

Strengths and pitfalls of NNT and NNH in early breast cancer escalation trials

Roberto Buonaiuto^{a,b,c,1} , Federica Pia Mangiacotti^{d,1} , Fabiola Giudici^e ,
Margherita Tafuro^{a,b} , Roberta Scarfetta^f , Alessandra Longobardi^d, Martina Pagliuca^{a,g} ,
Aldo Caltavuturo^a , Luca Arecco^c, Evandro de Azambuja^c, Mario Giuliano^d , Grazia Arpino^d,
Lucia Del Mastro^{h,i} , Carmine De Angelis^{a,b} , Michelino De Laurentiis^{a,b} ,
Fabio Puglisi^{j,k,*}

^a Clinical and Translational Oncology, Scuola Superiore Meridionale, Via Mezzocannone 4, Naples, 80134, Italy

^b Department of Breast and Thoracic Oncology, Istituto Nazionale Tumori IRCCS "Fondazione G. Pascale", Via Mariano Semmola 52, Napoli, 80131, Italy

^c Université Libre de Bruxelles (ULB), Hôpital Universitaire de Bruxelles (H.U.B.), Institut Jules Bordet, Bruxelles, Belgium

^d Oncology Unit, Department of Clinical Medicine and Surgery, University of Naples Federico II, Via Sergio Pansini 5, Naples, 80131, Italy

^e Unit of Cancer Epidemiology, CRO Aviano, National Cancer Institute, IRCCS, Via Franco Gallini 2, Aviano, 33081, Italy

^f Medical Oncology, Fondazione Policlinico Universitario Campus Bio-Medico, Via Alvaro Del Portillo 200, Roma, 00128, Italy

^g Molecular Predictors and New Targets in Oncology, Inserm U981, Université Paris-Saclay, Gustave Roussy, 114 Rue Edouard Vaillant, Villejuif, 94800, France

^h Department of Medical Oncology, IRCCS Ospedale Policlinico San Martino, Largo R. Benzi 10, Genova, 16132, Italy

ⁱ Dipartimento di Medicina Interna e Specialità Mediche (DIMI), University of Genova, Viale Benedetto XV 6, Genova, 16132, Italy

^j Department of Medical Oncology, CRO Aviano, National Cancer Institute, IRCCS, Via Franco Gallini 2, Aviano, 33081, Italy

^k Department of Medicine, University of Udine, Via Palladio 8, Udine, 33100, Italy

ARTICLE INFO

Keywords:

Number needed to treat
Number needed to harm
Early-stage breast cancer

ABSTRACT

Background: The management of patients with high-risk early breast cancer (eBC) requires balancing between clinical benefit and toxicity. Number Needed to Treat (NNT) and Number Needed to Harm (NNH) may provide a clinically intuitive framework to assess benefit–risk profiles, but their use in eBC has not been systematically explored.

Methods: We analysed 10 phase II–III randomized clinical trials (RCTs) of treatment escalation strategies in patients with mid to high-risk eBC. To ensure consistency and cross-trial comparison, invasive disease-free survival (IDFS), disease-free survival (DFS), event-free survival (EFS), and overall survival (OS) were extracted at 3 and 5 years (y). Safety outcomes included any-grade (G) and $G \geq 3$ adverse events (AEs), G5 AEs, and treatment discontinuation (TD). Absolute risk reduction (ARR) and absolute risk increase (ARI) were used to calculate NNT and NNH, respectively, reported with 95% CIs.

Results: In patients with the HER2+ eBC, KATHERINE showed the most favorable efficacy, with NNTs of 9 and 8 for IDFS at 3 and 5 years, respectively, and a NNT of 27 for OS at 5 years. APHINITY reported the most favorable safety profile, with the highest NNHs for $G \geq 3$ AEs (15) and TD (80). In patients with TNBC, CREATE-X reported the lowest NNTs for DFS (7 at both 3 and 5 years). The most favorable results in OS were observed in OlympiA (NNT 26) and KEYNOTE-522 (NNT 20). Regarding safety, KEYNOTE-522 achieved the most favorable NNH for $G \geq 3$ AEs (26), while OlympiA showed the best NNH for TD (16). In patients with HR+/HER2– eBC, OlympiA showed the strongest efficacy results, with the lowest NNTs for IDFS at 3 (10) and 5 years (9), as well as for OS at 5 years (26). The most favorable NNHs were reported for $G \geq 3$ AEs in OlympiA (8) and for TD in monarchE (19).

Conclusion: NNT and NNH provide intuitive metrics to communicate benefit–risk trade-offs in high-risk eBC, although their variability across trials, subgroups, and timepoints requires cautious interpretation.

* Corresponding author. Department of Medicine, University of Udine, via Palladio 8, 33100, Udine, Italy.

E-mail address: fabio.puglisi@uniud.it (F. Puglisi).

¹ Contributed equally.

1. Introduction

The management of early breast cancer (eBC) requires a careful balance between maximizing the therapeutic efficacy and minimizing treatment-related toxicity in the adjuvant setting. Patients identified as being at higher risk of recurrence, based on clinicopathological features such as nodal status, histological grade, Ki-67 index, and tumor size, as well as molecular profiling through multigene assays, are often candidates for treatment escalation strategies. These approaches aim to reduce the risk of recurrence by intensifying conventional therapies [1, 2] or by integrating novel targeted agents [1–3]. However, the potential benefits of these strategies must be carefully balanced against the higher risk of adverse events (AEs), underscoring the complexity of clinical decision-making in this setting. Clearly communicating the expected benefits and potential toxicities of treatment interventions to patients is crucial in clinical practice, particularly for those with early-stage disease. In this context, simple and intuitive measures that express the absolute magnitude of both benefits and harms may be especially valuable in guiding shared decision-making.

The Number Needed to Treat (NNT) and Number Needed to Harm (NNH) are quantitative metrics that offer a structured framework for evaluating the benefit-risk profile of therapeutic interventions [4–7]. NNT represents the number of patients who need to be treated to prevent one additional adverse clinical event, whereas NNH indicates the number of patients who must be treated for one additional patient to experience an AE. A lower NNT reflects greater treatment efficacy, with the theoretical optimal value of 1 indicating that every treated patient benefits from the new strategy, while no benefit is observed in the control group, while higher NNT values indicate reduced therapeutic effectiveness; in contrast, a lower NNH indicates a higher frequency of AEs, whereas higher NNH values reflect a lower risk of harm, corresponding to a more favorable safety profile. Although the NNT and the NNH provide a pragmatic framework for assessing the balance between efficacy and safety, these metrics are not without limitations. Their interpretation is highly context-dependent, influenced by the clinical setting, patient population, and temporal dynamics of treatment effects, and may therefore fail to fully capture the complexity of clinical outcomes.

In this study, we aim to critically appraise the application of NNT and NNH in the context of patients with high-risk eBC through a descriptive analysis of RCTs investigating treatment escalation strategies. By examining these metrics, our analysis seeks to elucidate their potential to inform clinical decision-making, while simultaneously underscoring their methodological constraints and interpretative challenges.

2. Methods

2.1. Research strategy, inclusion and exclusion criteria

We performed a descriptive, trial-level analysis of randomized phase II–III clinical trials evaluating therapeutic escalation strategies in early breast cancer (eBC). Treatment-escalation strategies for patients with intermediate-to high-risk eBC were identified from current international guidelines, including ESMO and NCCN [8,9]. For each strategy, we selected the pivotal randomized trial that supported EMA and/or FDA approval in the early breast cancer setting.

Specifically, we included trials that.

- I. Evaluated escalation of systemic therapy in the (neo)adjuvant setting;
- II. Enrolled patients with intermediate-to high-risk eBC, defined as stage II–III BC according to the AJCC Cancer Staging Manual, 7-8th Edition;
- III. Investigated therapeutic agents that received regulatory approval and are currently recommended by international guidelines [8, 9].

TRAIN-2 [10] and CREATE-X [11] were included despite not being registrational trials, as they established carboplatin-based chemotherapy and adjuvant capecitabine as treatment options currently recommended by international guidelines for HER2-positive and TNBC, respectively.

2.2. Data extraction

Two reviewers (RB, FM) independently extracted data using a standardized form. For each trial, we collected patient characteristics, median follow-up, and endpoints information, including type (primary vs secondary) and assessment timing (interim, final, or descriptive). Efficacy endpoints included invasive disease-free survival (IDFS), disease-free survival (DFS), event-free survival (EFS), and overall survival (OS). Safety data on adverse events (AEs) were collected, including any-grade AEs, grade ≥ 3 AEs, grade 5 events, AEs leading to treatment discontinuation (TD), and the most frequently reported grade ≥ 3 AEs.

2.3. Calculation of efficacy and toxicity metrics

Efficacy and toxicity measures were calculated for each trial using the most recent data available at the time of analysis. The Absolute Risk Reduction (ARR) was defined as the difference between the event (disease recurrence or death) rate in the control arm (CER) and those in the experimental arm (EER) ($ARR = CER - EER$). The Absolute Risk Increase (ARI) was defined as the difference in $G \geq 3$ adverse events rate between the experimental arm (ETR) and the control arm (CTR) ($ARI = ETR - CTR$).

The NNT and NNH were derived as the reciprocals of ARR and ARI, being $NNT = 1/ARR$ and $NNH = 1/ARI$, respectively.

For example, if the control recurrence rate is 0.20 and the treatment recurrence rate is 0.15, the ARR is 0.05, yielding an NNT of $1/0.05 = 20$. For instance, if the adverse event rate is 0.08 with treatment versus 0.04 with control, the ARI is 0.04, yielding an NNH of $1/0.04 = 25$.

To ensure methodological consistency and facilitate cross-trial comparisons, survival outcomes were assessed at standardized landmarks of 3 and 5 years. When survival probabilities were not directly reported, estimates were derived by reconstructing individual patient data (IPD) from latest available published Kaplan-Meier curves using the IPDfromKM Shiny app [12]. Reconstruction accuracy was considered acceptable when the root mean squared error (RMSE) was ≤ 0.05 , the mean absolute error ≤ 0.02 , and the maximum absolute error was ≤ 0.05 [12].

2.4. Confidence intervals for the NNTs and NNHs

The conventional method for calculating 95% confidence intervals (CIs) for NNTs/NNHs relies on the Wald method, which often yields overly narrow intervals. To address this limitation, we applied the Newcombe-Wilson Hybrid Method, which provides more accurate and

Table 1
Key study characteristics in phase II-III trials included in the analysis.

Year	Trial	Phase	Experimental arm	Control arm	Number of subjects in the Experimental arm	Number of subjects in the Control arm	Primary Endpoint	Secondary Endpoint	Median follow-up, months	Type of analysis primary endpoint	Type of analysis secondary endpoint
2025	KATHERINE [19]	3	T-DM1	Trast	743	743	IDFS	OS, DFS, DRFI, safety	101	Final	Interim analysis II (OS)
2024/2025	APHINITY (N+) [20,21]	3	Pert + Trast + CT	Plac + Trast + CT	1503	1502	IDFS	OS, DFS, RFI, DRFI, safety, QoL, IDFS ^a	100 (IDFS) 136 (OS)	Descriptive analysis	Final analysis (OS)
2024	KEYNOTE-522 [27]	3	Pembro + CT	Plac + CT	784	390	EFS, pCR	OS, pCR in ypT0 ypN0 and ypT0/Tis, pCR, EFS, and OS in the PD-L1 CPS ≥ 1 population, safety	75.1	Interim	Interim
2024	OlympiA HR+ [26]	3	Ola	Plac	168	157	IDFS	DDFS, OS	73.2	Interim III	Interim III (OS)
2024	OlympiA HR- [26]	3	Ola	Plac	751	758	IDFS	DDFS, OS	73.2	Interim III	Interim III (OS)
2017	CREATE-X [11]	3	Cape	Plac	443	444	DFS	OS	42	Final	Final (OS)
2021	ExteNET HR+ ≤ 1 y [23]	3	Neratinib	Plac	670	664	IDFS	DFS, DDFS, TDR, incidence of first occurrence of CNS, OS, safety	96 (OS)	Final	Final (OS)
2025	monarchE [31]	3	Abema + ET	ET	2808	2829	IDFS (STEEP criteria)	DRFS, OS, safety, PRO, PK	76.2	Descriptive	Interim (OS)
2025	NATALEE [25]	3	Ribo + ET	ET	2549	2552	IDFS (STEEP criteria)	RFS, DDFS, OS, safety, PK	55.4 (IDFS), 56.5 (OS)	Prespecified time triggered analysis	Descriptive analysis (OS)
2021	TRAIN-2 [10]	3	PTC + Pert	FEC + Trast + Pert	219	219	CR	EFS, OS, safety	36	Interim I	Interim II (EFS)
2016	NeoSphere [22]	2	Trast + Doce (A)	Trast + Pert + Doce (B) Trast + Pert (C) Pert + Doce (D)	101 (B)	103(A) 96(C) 92 (D)	pCR	PFS, DFS	60	Final	Final

Abbreviations: A, arm A; Abema, abemaciclib; AE, adverse event; B, arm B; C, arm C; Cape, capecitabine; CNS, central nervous system; CPS, combined positive score; CT, chemotherapy; D, arm D; DDFS, distant disease-free survival; DFS, disease-free survival; Doce, docetaxel; DRFI, distant relapse-free interval; DRFS, distant relapse-free survival; EFS, event-free survival; ET, endocrine therapy; FEC, fluorouracil, epirubicin hydrochloride, cyclophosphamide; IDFS, invasive disease-free survival; Obs, observation; Ola, Olaparib; OS, overall survival; pCR, pathologic complete response; PD-L1, programmed death-ligand 1; Pembro, pembrolizumab; Pert, pertuzumab; PK, pharmacokinetics; Plac, placebo; PRO, patient-reported outcome; PTC, paclitaxel, trastuzumab, carboplatin; QoL, quality of life; RFI, recurrence-free interval; RFS, recurrence-free survival; Ribo, ribociclib; STEEP, standardized definitions for efficacy end points; T-DM1, trastuzumab emtansine; TDR, time to distant recurrence; Trast, trastuzumab.

^a Including second primary non-breast cancer, per the STEEP definition.

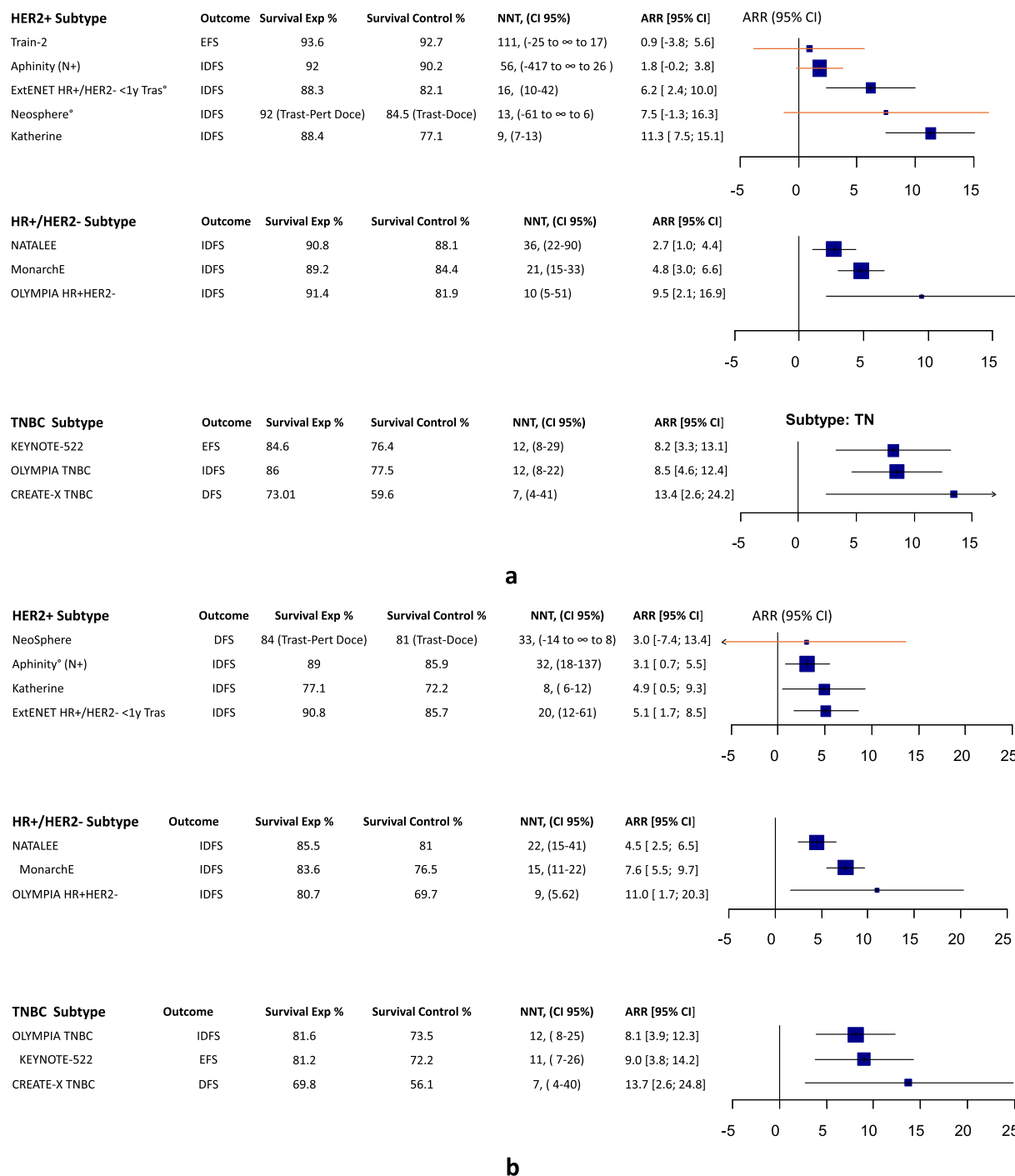


Fig. 1. Number needed to treat: IDFS/EFS/DFS.

a: 3-year IDFS/EFS/DFS; b: 5-year IDFS/EFS/DFS.

°: Synthetic individual patient data (IPD) were generated from the digitized curves

Abbreviations: ARR, absolute risk reduction; CI, confidence interval; Control, control arm; DFS, disease-free survival; EFS, event-free survival; Exp, experimental arm; HER2, Human epidermal growth factor receptor 2; HR, hormone receptor; IDFS invasive disease-free survival; ITT, intention-to-treat population; NNH, number needed to harm; NNT, number needed to treat; Tras, trastuzumab; y, years.

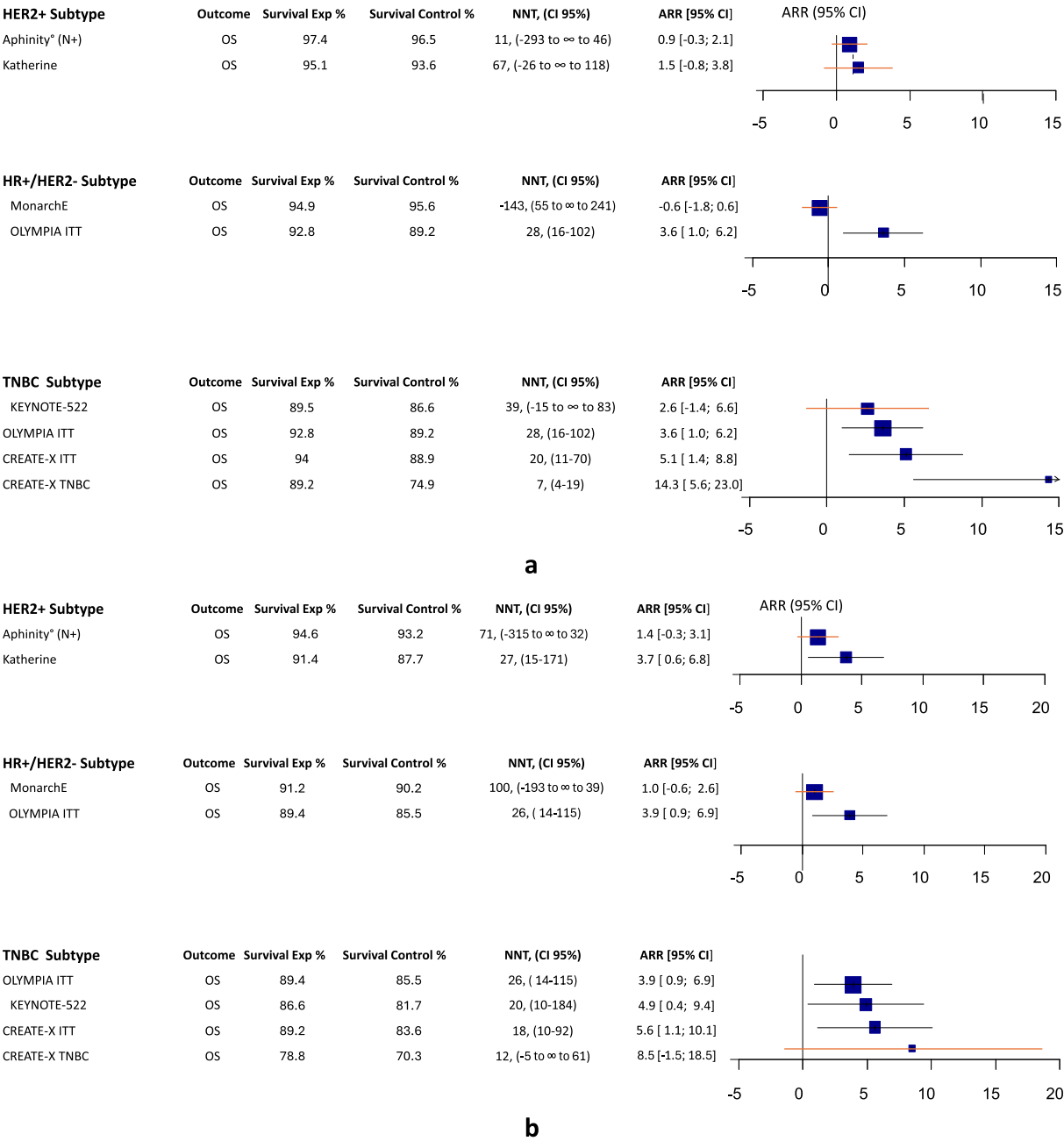


Fig. 2. Number needed to treat: OS.
a: 3-year OS; b: 5-year OS.
°: Synthetic individual patient data (IPD) were generated from the digitized curves
Abbreviations: ARR, absolute risk reduction; CI, confidence interval; Control, control arm; Exp, experimental arm; HER2, Human epidermal growth factor receptor 2; HR, hormone receptor; ITT, intention-to-treat population; NNH, number needed to harm; NNT, number needed to treat; OS, overall survival; TN, triple negative; y, years.

informative 95% CIs estimates [13–17]. While the point estimate of the NNT (or NNH) is simply the reciprocal of the ARR or ARI, calculating its confidence interval is more complex, particularly when the 95% CI for ARR includes zero. In such cases, the 95% CI for NNT necessarily includes infinity ($1/0 = \infty$), implying that an infinite number of patients would be required to observe a benefit, a concept that is counterintuitive.

To improve interpretability, Altman et al. [18] recommended distinguishing between NNTB (Number Needed to Treat for Benefit) and NNTH (Number Needed to Treat for Harm). For example, if the ARR is 0.10 (95% CI: -0.05 to 0.25), the result would be expressed as NNTB of 10 (95% CI: NNTH = 20 to ∞ to NNTB = 4), making explicit that the true

effect could plausibly range from harm to benefit.

In this example, the NNTH of 20 corresponds to the negative lower bound of the ARR confidence interval (-0.05) and indicates that, at one extreme, one additional patient may be harmed for every 20 patients treated, suggesting a potentially weaker or even detrimental treatment effect. The infinity (∞) reflects the crossing zero within the ARR CI, implying the possibility of no true difference between the treatment and control. The NNTB of 4 corresponds to the positive upper bound of the ARR CI (0.25) and suggests that, at the other extreme, only four patients would need to be treated for one additional patient to benefit, indicating a potentially stronger treatment effect. When the 95% CI of the ARR included zero, the corresponding NNT confidence interval was reported

Table 2
Reported adverse events in phase II-III trials included in the analysis.

Trial	Experimental arm	Control arm	Drug discontinuation in the Experimental arm (%)	Drug discontinuation in the Control arm (%)	Grade 5 AEs in the Experimental arm (%)	Grade 5 AEs in the Control arm (%)	Grade ≥ 3 AEs in the Experimental arm (%)	Grade ≥ 3 AEs in the Control arms (%)	Most common Grade ≥ 3 A E in the Experimental arm (%)	Most common Grade ≥ 3 AEs in the Control arm (%)	Most common any-grade AEs in the Experimental arm (%)	Most common any-grade AEs in the Control arm (%)
KATHERINE [19]	T-DM1	Trast	133 (18)	15 (2.1)	1 (0.1) ^a	0 ^a	190 (25.7)	111 (15.4)	Decreased platelet count 27 (5.7)	Hypertension 9 (1.2)	Fatigue 366 (49.5)	Fatigue 243 (33.8)
APHINITY [20,32,33]	Pert + Trast + CT	Plac + Trast + CT	186 (7.3)	155 (6.2)	23 (1%) ^c	32 (1.3) ^c	1517 (64.2)	1379 (57.3)	Neutropenia 385 (16.3)	Neutropenia 377 (15.7)	/	/
KEYNOTE-522 [27]	Pembro + CT	Plac + CT	216 (27.6) ^a	55 (14.1) ^a	4 (0.5) ^a	1 (0.3) ^a	604 (77.1) ^a	285 (73.3) ^a	Neutropenia 270 (34.6) ^a	Neutropenia 130 (33.2) ^a	Nausea 490 (62.7) ^a	Nausea 246 (63.2) ^a
OlympiA [26]	Ola	Plac	98 (10.8)	42 (4.6)	5 (<1%)	10 (1.1)	223 (24.5)	102 (11.3)	Anemia 79 (8.7)	Neutrophil count decreased 7 (0.8)	Nausea 520 (57.1)	Fatigue 248 (27.4)
CREATE-X [11]	Cape	Plac	101	0	0 ^a	0 ^a	/	/	Hand foot syndrome 49 (11.1)	Hypertransaminasemia 4 (0.8)	Hand foot syndrome 325 (73.4)	/
ExteNET HR+ ≤ 1y [23]	Neratinib	Plac	178 (27) ^b	30 (5) ^b	1 (0.1) ^b	0 ^b	327 (49) ^b	76 (12) ^b	Diarrhea 261 (39) ^b	Diarrhea 7 (1) ^b	/	/
monarchE [31,34,35]	Abema + ET	ET	180 (6.4)	30 (1.1)	2 ^a	0 ^a	1289 (46.2)	439 (15.7)	Neutropenia 548 (19.6) ^b	Neutropenia 24 (0.9) ^b	/	/
NATALEE [25,36,37]	Ribo + ET	ET	509 (20)	124 (4.9)	0 ^a	0 ^a	1471 (58.2)	437 (17.9)	Neutropenia 1121 (44.4)	Liver related AE 31 (1.3)	Neutropenia 1587 (62.8)	Arthralgia 1083 (44.4)
TRAIN-2 [10]	PTC + Pert	FEC + Trast + Pert	32 (15)	27 (12)	1	0	/	/	Neutropenia 118 (54)	Neutropenia 131 (60)	/	/
NeoSphere [22]	Trast + Doce (A)	Trast + Pert + Doce (B) Trast + Pert (C) Pert + Doce (D)	0	(B) 5 (5) (C) 8 (7) (D) 5 (5)	(B) 1 (1) ^a	(A) 0 ^a	(B) 78 (73)	(A) 87 (81) (C) 75 (60) (D) 74 (79)	Neutropenia (B) 59 (55)	Neutropenia (A) 71 (66) (C) 40 (37) (D) 60 (64)	(B) Alopecia 73(68)	(A) Alopecia 80 (75) (C) Alopecia 59 (55) (D) Neutropenia 69 (73)

Abbreviations: A, arm A; Abema, abemaciclib; AE, adverse event; B, arm B; C, arm C; Cape, capecitabine; CT, chemotherapy; D, arm D; Doce, docetaxel; ET, endocrine therapy; FEC, fluorouracil, epirubicin hydrochloride, cyclophosphamide; G, grade; Obs, observation; Ola, Olaparib; Pembro, pembrolizumab; Pert, pertuzumab; Plac, placebo; PTC, paclitaxel, trastuzumab, carboplatin; Ribo, ribociclib; T-DM1, trastuzumab emtansine; Trast, trastuzumab; y, year.

^a Treatment related AEs;/Not available.

^b Treatment emergent AEs..

^c Fatal AEs.

according to the Altman convention [18], in the following format: (−20 to ∞ to 4), distinguishing between the number needed to treat for benefit and the number needed to treat for harm. In this setting, the interval necessarily crosses infinity and indicates that the data are compatible with benefit, no effect, or harm. Therefore, these estimates were considered descriptive measures of uncertainty and were not interpreted as evidence of treatment benefit.

3. Results

Ten RCTs (KATHERINE [19], APHINITY [20,21], NeoSphere [22], TRAIN-2 [10], ExteNET [23], monarchE [24], NATALEE [25], CREATE-X [11], OlympiA [26] and KEYNOTE-522 [27]) were included in the analysis. In APHINITY [20,21], only patients with node-positive disease were considered, whereas in ExteNET [23] the analysis was restricted to the HR+/HER2+ subgroup who had completed adjuvant trastuzumab-based therapy within one year, in accordance with current international guidelines [8,28] and regulatory approval [29,30]. The detailed study characteristics, including treatment arms, are summarized in Table 1.

3.1. NNT

ARRs and NNTs with 95% CIs at 3 and 5 years are shown in Figs. 1 and 2. Additional NNT estimates at other time points are provided in the Supplementary Materials (Supplementary Table 1). Moreover, OS NNT estimates from trials without statistically significant OS results are reported in the Supplementary Material, because their NNT confidence intervals necessarily crossed zero, and should not be interpreted as

evidence of benefit (Supplementary Table 1).

3.1.1. HR+/HER2− disease

The OlympiA [26] trial reported the most favorable NNT for IDFS at 3 years (10; 95% CI, 5–51), followed by monarchE [24] (21; 95% CI, 15–33), and NATALEE [25] (36; 95% CI, 22–90). At 5 years, the benefit remained consistent, with OlympiA [26] again showing the lowest NNT (9; 95% CI, 5–62), followed by monarchE [24] (15; 95% CI, 11–22) and NATALEE [25] (22; 95% CI, 15–41). For OS at 5 years, OlympiA [26] showed the clearest absolute benefit (26; 95% CI, 14–115) followed by monarchE [24] (100; 95% CI, −193 to ∞ to 39).

3.1.2. HER2+ disease

KATHERINE [19] reported the most favorable NNT for IDFS at 3 years (9; 95% CI, 7–13), followed by NeoSphere [22] (13; 95% CI, −6 to ∞ to 66), ExteNET [23] (16; 95% CI, 10–42), APHINITY [20] (56; 95% CI, −417 to ∞ to 26) and TRAIN-2 [10] (111; 95% CI, −25 to ∞ to 17). At 5 years, KATHERINE [19] (8; 95% CI, 6–12) and ExteNET [23] (20; 95% CI, 12–61) maintained consistent benefit, followed by APHINITY [20] (32; 95% CI, 18–137), and NeoSphere [22] (33; 95% CI, −14 to ∞ to 8). For OS at 5 years, KATHERINE [19] showed a statistically supported absolute benefit, with an NNT of 27 (95% CI, 15–171), followed by APHINITY [21] (71; 95% CI, −315 to ∞ to 32). OS data were not available for NeoSphere.

3.1.3. Triple negative disease

At 3 years, the most favorable NNT for DFS was observed in the CREATE-X trial [11] (7; 95% CI, 4–41), followed by OlympiA [26] (12; 95% CI, 8–22) and KEYNOTE-522 [27] (12; 95% CI, 8–29) for IDFS and

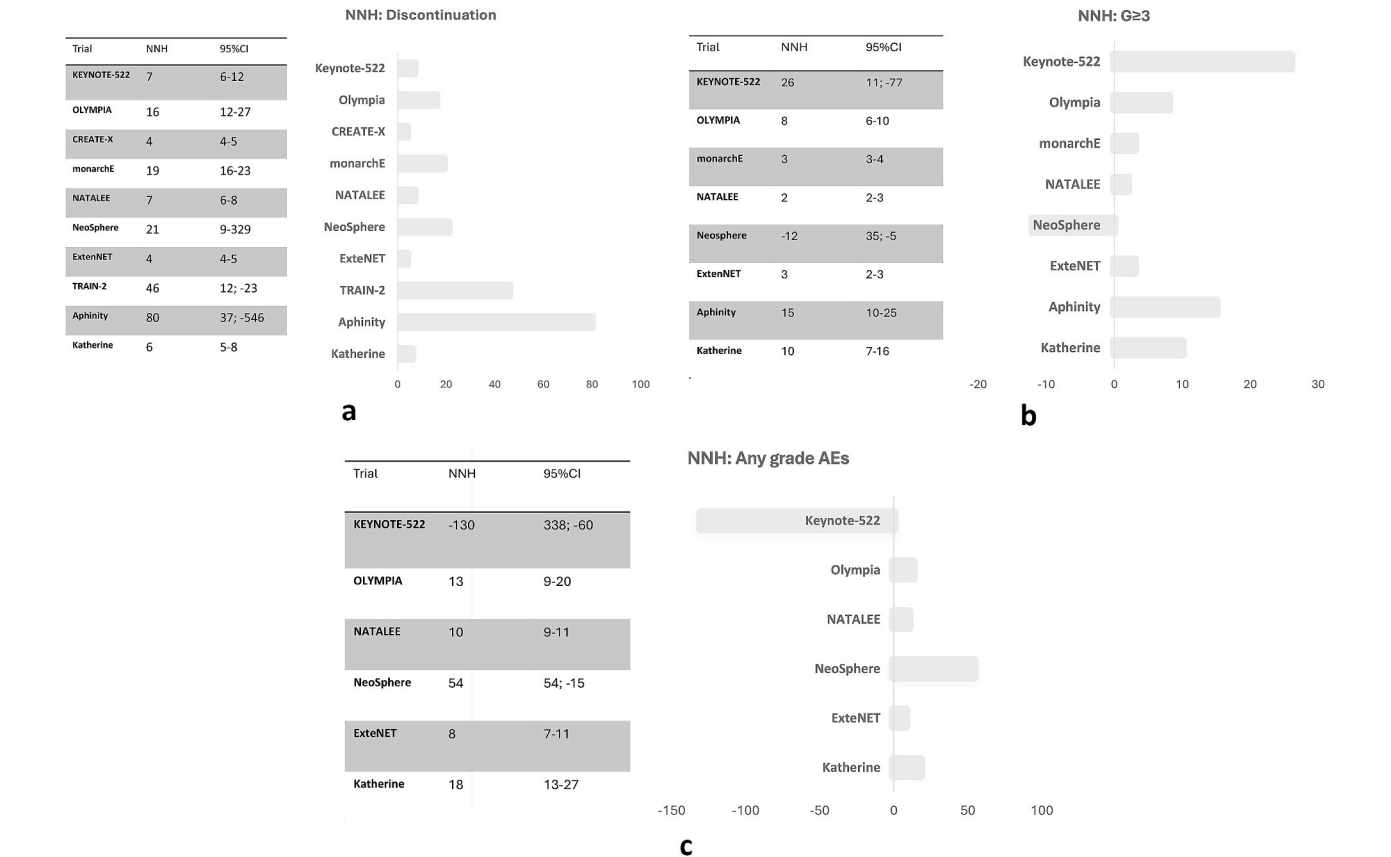


Fig. 3. Number needed to harm.
a: Discontinuation
b: G ≥ 3 AEs
c: Any grade AEs.

EFS, respectively. At 5 years, CREATE-X [11] demonstrated a sustained benefit (7; 95% CI, 4–40), with comparable results reported in KEYNOTE-522 [27] (11; 95% CI, 7–26) and OlympiA [26] (12; 95% CI, 8–25).

For OS, CREATE-X [11] reported the lowest NNT in the intention-to-treat (ITT) population (18; 95% CI, 10–92), followed by OlympiA [26] (26; 95% CI, 14–115) and KEYNOTE-522 [27] (20; 95% CI, 10–184).

3.2. NNH

NNHs were calculated for any-grade AEs, $G \geq 3$ AEs, G5 AEs, AEs leading to TD, and the most frequently reported $G \geq 3$ AE (Table 2).

3.2.1. HR+/HER2– disease

In patients with HR+/HER2– disease, the NATALEE [36] trial showed the highest toxicity risk, with NNHs of 10 (95% CI, 9–11) for any-grade AEs, 3 (95% CI, 2–3) for $G \geq 3$ AEs, and 7 (95% CI, 6–8) for TD (Fig. 3).

3.2.2. HER2+ disease

In patients with HER2+ disease, ExteNET [23] reported the lowest NNHs for any-grade AEs (8; 95% CI, 7–11), $G \geq 3$ AEs (3; 95% CI, 2–3), and TD (4; 95% CI, 4–5), indicating the highest toxicity risk among the available HER2-directed strategies. KATHERINE [38] was also associated with substantial toxicity, with NNHs of 10 (95% CI, 7–16) for $G \geq 3$ AEs and 6 (95% CI, 5–8) for TD (Fig. 3).

3.2.3. Triple negative disease

In patients with Triple Negative disease, OlympiA [26] showed the lowest NNHs for any-grade AEs (13; 95% CI, 9–20) and $G \geq 3$ AEs (8; 95% CI, 6–10). Additionally, CREATE-X [11] (4; 95% CI, 4–5) and KEYNOTE-522 [27] (7; 95% CI, 6–12) were associated with the highest risk of TD (Fig. 3). Detailed NNHs for most frequent $G \geq 3$ AEs and for G5 AEs are provided in Supplementary Table 2.

4. Discussion

In this study, we analysed RCTs evaluating therapeutic escalation strategies in patients with intermediate-to high-risk EBC and derived NNT and NNH to characterize the balance between treatment benefit and harm. Overall, our findings highlight substantial heterogeneity in benefit–risk profiles across disease subtypes and treatment strategies, with considerable variability in toxicity burden. As intuitive absolute measures, NNT and NNH may facilitate clinical interpretation and support patient-centered discussions by enabling the presentation of therapeutic trade-offs in a numerically grounded but user-friendly format. These metrics may also provide a simple framework for health-economic considerations, by linking the number of patients treated to the number of clinical events prevented or adverse events incurred [39,40]. For example, stating that “one in 12 patients treated will benefit over five years” may be more impactful and comprehensible than citing a “7–8% absolute risk reduction”. However, it is important to acknowledge that the NNT and NNH estimates are time-dependent and may not fully capture the dynamics of treatment effects. Hazard ratios derived from time-to-event analyses reflect the relative risk over the entire follow-up period, whereas NNTs provide a snapshot of the absolute benefit at a fixed time point, potentially overlooking more nuanced temporal patterns. This discrepancy may lead to divergent interpretations, particularly when comparing heterogeneous trials with varying follow-up durations. For example, treatments with delayed efficacy (e.g., immunotherapy) or late toxicities (e.g., CDK4/6 inhibitors) may have their benefit or harm underestimated (or overestimated) when assessed only at early time points. Therefore, meaningful interpretation of NNT and NNH requires standardized reporting of the outcome, comparator and time horizon, particularly when these metrics are used

for cross-trial comparisons [41].

To partially address this issue, we reported the NNT values separately by breast cancer subtype at fixed time points (3 and 5 years), thereby illustrating how the magnitude of benefit may evolve over time. Among the included trials, OlympiA [26], CREATE-X [11], and KEYNOTE-522 [27] in patients with TNBC disease, and KATHERINE [19] in patients with HER2+ disease, showed consistent NNT values over time, suggesting both early and sustained clinical benefits. However, the distinct biological behaviors of BC subtypes substantially affect the interpretation of these measures. TNBC is typically associated with a high risk of early recurrence, concentrating treatment effects within the first years, whereas HR+/HER2– tumors tend to relapse later, with a persistent long-term risk of recurrence [42]. These biological patterns are critical for the proper contextualization of NNT and NNH estimates. Consequently, early NNTs estimates may adequately reflect treatment benefit in TNBC, while in HR+/HER2– disease they may underestimate the magnitude of long-term benefit, underscoring the importance of extended follow-up.

In particular, in patients with HR+/HER2– EBC, both monarchE [24] and NATALEE [25] trials demonstrated a progressive reduction in NNT for IDFS even after treatment discontinuation, suggesting a potential carry-over effect, defined as a persistent or increasing treatment benefit after discontinuation of active therapy and previously described for adjuvant endocrine therapies [43,44]. At the same time, these trials highlighted the relevant toxicity burden associated with CDK4/6 inhibitor-based strategies. Neutropenia was the most frequent $G \geq 3$ AE, and TD rates were considerable. In high-risk early TNBC, treatment options commonly include capecitabine, pembrolizumab, or olaparib [8, 28]. The CREATE-X [11] trial demonstrated the most favorable 5-year NNT for OS; however, only ~31% of patients had TNBC, none had received prior carboplatin or neoadjuvant immunotherapy, and capecitabine was associated with substantial toxicity, with hand-foot syndrome being the most frequent severe AE. Moreover, the comparator consisted of observation without any additional systemic treatment, which may partially account for the relatively low NNT and limits direct comparability with more contemporary escalation strategies. By contrast, findings from OlympiA [26] and KEYNOTE-522 [27], which more closely reflect current patient profiles and modern treatment algorithms, may be more representative of present clinical practice. In patients with HER2+ disease, the KATHERINE [19] trial demonstrated consistent NNTs for both IDFS and OS, although the NNH for TD reflected a clinically relevant incidence of toxicity. However, the study population did not include patients who had received neoadjuvant pertuzumab. In contrast, the NeoSphere [22] and APHINITY [20] trials evaluated the addition of pertuzumab in the neoadjuvant and adjuvant settings, respectively, achieving comparable benefits in terms of NNT for IDFS, with a reassuring safety profile, particularly regarding grade ≥ 3 AEs and TD. On the contrary, in the ExteNET [23] trial, despite a modest reduction in recurrence risk, reflected by NNT estimates, escalation therapy with neratinib was associated with substantial gastrointestinal toxicity, resulting in unfavorable NNHs for both $G \geq 3$ AEs and TD. These findings raise concerns regarding the clinical applicability of this strategy in contemporary practice.

Despite the revision of the Consolidated Standards of Reporting Trials (CONSORT) statement, which recommends reporting RCT results using NNT and NNH for both binary and time-to-event outcomes [45], these metrics remain underused in the medical literature [46]. Moreover, when reported, CIs for NNTs are often omitted, which represents an important limitation [18].

We acknowledge the inherent limitations of the NNT and NNH estimates. In several cases, particularly NNTs for OS and NNHs for all-grade AEs, CI were wide and/or crossed zero, reflecting uncertainty arising from small absolute differences between treatment arms. In particular, a major methodological caveat concerns the interpretation of estimates derived from comparisons in which the 95% CI of the underlying absolute risk difference, either ARR or ARI, includes zero [47]. In

such cases, the data are compatible with both benefit and harm, and the corresponding NNT or NNH point estimate should not be interpreted in isolation as quantitative evidence of treatment effect or toxicity. We therefore reported these estimates using the Altman convention [18], which makes the bidirectional uncertainty explicit, and caution readers against drawing conclusions about the magnitude of benefit or harm from non-significant comparisons, particularly for OS endpoints, where several included trials had not yet reached statistical significance at the time of the most recent reported analysis.

Accordingly, although easily calculated and intuitively appealing, NNT and NNH should be considered alongside relative risk measures and relevant clinical context. In particular, these metrics are influenced by both the choice of comparator and the baseline risk of the study population. Trials employing less intensive control arms may yield more favorable NNT estimates, as in CREATE-X trial [11], where the comparator for TNBC consisted of observation alone. Furthermore, ongoing improvements in standard therapies progressively reduce baseline recurrence risk, potentially leading to higher NNTs over time even for effective interventions, particularly for OS outcomes. Because NNT is inherently dependent on baseline risk, patients at higher risk of recurrence tend to derive greater absolute benefit, as observed for the BRCA-mutated population enrolled in the OlympiA trial [26]. NNH estimates similarly have limitations, as they may not fully capture qualitative aspects of toxicity, including chronic treatment burden and quality-of-life impact, which are particularly relevant in the curative setting. Moreover, beyond their statistical interpretation, NNT and NNH should also be considered within the broader context of benefit–harm communication in oncology. The way treatment benefits and toxicities are presented to patients, including the format, framing, timing, and amount of information provided, may influence understanding, emotional response, and ultimately treatment choices [48,49]. In particular, numerical or visual formats may improve comprehension compared with purely verbal descriptions, while excessive or poorly timed information on treatment consequences may increase distress and undermine confidence [50]. In this context, NNT and NNH may help translate absolute benefit and harm into more clinically intuitive estimates, but they should not be presented as stand-alone measures. On the contrary, their use in clinical communication should remain individualized, avoiding both oversimplification and information overload. Additionally, they should support individualized discussions in which the numerical estimate is interpreted in light of the clinical relevance of the endpoint, the nature and burden of toxicity, the uncertainty of the estimate, and the patient's values and preferences. This study has several limitations that should be acknowledged. Heterogeneity across trials in terms of study design, patient populations, median follow-up, treatment duration, and endpoint definitions limits direct cross-trial comparability. In particular, DFS, IDFS, and EFS are distinct intermediate endpoints and should not be considered fully interchangeable. Accordingly, efficacy endpoints were extracted and reported as defined in each individual trial, as summarized in [Supplementary Table 3](#). Additionally, incomplete and variable AE reporting across trials, including differences in the reporting of AEs, treatment-related AEs, and treatment-emergent AEs, may have affected the accuracy of NNH estimates. Overall, these limitations reflect the heterogeneity of the available evidence and could not be fully harmonized across studies.

5. Conclusions

Our analysis highlights the value of NNT and NNH, as intuitive metric to characterize the benefit–risk balance of therapeutic escalation strategies in patients with intermediate-to high-risk eBC. By translating complex statistical results into clinically interpretable estimates, these measures may enhance both scientific communication and evidence-based decision-making. However, their interpretation requires careful consideration of trial design, patient population, follow-up duration, and the temporal dynamics of treatment efficacy and toxicity.

CI must always accompany these estimates to convey statistical precision and to mitigate the risk of misleading conclusions. Importantly, while NNT and NNH, provide a useful quantitative framework, they should complement, not replace, clinical judgment, integration of biological risk factors, and patient preferences.

In this regard, future studies incorporating patient partners or patient advocacy groups could help clarify how patients interpret and value these summary measures when weighing treatment benefit against potential harm. When applied responsibly, these measures can enhance shared decision-making and contextualize the trade-offs between therapeutic benefit and harm. Conversely, their presentation without appropriate contextualization risks oversimplifying the evidence and misrepresenting the true clinical implications of treatment.

Ethical approval

This study did not require ethical approval in accordance with the journal's policy and applicable regulations.

Funding source

This work was supported by the Italian Ministry of Health (Ricerca Corrente).

CRediT authorship contribution statement

Roberto Buonaiuto: Conceptualization, Data curation, Formal analysis, Methodology, Resources, Writing – original draft, Writing – review & editing. **Federica Pia Mangiacotti:** Data curation, Writing – original draft, Writing – review & editing. **Fabiola Giudici:** Data curation, Formal analysis, Methodology, Writing – original draft. **Margherita Tafuro:** Data curation, Writing – original draft, Writing – review & editing. **Roberta Scarfetta:** Conceptualization, Data curation, Writing – original draft, Writing – review & editing. **Alessandra Longobardi:** Writing – review & editing. **Martina Pagliuca:** Writing – review & editing. **Aldo Caltavuturo:** Writing – review & editing. **Luca Arecco:** Writing – review & editing. **Evandro de Azambuja:** Writing – review & editing. **Mario Giuliano:** Writing – review & editing. **Grazia Arpino:** Writing – review & editing. **Lucia Del Mastro:** Writing – review & editing. **Carmine De Angelis:** Methodology, Supervision, Writing – review & editing. **Michelino De Laurentiis:** Methodology, Supervision, Writing – review & editing. **Fabio Puglisi:** Data curation, Formal analysis, Methodology, Supervision, Validation, Visualization, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: **FP** reports research grants from AstraZeneca, Eisai, and Roche. **FP** has received consulting fees from AstraZeneca, Daiichi Sankyo, Eli Lilly, Menarini, MSD, Novartis, Pfizer, and Roche; and payments/honoraria from AstraZeneca, Daiichi Sankyo, Eli Lilly, Exact Sciences, Gilead, Italfarmaco, Menarini, MSD, Novartis, Pfizer, and Roche. **FP** has served in advisory roles for Amgen, Daiichi Sankyo, Eli Lilly, Gilead, GSK, Novartis, Pfizer, and Roche. **FP** has received travel support and accommodation from AstraZeneca, Daiichi Sankyo, Menarini, Pfizer, and Roche, and has participated in advisory boards for AstraZeneca, Daiichi Sankyo, Eli Lilly, Exact Sciences, Menarini, MSD, Novartis, Pfizer, and Roche. **MP** reports institutional funding from Gilead and travel reimbursement from Ipsen. **LDM** reports grants from Eli Lilly, Novartis, Roche, Daiichi Sankyo, Seagen, AstraZeneca, Gilead, Pierre Fabre, Pfizer, Menarini Stemline, and MSD. **LDM** has received consulting fees from Novartis, Roche, Daiichi Sankyo, Eli Lilly, Gilead, Pfizer, Menarini Stemline, Olema, and AstraZeneca; and honoraria from Eli Lilly, Novartis, Roche, Daiichi Sankyo, Seagen, AstraZeneca, Gilead, Pierre

Fabre, Pfizer, Menarini Stemline, MSD, Exact Sciences, GSK, Ipsen, and Eisai. LDM has received travel grants from Roche, Gilead, Menarini Stemline, AstraZeneca, and Daiichi Sankyo. LDM has served on advisory boards for Eli Lilly, Novartis, Roche, Daiichi Sankyo, Seagen, AstraZeneca, Gilead, Pierre Fabre, Pfizer, Menarini Stemline, MSD, Exact Sciences, GSK, Ipsen, and Eisai. LDM reports receipt of equipment from FoundationOne. **MDL** reports relationships with Roche, Novartis, Takeda, Eli Lilly, Pierre Fabre, AstraZeneca, MSD, Seagen, Gilead, Daiichi Sankyo, Tomalab, Genetic, Pfizer, Menarini, Sophos, Istituto Gentili, Sanofi, Ipsen, GSK, and Exact Sciences, including consulting/advisory roles and travel reimbursement. **MG** reports relationships with AstraZeneca, Daiichi Sankyo, Eisai, Gilead, Celgene, Exact Sciences, Eli Lilly, MSD, Novartis, Pfizer, Roche, and Seagen, including consulting/advisory roles and travel reimbursement. **GA** reports relationships with Roche, Pfizer, Eli Lilly, MSD, AstraZeneca, and Novartis, including consulting/advisory roles and research funding grants. **PV** has received honoraria and served in advisory roles for AstraZeneca, Daiichi Sankyo, Eli Lilly, Gilead, Incyte, Istituto Gentili, Novartis, Pfizer, Roche, Seagen, and Teva. **PV** has received travel support and accommodation from Daiichi Sankyo, Eli Lilly, and Novartis. **PV** reports research funding to the institution from Novartis and AstraZeneca. **CDA** reports consulting or advisory relationships with Novartis, GSK, Eli Lilly, and Pfizer. **EDA** reports honoraria and/or advisory board participation from Roche/GNE, Novartis, Seagen, Zodiac, Libbs, Pierre Fabre, Lilly, AstraZeneca, MSD, and Gilead Sciences; travel grants from AstraZeneca and Gilead; institutional research grants from Roche/GNE, AstraZeneca, GSK/Novartis, Gilead Sciences, and Seagen/Pfizer; and non-financial roles as ESMO Director of Membership (2023–2025) and BSMO President (2023–2026). **LA** reports travel grants from AstraZeneca and institutional research funding from Gilead Sciences. All other authors declare no competing financial interests or personal relationships that could have influenced the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.breast.2026.104820>.

References

- Wang J, Wu S-G. Breast cancer: an overview of current therapeutic strategies, challenge, and perspectives. *Breast Cancer* 2023;15:721–30. <https://doi.org/10.2147/BCTT.S432526>.
- Bekes I, Huober J. Extended adjuvant endocrine therapy in early breast cancer patients—review and perspectives. *Cancers (Basel)* 2023;15:4190. <https://doi.org/10.3390/cancers15164190>.
- Bourgier C, Ozsahin M, Azria D. Multidisciplinary approach of early breast cancer: the biology applied to radiation oncology. *Radiat Oncol* 2010;5:2. <https://doi.org/10.1186/1748-717X-5-2>.
- Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995;310:452–4. <https://doi.org/10.1136/bmj.310.6977.452>.
- Gensini GF, Conti AA. Evidence-based evaluation of benefits in therapeutic interventions: methodologically controlled and non-randomly assigned reflections on the number needed to treat. *Ital Heart J* 2003;4:80–3.
- Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med* 1988;318:1728–33. <https://doi.org/10.1056/NEJM198806303182605>.
- Bronte G. The likelihood of being helped or harmed obtained from clinical trial results for cancer therapy: can it really help? *Semin Oncol* 2025;52:7–9. <https://doi.org/10.1053/j.seminoncol.2024.12.001>.
- Loibl S, André F, Bachelot T, Barrios CH, Bergh J, Burstein HJ, et al. Early breast cancer: ESMO clinical practice guideline for diagnosis, treatment and follow-up. *Ann Oncol* 2024;35:159–82. <https://doi.org/10.1016/j.annonc.2023.11.016>.
- Gradishar, W. J., Moran, M. S., Abraham, J., et al. (2025). Breast cancer, version 5.2025 featured updates to the NCCN guidelines. *J Natl Compr Cancer Netw*, 23 (11), 426–436. doi: 10.6004/jnccn.2025.0053. n.d.
- van der Voort A, van Ramshorst MS, van Werkhoven ED, Mandjes IA, Kemper I, Vulink AJ, et al. Three-year Follow-up of neoadjuvant chemotherapy with or without anthracyclines in the presence of dual ERBB2 blockade in patients with ERBB2-Positive breast cancer. *JAMA Oncol* 2021;7:978. <https://doi.org/10.1001/jamaoncol.2021.1371>.
- Masuda N, Lee S-J, Ohtani S, Im Y-H, Lee E-S, Yokota I, et al. Adjuvant capecitabine for breast cancer after preoperative chemotherapy. *N Engl J Med* 2017;376:2147–59. <https://doi.org/10.1056/NEJMoa1612645>.
- Liu N, Zhou Y, Lee JJ. IPDfromKM: reconstruct individual patient data from published kaplan-meier survival curves. *BMC Med Res Methodol* 2021;21:111. <https://doi.org/10.1186/s12874-021-01308-8>.
- Andrade C. The numbers needed to treat and harm (NNT, NNH) statistics: what they tell us and what they Do not. *J Clin Psychiatry* 2015;76:e330–3. <https://doi.org/10.4088/JCP.15f09870>.
- Altman DG, Andersen PK. Calculating the number needed to treat for trials where the outcome is time to an event. *BMJ* 1999;319:1492–5. <https://doi.org/10.1136/bmj.319.7223.1492>.
- Wen L, Badgett R, Cornell J. Number needed to treat: a descriptor for weighing therapeutic options. *Am J Health Syst Pharm* 2005;62:2031–6. <https://doi.org/10.2146/ajhp040558>.
- Stang A, Poole C, Bender R. Common problems related to the use of number needed to treat. *J Clin Epidemiol* 2010;63:820–5. <https://doi.org/10.1016/j.jclinepi.2009.08.006>.
- McAlister FA. The “number needed to treat” turns 20 – and continues to be used and misused. *Can Med Assoc J* 2008;179:549–53. <https://doi.org/10.1503/cmaj.080484>.
- Altman DG. Confidence intervals for the number needed to treat. *BMJ* 1998;317:1309–12. <https://doi.org/10.1136/bmj.317.7168.1309>.
- Geyer CE, Untch M, Huang C-S, Mano MS, Mamounas EP, Wolmark N, et al. Survival with trastuzumab emtansine in residual HER2-Positive breast cancer. *N Engl J Med* 2025;392:249–57. <https://doi.org/10.1056/NEJMoa2406070>.
- Loibl S, Jassem J, Sonnenblick A, Parlier D, Winer E, Bergh J, et al. Adjuvant pertuzumab and trastuzumab in early human epidermal growth factor receptor 2-Positive breast cancer in the APHINITY trial: third interim overall survival analysis with efficacy update. *J Clin Oncol* 2024;42:3643–51. <https://doi.org/10.1200/JCO.23.02505>.
- Loibl S, Piccart M, Clark E, Viale G, Caballero C, Henry C, et al. LBA1 adjuvant pertuzumab or placebo + trastuzumab + chemotherapy (P or Pla + T + CT) in patients (pts) with early HER2-positive operable breast cancer in APHINITY: final analysis at 11.3 years' median follow-up. *ESMO Open* 2025;10:105112. <https://doi.org/10.1016/j.esmoop.2025.105112>.
- Gianni L, Pienkowski T, Im Y-H, Tseng L-M, Liu M-C, Lluch A, et al. 5-year analysis of neoadjuvant pertuzumab and trastuzumab in patients with locally advanced, inflammatory, or early-stage HER2-positive breast cancer (NeoSphere): a multicentre, open-label, phase 2 randomised trial. *Lancet Oncol* 2016;17:791–800. [https://doi.org/10.1016/S1470-2045\(16\)00163-7](https://doi.org/10.1016/S1470-2045(16)00163-7).
- Chan A, Moy B, Mansi J, Ejlersen B, Holmes FA, Chia S, et al. Final efficacy results of neratinib in HER2-positive hormone receptor-positive early-stage breast cancer from the phase III ExteNET trial. *Clin Breast Cancer* 2021;21:80–91.e7. <https://doi.org/10.1016/j.clbc.2020.09.014>.
- Johnston S, Martin M, O'Shaughnessy J, Hegg R, Tolane SM, Guarneri V, et al. Overall survival with abemaciclib in early breast cancer. *Ann Oncol* 2025. <https://doi.org/10.1016/j.annonc.2025.10.005>.
- Crown J, Stroyakovskii D, Yardley DA, Huang C-S, Fasching PA, Bardia A, et al. Adjuvant ribociclib plus nonsteroidal aromatase inhibitor therapy in patients with HR-positive/HER2-negative early breast cancer: 5-Year follow-up of NATALEE efficacy outcomes and updated overall survival. *ESMO Open* 2025;10:105858. <https://doi.org/10.1016/j.esmoop.2025.105858>.
- Garber J, Cameron D, Campbell C, Yothers G, Taboada M, El-Abed S, et al. Abstract GS1-09: ol ympia: a phase 3, multicenter, randomized, placebo-controlled trial of adjuvant olaparib after (Neo)adjuvant chemotherapy in patients w/germline BRCA1 & BRCA2 pathogenic variants & highrisk HER2-negative primary breast cancer. *Clin Cancer Res* 2025;31:GS1-09. <https://doi.org/10.1158/1557-3265.SABCS24-GS1-09>.
- Schmid P, Cortés J, Dent RA, McArthur HL, Pusztai L, Kummel S, et al. LBA4 neoadjuvant pembrolizumab or placebo plus chemotherapy followed by adjuvant pembrolizumab or placebo for high-risk early-stage TNBC: overall survival results from the phase III KEYNOTE-522 study. *Ann Oncol* 2024;35:S1204–5. <https://doi.org/10.1016/j.annonc.2024.08.2247>.
- Gradishar WJ, Moran MS, Abraham J, Abramson V, Aft R, Agnese D, et al. Breast cancer, version 3.2024, NCCN Clinical Practice guidelines in oncology. *J Natl Compr Cancer Netw* 2024;22:331–57. <https://doi.org/10.6004/jnccn.2024.0035>.
- EMA approval, 12.09.2018; Reference Number:EMA/455200/2018 n.d.
- Howie LJ, Scher NS, Amiri-Kordestani L, Zhang L, King-Kallimanis BL, Choudhry Y, et al. FDA approval summary: pertuzumab for adjuvant treatment of HER2-Positive early breast cancer. *Clin Cancer Res* 2019;25:2949–55. <https://doi.org/10.1158/1078-0432.CCR-18-3003>.
- Rastogi P, O'Shaughnessy J, Martin M, Boyle F, Cortes J, Rugo HS, et al. Adjuvant abemaciclib plus endocrine therapy for hormone receptor-positive, human epidermal growth factor receptor 2-Negative, high-risk early breast cancer: results from a preplanned monarchE overall survival interim analysis, including 5-Year efficacy outcomes. *J Clin Oncol* 2024;42:987–93. <https://doi.org/10.1200/JCO.23.01994>.
- von Minckwitz G, Procter M, de Azambuja E, Zardavas D, Benyunes M, Viale G, et al. Adjuvant pertuzumab and trastuzumab in early HER2-Positive breast cancer. *N Engl J Med* 2017;377:122–31. <https://doi.org/10.1056/NEJMoa1703643>.
- Piccart M, Procter M, Fumagalli D, de Azambuja E, Clark E, Ewer MS, et al. Adjuvant pertuzumab and trastuzumab in early HER2-Positive breast cancer in the APHINITY trial: 6 years' Follow-Up. *J Clin Oncol* 2021;39:1448–57. <https://doi.org/10.1200/JCO.20.01204>.

- [34] Johnston SRD, Toi M, O'Shaughnessy J, Rastogi P, Campone M, Neven P, et al. Abemaciclib plus endocrine therapy for hormone receptor-positive, HER2-negative, node-positive, high-risk early breast cancer (monarchE): results from a preplanned interim analysis of a randomised, open-label, phase 3 trial. *Lancet Oncol* 2023;24:77–90. [https://doi.org/10.1016/S1470-2045\(22\)00694-5](https://doi.org/10.1016/S1470-2045(22)00694-5).
- [35] Johnston SRD, Harbeck N, Hegg R, Toi M, Martin M, Shao ZM, et al. Abemaciclib combined with endocrine therapy for the adjuvant treatment of HR+, HER2-, Node-Positive, High-risk, early breast cancer (monarchE). *J Clin Oncol* 2020;38:3987–98. <https://doi.org/10.1200/JCO.20.02514>.
- [36] Fasching PA, Stroyakovskiy D, Yardley DA, Huang C-S, Crown J, Bardia A, et al. Ribociclib plus endocrine therapy in hormone Receptor-Positive/ERBB2-Negative early breast cancer. *JAMA Oncol* 2025;11:1364. <https://doi.org/10.1001/jamaoncol.2025.3700>.
- [37] Hortobagyi GN, Lacko A, Sohn J, Cruz F, Ruiz Borrego M, Manikhas A, et al. A phase III trial of adjuvant ribociclib plus endocrine therapy versus endocrine therapy alone in patients with HR-positive/HER2-negative early breast cancer: final invasive disease-free survival results from the NATALEE trial. *Ann Oncol* 2025;36:149–57. <https://doi.org/10.1016/j.annonc.2024.10.015>.
- [38] von Minckwitz G, Huang C-S, Mano MS, Loibl S, Mamounas EP, Untch M, et al. Trastuzumab emtansine for residual invasive HER2-Positive breast cancer. *N Engl J Med* 2019;380:617–28. <https://doi.org/10.1056/NEJMoa1814017>.
- [39] Wurcel V, Rojas Rojas M, Urrego-Reyes J, Medrano Rivera D, Acevedo R, Jiang R, et al. Number needed to treat (NNT) with pembrolizumab as an adjuvant therapy in resected patients with high-risk stage II (IIB and IIC) melanoma and its application to cost of preventing an event (COPE) in Mexico. *J Med Econ* 2025;28:346–53. <https://doi.org/10.1080/13696998.2025.2466365>.
- [40] Chanan-Khan A, Hanna K, Xue M, Massoudi M, Balk M, Gani R, et al. Number needed to treat and associated cost analysis of zanubrutinib vs ibrutinib in chronic lymphocytic leukemia. *J Manag Care Spec Pharm* 2025;31:482–90. <https://doi.org/10.18553/jmcp.2025.24330>.
- [41] Temple R. Reporting number needed to treat requires standardization. *Am Fam Physician* 2026;113:110–2.
- [42] van Maaren MC, de Munck L, Strobbe LJA, Sonke GS, Westenend PJ, Smidt ML, et al. Ten-year recurrence rates for breast cancer subtypes in the Netherlands: a large population-based study. *Int J Cancer* 2019;144:263–72. <https://doi.org/10.1002/ijc.31914>.
- [43] Chumsri S, Thompson EA. Carryover effects of aromatase inhibitors in prevention. *Lancet* 2020;395:91–2. [https://doi.org/10.1016/S0140-6736\(19\)33102-2](https://doi.org/10.1016/S0140-6736(19)33102-2).
- [44] Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. *Lancet* 2015;386:1341–52. [https://doi.org/10.1016/S0140-6736\(15\)61074-1](https://doi.org/10.1016/S0140-6736(15)61074-1).
- [45] Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;134:663–94. <https://doi.org/10.7326/0003-4819-134-8-200104170-00012>.
- [46] Hildebrandt M, Vervölgyi E, Bender R. Calculation of NNTs in RCTs with time-to-event outcomes: a literature review. *BMC Med Res Methodol* 2009;9:21. <https://doi.org/10.1186/1471-2288-9-21>.
- [47] Murad MH, Falck-Ytter Y, Ramachandran N, Davitkov P, Morgan RL. The number needed to treat: it is time to bow out gracefully. *J Clin Epidemiol* 2025;187:111924. <https://doi.org/10.1016/j.jclinepi.2025.111924>.
- [48] Stiefel F, Bourquin C, Salmon P, Ahtari Jeanneret L, Dauchy S, Ernstmann N, et al. Communication and support of patients and caregivers in chronic cancer care: ESMO clinical practice guideline. *ESMO Open* 2024;9:103496. <https://doi.org/10.1016/j.esmoop.2024.103496>.
- [49] van de Water LF, van Kleef JJ, Dijksterhuis WPM, Henselmans I, van den Boorn HG, Vaarzon Morel NM, et al. Communicating treatment risks and benefits to cancer patients: a systematic review of communication methods. *Qual Life Res* 2020;29:1747–66. <https://doi.org/10.1007/s11136-020-02503-8>.
- [50] Westendorp J, Geerse OP, van der Lee ML, Schoones JW, van Vliet MHM, Wit T, et al. Harmful communication behaviors in cancer care: a systematic review of patients and family caregivers perspectives. *Psychooncology* 2023;32:1827–38. <https://doi.org/10.1002/pon.6247>.