Outcomes Research Full Protocol Submission

The Causes and Consequences of Incomplete Paclitaxel Administration during the Neoadjuvant treatment of Early Triple negative and HER2 positive breast cancer (CIPNETH)

Principal Investigators

Full name, title	email	Cancer centre	Approval signature
Cédric VAN MARCKE, MD, PhD	Cedric.vanmarcke@saintluc.uclouvain.be	*Cliniques universitaires Saint-Luc, Brussels*	*Cedric van Marcke*
Eriseld KRASNIQI, MD	Eriseld.krasniqi@ifo.it	*Regina Elena National Cancer Institute*	*Eriseld Krasniqi*
Katarzyna POGODA, MD PhD	katarzyna.pogoda@pib-nio.pl	*Maria Sklodowska-Curie National Research Institute of Oncology, Warsaw, Poland*	*Katarzyna Pogoda*
Mahéva VALLET, M.eng, PhD	Maheva.vallet@ed.ac.uk	*University of Edinburgh/NHS Lothian*	*Maheva Vallet*
Hayley FENTON, MChem, PhD integ. MSc.	h.fenton2@nhs.net	*Leeds Teaching Hospitals Trust, Leeds, UK*	*Hayley Fenton*
Gaber PLAVC, MD	gplavc@onko-i.si	*Institute of Oncology Ljubljana, Slovenia*	
Michal UHER, MSc.	michal.uher@mou.cz	*Masaryk Memorial Cancer Institute, Brno, Czechia*	*Michal Uher*
Marina BORGES; MSc.	Marina.borges@ipoporto.min-saude.pt	*Portuguese Oncology Institute of Porto (IPO Porto), Porto, Portugal*	*Marina Borges*

Contents

1 List of Abbreviations

Abbreviation	Definition
BC	Breast Cancer
CDM	Common Data Model
EBC	Early stage Breast Cancer
DPO	Data Protection Officers
IBCFS	Invasive breast cancer-free survival
IHC	Immunohistochemistry
IPO Porto	Portuguese Oncology Institute of Porto
ISH	In-situ Hybridization
MMCI	Masaryk Memorial Cancer Institute
OS	Overall Survival
pCR	Pathological Complete Response
TBD	To Be Determined
TNBC	Triple Negative Breast Cancer

2 Abstract

Section	Description
Title	The Causes and Consequences of Incomplete Paclitaxel administration during the Neoadjuvant treatment of Early Triple negative and HER2-positive breast cancer (CIPNETH)
Rationale and Background	The backbone of neoadjuvant chemotherapy for early breast cancer (eBC) is the sequential administration of anthracyclines and taxanes. The administration of the initially planned dose-intensity of paclitaxel is frequently hampered by side effects, mainly chemotherapy-induced peripheral neuropathy. Importantly, there is no established strategy to treat or prevent this side effect. Despite these facts, no previous study has assessed the effects of reduced paclitaxel dose-intensity administration.
Research Questions and Objectives	Using a retrospective cohort of TNBC and HER2-positive EBC treated in the neoadjuvant setting with anthracyclines-cyclophosphamide and weekly paclitaxel or carboplatin-paclitaxel (with or without trastuzumab ± pertuzumab) collected at 8 different cancer centres across Europe, **we aim to**: **(Primary Objective 1)** highlight a potential impact of reduced paclitaxel dose-intensity on treatment effect (pCR rate at post-neoadjuvant surgery) and invasive breast cancer-free survival), **(Secondary Objective 2)** Estimate if this impact differs according to the breast cancer subtype, **(Secondary Objective 3)** Characterize TNBC/HER2+ early Breast Cancer Patients treated in the neoadjuvant setting with anthracyclines-cyclophosphamide and subsequent paclitaxel with respect to demographics and clinical characteristics at diagnosis index date and

	treatment received. **(Exploratory Objective 4)** highlight a potential impact of reduced paclitaxel dose-intensity on overall survival **(Exploratory Objective 5)** Characterize the clinical factors and side effects associated with early cessation or dose-intensity reduction of paclitaxel administration. **(Exploratory Objective 6)** Assess and quantify the frequency of early cessation and dose reduction of paclitaxel administration in patients presenting with treatment-induced neuropathy. The homogeneity of the treatment schedules across the different centres and the binary readout of the efficacy outcome (pCR at post-neoadjuvant surgery) will allow us to perform a reliable analysis of the datasets.
Study Design	In this retrospective study, we will collect data from all patients with early-stage TNBC or HER2-positive BC treated between 2018 and 2021 with a common neoadjuvant chemotherapy regimen at 8 different cancer centres.
Setting	Patients treated at 8 European centres will be included in this study. The overall study period will be from 01-January-2018 to 31-Jun-2022. The inclusion and exclusion criteria are: <u>Inclusion criteria</u> - Confirmed diagnosis of Stage I-III invasive breast cancer from 1st January 2018 up to 31st December 2021 - TNBC or HER2 positive BC - Female and male aged ≥ 18 years at the time of breast cancer diagnosis - Treated in the neoadjuvant setting with anthracyclines-cyclophosphamide (dose dense schedule or not) and weekly paclitaxel or carboplatin-paclitaxel, with or without trastuzumab ± pertuzumab - That underwent breast and axillary surgery with curative intent after the neoadjuvant systemic therapy - For which pathological report is available at diagnosis and at surgery - And for which data is available regarding each treatment cycle (schedule and administration) <u>Exclusion criteria</u> - Patients with a history of another primary malignancy, except for non-melanoma skin cancer or in situ carcinoma - Patients with bilateral invasive breast cancer (contralateral in situ carcinoma is allowed) - Administration of another anti-cancer systemic treatment (including investigational product and endocrine therapy) in the neoadjuvant setting. Use of LHRH agonists is allowed - Pregnancy at the time of breast cancer diagnosis or during the

	neoadjuvant treatment The estimated cohort size is 780 and 340 for TNBC and HER2 positive BC, respectively.
Link to the CDM?	The CDM will include (1) variables describing the population (patient's characteristics, comorbidities, tumour characteristics); (2) variables describing the treatments (details regarding the neoadjuvant chemotherapy, the surgery and the adjuvant therapies); and (3) variables describing the follow-up and toxicity during neoadjuvant chemotherapy. The complete CDM is in a separate file.
Databases	Cliniques universitaires Saint-Luc, Belgium Masaryk Memorial Cancer Institute, Czechia University of Edinburgh, Scotland Institute of Oncology Ljubljana, Slovenia Maria Sklodowska-Curie National Reseach Institute of Oncology, Poland IPO Porto, Portugal Leeds Teaching Hospital Trust, UK Regina Elena National Cancer Institute, Italy
Data Analysis	Patient-level data from each site will be anonymized and pooled at Masaryk Memorial Cancer Institute. Patient's characteristics will be summarized by descriptive statistics. Rate of pCR, as well as OS and IDFS data will be presented overall, by paclitaxel dose-intensity and by incidence of peripheral neuropathy. Logistic regression modeling will allow to estimate the impact of paclitaxel dose-intensity on pCR and the association between peripheral neuropathy and reduced paclitaxel dose-intensity.

3 Milestones

Milestone	Planned date
Protocol finalization (after LAB assessment)	15 Nov 2022
Other study documentation finalization	15 Nov 2022
DPO assessment (documents sent)	30 Nov 2022
Ethics Committee assessment (documents sent)	30 Nov 2022
Start of data collection	15 Jan 2023
End of data collection (including quality control)	30 Apr 2023
Data analysis	31 May 2023
Final report of study results	30 Jun 2023
First scientific output	Consideration for SABCS 23, ESMO 23 or ASCO 24
Article publication	TBD

4 Study Updates Since First Submission

We evaluated the possibility of performing data pooling; provided clear definitions of invasive breast cancer-free survival and overall survival; prioritized data and outcomes regarding paclitaxel dose-intensity over neuropathy.

5 Introduction

5.1 Rationale and background

Breast cancer (BC) is the most commonly diagnosed cancer and the leading cause of cancer-related deaths among women worldwide. Most patients are diagnosed with early-stage BC (EBC) in which the aim of treatment is to increase survival rates by reducing the risk of metastasis occurrence.

Despite tremendous efforts in the last decades, around 20-30% of patients with EBC still relapse months or years later and progress towards incurable, advanced disease. Patients with EBC considered at high risk of recurrence (one or more negative clinico-pathological prognostic features) are advised to receive chemotherapy, in order to improve long-term outcomes. Its administration in the neoadjuvant setting may allow a more conservative surgery, defy micro-metastatic disease in an early phase, and permit to assess treatment efficacy on the primary tumor, which has a prognostic value: pathological complete response (pCR) results in improved long-term outcomes, whereas significant residual disease leads to a significantly increased risk of recurrence (1).

The standard regimen, both in the neoadjuvant and adjuvant settings, consists of the sequential administration of anthracyclines and cyclophosphamide, preceded or followed by taxanes, with trastuzumab ± pertuzumab in case of HER2 overexpression (2). The addition of a taxane (an antimicrotubular agent) to the anthracycline regimen in the adjuvant setting leads to a significantly improved rate of disease-free survival, whereas studies in the neoadjuvant setting showed increased response rates and eligibility for breast conserving surgery, and decreased recurrence rates (3-5). The weekly administration of 12 cycles of paclitaxel is considered the standard of care due to better therapeutic characteristics (6). Furthermore, in triple negative breast cancer (TNBC), a recently published meta-analysis indicated better event-free survival when combining a platinum agent with the taxane (7). The concurrent administration of trastuzumab and pertuzumab with taxanes provides further clinical benefit to patients with HER2-positive EBC, increasing rates of pCR and disease-free survival (1,8). Several studies demonstrated that the activity between taxanes and those monoclonal antibodies targeting HER2 is synergistic (9,10).

Toxicity of weekly paclitaxel administration remains a major concern, with up to 70% of patients developing peripheral neuropathy, of which up to 30% are grade ≥ 2 (6,11). Furthermore, moderate to severe peripheral neuropathy may regress only slowly, or even persist for months to years, especially if the toxic compound's administration is pursued (12). Peripheral neuropathy is thus a main dose-limiting side effect of paclitaxel and a frequent reason for early cessation of chemotherapy administration. The other causes of paclitaxel discontinuation are neutropenia, allergic symptoms and liver toxicity. Overall, cohort studies suggest that dose reduction or discontinuation is required in 30 to 40% of patients, 30 to 50% of them due to peripheral neuropathy (13-14).

To the best of our knowledge, no study assessed the impact on treatment efficacy of shorter durations or early cessation of paclitaxel.

5.2 Research Question(s) and Objectives

Using a retrospective cohort of TNBC and HER2-positive EBC treated in the neoadjuvant setting with anthracyclines-cyclophosphamide and weekly paclitaxel or carboplatin-paclitaxel (with or without trastuzumab ± pertuzumab) collected at 8 different cancer centres across Europe, **we aim to**:

(Primary Objective 1) highlight a potential impact of reduced paclitaxel dose-intensity on treatment effect (pCR rate at post-neoadjuvant surgery) and invasive breast cancer-free survival),

(Secondary Objective 2) Estimate if this impact differs according to the breast cancer subtype,

(Secondary Objective 3) Characterize TNBC/HER2+ EBC patients treated in the neoadjuvant setting with anthracyclines-cyclophosphamide and subsequent paclitaxel with respect to main demographics and clinical characteristics at diagnosis index date and treatment received,

(Exploratory Objective 4) highlight a potential impact of reduced paclitaxel dose-intensity on overall survival,

(Exploratory Objective 5) Characterize the clinical factors and side effects associated with early cessation or dose-intensity reduction of paclitaxel administration,

(Exploratory Objective 6) Assess and quantify the frequency of early cessation and dose reduction of paclitaxel administration in patients presenting treatment-induced neuropathy.

The homogeneity of the treatment schedules across the different centres and the binary readout of the efficacy outcome (pCR at post-neoadjuvant surgery) will allow us to perform a reliable analysis of the datasets.

References

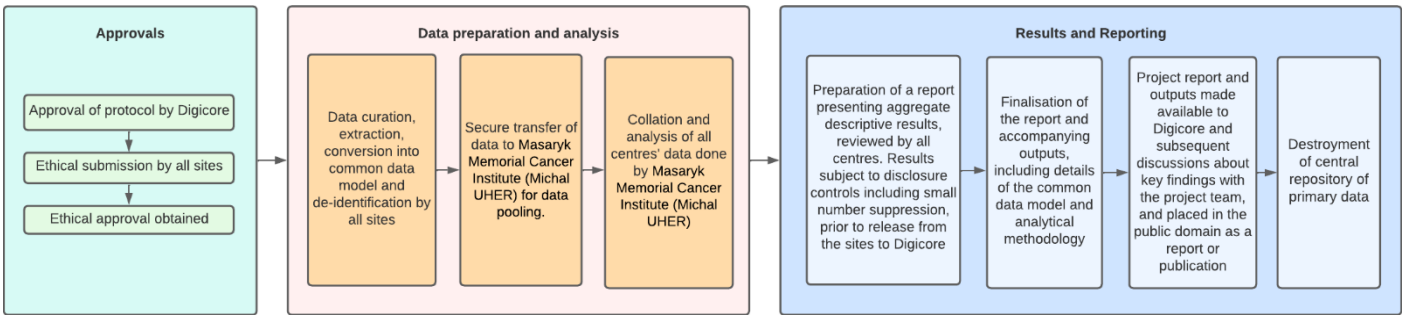
(1) Cortazar P et al. The Lancet. 2014;384(9938):164-72
(2) Anampa J et al. BMC Med. 2015;13(1):195
(3) Mamounas EP et al. J Clin Oncol. 2005;23(16):3686-96
(4) Bear HD et al. J Clin Oncol. 2006;24(13):2019-27
(5) Sparano JA et al. J Clin Oncol. 2015;33(21):2353-60
(6) Sparano JA et al. NEJM. 2008;358(16):1663-71
(7) Poggio F et al. Ann Oncol. 2022;33(3):347-349
(8) Untch M et al. J Clin Oncol. 2010;28(12):2024-31
(9) Perez EA et al. J Clin Oncol. 2011;29(34):4491–4497
(10) Di Modica M et al. Oncotarget. 2016;7(1): 255–265
(11) Staff NP et al. Exp Neurol. 2020;324:113121
(12) Seretny M et al. Pain. 2014;155(12):2461-2470
(13) Bhatnagar B et al. Springerplus. 2014;3:366
(14) Nyrop K et al. Cancer. 2019;125(17):2945-2954

6 Research Methods

6.1 Study Design

The study is a multi-national, retrospective, observational, real-world cohort study that includes patients with early-stage triple negative (TNBC) or HER2-positive breast cancer, treated in the neoadjuvant setting with anthracyclines-cyclophosphamide and weekly paclitaxel or carboplatin-paclitaxel. Only anonymised or pseudonymised data will be analysed.

General study methods:



6.2 Setting

6.2.1 Data sources

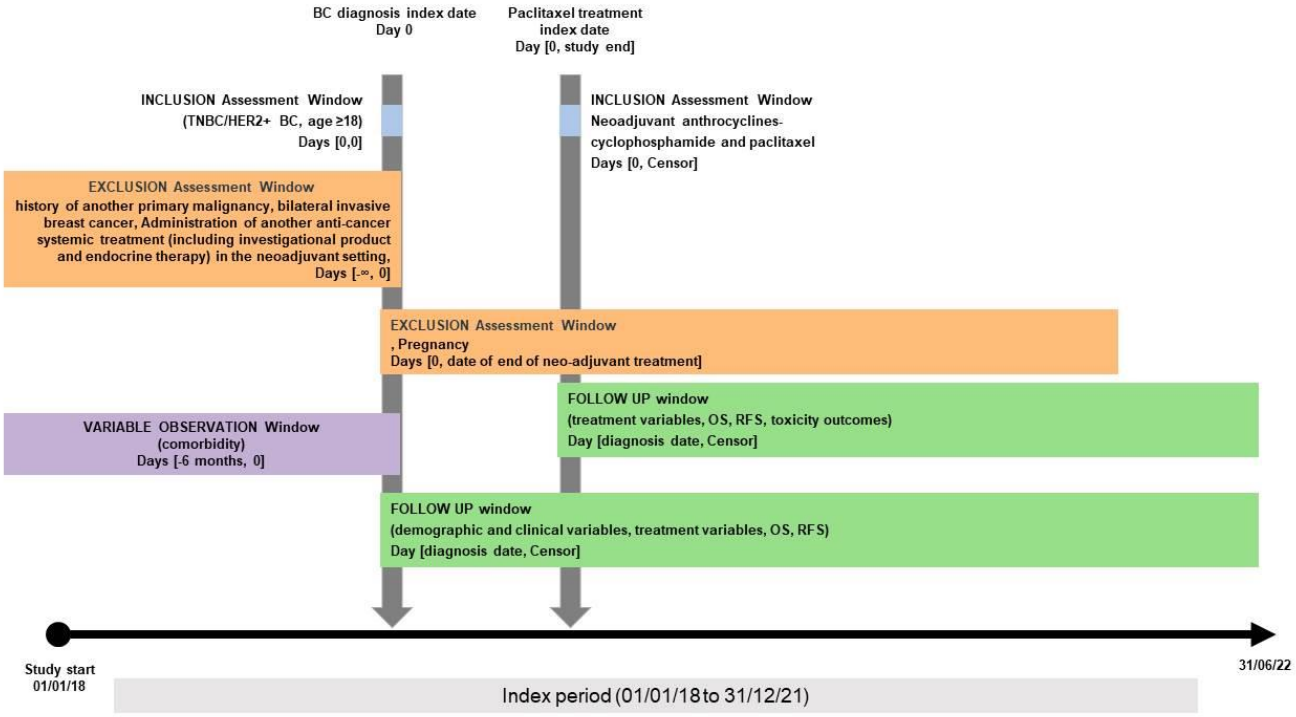
The study sites for inclusion are: Cliniques universitaires Saint-Luc; Regina Elena National Cancer Institute; University of Edinburgh; Maria Sklodowska-Curie National Research Institute of Oncology; Portuguese Oncology Institute of Porto; Institute of Oncology Ljubljana; Leeds Teaching Hospitals Trust; Masaryk Memorial Cancer Institute

- Edinburgh – Data specified fields will be extracted and stored on NHS Lothian secure servers and analysed by NHS Data Analytical staff within the NHS computing environment accessible by the Cancer Information Team (access already in place). The different data sources will be linked through patient identifier (Community Health Index - CHIs), before being de-identified for analysis. The data will be evaluated by an independent member of the Edinburgh Cancer Centre Cancer Information Team to confirm the data adheres to the specifications outlined in this application. Additional details about our datasets can be found here, and will include South East Scotland Database (SESCD), SMR06, SMR01, SMR00, ChemoCare, Prescribing Information System (PIS), National Records for Scotland Death Registrations and SCAN QPI audit.
- Leeds – Data specified fields will be extracted from the PPM system, the EMR in place at Leeds Teaching Hospital Trust using R4.0.3 or later and stored securely on the Leeds Teaching Hospital Trust server and analysed by the Real-Oncology team who are either employed by/have an Honorary Contract with Leeds Teaching Hospitals Trust. The different data sources will be linked via a PPM ID before being de-identified for analysis. Patients who have opted out of data sharing will not be included in the extract, which will be evaluated by the Information Governance Team prior to release for final analysis.
- Warsaw – Data fields will be extracted from the internal computer system in place at the Maria Sklodowska-Curie National Research Institute of Oncology including data from the software of the Pharmacy of the Institute and stored securely on the internal server and analysed by the collaborating staff (KP, MR, PH, MK) who are all employed by the Maria Sklodowska-Curie National Research Institute of Oncology.
- Ljubljana – Data specified fields will be extracted from the EMR in place at the Institute of Oncology Ljubljana, from the data software of the Pharmacy of the Institute of Oncology Ljubljana, and from the

National Cancer Registry of Slovenia and stored securely on the Institute of Oncology Ljubljana server and analysed by the collaborating staff (GP, AK, TŽ) who are all employed by the Institute of Oncology Ljubljana. The different data sources will be linked before being de-identified for analysis.

- Saint-Luc, Brussels – For this retrospective review, we will in one round of submission ask the ethics committee for waiver of informed consent and answer the questions regarding RGPD. Sending of de-identified data inside the European Union for non-commercial use will simultaneously be notified to the DPO. Data specified fields will be extracted from the EMR and the Pharmacy server and stored in a specifically designed RedCap form on an internal server. Data will be de-identified on RedCap.
- IPO Porto - Data Protection Impact Assessment is needed for submission to our Data Protection Officer (DPO); then the study needs to be submitted to the local Ethics Committee, including a request for Informed Consent waiver. Multiple datasets from internal IT systems will be combined to create a simple multi-table structure. Additionally, where applicable to the study, free text data from EMR will be collected and converted to structured variables. Data will be pseudonymised before being transferred for Masaryk Memorial Cancer Institute.
- Masaryk Memorial Cancer Institute - The available data sources are: the EMR in place at the Masaryk Memorial Cancer Institute, the local administrative claims database and selected parts of the National Health Information System (specifically the National Cancer Registry and the National Registry of Paid Health Services). Data specified fields will be extracted from the data sources, linked via a personal ID, anonymised and stored securely on MMCI's internal server. Data analysis will be performed by the research team who are all employed by MMCI.

6.2.2 Study time period



The study time period will be from the 1st January 2018 to 31st June 2022 (allowing six months for follow-up of treatment and outcomes post-diagnosis). The index period will be from the 1st January 2018 to 31st December 2021. The look back period will be 6 months.

6.2.3 Index date

The index date for the study will be the disease diagnosis index date: the date of diagnosis of early Breast Cancer

6.2.4 Follow-up period and censoring

Patient follow-up is defined as the time that accrues between index date (inclusive) and cohort exit (exclusive). Cohort exit is defined as:

- Date of death
- Date of loss to follow-up (e.g. emigration), defined as the last vital status date or date of last information available in patient record
- Study end date

6.2.5 Case definition

The study cohort will include all adult (≥18 years) patients with early stage or locally advanced triple negative/HER2+ breast cancer. In this study, early stage or locally advanced BC is defined as stage I-III breast cancer (ICD-10 codes in Table 1, Staging information in Table 2). Biological subtype information is given in Table 3.

Table 1: ICD-10 diagnostic codes for patient inclusion

ICD-10	Description
C50.0	Malignant neoplasm of nipple and areola
C50.1	Malignant neoplasm of central portion of breast
C50.2	Malignant neoplasm of upper-inner quadrant of breast
C50.3	Malignant neoplasm of lower-inner quadrant of breast
C50.4	Malignant neoplasm of upper-outer quadrant of breast
C50.5	Malignant neoplasm of lower-outer quadrant of breast
C50.6	Malignant neoplasm of axillary tail of breast

C50.8	Malignant neoplasm of overlapping sites of breast
C50.9	Malignant neoplasm of breast of unspecified site

Table 2: UICC 8th Edition Stage and TNM mapping

Overall Stage	T Category	N Category	M Category
Stage I	T1	N0	M0
Stage IIA	T0	N1	M0
	T1	N1	M0
	T2	N0	M0
Stage IIB	T2	N1	M0
	T3	N0	M0
Stage IIIA	T0	N2	M0
	T1	N2	M0
	T2	N2	M0
	T3	N1	M0
	T3	N2	M0
Stage IIIB	T4	Any N	M0
Stage IIIC	Any T	N3	M0

Table 3: Definition of TNBC and HER2+ BC

	Estrogen receptor status	Progesterone receptor status	HER2 status
TNBC BC	Negative	Negative	Negative, defined as negative ISH test or IHC status of 0 or 1+ as per local laboratory testing
HER2+ BC	Any	Any	Positive, defined as IHC status of 3+, or 2+ and positive ISH as per local laboratory testing

6.2.6 Patient selection

To address the study objectives, patients who meet the following eligibility criteria will be included in the study, unless they meet any of the exclusion criteria:

6.2.6.1 Inclusion criteria

- Confirmed diagnosis of Stage I-III invasive breast cancer from 1st January 2018 up to 31st December 2021
- Patient has TNBC or HER2 positive BC
- Female and male aged ≥ 18 years at the time of breast cancer diagnosis
- Treated in the neoadjuvant setting with anthracyclines-cyclophosphamide (dose dense schedule or not) and weekly paclitaxel or carboplatin-paclitaxel, with or without trastuzumab ± pertuzumab
- That underwent breast and axillary surgery with curative intent after the neoadjuvant systemic therapy
- For which pathological report is available at diagnosis and at surgery
- And for which data is available regarding each treatment cycle (schedule and administration)

6.2.6.2 Exclusion criteria

- Patients with a history of another primary malignancy, except for non-melanoma skin cancer or *in situ* carcinoma
- Patients with bilateral invasive breast cancer (contralateral in situ carcinoma is allowed)

- Administration of another anti-cancer systemic treatment (including investigational product and endocrine therapy) in the neoadjuvant setting. Use of LHRH agonists is allowed
- Pregnancy at the time of breast cancer diagnosis or during the neoadjuvant treatment

6.2.7 Analytical sub-populations

Key subgroups will be the BC subtypes (TNBC vs HER2-positive); paclitaxel dose-intensity (full dose vs reduced dose); and incidence or not of chemotherapy-induced peripheral neuropathy.

6.3 Variables

6.3.1 Treatment events

The variables used to derive and describe surgical and SACT treatment events are defined in the accompanying CDM. The following variables will be presented as outputs.

Table 4: Variables describing surgical events

Variable	*Description*	*Summary Statistics*
Mastectomy	Variable to describe whether patient underwent mastectomy including total mastectomy, radical mastectomy. Relevant example OPCS-4 codes can be found in the CDM. 0 = procedure not performed 1 = procedure performed	*N, %*
Wide Local Excision	Variable to describe whether patient underwent wide local excision, including segmental/partial mastectomy, lumpectomy. Relevant example OPCS-4 codes can be found in the CDM. 0 = procedure not performed 1 = procedure performed	*N, %*
Axillary Lymph Node Dissection	Relevant example OPCS-4 codes can be found in the CDM. *0 = procedure not performed* *1 = procedure performed*	*N, %*
Targeted Axillary Dissection	Relevant example OPCS-4 codes can be found in the CDM. *0 = procedure not performed* *1 = procedure performed*	*N, %*
Sentinel Lymph Node Biopsy	Relevant example OPCS-4 codes can be found in the CDM. *0 = procedure not performed* *1 = procedure performed*	*N, %*

Table 5: Variables describing SACT treatment data

Variable	Description	Summary Statistics
Neoadjuvant treatment		
Duration of anthracycline-cyclophosphamide treatment	Number of days from first administration of anthracycline-cyclophosphamide treatment until date of last administration during neoadjuvant chemotherapy	*N, %, Mean (SD), Median (IQR), Minimum-Maximum*
Number of Anthracyclines-Cyclophosphamide Cycles	Number of individual treatment administrations during neoadjuvant chemotherapy	*N, %, Mean (SD), Median (IQR), Minimum-Maximum*
Total dose of doxorubicin	Total cumulative dose of doxorubicin during neoadjuvant chemotherapy in milligrams (mg/m^2).	*N, %, Mean (SD), Median (IQR), Minimum-Maximum*
Total dose of epirubicin	Total cumulative dose of epirubicin during neoadjuvant chemotherapy in milligrams (mg/m^2).	*N, %, Mean (SD), Median (IQR), Minimum-Maximum*
Total dose of cyclophosphamide	Total cumulative dose of cyclophosphamide during neoadjuvant chemotherapy in milligrams (mg/m^2).	*N, %, Mean (SD), Median (IQR), Minimum-Maximum*
Duration of paclitaxel treatment	Number of days from first administration of paclitaxel treatment until date of last administration during neoadjuvant chemotherapy	*N, %, Mean (SD), Median (IQR), Minimum-Maximum*
Total dose of paclitaxel	Total cumulative dose of paclitaxel during neoadjuvant chemotherapy in milligrams (mg/m^2).	*N, %, Mean (SD), Median (IQR), Minimum-Maximum*
Number of Paclitaxel Cycles	Number of individual treatment administrations during neoadjuvant chemotherapy	*N, %, Mean (SD), Median (IQR), Minimum-Maximum*
Reduction in dose of Paclitaxel	0 = no dose reduction 1 = dose reduction	*N, %*
Dose Reduction of Neoadjuvant Paclitaxel due to Neuropathy	0 = No 1 = Yes, NA = Missing/Unknown	*N, %*
Dose Reduction of Neoadjuvant Paclitaxel due to Neutropenia	0 = No 1 = Yes, NA = Missing/Unknown	*N, %*
Dose Reduction of Neoadjuvant Paclitaxel due to Liver Toxicity	0 = No 1 = Yes, NA = Missing/Unknown	*N, %*
Dose Reduction of Neoadjuvant Paclitaxel due to Other Reason	0 = No 1 = Yes, NA = Missing/Unknown	*N, %*
Early cessation of neoadjuvant paclitaxel	0 = completed treatment 1 = early cessation	*N, %*

Early Cessation of Neoadjuvant Paclitaxel due to Neuropathy	0 = No 1 = Yes, NA = Missing/Unknown	*N, %*
Early Cessation of Neoadjuvant Paclitaxel due to Neutropenia	0 = No 1 = Yes, NA = Missing/Unknown	*N, %*
Early Cessation of Neoadjuvant Paclitaxel due to Liver Toxicity	0 = No 1 = Yes, NA = Missing/Unknown	*N, %*
Early Cessation of Neoadjuvant Paclitaxel due to Allergy	0 = No 1 = Yes, NA = Missing/Unknown	*N, %*
Early Cessation of Neoadjuvant Paclitaxel due to Disease Progression	0 = No 1 = Yes, NA = Missing/Unknown	*N, %*
Early Cessation of Neoadjuvant Paclitaxel due to other reason	0 = No 1 = Yes, NA = Missing/Unknown	*N, %*
Neoadjuvant Carboplatin	0 = not received 1 = received	*N, %*
Neoadjuvant Trastuzumab without Pertuzumab	0 = not received 1 = received	*N, %*
Neoadjuvant Trastuzumab and Pertuzumab	0 = not received 1 = received	*N, %*
Adjuvant treatment		
Adjuvant Capecitabine	0 = not received 1 = received	*N, %*
Adjuvant Trastuzumab Emtansine	0 = not received 1 = received	*N, %*
Adjuvant trastuzumab without pertuzumab	0 = not received 1 = received	*N, %*
Adjuvant Trastuzumab + Pertuzumab	0 = not received 1 = received	*N, %*
Adjuvant radiotherapy	0 = not received 1 = received	*N, %*
Adjuvant endocrine therapy	0 = not received 1 = received	*N, %*

6.3.2 Outcomes

Table 6: Time to event outcomes

Outcome	*Index Date*	*Event*	*Censoring events*	*Key statistics of interest*
Overall survival from diagnosis	*Diagnosis index date*	*Date of death*	*Earliest of:* • *End of study period* • *Date of loss to follow up*	Number of patients, Number of deaths, Number of censored patients, Median survival with IQR. Also the Number of patients at risk and Survival probability with 95%CI at 3, 6, 9, 12, 15, 18, 21 and 24 months post specified index date

Invasive breast cancer-free survival from diagnosis	*Diagnosis index date*	Earliest of: • Next recorded recurrence event • Date of death	Earliest of: • End of study period • Date of loss to follow up	Number of patients, Number of events, Number of censored patients, Median DFS with IQR. Also the Number of patients at risk and Survival probability with 95%CI at 3, 6, 9, 12, 15, 18, 21 and 24 months post specified index date

Table 7: Other outcome variables

Variable	***Definition***	***Summary Statistics***
Pathological Complete Response	Secondarily derived from pT and pN classification and defined as pT0/is and pN0, meaning the absence of residual invasive cancer on hematoxylin and eosin evaluation of the complete resected breast specimen and all sampled regional lymph nodes after neoadjuvant systemic therapy *1 = Yes* *0 = No* *NA = Missing/Unknown*	*N, %*
Neuropathy due to paclitaxel	*1 = Yes* *0 = No* *NA = Missing/Unknown*	*N, %*
Maximum grade of neuropathy	*WHO definition* *1 =* paresthesias (a tingling, tickling or prickling sensation) and/or decreased tendon reflexes *2 =* severe paresthesias and/or mild weakness *3 = i*ntolerable paresthesias and/or marked motor loss *4 = paralysis* *NA = Missing/Unknown*	*N, %*
Neutropenia G3-4 during paclitaxel	*Neutropenia grades 3 to 4 (neutrophil count <1000) due to paclitaxel* *1 = Yes* *0 = No* *NA = Missing/Unknown*	*N, %*
Allergic reaction during paclitaxel	*1 = Yes* *0 = No* *NA = Missing/Unknown*	*N, %*
Liver toxicity during paclitaxel	*1 = Yes*	*N, %*

	0 = No *NA = Missing/Unknown*	

6.3.3 Other variables

Demographic and clinical characteristics will be presented using descriptive statistics. The variables required are listed in Table 8. All other variables are listed in the CDM:

Table 8: Demographic and clinical characteristics

Variable	Definition	Summary Statistics
Age categories at diagnosis index date	Derived as above Age (in years) in categories: 18-24; 25-44; 45-64;65-79; 80+	N, %
Gender	1=Female 2=Male 9=Other NA=Missing/Unknown	N, %
Menopausal Status	0=Pre-menopausal 1=Peri/Post-menopausal NA=Missing/Unknown	N, %
ECOG Performance Status	Most recent record prior to or on diagnosis index date 0 = Fully active, able to carry on all pre-disease performance without restriction. Equivalent to Karnofsky Performance status of 90 or 100. 1 = Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light house work, office work. Equivalent to Karnofsky Performance Status of 70 or 80. 2 = Ambulatory and capable of all self-care but unable to carry out any work activities. Up and about more than 50% of waking hours. Equivalent to Karnofsky Performance Status of 50 or 60. 3 = Capable of only limited self-care, confined to bed or chair for more than 50% of waking hours. Equivalent to Karnofsky Performance Status of 30 or 40. 4 = Completely disabled. Cannot carry on any self-care. Totally confined to bed or chair. Equivalent to Karnofsky Performance Status of 10 or 20. NA = Missing/Unknown	N, %
Body Surface Area	Continuous variable derived from the	N, %, Mean (SD), Median (IQR),

	most recent record of height and weight of patient, prior to or on 1st Line of Therapy. May be Missing/Unknown	Minimum-Maximum
Comorbidities	One variable for each of the following comorbidities listed below, describing whether comorbidity is present at any point in the six months prior to diagnosis index date. (Example ICD-10 codes in Comorbidities tab of CDM): Peripheral Vascular Disease Cerebrovascular Disease Dementia Diabetes Congestive Heart Failure Liver Disease Active Smoker Previous or present alcohol abuse Myocardial infarction Variable coded as: 1 = Present 0 = Absent NA = Missing/Unknown	N, %
Stage	I IIA IIB IIIA IIIB IIIC	N, %
Mode of presentation	Mode of presentation at diagnosis index date: 1 = Screening 2 = Symptomatic 9 = Other NA = Missing/Unknown	N, %
Location	Location of primary breast cancer. Use ICD-10 codes and descriptors in Table 1	N, %
Laterality	Laterality of primary bresat cancer. 1 = right/dex 2 = left/sin 9 = other NA = Missing/Unknown	N, %
Histological type	1 = invasive ductal carcinoma; no special type) 2 = any special type 9 = other NA = Missing/Unknown	N, %
Grade	Coded using ICD-O-3 grade/differentiation codes 1 2	N, %

	3 NA = Missing/Unknown	
Inflammatory breast cancer	Secondarily derived from T stage and defined as T4d. 1 = Yes 0 = No NA = Missing/Unknown	N, %
Ki-67 Marker of Proliferation	The Ki-67 percentage score is the percentage of positively stained tumor cells among the total number of malignant cells assessed. 1 = <20 2 = ≥20 - <50 3 = ≥ 50 NA = Missing/Unknown	
BRCA germline mutation	0 = both BRCA1 and BRCA2 negative, 1 = BRCA1 or BRCA2 class 4 or 5 mutation positive NA = Missing/Unknown	N, %
Biological Subtype	Derived from ER/PR/HER2 status 1 =TNBC (negative estrogen/progesterone receptors as per local laboratory testing and negative HER2 defined as negative ISH test or IHC status of 0 or 1+ as per local laboratory testing) 2 = Luminal HER2+ (HER2 IHC status of 3+ or 2+ and positive ISH as per local laboratory testing, estrogen receptor positive status as per local laboratory testing) 3 = Non luminal HER2+ (HER2 IHC status of 3+ or 2+ and positive ISH as per local laboratory testing, and estrogen receptor status negative as per local laboratory testing)	N, %

6.3.4 Possible confounders

The logistic regression models will be adjusted on potential confounding factors at baseline (age, breast cancer subtype, menopausal status, tumor size, nodal status, tumor grade, anthracycline regimen (dose dense vs standard), therapy with other agents: carboplatin, pertuzumab, presence of a chemotherapy-induced side effect (neuropathy, grade 3-4 neutropenia, allergic symptoms, liver toxicity), and presence of a main comorbidity (diabetes, heart failure, vascular or liver disease).

6.4 Data Management

The process will take into account any data governance imposed on the data source including any plans to handle the data outside of the institution or country of origin. A data management plan will be created before data collection begins and will describe all functions, processes and specifications for data collection, cleaning and validation.

Data Cleaning

Data cleaning will be carried out by the sites before the attrition step to remove patients with missing variables in fields necessary for analysis. In particular, the following incomplete records will be removed:

- Patients without recorded age and gender information
- Patients with missing vital status information

Data masking

For the harmonised data analysis, the data will be pooled and then masked at a site-level and for the overall aggregated outputs. The following masking criteria apply at each site.

Table 9: Requirement for suppression of small numbers at each centre

	Scotland	Portugal	England	Slovenia	Belgium	Czech Republic	Poland	Italy
Numbers requiring suppression	<10	<5	<6	NA	<10	0	NA	<10

6.5 Study Size

The planned cohort size for this study is approximately **780 TNBC cases and 340 HER2+ cases**, with the following distribution by participating centre:

Table 10: Cohort size at each centre

	Scotland	Portugal	England	Slovenia	Belgium	Czech Republic	Poland	Italy
TNBC	60	160	40	40	120	90	200	70
HER2+	20	5 (*)	10	40	70	90	5	100

(*) Because we have only 5 patients, we will not include them in the study.

Based on our own clinician experience, it is estimated that up to 50% of patients, both TNBC and HER2+ will have pCR. It is estimated that for TNBC patients, a reduction in dose-intensity may lead to a smaller proportion of patients with pCR (40% or less). For HER2+ patients, we hypothesize a reduction in dose intensity will not lead to a reduction in pCR, given the high sensitivity to trastuzumab/pertuzumab.

An a priori power analysis was conducted using the pwr package in R, for a two-sample proportion test, with a small effect size (d= .15), and an alpha of .05. Results showed that with the expected total sample of 780 TNBC patients, a power of .842 could be achieved. For the same analysis on the expected total sample of 340 HER2+ patients, a power of .498 could be achieved.

6.6 Data Analysis

General Approach

All analysis will be conducted using R Studio version 4.0.3 or later. Descriptive statistics will be generated for all study variables. The descriptive statistics will include mean, standard deviation, median, minimum, maximum, first and third quartile values for continuous variables (e.g. age at diagnosis) and frequencies and percentages for categorical variables (e.g. ECOG performance status). Missing data will be categorized and reported in statistical tables and the final study report. All time-to-event analyses will be depicted graphically by KM survival curves with number of patients still at risk and their probabilities of remaining event-free with 95% CIs at 3, 6, 9,

12, 15, 18, 21 and 24 months post their specified index date reported. KM survival curves will be compared by the Log-rank test. A study attrition table will be presented as per Table X.

Table X: Attrition Table

Attrition Criteria	N (%)
Inclusion Criteria	
A. Confirmed diagnosis of Stage I-III invasive breast cancer during the study period (from 1st January 2018 up to 31st December 2021)	N (100%)
B. Patient has TNBC (negative estrogen/progesterone receptors as per local laboratory testing and negative HER2 defined as negative ISH test or IHC status of 0 or 1+ as per local laboratory testing) or HER2 positive BC (defined as IHC status of 3+ or 2+ and positive ISH as per local laboratory testing, irrespective of the estrogen receptor status)	N (% of A)
C. Female and male aged ≥ 18 years at the time of breast cancer diagnosis	N (% of B)
D. Treated in the neoadjuvant setting with anthracyclines-cyclophosphamide (dose dense schedule or not) and weekly paclitaxel or carboplatin-paclitaxel, with or without trastuzumab ± pertuzumab	N (% of C)
E. That underwent breast and axillary surgery with curative intent after the neoadjuvant systemic therapy	N (% of D)
F. For which pathological report is available at diagnosis and at surgery	N (% of E)
G. And for which data is available regarding each treatment cycle (schedule and administration)	N (% of F)
H. Total meeting inclusion criteria	N (% of A)
Exclusion Criteria	
I. Patients with a history of another primary malignancy, except for non-melanoma skin cancer or in situ carcinoma	N (% of H)
J. Patients with bilateral invasive breast cancer (contralateral in situ carcinoma is allowed)	N (% of I)
K. Administration of another anti-cancer systemic treatment (including investigational product and endocrine therapy) in the neoadjuvant setting. Use of LHRH agonists is	N (% of J)

allowed	
L. Pregnancy at the time of breast cancer diagnosis or during the neoadjuvant treatment	N (% of K)
M. Total meeting exclusion criteria	N (% of H)
N. Final patient cohort	N (% of A)

Primary Objective 1, Secondary Objective 2 & Exploratory objective 4

pCR at post-neoadjuvant surgery, overall survival and invasive breast cancer-free survival will be presented descriptively overall and by early cessation of paclitaxel / paclitaxel-dose-reduction and by the presence of neuropathy.

The impact of reduced paclitaxel dose-intensity on pCR at post-neoadjuvant surgery will be estimated by odds ratio (OR) derived from a logistic regression model. This model will be adjusted on potential confounding factors at baseline (such as age, breast cancer subtype, tumor size, nodal status, tumor grade, anthracycline regimen (dose dense vs standard), therapy with other agents: carboplatin, pertuzumab, presence of a chemotherapy-induced side effect (neuropathy, grade 3-4 neutropenia, allergic symptoms, liver toxicity), and main comorbidities (diabetes, heart failure, vascular or liver disease).

The impact of reduced paclitaxel dose-intensity on time to event outcome (overall survival and invasive breast cancer-free survival) will be estimated by hazard ratios (HR) derived from a Cox proportional hazard model, adjusted by potential confounding factors at baseline (as indicated in pCR analysis). HR with 95% confidence intervals and corresponding p-values (using Wald test) will be presented. The proportional hazard assumption will be checked graphically by Schoenfeld residuals plots. A further graphical method which may be used for categorical variables is the log-log plot in which the log-log KM survival estimates are plotted against time.

Secondary Objective 3

To better understand our population, descriptive statistics as described for the general approach will be presented for all demographic characteristics (age, gender, year of diagnosis, ECOG/Performance Status, surface area, main comorbidities) and clinical characteristics (mode of presentation, menopausal status, laterality, histological type, AJCC staging, tumor size and nodal status, tumor grade, Ki-67 marker of proliferation, *BRCA1/2* mutation) at diagnosis.

Exploratory Objectives 5 and 6

In order to characterize the factors associated with reduced paclitaxel dose-intensity and its association with treatment-induced neuropathy, the frequency of patients experiencing neuropathy and the grade assessed using CTCAE version 5 (where available) will be summarized using descriptive statistics as outlined above, stratified by whether patients had early cessation/dose reduction of taxanes or completed the therapy. A logistic regression model will be constructed to estimate the associations between paclitaxel-induced neuropathy, grade of paclitaxel neuropathy (where available) and early cessation of paclitaxel or reduced dose-intensity, adjusted for confounding factors at baseline. The frequency of patients experiencing the other main paclitaxel-related toxicities (grade 3-4 neutropenia, allergic reactions and liver toxicity) will also be described.

6.7 Quality Control

Within-site quality will be checked by performing manual data curation: each site will verify data from at least 20 randomly chosen patients. For sites with a more automated approach to data extraction, script files will also undergo checking and QC.

Between-site consistency of data will be checked by descriptive statistical analyses, principal component analyses and graphical approaches to highlight outliers.

A CDM is going to be in place and during data extraction teleconferences will be scheduled regularly to answer any questions related with the protocol to ensure alignment between sites.

The data will be checked in a series of completeness, logic and sense checks, before and after the transference of the dataset.

6.8 Limitations of Research Methods

This is a retrospective, non-interventional database study, which reflects clinical practice and provides a good overview of paclitaxel treatment in the neoadjuvant breast cancer setting. The limitations of the study include its retrospective observational nature. Data entry errors may be present in the patient records, however a chart review and/or curation of the patients will be conducted, which may allow identification and correction of these errors. Some measures may not be consistently captured and thus may be over/under represented in analyses. For variables such as comorbidities and adverse events, only documented instances of these variables will be included within the variable assessment and an absence of a record will be interpreted as absence of disease or event. Care must be taken in the interpretation as they are likely to be under-reported in the data.

The smaller cohort size at some of the centres may result in the data for that centre not being representative of treatment at the country level. For this reason, data will be pooled, after checks are carried out to understand differences between the cohorts at each centre.

The duration of follow-up may not be sufficient to capture differences in overall survival. Also patients towards the later part of the cohort will be followed up for less time than those near the beginning of the cohort.

There may be a high degree of heterogeneity in patient demographics and characteristics between different centres and also differences in the completeness of data, which will need to be understood prior to the pooled data analysis. Exploration of this heterogeneity will be conducted using principal components analysis alongside graphical approaches where applicable, so that these confounding factors can be taken into account during interpretation of the data.

As this study will use real-world data, where patients are seeing a physician as part of their Standard of Care, outcome definitions and censoring for time-to-event outcomes may differ from those used in interventional studies.

7 Research Outputs

An abstract of this study could be presented at ESMO 2023, San Antonio Breast Cancer Symposium 2023 (abstract submission between May and July) or later at ASCO 2024 (abstract submission between November and February). Submission to a peer-reviewed journal will be discussed based on the obtained data.

Confirmation of any finding could first be sought by asking for re-exploration of available randomized data (e.g. CTNeoBC that corresponds to pooled data of 12 large randomized trials; BrighTNess trial; CALGB 40603; Keynote-522).

Non-significantly different pCR outcomes according to full or reduced paclitaxel dose-intensity in any of the different BC subtypes could allow to question shorter treatment duration as de-escalation method in a future trial. By taking into account several clinical and pathological co-factors, we could help define the patient population that could safely take part in this kind of study.