Article

Early versus deferred use of CDK4/6 inhibitors in advanced breast cancer

https://doi.org/10.1038/s41586-024-08035-2

Received: 9 January 2024

Accepted: 10 September 2024

Published online: 27 November 2024

Check for updates

Gabe S. Sonke[1,21]✉, Annemiek van Ommen-Nijhof[1,20], Noor Wortelboer[2,20], Vincent van der Noort[3], Astrid C. P. Swinkels[4], Hedwig M. Blommestein[5], Cristina Guerrero Paez[6], Linda Mol[4], Aart Beeker[7], Karin Beelen[8], Lisanne C. Hamming[9], Joan B. Heijns[10], Aafke H. Honkoop[11], Paul C. de Jong[12], Quirine C. van Rossum-Schornagel[13], Christa van Schaik-van de Mheen[14], Jolien Tol[15], Cathrien S. Tromp-van Driel[16], Suzan Vrijaldenhoven[17], A. Elise van Leeuwen-Stok[18], Inge R. Konings[19,21], Agnes Jager[2,21] & SONIA Study Consortium*

Cyclin-dependent kinase 4 and 6 inhibitors (CDK4/6i) in combination with endocrine therapy improve the outcomes of patients with hormone-receptor (HR)-positive, HER2-negative advanced breast cancer and can be used early as first-line treatment or deferred to second-line treatment[1–7]. Randomized data comparing the use of CDK4/6i in the first- and second-line setting are lacking. The phase 3 SONIA trial (NCT03425838) randomized 1,050 patients who had not received previous therapy for advanced breast cancer to receive CDK4/6i in the first- or second-line setting[8]. All of the patients received the same endocrine therapy, consisting of an aromatase inhibitor for first-line treatment and fulvestrant for second-line treatment. The primary end point was defined as the time from randomization to disease progression after second-line treatment (progression-free survival 2 (PFS2)). We observed no statistically significant benefit for the use of CDK4/6i as a first-line compared with second-line treatment (median, 31.0 versus 26.8 months, respectively; hazard ratio = 0.87; 95% confidence interval = 0.74–1.03; $P = 0.10$). The health-related quality of life was similar in both groups. First-line CDK4/6i use was associated with a longer CDK4/6i treatment duration compared with second-line use (median CDK4/6i treatment duration of 24.6 versus 8.1 months, respectively) and more grade ≥3 adverse events (2,763 versus 1,591, respectively). These data challenge the need for first-line use of a CDK4/6i in all patients.

Treatment with CDK4/6i has demonstrated significant improvements in the outcome of patients with HR^+HER2^- advanced breast cancer (ABC). PFS benefit in the pivotal phase 3 trials was observed in the first- and second-line setting and was highly similar with three different CDK4/6 inhibitors (palbociclib, ribociclib and abemaciclib)[1–7]. Abemaciclib demonstrated an overall survival (OS) benefit of 9.4 months in the second-line setting[9], and ribociclib in both the first-line (12.5 months) and second-line (7.7 months) settings[10–12]. On the basis of the landmark trials, there are two viable strategies for these patients: adding CDK4/6i early to first-line endocrine treatment (that is, an aromatase inhibitor (AI)) or deferred to second-line treatment (that is, fulvestrant).

Absolute PFS improvements with the use of CDK4/6i in first-line studies exceeded those in second-line studies and most guidelines therefore recommend first-line use as the preferred strategy[13]. However, this strategy also comes with added toxicity and cost as the treatment duration is longer for first-line compared with second-line use[1–7]. Moreover, many patients have lasting disease remission of many years with an AI alone in the first-line setting[14]. Conversely, there may be concerns about the retention of endocrine sensitivity in second-line treatment after first-line exposure to CDK4/6i treatment, although this has not been formally demonstrated[15,16]. Finally, when testing the efficacy of a cancer drug in an earlier line of therapy, it is important to build crossover into the trial design if the drug has already been proven to be effective in a later line of therapy[17]. However, only a minority of patients in the control arm of first-line studies crossed-over to receive CDK4/6i in later lines.

[1]Department of Medical Oncology, Netherlands Cancer Institute, Amsterdam, The Netherlands. [2]Department of Medical Oncology, Erasmus Medical Center Cancer Institute, Rotterdam, The Netherlands. [3]Department of Biometrics, Netherlands Cancer Institute, Amsterdam, The Netherlands. [4]Clinical Research Department, Netherlands Comprehensive Cancer Organisation, Utrecht, The Netherlands. [5]Erasmus School of Health Policy & Management, Erasmus University Rotterdam, Rotterdam, The Netherlands. [6]Dutch Breast Cancer Society, Utrecht, The Netherlands. [7]Department of Medical Oncology, Spaarne Gasthuis, Hoofddorp, The Netherlands. [8]Department of Medical Oncology, Rijnstate Hospital, Arnhem, The Netherlands. [9]Department of Medical Oncology, Medical Center Leeuwarden, Leeuwarden, The Netherlands. [10]Department of Medical Oncology, Amphia, Breda, The Netherlands. [11]Department of Medical Oncology, Isala Clinics, Zwolle, The Netherlands. [12]Department of Medical Oncology, St Antonius Hospital, Nieuwegein, The Netherlands. [13]Department of Medical Oncology, Franciscus Gasthuis & Vlietland, Schiedam, The Netherlands. [14]Department of Medical Oncology, Meander Medical Center, Amersfoort, The Netherlands. [15]Department of Medical Oncology, Jeroen Bosch Hospital, 's-Hertogenbosch, The Netherlands. [16]Department of Medical Oncology, Gelre Hospital, Apeldoorn, The Netherlands. [17]Department of Medical Oncology, Northwest Hospital Group, Alkmaar, The Netherlands. [18]Dutch Breast Cancer Research Group, BOOG Study Center, Utrecht, The Netherlands. [19]Department of Medical Oncology, Amsterdam University Medical Center, Cancer Center Amsterdam, Amsterdam, The Netherlands. [20]These authors contributed equally: Annemiek van Ommen-Nijhof, Noor Wortelboer. [21]These authors jointly supervised this work: Gabe S. Sonke, Inge R. Konings, Agnes Jager. *A full list of members and their affiliations appears in the Supplementary Information. ✉e-mail: g.sonke@nki.nl

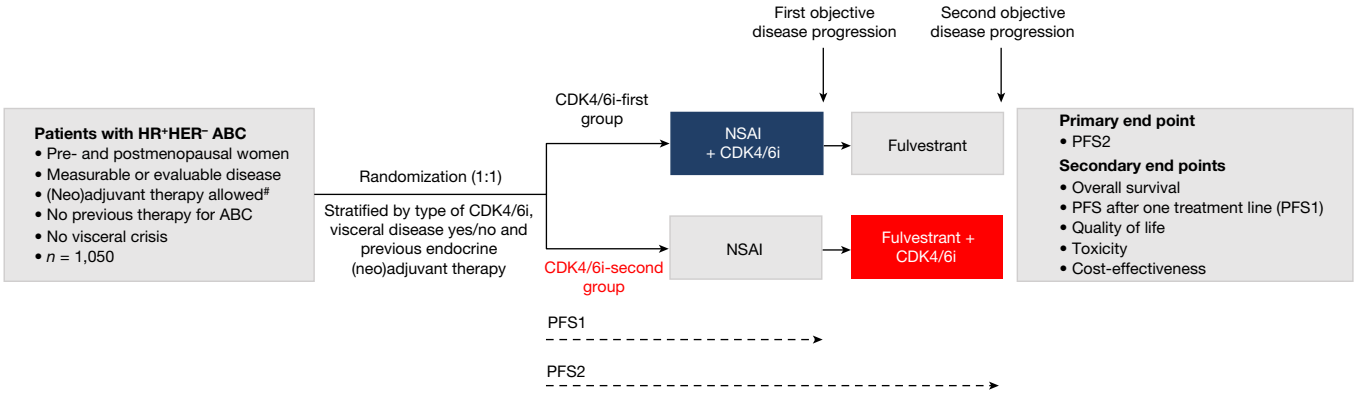


Fig. 1 | Trial design. Imaging occurred every 12 weeks until cession of the two protocol-defined treatment lines. The switch to second-line treatment was indicated at first objective disease progression. #The disease-free interval after NSAI > 12 months. NSAI, non-steroidal AI.

Given the absence of randomized data comparing first- and second-line use, the lack of predictive biomarkers and inspired by the commitment of the American Society of Clinical Oncology (ASCO) and the European Society for Medical Oncology (ESMO) to promote equitable and sustainable cancer care, we designed the investigator-initiated SONIA trial (Fig. 1) to determine the optimal strategy for using CDK4/6i in clinical practice[18,19]. SONIA received no commercial funding.

Baseline characteristics

Between 23 November 2017 and 1 September 2021, 1,050 patients were randomized to either the CDK4/6i-first group (n = 524) or the CDK4/6i-second group (n = 526) (Extended Data Fig. 1). Baseline characteristics were well balanced across both study groups (Table 1). The median age was 64 years, 583 (56%) patients presented with visceral disease and 182 (17%) patients with bone-only disease. Approximately half of the study population (512 patients) had received previous (neo)adjuvant endocrine therapy, of whom 327 patients (64%) had disease progression ≥24 months since the end of endocrine therapy. A total of 40% of patients (n = 422) had received previous (neo)adjuvant chemotherapy. In total, 35% of patients (n = 364) presented with de novo metastatic disease. Most patients (n = 958, 91%) were intended to receive palbociclib as initial CDK4/6i at the time of inclusion in the study, due to the later reimbursement of ribociclib and abemaciclib in The Netherlands.

Trial overview

At the data cut-off date (1 December 2022), 329 patients were still on first-line treatment and 106 patients were on second-line treatment. A total of 717 PFS1 events (defined as the time from randomization to first progression, initiation of chemotherapy or death, whichever came first) had occurred, 310 in the CDK4/6i-first group and 407 in the CDK4/6i-second group. In total, 719 patients discontinued first-line treatment, 315 in the CDK4/6i-first group and 404 in the CDK4/6i second group. Most of these patients (295 patients (94%) in the CDK4/6i-first and 383 (95%) in the CDK4/6i-second group) discontinued first-line treatment due to progressive disease and usually switched to the protocol-defined second-line treatment (246 out of 315 (78%) patients in the CDK4/6i-first group and 349 out of 404 (86%) in the CDK4/6i-second group). In total, 31 patients in the CDK4/6i-first group and 21 in the CDK4/6i-second group started chemotherapy as second-line treatment, whereas 9 and 4, respectively, started off-protocol endocrine therapy. In the CDK4/6i-first and CDK4/6i-second group, 29 and 30 patients who discontinued first-line treatment did not start any second-line therapy at the data cut-off date, respectively, mostly due to death, consent withdrawal and poor performance status. During the study period, six patients withdrew informed consent (four in the CDK4/6i-first group and two in the CDK4/6i-second group) and seven patients were rendered not evaluable and were therefore censored at the time of diagnosis of a second malignancy that impacted treatment or prognosis. We observed 281 PFS2 events in the CDK4/6i-first group and 310 in the CDK4/6i-second group. The most common reason for discontinuation of second-line endocrine treatment was progressive disease in both groups (in 226 out of the 235 patients (96%) in the CDK4/6i-first group and in 245 out of the 267 patients (92%) in the CDK4/6i-second group). The median and restricted mean CDK4/6i treatment durations in the CDK4/6i-first group were 24.6 months (95% confidence interval (CI) = 21.9–27.8) and 28.5 months (95% CI = 26.7–30.4), respectively. In the CDK4/6i-second group, the median and restricted mean CDK4/6i treatment durations were 8.1 months (95% CI = 7.7–10.0) and 12.4 months (95% CI = 11.0–13.9), respectively.

Primary outcome

After a median follow-up of 37.3 months (95% CI = 35.7–38.3), median PFS2 was 31.0 months in the CDK4/6i-first group (95% CI = 26.4–34.5) versus 26.8 months (95% CI = 24.1–30.4) in the CDK4/6i-second group (stratified hazard ratio = 0.87, 95% CI = 0.74–1.03, P = 0.10; Fig. 2). The H1 hypothesis of superiority of first-line over second-line use cannot be established based on these numbers. As superiority could not be established, non-inferiority was tested next as defined in the protocol. On the basis of the upper limit of the 95% CI of 1/0.74 (1.35), non-inferiority of the CDK4/6i-second line strategy over the CDK4/6i-first line strategy was established based on the ESMO–MCBS margin of 1/0.65 (1.54).

PFS2 results were consistent across predefined subgroups, with no test for interaction reaching statistical significance. Figure 3 shows the effect of the treatment strategy on PFS2 among these predefined as well as several exploratory post hoc subgroups. The preplanned sensitivity analyses were all consistent with the primary outcome analysis (Extended Data Table 1).

Secondary efficacy outcomes

At the time of data cut-off with a median follow-up of 37.3 months and 56% of patients with at least 36 months of follow-up, a total of 372 deaths (35%) had occurred—184 out of the 524 patients in the CDK4/6i-first group and 188 out of the 526 patients in the CDK4/6i-second group. Exploratory analysis showed a median OS of 45.9 months (95% CI = 42.2–not reached) and 53.7 months (95% CI = 44.7–not reached), respectively (hazard ratio = 0.98; 95% CI = 0.80–1.20; P = 0.83) (Fig. 4). The median PFS1 in the CDK4/6i-first- and CDK4/6i-second groups was

Table 1 | Patient and baseline characteristics

Characteristic	CDK4/6i-first group *n*=524	CDK4/6i-second group *n*=526
Median age, years (interquartile range)	64 (56–71)	63 (54–71)
ECOG performance status		
0	257 (49)	257 (49)
≥1	267 (51)	269 (51)
Hormone receptor status[a]		
ER^+	523 (100)	525 (100)
PR^+	350 (67)	353 (67)
Menopausal status		
Premenopausal/perimenopausal	69 (13)	76 (14)
Postmenopausal	455 (87)	450 (86)
Disease setting[b]		
De novo metastatic	182 (35)	182 (35)
Metastatic recurrent	342 (65)	343 (65)
Locoregionally recurrent (not curable)	0 (0)	1 (0)
Treatment-free interval[c]		
≤24 months	94 (36)	91 (36)
>24 months	164 (64)	163 (64)
Previous (neo)adjuvant therapy		
Chemotherapy	212 (40)	210 (40)
Endocrine therapy	258 (49)	254 (48)
Tamoxifen only	105 (20)	104 (20)
AI only	32 (6)	29 (6)
Sequential	121 (23)	120 (23)
Unknown	0 (0)	1 (0)
Metastatic site		
Visceral disease	291 (56)	292 (56)
Bone only disease	91 (17)	91 (17)
Measurable disease	315 (60)	312 (59)
Type of CDK4/6i[d]		
Palbociclib	479 (91)	479 (91)
Ribociclib	42 (8)	44 (8)
Abemaciclib	3 (1)	3 (1)

Data are numbers (%), unless otherwise specified. ECOG, Eastern Cooperative Oncology Group.
[a]ER status was available for all patients, PR status was missing for six patients in the CDK4/6i-first group and eight in the CDK4/6i-second group.
[b]De novo metastatic disease (referred to as synchronous disease) is defined as the presence of metastatic disease at the time of initial diagnosis (within 6 months).
[c]Calculated for patients with previous endocrine therapy only, defined as the time from the end of adjuvant endocrine therapy until diagnosis of metastatic disease.
[d]Type of CDK4/6i intended to prescribe at the time of randomization.

24.7 (95% CI = 22.1–28.3) and 16.1 (95% CI = 14.4–17.1) months, respectively (hazard ratio = 0.59; 95% CI = 0.51–0.69; $P < 0.0001$) (Extended Data Fig. 2).

Health-related quality of life

All 921 patients (88% of the 1,050 randomized patients) who completed at least the baseline FACT-B questionnaire were included in the health-related quality of life (HRQOL) analyses. The baseline mean FACT-B total scores were similar for both groups (mean ± s.d., 105 ± 17 in the CDK4/6i-first group versus 103 ± 18 in the CDK4/6i-second group) and remained similar over the different timepoints thereafter (Extended Data Table 2). Completion rates remained high over time and no significant difference was observed in the FACT-B total score over time between the two groups (regression coefficient = −0.91 with the CDK4/6i-seond group as the reference; s.e. = 2.48; $P > 0.24$) (Extended Data Table 3).

Safety

The total number of grade ≥3 adverse events in the safety population was 2,763 in the CDK4/6i-first group and 1,591 in the CDK4/6i-second group (Table 2). In the CDK4/6i-first group, 433 patients (83%) experienced at least one grade ≥3 event, and 338 patients (64%) in the CDK4/6i-second group experienced at least one grade ≥3 event. The safety profile was characteristic for CDK4/6i, with no new safety issues observed (Extended Data Table 4). A total of 299 and 273 serious adverse events from any cause occurred in the CDK4/6i-first and CDK4/6i-second group, respectively. CDK4/6i dose reductions occurred in 225 (43%) patients in the CDK4/6i-first and 134 (39%) patients in the CDK4/6i-second group. Moreover, 47 (9%) patients in the CDK4/6i-first group and 20 (6%) patients in the CDK4/6i-second group discontinued CDK4/6i treatment while continuing endocrine therapy.

Drug costs

The total CDK4/6i drug costs (based on the Dutch list prices of 2023[20]) up to the data cut-off date were €24,890,620 in the CDK4/6i-first and €7,204,034 in the CDK4/6i-second group. The average CDK4/6i costs per patient were €47,959 in the CDK4/6i-first and €20,881 in the CDK4/6i-second group, resulting in additional CDK4/6i costs of €27,078 per patient with first-line use (Extended Data Table 5). Total and incremental drug costs were slightly higher using earlier reference prices (Dutch list prices of 2019) and substantially higher using US reference prices[21,22].

Subsequent therapy

Of the 502 patients who discontinued second-line endocrine therapy, 189 patients in the CDK4/6i-first group and 217 patients in the CDK4/6i-second group started subsequent treatment. Subsequent use of chemotherapy (86%), endocrine monotherapy (25%) and endocrine therapy plus targeted therapy (22%) was similar in both groups (Extended Data Table 6).

Discussion

SONIA is to our knowledge the first randomized trial directly comparing the use of CDK4/6i in first- versus second-line HR^+HER2^- ABC. We identified no statistically significant, nor clinically relevant PFS2 benefit according to the ESMO–MCBS for first-line compared with second-line use. Similarly, there was no difference in HRQOL during SONIA treatment between the two strategies. The median treatment duration with CDK4/6i was 16.6 months longer for first-line use than for second-line use and was associated with substantially more toxicity in terms of adverse events. No difference in serious adverse events was observed between the treatment groups. These results from SONIA challenge the general oncologic principle of always using the most effective drugs first and highlight the need for sequencing studies to generate evidence for using effective treatments in earlier lines of therapy.

When comparing SONIA to the first-line landmark trials, PFS1 results are very similar, with an approximate doubling of the median PFS1 with the addition of CDK4/6i[1–3]. Patients in SONIA were slightly older and had a shorter treatment-free interval, reflecting the real-world patient population enrolled in the trial. The key difference between SONIA and the first-line landmark trials is the protocol-defined crossover to second-line CDK4/6i for all patients who progress on endocrine

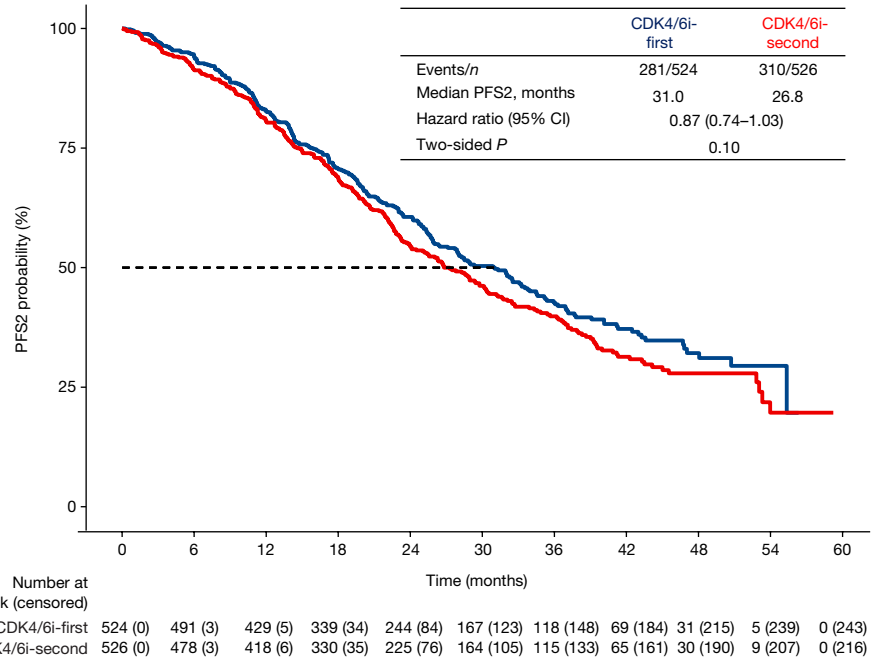


Fig. 2 | PFS2 in the ITT population. Kaplan–Meier plots for PFS after two treatment lines (PFS2) in the CDK4/6i-first and CDK4/6i-second group in the ITT population. Cox proportional hazard models were used to calculate hazard ratios between the study groups and were stratified according to the stratification factors used in randomization. The difference was assessed using the stratified log-rank tests; *P* values are two sided. Events, number of PFS2 events; *n*, number of patients randomized.

monotherapy. By contrast, only 12–34% of patients treated with endocrine monotherapy in the landmark trials received any CDK4/6i at progression[10,23,24]. Our trial results underscore the importance of trials that determine the optimal sequence of proven effective treatments, with mandatory crossover for all patients[16].

We did not find any difference in HRQOL over time between the two strategies. Our results showed that the addition of CDK4/6i to endocrine therapy neither improves nor reduces HRQOL, confirming previous findings in literature[25].

SONIA allowed use of each of the three approved CDK4/6 inhibitors according to availability and physician's choice. After successive reimbursement moments in the Netherlands, palbociclib was used the most, followed by ribociclib and abemaciclib. Although all three CDK4/6 inhibitors perform equally well in terms of PFS, recent OS data question the comparability of the three drugs[9–12,26,27]. However, it is important to distinguish between studies comparing different drugs and studies comparing different sequences as the two treatment lines in SONIA. We do not expect that a different distribution of type of CDK4/6i in our trial would have resulted in a different disease progression rate given the very similar PFS results with all three agents in first- and second-line in the landmark trials. Indeed, the subgroup analysis by type of CDK4/6i in our trial shows no evidence of a different effect for palbociclib or ribociclib. The numbers of individuals on abemaciclib were too small for comparison.

Since the start of the SONIA trial, new targeted therapies have emerged for patients with HR^+HER2^- ABC. Patients with a *PIK3CA* mutation and patients with activated AKT pathway alterations (both ±40%) can be candidates for fulvestrant plus alpelisib or capivasertib, respectively[28,29]. Phase 3 data on the use of these combinations after previous treatment with fulvestrant are lacking. Treatment beyond second progression in SONIA was similar between both arms of the trial and these new options did not affect the results of our trial. Developing combination treatments of targeted agents with different endocrine backbones (for example, AI + alpelisib, SERD + alpelisib, tamoxifen + alpelisib) would allow more flexibility in deciding the optimal treatment sequence and combination in individual patients[30,31].

PFS2 is a new and increasingly popular end point in oncology clinical trials. While its surrogacy for OS is not fully established, PFS2 spans a longer period of a patient's life and, as such, is potentially a better surrogate for OS than PFS[32]. The definition of PFS2 used in our study is based on the guidelines of the European Medicines Agency (EMA) and limits bias due to informative censoring[33,34]. Sensitivity analyses exploring the effect of different PFS2 definitions were all consistent with the primary analysis (Extended Data Table 1).

The SONIA trial has some limitations. First, the trial was designed as a superiority trial as there was equipoise regarding the effects of adding CDK4/6i in the first- and second-line context, given the absence of comparative data. As the primary analysis in our study did not show superiority of a first-line strategy over a second-line strategy, the protocol specified to also test for non-inferiority. This analysis demonstrated non-inferiority of the second-line strategy according to the prespecified non-inferiority margin of 1/0.65, but this margin may be considered to be too lenient to exclude a clinically relevant effect[35]. Conversely, very strict margins place a disproportionate burden of proof on alternative, less intensive treatments even when they provide very similar patient benefits compared to more intensive strategies, simply because the latter were developed first[36,37]. Clearly, there is a need to develop guidelines for balancing between too lenient and too strict non-inferiority margins. The EMA recognizes this knowledge gap and is currently working on a new guideline on this topic[38]. Second, the trial was open-label and we cannot rule out bias related to physician or patient awareness of the assigned treatment. We believe that the effect of this potential bias on the primary end point is limited, because all of the patients underwent protocol-defined tumour assessments. For the patients who did not undergo all protocol-defined tumour assessments, we performed sensitivity analyses to investigate the effect of such protocol deviations on the primary outcome. The results of these sensitivity analyses were all consistent with the intention-to-treat (ITT) analysis. It is also unlikely that the open-label design affected the OS results. Third, SONIA excluded patients with visceral crisis and the results can therefore not be extrapolated to these patients. Fourth, the fact that 85% of the SONIA study population was post-menopausal at

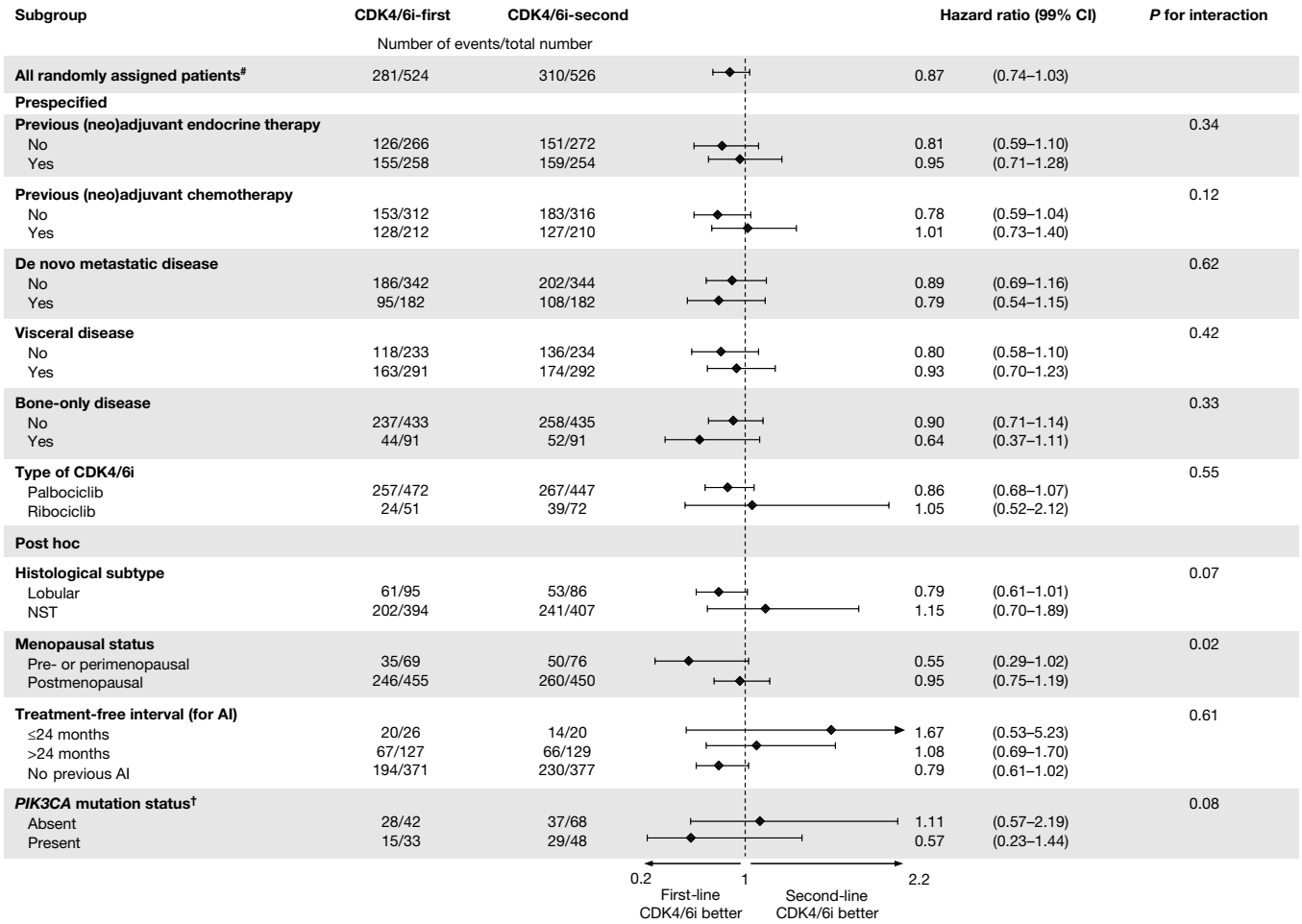


Fig. 3 | Subgroup analysis of PFS2 in the ITT population. The central points indicate hazard ratios for the primary end point PFS2 in various prespecified and post hoc subgroups in the ITT population. The horizontal lines indicate the 95% CI. #The hazard ratio of all randomly assigned patients is denoted with the 95% CI. Previous (neo)adjuvant endocrine therapy and chemotherapy include all therapy types in the specific categories. De novo metastatic disease is defined as the presence of metastatic disease at the time of initial diagnosis (within 6 months). Type of CDK4/6i refers to the CDK4/6i that the patient first received or the type of CDK4/6i that it was intended for the patient to receive in cases in which a patient did not receive a CDK4/6i. Invasive carcinoma 'no special type' (NST) includes large-cell carcinoma, adenocarcinoma and ductal carcinoma. Data from patients with a miscellaneous histological subtype are not reported owing to the small sample size. Treatment-free interval (for AI) is defined as the time from the end of adjuvant endocrine therapy with AI until diagnosis of metastatic disease. †Only patients for whom the mutation status could be retrieved through Palga[41] were included. The presence of a *PIK3CA* mutation was determined based on the 11 hotspot mutations in the SOLAR-1 study[28].

baseline and the majority used palbociclib may limit the generalizability of the SONIA results, and the validity for premenopausal women and non-palbociclib treatments is less certain. The results of the post hoc subgroup analysis according to menopausal status signify a potential effect of menopausal status on CDK4/6i timing, but these results should be interpreted with caution given the post hoc nature of the analysis and the small patient numbers. Fifth, SONIA included only patients in the Netherlands, and Dutch guidelines do not recommend endocrine therapy for all patients with stage I disease. This may have resulted in a relatively large number of treatment-naive, recurrent patients. As there is no indication that the primary outcome PFS2 differs between patients who did or did not receive previous adjuvant treatment, we do not believe that this has had an impact on the overall conclusions of the study. It is also uncertain whether these results are valid for tumours with 1–10% oestrogen receptor (ER)/progesterone receptor (PR) expression. Lastly, single-agent fulvestrant is no longer the preferred second-line therapy after CDK4/6i. Patients may also be candidates for other second-line therapies, such as fulvestrant in combination with alpelisib or capivasertib, PARP inhibition in the case of *BRCA1/2* mutation or capecitabine. Importantly, most of these other treatment options were available to patients in both arms of the SONIA trial after second progression, making a differential effect on OS unlikely, also given the fact that most of these therapies have not shown OS benefit[39,40]. Moreover, the limited efficacy of single-agent endocrine therapy in later lines of treatment highlights the importance of considering it as first-line option. Starting patients on combination therapy in the first-line context deprives them of the opportunity to benefit from this simple and often very effective single-agent treatment modality.

Biomarker research is of vital importance to predict responses to both targeted and non-targeted therapy, and many patients with ABC receive therapies that ultimately provide little to no benefit. Although the absolute benefit of CDK4/6i seems smaller for tumours with *ESR1* or *PIK3CA* mutations compared with their wild-type counterparts, there is currently no clinically validated biomarker beyond ER to predict benefit from CDK4/6i[18]. Given the results of the SONIA trial, research efforts in this area should focus on identifying patients who could benefit from early use of CDK4/6i rather than identifying patients who do not.

SONIA yields promise for both patients and societies, due to the shorter duration of CDK4/6i use in second-line treatment with similar

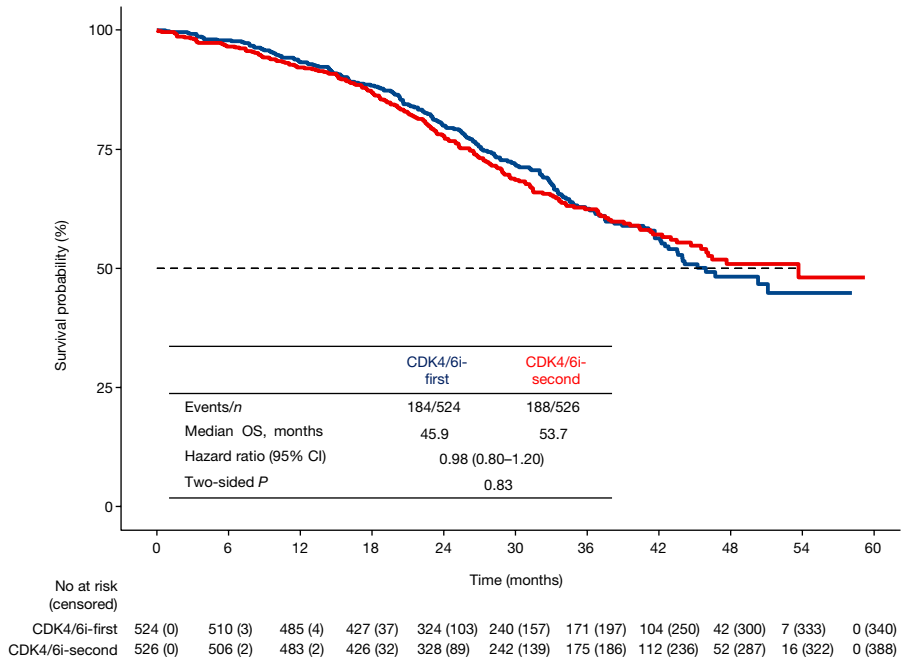


Fig. 4 | OS in the ITT population. Kaplan–Meier plots for OS in the CDK4/6i-first and CDK4/6i-second group in the ITT population. Cox proportional hazard models were used to calculate hazard ratios between the study groups and were stratified according to the stratification factors used in randomization. The difference was assessed using the stratified log-rank test; *P* values are two sided. Events, number of OS events; *n*, number of patients randomized.

oncological outcomes and HRQOL. For patients, the most obvious benefit is the lower toxicity, while a shorter treatment duration also requires fewer hospital visits and less diagnostic assessments. As such, a second-line strategy reduces 'hospital focused' time in a context in which time is precious. For patients and societies, implementation of the SONIA results reduces the burden on the health-care system, including costs. The shorter duration of CDK4/6i with second-line use cuts drug costs by €27,078 per patient based on Dutch reference prices of 2023. In the US setting, the difference in drug costs between a first- and second-line strategy rises to US$184,780 per patient using US list prices. These savings can accumulate to billions of dollars per country, contribute to equitable and sustainable cancer care and expand access to new anticancer therapies for those who would otherwise find them unaffordable.

As half of the SONIA trial population received a shorter treatment duration than recommended in international guidelines, execution of the trial alone (regardless of the results) saved an approximate 20.3 million euros in drug expenditures in the Netherlands (Dutch list prices of 2019). Self-funded trials like SONIA provide a new route for independent academic research aimed at optimizing drug use across all types of cancer and beyond oncology.

SONIA challenges the general oncologic principle of always using the most effective drugs first. Most oncologic drugs are developed in later lines of treatment and are then quickly moved to earlier lines and even the (neo)adjuvant setting, based on clinical trials showing progression-free survival benefits compared to older drugs. Cross-over in studies that aim to establish patient benefit from earlier use of a drug is often limited but should be mandatory. If this crossover nullifies an OS benefit, despite a PFS gain, this finding in itself indicates that early use is not benefitting patients.

The SONIA results stimulate physicians to reconsider the current standard of first-line CDK4/6i for their patients, as first-line use does not lead to better oncologic outcomes or better HRQOL, and is associated with more toxicity and higher drug costs. The current widespread first-line recommendation was not based on randomized data but on concerns about depriving patients of an effective therapy. SONIA reaffirms that robust sequencing trial data are needed to determine the optimal use of our available therapies, and that this type of mostly post-marketing academic research can greatly benefit patients and societies.

Table 2 | Total number of serious adverse events and grade ≥3 adverse events per treatment group in the safety population[a]

	CDK4/6i-first group *n*=520	CDK4/6i-second group *n*=528
Serious adverse events		
Events, total	299	273
Events, average per patient	0.6	0.5
Number of patients with at least one event	165	162
Adverse events, all		
Events, total	2,763	1,591
Events, average per patient	5.3	3.0
Number of patients with at least one event	433	338
Haematological adverse events[b]		
Events, total	2,129	1,066
Events, average per patient	4.1	2.0
Non-haematological adverse events[c]		
Events, total	634	525
Events, average per patient	1.2	1.0

[a]Including all patients who received at least one dose of study treatment, according to actual study treatment received regardless of randomization.
[b]Includes the following terms: anaemia, leukopenia, neutropenia, pancytopenia and thrombocytopenia.
[c]Includes all other event terms according to National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) version 4.0.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-024-08035-2.

1. Finn, R. S. et al. Palbociclib and letrozole in advanced breast cancer. *N. Engl. J. Med.* **375**, 1925–1936 (2016).
2. Hortobagyi, G. N. et al. Updated results from MONALEESA-2, a phase III trial of first-line ribociclib plus letrozole versus placebo plus letrozole in hormone receptor-positive, HER2-negative advanced breast cancer. *Ann. Oncol.* **29**, 1541–1547 (2018).
3. Johnston, S. et al. MONARCH 3 final PFS: a randomized study of abemaciclib as initial therapy for advanced breast cancer. *NPJ Breast Cancer* **5**, 5 (2019).
4. Tripathy, D. et al. Ribociclib plus endocrine therapy for premenopausal women with hormone-receptor-positive, advanced breast cancer (MONALEESA-7): a randomised phase 3 trial. *Lancet Oncol.* **19**, 904–915 (2018).
5. Cristofanilli, M. et al. Fulvestrant plus palbociclib versus fulvestrant plus placebo for treatment of hormone-receptor-positive, HER2-negative metastatic breast cancer that progressed on previous endocrine therapy (PALOMA-3): final analysis of the multicentre, double-blind, phase 3 randomised controlled trial. *Lancet Oncol.* **17**, 425–439 (2016).
6. Slamon, D. J. et al. Phase III randomized study of ribociclib and fulvestrant in hormone receptor-positive, human epidermal growth factor receptor 2-negative advanced breast cancer: MONALEESA-3. *J. Clin. Oncol.* **36**, 2465–2472 (2018).
7. Sledge, G. W. et al. MONARCH 2: abemaciclib in combination with fulvestrant in women with HR+/HER2– advanced breast cancer who had progressed while receiving endocrine therapy. *J. Clin. Oncol.* **35**, 2875–2884 (2017).
8. van Ommen-Nijhof, A. et al. Selecting the optimal position of CDK4/6 inhibitors in hormone receptor-positive advanced breast cancer—the SONIA study: study protocol for a randomized controlled trial. *BMC Cancer* **18**, 1146 (2018).
9. Sledge, G. W. et al. The effect of abemaciclib plus fulvestrant on overall survival in hormone receptor-positive, ERBB2-negative breast cancer that progressed on endocrine therapy-MONARCH 2: a randomized clinical trial. *JAMA Oncol.* **6**, 116–124 (2020).
10. Hortobagyi, G. N. et al. Overall survival with ribociclib plus letrozole in advanced breast cancer. *N. Engl. J. Med.* **386**, 942–950 (2022).
11. Slamon, D. J. et al. Overall survival with ribociclib plus fulvestrant in advanced breast cancer. *N. Engl. J. Med.* **382**, 514–524 (2020).
12. Lu, Y. S. et al. Updated overall survival of ribociclib plus endocrine therapy versus endocrine therapy alone in pre- and perimenopausal patients with $HR^+/HER2^-$ advanced breast cancer in MONALEESA-7: a phase III randomized clinical trial. *Clin. Cancer Res.* **28**, 851–859 (2022).
13. Gradishar, W. J. et al. NCCN Guidelines Insights: Breast cancer, Version 4.2023. *J. Natl Compr. Cancer Netw.* **21**, 594–608 (2023).
14. Kümler, I., Knoop, A. S., Jessing, C. A., Ejlertsen, B. & Nielsen, D. L. Review of hormone-based treatments in postmenopausal patients with advanced breast cancer focusing on aromatase inhibitors and fulvestrant. *ESMO Open* **1**, e000062 (2016).
15. Yang, C. et al. Acquired CDK6 amplification promotes breast cancer resistance to CDK4/6 inhibitors and loss of ER signaling and dependence. *Oncogene* **36**, 2255–2264 (2017).
16. Park, Y. H. et al. Longitudinal multi-omics study of palbociclib resistance in HR-positive/HER2-negative metastatic breast cancer. *Genome Med.* **15**, 55 (2023).
17. Gyawali, B. et al. Problematic crossovers in cancer drug trials. *Nat. Rev. Clin. Oncol.* **20**, 815–816 (2023).
18. Spring, L. M. et al. Cyclin-dependent kinase 4 and 6 inhibitors for hormone receptor-positive breast cancer: past, present, and future. *Lancet* **395**, 817–827 (2020). A.
19. Cherny, N. I. et al. A standardised, generic, validated approach to stratify the magnitude of clinical benefit that can be anticipated from anti-cancer therapies: the European Society for Medical Oncology Magnitude of Clinical Benefit Scale (ESMO-MCBS). *Ann. Oncol.* **26**, 1547–1573 (2015).
20. *G-standaard, Tarieven Januari* (Z-index, 2023).
21. *G-standaard, Tarieven Januari* (Z-index, 2019).
22. Centers for Medicare & Medicaid Services. *Medicare Part D Spending by Drug* (Centers for Medicare & Medicaid Services, accessed 8 May 2024); data.cms.gov/summary-statistics-on-use-and-payments/medicare-medicaid-spending-by-drug/medicaid-spending-by-drug.
23. Johnston, S. et al. Abemaciclib as initial therapy for advanced breast cancer: MONARCH 3 updated results in prognostic subgroups. *NPJ Breast Cancer* **7**, 80 (2021).
24. Rugo, H. S. et al. Palbociclib plus letrozole as first-line therapy in estrogen receptor-positive/human epidermal growth factor receptor 2-negative advanced breast cancer with extended follow-up. *Breast Cancer Res. Treat.* **174**, 719–729 (2019).
25. Di Lauro, V. et al. Health-related quality of life in breast cancer patients treated with CDK4/6 inhibitors: a systematic review. *ESMO Open* **7**, 100629 (2022).
26. Turner, N. C. et al. Overall survival with palbociclib and fulvestrant in advanced breast cancer. *N. Engl. J. Med.* **379**, 1926–1936 (2018).
27. Finn, S. R. et al. Overall survival (OS) with first-line palbociclib plus letrozole (PAL+LET) versus placebo plus letrozole (PBO+LET) in women with estrogen receptor–positive/human epidermal growth factor receptor 2–negative advanced breast cancer (ER+/HER2–ABC): Analyses from PALOMA-2. *J. Clin. Oncol.* **40**, LBA1003 (2022).
28. André, F. et al. Alpelisib for PIK3CA-mutated, hormone receptor-positive advanced breast cancer. *N. Engl. J. Med.* **380**, 1929–1940 (2019).
29. Turner, N. C. et al. Capivasertib in hormone receptor-positive advanced breast cancer. *N. Engl. J. Med.* **388**, 2058–2070 (2023).
30. Bidard, F. C. et al. Elacestrant (oral selective estrogen receptor degrader) versus standard endocrine therapy for estrogen receptor-positive, human epidermal growth factor receptor 2-negative advanced breast cancer: results from the randomized phase III EMERALD trial. *J. Clin. Oncol.* **40**, 3246–3256 (2022).
31. Garcia-Fructuoso, I., Gomez-Bravo, R. & Schettini, F. Integrating new oral selective oestrogen receptor degraders in the breast cancer treatment. *Curr. Opin. Oncol.* **34**, 635–642 (2022).
32. Woodford, R. et al. Validity and efficiency of progression-free survival-2 as a surrogate end point for overall survival in advanced cancer randomized trials. *JCO Precis. Oncol.* **8**, e2300296 (2024).
33. European Medicines Agency. *Appendix 1 to the Guideline on the Evaluation of Anticancer Medicinal Products in Man* (EMA, 2012).
34. Fojo, T. & Simon, R. M. Inappropriate censoring in Kaplan-Meier analyses. *Lancet Oncol.* **22**, 1358–1360 (2021).
35. Gyawali, B. et al. Biases in study design, implementation, and data analysis that distort the appraisal of clinical benefit and ESMO-Magnitude of Clinical Benefit Scale (ESMO-MCBS) scoring. *ESMO Open* **6**, 100117 (2021).
36. Committee for Medicinal Products for Human Use. *Guideline on the Choice of the Non-Inferiority Margin* (EMA, 2005).
37. Tannock, I. F. et al. The tyranny of non-inferiority trials. *Lancet Oncol.* **25**, e520–e525 (2024).
38. Committee for Medicinal Products for Human Use. *Concept Paper for the Development of a Guideline on Non-Inferiority and Equivalence Comparisons in Clinical Trials* (EMA, 2024).
39. André, F. et al. Alpelisib plus fulvestrant for PIK3CA-mutated, hormone receptor-positive, human epidermal growth factor receptor-2-negative advanced breast cancer: final overall survival results from SOLAR-1. *Ann. Oncol.* **32**, 208–217 (2021).
40. Robson, M. E. et al. OlympiAD final overall survival and tolerability results: Olaparib versus chemotherapy treatment of physician's choice in patients with a germline BRCA mutation and HER2-negative metastatic breast cancer. *Ann. Oncol.* **30**, 558–566 (2019).
41. Casparie, M. et al. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cell Oncol.* **29**, 19–24 (2007).

SONIA Study Consortium

Gabe S. Sonke[1,21], Aart Beeker[7], Lisanne C. Hamming[9], Joan B. Heijns[10], Aafke H. Honkoop[11], Paul C. de Jong[12], Quirine C. van Rossum-Schornagel[13], Christa van Schaik-van de Mheen[14], Jolien Tol[15], Cathrien S. Tromp-van Driel[16], Suzan Vrijaldenhoven[17], Inge R. Konings[19,21] & Agnes Jager[2,21]

Methods

Study design and participants

This phase 3, randomized, open-label, investigator-initiated academic trial was conducted in 74 of the 75 hospitals in the Netherlands. The trial compared two strategies: the sequence of an AI plus CDK4/6i for first-line treatment followed by fulvestrant for second-line treatment (the CDK4/6i-first group) versus the sequence of an AI for first-line treatment followed by fulvestrant plus CDK4/6 for second-line treatment (the CDK4/6i-second group) (Fig. 1). Women with HR^+HER2^- ABC who had not received previous systemic therapy for advanced disease were eligible. HR^+ was defined as ER expression greater than 10%. Any menopausal status was allowed, with pre- and perimenopausal women requiring ovarian ablation (surgical or chemical). Previous (neo)adjuvant treatment with a non-steroidal AI (NSAI) was allowed unless the disease had recurred ≤12 months after completing this treatment. Other eligibility criteria included adequate organ function and an Eastern Cooperative Oncology Group performance status of 0 to 2. Tumour lesions had to be measurable or evaluable according to Response Evaluation Criteria in Solid Tumours (RECIST) version 1.1[42]. Key exclusion criteria were the presence of visceral disease requiring chemotherapy, symptomatic central nervous system metastases and previous treatment with CDK4/6i. All of the patients provided written informed consent. The trial was conducted in accordance with the Declaration of Helsinki and the International Guidelines for Good Clinical Practice. The trial protocol and statistical analysis plan (SAP) are available as full text online and were approved by the accredited Medical Ethics Committee of the Netherlands Cancer Institute as were all protocol amendments. An overview of the relevant protocol amendments and adjustments to the SAP is provided in Supplementary Tables 1 and 2, as well as all participating centres. The trial was designed and conducted in close collaboration with the Dutch Breast Cancer Association (BVN). A steering committee and an independent data and safety monitoring committee reviewed the conduct of the trial. The SONIA trial was registered at ClinicalTrials.gov (NCT03425838).

Randomization and masking

Patients were randomized in a 1:1 ratio to receive a CDK4/6i either in first-line or in second-line treatment. All of the patients received the same endocrine backbone, consisting of a NSAI in the first-line context followed at progression by fulvestrant in second-line treatment. Patients were enrolled by the local investigator or research nurse, as specified on the delegation log of each participating centre. Randomization was stratified by site of disease (visceral versus non-visceral), previous endocrine treatment in the (neo)adjuvant setting (yes versus no), hospital and type of CDK4/6i, and was managed through a centralized internet/telephone registration system (ALEA). The choice of type of NSAI (letrozole or anastrozole) was left to the discretion of the treating physician as was the choice of type of CDK4/6i once it was reimbursed. Patients, physicians and investigators were not masked to the treatment received. Data were collected locally in all of the participating hospitals by certified oncology data managers. A centralized electronic case record form (eCRF) was used.

Procedures

Patients received all study treatments according to the standard of care. Dose reductions for CDK4/6i were allowed to manage treatment-related adverse events; no dose reductions were allowed for NSAIs and fulvestrant. Switching to a different drug within the same drug class was also permitted for both NSAIs as well as CDK4/6i to manage side effects. Patients were allowed to discontinue CDK4/6i while continuing endocrine therapy, not vice versa. The protocol-defined switch to second-line treatment was mandated within 28 days of disease progression according to RECIST on first-line treatment. After the two protocol-defined treatment lines, patients in both arms of the trial could receive any available treatment in subsequent treatment lines at the discretion of the treating physician, including alpelisib in the case of a *PIK3CA* mutation.

Tumour assessments according to RECIST were performed locally, at baseline and every 12 weeks thereafter until cessation of the protocol defined treatment sequence. In case of stable disease for at least 2 years during one line of therapy and in the presence of a reliable tumour marker to evaluate the disease (based on the study team's discretion), the scan interval could be prolonged to 24 weeks with tumour marker assessment every 12 weeks (in case of tumour marker rise, radiological evaluation had to be performed within 14 days). The complete assessment schedule has been published previously[8]. Adverse events of grade 3 or higher were recorded throughout the study treatment period according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE), version 4.0[43]. Patient-reported HRQOL outcomes were assessed using the validated Functional Assessment of Cancer Therapy-Breast (FACT-B) and the EuroQol 5-Dimensions (EQ-5D-5L) questionnaire, at a maximum of 11 different timepoints, ranging from baseline until the end of second-line treatment[44]. Data on EQ-5D-5L are not included in this Article. After completion of protocol treatment, all of the patients were followed until death or withdrawal of informed consent.

Primary end point

The primary end point was locally assessed progression-free survival after two treatment lines (PFS2), predefined as the time from randomization until the second objective disease progression (according to RECIST version 1.1), objective disease progression on second-line treatment, symptomatic deterioration on second-line therapy leading to discontinuation of second-line therapy, initiation of chemotherapy for ABC or death, whichever occurred first[33].

We included 'second objective disease progression' and 'objective disease progression on second-line' in the PFS2-definition to account for patients who receive first-line treatment beyond progression (a clinical scenario that may occur, for example, when, according to RECIST, progression is not unequivocal and requires confirmation on the next planned evaluation scan) and for patients who switch to second-line therapy in the absence of objective first disease progression (for example, if the treating physician deems that this is in the best interest of the patient). For these patients, second objective disease progression and objective disease progression on second-line are not necessarily the same. Initiation of chemotherapy for ABC was included as an event because it indicates a failure of the treatment strategy. By contrast, initiation of a new endocrine treatment does not affect PFS2 in our definition. To assess the effect of these choices in the definition of PFS2 on outcome, we performed prespecified sensitivity analyses, as described in the SAP (Supplementary Information).

Secondary end points

Secondary end points included OS (defined as the time from randomization to death from any cause), progression-free survival after one treatment line (PFS1, defined as time from randomization to first progression, initiation of chemotherapy or death, whichever occurred first), HRQOL, toxicity and cost-effectiveness. Analyses of OS, PFS1, HRQOL (as measured by FACT-B), toxicity and the actual CDK4/6i drug costs are described in this Article. Separate publications are scheduled focusing on the remaining prespecified secondary end points, including objective response rate (defined as the proportion of patients with a best overall response of complete or partial response), biomarkers (in tumour tissue, circulating tumour DNA and FES/FDG-PET scans), pharmacokinetics, cognitive functioning, HRQOL and cost-effectiveness.

Statistical analysis

On the basis of the literature and statistical simulations (details are provided in the SAP in the Supplementary Information), we calculated

that inclusion of 500 patients per study group in a period of 42 months and an additional 18 months follow-up will yield an expected number of 574 PFS2 events and 89% power to show superiority of the CDK4/6i-first strategy over the CDK4/6i-second strategy in terms of PFS2 in a log-rank test at the two-sided 95% CI. The threshold for a clinically relevant difference in PFS2 was set according to the European Society for Medical Oncology–Magnitude of Clinical Benefit Scale (ESMO-MCBS)[19], with a clinical benefit grade of 3, indicating an absolute difference in median PFS2 of at least 3 months and the lower bound of the 95% CI around the hazard ratio of ≤0.65. The protocol specified that, in case superiority could not be established, non-inferiority would be tested next. The non-inferiority margin was set according to the ESMO–MCBS at 1.54 (that is, the reciprocal of 0.65).

The Kaplan–Meier method was used to obtain estimates of PFS2. Patients who had not yet reached their PFS2 event at the time of data cut-off were censored at the last moment of contact. Patients who withdrew informed consent during the study were censored at the moment of withdrawal. Patients who were not evaluable due to local treatment or a second malignancy were censored at that moment. Cox proportional hazard models were used to calculate hazard ratios between study groups and were stratified according to the stratification factors used in randomization. The difference in the primary end point PFS2 between groups was assessed using a log-rank test.

Efficacy analyses for both the primary end point PFS2 and secondary end points OS and PFS1 were performed in the ITT population, analysing all patients according to their randomly assigned study group. To assess the robustness of the primary analysis, multiple sensitivity analyses were performed according to the SAP (Supplementary Information). These explored the effect of protocol deviations and the effect of alternative definitions and censoring rules for PFS2 (for example, censoring patients at the initiation of chemotherapy instead of counting this as an event).

Several prespecified subgroup analyses of PFS2 were performed according to the stratification factors used in the analysis (previous (neo)adjuvant endocrine therapy (yes/no), visceral disease (yes/no), type of CDK4/6i) as well as the factors previous (neo)adjuvant chemotherapy (yes/no), de novo metastatic disease (yes/no) and bone-only disease (yes/no). In addition to the prespecified subgroup analyses, we performed post hoc analyses to examine differences between study arms in subgroups based on histological subtype, menopausal status, AI treatment-free interval and *PIK3CA* mutation status. Stratified cox proportional hazards models were used to estimate hazard ratios and 95% CI comparing PFS2 between study groups across the different subgroups, presented in a forest plot. *P* values for interaction were computed to evaluate whether the effect of the treatment strategy is statistically significant different between subgroups. The secondary end points OS and PFS1 were analysed using the Kaplan–Meier method, and with Cox proportional hazards models, stratified according to the stratification factors used in the randomization. Patients who were still alive at the data cut-off were censored for OS at their last contact date at which the patient was known to be alive. The only other patients censored for OS are those who withdrew informed consent. Censoring for PFS1 followed the same rules as for PFS2 specified above. In the safety analysis, all of the patients who received at least one dose of study treatment were analysed according to the treatment they had actually received.

In a prespecified interim analysis, an independent data safety monitoring board evaluated the percentage of patients who started second-line treatment according to protocol after 100 patients in the first-line group had reached PFS1. On the basis of these results, the DSMB recommended to continue the trial according to the protocol.

To provide insights in HRQOL, we present, per study arm, FACT-B total scores at baseline as well as over the different timepoints. Questionnaires were included for the timepoints when they were filled in within 28 days of the required time frame. In addition to this descriptive analysis, we compare HRQOL using longitudinal linear regression modelling, using the FACT-B total score as dependent variable. In this analysis, all questionnaires were included of patients who filled out at least the baseline questionnaire. The variables used in the model were study arm, time since randomization (in days) and HRQOL at baseline. A difference of 7–8 points in FACT-B total score was defined as a clinically meaningful difference[45].

Median and restricted mean CDK4/6i treatment duration were calculated for both study arms based on the start date of the first CDK4/6i cycle and the end date of the last CDK4/6i cycle. Patients still on first- or second-line treatment were censored at data cut-off date.

Costs for CDK4/6i were calculated by multiplying the actual use of CDK4/6i (including dose adjustments and interruptions) per patient as observed with Dutch reference prices from January 2023[20]. The total, average, minimum and maximum costs per study arm were calculated over all patients with CDK4/6i use.

Statistical analyses were performed using SAS (v.9.4) and R (v.4.2.2).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

De-identified patient clinical data underlying the results reported in this Article will be made available to other researchers on reasonable request for academic use, within the limitations of the informed consent and the study's consortium agreement. A detailed data proposal is required and will be considered on a case-by-case basis. Requests should be directed to BOOG study Center (info@boogstudycenter.nl) and will be reviewed by the study's principal investigators. A response will be provided within 90 days. A signed data-access agreement with the sponsor is required before accessing shared data. The study protocol and SAP are provided with the paper.

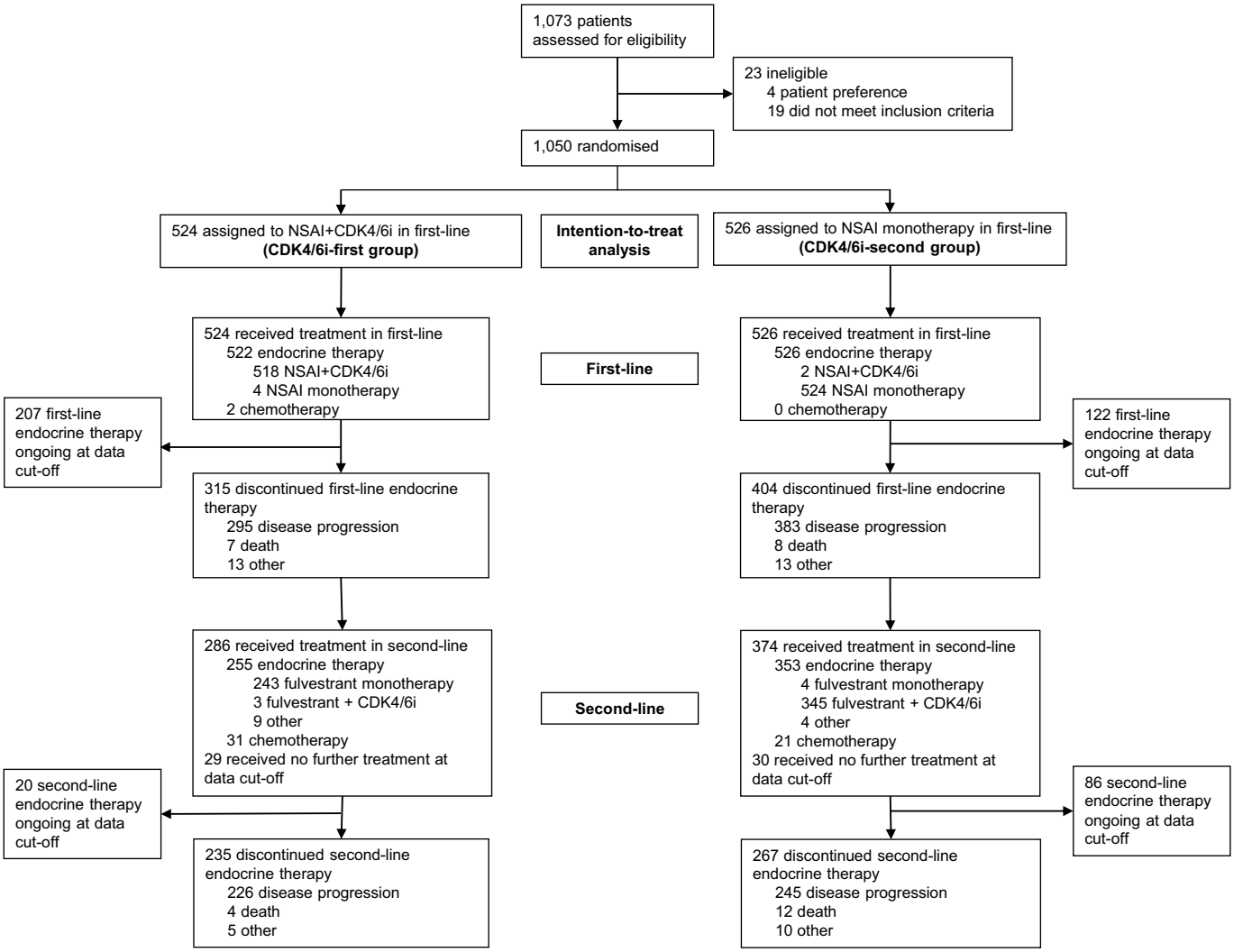
42. Eisenhauer, E. A. et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur. J. Cancer* **45**, 228–247 (2009).
43. Cancer Therapy Evaluation Program. *Common Terminology Criteria for Adverse Events (CTCAE) v.4.0* (National Cancer Institute, 2010).
44. Brady, M. J. et al. Reliability and validity of the functional assessment of cancer therapy-breast quality-of-life instrument. *J. Clin. Oncol.* **15**, 974–986 (1997).
45. Eton, D. T. et al. A combination of distribution- and anchor-based approaches determined minimally important differences (MIDs) for four endpoints in a breast cancer scale. *J. Clin. Epidemiol.* **57**, 898–910 (2004).

Acknowledgements We thank all the patients who participated in this trial, as well as their families; all members of the steering committee and the study staff at all 74 hospitals of the SONIA trial; the independent Data Safety Monitoring Board; the Dutch Breast Cancer research Group (BOOG) for their support and sponsorship; the registration team of the Netherlands Comprehensive Cancer Organization (IKNL) for the collection of data for the Netherlands Cancer Registry; and all data managers for data collection throughout the trial. The Dutch Nationwide Pathology Databank (Palga) provided histopathology data. The Dutch Breast Cancer Association (BVN) was involved in the development and the execution of the trial. The trial was funded by the Netherlands Organization for Health Research (ZonMw) and Development and Dutch Health Insurers (ZN).

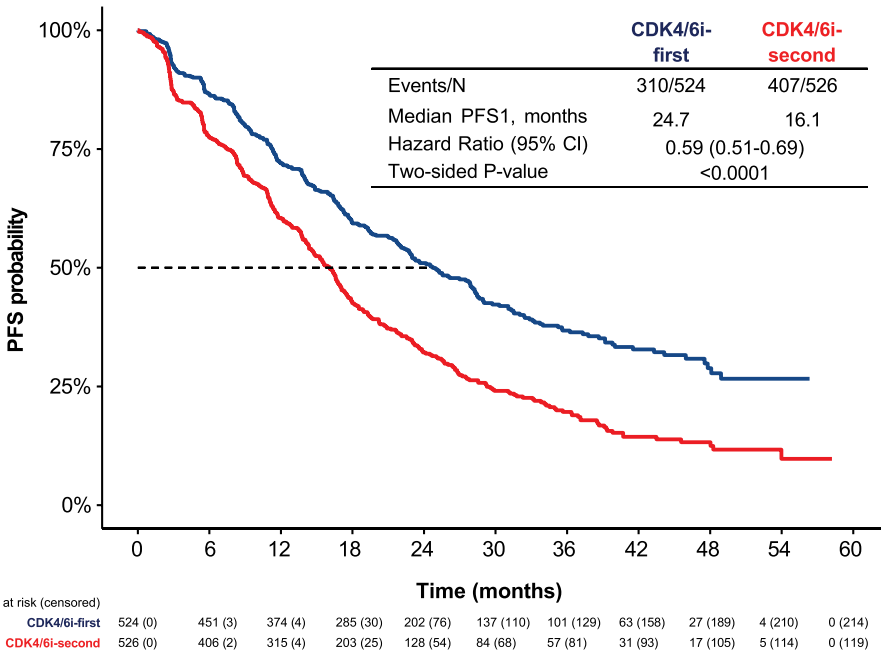
Author contributions A.J., G.S.S. and I.R.K. initiated the study. A.J., G.S.S., I.R.K. and V.v.d.N. designed the study with support from the study steering committee and The Dutch Breast Cancer Association (BVN). C.G.P. as director of BVN was involved in the development and the execution of the trial. A.E.v.L.-S. as director of the Dutch Breast Cancer research Group (BOOG) was involved in the project administration, funding acquisition and in the development and execution of the trial. A.B., A.v.O.-N., A.H.H., A.J., C.v.S.-v.d.M., C.S.T.-v.D., G.S.S., I.R.K., J.B.H., J.T., K.B., L.C.H., N.W., P.C.d.J., Q.C.v.R.-S. and S.V. contributed to recruitment of patients. A.C.P.S. and L.M. contributed to data collection and validation on behalf of the Netherlands Comprehensive Cancer Organization (IKNL). V.v.d.N. was the trial statistician. A.v.O.-N., N.W. and V.v.d.N. accessed and verified the data and contributed to the data analysis. H.M.B. contributed to the data analysis of the health-related quality-of-life and costing. A.J., G.S.S. and I.R.K. contributed to supervision of the study. A.J., A.v.O.-N., G.S.S., I.R.K. and N.W. wrote the initial draft of the article and decided to submit the manuscript. All of the authors participated in interpretation of the data and drafting of the manuscript. All of the authors reviewed and approved the final, submitted version.

Competing interests G.S.S. reports institutional research support from Agendia, AstraZeneca, Merck, Novartis, Roche and Seagen; and consultancy for Biovica, Novartis and Seagen. H.M.B. received grants from CADTH, ZIN and Medical Delta; and participated in a data safety monitoring board or advisory board for Pfizer. A.H.H. received consulting fees from Gilead and Lilly; and received payment or honoraria from Lilly. Q.C.v.R.-S. has participated in a data safety monitoring board or advisory board for Roche. I.R.K. reports institutional research grant support from Novartis and Gilead. The other authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-024-08035-2.
Correspondence and requests for materials should be addressed to Gabe S. Sonke.
Peer review information *Nature* thanks Suzette Delaloge, Debu Tripathy and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.



Extended Data Fig. 1 | Trial profile. The figure presents an overview of the course of the trial for all study participants. Of note, the number of deaths reported here are only those deaths that were reported to be the reason for discontinuation of first- or second-line therapy (i.e., patients who died in the absence of objective disease progression while on study treatment). The number of PFS2 events (n = 591) is different from the number of patients that discontinued second-line endocrine therapy (n = 502), since not all patients with a PFS2 event discontinued second-line treatment and not all patients who discontinued second-line treatment experienced a PFS2 event. NSAI, non-steroidal aromatase inhibitor; CDK4/6i, cyclin-dependent kinase 4 and 6 inhibitor.



Extended Data Fig. 2 | PFS1 in ITT population. Kaplan–Meier plots for PFS after one treatment line (PFS1) in the CDK4/6i-first and CDK4/6i-second group in the ITT population. Cox proportional hazard models were used to calculate hazard ratios between the study groups and were stratified according to the stratification factors used in randomization. The difference was assessed using the stratified log-rank test. *P* values are two-sided. Events, number of PFS1 events; *n*, number of patients randomized.

Extended Data Table 1 | Overview of sensitivity analysis

	Median PFS2 (months)		HR (stratified)	95% CI of the HR
	CDK4/6i-first group	CDK4/6i-second group		
Intention-to-treat analysis	31.0	26.8	0.87	0.74-1.03
As treated analysis	31.0	26.8	0.87	0.74-1.03
Sensitivity analyses				
Treatment beyond progression				
Sensitivity 1	33.1	29.8	0.89	0.74-1.05
Sensitivity 2	30.8	26.6	0.88	0.75-1.04
Sensitivity 3	32.0	28.6	0.91	0.77-1.07
Time-on-SONIA-analysis	29.1	28.1	0.93	0.80-1.10
Clinical progression				
Sensitivity 4	32.3	28.6	0.86	0.73-1.01
Start a different second-line treatment				
Sensitivity 5	33.1	28.9	0.84	0.70-0.99
Sensitivity 6	32.0	27.3	0.86	0.73-1.01
Sensitivity 7	31.3	27.3	0.87	0.73-1.02
Prolonged scan interval before progression				
Sensitivity 8	30.9	26.8	0.88	0.75-1.03
Sensitivity 9	31.0	26.8	0.87	0.74-1.03
Sensitivity 10	31.0	27.3	0.88	0.74-1.03

Overview of sensitivity analyses and their results compared to the primary ITT analysis. PFS2 in the ITT analysis is defined as time from randomization until second objective disease progression (according to RECIST), objective disease progression on second-line treatment, initiation of chemotherapy for advanced breast cancer, or death, whichever occurred first. In the as treated analysis, patients are analysed according to the study medication they actually received. Several prespecified sensitivity analyses were performed to assess the robustness of the ITT analysis of the primary end point, as detailed in the statistical analysis plan (SAP). We performed several sensitivity analysis (sensitivity 1, 2, and 3) regarding the protocol deviation of 'treatment beyond progression', i.e. continuing first-line treatment ≥28 days beyond objective progression. In sensitivity analysis 1, patients with treatment beyond progression are censored at their first objective progression. In sensitivity analysis 2, PFS2 is defined as time from randomization until first progression plus the time from start second line treatment to progression on second-line, start of chemotherapy or death, whichever occurred first (definition applied to all patients). In sensitivity analysis 3, PFS2 is defined as time from randomization until progression on second-line therapy, initiation of chemotherapy for advanced breast cancer, or death, whichever occurred first (definition applied to all patients). In the 'time-on-SONIA-analysis', PFS2 is defined as time from randomization until discontinuation of second-line therapy (or first-line therapy in case first-line treatment was discontinued and no second-line according to the SONIA protocol was started) (definition applied to all patients). In sensitivity analysis 4, patients that switch to second-line therapy in the absence of objective progression, are censored at their first objective progression (i.e. in case of clinical progression). In sensitivity analysis 5, 6, and 7 we analysed the effect of patients starting a different second-line therapy as specified in the SONIA protocol. In sensitivity analysis 5, patients are censored at the initiation of chemotherapy instead of counting this as event. In sensitivity analysis 6, PFS2 is defined as time from randomization until second objective disease progression, progression on second-line or death, whichever occurred first. The initiation of chemotherapy is no event/censor moment in this sensitivity analysis (definition applied to all patients). In sensitivity analysis 7, patients are censored at the initiation of a different endocrine treatment in case no PFS2 event had occurred until that moment. To assess the effect of the ITT analysis of patients that had their PFS2 event after prolonged scan interval (i.e. objective progression after a scan interval of >24 weeks), we performed sensitivity analysis 8, 9, and 10. In sensitivity analysis 8, PFS2 is defined as time from randomization until their last known tumour assessment plus 12 weeks (definition applied to the patients with a prolonged scan interval). In sensitivity analysis 9, PFS2 is defined as time from randomization until the actual progression date minus 12 weeks (definition applied to the patients with a prolonged scan interval). In sensitivity analysis 10, patients are censored at their last known tumour assessment (definition applied to the patients with a prolonged scan interval). ITT, intention-to-treat; HR, hazard ratio; CI, confidence interval.

Extended Data Table 2 | HRQOL over time, descriptive analysis

	Line 1						Line 2					
	baseline 1	3 months	6 months	12 months	18 months	end line 1	baseline 2	3 months	6 months	12 months	18 months	end line 2
Number of questionnaires	921	751	684	532	378	473	409	345	212	108	55	283
	FACT-B total score (n)						**FACT-B total score (n)**					
CDK4/6i first-group												
All questionnaires	105 (464)	108 (396)	108 (357)	109 (293)	108 (220)	99 (197)	101 (157)	100 (129)	103 (52)	101 (23)	112 (9)	97 (134)
Not representing EOT[a]	105 (462)	109 (376)	108 (347)	109 (290)	109 (211)	NA (NA)	102 (143)	105 (55)	108 (38)	103 (16)	118 (6)	NA (NA)
Representing EOT[b]	104 (2)	94 (20)	106 (10)	102 (3)	98 (9)	99 (197)	95 (14)	97 (74)	89 (14)	98 (7)	102 (3)	97 (134)
CDK4/6i second-group												
All questionnaires	103 (457)	107 (355)	109 (327)	109 (239)	109 (158)	102 (276)	103 (252)	103 (216)	106 (160)	105 (85)	108 (46)	98 (149)
Not representing EOT[a]	103 (450)	108 (342)	109 (310)	109 (228)	109 (151)	NA (NA)	103 (245)	105 (175)	107 (134)	107 (76)	110 (36)	NA (NA)
Representing EOT[b]	90 (7)	101 (13)	100 (17)	96 (11)	99 (7)	102 (276)	77 (7)	96 (41)	99 (26)	90 (9)	101 (10)	98 (149)
	Mean difference FACT-B total score compared to baseline (n)						**Mean difference FACT-B total score compared to baseline (n)**					
CDK4/6i first-group												
All questionnaires		2.3 (396)	2.2 (357)	2.7 (293)	1.2 (220)	-4.7 (197)	-3.7 (157)	-5.2 (129)	-4 (52)	-6.5 (23)	0.6 (9)	-8.4 (134)
Not representing EOT[a]		2.7 (376)	2.4 (347)	2.7 (290)	1.6 (211)	NA (NA)	-3.0 (143)	-0.8 (55)	-1.1 (38)	-8.8 (16)	4.5 (6)	NA (NA)
Representing EOT[b]		-6 (20)	-5.4 (10)	-3.8 (3)	-9.3 (9)	-4.7 (197)	-10.9 (14)	-8.5 (74)	-11.6 (14)	-1.6 (7)	-7.2 (3)	-8.4 (134)
CDK4/6i second-group												
All questionnaires		3 (355)	4.1 (327)	3.4 (239)	2.8 (158)	-1.3 (276)	-0.9 (252)	-0.5 (216)	1.1 (160)	-0.2 (85)	0.1 (46)	-6.6 (149)
Not representing EOT[a]		3.1 (342)	4.5 (310)	3.5 (228)	3.4 (151)	NA (NA)	-0.5 (245)	1.6 (175)	2.4 (134)	1.3 (76)	-0.9 (36)	NA (NA)
Representing EOT[b]		-0.3 (13)	-3.2 (17)	1 (11)	-9.5 (7)	-1.3 (276)	-13.4 (7)	-9.8 (41)	-5.5 (26)	-14.6 (9)	3.6 (10)	-6.6 (149)

To provide insight in HRQOL, we present, per study arm, FACT-B total scores at baseline as well as over the different timepoints. For each patient, up to 11 identical questionnaires (representing distinct timepoints) were completed. Questionnaires were included in the time points (e.g., 3 months, 6 months from randomization. etc.) when they were filled in within 28 days of the required time frame. The baseline second-line questionnaire comprises patients who started second-line SONIA treatment and completed a questionnaire within 28 days of starting second-line treatment. The End of treatment (EOT) timeframe was defined from 28 days before and up to 60 days after the discontinuation of first- or second-line treatment. Extended Data Table 2 shows the number of completed questionnaires (n), mean FACT-B total scores, and mean difference in FACT-B total scores over time. The difference per patient was calculated by subtracting baseline FACT-B total score from the score per timepoint; positive values mean improvement. Results are presented separately for the CDK4/6i-first and CDK4/6i second groups, and for patients meeting the EOT timeframe during a timepoint. A difference of 7-8 points in FACT-B total score was defined as a clinically meaningful difference[45]. [a]Questionnaire not within 28 days before or up to 60 days after the discontinuation of treatment. [b]Questionnaire within 28 days before and up to 60 days after the discontinuation of treatment.

Extended Data Table 3 | HRQOL over time, regression analysis

	Estimate	Standard error	Df	Z-value	Pr(>|z|)
(Intercept)	37,950	2,484	866,602	15,280	< 2e-16***
Arm CDK4/6i second-group	-0,910	0,807	846,026	-1,128	0,26
as.numeric(time[a])	-0,008	0,001	3048,066	-12,159	<2e-16***
FACT-B total score at baseline	0,676	0,023	852,488	29,103	< 2e-16***

Results of the longitudinal linear regression modelling showing the estimates (coefficients), standard errors, degrees of freedom (df), z-values and two-sided *p* values for all variables in the model. [a]Time since randomization in days (continuous variable). Of note, all questionnaires up until a maximum of 60 days after end of SONIA treatment were included of patients who filled out at least the baseline questionnaire (no restrictions on timepoints as compared to the descriptive analysis), resulting in 869 patients contributing to 3,634 observations). ***P value ≤ 0.001.

Extended Data Table 4 | Total number of adverse events per treatment group in the safety population

	CDK4/6i-first group N=520	CDK4/6i-second group N=528
Any adverse event	433 (83)	338 (64)
Neutrophil count decreased	313 (60)	193 (37)
White blood cell decreased	68 (13)	49 (9)
Hypertension	43 (8)	32 (6)
GGT increased	39 (8)	30 (6)
Anemia	31 (6)	39 (7)
Platelet count decreased	31 (6)	17 (3)
Dyspnea	20 (4)	29 (5)
Thromboembolic event	17 (3)	9 (2)
Aspartate aminotransferase increased	17 (3)	14 (3)
Bone pain	15 (3)	21 (4)
Diarrhea	15 (3)	9 (2)
Fatigue	15 (3)	9 (2)
Alanine aminotransferase increased	14 (3)	8 (2)
Fracture	14 (3)	11 (2)
Back pain	13 (3)	8 (2)
Pleural effusion	11 (2)	15 (3)
Urinary tract infection	11 (2)	5 (1)
Vomiting	11 (2)	7 (1)
Lung infection	10 (2)	6 (1)
Nausea	7 (1)	10 (2)

Data are numbers (%). Listed are grade ≥3 adverse events that were reported in at least 2% of the patients in any treatment group. [a]Including all patients who received at least one dose of study treatment, according to actual study treatment received regardless of randomization.

Extended Data Table 5 | Total CDK4/6i drug costs per study group with drug prices in the Netherlands derived from both list prices of 2023 (a) and 2019 (b) and the United States (c)

a. Dutch list price of 2023	CDK4/6i-first group N=524	CDK4/6i-second group N=526	Incremental differences
Number of patients with CDK4/6i use at data cut-off	519	345	
Total costs	€ 24,890,620	€7,204,034	€17,686,586
Average costs per patient	€47,959	€20,881	€27,078
Minimum costs per patient	€725	€612	€113
Maximum costs per patient	€139,626	€101,602	€38,024

b. Dutch list price of 2019	CDK4/6i-first group *n*=524	CDK4/6i-second group *n*=526	Incremental differences
Number of patients with CDK4/6i use at data cut-off	519	345	
Total costs	€ 28,577,724	€ 8,292,270	€20,285,454
Average costs per patient	€ 55,063	€ 24,036	€31,027
Minimum costs per patient	€ 868	€ 700	€168
Maximum costs per patient	€ 158,465	€ 114,777	€43,688

c. US list price of 2022	CDK4/6i-first group *n*=524	CDK4/6i-second group *n*=526	Incremental differences
Number of patients with CDK4/6i use at data cut-off	519	345	
Total costs	$170,792,867	$49,783,721	$121,009,146
Average costs per patient	$329,081	$144301	$184,780
Minimum costs per patient	$5,552	$4,164	$1,388
Maximum costs per patient	$909,644	$670,346	$239,298

d. Input drug costs	Mg	Price NL 2023[20]	Price NL 2019[21]	Price US 2022[22]
Abemaciclib	50	€47	€53	$259
Abemaciclib	100	€47	€53	$259
Abemaciclib	150	€47	€53	$259
Palbociclib	75	€102	€117	$694
Palbociclib	100	€102	€117	$694
Palbociclib	125	€102	€117	$694
Ribociclib	200	€35	€41	$272

a. Drug prices in the Netherlands in euros (€), derived from the Dutch list price Z-index of 2023[20]. The total, average, minimum and maximum costs per study group were calculated over all patients with CDK4/6i use at data cut-off date.
b. Drug prices in the Netherlands in euros (€), derived from the Dutch list price Z-index of 2019[21]. The total, average, minimum and maximum costs per study group were calculated over all patients with CDK4/6i use at data cut-off date.
c. Drug prices in the United States (US) in dollars ($), derived from Medicare Part D from 2022[22]. The total, average, minimum and maximum costs per study group were calculated over all patients with CDK4/6i use at data cut-off date.
d. List prices of CDK4/6i for the Netherlands (NL) and the United States (US). These input prices are used to calculate the actual CDK4/6i per patient: the type of CDK4/6i (abemaciclib, palbociclib, ribociclib), dose per administration in milligram (mg), and number of administrations are multiplied with these prices.

Extended Data Table 6 | Overview subsequent treatments in third line and higher

	CDK4/6i-first group N=524	CDK4/6i-second group N=526
Patients starting subsequent treatment / patients discontinuing 2nd line endocrine therapy, n/N (%)	189[a]/235 (80)	217[a]/267 (81)
Chemotherapy, n (%)	158 (84)	190 (88)
Endocrine monotherapy, n (%)	50 (26)	50 (23)
Endocrine monotherapy + targeted therapy, n (%)	36[b] (19)	53[b] (24)
CDK4/6i	9 (25)	5 (9)
Alpelisib	6 (17)	9 (17)
Everolimus	23 (64)	41 (77)
Other	0 (0)	1[c] (4)
Targeted therapy alone, n (%)	0 (0)	1[d] (0)
Other systemic therapy, n (%)	6[e1] (3)	9[e2] (4)

Overview of subsequent treatments among patients who discontinued second-line endocrine therapy. The table contains numbers of all subsequent treatment lines, meaning third-line of treatment and thereafter. Numbers can add to more than 100% as patients could receive more than 1 subsequent treatment line. A patient was counted only once in one of the medication types. [a]Used as the denominator for the percentages in the following sub-treatments. [b]Used as the denominator for the percentages in the following sub-treatments. [c]Taselisib in combination with tamoxifen. [d]Abemaciclib monotherapy. [e1]Atezolizumab, Bevacizumab, Olaparib, Lenvatinib, Trastuzumab. [e2]Bevacizumab, Olaparib, Pembrolizumab, Pertuzumab + Trastuzumab.

nature portfolio

Corresponding author(s): Gabe S Sonke

Last updated by author(s): 8th of August, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	electronic Case Report Forms (eCRFs) in ALEA datamanagement (current version 18.6)
Data analysis	The statistical analysis were performed using SAS software (version 9.4) and R (version 4.2.2)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

De-identified patient clinical data that underlie the results reported in this article will be made available to other researchers on reasonable request for academic use, within the limitations of the informed consent and the study's consortium agreement. A detailed data proposal is required and will be considered on a case-by-case basis. Requests should be directed to BOOG study Center (info@boogstudycenter.nl) and will be reviewed by the study's principle investigators. A response will be provided within 90 days. A signed data access agreement with the sponsor is required before accessing shared data. The study protocol and statistical analysis plan are provided with the paper.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	The SONIA study is a phase 3 trial in adult women (18 years or older) with proven diagnosis of ER+, HER2- advanced breast cancer.
Reporting on race, ethnicity, or other socially relevant groupings	The SONIA study did not collect data on race or ethnicity. In this manuscript, we did not use self-reported socioeconomic data.
Population characteristics	SONIA participants are women aged 18 years and older whom are treated for advanced breast cancer in the Netherlands.
Recruitment	From 23 November 2017 and 1 September 2021, participants were recruited from 74 Dutch hospitals using eligibility criteria pre-specified in the study protocol. Eligible patients were enrolled without selection. The eligibility criteria were designed to closely reflect a real-world patient population, thereby minimizing the potential of selection bias. Patients were enrolled by the local investigator or research nurse of each participating center. Randomisation was stratified by hospital. The investigator at each center ensured that all patients were given full and adequate oral and written information about the nature, purpose, possible risk, and benefit of the study. Patients were recruited according to the study protocol and provided written informed consent before recruitment. Patients and investigators were not masked to the treatment assigned to. The risk of bias due to attrition is minimal as the number of patients not starting second-line treatment was similar in both treatment groups.
Ethics oversight	All patients provided written informed consent prior to enrollment. The trial was conducted in accordance with the Declaration of Helsinki and the International Guidelines for Good Clinical Practice. The trial protocol and statistical analysis plan (SAP) are available as full text online and were approved by the accredited Medical Ethics Committee of the Netherlands Cancer Institute as were all protocol amendments

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The trial is designed to have 89% power to show a HR for CDK4/6i in first-line / CDK4/6i in second-line of 0.765 and a 95% CI of 0.648-0.902 as produced by Cox' proportional hazards model. At a type I error probability of 0.05, 1000 patients are required over a period of 42 months and followed over an additional 18 months, yielding an expected number of 574 events. Anticipating 5% of patients to be non-evaluable, a total of 1,050 patients will need to be randomised. Please see the statistical analysis plan for more details.
Data exclusions	No data were excluded from the analysis (data cut-off date December 1, 2022)

Replication

This was a prospective clinical trial and replication was not within the scope of the trial.

Randomization

This was a phase 3 randomised controlled trial: randomisation was managed through a centralized internet/telephone registration system (ALEA)

Blinding

Patients and study personnel were not blinded to the treatment assgined to (either the CDK4/6i-first or CDK4/6i-second group). Rationale is that the SONIA study is not comparing the effectivity of different drugs, but different strategies. However, we cannot rule out bias related to physician or patient awareness of the assigned treatment. We believe that the effect of this potential bias on the primary endpoint is limited, because all patients underwent protocol-defined tumour assessments. For the patients that did not undergo all protocol-defined tumour assessments, we performed sensitivity analyses to investigate the effect of such protocol deviations on the primary outcome. It is also unlikely that the open-label design affected OS results.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration

ClinicalTrials.gov Identificer, NCT03425838

Study protocol

The first approved protocol (version 1.2) and the last protocol version (version 1.11) are attached as supplemented data

Data collection

The SONIA study was conducted between November 23, 2017 and September 1, 2021 at 74 sites in The Netherlands. Tumour assessments according to RECIST were performed locally, at baseline and every 12 weeks thereafter until cessation of the protocol defined treatment sequence. Adverse events grade 3 or higher were recorded throughout the study treatment period according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE), version 4.042. Patient-reported health-related quality of life (HRQoL) outcomes were assessed using the validated Functional Assessment of Cancer Therapy-Breast (FACT-B) and the EuroQol 5-Dimensions (EQ-5D-5L) questionnaire, at a maximum of 11 different time points, ranging from baseline until end of second-line treatment. The clinical data were collected locally in electronic Case Report Forms (using ALEA Clinical) in all participating hospitals by certified oncology data managers. Data cleaning occurred centrally by the Netherlands Comprehensive Cancer Organization (IKNL). The data cut-off date was December 1, 2022. All participating centers are listed in the supplementary information.

Outcomes

The primary and secondary endpoints were predefined in the protocol and SAP. The primary endpoint was locally assessed progression-free survival after two treatment lines (PFS2), defined as time from randomisation until second objective disease progression (according to RECIST version 1.1), objective disease progression on second-line treatment, symptomatic deterioration on second-line therapy leading to discontinuation of second-line therapy, initiation of chemotherapy for ABC or death, whichever occurred first. This definition was based on EMA guidelines and recommendations from the Data Safety Monitoring Board and was finalized before the data cut-off date, without knowledge of the data and before any analysis were performed. Secondary endpoints included overall survival (OS, defined as time from randomisation to death from any cause), progression-free survival after one treatment line (PFS1, defined as time from randomisation to first progression, initiation of chemotherapy or death, whichever occurred first), HRQoL, toxicity and cost-effectiveness. Adverse events grade 3 or higher were recorded throughout the study treatment period according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE), version 4.042. Patient-reported health-related quality of life (HRQoL) outcomes were assessed using the validated Functional Assessment of Cancer Therapy-Breast (FACT-B) and the EuroQol 5-Dimensions (EQ-5D-5L) questionnaire, at a maximum of 11 different time points, ranging from baseline until end of second-line treatment. Cost-effectiveness was assessed by reporting health-care resource use on a patient-level while on SONIA treatment.

Plants

Seed stocks	Not applicable.
Novel plant genotypes	Not applicable.
Authentication	Not applicable.