO determination. These attempts have gone in two directions, namely the identification of IHC surrogates and the application of GEP (either high-throughput or low-throughput) to formalin-fixed and paraffin-embedded (FFPE) samples.

Several IHC surrogate protocols use an algorithm to identify the COO in FFPE samples of DLBCL. Several algorithms have been published, including those by Colomo et al. [70], Hans et al. [71], Muris et al. [72], Choi et al. [73], Nyman et al. [74], Natkunam et al. [75], Meyer (better known as 'Tally') et al. [76] and Visco et al. [77]. Although these seem to work well as survival predictors when samples are stained and analysed in a single centre, the results are not easily transferrable to other laboratories [78–80]. Indeed, data from a large randomised clinical trial [81] and a meta-analysis have shown a limited role for IHC algorithms [82].

Given the limitations of IHC in terms of COO signature reproducibility, several groups have attempted to use FFPE as a source of RNA to identify the COO signature by high-throughput [83, 84] or low- to medium-throughput GEP techniques [85–91]. The results have been much more robust than those obtained by IHC, and findings from a meta-analysis have confirmed the usefulness of GEP approaches [82]. Most studies were retrospective, but two phase III clinical trials incorporating COO GEP on FFPE samples, namely the REMoDL-B study (ClinicalTrials.gov identifier NCT01324596 [92]), which uses the Illumina DASL platform (Illumina Inc., San Diego, CA), and the ROBUST study (ClinicalTrials.gov identifier NCT02285062 [93]), which uses the Nanostring nCounter-based Lymph2Cx platform (NanoString Technologies, Seattle, WA), are ongoing [94].

Against this background, the panel members discussed the adequacy for clinical use of COO-determining methods in DLBCL by both IHC and GEP.

methodological considerations

broad availability. IHC is widely available. Conversely, GEP technologies are currently limited to very specialised laboratories. The introduction of more user-friendly technologies (such as Nanostring nCounter) might render GEP more widely available and applicable in routine clinical practice in the near future.

reproducibility. IHC suffers from major reproducibility issues, which include inter-laboratory and inter-observer concordance, varying degrees of overlap with the gold standard GEP techniques and often poor correlation between the various algorithms available [78–82]. GEP using well-established high-throughput commercial chips is robust; however, inter-laboratory variability needs to be explored. So far, only one study using the Nanostring-based Lymph2Cx assay has assessed inter-laboratory agreement, with excellent results; the same test also showed excellent concordance for resampled biopsies and between different reagent lots [88, 91]. However, processing of samples would critically influence the outcome of GEP results and so particular care should be devoted to pre-analytical variables.

clarity of reporting systems. The reporting system for IHC has been standardised, with algorithms to clearly specify thresholds and procedures (e.g. the Hans classifier uses a 30% positive cell cut-off and a step-by-step algorithm), although few pathology reports specify the exact percentage of positive cells or even the algorithm used. For GEP, no standardised system for the interpretation of data or reporting of results is available.

prognostic value

The limitations of IHC algorithms have been described earlier. The panel also raised substantial concerns regarding the prognostic value of IHC. Conversely, several published studies support the general prognostic value of COO assessment by GEP in DLBCL, and so the panel was more confident in supporting this technical approach for prognostication [I, A] [82]. These considerations are particularly relevant with regard to drugs, which promise differential activity in germinal centre B-cell-like versus activated B-cell-like DLBCLs.

panel recommendations for the use of COO identification by IHC and GEP in clinical trials for monitoring and intervention

recommendation 3.1. Given the limitations of IHC, the panel does not encourage its use in prospective clinical trials for prognostication.

Level of evidence: I
Strength of recommendation: C
Consensus: 100% yes (23 voters)

recommendation 3.2. Given the limitations of IHC, the panel discourages its use in prospective clinical trials to guide intervention.

Level of evidence: I
Strength of recommendation: D
Consensus: 100% yes (23 voters)

recommendation 3.3. The panel strongly encourages the use of GEP in prospective clinical trials for prognostication.

Level of evidence: I
Strength of recommendation: A
Consensus: 100% yes (23 voters)

recommendation 3.4. Clinical trials of interventions based on GEP results are encouraged.

Level of evidence: I
Strength of recommendation: B
Consensus: 100% yes (23 voters)

panel recommendations for the use of COO identification by IHC and GEP for monitoring and intervention in routine clinical practice

recommendation 3.5. Based on inadequate standardisation, and a lack of well-designed interventional studies, the use of COO determination by IHC or GEP in DLBCL is generally not recommended in routine clinical practice outside of clinical trials.

Level of evidence: V
Strength of recommendation: D
Consensus: 100% yes (23 voters)

4. MRD evaluation by polymerase chain reaction-based methods and flow cytometry

MRD assessment can be used for the identification of different prognostic subgroups in patients with B-cell lymphomas and CLL, and is an excellent surrogate for treatment outcome [95–99]. Published evidence for the prognostic impact of MRD exists for MCL [96, 100, 101], FL [95, 96, 98, 102, 103] and CLL [97, 104–107]. In these entities, achievement of MRD response by conventional or intensified treatment is associated with prolonged progression-free survival (PFS) and OS independent of categorical response assessment and a favourable prognosis. Several prospective phase III trials using standardised approaches for MRD assessment have been published and demonstrate the prognostic relevance of MRD response in FL, MCL and CLL independently of treatment regimen or strategy and clinical risk parameters [95–99]. Indeed, the prognostic impact of MRD status has led to MRD being proposed as a secondary end point in ongoing clinical trials. In CLL, recent evidence suggests that MRD might also be used to identify candidates for dose de-escalations. Therefore, polymerase chain reaction (PCR)-based MRD evaluation is considered a promising prognosticator in MCL and FL, whereas MRD evaluation by flow cytometry is preferred in CLL.

methodological considerations

broad availability. Flow cytometry is generally available in Europe for CLL, but standardised four-colour flow to detect MRD at a level of 10-4 is only available in specialised institutions; real-time quantitative (RQ)-PCR is only available in specialised centres (EURO MRD network; www.euromrd.org).

reproducibility and standardisation. For RQ-PCR, reproducibility is excellent and methods are standardised and subjected to periodic quality controls at specialised institutions involved in the EURO MRD network. Flow-based MRD methods are currently harmonised, but not standardised, and inter-laboratory reproducibility has not been systematically assessed.

clarity of reporting systems. Reporting of molecular MRD results is standardised within established networks. Flow cytometry standardisation is currently ongoing within the EuroFlow network (http://www.euroflow.org).

prognostic value

The panel is confident of the general prognostic value of MRD evaluation in MCL [I, A], FL [I, A] and CLL [I, A].

Several phase III clinical trials have been carried out in FL [95, 102, 108, 109], CLL [97, 110, 111] and MCL [96] that clearly demonstrate the usefulness of MRD as a surrogate end point for monitoring treatment efficiency and for its prognostic value. Remarkably, in all three entities, the prognostic impact of MRD response on PFS and OS has been documented independent of treatment regimen, mostly in both peripheral blood and bone marrow.

panel recommendations for the use of MRD evaluation in clinical trials for monitoring and intervention

The panel felt confident that MRD monitoring of treatment response as an end point in clinical trials might facilitate the interpretation of results. Whether trials investigating MRD-based treatment tailoring might lead to substantial therapeutic improvement and treatment optimisation is an attractive but as yet unproven possibility. MRD assessment post-induction therapy is the most frequently assessed time point for MRD response as it is associated with a high prognostic impact and is therefore suitable to guide treatment intervention. Later time points during treatment are also of prognostic value and are suitable to

guide treatment intervention. However, so far, only one clinical trial in CLL has been published, which showed that MRD-based intervention (in terms of discontinuation of treatment once MRD negativity was seen) was associated with comparable PFS and OS independent of the number of courses of treatment received [110]. The panel therefore decided that more data are required to support the clinical benefit of treatment modification based on efficacy, as determined by MRD negativity.

recommendation 4.1. The use of MRD evaluation for monitoring and intervention in clinical trials is encouraged in MCL, FL and CLL:

(i) MCL: for monitoring [112, 113]:
Level of evidence: I
Strength of recommendation: B
(ii) MCL: for intervention [112, 113]:
Level of evidence: III
Strength of recommendation: C
(iii) FL: for monitoring [113]:
Level of evidence: I
Strength of recommendation: B
(iv) FL: for intervention [113]:
Level of evidence: IV
Strength of recommendation: C
(v) CLL: for monitoring (depends on the drug used):
Level of evidence: I
Strength of recommendation: B
(vi) CLL: for intervention:
Level of evidence: IV
Strength of recommendation: C
Consensus: 100% yes (23 voters)

panel recommendations for the use of MRD for monitoring and intervention in routine clinical practice

The panel does not support MRD evaluation for monitoring or intervention in routine practice outside of clinical trials as there is no general recommendation for treatment modification currently published. However, results of ongoing trials might modify this attitude in the coming years. The only exception is MRD assessment after allo-SCT, where it is a useful tool to monitor lymphoma regrowth and is more sensitive than currently used short tandem repeat analysis. In this setting, MRD can be used for discontinuation or intensification of immunosuppression [114].

recommendation 4.2. The use of MRD evaluation for monitoring and intervention in clinical practice is not recommended in MCL, FL and CLL, with the exception of monitoring after allo-SCT.

Level of evidence: V
Strength of recommendation: D
Consensus: 100% yes (23 voters)

acknowledgements

The authors thank Jennifer Lamarre, Claire Bramley, Matthew Wallace, Aude Galli and all ESMO staff for their support throughout the whole consensus process. Angela Corstorphine of Kstorfin Medical Communications Ltd provided medical writing support with the preparation of this manuscript. This support was funded by ESMO.

funding

All costs relating to the consensus conference were covered from the European Society for Medical Oncology central funds. There was no external funding of the event or manuscript production.

disclosure

CB (honoraria from Roche, Pfizer, Celgene, Pharmacyclics and Janssen; research grants from Roche and Janssen), MD (advisory boards for Bayer, Celgene, Gilead, Janssen and Pfizer; support for academic trials from Celgene, Janssen, Mundipharma, Pfizer and Roche; speaker's honoraria from Celgene, Janssen, Mundipharma, Pfizer and Roche), AJMF (advisory boards for Celgene, Gilead and Mundipharma; speaker honoraria from Italfarmaco, Mundipharma and Gilead; research support from Celgene, Roche, Mundipharma and Rhizen), GG (advisory boards for Roche, Janssen, Amgen, Novartis, GlaxoSmithKline, Karyopharm and Morphosys; speaker honoraria from Roche, Janssen, Gilead and GlaxoSmithKline), VG (speaker honoraria, advisory boards, consultancy and travel grants from Roche; speaker honoraria from Mundipharma, GlaxoSmithKline and Bristol-Myers Squibb), MH (speaker honoraria, advisory boards and consultancy from Takeda; advisory boards and consultancy from Janssen, Celgene and Roche/Genentech), ML (honoraria from Celgene, Janssen-Cilag, Roche, Amgen, Mundipharma and Teva; research contracts with Celgene, Pfizer, Mundipharma and Roche; research grants from Amgen, Roche and Takeda), SLG (speaker honoraria, advisory boards, consultancy and travel grants from Roche, Janssen-Cilag and Celgene; research programme funding from Roche, Janssen and Servier), UM (honoraria for advisory board activities from Roche, Celgene, Amgen and Janssen-Cilag; research grants from Mundipharma and Celgene), PdNB (speaker's bureau/lectures for Roche and Bayer), MP (advisory boards for Celgene, Roche and Spectrum; research support from Amgen, Roche and Spectrum), NS (travel grants, advisory boards and research support from Roche; travel grants and advisory boards from Celgene; travel grants and honoraria for company symposia from Riemser Pharma and CTI Life Sciences), PS (advisory boards for Celgene and Teva; honoraria from Spectrum Pharmaceuticals and Pierre Fabre Oncology; travel grants from Hospira, Celgene and Teva), MS (speaker's bureau and advisory boards for Teva, Mundipharma, Janssen-Cilag, Gilead, CTI BioPharma and Servier; research support, consultant, speaker's bureau and advisory boards for Menarini; speaker's bureau for Roche and Takeda), RS (research support from Celgene, Novartis and Teva; honoraria from Celgene, Novartis and Teva; scientific advisory board for Celgene), ASB (consultancy fees from Takeda), MT (consultancy for Roche, Gilead, Janssen, Takeda and Celgene), JW (advisory boards for Roche, Celgene, Janssen-Cilag and Takeda; research grants from Roche, Celgene, Janssen-Cilag, Takeda, GlaxoSmithKline and Seattle Genetics), UW (speaker honoraria from Amgen, Novartis, Roche and Chugai), EZ (advisory honoraria and/or support for investigator-initiated studies for the

institution from Celgene, Johnson and Johnson/Janssen, Gilead, Mundipharma, Roche and Bayer). All remaining authors have declared no conflicts of interest.

references

1. Dykewicz CA. Summary of the guidelines for preventing opportunistic infections among hematopoietic stem cell transplant recipients. Clin Infect Dis 2001; 33: 139–144.
2. Dreyling M, Ghielmini M, Marcus R et al. Newly diagnosed and relapsed follicular lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2014; 25(Suppl 3): iii76–iii82.
3. Eichenauer DA, Engert A, André M et al. Hodgkin's lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2014; 25(Suppl 3): iii70–iii75.
4. Tilly H, Gomes da Silva M, Vitolo U et al. Diffuse large B-cell lymphoma (DLBCL): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2015; 26(Suppl 5): v116–v125.
5. Eichhorst B, Robak T, Montserrat E et al. Chronic lymphocytic leukaemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2015; 26(Suppl 5): v78–v84.
6. Dreyling M, Geisler C, Hermine O et al. Newly diagnosed and relapsed mantle cell lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2014; 25(Suppl 3): iii83–iii92.
7. Barrington SF, Mikhaeel NG, Kostakoglu L et al. Role of imaging in the staging and response assessment of lymphoma: consensus of the International Conference on Malignant Lymphomas Imaging Working Group. J Clin Oncol 2014; 32: 3048–3058.
8. Gallamini A, Hutchings M, Rigacci L et al. Early interim 2-[18F]fluoro-2-deoxy-D-glucose positron emission tomography is prognostically superior to international prognostic score in advanced-stage Hodgkin's lymphoma: a report from a joint Italian-Danish study. J Clin Oncol 2007; 25: 3746–3752.
9. Smeltzer JP, Cashen AF, Zhang Q et al. Prognostic significance of FDG-PET in relapsed or refractory classical Hodgkin lymphoma treated with standard salvage chemotherapy and autologous stem cell transplantation. Biol Blood Marrow Transplant 2011; 17: 1646–1652.
10. Moskowitz AJ, Yahalom J, Kewalramani T et al. Pretransplantation functional imaging predicts outcome following autologous stem cell transplantation for relapsed and refractory Hodgkin lymphoma. Blood 2010; 116: 4934–4937.
11. Safar V, Dupuis J, Itti E et al. Interim [18F]fluorodeoxyglucose positron emission tomography scan in diffuse large B-cell lymphoma treated with anthracycline-based chemotherapy plus rituximab. J Clin Oncol 2012; 30: 184–190.
12. Hutchings M, Loft A, Hansen M et al. FDG-PET after two cycles of chemotherapy predicts treatment failure and progression-free survival in Hodgkin lymphoma. Blood 2006; 107: 52–59.
13. Moskowitz CH, Schöder H, Teruya-Feldstein J et al. Risk-adapted dose-dense immunochemotherapy determined by interim FDG-PET in advanced-stage diffuse large B-cell lymphoma. J Clin Oncol 2010; 28: 1896–1903.
14. Meignan M, Gallamini A, Haioun C, Polliack A. Report on the Second International Workshop on interim positron emission tomography in lymphoma held in Menton, France, 8–9 April 2010. Leuk Lymphoma 2010; 51: 2171–2180.
15. Gallamini A, Barrington SF, Biggi A et al. The predictive role of interim positron emission tomography for Hodgkin lymphoma treatment outcome is confirmed using the interpretation criteria of the Deauville five-point scale. Haematologica 2014; 99: 1107–1113.
16. Itti E, Meignan M, Berriolo-Riedinger A et al. An international confirmatory study of the prognostic value of early PET/CT in diffuse large B-cell lymphoma: comparison between Deauville criteria and ΔSUVmax. Eur J Nucl Med Mol Imaging 2013; 40: 1312–1320.
17. Biggi A, Gallamini A, Chauvie S et al. International validation study for interim PET in ABVD-treated, advanced-stage Hodgkin lymphoma: interpretation criteria and concordance rate among reviewers. J Nucl Med 2013; 54: 683–690.
18. Casasnovas RO, Meignan M, Berriolo-Riedinger A et al. SUVmax reduction improves early prognosis value of interim positron emission tomography scans in diffuse large B-cell lymphoma. Blood 2011; 118: 37–43.
19. Meignan M, Gallamini A, Haioun C et al. Report on the 5th International Workshop on Positron Emission Tomography in Lymphoma held in Menton, France, 19–20 September 2014. Leuk Lymphoma 2015; 56: 1229–1232.
20. Johnson PW, Federico M, Fossa A et al. Response-adapted therapy based on interim FDG-PET scans in advanced Hodgkin lymphoma: first analysis of the safety of de-escalation and efficacy of escalation in the international RATHL study (CRUK/07/033). Hematol Oncol 2015; 33(Suppl 1): 102 (abstr 008).
21. Raemaekers J. Early FDG-PET adapted treatment improved the outcome of early FDG-PET positive patients with stages I/II Hodgkin lymphoma (HL): final results of the randomized Intergroup EORTC/LYSA/FIL H10 trial. In 13th International Conference on Malignant Lymphoma, Lugano, Switzerland, 2015.
22. Zinzani PL, Broccoli A, Gioia DM et al. Interim positron emission tomography response-adapted therapy in advanced-stage Hodgkin lymphoma: final results of the phase II part of the HD0801 study. J Clin Oncol 2016; 34: 1376–1385.
23. Hutchings M, Kostakoglu L, Zaucha JM et al. In vivo treatment sensitivity testing with positron emission tomography/computed tomography after one cycle of chemotherapy for Hodgkin lymphoma. J Clin Oncol 2014; 32: 2705–2711.
24. Hutchings M, Mikhaeel NG, Fields PA et al. Prognostic value of interim FDG-PET after two or three cycles of chemotherapy in Hodgkin lymphoma. Ann Oncol 2005; 16: 1160–1168.
25. Radford J, Illidge T, Counsell N et al. Results of a trial of PET-directed therapy for early-stage Hodgkin's lymphoma. N Engl J Med 2015; 372: 1598–1607.
26. Mikhaeel NG, Hutchings M, Fields PA et al. FDG-PET after two to three cycles of chemotherapy predicts progression-free and overall survival in high-grade non-Hodgkin lymphoma. Ann Oncol 2005; 16: 1514–1523.
27. Haioun C, Itti E, Rahmouni A et al. [18F]fluoro-2-deoxy-D-glucose positron emission tomography (FDG-PET) in aggressive lymphoma: an early prognostic tool for predicting patient outcome. Blood 2005; 106: 1376–1381.
28. Cashen AF, Dehdashti F, Luo J et al. 18F-FDG PET/CT for early response assessment in diffuse large B-cell lymphoma: poor predictive value of international harmonization project interpretation. J Nucl Med 2011; 52: 386–392.
29. Itti E, Lin C, Dupuis J et al. Prognostic value of interim 18F-FDG PET in patients with diffuse large B-Cell lymphoma: SUV-based assessment at 4 cycles of chemotherapy. J Nucl Med 2009; 50: 527–533.
30. Mocikova H, Pytlik R, Markova J et al. Pre-transplant positron emission tomography in patients with relapsed Hodgkin lymphoma. Leuk Lymphoma 2011; 52: 1668–1674.
31. Becherer A, Mitterbauer M, Jaeger U et al. Positron emission tomography with [18F]2-fluoro-D-2-deoxyglucose (FDG-PET) predicts relapse of malignant lymphoma after high-dose therapy with stem cell transplantation. Leukemia 2002; 16: 260–267.
32. Terasawa T, Dahabreh IJ, Nihashi T. Fluorine-18-fluorodeoxyglucose positron emission tomography in response assessment before high-dose chemotherapy for lymphoma: a systematic review and meta-analysis. Oncologist 2010; 15: 750–759.
33. Dickinson M, Hoyt R, Roberts AW et al. Improved survival for relapsed diffuse large B cell lymphoma is predicted by a negative pre-transplant FDG-PET scan following salvage chemotherapy. Br J Haematol 2010; 150: 39–45.
34. Cremerius U, Fabry U, Wildberger JE et al. Pre-transplant positron emission tomography (PET) using fluorine-18-fluoro-deoxyglucose (FDG) predicts outcome in patients treated with high-dose chemotherapy and autologous stem cell transplantation for non-Hodgkin's lymphoma. Bone Marrow Transplant 2002; 30: 103–111.
35. Spaepen K, Stroobants S, Dupont P et al. Prognostic value of pretransplantation positron emission tomography using fluorine 18-fluorodeoxyglucose in patients with aggressive lymphoma treated with high-dose chemotherapy and stem cell transplantation. Blood 2003; 102: 53–59.
36. Filmont JE, Czernin J, Yap C et al. Value of F-18 fluorodeoxyglucose positron emission tomography for predicting the clinical outcome of patients with aggressive lymphoma prior to and after autologous stem-cell transplantation. Chest 2003; 124: 608–613.
37. Raemaekers JM, André MP, Federico M et al. Omitting radiotherapy in early positron emission tomography-negative stage I/II Hodgkin lymphoma is associated with an increased risk of early relapse: clinical results of the preplanned interim analysis of the randomized EORTC/LYSA/FIL H10 trial. J Clin Oncol 2014; 32: 1188–1194.
38. Gallamini A, Rossi A, Patti C et al. Interim PET-adapted chemotherapy in advanced Hodgkin lymphoma (HL). Results of the second interim analysis of the Italian GITIL/FIL DH0607 trial. Hematol Oncol 2015; 33(Suppl S1): 163–164 (abstr 118).
39. Johnson P, Federico M, Kirkwood A et al. Adapted treatment guided by interim PET-CT scan in advanced Hodgkin's lymphoma. N Engl J Med 2016; 374: 2419–2429.

40. Press OW, Li H, Schöder H et al. US Intergroup trial of response-adapted therapy for stage III to IV Hodgkin lymphoma using early interim fluorodeoxyglucose-positron emission tomography imaging: Southwest Oncology Group S0816. J Clin Oncol 2016; 34: 2020–2027.
41. Soussi T. The TP53 gene network in a postgenomic era. Hum Mutat 2014; 35: 641–642.
42. Soussi T, Wiman KG. TP53: an oncogene in disguise. Cell Death Differ 2015; 22: 1239–1249.
43. Zenz T, Vollmer D, Trbusek M et al. TP53 mutation profile in chronic lymphocytic leukemia: evidence for a disease specific profile from a comprehensive analysis of 268 mutations. Leukemia 2010; 24: 2072–2079.
44. Foà R, Del Giudice I, Guarini A et al. Clinical implications of the molecular genetics of chronic lymphocytic leukemia. Haematologica 2013; 98: 675–685.
45. Hallek M, Cheson BD, Catovsky D 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.
46. Oscier D, Dearden C, Eren E et al. Guidelines on the diagnosis, investigation and management of chronic lymphocytic leukaemia. Br J Haematol 2012; 159: 541–564.
47. Pospisilova S, Gonzalez D, Malcikova J et al. ERIC recommendations on TP53 mutation analysis in chronic lymphocytic leukemia. Leukemia 2012; 26: 1458–1461.
48. Stilgenbauer S. Prognostic markers and standard management of chronic lymphocytic leukemia. Hematology Am Soc Hematol Educ Program 2015; 2015: 368–377.
49. Döhner H, Stilgenbauer S, Benner A et al. Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med 2000; 343: 1910–1916.
50. Catovsky D, Richards S, Matutes E et al. Assessment of fludarabine plus cyclophosphamide for patients with chronic lymphocytic leukaemia (the LRF CLL4 Trial): a randomised controlled trial. Lancet 2007; 370: 230–239.
51. Hallek M, Fischer K, Fingerle-Rowson G et al. Addition of rituximab to fludarabine and cyclophosphamide in patients with chronic lymphocytic leukaemia: a randomised, open-label, phase 3 trial. Lancet 2010; 376: 1164–1174.
52. Zenz T, Eichhorst B, Busch R et al. TP53 mutation and survival in chronic lymphocytic leukemia. J Clin Oncol 2010; 28: 4473–4479.
53. Gonzalez D, Martinez P, Wade R et al. Mutational status of the TP53 gene as a predictor of response and survival in patients with chronic lymphocytic leukemia: results from the LRF CLL4 trial. J Clin Oncol 2011; 29: 2223–2229.
54. Rossi D, Khiabanian H, Spina V et al. Clinical impact of small TP53 mutated subclones in chronic lymphocytic leukemia. Blood 2014; 123: 2139–2147.
55. Stilgenbauer S, Schnaiter A, Paschka P et al. Gene mutations and treatment outcome in chronic lymphocytic leukemia: results from the CLL8 trial. Blood 2014; 123: 3247–3254.
56. O'Shea D, O'Riain C, Taylor C et al. The presence of TP53 mutation at diagnosis of follicular lymphoma identifies a high-risk group of patients with shortened time to disease progression and poorer overall survival. Blood 2008; 112: 3126–3129.
57. Xu-Monette ZY, Wu L, Visco C et al. Mutational profile and prognostic significance of TP53 in diffuse large B-cell lymphoma patients treated with R-CHOP: report from an International DLBCL Rituximab-CHOP Consortium Program Study. Blood 2012; 120: 3986–3996.
58. Pospisilova S, Sutton LA, Malcikova J et al. Innovation in the prognostication of chronic lymphocytic leukemia: how far beyond TP53 gene analysis can we go? Haematologica 2016; 101: 263–265.
59. Byrd JC, Brown JR, O'Brien S et al. Ibrutinib versus ofatumumab in previously treated chronic lymphoid leukemia. N Engl J Med 2014; 371: 213–223.
60. Dreger P, Schetelig J, Andersen N et al. Managing high-risk CLL during transition to a new treatment era: stem cell transplantation or novel agents? Blood 2014; 124: 3841–3849.
61. Furman RR, Sharman JP, Coutre SE et al. Idelalisib and rituximab in relapsed chronic lymphocytic leukemia. N Engl J Med 2014; 370: 997–1007.
62. Burger JA, Tedeschi A, Barr PM et al. Ibrutinib as initial therapy for patients with chronic lymphocytic leukemia. N Engl J Med 2015; 373: 2425–2437.
63. Stein H, Warnke R, Chan WC et al. Diffuse large B-cell lymphoma, not otherwise specified. In SH Swerdlow, E Campo, NL Harris et al. (eds), WHO Classification of Tumours of Hematopoietic and Lymphoid Tissues. Lyon: IARC 2008; 233–237.
64. Alizadeh AA, Eisen MB, Davis RE et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 2000; 403: 503–511.
65. Monti S, Savage KJ, Kutok JL et al. Molecular profiling of diffuse large B-cell lymphoma identifies robust subtypes including one characterized by host inflammatory response. Blood 2005; 105: 1851–1861.
66. Shipp MA, Ross KN, Tamayo P et al. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. Nat Med 2002; 8: 68–74.
67. Lenz G, Wright G, Dave SS et al. Stromal gene signatures in large-B-cell lymphomas. N Engl J Med 2008; 359: 2313–2323.
68. Rosenwald A, Wright G, Chan WC et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. N Engl J Med 2002; 346: 1937–1947.
69. Wright G, Tan B, Rosenwald A et al. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. Proc Natl Acad Sci USA 2003; 100: 9991–9996.
70. Colomo L, López-Guillermo A, Perales M et al. Clinical impact of the differentiation profile assessed by immunophenotyping in patients with diffuse large B-cell lymphoma. Blood 2003; 101: 78–84.
71. Hans CP, Weisenburger DD, Greiner TC et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. Blood 2004; 103: 275–282.
72. Muris JJ, Meijer CJ, Vos W et al. Immunohistochemical profiling based on Bcl-2, CD10 and MUM1 expression improves risk stratification in patients with primary nodal diffuse large B cell lymphoma. J Pathol 2006; 208: 714–723.
73. Choi WW, Weisenburger DD, Greiner TC et al. A new immunostain algorithm classifies diffuse large B-cell lymphoma into molecular subtypes with high accuracy. Clin Cancer Res 2009; 15: 5494–5502.
74. Nyman H, Adde M, Karjalainen-Lindsberg ML et al. Prognostic impact of immunohistochemically defined germinal center phenotype in diffuse large B-cell lymphoma patients treated with immunochemotherapy. Blood 2007; 109: 4930–4935.
75. Natkunam Y, Farinha P, Hsi ED et al. LMO2 protein expression predicts survival in patients with diffuse large B-cell lymphoma treated with anthracycline-based chemotherapy with and without rituximab. J Clin Oncol 2008; 26: 447–454.
76. Meyer PN, Fu K, Greiner TC et al. Immunohistochemical methods for predicting cell of origin and survival in patients with diffuse large B-cell lymphoma treated with rituximab. J Clin Oncol 2011; 29: 200–207.
77. Visco C, Li Y, Xu-Monette ZY et al. Comprehensive gene expression profiling and immunohistochemical studies support application of immunophenotypic algorithm for molecular subtype classification in diffuse large B-cell lymphoma: a report from the International DLBCL Rituximab-CHOP Consortium Program Study. Leukemia 2012; 26: 2103–2113.
78. de Jong D, Rosenwald A, Chhanabhai M et al. Immunohistochemical prognostic markers in diffuse large B-cell lymphoma: validation of tissue microarray as a prerequisite for broad clinical applications—a study from the Lunenburg Lymphoma Biomarker Consortium. J Clin Oncol 2007; 25: 805–812.
79. Coutinho R, Clear AJ, Owen A et al. Poor concordance among nine immunohistochemistry classifiers of cell-of-origin for diffuse large B-cell lymphoma: implications for therapeutic strategies. Clin Cancer Res 2013; 19: 6686–6695.
80. Gutiérrez-García G, Cardesa-Salzmann T, Climent F et al. Gene-expression profiling and not immunophenotypic algorithms predicts prognosis in patients with diffuse large B-cell lymphoma treated with immunochemotherapy. Blood 2011; 117: 4836–4843.
81. Ott G, Ziepert M, Klapper W et al. Immunoblastic morphology but not the immunohistochemical GCB/nonGCB classifier predicts outcome in diffuse large B-cell lymphoma in the RICOVER-60 trial of the DSHNHL. Blood 2010; 116: 4916–4925.
82. Read JA, Koff JL, Nastoupil LJ et al. Evaluating cell-of-origin subtype methods for predicting diffuse large B-cell lymphoma survival: a meta-analysis of gene expression profiling and immunohistochemistry algorithms. Clin Lymphoma Myeloma Leuk 2014; 14: 460–467.e2.
83. Barrans SL, Crouch S, Care MA et al. Whole genome expression profiling based on paraffin embedded tissue can be used to classify diffuse large B-cell lymphoma and predict clinical outcome. Br J Haematol 2012; 159: 441–453.
84. Williams PM, Li R, Johnson NA et al. A novel method of amplification of FFPET-derived RNA enables accurate disease classification with microarrays. J Mol Diagn 2010; 12: 680–686.
85. Collie AM, Nölling J, Divakar KM et al. Molecular subtype classification of formalin-fixed, paraffin-embedded diffuse large B-cell lymphoma samples on the ICEPlex® system. Br J Haematol 2014; 167: 281–285.
86. Masqué-Soler N, Szczepanowski M, Kohler CW et al. Molecular classification of mature aggressive B-cell lymphoma using digital multiplexed gene expression on

formalin-fixed paraffin-embedded biopsy specimens. Blood 2013; 122: 1985–1986.
87. Roberts RA, Sabalos CM, LeBlanc ML et al. Quantitative nuclease protection assay in paraffin-embedded tissue replicates prognostic microarray gene expression in diffuse large-B-cell lymphoma. Lab Invest 2007; 87: 979–997.
88. Scott DW, Wright GW, Williams PM et al. Determining cell-of-origin subtypes of diffuse large B-cell lymphoma using gene expression in formalin-fixed paraffin-embedded tissue. Blood 2014; 123: 1214–1217.
89. Veldman-Jones MH, Lai Z, Wappett M et al. Reproducible, quantitative, and flexible molecular subtyping of clinical DLBCL samples using the NanoString nCounter System. Clin Cancer Res 2015; 21: 2367–2378.
90. Xue X, Zeng N, Gao Z, Du MQ. Diffuse large B-cell lymphoma: sub-classification by massive parallel quantitative RT-PCR. Lab Invest 2015; 95: 113–120.
91. Scott DW, Mottok A, Ennishi D et al. Prognostic significance of diffuse large B-cell lymphoma cell of origin determined by digital gene expression in formalin-fixed paraffin-embedded tissue biopsies. J Clin Oncol 2015; 33: 2848–2856.
92. NCT01324596. A randomised evaluation of molecular guided therapy for diffuse large B-cell lymphoma with bortezomib (REMoDL-B). 2015. https://clinicaltrials.gov/ct2/show/NCT01324596 (10 August 2016, date last accessed).
93. NCT02285062. Efficacy and safety study of lenalidomide plus R-CHOP chemotherapy versus placebo plus R-CHOP chemotherapy in untreated ABC type diffuse large B-cell lymphoma (ROBUST). 2015. https://clinicaltrials.gov/ct2/show/NCT02285062 (10 August 2016, date last accessed).
94. Scott DW. Cell-of-origin in diffuse large B-cell lymphoma: are the assays ready for the clinic? Am Soc Clin Oncol Educ Book 2015; 35: e458–e466.
95. Galimberti S, Luminari S, Ciabatti E et al. Minimal residual disease after conventional treatment significantly impacts on progression-free survival of patients with follicular lymphoma: the FIL FOLL05 trial. Clin Cancer Res 2014; 20: 6398–6405.
96. Pott C, Hoster E, Delfau-Larue MH et al. Molecular remission is an independent predictor of clinical outcome in patients with mantle cell lymphoma after combined immunochemotherapy: a European MCL intergroup study. Blood 2010; 115: 3215–3223.
97. Böttcher S, Ritgen M, Fischer K et al. Minimal residual disease quantification is an independent predictor of progression-free and overall survival in chronic lymphocytic leukemia: a multivariate analysis from the randomized GCLLSG CLL8 trial. J Clin Oncol 2012; 30: 980–988.
98. Ladetto M, Lobetti-Bodoni C, Mantoan B et al. Persistence of minimal residual disease in bone marrow predicts outcome in follicular lymphomas treated with a rituximab-intensive program. Blood 2013; 122: 3759–3766.
99. Lobetti-Bodoni C, Mantoan B, Monitillo L et al. Clinical implications and prognostic role of minimal residual disease detection in follicular lymphoma. Ther Adv Hematol 2013; 4: 189–198.
100. Liu H, Johnson JL, Koval G et al. Detection of minimal residual disease following induction immunochemotherapy predicts progression free survival in mantle cell lymphoma: final results of CALGB 59909. Haematologica 2012; 97: 579–585.
101. Pott C, Schrader C, Gesk S et al. Quantitative assessment of molecular remission after high-dose therapy with autologous stem cell transplantation predicts long-term remission in mantle cell lymphoma. Blood 2006; 107: 2271–2278.
102. Ladetto M, De Marco F, Benedetti F et al. Prospective, multicenter randomized GITMO/IIL trial comparing intensive (R-HDS) versus conventional (CHOP-R) chemoimmunotherapy in high-risk follicular lymphoma at diagnosis: the superior disease control of R-HDS does not translate into an overall survival advantage. Blood 2008; 111: 4004–4013.
103. Rambaldi A, Carlotti E, Oldani E et al. Quantitative PCR of bone marrow BCL2/IgH+ cells at diagnosis predicts treatment response and long-term outcome in follicular non-Hodgkin lymphoma. Blood 2005; 105: 3428–3433.
104. Böttcher S, Fischer K, Stilgenbauer S et al. Quantitative MRD assessments predict progression free survival in CLL patients treated with fludarabine and cyclophosphamide with or without rituximab—a prospective analysis in 471 patients from the randomized GCLLSG CLL8 trial. Blood 2008; 112: 125–126 (abstr 326).
105. Bosch F, Ferrer A, Villamor N et al. Fludarabine, cyclophosphamide, and mitoxantrone as initial therapy of chronic lymphocytic leukemia: high response rate and disease eradication. Clin Cancer Res 2008; 14: 155–161.
106. Fink AM, Böttcher S, Ritgen M et al. Prediction of poor outcome in CLL patients following first-line treatment with fludarabine, cyclophosphamide and rituximab. Leukemia 2013; 27: 1949–1952.
107. Hillmen P, Cohen DR, Cocks K et al. A randomized phase II trial of fludarabine, cyclophosphamide and mitoxantrone (FCM) with or without rituximab in previously treated chronic lymphocytic leukaemia. Br J Haematol 2011; 152: 570–578.
108. Hirt C, Schüler F, Kiefer T et al. Rapid and sustained clearance of circulating lymphoma cells after chemotherapy plus rituximab: clinical significance of quantitative t(14;18) PCR monitoring in advanced stage follicular lymphoma patients. Br J Haematol 2008; 141: 631–640.
109. Ladetto M, Corradini P, Vallet S et al. High rate of clinical and molecular remissions in follicular lymphoma patients receiving high-dose sequential chemotherapy and autografting at diagnosis: a multicenter, prospective study by the Gruppo Italiano Trapianto Midollo Osseo (GITMO). Blood 2002; 100: 1559–1565.
110. Strati P, Keating MJ, O'Brien SM et al. Eradication of bone marrow minimal residual disease may prompt early treatment discontinuation in CLL. Blood 2014; 123: 3727–3732.
111. Goede V, Fischer K, Busch R et al. Obinutuzumab plus chlorambucil in patients with CLL and coexisting conditions. N Engl J Med 2014; 370: 1101–1110.
112. Andersen NS, Pedersen LB, Laurell A et al. Pre-emptive treatment with rituximab of molecular relapse after autologous stem cell transplantation in mantle cell lymphoma. J Clin Oncol 2009; 27: 4365–4370.
113. Ferrero S, Monitillo L, Mantoan B et al. Rituximab-based pre-emptive treatment of molecular relapse in follicular and mantle cell lymphoma. Ann Hematol 2013; 92: 1503–1511.
114. Ritgen M, Böttcher S, Stilgenbauer S et al. Quantitative MRD monitoring identifies distinct GVL response patterns after allogeneic stem cell transplantation for chronic lymphocytic leukemia: results from the GCLLSG CLL3X trial. Leukemia 2008; 22: 1377–1386.

appendix

ESMO Lymphoma Consensus Conference Panel Members

C Buske, Comprehensive Cancer Center Ulm and Department of Internal Medicine III, Institute of Experimental Cancer Research, University Hospital, Ulm, Germany; M Dreyling, Medizinische Klinik III, Klinikum der Universität München/LMU, Munich, Germany; AJM Ferreri, Unit of Lymphoid Malignancies, Department of Onco-Hematology, IRCCS San Raffaele Scientific Institute, Milan, Italy; P Fields, Department of Haematology, Guys and St Thomas' and King's College Hospitals, London, UK; G Gaidano, Division of Hematology, Department of Translational Medicine, Amedeo Avogadro University of Eastern Piedmont, Novara, Italy; V Goede, Department of Internal Medicine, University Hospital Cologne, Germany; M Hutchings, Department of Hematology, Rigshospitalet, Copenhagen, Denmark; M Ladetto, Hematology Division, Azienda Ospedaliera Santi Antonio e Biagio e Cesare Arrigo, Alessandria, Italy; SL Gouill, Centre Hospitalo-Universitaire de Nantes, UMR892 team 10, CIC Nantes, France; S Luminari, Hematology, Arcispedale S. Maria Nuova, IRCCS Reggio Emilia, and Department of Diagnostic, Clinical and Public Health Medicine, University of Modena and Reggio Emilia, Modena, Italy; U Mey, Department of Oncology and Haematology, Kantonsspital Graubünden, Chur, Switzerland; PdN Brown, Department of Hematology, Rigshospitalet, Copenhagen, Denmark; M Pfreundschuh, Innere Medizin I, University Klinik des Saarlandes, Hamburg, Germany; C Pott, Second Medical Department, University Hospital Schleswig-Holstein, Kiel, Germany; N Schmitz, Department of Hematology, Oncology and Stem Cell Transplantation, Asklepios Klinik St Georg, Hamburg, Germany; P Soubeyran,

Department of Medical Oncology, Institut Bergonié, Comprehensive Cancer Centre, Bordeaux, France; M Spina, Division of Medical Oncology A, National Cancer Institute, Aviano, Italy; R Stauder, Haematology and Oncology Department, Innsbruck Medical University, Innsbruck, Austria; AS Balari, Servei d'Hematologia, Institut Català d'Oncologia—Hospital Duran i Reynals, Barcelona, Spain; M Trněný, Institute of Hematology and Blood Transfusion, 1st Department of Medicine, 1st Faculty of Medicine, Charles University, General Hospital, Prague, Czech Republic; Gv Imhoff, Section of Hematology, University of Groningen, Groningen, The Netherlands; J Walewski, Department of Lymphoid Malignancies, Maria Sklodowska-Curie Memorial Cancer Centre and Institute of Oncology, Warsaw, Poland; U Wedding, Department of Palliative Care, University Hospital, Jena, Germany; A Zamò, Department of Diagnostics and Public Health, University of Verona, Verona, Italy; E Zucca, Lymphoma Unit, Oncology Institute of Southern Switzerland, Ospedale San Giovanni, Bellinzona, Switzerland.

Annals of Oncology 27: 2160–2167, 2016
doi:10.1093/annonc/mdw307

Statistical controversies in clinical research: prognostic gene signatures are not (yet) useful in clinical practice

S. Michiels[1,2]*, N. Ternès[1,2] & F. Rotolo[1,2]

[1]*Gustave Roussy, Service de Biostatistique et d'Epidémiologie, Villejuif;* [2]*Université Paris-Saclay, Université Paris-Sud, UVSQ, CESP, INSERM U1018, Villejuif, France*

Received 5 January 2016; revised 4 May 2016 and 20 June 2016; accepted 25 July 2016

With the genomic revolution and the era of targeted therapy, prognostic and predictive gene signatures are becoming increasingly important in clinical research. They are expected to assist prognosis assessment and therapeutic decision making. Notwithstanding, an evidence-based approach is needed to bring gene signatures from the laboratory to clinical practice. In early breast cancer, multiple prognostic gene signatures are commercially available without having formally reached the highest levels of evidence-based criteria. We discuss specific concepts for developing and validating a prognostic signature and illustrate them with contemporary examples in breast cancer. When a prognostic signature has not been developed for predicting the magnitude of relative treatment benefit through an interaction effect, it may be wishful thinking to test its predictive value. We propose that new gene signatures be built specifically for predicting treatment effects for future patients and outline an approach for this using a cross-validation scheme in a standard phase III trial. Replication in an independent trial remains essential.

Key words: gene signature, prognostic, predictive, evidence based, clinical utility

introduction

Molecular signatures are becoming increasingly important for anticipating the prognosis of individual patients ('prognostic' biomarkers) or for predicting how individual patients will respond to specific treatments ('predictive' biomarkers, more generally called 'treatment-effect modifiers'). A voluminous literature of >150 000 papers documenting thousands of claimed biomarkers has been produced in medicine of which fewer than 100 have been validated for routine clinical practice [1]. Indeed, <20 prognostic or predictive biomarkers are recognized with variable levels of evidence in the 2014 European Society of Medical Oncology (ESMO) clinical practice guidelines for lung, breast, colon and prostate cancer [2].

In early breast cancer, while several clinical prediction models exist based on clinical and pathological (CP) characteristics, such as age, tumor size, nodal status, tumor grade, estrogen receptor, at least six different gene signatures are commercially available (Onco*type* DX, MammaPrint, Genomic Grade Index, PAM50, Breast Cancer Index and EndoPredict). The concordance of predicted risk categories of the different gene signatures for individual patients is moderate [3, 4], as illustrated by recent OPTIMA study which evaluated—among others—the two well-known tests Mammaprint (low/high) and Onco*type* Dx (≤25 versus >25) on 302 patients in a head-to-head comparison and found a low level of agreement, i.e. a kappa value of 0.40 (95% CI 0.30–0.49) [5]. Of course, even when repeating the same assay twice on a single tumor sample, some inherent degree of inaccuracy would be expected but unlikely to this extent. This has led to a pretty awkward situation where the treatment decision for adjuvant chemotherapy does not depend anymore on the clinician but on the

*Correspondence to: Dr. Stefan Michiels, Service de Biostatistique et d'Epidémiologie, Gustave Roussy, B2M RDC, 114 rue Edouard Vaillant, 94805 Villejuif Cedex, France. Tel: +33-1-42-11-41-44; E-mail: stefan.michiels@gustaveroussy.fr