**The road map to safe and well-designed de-escalation trials of systemic adjuvant therapy for solid tumors**

Martine J. Piccart1, Florentine S. Hilbers2, Judith M. Bliss3, Carmela Caballero2, Elizabeth S. Frank4, Patrick Renault5, Rachida Naït Kaoudjt6, Eva Schumacher7, Patricia A. Spears 8, Meredith M. Regan9, Richard D. Gelber10, Nancy E. Davidson11,12, Larry Norton13, Eric P. Winer4, On Behalf of the BIG-NABCG Collaboration

1 Institut Jules Bordet, Department of Medicine, Brussels, Belgium

2Breast International Group, Brussels, Belgium

3The Institute of Cancer Research, Clinical Trials & Statistics Unit (ICR-CTSU), London, UK

4 Dana Farber Cancer Institute, Department of Medical Oncology, Boston, USA

5 Patient advocate, Belgium

6 Patient advocate, France

7 Mamma Mia! Das Brustkrebsmagazin, Kronberg, Germany

8 UNC Lineberger Comprehensive Cancer Center, Chapel Hill,USA

9 International Breast Cancer Study Group Statistical Center, Department of Data Sciences, Dana-Farber Cancer Institute, Harvard Medical School, Boston, USA

10 International Breast Cancer Study Group Statistical Center, Department of Data Sciences, Dana-Farber Cancer Institute, Harvard Medical School, Harvard T.H. Chan School of Public Health and Frontier Science and Technology Research Foundation, Boston, USA

11 Fred Hutchinson Cancer Research Center, Seattle, USA

12 University of Washington, Seattle, USA

13 Memorial Sloan Kettering Cancer Center, Department of Medicine, New York, NY

**Funding**

This work was funded by the Breast Cancer Research Foundation (BCRF).

**Running head:**

Design of de-escalation trials of systemic adjuvant therapy

**Corresponding author:**

Prof. Martine Piccart, Institut Jules Bordet, Department of Medicine,

Boulevard de Waterloo 125, B-1000 Brussels, Belgium.

Phone: 32-2-541-3206;

E-mail: [martine.piccart@bordet.be](mailto:martine.piccart@bordet.be)

**Context Summary**

*Key objective:* While the scientific community and society at large welcome de-escalation of systemic adjuvant therapy in selected cancer patients in view of quality of life and health economics benefits, more attention should be paid to rigorous de-escalation trial design.

*Knowledge generated:* A review of de-escalation trials in the last few years highlighted a number of weaknesses which can effectively be addressed in the future. The BIG-NABCG Collaboration developed a roadmap for improving the design and implementation of de-escalation trials guided by patients’ insights, with recommendations on how to minimize treatment non-adherence, heterogeneity of results and risks of undertreatment, and support the selection of RFI, RFS and DMFS as meaningful endpoints.

*Relevance:* This roadmap can help investigators conduct de-escalation trials with robust, practice-changing, and truly patient-centered results.

**Abstract**

An important challenge in the field of cancer is finding the balance between delivering effective treatments, while avoiding side effects and financial toxicity caused by innovative yet expensive drugs. To address this, several treatment de-escalation trials have been conducted in the past but only a few of these have provided clear answers. A few of them suffered from poor accrual or had study design flaws that led to conflicting results. Members of the Breast International Group (BIG) and North American Breast Cancer Group (NABCG) believe that the way forward would be to understand the lessons from these trials and listen more carefully to what truly matters to our patients. We reviewed several adjuvant trials from different cancer types and developed a roadmap for improving the design and implementation of future de-escalation trials. This incorporates insights from patients obtained through focused group discussions across the BIG-NABCG networks. Considerations for the development of de-escalation trials for systemic adjuvant treatment including non-inferiority trial design, choice of endpoints and prioritization of patient’s perspectives are presented in this consensus paper.

**Introduction**

Progress against micro-metastases in diseases like breast or colorectal cancer reflects the introduction of effective cytotoxic agents and targeted drugs, given in sequence or in combination. Large, prospective, randomized clinical trials have been instrumental to this progress, which translates to 5-year survival rates in excess of 90% for many patients, such as those with HER2-positive breast cancer. Earlier diagnosis and improved loco-regional treatment modalities – e.g. surgery with or without radiotherapy – have also been pivotal in the “race for curing cancer”.

While oncologists are proud of the high rates of cure as a result of systemic therapy trials, they have become aware of its limitations: overtreatment of patients already cured by loco-regional therapy, short-term toxicities, emergence of chronic, long-term and sometimes life-threatening side effects and financial toxicity. Following the example of pediatric oncologists, who have been pioneers in careful treatment de-escalation for highly curable childhood malignancies1,2, oncologists treating adult tumors, in particular breast cancer, are now exploring safer or shorter drug regimens with the hope of little or no loss in rates of long term survival with gains in quality of life.

The rapid growth in the number of drug de-escalation trials in the early disease setting has not always been accompanied by in depth consultation with patients, or statisticians. As a consequence, trial results have often been left with a worrisome heterogeneity in designs and hypotheses. Another concern is the under-funding of some of these trials – given the lack of interest from the pharmaceutical industry – leading to inadequate sample sizes and, consequently, insufficient statistical power for a “robust” practice-changing result.

For all these reasons, academic investigators and experts in clinical trial design from BIG and NABCG which have been collaborating for more than a decade came together to develop a consensus to 1) highlight lessons from important trials on de-escalation 2) build a road map for future drug de-escalation studies in the curative setting incorporating insights from our patients and expert statisticians.

**Lessons from past drug de-escalation trials in the adjuvant setting**

**Shortening the duration of adjuvant systemic therapy**

Two large efforts in the field of breast and colon cancer illustrate the lessons from studies that aimed to de-escalate the duration of adjuvant systemic therapy (summarized in table 1).

In colon cancer, oxaliplatin combined with 5-fluorouracil or capecitabine is a standard adjuvant chemotherapy in patients with stage III disease3–5. However, this drug can induce cumulative dose-dependent long-term sensory neurotoxicity6. This led to investigations into shorter duration of oxaliplatin to reduce the risk of neurotoxicity without compromising DFS. The IDEA collaboration was established to analyze data from six randomized trials that evaluated whether a three-month course of oxaliplatin-based adjuvant therapy (FOLFOX4/modified FOLFOX6 or CAPOX) is non-inferior to the standard six-month treatment for patients with stage III colon cancer, with a primary endpoint of DFS at three years7. A modified intention-to-treat method was used for the primary analysis, which included all the patients who had undergone randomization and had received at least one dose of a trial drug.

Non-inferiority of 3 months of treatment versus 6 months was not confirmed in the overall study population; non-inferiority of the shorter regimen was seen for CAPOX but not for FOLFOX8. As this difference between the treatments was unexpected, no randomization between them took place. Moreover, an exploratory analysis showed that for the combined regimen, 3 months of treatment was non-inferior in lower risk stage III patients (T1-3, N1), but not in higher risk (T4 and/or N2) patients.

Lessons learned from IDEA include: 1) a later time for randomization would have been better to test the efficacy of the two regimens, as the selected early point for randomization increased the risk of suboptimal treatment adherence in the arm with longer treatment duration; 2) inclusion of patients with higher disease burden should be avoided as it can jeopardize a “safe” de-escalation approach. 3) Pooling data from similar albeit different trials contributes to heterogeneity of results and complicates their interpretation.

In HER2-positive breast cancer, five randomized clinical trials have been conducted to investigate non-inferiority of adjuvant trastuzumab administration durations of less than the “standard” one-year treatment. A meta-analysis showed important differences in the selection of “non-inferiority margin”.9,10 It suggests inferior outcomes for shorter duration trastuzumab for DFS but a significantly higher risk of cardiac events with 12 months of adjuvant. The only “positive” trial, PERSEPHONE (NCT00712140), concluded that 6 months of trastuzumab is equivalent to 12 months, but raises doubts as to the applicability of its results10–12. PERSEPHONE and PHARE13,14 (NCT00381901), scheduled randomization between 6 and 12 months of adjuvant trastuzumab, near the point of treatment divergence. Unfortunately, during their very long recruitment period, clinical practice changed such that virtually all patients today receive concomitant trastuzumab and chemotherapy. Only 47% of the PERSEPHONE patients were treated in this manner, complicating the interpretation of its results. Moreover, only 15% of patients were treated in the neoadjuvant setting, a common approach today. Optimal randomized de-escalation trials need international collaboration to be robust and generalizable. Careful planning with attention to statistical considerations, specifically the non-inferiority margin, are critical to obtain practice-changing results.

**Trial designs to de-escalate treatment**

Reducing the burden of chemotherapy can be accomplished by removing the most toxic drug(s), replacing a cytotoxic drug by a targeted drug or omitting chemotherapy altogether with the aid of specific biomarkers to allow tailoring of treatment.

In Hodgkin disease, early FDG-PET scan was tested to identify patients sensitive to the eBEACOPP regimen. Identification of patients with negative PET scans at cycle 2 allowed reduction from six to four cycles of eBEACOPP without loss of tumor control15.

Positive prognostic role of Human Papilloma Virus status prompted randomized de-escalation trials for chemoradiation in head and neck cancer. However, despite positive phase II trials, substituting cisplatin for cetuximab in combination with radiotherapy led to inferior results in two randomized phase III trials16,17.

In HER2 positive breast cancer, the single arm APT trial (NCT00542451) prospectively tested omission of doxorubicin and cyclophosphamide in a carefully selected patient population with a low tumor burden (mostly T1N0 tumors)18. It reported a 3-year iDFS of 98.7% (CI: 97.6-99.8%) which was well above its predefined non-inferiority boundary of 90.8%. At 7 years median follow-up results remained excellent with only four distant metastases reported19. Most national guidelines have endorsed APT, which shortens duration of chemotherapy and alleviates concerns for irreversible treatment-induced cardiotoxicity and substantially reduces acute toxicity among patients with an excellent prognosis.

**Tailoring chemotherapy: Evidence from biomarker-based trials**

The decision to forego chemotherapy is challenging. The largest experience is in luminal breast cancer for which effective targeted endocrine therapy exists. With the development of multigene expression signatures, such as Oncotype-DX or Mammaprint®, it has become easier to identify patients with excellent outcomes when treated with adjuvant endocrine therapy only (Oncotype-DX) or no adjuvant chemotherapy (Mammaprint®)20,21. The clinical utility of these signatures in selected patients with luminal breast cancer who may safely forego chemotherapy has been tested in two large, prospective trials, TAILORX (NCT00310180) and MINDACT (NCT00433589), which differ in terms of patient populations, design and length of follow-up (table 2). Both trials concluded that chemotherapy could be avoided in selected patients without an unacceptable loss in efficacy. TAILORX did not show a benefit from adding chemotherapy to endocrine therapy in women with node-negative disease and an intermediate recurrence score, according to a fully powered randomized design, while MINDACT demonstrated the per protocol predefined excellent outcome at 5 years in women treated with hormonal therapy only, whose tumors were classified as high clinical risk but found to have a low genomic risk signature (the key “target” group of MINDACT)22,23. Although treatment allocation for women with discordant risks (as determined clinically or genomically) was the result of randomization in MINDACT, there was not enough statistical power for a reliable comparison of these “randomized groups”. As a result, a small distant metastasis free survival (DMFS) advantage from the addition of chemotherapy cannot be formally excluded in the target subgroup of MINDACT but if it exists it does not exceed 2% at 5 years in absolute terms, a benefit that women should be free to take or to reject based on their own perception of benefit versus risk.

There are important differences in the use of the two genomic signatures in the trials. In TAILORX, Oncotype-DX was used at time of primary surgery, independently from “clinical risk” in N0, hormone receptor-positive disease, while in MINDACT a large proportion of clinical high-risk patients were recruited, with 21% with N1 disease overall and 47.6 % in the “high clinical-low genomic risk” subgroup. Controversy remains as to whether Oncotype-Dx and/or Mammaprint® can be safely used in patients with 1 to 3 positive nodes. Registry data would support the use of Oncotype-Dx in such patients while MINDACT can claim the prospective inclusion of such patients, although only 226 patients women with 2 or 3 positive nodes were part of the target group.22,24 This is why results from RxPONDER (NCT01272037), a US trial exploring chemotherapy de-escalation with the use of Oncotype-DX in women with 1 to 3 positive nodes, are eagerly awaited.

In spite of the huge enthusiasm of patients and doctors for conduct of and application of results from these well designed and expensive de-escalation trials, health technology assessment specialists from outside the US have remained “skeptical”. They wonder if the immediate cost linked to the use of these genomic tests will be offset by a reduction in indirect costs such as those induced by short- and long-term chemotherapy side effects. They argue that the “good results” will diminish in time given the persistent risk of late recurrence in this patient group. Recently, more mature results of MINDACT, at 8.7 years median follow-up, have shown an intact performance of the 70 gene assay to guide adjuvant chemotherapy decision in women older than 50 years, with no hint of a chemotherapy benefit in the group with a high clinical risk and a low genomic risk.25 In contrast, for women younger than 50 years of age in this target group, there is an estimated 5% absolute gain in DMFS with adjuvant chemotherapy, which will need to be communicated as part of patient-doctor shared decision making. A very similar observation was made in the TailorX trial.23 In countries where genomic assays are not available, clinical parameters can be used to de-escalate chemotherapy, but may be less reliable.

**The way forward**

**Involving patients in the design and conduct of de-escalation trials**

In the context of the BIG-NABCG meeting 2019, two workshops were held in Belgium and the U.S. to discuss perspectives on various de-escalation strategies and to evaluate patient priorities for treatment outcome and quality of life. The participants were a small, selected group of breast cancer patients, with a variety in age, gender, nationality, work background, family situation and disease stage. They were oriented to principles of trial design and outcomes of various de-escalation studies as well as basic statistical concepts. Subsequently, they were asked whether they would have been willing to participate in these trials at the time of their diagnosis and whether their decision would have changed following the focused group discussion.

While breast cancer patients are cognizant of the benefits of chemotherapy and radiotherapy, they may be less aware of or less worried about the long-term side effects of such treatments and why appropriate tailoring of treatment is so important. Many factors could affect a patient’s view on treatment and de-escalation. A physician’s aversion to uncertainty will also affect patient’s acceptance of risk. Patient’s expectations are further tempered by their feelings of vulnerability at the time of diagnosis, personal values, influences from media, social networks, family members, experiences of other patients and members of their support group.

Risk tolerance discussed in the context of a shared decision-making process

De-escalation can be a difficult concept for patients to understand, particularly at the time of diagnosis when patients feel most vulnerable. Treatment decisions are difficult and are particularly hard under the psychological distress of a cancer diagnosis. Most patients think of chemotherapy therapy as lifesaving and believe that it will eliminate their cancer (therapeutic optimism). There is a high probability that patients will misunderstand the intention of a de-escalation clinical trial. Therefore, it is the responsibility of doctors to convey clearly to patients the reason why less treatment is being tested and how it can result in excellent outcomes in terms of efficacy, safety and quality of life. It must also be made clear whether the de-escalation strategy might benefit patients by reducing side effects, financial toxicity, or both.

Not surprisingly, patients are more willing to accept de-escalation for treatments that have worse side effects. In the case of the PRIMETIME trial26 where patients with low-grade, node-negative tumors are identified as low-risk by a histochemical (IHC4) biomarker score to avoid radiotherapy, more patients responded that they would consider de-escalation despite the potential increased risk of a rare but treatable local recurrence. The same was true when asked to consider the MINDACT trial where patients identified as low-risk by genomic markers would be randomized to forgo chemotherapy.

There was less tolerance for risk when considering reducing trastuzumab treatment, such as in PHARE and PERSEPHONE, since trastuzumab has fewer uncomfortable side effects. This was true even after being educated about the fact that trastuzumab increases risk of cardiovascular events.

Risk tolerance was higher for older patients who were less willing to exchange quality of life for a limited increase in survival. Maintaining physical function and independence is even more important to older patients. This is in contrast to younger patients who were willing to exchange a lower quality of life for any increase in survival since they are usually looking toward a long future of work and family obligations. Importantly, at each workshop, patients were reticent to consider greater than 3% decrease in DMFS when considering a treatment de-escalation.

For de-escalation trials it is even more important that physicians take the time to prepare patients to make informed treatment decisions. Patients expressed that they rely on the expertise of the doctor to lead them to the “right choice.” Patients with metastatic or high-risk cancer tend to be more hesitant to accept the risk of recurrence in a de-escalation trial. It is paramount that risks and potential therapeutic benefits are fully explained to them before they make the commitment to participate. If they perceive that they may regret their decision later, they would rather not take the risk. The patients emphasized that it is important that physicians are honest and that “I don’t know,” is a valid medical opinion.

To determine the feasibility of a de-escalation trial, it might be useful to conduct a patient survey to better understand patient’s acceptance of the risks and benefits of the proposed trial. Patient Advocates are also willing to be involved in the review of trial protocols, as well as patient information documents to ensure that other patients will thoroughly understand it.

Facilitating communication and patient education about de-escalation trials

The patient workshops highlighted opportunities for improved patient-doctor communication. Often, doctors have a limited amount of time to talk with patients to hear their fears, personal values, questions or offer an easily digestible explanation of what is entailed in a de-escalation trial. Deciding to participate in a clinical trial is a decision that is not only based on the science but also on values and personal experiences of the patient. As these are complex decisions, there is a need for clear communication and understanding of the individual patient’s values and situation. Tools such as pamphlets and videos that can be shared with friends and family go a long way to make up for the brevity of a typical consultation. Visual aids can be especially helpful for the communication regarding complex aspects of a trial, such as expected absolute risk and the maximum acceptable loss in efficacy (non-inferiority threshold). Having access to other healthcare professionals (such as nurses or hospital volunteers) who can address their questions would be extremely beneficial and helpful in making a shared and informed treatment decision.

**Improving the statistical design of drug de-escalation trials**

Careful statistical design is critical for treatment de-escalation trials, as the scientific community often only gets one opportunity to test a de-escalation hypothesis in a clinical trial; it is often difficult to fund de-escalation trials as they require a large sample size, and they might suffer from slow accrual. Moreover, as the underlying hypothesis goes against the body of evidence supporting the current standard of care, a single trial not being able to show non-inferiority is typically sufficient to lose clinical equipoise, and thereby undermine the ethical foundation for future trials assessing the same or a similar hypothesis. Additionally, while a poorly designed superiority trial may fail to pass the threshold of statistical significance and have no influence on clinical practice, a poorly designed non-inferiority trial, has the potential to claim non-inferiority where there is none, and so adversely affect guidelines for standard of care.

Key aspects in the statistical design of non-inferiority trials

The goal of a non-inferiority trial is to show that an experimental, de-escalated treatment is no worse than standard of care. However, as determining if two treatment arms are exactly equally effective would require an infinite sample size, in practice a predefined threshold is set based on the maximum absolute or relative loss in effectiveness one is willing to accept. This threshold cannot be crossed by the confidence interval surrounding the estimate for the de-escalated treatment arm. The non-inferiority threshold is critical for the design of a de-escalation trial and determines whether the results of a trial will be practice changing. The choice of threshold might be affected by factors such as the absolute risk of disease-related events and the severity of the toxicities associated with the treatment that is being de-escalated. Moreover, input from patients on how they weigh the risk of disease recurrence against the overall burden of the de-escalated treatment is critical.

Typically, the non-inferiority threshold is either defined as an absolute or as a relative loss in effectiveness compared to the standard of care. However, when assumptions about the effectiveness of the standard of care differ substantially from the observed effectiveness in the study, the relation between absolute and relative difference in effectiveness will shift. This can lead to results that are difficult to interpret. For example, in the case of a much better than expected event-free survival, the 95% CI interval of the de-escalated treatment arm might not cross the non-inferiority threshold based on absolute risk, but the corresponding relative risk might be higher than expected and seem difficult to accept.

Although non-adherence to treatment is a problem in most trials with systemic treatments, it’s effects can have particularly large consequences in non-inferiority trials. As non-adherence to the assigned treatment will make the standard of care arm more similar to the de-escalation arm, the difference in effect between the two arms might be underestimated. Therefore, an “intention to treat” analysis in a de-escalation trial with substantial non-adherence will be biased towards concluding non-inferiority. A “per protocol analysis” reflects a less conservative alternative.27 In de-escalation studies assessing treatment duration, it might be beneficial to randomize patients when they are already on treatment and have almost completed the shorter of the two durations being compared (i.e., the point at which the two treatments diverge). By that time, it is known how well the patient is tolerating the treatment and whether the patient is likely to be able to complete the longer regimen. Moreover, as the initial anxiety at diagnosis might have decreased, the patient may be more willing to consider a treatment de-escalation trial. Finally, adherence to treatment must be part of the IDMC monitoring of the trial.

The single-arm cohort design

Non-inferiority designs typically require narrow confidence intervals and therefore large sample sizes, often leading to prohibitively high costs. In some settings, the use of a single arm cohort design, where the outcomes of a group of patients treated with the de-escalated treatment is compared with an absolute threshold of disease outcomes, might be a way to conduct a trial with a smaller sample size. However, inappropriate use of this design can easily result in the incorrect conclusion about the non-inferiority of the de-escalated treatment. A key requirement for a single-arm cohort design is the availability of high-quality data from one or more a well-matched cohort(s) treated with the current standard of care, which can be used to determine the expected event rate. However, even if these data are available, a single arm cohort design is only appropriate when the expected event rate is relatively low (e.g. <10% events at the main analysis timepoint). For higher event rates, a small difference in the risk profile of study population compared to the historical control will result in a relatively large difference in absolute risk of events, making the comparison invalid. Importantly, even though the inclusion criteria of a single arm cohort study might match a historical control population, the risk profile of patients actually enrolled in the trial might be skewed, typically towards lower risk patients. This might be particularly true for studies that are relatively small or have broad inclusion criteria.

Most appropriate endpoints of non-inferiority trials

Overall survival (OS) is still the gold standard for outcome measures in clinical trials and the most important outcome for patients. However, assessing OS requires large trials and long-term follow-up. This might mean that a drug has become obsolete by the time the results of a clinical study are available. Moreover, the effect of a treatment on OS is confounded by the effect of subsequent treatment lines given in the case of disease recurrence. Therefore, composite endpoints, for example invasive disease-free survival (iDFS), which takes into account multiple clinical endpoints such as disease recurrence, new primary cancers and death of any cause, are commonly used as the primary endpoint in superiority trials28,29. However, their use in the context of non-inferiority trials is disputed.

Composite endpoints often consider events that cannot be prevented by the treatment that is being de-escalated, such as non-breast cancer-related deaths and new primary tumors in organs other than the breast. These events are expected to occur at an approximately equal rate in both study arms and will dilute the efficacy of treatment de-escalation on preventable events. Therefore, use of endpoints related to disease recurrence events such as recurrence-free interval (RFI) or recurrence free-survival (RFS), are more appropriate in treatment de-escalation trials. Of note, these do not include treatment related deaths, which might provide important insights in the difference in toxicity between the de-escalated treatment and the standard of care. As these events are relatively rare, they should be carefully annotated and analyzed separately. Finally, DMFS should be considered as a primary endpoint as there is no salvage therapy yet for such events that patients fear the most.

Apart from efficacy endpoints, patient reported outcomes (PROs) and the evaluation of short and long-term toxicities are key in demonstrating the impact of de-escalation treatment. However, it is critical for PROs to be collected in a scientifically rigorous manner for them to truly inform clinical practice30–32.

Using an intermediate endpoint with established individual patient level surrogacy

An elegant way of enhancing the probability of “success” of a de-escalation trial is to use a surrogate endpoint that has robustly shown to have individual-level surrogacy, that is to be significantly associated with an excellent long term DFS.

One of the best examples of this strategy is the use of a drop in proliferation marker Ki67 after 2 weeks of exposure to endocrine therapy given prior to surgery in early luminal breast cancer.

In the POETIC trial (NCT02338310), this dynamic marker, centrally measured before and after a 2 week preoperative exposure to an aromatase inhibitor (AI), was able to divide the patient population into 3 prognostic subgroups: a very low risk group (low Ki67 before and after 2 weeks of AI) with a 4.5% risk of recurrence by 5 years, a median risk group (high Ki67 before and low Ki67 after the 2 weeks of AI) with an 8,9% risk of recurrence by 5 years and a high risk group (high Ki67 before and high Ki67 after the treatment window) with a 19.6% risk of recurrence by 5 years. All patients continued AI therapy after surgery; a small proportion also received post-operative chemotherapy at the discretion of the treating oncologist. This trial represents the largest reported effort so far using a dynamic biomarker to identify low risk patients who should no longer be entered in drug escalation trials.

In HER2-positive breast cancer, the excellent outcome in the standard arm of APHINITY (NCT01358877), which shows an iDFS rate in excess of 90% at 45 months median follow-up, provides an encouragement for developing careful de-escalation strategies beyond the APT regimen. Two such trials are due to start, on both sides of the Atlantic Ocean, COMPASS and DECRESCENDO. Both will use the endpoint of pathological complete response (pCR), which has demonstrated individual-level surrogacy, to remove anthracyclines from the treatment scheme in case the tumor has pCR after 3 months of exposure to a taxane combined with dual HER2-blockade using trastuzumab and pertuzumab. Although the two trials, differ in terms of patient populations and salvage regimens for non-pCR patients (table 3), both define “non-inferiority” as no less than 94% RFS at 3 years, an ambitious threshold that is needed if these trials are going to be practice changing.

**Conclusions**

While the scientific community and society at large welcome de-escalation of systemic adjuvant therapy in selected cancer patients in view of quality of life and health economics benefits, more attention should be paid to rigorous de-escalation trial design if robust, practice-changing, and truly patient-centered results are expected. A review of de-escalation trials in the last few years highlighted a number of weaknesses which can effectively be addressed in the future. The BIG-NABCG Collaboration developed a roadmap for improving the design and implementation of de-escalation trials guided by patients’ insights, with recommendations on how to minimize treatment non-adherence, heterogeneity of results and risks of undertreatment, and support the selection of RFI, RFS and DMFS as meaningful endpoints.

**Acknowledgments**

The BIG-NABCG annual meeting 2019 and associated patient workshops were funded by the Breast Cancer Research Foundation (BCRF). We want to thank the patients advocates and all participants of the BIG-NABCG annual meeting for their valuable input.

**References**

1. Twist CJ, Schmidt M Lou, Naranjo A, et al. Maintaining Outstanding Outcomes Using Response- and Biology-Based Therapy for Intermediate-Risk Neuroblastoma: A Report From the Children’s Oncology Group Study ANBL0531. *J Clin Oncol*. 2019;37(34):3243-3255. doi:10.1200/JCO.19.00919

2. Tolbert VP, Matthay KK. Neuroblastoma: clinical and biological approach to risk stratification and treatment. *Cell Tissue Res*. 2018;372(2):195-209. doi:10.1007/s00441-018-2821-2

3. André T, Boni C, Navarro M, et al. Improved overall survival with oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment in stage II or III colon cancer in the MOSAIC trial. *J Clin Oncol*. 2009;27(19):3109-3116. doi:10.1200/JCO.2008.20.6771

4. Schmoll H-J, Tabernero J, Maroun J, et al. Capecitabine Plus Oxaliplatin Compared With Fluorouracil/Folinic Acid As Adjuvant Therapy for Stage III Colon Cancer: Final Results of the NO16968 Randomized Controlled Phase III Trial. *J Clin Oncol*. 2015;33(32):3733-3740. doi:10.1200/JCO.2015.60.9107

5. Kuebler JP, Wieand HS, O’Connell MJ, et al. Oxaliplatin Combined With Weekly Bolus Fluorouracil and Leucovorin As Surgical Adjuvant Chemotherapy for Stage II and III Colon Cancer: Results From NSABP C-07. *J Clin Oncol*. 2007;25(16):2198-2204. doi:10.1200/JCO.2006.08.2974

6. Land SR, Kopec JA, Cecchini RS, et al. Neurotoxicity From Oxaliplatin Combined With Weekly Bolus Fluorouracil and Leucovorin As Surgical Adjuvant Chemotherapy for Stage II and III Colon Cancer: NSABP C-07. *J Clin Oncol*. 2007;25(16):2205-2211. doi:10.1200/JCO.2006.08.6652

7. André T, Iveson T, Labianca R, et al. The IDEA (international duration evaluation of adjuvant chemotherapy) collaboration: Prospective combined analysis of phase III trials investigating duration of adjuvant therapy with the FOLFOX (FOLFOX4 or modified FOLFOX6) or XELOX (3 versus 6 months) re. *Curr Colorectal Cancer Rep*. 2013;9(3):261-269. doi:10.1007/s11888-013-0181-6

8. Grothey A, Sobrero AF, Shields AF, et al. Duration of Adjuvant Chemotherapy for Stage III Colon Cancer. *N Engl J Med*. 2018;378(13):1177-1188. doi:10.1056/NEJMoa1713709

9. Niraula S, Gyawali B. Optimal duration of adjuvant trastuzumab in treatment of early breast cancer: a meta-analysis of randomized controlled trials. *Breast Cancer Res Treat*. 2019;173(1):103-109. doi:10.1007/s10549-018-4967-8

10. Pondé N, Gelber RD, Piccart M. PERSEPHONE: are we ready to de-escalate adjuvant trastuzumab for HER2-positive breast cancer? *npj Breast Cancer*. 2019;5(1):1. doi:10.1038/s41523-018-0098-y

11. Earl H, Cameron D, Miles D, al. et. PERSEPHONE: duration of trastuzumab with chemotherapy in women with HER2-positive early breast cancer—six versus twelve months. *J Clin Oncol*. 2014;32.

12. Earl HM, Hiller L, Vallier AL, 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. doi:10.1016/S0140-6736(19)30650-6

13. Pivot X, Romieu G, Debled M, al. et. 6 months versus 12 months of adjuvant trastuzumab for patients with HER2-positive early breast cancer (PHARE): a randomised phase 3 trial. *Lancet Oncol*. 2013;14:741-748.

14. Pivot X, Romieu G, Debled M, et al. 6 months versus 12 months of adjuvant trastuzumab in early breast cancer (PHARE): final analysis of a multicentre, open-label, phase 3 randomised trial. *Lancet*. 2019;393(10191):2591-2598. doi:10.1016/S0140-6736(19)30653-1

15. Borchmann P, Goergen H, Kobe C, et al. PET-guided treatment in patients with advanced-stage Hodgkin’s lymphoma (HD18): final results of an open-label, international, randomised phase 3 trial by the German Hodgkin Study Group. *Lancet*. 2017;390(10114):2790-2802. doi:10.1016/S0140-6736(17)32134-7

16. 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 (London, England)*. 2019;393(10166):51-60. doi:10.1016/S0140-6736(18)32752-1

17. Gillison ML, Trotti AM, Harris J, et al. Radiotherapy plus cetuximab or cisplatin in human papillomavirus-positive oropharyngeal cancer (NRG Oncology RTOG 1016): a randomised, multicentre, non-inferiority trial. *Lancet*. 2019;393(10166):40-50. doi:10.1016/S0140-6736(18)32779-X

18. Tolaney S, Barry W, Dang C, al. et. Adjuvant paclitaxel and trastuzumab for node-negative, HER2-positive breast cancer. *N Engl J Med*. 2015;372:134-141.

19. 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. In: *Journal of Clinical Oncology*. Vol 37. American Society of Clinical Oncology; 2019:1868-1875. doi:10.1200/JCO.19.00066

20. Ras-dependent MDA, Veer LJ Van, Dai H, et al. Gene expression pro ® ling predicts clinical outcome of breast cancer. 2002;415(345).

21. Paik S, Shak S, Tang G, et al. A Multigene Assay to Predict Recurrence of Tamoxifen-Treated, Node-Negative Breast Cancer. *N Engl J Med*. 2004;351(27):2817-2826. doi:10.1056/NEJMoa041588

22. Cardoso F, Van’t Veer LJ, Bogaerts J, et al. 70-Gene signature as an aid to treatment decisions in early-stage breast cancer. *N Engl J Med*. 2016;375(8):717-729. doi:10.1056/NEJMoa1602253

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. doi:10.1056/NEJMoa1804710

24. Petkov VI, Kurian AW, Jakubowski DM, Shak S. Abstract P3-07-01: Breast cancer-specific mortality (BCSM) in patients age 50 years or younger with node-positive (N+) breast cancer (BC) treated based on the 21-gene assay in clinical practice. *Cancer Res*. 2020;80(4 Supplement):P3-07-01 LP-P3-07-01. doi:10.1158/1538-7445.SABCS19-P3-07-01

25. Cardoso F, van ’t Veer L, Poncet C, et al. MINDACT: Long-term results of the large prospective trial testing the 70-gene signature MammaPrint as guidance for adjuvant chemotherapy in breast cancer patients. *J Clin Oncol*. 2020;38(15\_suppl):506-506. doi:10.1200/jco.2020.38.15\_suppl.506

26. Bhattacharya IS, Kirby AM, Bliss JM, Coles CE. Can Interrogation of Tumour Characteristics Lead us to Safely Omit Adjuvant Radiotherapy in Patients with Early Breast Cancer ? Statement of Search Strategies Used and Sources of Information What are the Factors Contributing to the What are the Bene fi ts. *Clin Oncol*. 2018;30(3):158-165. doi:10.1016/j.clon.2017.12.022

27. D’Agostino RB, Massaro JM, Sullivan LM. Non-inferiority trials: Design concepts and issues - The encounters of academic consultants in statistics. *Stat Med*. 2003;22(2):169-186. doi:10.1002/sim.1425

28. Hudis C, Barlow W, Costantino J, al. et. Proposal for standardized definitions for efficacy end points in adjuvant breast cancer trials: the STEEP system. *J Clin Oncol*. 2007;25:2127-2132.

29. 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. doi:10.1093/annonc/mdv106

30. Calvert M, Blazeby J, Altman DG, Revicki DA, Moher D, Brundage MD. Reporting of patient-reported outcomes in randomized trials: The CONSORT PRO extension. *JAMA - J Am Med Assoc*. 2013;309(8):814-822. doi:10.1001/jama.2013.879

31. Calvert M, Kyte D, Mercieca-Bebber R, Slade A, Chan AW, King MT. Guidelines for inclusion of patient-reported outcomes in clinical trial protocols the spirit-pro extension. *JAMA - J Am Med Assoc*. 2018;319(5):483-494. doi:10.1001/jama.2017.21903

32. Coens C, Pe M, Dueck AC, et al. International standards for the analysis of quality-of-life and patient-reported outcome endpoints in cancer randomised controlled trials: recommendations of the SISAQOL Consortium. *Lancet Oncol*. 2020;21(2):e83-e96. doi:10.1016/S1470-2045(19)30790-9

Table 1

|  |  |  |
| --- | --- | --- |
|  | **IDEA** | **PERSEPHONE** |
| **Population** | Patients with stage III colon cancer were enrolled in six concurrently conducted phase 3 trials (n = 12, 834) | HER2-positive early breast cancer with indication for chemotherapy (n= 4088) |
| **Treatment** | 3 months vs 6 months of adjuvant therapy with either FOLFOX (fluorouracil, leucovorin, and oxaliplatin) or CAPOX (capecitabine and oxaliplatin). | 6 vs 12 months of trastuzumab, given concurrently or sequentially with neoadjuvant or adjuvant chemotherapy. |
| **Moment of randomization** | The six trials were conducted in heterogeneous settings. The study design did not call for patients to undergo a secondary randomization to CAPOX or FOLFOX. | Initially randomized before start of trastuzumab treatment, after protocol amendment any time up to and including the ninth cycle of trastuzumab. |
| **Non-inferiority margin** | Noninferiority of 3 months vs 6 months could be claimed if the upper limit of the two-sided 95% confidence interval of the hazard ratio did not exceed 1.12. The predefined noninferiority margin at a one-sided type I error rate of 0.025 assumes a 3-year disease-free survival rate of 72% in the 6-month therapy group. | 3% below the 4-year DFSl of the 12 months group. |
| **Analysis population** | Modified ITT for the primary analysis that includes all patients who had undergone randomization and had received at least one dose of a trial drug | ITT for the efficacy analysis; modified ITT, patients who received at least one dose of trastuzumab, for the safety analysis |
| **Efficacy results** | Noninferiority of 3 months of treatment vs 6 months was not confirmed in the ITT population (hazard ratio, 1.07; 95% confidence interval [CI], 1.00 to 1.15).  Noninferiority of the shorter regimen was seen for CAPOX (hazard ratio, 0.95; 95% CI, 0.85 to 1.06) but not for FOLFOX (hazard ratio, 1.16; 95% CI, 1.06 to 1.26). | 4-year disease-free survival was 89·4% (95% CI 87·9–90·7) in the 6 month group and 89·8% (95% CI 88·3–91·1) in the 12 month group (HR 1·07 (90% CI 0·93–1·24) |
| **Safety results** | 3 month group was associated with significantly lower rates of adverse events, independent of the chemotherapy regimen  Neurotoxicity of grade 2 or higher during active therapy and in the month after cessation of treatment was substantially lower in the 3-month group (16.6% with FOLFOX and 14.2% with CAPOX) than in the 6-month therapy group (47.7% with FOLFOX and 44.9% with CAPOX). | 6-month group reported SAEs in 373 patients (19%) vs 459 (24%) in the 12 month group, p=0·0002;  Number of patients in the 6-month group stopping early because of cardiotoxicity: 61 (3%) vs 146 (8%) in the 12 month group, p<0·0001. |
| **Comparison with current SOC** |  | 15% received neoadjuvant treatment  47% received concurrent trastuzumab and chemotherapy treatment |

Table 2

|  |  |  |
| --- | --- | --- |
|  | **MINDACT** | **TAILORX** |
| **Study population** | Overall population: histologically confirmed primary invasive breast cancer (stage T1 or T2 or operable T3) with up to three positive axillary nodes, n=6693 | Hormone-receptor-positive, HER2-negative, axillary node-negative breast cancer with recommendation or consideration of adjuvant chemotherapy based on the NCCN guidelines, n=10,273 |
| **Genomic test** | The “MammaPrint” 70-gene signature. | The “Oncotype DX” 21-gene signature. |
| **Clinical risk** | Low risk was defined as a 10-year probability of breast-cancer–specific survival without systemic therapy >88% for ER+ and >92% for ER- breast cancer based on Adjuvant! Online. | No further classification based on clinical risk was made. |
| **Treatment** | Patients with discordant results were randomly assigned to the chemotherapy group or the no-chemotherapy group on the basis of either the clinical result or the genomic result. | Women with a Oncotype DX score of 11 to 25 (mid-range) were randomized to receive endocrine therapy alone or endocrine therapy and chemotherapy. |
| **Statistical test** | Lower-boundary of the 95% confidence interval for the 5-year distant metastasis free survival 92% or higher. | Null hypothesis of no difference in invasive disease-free survival (STEEP) between the arms with a one-sided type I error of 10% and a 5% type II error. |
| **Primary analysis population** | Patients with high-risk clinical features and low-risk gene-expression profile who did not receive chemotherapy (per protocol), n=644 | Women with a Oncotype DX score of 11 to 25 (intention-to-treat), n =6711 |
| **Results** | In the clinical high-risk, genomic low-risk group who did not receive chemotherapy 5-year distant metastasis free survival was 94.7% (95% CI 92.5-96.2%). | The hazard ratio for invasive disease-free survival was 1.08 (95% CI, 0.94-1.24, p=0.26) for the endocrine alone arm compared to the chemotherapy + endocrine therapy arm. |

Table 3

|  |  |  |
| --- | --- | --- |
|  | **COMPASS** | **DECRESCENDO** |
| **Study population** | Stage 2 or 3a HER2-positive breast cancer, n=1250 | HER2-positiver, ER-negative, breast cancer, >=15mm and =<50mm, axillary node negative, non-metastatic, n=1065 |
| **Neoadjuvant treatment** | Paclitaxel or docetaxel plus trastuzumab and pertuzumab for 12 weeks. | Paclitaxel or docetaxel plus (subcutaneous) trastuzumab and pertuzumab for 12 weeks |
| **Adjuvant treatment if pCR achieved** | Complete 1 year of trastuzumab and pertuzumab, no further chemotherapy. | Complete 1 year of (subcutaneous) trastuzumab and pertuzumab, no further chemotherapy |
| **Adjuvant treatment if no pCR** | Randomization between T-DM1 alone or T-DM1 + tucatinib for 1 year. | T-DM1 to complete 1 year of anti-HER2 therapy, patients with a recidual burden score of 2 or larger can receive 3-4 cycles of anthracycline-based chemotherapy at the investigator’s discretion. |
| **Non-inferiority margin** | 3-year recurrence-free survival >=95%, lower 95% CI limit > 92% | 3-year recurrence-free survival >=94%, lower 95% CI limit > 92%  Hierachical analysis: first patients with HER2-enriched tumors by PAM50, then overall population |

**Figures**

**Figure 1.** The proposed roadmap to safe and well-designed de-escalation trials of systemic adjuvant therapies in solid tumors