Supplementary Online Content

Steenbruggen TG, Steggink LC, Seynaeve CM, et al. High-dose chemotherapy with hematopoietic stem cell transplant in patients with high-risk breast cancer and 4 or more involved axillary lymph nodes: 20-year follow-up of a phase 3 randomized clinical trial. *JAMA Oncol.* Published online January 30, 2020. doi:10.1001/jamaoncol.2019.6276

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods 1. Additional Details on Data Sources and Collection of Data

Central review of ER, PR and ERBB2

Central review of estrogen receptor (ER), progesterone receptor (PR), and ERBB2, Erb-B2 receptor tyrosine kinase 2 (ERBB2) on all available pathological specimens was performed in 2006.[1] Tumors were considered positive for hormone receptors if ≥10% of the tumor cells showed nuclear staining. ER-positivity was used to define breast cancer subtypes. Tumors were considered ERBB2-positive in case of strong homogeneous positive staining (3+) by immunohistochemistry or gene amplification by *in situ* hybridization in case of 2+ by immunohistochemistry analysis.[2]

Collection of data

Collected data included duration of endocrine therapy, including start of an aromatase inhibitor, occurrence, localization and date of relapse, treatment for relapse, date and cause of death, occurrence and date of second malignancies, treatment for second malignancies, history of cardiac disease before breast cancer diagnosis, occurrence and date of diagnosis of major cardiovascular events (i.e., coronary artery disease, myocardial infarction, transient ischemic attack, cerebrovascular accident, cardiac valve dysfunction), minor cardiovascular events (i.e., dysrhythmias, pericarditis, peripheral vascular events), cardiovascular risk factors (i.e., hypertension, hypercholesterolemia, diabetes mellitus type 2, body weight, smoking status), and medication at last follow-up. Second malignancies were classified according to International Statistical Classification of Diseases and Related Health Problems (ICD-10). Myelodysplastic syndrome was included as second malignancy, in situ carcinoma was not. Database cut-off was set on September 16th, 2018.

The general practitioners (GP) of each patient was sent a questionnaire, requesting medical correspondence regarding disease recurrence, second malignancies, cardiovascular outcomes, and vital status. In addition, physicians in referral hospitals received questionnaires and were requested to submit medical correspondence if patients had received treatment for cancer or cardiovascular disease in hospitals other than the ten hospitals participating in the original trial. Response rates from GPs and treating physicians were 68% and 77%, respectively, without difference between treatment arms. Data on BC recurrence status and incidence of cardiovascular events (and risk factors) was deemed complete if information was available from 2016 onwards or until death.

Additionally, all patients were linked via unique, pseudo-anonymized patient identification numbers with the population-based municipal Personal Records Database for vital status and with Statistics Netherlands for the cause of death. To complete and verify receptor-status from the primary tumor and data on second malignancies (including contralateral BC), patient-level data from the Netherlands Cancer Registry (NCR) and the nationwide network and registry of

histopathology and cytopathology in the Netherlands (PALGA) were extracted in December 2017.

Personal Records Database

The Personal Records Database records vital status of everyone born in the Netherlands, or who enters the country and will remain longer than four months. Personal Records Database data were used to acquire the exact date of death. Patients who did not receive active follow-up at database cut-off were censored for overall survival at the last Personal Records Database update, which was February 1st, 2017.

Statistics Netherlands

Statistics Netherlands registers the underlying cause of death for each decedent in the Netherlands. Detailed information on methods of registration of causes of death in the Netherlands has been published previously.[3] In short, for each deceased person, attending physicians or medical examiners are obliged to fill out a death certificate, which is sent to Statistics Netherlands. Statistics Netherlands codes the death certificates and an underlying cause is selected according to the 10th revision of the International Statistical Classification of Diseases and Related Health Problems (ICD-10). A maximum of four diseases can be coded per decedent. The ICD-10 defines the underlying cause of death as: 1) the disease or injury that initiated the train of morbid events leading directly to death or 2) the circumstances of the accident or violence that produced the fatal injury. Volume 2 of ICD-10 provides several instructions for the application of definition 2.

According to Statistics Netherlands, four patients died due to breast cancer although no relapse of breast cancer was known at last follow-up when patients' records were reviewed. For these four patients, the missing date of first relapse was substituted by the median time between relapse and death within the sample set, for the ER-negative and ER-positive subsets separately.

The primary cause of death was based on information acquired during medical record abstraction. The cause of death was categorized as (1) breast cancer, (2) non-breast malignancy, (3) related to the protocol treatment and within six months after completion of protocol treatment, (4) cardiovascular disease, or (5) other causes. To obtain cause-of-death data for patients no longer under medical surveillance, causes of death through December 2016 were obtained through linkage with Statistics Netherlands, classified according to International Classification of Diseases, 9th or 10th revision, depending on date of death.[4] Linkage of deceased patients was successful in 98% (495/505) of cases. Cases with inconsistently reported causes of death were discussed within the study team and consensus was reached based on all information available from the medical records. A recent study showed 95% concordance between the registered cause of death by Statistics Netherlands

and the cause of death obtained from the medical records in patients previously treated for breast cancer.[5]

Netherlands Cancer Registry (NCR) and the nationwide network and registry of histopathology and cytopathology in the Netherlands (PALGA)

The NCR has almost complete nationwide coverage (>96%) of invasive malignant neoplasms and ductal carcinoma in situ of the breast occurring in the Netherlands since 1989.[3] Data was used to complete data on incidence of second malignancies.

PALGA is a nationwide network of all pathology laboratories in the Netherlands with national coverage since 1991. Patient-level data obtained through linkage with the PALGA used to confirm and complete data on the ER, PR and ERBB2 status from the primary tumor. Additionally, pathology data were used to eliminate uncertainties about second malignancies or distant metastases. In case of discrepancies regarding date of recurrence or second malignancy, the date obtained from the NCR was used, or, if unavailable, the earliest date as reported in the other data sources.

eMethods 2. Additional Details on Statistical Analysis

The reverse Kaplan-Meier method was used to determine median follow-up and its interquartile range (IQR). The Kaplan-Meier method was used to calculate OS, and BCSS probabilities, and these were compared with log-rank test. Hazard ratios (HR) with 95% confidence intervals (CI) are based on fitted Cox proportional hazard models, including p-values for interaction tests for the pre-planned subgroup analyses for this update.

The cumulative incidence of second malignancies between the two treatment arms was compared, treating death as competing risk. The p-value was based on the K-sample test.[6] For cardiovascular events the event dates were not available for all patients. Therefore, occurrence of cardiovascular events between groups was evaluated using Chi^2-test.

eReferences

1. Rodenhuis S, Bontenbal M, van Hoesel QGCM, et al. Efficacy of high-dose alkylating chemotherapy in HER2/neu-negative breast cancer. *Ann Oncol*. 2006;17(4):588-596. doi:10.1093/annonc/mdl001
2. Senkus E, Kyriakides S, Ohno S, et al. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2015;26 Suppl 5:v8-30. doi:10.1093/annonc/mdv298
3. van der Sanden GA, Coebergh JW, Schouten LJ, Visser O, van Leeuwen FE. Cancer incidence in the Netherlands in 1989 and 1990: first results of the nationwide Netherlands cancer registry. Coordinating Committee for Regional Cancer Registries. *Eur J Cancer*. 1995;31A(11):1822-1829.
4. Harteloh P, de Bruin K, Kardaun J. The reliability of cause-of-death coding in The Netherlands. *Eur J Epidemiol*. 2010;25(8):531-538. doi:10.1007/s10654-010-9445-5
5. Vries S de, Schaapveld M, Kardaun JWPF, et al. Comparing causes of death of Hodgkin lymphoma and breast cancer patients between medical records and cause-of-death statistics. *Clin Epidemiol*. 2018;10:1523-1531.
6. Gray RJ. A class of K-sample tests for comparing the cumulative incidence of a competing risk. *Ann Stat*. 1988;16(3):1141-1154.

eTable 1. Baseline Characteristics			
		HDCT Group (n=442)	**CDCT Group (n=443)**
Age at randomization, median (IQR), y		46 (23–55)	45 (25–55)
Tumor size classification, no. (%)			
	T1	90 (20%)	105 (24%)
	T2	274 (62%)	260 (59%)
	T3	70 (16%)	73 (16%)
	Unknown	8 (2%)	5 (1%)
Involved axillary lymph nodes, no. (%)			
	4-9 nodes	284 (64%)	284 (64%)
	≥10 nodes	158 (36%)	159 (36%)
Histologic grade, no. (%)			
	I	78 (18%)	76 (17%)
	II	138 (31%)	154 (33%)
	III	184 (42%)	184 (41%)
	Unknown	42 (10%)	31 (8%)
Breast Cancer subtype, no. (%)			
	ER–positive/ERBB2–negative	242 (55%)	231 (52%)
	ERBB2–positive	96 (22%)	109 (25%)
	Triple negative	68 (15%)	72 (16%)
	Unknown	36 (8%)	31 (7%)
Length of tamoxifen therapy*, no. (%)			
	Never or discontinued within 2 years, with relapse within 2 years	81 (18%)	88 (20%)
	< 2 years (without relapse within 2 years)	58 (13%)	57 (13%)
	≥ 2 and < 5 years	222 (50%)	202 (46%)
	≥ 5 years	77 (17%)	89 (20%)
	Never (without relapse within 2 years)	4 (1%)	6 (1%)
	Unknown	0 (0%)	1 (0%)
Menopausal status at randomization, no. (%)			
	Premenopausal	367 (83%)	369 (83%)
	Postmenopausal	49 (11%)	58 (13%)
	Uncertain	26 (6%)	16 (4%)
Menopausal status 30 months after randomization, no. (%)			
	Premenopausal	26 (5%)	93 (21%)
	Postmenopausal	335 (76%)	262 (59%)
	Uncertain	85 (19%)	88 (20%)

Abbreviations: ER, estrogen receptor; ERBB2, Erb-B2 receptor tyrosine kinase2; CDCT, conventional-dose chemotherapy; HDCT, high-dose chemotherapy; IQR, interquartile range.

* Data are for approximate duration of adjuvant tamoxifen. In total, 14 patients also received adjuvant aromatase inhibitor (AI) after tamoxifen: n=4 and n=7 started AI after ≥ 5 years of tamoxifen in the HDCT and CDCT group, respectively; n=2 started AI after ≥ 2 and < 5 years of tamoxifen in the CDCT group; n=1 started AI after < 2 years of tamoxifen (without relapse) in the CDCT group. Data of adjuvant AI treatment were missing in 26 and 39 patients in the HDCT and CTCT groups, respectively.

eTable 2. Distribution Number of OS Events in Breast Cancer Subtypes Split for Number of Involved Axillary Lymph Nodes and Between Treatment Groups			
Subgroups		**HDCT Group (n=442)**	**CDCT Group (n=443)**
ER-positive/ERBB2-negative, no.			
	4-9 nodes involved	160	149
	≥10 nodes involved	82	82
ERBB2-positive, no.			
	4-9 nodes involved	56	69
	≥10 nodes involved	40	40
Triple negative, no.			
	4-9 nodes involved	46	47
	≥10 nodes involved	22	25

Abbreviations: ER, estrogen receptor; ERBB2, Erb-B2 receptor tyrosine kinase 2; CDCT, conventional-dose chemotherapy; HDCT, high-dose chemotherapy.

Breast cancer subtypes were unknown for 36 patients (8%) in the HDCT group and 31 patients (7%) in the CDCT group.

eTable 3. Second Malignant Neoplasms[a]		
	HDCT Group (n=442)	**CDCT Group (n=443)**
Second malignant neoplasms – no.		
Second breast cancer	33	35
Lip, oral cavity and pharynx	3	3
Digestive organs	6	14
Respiratory organs	7	11
Bone and articular cartilage	0	2
Myelodysplasia or leukemia	4	3
Melanoma	3	1
Non–melanoma skin cancer	15	17
Mesothelial and soft tissue	0	3
Female genital cancer	6	7
Urinary tract cancer	2	2
Other or unspecified	2	0

Abbreviations: CDCT, conventional-dose chemotherapy; HDCT, high-dose chemotherapy.

[a] Data are the number of patients per type of malignancy. Note that 31 patients had more than one second primary malignancy: 14 patients in the CDCT group and 17 in the HDCT group.

eTable 4. Cardiovascular Events and Cardiovascular Risk Factors[a]			
	HDCT Group (n=442)	**CDCT Group (n=443)**	***P* Value**
Major cardiovascular events, no. (%)			
Coronary artery diseases	13 (3%)	13 (3%)	.60
Heart failure	27 (6%)	33 (8%)	.26
TIA/CVA	13 (3%)	9 (2%)	.44
Cardiac valve dysfunction	18 (4%)	12 (3%)	.57
Minor cardiovascular events, no. (%)			
Dysrhythmia [b]	36 (9%)	19 (5%)	.005
Pericarditis	4 (1%)	3 (1%)	.58
Cardiovascular Risk Factors, no. (%)			
Hypertension	91 (22%)	59 (14%)	.02
Hypercholesterolemia	66 (16%)	44 (11%)	.04
Diabetes mellitus type II	30 (7%)	18 (4%)	.14
BMI ≥25 [c]	164 (51%)	144 (46%)	.19
Smoking ever	136 (46%)	117 (45%)	.89

Abbreviations: CDCT, conventional-dose chemotherapy; HDCT, high-dose chemotherapy; BMI, body mass index; CVA, cerebrovascular accident; TIA, transient ischemic attack.

[a] P is are based on chi-square test.

[b] Including atrial fibrillation (n=15), heart block (n=7), sinus tachycardia (n=6), supraventricular extra systole (n=4), atrioventricular nodal re-entry tachycardia (n=1) and unknown (n=22).

[c] At last follow-up.

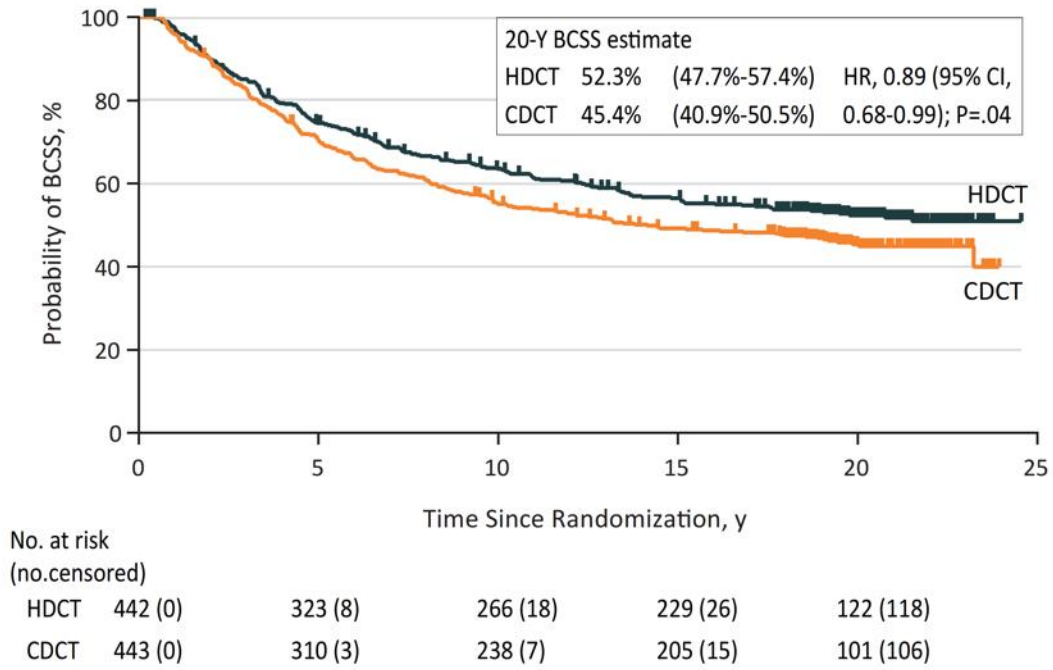
eTable 5. Distribution Number of OS Events in Breast Cancer Subtypes Split for Menopausal Status and Between Treatment Groups			
Subgroups		**HDCT Group (n=442)**	**CDCT Group (n=443)**
ER-positive/ERBB2-negative, no.			
	Premenopausal	201	185
	Postmenopausal	32	36
ERBB2-positive, no.			
	Premenopausal	81	91
	Postmenopausal	8	9
Triple negative, no.			
	Premenopausal	55	60
	Postmenopausal	11	11

Abbreviations: ER, estrogen receptor; ERBB2, Erb-B2 receptor tyrosine kinase 2; CDCT, conventional-dose chemotherapy; HDCT, high-dose chemotherapy.

Breast cancer subtypes were unknown for 36 patients (8%) in the HDCT group and 31 patients (7%) in the CDCT group. Menopausal status at baseline was unknown for 20 patients (4.7%) in the HDCT group and 22 patients (4.9%) in the CDCT group.

eFigure 1

Breast-cancer specific survival in all patients



No. at risk (no.censored)

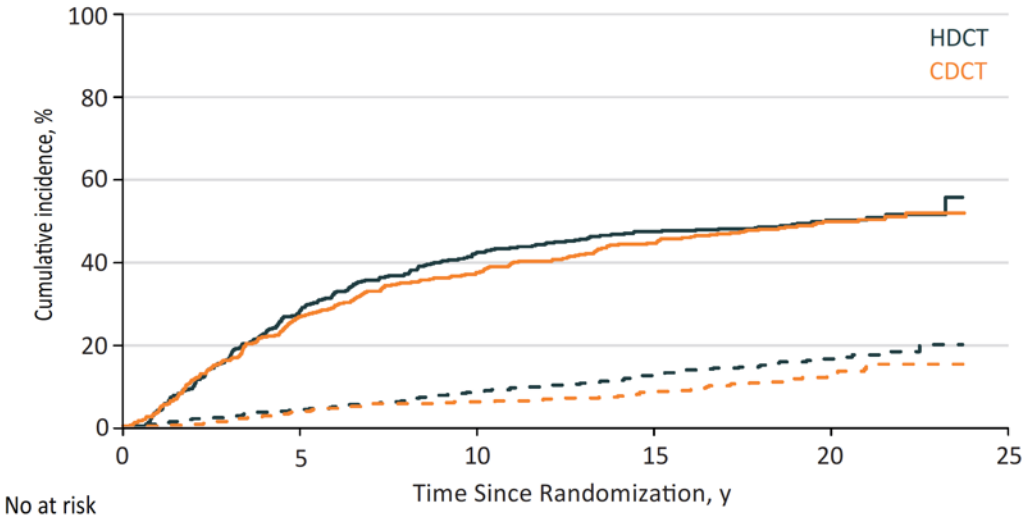
HDCT	442 (0)	323 (8)	266 (18)	229 (26)	122 (118)
CDCT	443 (0)	310 (3)	238 (7)	205 (15)	101 (106)

20-year BCSS estimates with corresponding 95% confidence interval and hazard ratios with corresponding 95% confidence interval are reported.

Abbreviations: CDCT, conventional-dose chemotherapy; HDCT, high-dose chemotherapy; HR, hazard ratio; BCSS, breast cancer-specific survival.

eFigure 2

Cumulative incidence of a second malignancy treating death as competing risk in all patients



No at risk (no.censored)					
HDCT	442 (0)	309 (0)	251 (1)	209 (1)	106 (68)
CDCT	443 (0)	302 (0)	220 (0)	179 (2)	80 (75)

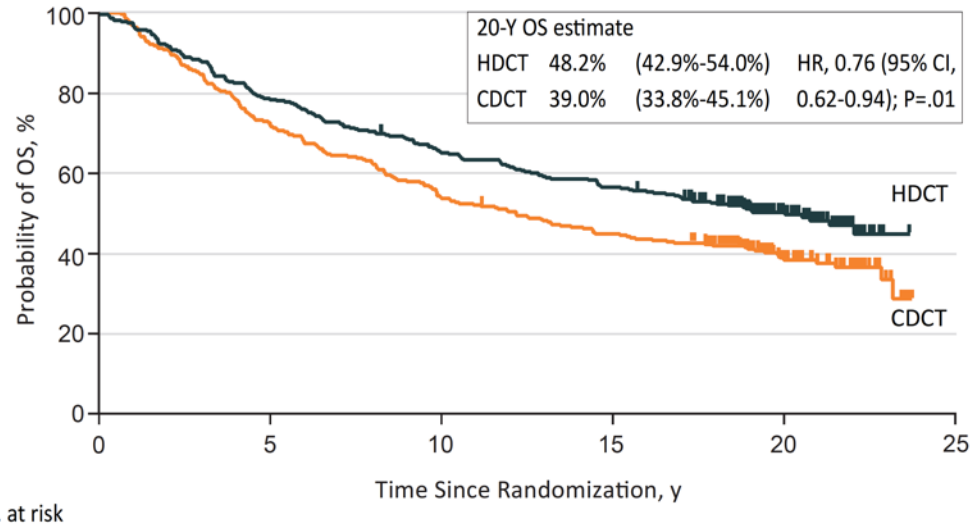
* Only the first second malignancy per patient is taken into account.

Solid lines represent cumulative incidence for death. Dashed lines represent cumulative incidence for second malignancy.

Abbreviations: CDCT, conventional-dose chemotherapy; HDCT, high-dose chemotherapy.

eFigure 3

Overall survival in patients with ERBB2-negative breast cancer



No. at risk (no.censored)

HDCT	323 (0)	252 (0)	210 (1)	182 (1)	100 (58)
CDCT	307 (0)	220 (0)	165 (0)	137 (1)	66 (57)

20-year OS estimates with corresponding 95% confidence interval and hazard ratios with corresponding 95% confidence interval are reported.

Abbreviations: ERBB2, Erb-B2 receptor tyrosine kinase 2; CDCT, conventional-dose chemotherapy; HDCT, high-dose chemotherapy; HR, hazard ratio; OS, overall survival