

## ORIGINAL ARTICLE

## Risk-adapted modulation through de-intensification of cancer treatments: an ESMO classification

D. Trapani<sup>1,2†</sup>, M. A. Franzoi<sup>3†</sup>, H. J. Burstein<sup>2</sup>, L. A. Carey<sup>4</sup>, S. Delaloge<sup>5</sup>, N. Harbeck<sup>6</sup>, D. F. Hayes<sup>7</sup>, K. Kalinsky<sup>8</sup>, L. Pusztai<sup>9</sup>, M. M. Regan<sup>10</sup>, I. Sestak<sup>11</sup>, T. Spanic<sup>12</sup>, J. Sparano<sup>13</sup>, S. Jezdic<sup>14</sup>, N. Cherny<sup>15</sup>, G. Curigliano<sup>16\*</sup> & F. Andre<sup>3\*</sup>

<sup>1</sup>New Drugs Development for Innovative Therapies, European Institute of Oncology, IRCCS, Milan, Italy; <sup>2</sup>Department of Medical Oncology, Dana-Farber Cancer Center, Boston, USA; <sup>3</sup>INSERM Unit 981 – Molecular Predictors and New Targets in Oncology, PRISM Center for Precision Medicine, Gustave Roussy, Villejuif, France; <sup>4</sup>University of North Carolina at Chapel Hill, Lineberger Comprehensive Cancer Center, Chapel Hill, USA; <sup>5</sup>Breast Cancer Unit, Department of Medical Oncology, Gustave Roussy, Villejuif, France; <sup>6</sup>Breast Center, Department of Obstetrics & Gynecology and Comprehensive Cancer Center Munich, LMU University Hospital, Munich, Germany; <sup>7</sup>University of Michigan Rogel Cancer Center, Ann Arbor; <sup>8</sup>Department of Hematology and Medical Oncology, Winship Cancer Institute at Emory University, Atlanta; <sup>9</sup>Yale Cancer Center Genetics and Genomics Program, Yale Cancer Center, Yale School of Medicine, New Haven; <sup>10</sup>Division of Biostatistics, Dana-Farber Cancer Institute, Harvard Medical School, Boston, USA; <sup>11</sup>Wolfson Institute of Preventive Medicine – Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK; <sup>12</sup>ESMO Patient Advocates Working Group, Ljubljana, Slovenia; <sup>13</sup>Division of Hematology/Oncology, Icahn School of Medicine at Mt. Sinai, Tisch Cancer Institute, New York, USA; <sup>14</sup>Scientific and Medical Division, European Society for Medical Oncology, Lugano, Switzerland; <sup>15</sup>Department of Medical Oncology, Shaare Zedek Medical Center, Jerusalem, Israel; <sup>16</sup>Department of Oncology and Hemato-Oncology, University of Milan, European Institute of Oncology, IRCCS, Milan, Italy

Available online XXX

**Background:** The landscape of clinical trials testing risk-adapted modulations of cancer treatments is complex. Multiple trial designs, endpoints, and thresholds for non-inferiority have been used; however, no consensus or convention has ever been agreed to categorise biomarkers useful to inform the treatment intensity modulation of cancer treatments.

**Methods:** An expert subgroup under the European Society for Medical Oncology (ESMO) Precision Medicine Working Group shaped an international collaborative project to develop a classification system for biomarkers used in the cancer treatment de-intensification, based on a tiered approach. A group of disease-oriented clinical, translational, methodology and public health experts, and patients' representatives provided an analysis of the status quo, and scanned the horizon of ongoing clinical trials. The classification was developed through multiple rounds of expert revisions and inputs.

**Results:** The working group agreed on a univocal definition of treatment de-intensification. Evidence of reduction in the dose-density, intensity, or cumulative dose, including intermittent schedules or shorter treatment duration or deletion of segment(s) of the standard regimens, compound(s), or treatment modality must be demonstrated, to define a treatment de-intensification. De-intensified regimens must also portend a positive impact on toxicity, quality of life, health system burden, or financial toxicity. ESMO classification categorises the biomarkers for treatment modulation in three tiers, based on the level of evidence. Tier A includes biomarkers validated in prospective, randomised, non-inferiority clinical trials. The working group agreed that in non-inferiority clinical trials, boundaries are highly dependent upon the disease scenario and endpoint being studied and that the absolute differences in the outcomes are the most relevant measures, rather than relative differences. Biomarkers tested in single-arm studies with a threshold of non-inferiority are classified as Tier B. Tier C is when the validation occurs in prospective-retrospective quality cohort investigations.

**Conclusions:** ESMO classification for the risk-guided intensity modulation of cancer treatments provides a set of evidence-based criteria to categorise biomarkers deemed to inform de-intensification of cancer treatments, in risk-defined patients. The classification aims at harmonising definitions on this matter, therefore offering a common language for all the relevant stakeholders, including clinicians, patients, decision-makers, and for clinical trials.

**Key words:** de-intensification, treatment modulation, risk-adapted, precision medicine, biomarkers, classification system

\*Correspondence to: Prof. Giuseppe Curigliano, ESMO Head Office – Scientific and Medical Division, Via Ginevra 4, Lugano CH-6900, Switzerland. Tel: +41-91-973-1999; Fax: +41-91-973-1902

E-mail: [education@esmo.org](mailto:education@esmo.org) (G. Curigliano).

\*Prof. Fabrice Andre, ESMO Head Office – Scientific and Medical Division, Via Ginevra 4, Lugano CH-6900, Switzerland. Tel: +41-91-973-1999; Fax: +41-91-973-1902

E-mail: [education@esmo.org](mailto:education@esmo.org) (F. Andre).

†Co-first authors.

0923-7534/© 2022 The Author(s). Published by Elsevier Ltd on behalf of European Society for Medical Oncology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

### INTRODUCTION

The steep increase in cure rates and prolonged overall survival (OS) achieved in various disease settings in recent years has brought attention to an important facet of personalised medicine: the opportunity to individualise and tailor treatments without compromising efficacy.<sup>1,2</sup> A better knowledge of biology and prognostic determinants of single cancers in single patients allows for a stratified modulation of the treatment plan. Modulation can take the form of

**Table 1. Overview of selected studies evaluating risk-adapted intensity modulation of cancer treatments**

Disease (setting)	Risk definition	Standard treatment	Alternative treatment
Hodgkin's lymphoma (stage I-III, favourable) <sup>17-19</sup>	PET response	CT and RT	CT
Childhood acute lymphoblastic leukaemia <sup>a</sup> (standard risk) <sup>20-22</sup>	MRD	Standard induction and maintenance CT	De-intensified post-induction CT
Breast cancer (adjuvant) <sup>23,24</sup>	GEP test	CT and ET	ET
Breast cancer (adjuvant) <sup>25</sup>	Clinical and pathological factors	Standard duration, anthracycline and taxane-containing CT and trastuzumab	Shorter duration, taxane-only containing CT and trastuzumab
Breast cancer (adjuvant) <sup>7</sup>	Clinical and pathological factors	12-Month trastuzumab	6-Month trastuzumab
Colorectal cancer (adjuvant) <sup>26</sup>	Clinical and pathological factors	6-Month CT	3-Month CT <sup>b</sup>
HPV+ oropharyngeal cancer (LA) <sup>27</sup>	p16 IHC	Standard dose PORT	Reduced dose PORT
Testicular germ cell tumour (good risk) <sup>28</sup>	Clinical and pathological factors	BEP	EP

BEP, bleomycin, etoposide and platinum regimen; CT, chemotherapy; EP, etoposide and platinum regimen; ET, endocrine therapy; GEP, gene expression profile; HPV+, human papillomavirus-associated cancer; IHC, immunohistochemistry; LA, locally advanced; MRD, minimal residual disease; PET, positron emission tomography; PORT, post-operative radiation therapy; RT, radiation therapy.

<sup>a</sup>Patients 1-18 years old.

<sup>b</sup>Non-inferiority not confirmed in the overall population, but in a subgroup of risk-defined patients.

treatment intensification for patients at very high risk or with adverse biological features, or de-escalation (or de-intensification) for patients with lower risk or favourable biology. This form of de-escalating modulation aims to provide equivalent benefits in terms of survival benefits but with reduced patient burden and adverse effects.

De-escalating treatment modulation generates important patient benefits for suitable selected patients. It can reduce treatment burden, rates of short-term and life-threatening adverse events, chronic, long-term toxicities, and lessen health-related financial distress.<sup>3</sup> Since long-term or late adverse events can adversely impact overall life expectancy,<sup>4</sup> there may be long-term survival benefits. Importantly, this approach could increase or preserve the quality of life of patients, with equivalent oncology outcomes or with an acceptable trade-off.<sup>5,6</sup> At a health care system level, de-escalated modulation of treatments can have economic benefits of enhanced sustainability, and broader access to cancer care.<sup>7,8</sup>

Historically, some of the original and groundbreaking clinical trials testing de-intensified treatments in patients with solid tumours were first carried out in patients with breast cancer, with efforts to reduce the duration of adjuvant chemotherapy<sup>9</sup> and the extension of local therapies,<sup>10</sup> whereas the concept of reducing treatment intensity is well established in haematological and paediatric oncology. Table 1 presents some examples of risk-adapted intensity modulation of cancer treatments.

An enhanced knowledge on tumour subtypes and other prognostic factors facilitates the identification of patients who may benefit from less intense treatments.<sup>11</sup> Rigorous research methodology is critical to identify de-escalation modulation strategies which minimise toxicity without compromising survival. Hitherto, the evaluation of such regimens has been inconsistent, using different definitions and supported by diverse research methodologies.<sup>12-30</sup> Although frameworks for biomarker-driven cancer treatment choices have been developed,<sup>13-16</sup> none has focused

specifically on the topic of treatment modulation through de-intensification.<sup>15-26</sup>

This classification aims to introduce a standard nomenclature and taxonomy and to provide guidance for clinical research design to provide adequate evidence of the safety and efficacy of proposed de-escalated treatment modulations.

## METHODS

Responding to a perceived need, the European Society for Medical Oncology (ESMO) Precision Medicine Working Group, supported by the ESMO leadership, established a dedicated and interdisciplinary multi-stakeholder expert panel, with disease-oriented and public health experts, biostatisticians, and patient advocacy representatives.

The aim of this ESMO expert subgroup for de-intensified treatment modulation was to systematise and categorise the research evidence to inform therapeutic decisions of treatment de-intensification, based on a tiered approach. Preliminary research of the literature was carried out, to understand the landscape and scan the horizon of ongoing clinical studies, and presented to the group. The expert subgroup met periodically online to develop the concept of a classification, ultimately consolidating the final draft.

## RESULTS

### Treatment de-intensification definition

The definitions retrieved in literature and in clinical guidelines for treatment modulation are heterogeneous, and substantially variable. Different forms of treatment de-intensifications are context-pertinent, based on the specific setting of care.

**Deletions.** One common way to de-intensify treatments is to forgo one segment of the standard treatment. For example, a regimen composed of sequential chemotherapy

Table 2. Definition of treatment modulation	
A: Treatment attributes	Treatment (or a segment of) is <b>omitted</b> <b>Dose</b> density/dose intensity or cumulative dose is lower than SoC <b>Duration</b> is shorter than SoC treatment
B: Treatment impact and patient burden <sup>a</sup>	Toxicity is estimated to be lower: - randomised trials: there are clinically and statistically significantly less grade 3-4 toxicities or low-grade persisting AEs impacting on daily well-being - non-randomised trials/single-arm studies: the comparison with the SoC toxicity profile (historical arm/extrapolated) is favourable QoL is expected to be improved: - endpoint QoL showed a clinically meaningful improvement Treatment costs or health care system burden is <b>significantly reduced</b>
Treatment modulation through de-escalation applies if $\geq 1$ item(s) are demonstrated in cluster A (treatment attributes) and cluster B (treatment impact) items in the context of not worsening clinical outcome.	

AEs, adverse events; QoL, quality of life; SoC, standard of care.

<sup>a</sup>These analyses should not be *post hoc* or exploratory endpoints. The de-intensified regimens must provide non-inferior survival outcomes, and overall demonstrate reducing the treatment burden for patients.

segments or polychemotherapy can be de-escalated by providing only a portion of the entire treatment.<sup>29,30</sup>

**Shortened duration of treatment.** De-intensified treatment modulation may occur with reduced treatment duration, leading to less treatment-related burden and risk of toxicities.<sup>7,9,26</sup> An example of this is the use of 3 months adjuvant therapy rather than 6 months for patients with lower risk stage III colon cancer.<sup>26</sup>

**Combined (deleted and foreshortened).** One example of this approach occurs when a treatment modality is entirely forgone or skipped. For instance, in a disease setting in which a combination of chemo- and radiotherapy is a standard approach, selected patients can report non-inferior survival outcomes with a single treatment modality. This can shorten treatment, reduce therapeutic burden, lessen toxicities, improve tolerability, and reduce risk of long-term sequelae. An example is in the response-adapted treatment of patients with Hodgkin's lymphoma, based on *ad interim* imaging: patients with a complete response can receive less intensive chemotherapy regimens, with non-inferior outcomes.<sup>17-19</sup> In the management of solid tumours, this is well illustrated in the adjuvant treatment of breast cancer, where validated genomic prognostic and/or predictive tools can identify patients in whom it is safe and reasonable to forgo the systemic cytotoxic treatments.<sup>23,24,31,32</sup>

The working group criteria for the designation of a modulation of cancer treatment through de-intensification require that two conditions are met: (i) clear evidence of treatment de-escalation along with (ii) tangible evidence of improved patient outcomes (Table 2). Clear evidence of treatment de-escalation requires evidence of reduction in the dose-density, intensity, or cumulative dose, including intermittent schedules or shorter treatment duration or omission of one or more segments of the standard

regimens or one or more compounds or treatment modality. Tangible evidence of improved impact requires evidence of non-inferior survival outcomes along with statistically significant and clinically meaningful patient benefits such as reduced burden and/or clinically relevant toxicities, and/or improved quality of life, and/or reduced costs (Table 2).

### Research methodologies for de-intensification of cancer treatments

**General.** This expert group advocates a stepwise clinical research approach: (i) harm-minimisation regimens could be initially assessed using stand-alone randomised phase II studies that recruit and report before proceeding to phase III studies; (ii) if no detriment in clinical outcome is identified, then a phase III trial could follow.

### Non-inferiority clinical trials comparing the standard of care treatments with the alternative, de-intensified regimens

The gold standard study design to prove that a de-escalated regimen is advantageous and safe is a randomised, controlled, non-inferiority clinical trial comparing the standard of care treatments with the alternative, de-intensified regimens. Non-inferiority trials must capture clinical equivalence (Figure 1).

**General comments.** These studies present challenges, starting with the definition of the non-inferiority margin and of clinical equivalence. They usually require large sample size, international collaboration, long recruitment period, and heavy financial investments.<sup>12,33-36</sup>

**Non-inferiority margin.** By convention, the minimum acceptable difference is expressed as a relative risk [e.g. hazard ratio (HR)] or absolute risk (e.g. delta), that has been largely addressed in the literature, and hardly reconciled.<sup>33</sup> In general, the decision to set a specific threshold of non-inferiority is based on a consultative process of multiple stakeholders, and includes pragmatic considerations on feasibility.<sup>12,35</sup> Recently, in a review of factors contributing to bias in clinical research, the ESMO-Magnitude of Clinical Benefit Scale (ESMO-MCBS) Working Group has noted that the defined non-inferiority margins might be too lenient, especially in the context of non-inferiority trials, expressing a concern that treatments with true inferiority may seem non-inferior in inappropriate investigations.<sup>33</sup>

When enrolling lower risk populations, this working group would suggest that the non-inferiority margin be defined in absolute terms (i.e. delta), rather than the relative risk reduction, and that the absolute differences should always be reported.

**Evaluating harms and benefits.** The reduction in toxicity and/or improvement of quality of life should be based on a rigorous methodology, utilising evidence-based and validated tools. For toxicity, the classification will consider chronic, persistent, or disabling adverse events, the burden

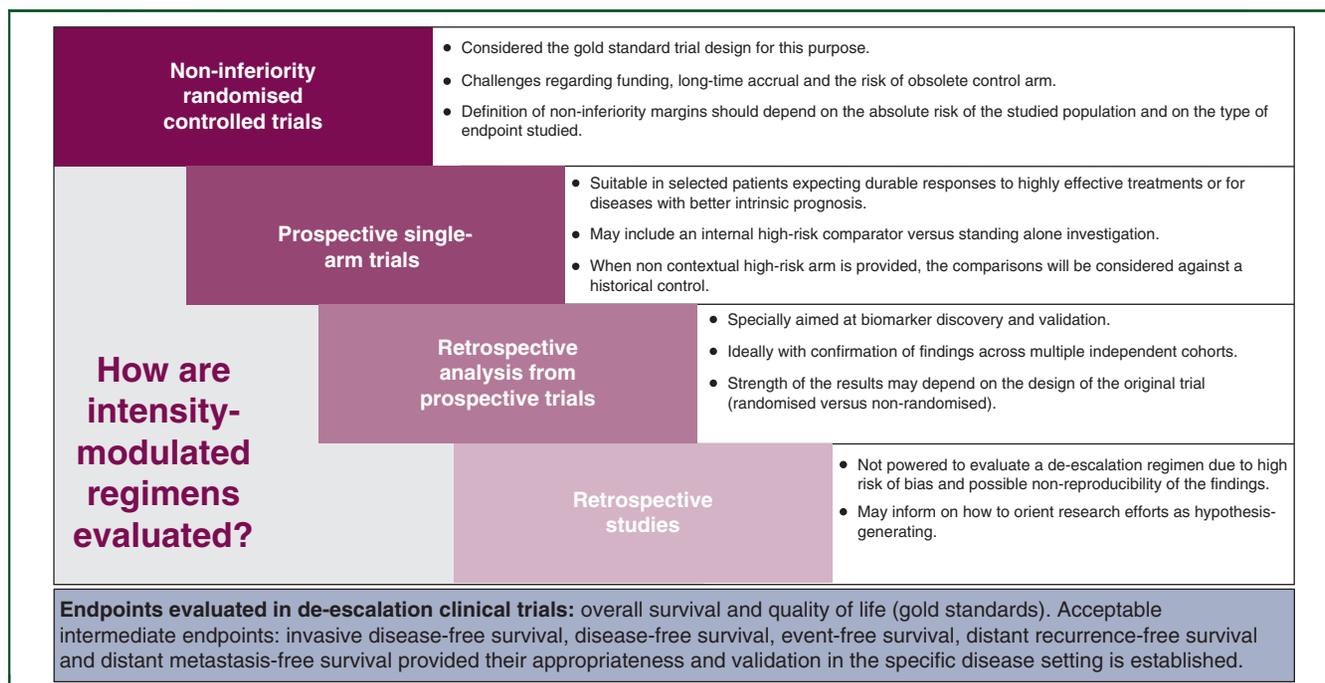


Figure 1. Clinical trial designs selected for investigating intensity-modulated cancer treatments.

of low-grade toxicities affecting the daily well-being, and the severe acute toxic events, including if transitory and self-limiting (e.g. some haematological toxicities). Of note, physicians may underestimate treatment-related toxicities.<sup>37,38</sup> Therefore, the working group believes that patient-reported outcomes (PROs) are of preeminent importance and should be part of the clinical trial endpoints: de-intensified regimens must ultimately prove an improved treatment experience as assessed through PROs.

**Challenges.** Non-inferiority trials are generally not seen as attractive to the pharmaceutical industry, and often need to be conducted through academic efforts, often with public funding.

**Examples.** The International Duration Evaluation of Adjuvant Chemotherapy (IDEA) research protocol, which conglomerated six distinct datasets from prospective trials for the modulation of adjuvant chemotherapy in patients with resected colorectal cancer, based on a non-inferiority hypothesis.<sup>26</sup> In this study, the investigators agreed on an acceptable upper limit of the two-sided 95% confidence interval of the HR (HR upper) not exceeding 1.12 for the disease-free survival (DFS).

Another paradigmatic example is the design of the TAILORx trial, for the cohort of patients with intermediate risk of disease recurrence, as estimated with the genomic tool Oncotype Dx® [i.e. recurrence score (RS) 11-25]. In this study, the investigators posed a non-inferiority margin at HR = 1.322 for invasive DFS (iDFS) to state an acceptable risk by omitting systemic cytotoxic chemotherapy in the adjuvant setting for node-negative disease.<sup>23</sup>

In the setting of early human epidermal growth factor receptor 2 (HER2)-positive breast cancer, the non-inferiority

trial of 6 versus 12 months of adjuvant trastuzumab (PERSEPHONE trial) was based on an acceptable risk of absolute non-inferiority margin of 3% of DFS.<sup>7</sup>

**Derived lessons.** These examples show that the decision for non-inferiority margins across disease types, subtypes, and settings is highly variable, and dependent on the absolute risk of the studied population. In addition, this working group acknowledges that the absolute non-inferiority threshold should also depend on the type of endpoint studied. The consequences of taking a potentially 3% higher risk due to treatment modulation are quite different if the primary endpoint for comparison is (i) OS rate in a curable cancer versus, (ii) surrogate for cure at a specific time point (i.e. 3-year DFS), (iii) median survival with metastatic cancer, or (iv) a surrogate for survival in metastatic cancer such as PFS.

#### Single-arm prospective studies

**General comments.** This approach presents a lower level of evidence. It may be suitable for the evaluation of a de-intensified treatment of a selected patient population expecting durable responses to highly effective treatments or for diseases with intrinsic better prognosis.<sup>12,16</sup>

**Single arm design.** In single-arm designs, the aim of the study would be based on a prognostic hypothesis of non-inferiority, setting a threshold for non-inferiority of the relevant outcome based on historic control. Single-arm trials with no high-risk comparisons (i.e. compared only with a historical control) lack external validation, and consequently the clinical implementation is sometimes debated, or conditional on the performance of a confirmatory randomised, controlled trial (RCT).

**Contextual.** Single-arm studies can be contextual to larger clinical trials with multiple risk-defined cohorts of patients such that the higher-risk population may serve as an internal control, to assure consistency of the findings and reproducibility of the prognostic biomarker.

**Examples.** The TAILORx clinical trial enrolled patients with low-risk breast cancer [i.e. recurrence score < 11] in a prospective, non-randomised cohort, in the context of high-risk arms of the same trial (i.e. recurrence score  $\geq$  11). The strategy of a chemotherapy-free adjuvant treatment of patients with node-negative, hormone receptor-positive breast cancer with low risk of recurrence was tested based on the hypothesis of non-inferiority, with a margin of distant relapse-free interval at 10 years of  $\geq$ 95%, in a single-arm investigation.<sup>23</sup>

In the MINDACT clinical trial, using the prognostic signature MammaPrint™, for patients with high-risk clinical features and low-risk gene expression profile, the hypothesis of non-inferiority was tested against a minimum acceptable risk, at a lower boundary of the 95% confidence interval for the 5-year distant metastasis-free survival (DMFS) of  $\geq$ 92%.<sup>24</sup>

Also, for patients with HER2-positive early breast cancer, the opportunity to de-intensify chemotherapy by sparing anthracyclines and alkylators, reducing the total treatment duration and toxicities was tested in the APT single-arm study. The investigators provided a maximally acceptable risk of iDFS at 3 years of 95%.<sup>25</sup>

**Derived lessons.** There is no consensus on the best thresholds of non-inferiority in single-arm investigations, resulting in large variabilities of endpoints and thresholds across the studies. Most of these studies enrolled low-risk populations, unlikely to derive great benefit by treatment intensifications.

#### Retrospective analyses of prospective cohorts

Non-inferiority can also be retrieved from retrospective analyses of prospective cohorts, aimed at biomarker discovery and validation.<sup>15</sup> In the most accepted design, multiple confirmations of the findings are requested, across independent cohorts. Therefore, a retrospective study for the validation of a biomarker for treatment modulation from a single prospective cohort is considered at high risk of bias, and not recommended to be used to support implementation in clinical practice by this working group.

**Example.** The validation of the EndoPredict® multi-parametric assay occurred in prospective-retrospective analyses of women enrolled in two phase III clinical trials.<sup>39</sup> The authors demonstrated an independent prognostic significance of EndoPredict®-associated risk score (i.e. EPclin score), capable of discriminating between low and high risk of recurrence (i.e. EndoPredict® EPclin score <3.3287). This tool was granted approval for clinical use in the USA, and intended to inform on the prognosis of patients and aid in treatment decision making.

#### The ESMO classification for grading the evidence supporting de-intensified treatment modulation: a tiered approach

This working group acknowledges the past efforts of ESMO in delivering clear and easy-to-use tools, aiming to improve clinical practice, patient's self-empowerment, and orient clinical research. Based on the past experiences with the ESMO-MCBS (which scores the clinical benefit of a therapeutic strategy, based on two distinct scales on the curative and non-curative settings) and the ESMO Scale for Clinical Actionability of molecular Targets (ESCAT) (the level of evidence supporting proposed therapies based on identified tumour genomic alterations), and the current clinical trial methodology, patient-centred dialogues, potential social and health economy implications,<sup>40-63</sup> a tiered approach to evaluate the level of evidence in support of de-escalation treatment modulation was selected.

Three categories of tiers were created, from A to C, based on the robustness of the evidence to prove the non-inferiority [Table 3](#).

- Tier A accommodates the biomarkers clinically validated in prospective, randomised, non-inferiority clinical trials, considered to be the gold standard trial designs for this purpose. Intensity-modulated regimens are compared head-to-head with standard of care treatments, to demonstrate non-inferior OS or other intermediate acceptable outcome metrics, along with improved safety and tolerability, and/or better quality of life. Such a difference in safety/tolerability or quality of life must be statistically significant, based on the original statistical plan, as components of the endpoints of the study. Oncotype Dx® for patients with node-negative recurrence score 11-25 and the clinical-pathological factors guiding the reduction of the adjuvant treatment duration for colorectal cancer would fit in this definition of biomarkers for treatment modulation.<sup>23,26</sup>
- Tier B is intended for biomarkers tested in the single-arm investigations, contextual or not to larger clinical trials with external high-risk group comparisons. The control is exploited against a historical control, as agreed by the steering committees of the studies, and derived from quality datasets. Tier B biomarkers carry the risk to investigate hyper-selected populations, especially when there is no trial-contextual high-risk cohort for control, suggesting that caveats could be raised for the universal clinical implementation, in specific settings. The use of Oncotype Dx® in patients with recurrence score < 11, of MammaPrint™ for women with high-risk clinical features and low genomic risk, of metabolic imaging-guided omission of radiation therapy in non-bulky, early-stage Hodgkin's lymphoma setting, and of the minimal residual-disease triggered treatment de-intensification for children with acute lymphoblastic leukaemia would fit in this tier.<sup>17,21,23,24</sup>
- Tier C biomarkers are validated in quality, prospectively defined retrospective reanalyses of prospective cohorts

**Table 3. ESMO classification for intensity modulation of cancer treatments**

Tier	Level of evidence	Definition of level of evidence	Outcomes	Magnitude of difference in the primary endpoint
A	Level IA	Prospective, randomised, non-inferiority clinical trials	De-escalated regimens provide non-inferior outcomes versus SoC therapy	The benefits have been demonstrated with strong evidence (statistically significant, clinically meaningful).
B	Level IB	Single-arm, prospective trial	De-escalated regimens provide non-inferior outcomes versus historical control or contextual higher-risk arm	The comparator is deemed to be adequate, mirroring the population receiving an SoC therapy. The outcome in the single-arm population is excellent. The risk for hyper-selection should be highlighted.
C	Level IC	Prospective-retrospective validation from prospective cohorts	The biomarker is validated in quality retrospective analysis from adequate samples from prospective cohorts	The risk of a sub-treatment in the evaluated population is deemed low. <ul style="list-style-type: none"> <li>• C1: from prospective, randomised cohorts</li> <li>• C2: from prospective, non-randomised cohorts</li> </ul>

SoC, standard of care.

(typically clinical trial cohorts). The principal condition to qualify in this tier is that the analysis is robust, and involves most of the original population, with a sample adequately sized to prove the non-inferiority, in more than a single cohort, for validation. Also, there must be agreement that the risk of a substandard treatment in the selected population for the de-escalated regimens is reasonably low. Some Tier C biomarkers have granted, in the past, enough evidence to receive regulatory approval for human use. Regardless of the position of the decision makers, however, challenges can still emerge from this group of biomarkers, and multiple confirmations in independent cohorts are recommended. The use of EPclin score to inform the direction of treatments and Breast Cancer Index to guide extended adjuvant endocrine therapy would fit in this definition.<sup>32,39,64</sup> Tier C can be subdivided in a sub-tier C1, if the cohorts were initially randomised, and C2, in case they were not. In principle, Tier C can also accommodate non-inferiority exploratory secondary reanalysis of clinical trials. This can be the case of studies that are negative for superiority, as in the original statistical design, but proving to be non-inferior under another hypothesis, provided the tested treatment(s) fulfils the criteria for treatment de-intensification in this classification.<sup>65</sup>

The classification can be applied for several biomarkers informing on treatment modulation, including predictive biomarkers, treatment resistance biomarkers, and pharmacogenomic biomarkers supporting dose reductions.

Retrospective only or *post hoc* exploratory analyses do not qualify for this classification but can inform on how to orient the research efforts as hypothesis generating: the toxicity reduction or impact on quality of life should be part of the initial trial design, as primary or secondary endpoints consistent with ESMO-MCBS rules.<sup>62</sup>

The opportunity to compare the population with a synthetic control arm is still open, and not part of the discussions of this classification. Currently, whether a biomarker is validated on the base of a synthetic control arm, it cannot qualify for the present classification. The same assumption can be provided for systematic review and/or meta-analysis or clinical trials originally designed to define a superiority of

a regimen over another one, under a non-inferiority hypothesis—if qualifying for Tier C.

## DISCUSSION

### *Practical issues regarding the trial methodology*

**Trial endpoints.** This working group agreed that relevant endpoints for de-intensification studies are survival endpoints, safety, and quality of life. While OS is largely seen as the gold standard, many clinical trials testing de-escalated regimens use surrogates of OS to assess the hypothesis of non-inferiority.<sup>38-40</sup> Despite the acknowledged limitations of their predicative reliability for improved survival and or quality of life,<sup>33</sup> their use as primary outcome is arguably desirable to ensure that the study results are not published with obsolete comparators.<sup>43-46</sup> Ultimately, most of the clinical trials of treatment modulation will report OS among the secondary endpoints; the majority will also account safety and quality of life, which are crucial endpoints, in this setting.

In the curative setting, the working group considered DFS, event-free survival, distant recurrence-free survival, and DMFS and iDFS as acceptable surrogate endpoints for the consideration in this classification, provided their appropriateness and validation in the specific disease setting is established. DMFS may be preferred due to its clinical relevance, since most solid tumours are not curable in the metastatic setting.<sup>2</sup> DMFS does not, however, consider local-regional recurrences, which can be impacted by de-escalated treatment regimens and have considerable impact in patients' quality of life, usually requiring subsequent invasive procedures. Also, these surrogate endpoints may be more appropriate in high-income settings, where competitive causes of mortality are less pronounced, and are minimised through efficient population health interventions within functioning and well-resourced health systems.

### *Biomarker validation*

Molecular biomarker assessment deserves robust and validated methodology. While validating biomarkers in the context of clinical trials based on robust endpoints is critical, the process to identify valid biomarkers might be more

complex. Analytical validity should be pursued to understand the technical performance of a biomarker, across the value chain, from discovery to clinical use. In fact, having some correlations of possible biomarkers with the prognosis or treatment response is not commonly sufficient, and good trial designs must be pursued, to clinically implement new biomarkers and support their clinical utility.<sup>15,66</sup>

**Public health outcomes.** Treatment modulation may impact on public health outcomes by reducing treatment costs or health care social and economic burden. While clinical trials infrequently include economic metrics as endpoints, these data could impact on the uptake of the clinical biomarkers for treatment modulation.

One traditional economic method of measuring value in health care is the quality-adjusted life year (QALY), which combines morbidity (quality of life) with mortality (quantity of life) to estimate the value of a specific health intervention. More recently, the incremental cost-effectiveness ratio (ICER), has been also emphasised by Health Technology Assessment bodies, to estimate the economic value of new health interventions. The analysis of quality-adjusted time without symptoms of disease progression or toxicity of treatment (Q-TWiST) has eventually been proposed to better portray the toxicity-efficacy trade-off, although not inclusive of economic considerations.<sup>48</sup> While a universal metric to deliberate around the opportunity to uptake, implement, or disengage from clinical practices based on QALYs or ICERs is unlikely to exist, the context/pertinence of the underlying models for economic estimations may inform on a national and regional health priority setting.<sup>49</sup> Presently, there is no obvious consensus about a threshold of cost-effectiveness to warrant additional expenditure.<sup>50</sup> Countries, however, may set their own rules based on their Gross Domestic Product and budget impact, most commonly, and in conjunction with broader considerations under multi-stakeholder, multidisciplinary Health Technology Assessment criteria.<sup>51,52</sup> Economic considerations are solicited more commonly when the treatment modulation requires the evaluation of new biomarkers, especially if based on high-cost technologies or requiring structural investments, or for more expensive medicines, with consequential impact on resource utilisations. In the absence of common rules to accept or reject new health technologies for treatment modulation, including a trade-off between lives saved, financial costs, and economic implications, the decisions across the countries are very variable, and based on different criteria.<sup>52</sup> Efforts to validate clinical endpoints for the description of the benefits of de-escalated regimens should occur in parallel to efforts to define pathways and of more coherent decision making, prompting harmonisation, to enhance high-value investments and ensure value for money.

**Patient selection.** This working group agreed that in clinical trials investigating treatment modulation through de-intensification of standard therapy, patients should be carefully selected and eligibility criteria tightly defined. In non-inferiority RCTs and contextual single-arm studies,

different patient groups can be studied in the same trial by ensuring stratification and adequate power for analysis of the subgroups.

**Data monitoring and early stopping.** Any trial assessing the impact of treatment modulation should have stringent stopping criteria and frequent monitoring by an independent data and safety monitoring committee triggering interim analysis, to avoid undue harm to patients.

**Research perspectives and unresolved questions.** Treatment modulation deserves dedicated research, to elucidate the caveats and benefits of certain trial designs and improve efficiency in research while assuring rigorosity. This process should include the input of patient and advocacy group stakeholders.

There remain many unresolved issues in this field of interest which include optimal study endpoints, criteria for the selection of a non-inferiority margin, the threshold for minimal significant difference in health-related quality of life,<sup>67,68</sup> the pace of 'real world evidence', and research on innovative endpoints in the era of immunotherapy is also critical.

### *Clinical and decisional issues*

**Acknowledging uncertainty.** With a few notable exceptions, current strategies investigating modulation of treatment through de-intensification are based on identifying patients who are at low risk for an outcome, rather than biomarkers predictive of diminished sensitivity to a particular standard of care therapy. Consequently, in clinical settings where low risk is defined based on tumour size and nodal status, treatment modulation inevitably carries a small risk of missing out on potentially beneficial treatment. The magnitude of this risk depends on the absolute risk of an event with standard of care therapy and the inferiority margin associated with the proposed de-intensification strategy.

The personal implications of having a potentially slightly higher risk for an adverse clinical outcome due to treatment modulation are dependent on the clinical endpoint. For these reasons, it is incumbent on physicians to communicate uncertainties when treatment modulation is proposed, particularly in the absence of larger randomised clinical trial data and mature OS data. Risk-averse patients may sometimes reasonably forgo the option of de-intensification.

**Shared decision making for treatment modulation.** Patients may perceive the term 'treatment de-escalation', differently from clinicians and researchers. Indeed, some previous reports showed that the term 'de-escalation' may not be the most preferred term by some patients, who may feel that they will receive substandard treatments.<sup>53,54</sup>

Patients can hold high expectations on adjuvant therapy and may prefer chemotherapy, albeit providing small benefits, when the communication of the absolute benefit is imperfect or the benefits of alternative approaches are not comprehensively articulated.<sup>55-57</sup> Fear surrounding

recurrence has been reported as a major barrier to participation in clinical trials of treatment modulation.<sup>53</sup>

At the same time, the reduction of physical and financial toxicities and the description of a positive impact on daily life were attractive aspects of treatment with reduced intensity.<sup>53</sup>

Importantly, patients' views on de-escalated treatment regimens may vary according to individual values and priorities. Some patients will focus on reducing toxicity and burden; others may be more concerned about not compromising disease control and future outcomes. Consequently, shared decision making requires a thoughtful discussion about benefits, risks, the limits of certainty, and practical and financial burdens of de-intensified versus standard treatments.<sup>58,59</sup> Prior research has shown that women with breast cancer with a low risk of recurrence are more likely to decline chemotherapy if they are aware of the small impact of this approach, underlining the need for effective communication through education, rather than persuasion.<sup>60</sup>

Semantics matters and choosing the appropriate language to explain a treatment modulation approach may have an impact. The use of patient-centred rather than treatment-centred language has been preferred by patients, with positive reactions to words such as 'personalised', 'customised', 'tailored', and 'risk-modulated' therapies.<sup>53</sup> In this regard, this working group favour the terms 'modulated' or 'tailored' with the explanation that this approach will provide the best balance between optimal disease control with the least risk of harm and burden.

The use of unbiased decision aids using high-quality visual and numerical information could supplement physicians counselling about the patient's health condition, the options for care, associated benefits, harms, probabilities, and scientific uncertainties related to a treatment modulation approach.<sup>61</sup>

### Conclusion

The results of de-intensified treatment modulation studies to date have brought into focus the need for heightened caution when considering modulation paradigms, even in a disease that may appear to have favourable outcomes. These paradigms should be evaluated in phase II studies, and results should be awaited before proceeding to phase III studies. Implementation into clinical practice before high-level evidence is available risks patient harms and is not advised. Furthermore, trials of treatment modulation should only be considered in well-defined, low-risk groups and when there is a strong rationale for investigating a particular treatment strategy. Recently, the COVID-19 pandemic brought us a practical 'window of opportunity' to study and communicate more about tailored and individualised cancer treatment strategies.<sup>69</sup> In the era of value-based health care and a patient-centred goal-setting, treatment intensity-modulation is a research priority in oncology.

### ACKNOWLEDGEMENTS

This is a project initiated by the ESMO Precision Medicine Working Group (PMWG). We thank ESMO leadership for their support and expert inputs in this manuscript, including the ESMO CMO Dr George Pentheroudakis and the ESMO PMWG Chair Dr Joaquin Mateo.

### FUNDING

This work was supported by the European Society for Medical Oncology (no grant number). There was no external funding of the event or manuscript production.

### DISCLOSURE

HB reports non-remunerated activities as a principal investigator (PI) within grant for clinical research from the National Cancer Institute, and is a non-remunerated member of Board of Directors in the Alliance for Clinical Trials in Oncology; LC reports receipt of institutional research funding from AbbVie, Immunomedics, NanoString Technologies, Novartis, Seattle Genetics, Syndax, Veracyte, advisory role in Aptitude Health, AstraZeneca/Daiichi Sankyo, Exact Sciences, G1 Therapeutics, Genentech/Roche, GlaxoSmithKline, Novartis, and Sanofi; SD reports receipt of honoraria for participation in Advisory Board from AstraZeneca, Cellectis, Isis/Servier, Novartis, Orion, Pierre Fabre, Rapta, Sanofi, receipt of honoraria as invited speaker from AstraZeneca, Exact Sciences, Eli Lilly, Merck Sharp & Dohme (MSD), Pfizer, Seagen, institutional financial interest for research as a Steering Committee Member from AstraZeneca, Bristol Myers Squibb (BMS), Pfizer, Puma, Roche/Genentech, Sanofi, institutional research funding from GE, institutional financial interest for clinical research funding as a coordinating PI from Taiho, institutional financial interest as a local PI from Orion, and a non-remunerated member of Board of Directors of the Société Française de Sénologie et Pathologie Mammaire; NH reports receipt of honoraria for participation in Advisory Board from Aptitude Health, AstraZeneca, Daiichi Sankyo, Eli Lilly, MSD, Novartis, Pfizer, Roche, Sandoz-Hexal, Seattle Genetics, receipt of honoraria as invited speaker from Art Temp, AstraZeneca, Daiichi Sankyo, Eli Lilly, Medscape, MSD, Novartis, Onkowsen, Pierre Fabre, Roche, Seattle Genetics, receipt of honoraria from West German Study Group (WSG) (husband), ownership interest in West German Study Group, institutional research funding from AstraZeneca, BMS, Daiichi Sankyo, Eli Lilly, MSD, Novartis, Odonate, Palleos, Pierre Fabre, Roche, Seattle Genetics, TRIO, WSG, a member of German AGO Breast Guideline Committee, a member of European School of Oncology/College of the European School of Oncology Breast Cancer Educational Programmes, Editor-in-Chief of *Breast Care* journal; DH reports receipt of honoraria for consulting in Advisory Board from Agendia, BioVeca, Cellworks, Cepheid, Epic Sciences, Freenome, L-Nutra, OncoCyte, Turnstone Biologics, owns stock shares in Inbiomotion SL, receipt of royalties from an institutional patent on which he is named investigator from Menarini Silicon Biosystems, institutional sponsored laboratory research as a

coordinating PI from AstraZeneca, Pfizer, institutional sponsored laboratory and clinical research as a coordinating PI from Menarini Silicon Biosystems, non-remunerated activities as a past president and past member of Board of Directors of ASCO, non-remunerated Chair of Translational Research Advisory Committee in EORTC, and a non-remunerated member of the ESMO Working Group; KK reports receipt of honoraria for participation in Advisory Board from 4D Pharma, AstraZeneca, Cyclacel, Daiichi Sankyo, Eisai, Eli Lilly, Immunomedics, Merck, Novartis, OncoSec, Pfizer, Puma, Seattle Genetics, holding stocks/shares from Grail, spouse employment in Grail, no financial interest to institution as local PI from Acetylon, Amgen, Calithera, CytomX Therapeutics, Eli Lilly, Genentech, Immunomedics, Incyte, Novartis, Pfizer, Seattle Genetics, Zentalis, participation in Steering Committee of Ambrx, AstraZeneca, Genentech, Immunomedics, and support for attending meetings and/or travel from AstraZeneca, Eli Lilly, and Pfizer; LP reports receipt of honoraria for participation in Advisory Board and trial Steering Committee from AstraZeneca, Merck, honoraria for participation in Advisory Board from Novartis, Pfizer, Roche Genentech, no financial interest for clinical trial support to institution as local PI from AstraZeneca, Seagen, no financial interest for clinical trial support to institution as trial chair from Merck, Pfizer, research grant to institution from BMS, and no financial interest for advisory role in Breast Cancer Research Foundation, as a member of the scientific Advisory Board of Pfizer, for product samples, material transfer agreement; MR reports receipt of honoraria for participation in Advisory Board from BMS, Tolmar Pharmaceuticals, honoraria as invited speaker from WebMD, institutional financial interest for investigator-initiated clinical trials supported or drug supply from AstraZeneca, Roche, research grant to institution from Bayer, BMS, institutional financial interest for investigator-initiated clinical trials supported by Ipsen, Novartis, Pfizer, TerSera, and non-financial interest for advisory role from BMS; IS reports receipt of honoraria as invited speaker from Myriad Genetics and Nanostring Technologies; TS reports receipt of honoraria for participation in Advisory Board as a member of Patient Advisory Board from Pfizer and Roche; JS reports receipt of honoraria for participation in Advisory Board from Arvinas, AstraZeneca, Certara, Eisai, GlaxoSmithKline, honoraria as a Data and Safety Monitoring Board member from BMS, Novartis, Roche, and no financial interest to institution as coordinating PI from Merck, Olema, Radius Pharmaceuticals; GC reports receipt of honoraria for participation in Advisory Board from AstraZeneca, BMS, Daiichi Sankyo, Ellipsis, Exact Sciences, Lilly, Merck, Pfizer, Roche, Veracyte, receipt of honoraria as invited speaker from AstraZeneca, Daiichi Sankyo, Novartis, Pfizer, Roche, receipt of honoraria for writing engagement from Pfizer, institutional research grant for investigator-initiated clinical trial from Merck, institutional funding for conduct of phase I studies from Astellas, AstraZeneca, Blueprint Medicine, BMS, Daiichi Sankyo, Kymab, Novartis, Philogen, Roche, Sanofi, non-remunerated activities as an officer of the Italian National Health Council,

non-remunerated advisory role as a member of the Scientific Council of Europa Donna, non-remunerated advisory role in Fondazione Beretta, and non-remunerated member of Board of Directors of the Lega Italiana Lotta ai Tumori; FA reports receipt of institutional research grants from AstraZeneca, Daiichi Sankyo, Eli Lilly, Novartis, Pfizer, Roche, and being a founder of Pegacsy. DT has been a past consultant for World Health Organization, noncommunicable diseases, cancer management team. All other authors have declared no conflicts of interest.

## REFERENCES

- Shapiro CL. De-escalation yes, but not at the expense of efficacy: in defense of better treatment. *NPJ Breast Cancer*. 2019;5(1):1-3.
- Piccart MJ, Hilbers FS, Bliss JM, et al. Road map to safe and well-designed de-escalation trials of systemic adjuvant therapy for solid tumors. *J Clin Oncol*. 2020;38(34):4120-4129.
- Burstein HJ, Curigliano G, Thürlimann B, et al. Customizing local and systemic therapies for women with early breast cancer: the St. Gallen International Consensus Guidelines for treatment of early breast cancer 2021. *Ann Oncol*. 2021;32(10):1216-1235.
- Hellesnes R, Myklebust TÅ, Fosså SD, et al. Testicular cancer in the cisplatin era: causes of death and mortality rates in a population-based cohort. *J Clin Oncol*. 2021;39(32):3561-3573.
- Chevalier H, Vienot A, Lièvre A, et al. FOLFIRINOX de-escalation in advanced pancreatic cancer: a multicenter real-life study. *Oncologist*. 2020;25(11):e1701-e1710.
- Weiser R, Polychronopoulou E, Kuo Y-F, et al. De-escalation of endocrine therapy in early hormone receptor-positive breast cancer: when is local treatment enough? *Ann Surg*. 2021;274(4):654-663.
- Earl HM, Hiller L, Vallier A-L, et al. 6 versus 12 months of adjuvant trastuzumab for HER2-positive early breast cancer (PERSEPHONE): 4-year disease-free survival results of a randomised phase 3 non-inferiority trial. *Lancet*. 2019;393(10191):2599-2612.
- Kunst N, Wang S-Y, Hood A, et al. Cost-Effectiveness of neoadjuvant-adjuvant treatment strategies for women with ERBB2 (HER2)-positive breast cancer. *JAMA Netw Open*. 2020;3(11):e2027074.
- Henderson IC, Gelman RS, Harris JR, Canellos GP. Duration of therapy in adjuvant chemotherapy trials. *NCI Monogr*. 1986;1:95-98.
- Early Breast Cancer Trialists' Collaborative Group. Effects of radiotherapy and surgery in early breast cancer. An overview of the randomized trials. *N Engl J Med*. 1995;333(22):1444-1455.
- Mehanna H, Robinson M, Hartley A, et al. Radiotherapy plus cisplatin or cetuximab in low-risk human papillomavirus-positive oropharyngeal cancer (De-ESCALaTE HPV): an open-label randomised controlled phase 3 trial. *Lancet*. 2019;393(10166):51-60.
- Regan MM, Barry WT. Trial designs and results supporting treatment de-escalation and escalation. *Breast*. 2017;34(suppl 1):S10-S12.
- Sparano JA, Solin LJ. Defining the clinical utility of gene expression assays in breast cancer: the intersection of science and art in clinical decision making. *J Clin Oncol*. 2010;28(10):1625-1627.
- Dinan MA, Lyman GH, Schilsky RL, Hayes DF. Proposal for value-based, tiered reimbursement for tumor biomarker tests to promote innovation and evidence generation. *JCO Precis Oncol*. 2019;3:1-10.
- Hayes DF. Defining clinical utility of tumor biomarker tests: a clinician's viewpoint. *J Clin Oncol*. 2021;39(3):238-248.
- Kalinsky K, Thomas A, Cescon DW. On the road to precision: understanding the biology driving genomic assays. *J Clin Oncol*. 2021;39(2):100-102.
- Straus DJ, Jung SH, Pitcher B, et al. CALGB 50604: risk-adapted treatment of nonbulky early-stage Hodgkin lymphoma based on interim PET. *Blood*. 2018;132(10):1013-1021.
- Barrington SF, Phillips EH, Counsel N, et al. Positron emission tomography score has greater prognostic significance than pretreatment

- risk stratification in early-stage Hodgkin lymphoma in the UK RAPID study. *J Clin Oncol*. 2019;37(20):1732-1741.
19. 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(25):2419-2429.
  20. Yeoh AEJ, Ariffin H, Chai ELL, et al. Minimal residual disease-guided treatment deintensification for children with acute lymphoblastic leukemia: results from the Malaysia-Singapore acute lymphoblastic leukemia 2003 study. *J Clin Oncol*. 2012;30(19):2384-2392.
  21. Pieters R, de Groot-Kruseman H, Van der Velden V, et al. Successful therapy reduction and intensification for childhood acute lymphoblastic leukemia based on minimal residual disease monitoring: study ALL10 from the Dutch Childhood Oncology Group. *J Clin Oncol*. 2016;34(22):2591-2601.
  22. Vora A, Goulden N, Wade R, et al. Treatment reduction for children and young adults with low-risk acute lymphoblastic leukaemia defined by minimal residual disease (UKALL 2003): a randomised controlled trial. *Lancet Oncol*. 2013;14(3):199-209.
  23. Sparano JA, Gray RJ, Makower DF, et al. Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. *N Engl J Med*. 2018;379(2):111-121.
  24. Piccart M, van 't Veer LJ, Poncet C, et al. 70-gene signature as an aid for treatment decisions in early breast cancer: updated results of the phase 3 randomised MINDACT trial with an exploratory analysis by age. *Lancet Oncol*. 2021;22(4):476-488.
  25. Tolaney SM, Guo H, Pernas S, et al. Seven-year follow-up analysis of adjuvant paclitaxel and trastuzumab trial for node-negative, human epidermal growth factor receptor 2-positive breast cancer. *J Clin Oncol*. 2019;37(22):1868-1875.
  26. André T, Meyerhardt J, Iveson T, et al. Effect of duration of adjuvant chemotherapy for patients with stage III colon cancer (IDEA collaboration): final results from a prospective, pooled analysis of six randomised, phase 3 trials. *Lancet Oncol*. 2020;21(12):1620-1629.
  27. Ferris RL, Flamand Y, Weinstein GS, et al. Phase II randomized trial of transoral surgery and low-dose intensity modulated radiation therapy in resectable p16+ locally advanced oropharynx cancer: an ECOG-ACRIN Cancer Research Group Trial (E3311). *J Clin Oncol*. 2022;40(2):138-149.
  28. Kondagunta GV, Bacik J, Bajorin D, et al. Etoposide and cisplatin chemotherapy for metastatic good-risk germ cell tumors. *J Clin Oncol*. 2005;23(36):9290-9294.
  29. Sawaki M, Taira N, Uemura Y, et al. Randomized controlled trial of trastuzumab with or without chemotherapy for HER2-positive early breast cancer in older patients. *J Clin Oncol*. 2020;38(32):3743-3752.
  30. Bajetta E, Floriani I, Di Bartolomeo M, et al. Randomized trial on adjuvant treatment with FOLFIRI followed by docetaxel and cisplatin versus 5-fluorouracil and folinic acid for radically resected gastric cancer. *Ann Oncol*. 2014;25(7):1373-1378.
  31. Buus R, Sestak I, Kronenwett R, et al. Molecular drivers of Oncotype DX, Prosigna, EndoPredict, and the Breast Cancer Index: a TransATAC study. *J Clin Oncol*. 2021;39(2):126-135.
  32. Sestak I, Buus R, Cuzick J, et al. Comparison of the performance of 6 prognostic signatures for estrogen receptor-positive breast cancer: a secondary analysis of a randomized clinical trial. *JAMA Oncol*. 2018;4(4):545-553.
  33. Gyawali B, de Vries EGE, Dafni U, et al. Biases in study design, implementation, and data analysis that distort the appraisal of clinical benefit and ESMO-Magnitude of Clinical Benefit Scale (ESMO-MCBS) scoring. *ESMO Open*. 2021;6(3):100117.
  34. Mauri L, D'Agostino RB. Challenges in the design and interpretation of noninferiority trials. *N Engl J Med*. 2017;377(14):1357-1367.
  35. Aberegg SK, Hersh AM, Samore MH. Empirical consequences of current recommendations for the design and interpretation of noninferiority trials. *J Gen Intern Med*. 2018;33(1):88-96.
  36. Rehal S, Morris TP, Fielding K, et al. Non-inferiority trials: are they inferior? A systematic review of reporting in major medical journals. *BMJ Open*. 2016;6(10):e012594.
  37. Fromme EK, Eilers KM, Mori M, et al. How accurate is clinician reporting of chemotherapy adverse effects? A comparison with patient-reported symptoms from the Quality-of-Life Questionnaire C30. *J Clin Oncol*. 2004;22(17):3485-3490.
  38. Di Maio M, Gallo C, Leighl NB, et al. Symptomatic toxicities experienced during anticancer treatment: agreement between patient and physician reporting in three randomized trials. *J Clin Oncol*. 2015;33(8):910-915.
  39. Filipits M, Rudas M, Jakesz R, et al. A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. *Clin Cancer Res*. 2011;17(18):6012-6020.
  40. Tolaney SM, Garrett-Mayer E, White J, et al. Updated standardized definitions for efficacy end points (STEEP) in adjuvant breast cancer clinical trials: STEEP version 2.0. *J Clin Oncol*. 2021;39(24):2720-2731.
  41. Gourgou-Bourgade S, Cameron D, Poortmans P, et al. Guidelines for time-to-event end point definitions in breast cancer trials: results of the DATECAN initiative (Definition for the Assessment of Time-to-event Endpoints in CANcer trials). *Ann Oncol*. 2015;26(5):873-879.
  42. Hudis CA, Barlow WE, Costantino JP, et al. Proposal for standardized definitions for efficacy end points in adjuvant breast cancer trials: the STEEP system. *J Clin Oncol*. 2007;25(15):2127-2132.
  43. Gyawali B, Hey SP, Kesselheim AS. Evaluating the evidence behind the surrogate measures included in the FDA's table of surrogate endpoints as supporting approval of cancer drugs. *EclinicalMedicine*. 2020;21:100332.
  44. Saad ED, Squifflet P, Burzykowski T, et al. Disease-free survival as a surrogate for overall survival in patients with HER2-positive, early breast cancer in trials of adjuvant trastuzumab for up to 1 year: a systematic review and meta-analysis. *Lancet Oncol*. 2019;20(3):361-370.
  45. Carey LA, Winer EP. Defining success in neoadjuvant breast cancer trials. *Lancet*. 2014;384(9938):115-116.
  46. Korn EL, Freidlin B, Abrams JS. Overall survival as the outcome for randomized clinical trials with effective subsequent therapies. *J Clin Oncol*. 2011;29(17):2439-2442.
  47. Harbeck N, Schneeweiss A, Thuss-Patience P, et al. Neoadjuvant and adjuvant end-points in health technology assessment in oncology. *Eur J Cancer*. 2021;147:40-50.
  48. Gelber RD, Goldhirsch A. A new endpoint for the assessment of adjuvant therapy in postmenopausal women with operable breast cancer. *J Clin Oncol*. 1986;4(12):1772-1779.
  49. Bertram MY, Lauer JA, De Joncheere K, et al. Cost-effectiveness thresholds: pros and cons. *Bull World Health Organ*. 2016;94(12):925-930.
  50. Azimi NA, Welch HG. The effectiveness of cost-effectiveness analysis in containing costs. *J Gen Intern Med*. 1998;13(10):664-669.
  51. Giorgi Rossi P, Lebeau A, Canelo-Aybar C, et al. Recommendations from the European Commission Initiative on Breast Cancer for multigene testing to guide the use of adjuvant chemotherapy in patients with early breast cancer, hormone receptor positive, HER-2 negative. *Br J Cancer*. 2021;124(9):1503-1512.
  52. Trapani D, Curigliano G, Eniu A. Breast cancer: reimbursement policies and adoption of new therapeutic agents by National Health Systems. *Breast Care (Basel)*. 2019;14(6):373-381.
  53. Rocque GB, Williams CP, Andrews C, et al. Patient perspectives on chemotherapy de-escalation in breast cancer. *Cancer Med*. 2021;10(10):3288-3298.
  54. Wang T, Mott N, Miller J, et al. Patient perspectives on treatment options for older women with hormone receptor-positive breast cancer: a qualitative study. *JAMA Netw Open*. 2020;3(9):e2017129.
  55. Blinman P, Duric V, Nowak AK, et al. Adjuvant chemotherapy for early colon cancer: what survival benefits make it worthwhile? *Eur J Cancer*. 2010;46(10):1800-1807.
  56. Blinman P, King M, Norman R, et al. Preferences for cancer treatments: an overview of methods and applications in oncology. *Ann Oncol*. 2012;23(5):1104-1110.
  57. Blinman P, Martin A, Jefford M, et al. Patients' preferences for 3 months vs 6 months of adjuvant chemotherapy for colon cancer. *JNCI Cancer Spectr*. 2020;5(1):pkaa107.

58. Nicolai J, Buchholz A, Seefried N, et al. When do cancer patients regret their treatment decision? A path analysis of the influence of clinicians' communication styles and the match of decision-making styles on decision regret. *Patient Educ Couns*. 2016;99(5):739-746.
59. Budden LM, Hayes BA, Buettner PG. Women's decision satisfaction and psychological distress following early breast cancer treatment: a treatment decision support role for nurses. *Int J Nurs Pract*. 2014;20(1):8-16.
60. Belkora JK, Volz S, Teng AE, et al. Impact of decision aids in a sustained implementation at a breast care center. *Patient Educ Couns*. 2012;86(2):195-204.
61. Vromans R, Tenfelde K, Pauws S, et al. Assessing the quality and communicative aspects of patient decision aids for early-stage breast cancer treatment: a systematic review. *Breast Cancer Res Treat*. 2019;178(1):1-15.
62. Cherny NI, Dafni U, Bogaerts J, et al. ESMO-Magnitude of Clinical Benefit Scale version 1.1. *Ann Oncol*. 2017;28(10):2340-2366.
63. Mateo J, Chakravarty D, Dienstmann R, et al. A classification to rank genomic alterations as targets for cancer precision medicine: the ESMO Scale for Clinical Actionability of molecular Targets (ESCAT). *Ann Oncol*. 2018;29(9):1895-1902.
64. Noordhoek I, Treuner K, Putter H, et al. Breast Cancer Index predicts extended endocrine benefit to individualize selection of patients with HR+ early-stage breast cancer for 10 years of endocrine therapy. *Clin Cancer Res*. 2021;27(1):311-319.
65. Freidlin B, Korn EL, George SL, Gray R. Randomized clinical trial design for assessing noninferiority when superiority is expected. *J Clin Oncol*. 2007;25(31):5019-5023.
66. Pletcher MJ, Pignone M. Evaluating the clinical utility of a biomarker: a review of methods for estimating health impact. *Circulation*. 2011;123(10):1116-1124.
67. Osoba D, Rodrigues G, Myles J, et al. Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol*. 1998;16(1):139-144.
68. Regan MM, Werner L, Rao S, et al. Treatment-free survival: a novel outcome measure of the effects of immune checkpoint inhibition—a pooled analysis of patients with advanced melanoma. *J Clin Oncol*. 2019;37(35):3350-3358.
69. Hanna TP, Evans GA, Booth CM. Cancer, COVID-19 and the precautionary principle: prioritizing treatment during a global pandemic. *Nat Rev Clin Oncol*. 2020;17(5):268-270.