NAoMe
National Audit of Metastatic Breast Cancer

NATCAN
National Cancer Audit Collaborating Centre

National Audit of Metastatic Breast Cancer State of the Nation Report:

Methodology Supplement

An audit of care received by people diagnosed with metastatic breast cancer between 1 January 2020 and 31 December 2022 in England and Wales.

Published September 2025

Citation for this document:
National Audit of Metastatic Breast Cancer (NAoMe) State of the Nation Report 2025. London:
National Cancer Audit Collaborating Centre, Royal College of Surgeons of England, 2025.

Commented [GH1]: Is this intended to be part of a full address?

This document was prepared by members of the NAoMe project team:
Kieran Horgan, Surgical Lead
David Dodwell, Clinical Oncology Lead
Mark Verrill, Medical Oncology Lead
Sarah Blacker, Clinical Research Fellow
Jemma Boyle, Senior Clinical Research Fellow
Liyang Wang, Clinical Research Fellow
Christine Delon, Data Scientist
Diana Withrow, Methodologist
David Cromwell, Senior Methodologist

With review and input from:
NAoMe Audit Advisory Committee (Clinical Reference Group)
NATCAN Executive Team

Acknowledgements: Special thanks to Jibby Medina for her many contributions as Programme Manager for NAoPri, NAoMe and NABCOP from April 2016 until May 2025

The Royal College of Surgeons of England is an independent professional body committed to enabling surgeons to achieve and maintain the highest standards of surgical practice and patient care. As part of this it supports audit and the evaluation of clinical effectiveness for surgery. Registered Charity no: 212808.

The National Cancer Audit Collaborating Centre (NATCAN) is commissioned by the Healthcare Quality Improvement Partnership (HQIP) and funded by NHS England and Welsh Government as part of the National Clinical Audit and Patient Outcomes Programme (NCAPOP). NATCAN delivers national audits in bowel, breast (primary and metastatic), kidney, lung, non-Hodgkin lymphoma, oesophago-gastric, ovarian, pancreatic and prostate cancers.

The Association of Breast Surgery is a registered charity dedicated to advancing the practice of breast surgery and the management of breast conditions for the benefit of the public. It is a multi-professional membership association, which promotes training, education, clinical trials and guideline composition and adoption. For further information, please refer to the website www.associationofbreastsurgery.org.uk. Registered charity no: 1135699

The UK Breast Cancer Group (UKBCG) is a forum for Clinical and Medical Oncologists. The UKBCG acts as a stakeholder to NICE, NHS England and other organisations; and undertakes key pieces of work, at times in collaboration with other bodies, with the overriding endpoint of improving patient care.
The Group's objectives include advancing the education of clinical and medical oncologists in the subject of breast cancer, concerning its identification, diagnosis and treatment; promoting research for the public benefit in all aspects of breast cancer and publishing the results; and assisting in the treatment and care of persons suffering from breast cancer, or in need of rehabilitation, by the provision of education for healthcare professionals.
Further information on the work of the UKBCG is communicated via this website on a regular basis https://ukbcg.org/. Registered charity no: 1177296

This work uses data that have been provided by patients and collected by the NHS as part of their care and support. For patients diagnosed in England, the data are collated, maintained and quality assured by the National Disease Registration Service (NDRS), which is part of NHS England. Access to the data was facilitated by the NHS England Data Access Request Service.

NHS Wales is implementing a new cancer informatics system. As a result, the quality and completeness of data from Wales is likely to have been impacted due to implementation of this new system across multiple NHS organisations (Health Boards), which has resulted in data being supplied by both old and new systems. Additionally, and reflecting the uncertainty of data quality, the data submitted to the audit may not have undergone routine clinical validation prior to submission to the Wales Cancer Network (WCN), Public Health Wales.

Contents

1. Introduction

This document provides supporting material to the 2025 State of the Nation (SotN) Report for the National Audit of Metastatic Breast Cancer (NAoMe) and its data tables and data viewer. The document describes the data used in the report with details on sources of data, criteria for inclusion and how data completeness, patient characteristics and performance indicators are derived and reported.

2. Sources of Data

The audit uses information from routine national health care datasets in England and Wales. These datasets capture details on the diagnosis, management, treatment and outcome of every patient newly diagnosed with cancer in the NHS in England and Wales.

For England, the audit received information from the National Disease Registration Service (NDRS) at a tumour level for this State of the Nation report. The information held in the National Cancer Registration Data (NCRD) is compiled from a variety of sources including the Cancer Outcomes and Services Dataset (COSD), Hospital Episode Statistics admitted patient care (HES APC) records, the Systemic Anti-Cancer Therapy dataset (SACT), Radiotherapy Data Set (RTDS) and data submitted by pathology laboratories. The audit also received linked information from COSD (linked at tumour level), HES APC, HES Outpatients data (HES OP), SACT, RTDS and the Primary Care Prescription Database (PCPD) (all linked at patient level). **Appendix 1** provides more detail on the data sources listed below and the information they contain.

The English data received by the National Cancer Audit Collaborating Centre (NATCAN) included data on patients registered with cancer from 01 January 2015 up to 31 December 2022.

As with cancer registries in other countries, cancer registrations in England can take up to 5 years after the end of a given calendar year to approach 100% completeness and stability. NDRS uses an active system of gathering information on cancer diagnoses from multiple sources across the patient pathway. Completeness varies by tumour type because different patient pathways provide different opportunities for data flows into NDRS. The 'Gold standard' cancer registration dataset that is used in cancer statistics bulletins and available for analysis outside of NDRS contains over 98% of all the people that will eventually be found by the registration process, and the completeness for a calendar year of data increases over time. More information about the cancer registration process can be found here.

For Wales, the audit was provided with a registration dataset at patient level for patients diagnosed with cancer from 01 January 2015 up to 31 December 2022. Welsh cancer registration data is captured through a national system, Cancer Information System for Wales (CaNISC) and the new Welsh Clinical Portal (WCP). The audit also received linked datasets of records from the Patient Episode Database for Wales (PEDW) containing information on inpatient and day case activity, and mortality data from the Office for National Statistics (ONS).

3. Inclusion and Exclusion Criteria

The data submitted by NDRS and Wales Cancer Network (WCN) is checked and filtered for eligible participants, tables 3.1 and 3.2 explain the process in defining the final cohort to be used in the audit. The NAoMe aspires to include all people diagnosed with metastatic breast cancer defined as spread beyond the breast and regional lymph nodes. There are two distinct cohorts of people with metastatic breast cancer identified within the NAoMe (Table 3.1):

1) *De novo* cohort – metastatic disease identified at diagnosis.

2) Recurrent cohort – metastatic disease identified following initial treatments for primary breast cancer.

Currently, information regarding the date and type of recurrent disease is largely missing in both English and Welsh cancer datasets. The methodology outlined in Table 3.1 is being used as an interim solution. This approach does not identify all patients with recurrent metastatic breast cancer. In particular, those who did not have a record in the HES-Admitted Patient Care (APC) dataset will be excluded. HES APC includes "any secondary care based activity that requires a hospital bed, thus including both emergency and planned admission, day cases, births and associated deliveries. HES APC does not cover accident and emergency (A&E, emergency department) attendances or outpatient bookings; these data are held in separate HES databases."[1]

[1] https://pmc.ncbi.nlm.nih.gov/articles/PMC5837677/

People were included for analysis within the SotN Report if they met the following inclusion and not the exclusion criteria:

Table 3.1: Audit Inclusion Criteria	
Inclusion Criteria	**Details**
Type of cancer	A primary diagnosis with ICD-10 diagnosis codes: C50 (invasive breast cancer) or D05 (stage 0)
Adults	Age >=18
Valid Diagnosis Date	***De novo* cohort:** Primary breast cancer diagnosed 01 January 2020 to 31 December 2022 **Recurrent cohort:** Primary breast cancer diagnosed 01 January 2015 to 30 June 2022. Recurrence identified between 01 January 2020 to 31 December 2022
Cancer Stage	***De novo* cohort:** Stage 4 at diagnosis according to Cancer Registration/COSD for England or CaNISC for Wales OR Stage 0-3 or unknown staging at initial diagnosis with metastatic disease recorded at a distant site in HES-APC or PEDW (ICD10: C78, 79) within 6 months of date of diagnosis **Recurrent cohort:** Stage 0-3 or unknown staging at initial diagnosis AND secondary malignancy recorded at a distant site in HES-APC or PEDW (ICD10: C78, 79) more than 6 months after date of diagnosis (with admission date between 01 January 2020 and 31 December 2022)

Table 3.2: Audit Exclusion Criteria	
Exclusion Criteria	**Details**
Reported by death certificate only or date of diagnosis corresponds to date of death	**For English data:** Using NCRD: basisofdiagnosis = 0 (Death certificate) and/or dco = Y (tumour registered from a death certificate only) and/or date_diagnosis=ONS_date_death **For Welsh data:** Date_diagnosis = ONS_date_death and/or basisofdiagnosis=0
Diagnosed outside of an eligible NHS organisation in England or Wales	Diagnosed outside of an NHS organisation in England or Wales OR Place of diagnosis not provided OR Diagnosed at an NHS organisation with no active breast unit OR Diagnosed within an NHS organisation with less than 30 allocated registrations of breast cancer per year OR Diagnosed at a tertiary centre
Multiple cancer registrations during the audit period	>1 tumour ID in the cancer registration dataset (e.g. bilateral cancer)

4. Key Data Items

Details of the variables and datasets used to compile the data completeness information are shown below in Table 4.1. Data completeness is reported for only the *de novo* cohort of patients.

Table 4.1: Data Completeness Variables for the *de novo* cohort				
Data Item	**Source**			
	England		**Wales**	
	Data field	**Dataset**	**Data field**	**Dataset**
Performance status	CR0510	COSD	PerformanceStatus	CaNISC
Clinical Nurse Specialist	CR2050	COSD	HasSeenClinicalNurseSpecialist	CaNISC
ER status[A]	ER_STATUS; ER_SCORE / pBR4220	NCRD /COSD	ErStatus	CaNISC
HER2 status	HER2_STATUS; pBR4280	NCRD/COSD	Her2Status; Her2FishStatus	CaNISC
PR status	PR_STATUS; PR_SCORE / pBR4290	NCRD/COSD	PrStatus	CaNISC
Stage (overall)	STAGE_BEST, T-stage, N-stage, M-stage	NCRD	StageGroupIntegrated	CaNISC
T-stage	T_BEST	NCRD	FirstTStagePathological	CaNISC
N-stage	N_BEST	NCRD	FirstNumberOfNodesExamined FirstNumberOfNodesPositive	CaNISC
Grade	GRADE	NCRD	FirstGradeOfDifferentiationPathology	CaNISC

[A] The percentage reported for data completeness reflects the data quality as received by the NAoPri, without augmentation with data for endocrine therapy prescription, to highlight the need for improved data quality. When oestrogen receptor status was used for risk adjustment or for subgroup analyses, this was augmented with data from Primary Care Prescription Database, where persons receiving endocrine therapy were assumed to be ER-positive.

Details of the variables and datasets used to compile the patient and tumour characteristics for each cohort are shown below in Tables 4.2 and 4.3.

Table 4.2: Patient and Tumour Characteristics Variables for the *de novo* cohort				
Data Item	**Source**			
	England		**Wales**	
	Data field	**Dataset**	**Data field**	**Dataset**
Year of Diagnosis	DIAGNOSISDATEBEST / CR2030	NCRD/COSD	DateOfPrimaryDiagnosis	CaNISC
Age at diagnosis	Birthmonth; birthyear / CR0100	NCRD/COSD	AgeAtDiagnosis	CaNISC
Sex	SEX	NCRD	Gender	CaNISC
Screened	SCREEN_DETECTED / CR1600	NCRD/COSD	CancerReferralSource	CaNISC
Ethnic Group	ETHNICITY / CR0150	NCRD/COSD	PatientEthnicGroupDescription	CaNISC
Index of Multiple Deprivation Quintile	QUINTILE_2019 / CR0080	NCRD/COSD	LSOA_ Deprivation Quintile Group	CaNISC
SCARF index	DIAG_nn	HES APC	Diagnosis_n	PEDW
Charlson comorbidity score	DIAG_nn	HES APC	Diagnosis_n	PEDW
Performance status	CR0510	COSD	PerformanceStatus	CR0510
Metastatic disease	DIAG_nn	HES APC	Diagnosis_n	PEDW
Stage (overall)	STAGE_BEST, M_Best	NCRD	StageGroupIntegrated	CaNISC

Date of Death	Vitalstatus, vitalstatusdate, deathdatebest	NCRS/ONS	DeathDate	CaNISC/ONS

Table 4.3: Patient and Tumour Characteristics Variables for the recurrent cohort				
Data Item	**Source**			
	England		**Wales**	
	Data field	**Dataset**	**Data field**	**Dataset**
Date of recurrence	Date of first ICD-10 code for metastatic breast cancer (DIAG_nn)	HES APC	Date of first ICD-10 code for metastatic breast cancer.	PEDW
Age at first metastases recorded	Birthmonth; birthyear / CR0100 & Date of recurrence	NCRD/COSD	AgeAtDiagnosis (first diagnosis), DateofDiagnosis (first diagnosis) & Date of recurrence	CaNISC
Sex	SEX	NCRD	Gender	CaNISC
Ethnic Group	ETHNICITY / CR0150	NCRD/COSD	N/A	CaNISC
Site of metastases	DIAG_nn	HES APC	Diagnosis_n	PEDW

5. Indicator Definitions

The audit uses key indicators to monitor progress against its healthcare improvement goals. These indicators align with national guidelines and standards. Definitions of how the indicators included in the SotN report were derived from data for England and Wales are described below.

Some indicators are focused on subgroups of patients as defined by histopathology or treatment, as these factors are important determinants of whether particular treatments or processes are suitable for patients.

Performance Indicator 1: Discussed in a multidisciplinary team (MDT)

Percentage of patients with newly diagnosed metastatic breast cancer (MBC) (*de novo*) disease who have their care discussed within a multidisciplinary team (MDT) meeting.

Table 5.1: Percentage of patients with newly diagnosed metastatic breast cancer (MBC) discussed in a multidisciplinary team (MDT).		
	England	**Wales**
Dates of diagnosis:	1/1/2020 to 31/12/2022	1/1/2020 to 31/12/2022
Numerator: Number of people who were discussed at an MDT.	COSD_date_firstmdtmeeting1 is completed.	Non missing value for CancerCarePlanMDTDate
Denominator: Number of people with invasive breast cancer and evidence of *de novo* metastatic disease.	Final ***de novo* cohort** as described in patient inclusion / exclusion.	Final ***de novo* cohort** as described in patient inclusion / exclusion.
Construction notes	Recorded MDT date is within 30 days of date of diagnosis.	Recorded MDT date is within 30 days of date of diagnosis.
Country reporting:	England & Wales – combined and separate.	
Organisational Reporting level:	English NHS Trusts	Welsh Health Boards

Subgroup Reporting:	Gender, age-group and diagnosis year.	Gender, age-group and diagnosis year.
Risk adjusted:	No – not required as should be near 100% regardless of case-mix.	
Outlier reporting:	No	

Performance Indicator 2: Biopsy in people with recurrent disease

Percentage of people with recurrent metastatic breast cancer who had a biopsy to inform care.

This performance indicator is reported for England only due to differences in data availability. There is an absence of treatment data available for the recurrent cohort for Wales, precluding any reporting.

Table 5.2: Percentage of people with recurrent metastatic breast cancer who had a lesion biopsied to inform care.		
	England	**Wales**
Dates of diagnosis:	1/1/2020 to 31/12/2022	Not available
Numerator: Number of people who have a lesion biopsied.	Biopsy in COSD no more than 6 months prior to, or any time after recurrence[A]	Not available
Denominator: Number of people with invasive breast cancer and evidence of recurrent metastatic disease.	Final **recurrent cohort** as described in patient inclusion / exclusion.	Not available
Construction notes	Date for biopsy in COSD-pathology file: no more than six months before date of recurrence and at least six months after the date of primary diagnosis	Not available
Country reporting:	England only	
Organisational Reporting level:	English NHS Trusts	Not available
Subgroup Reporting:	Gender, age-group and diagnosis year.	Not available
Risk adjusted:	No	
Outlier reporting:	No	

[A] A wide time interval is used to account for the fact that date of recurrence from HES is not likely to be the true date a recurrence was discovered.

Performance Indicator 3: Receipt of CDK4/6 inhibitors.

Percentage of patients with ER positive MBC who received CDK4/6 inhibitors.

This performance indicator is reported for England only due to differences in data availability. There is an absence of data for CDK4/6 inhibitors for Wales, precluding reporting.

Table 5.3: Percentage of patients with ER positive MBC who received CDK4/6 inhibitors.		
	England	**Wales**
Dates of diagnosis:	1/1/2020 to 31/12/2022	Not available
Numerator: Number of people who have treatment with a CDK4/6 inhibitor initiated.	Record of CDK4/6 inhibitor within SACT dataset.	Not available
Denominator: Number of people with ER positive, HER2 negative or HER2 "unknown", invasive	Final **de novo cohort** as described in patient inclusion / exclusion. Restricted to those with ER positive, HER2 negative or HER2 unknown disease. Excluding	Not available

breast cancer with evidence of *de novo* metastatic disease (excluding those who die within 30 days of diagnosis).	those who die within 30 days of diagnosis.	
Construction notes	If SACT drug name is one of abemaciclib, palbociclib, ribociclib, and using matching date_SACT_administration	Not available
Country reporting:	England only	
Organisational Reporting level:	English NHS Trusts	Not available
Subgroup Reporting:	Gender, age-group and diagnosis year.	Not available
Risk adjusted:	Yes – age, tumour grade, Charlson co-morbidity score, SCARF frailty index, source of metastases and diagnosis year.	
Outlier reporting:	No	

Performance Indicator 4: Receipt of anti-HER2 therapy.

Percentage of people with HER2 positive MBC who received anti-HER2 therapy within six months of diagnosis.

This performance indicator is reported for England only due to differences in data availability. There is an absence of data for dates of administration of anti-HER2 therapy for Wales, precluding reporting.

Table 5.4: Percentage of people with HER2 positive MBC who received anti-HER2 therapy within six months of diagnosis		
	England	**Wales**
Dates of diagnosis:	1/1/2020 to 31/12/2022	Not available
Numerator: Number of people who have treatment with an anti-HER2 therapy initiated.	Record of anti-HER2 therapy within the SACT dataset with an administration date within 6 months of the date of diagnosis.	Not available
Denominator: Number of people with HER2-positive invasive breast cancer and evidence of metastatic disease (excluding those who die within 30 days of diagnosis).	Final **de novo cohort** as described in patient inclusion / exclusion. Restricted to those with HER2 positive disease. Excluding those who die within 30 days of diagnosis.	Not available
Construction notes	If SACT-name is one of the following: Lapatinib (Tykerb, Tyverb), Neratinib (Nerlynx), Pertuzumab (Perjeta), trastuzumab (Herceptin, Herzuma, Ontruzant, Trazimera), Phesgo, Trastuzumab duocarmazine (Syd985), Tucatinib and using matching date_SACT_administration	Not available
Country reporting:	England only	
Organisational Reporting level:	English NHS Trusts	Not available
Subgroup Reporting:	Gender, age-group and diagnosis year.	Not available
Risk adjusted:	Yes - age, tumour grade, Charlson co-morbidity score, SCARF frailty index, source of metastases, ER-status and diagnosis year.	
Outlier reporting:	No	

Performance Indicator 5: Chemotherapy.

Percentage of people with *de novo* disease who received any chemotherapy. This indicator is presented for both the *de novo* and recurrent population, separately. Treatment for the recurrent population is not available for Wales.

Table 5.5: Percentage of people with de novo MBC who received chemotherapy.		
	England	**Wales**
Dates of diagnosis:	1/1/2020 to 31/12/2022	1/1/2020 to 31/12/2022
Numerator: Number of people who have treatment with chemotherapy initiated.	Record of chemotherapy within the SACT dataset.	Record of chemotherapy within the CaNISC dataset.
Denominator: Number of people with invasive breast cancer and evidence of *de novo* metastatic disease.	Final **de novo cohort** as described in patient inclusion / exclusion.	Final **de novo cohort** as described in patient inclusion / exclusion.
Construction notes		
Country reporting:	England & Wales – combined and separate.	
Organisational Reporting level:	English NHS Trusts	Welsh Health Boards
Subgroup Reporting:	Gender, age-group, hormone receptor status and diagnosis year.	Gender, age-group, hormone receptor status and diagnosis year.
Risk adjusted:	Yes - age, tumour grade, Charlson co-morbidity score, SCARF frailty index, source of metastases, ER-status, HER2 status and diagnosis year.	
Outlier reporting:	No	

Table 5.6: Percentage of patients with recurrent MBC who received chemotherapy.		
	England	**Wales**
Dates of diagnosis:	1/1/2020 to 31/12/2022	Not available
Numerator: Number of people who have treatment with chemotherapy initiated.	Record of chemotherapy within the SACT dataset.	Not available
Denominator: Number of people with invasive breast cancer and evidence of recurrent metastatic disease.	Final **recurrent cohort** as described in patient inclusion / exclusion.	Not available
Construction notes	Any record of a date within SACT occurring after the recurrence date or up to 30 days before, but at least six months after the original date of diagnosis with primary breast cancer.	Not available
Country reporting:	England only	
Organisational Reporting level:	English NHS Trusts	Not available
Subgroup Reporting:	Gender, age-group, hormone receptor status and diagnosis year.	Not available
Risk adjusted:	Yes - age, tumour grade, Charlson co-morbidity score, SCARF frailty index, source of metastases, ER-status, HER2 status and diagnosis year.	

Outlier reporting:	No

Performance Indicator 6: Receipt of a bisphosphonate or denosumab.

Percentage of patients with bone metastases who received a bisphosphonate or denosumab.

This performance indicator remains under development and is not yet reported on.

Performance Indicator 7: Radiotherapy.

Percentage of MBC patients who received radiotherapy.

This performance indicator remains under development and is not yet reported on.

Performance Indicator 8: Clinical nurse specialist (CNS) contact.

Percentage of patients with clinical nurse specialist (CNS) contact recorded as "Yes".

Table 5.7: Percentage of patients with clinical nurse specialist (CNS) contact recorded as "Yes".		
	England	**Wales**
Dates of diagnosis:	1/1/2020 to 31/12/2022	1/1/2020 to 31/12/2022
Numerator: Number of people who have contact with a breast CNS.	Number of patients who have COSD data item CR2050 (*Clinical Nurse Specialist Indication Code)* recorded as yes.	Number of patients who have CaNISC data item HasSeenClinicalNurseSpecialist recorded as yes.
Denominator: Number of people with invasive breast cancer and evidence of *de novo* metastatic disease.	Final **de novo cohort** as described in patient inclusion / exclusion.	Final **de novo cohort** as described in patient inclusion / exclusion.
Construction notes		
Country reporting:	England & Wales – combined and separate.	
Organisational Reporting level:	English NHS Trusts	Welsh Health Boards
Subgroup Reporting:	Gender, age-group and diagnosis year.	Gender, age-group and diagnosis year.
Risk adjusted:	No – not required as we expect this to be yes, regardless of case-mix.	
Outlier reporting:	No	

Performance Indicator 9: Death recorded within 30 days of a chemotherapy cycle.

Percentage of patients with death recorded within 30 days of a chemotherapy cycle. Patients in the recurrent cohort who die within 30 days of a chemotherapy cycle. This indicator is presented for both the de novo and recurrent population, separately. This metric is not available for Wales as only the first course of chemotherapy was available.

Table 5.8: Percentage of patients with death recorded within 30 days of a chemotherapy cycle (*de novo* cohort).		
	England	**Wales**

Dates of diagnosis:	1/1/2020 to 31/12/2022	Not available
Numerator: Number of people who die within 30 days of a chemotherapy cycle.	Record of mortality up to 30 days after chemotherapy recorded in SACT.	Not available
Denominator: Number of people with invasive breast cancer and evidence of *de novo* metastatic disease who have received chemotherapy.	Final **de novo cohort** as described in patient inclusion / exclusion, restricted to those who have had any record of chemotherapy as described in indicator 5.	Not available
Construction notes	Record of date of death within 30 days of last date recorded in SACT dataset.	Not available
Country reporting:	England & Wales – combined and separate.	
Organisational Reporting level:	English NHS Trusts	Not available
Subgroup Reporting:	Age-group and diagnosis year	Not available
Risk adjusted:	Yes - age, tumour grade, Charlson co-morbidity score, SCARF frailty index, source of metastases, ER-status, HER2 status and diagnosis year.	Not available
Outlier reporting:	No	

Table 5.9: Percentage of patients with death recorded within 30 days of a chemotherapy cycle (recurrent cohort).		
	England	**Wales**
Dates of diagnosis:	1/1/2020 to 31/12/2022	Not available
Numerator: Number of people who die within 30 days of a chemotherapy cycle.	Record of mortality up to 30 days after chemotherapy recorded in SACT.	Not available
Denominator: Number of people with invasive breast cancer and evidence of *de novo* metastatic disease who have received chemotherapy.	Final **recurrent cohort** as described in patient inclusion / exclusion. Restricted to people who have had any record of chemotherapy as described in indicator 5.	Not available
Construction notes:	Record of date of death within 30 days of last date recorded in SACT dataset, where that date occurs after the date of recurrence as described in indicator 5.	Not available
Country reporting:	England only	
Organisational Reporting level:	English NHS Trusts	Not available
Subgroup Reporting:	Age-group and diagnosis year.	Not available
Risk adjusted:	Yes - age, tumour grade, SCARF frailty index, ER-status, HER2 status and diagnosis year.	Not available
Outlier reporting:	No	

Commented [DW2]: Is there a reason that Charlson is excluded here?

Performance Indicator 10: 1 and 3 year survival.

Percentage of patients who survived at least 1, 3 or 5 years after diagnosis. Mortality data for England and Wales is derived from Office for National Statistics mortality data.

Commented [GH3]: Is 5 year survival included? If so perhaps we add this to the title line and table also?

Table 5.10: Percentage of patients who survived at least 1 or 3 years from their initial breast cancer diagnosis.		
	England	**Wales**
Dates of diagnosis:	1/1/2020 to 31/12/2022	1/1/2020 to 31/12/2022
Numerator:	Number of people who survive for at least 1 or 3 years from the date of diagnosis.	Number of people who survive for at least 1 or 3 years from the date of diagnosis.
Denominator: Number of people with invasive breast cancer and evidence of *de novo* metastatic disease.	Final **de novo cohort** as described in patient inclusion / exclusion.	Final **de novo cohort** as described in patient inclusion / exclusion.
Construction notes:	Used Kaplan-Meier to estimate overall survival. See Statistics section for more information.	
Country reporting:	England & Wales – combined and separate.	
Organisational Reporting level:	National	National
Subgroup Reporting:	No	No
Risk adjusted:	N/A – reported at the national level	
Outlier reporting:	No	

6. NHS organisations

The audit presents organisation-level findings by the NHS organisation of diagnosis. This is because this is the organisation where diagnosis and multidisciplinary team decisions are likely to be made.

There are some tertiary centres that mainly provide oncological treatment for people with breast cancer. This includes the Christie NHS Foundation Trust, Clatterbridge Cancer Centre NHS Foundation Trust, and Velindre NHS Trust. These tertiary centres are not included directly within audit outputs where findings are reported by the diagnosing NHS organisation. This is because patients are not diagnosed at these centres.

7. Statistical Analysis

All statistical analyses were conducted using *Stata version 17.*

Most results in the SotN Report are descriptive. The results of categorical data items are reported as percentages (%). Results are typically provided as an overall figure and broken down by NHS organisation of diagnosis (see NHS organisations section). Note that within tables in the SotN Report, the total percentage may not equal 100%, due to rounding.

1- and 3-year overall survival

Overall survival was calculated within Stata using Kaplan-Meier survival analysis methods. 1- and 3-year overall survival was calculated from the date of breast cancer diagnosis using ONS mortality data. For those patients with no ONS date of death, a “date last known alive” or censoring date was calculated for use in survival analyses. For English patients provided by the NCRAS, this was taken to be the vital status date provided; where this date was missing, the day after the last reported date of death was used. For Welsh patients, the day after the last reported date of death was used.

7.1 Suppression

Data suppression is primarily used to protect patient privacy by preventing the disclosure of sensitive information when dealing with small sample sizes or specific combinations of data points. It involves withholding certain data elements to minimize the risk of identifying individuals. For data quality and completeness, disclosure risk is deemed negligible, and the results have not been suppressed. For performance indicators, organisations with indicator denominator values of less than 10 have been suppressed. Risk-adjusted percentages for indicators with fewer than 10 in the denominator have also been suppressed as small numbers may lead to spurious conclusions if not provided within a wider context and with appropriate uncertainties acknowledged.

To avoid residual disclosure which could occur by summing across subcategories, where only a single trust within a cancer alliance has been supressed, a second trust has had the number and unadjusted percentage supressed. This prevents residual disclosure however the risk-adjusted percentage is provided.

In the patient characteristics table, for categories with small numbers (1-4) the number has been supressed by converting it to 2 and editing the numbers of a larger group to maintain the overall total. Therefore, the percentage provided for <5 is similar to the real value and the other numbers in the indicator group may contain a small amount of 'jitter'. Where the category value is 'Unknown', 'Not in HESAPC', 'Not in PEDW' or 'Not Recorded', i.e. it represents missing data, we have not suppressed small numbers.

7.2 Risk-adjustment of indicators

Multivariable logistic regression was performed to 'risk adjust' relevant indicators. The regression model is used to estimate the probability of a patient having an outcome (e.g., a given treatment) given patient and tumour characteristics. When summed, these individual probabilities produce the expected number of events at an organisation. The adjusted indicator value for an organisation is then calculated as: the observed number of events at an institution divided by the expected number (based on the national model), multiplied by the overall national average. This is called indirect standardisation. The tables of performance indicators state whether risk adjustment has been performed. Table 7.1 provides details on the datasets and variables used for risk adjustment.

Table 7.1: Risk Adjustment Variables

Data Item	**Additional detail**	
	England	**Wales**
Age at diagnosis	Age, included as a spline (birthmonth; birthyear from NCRD, date of diagnosis from NCRD) For the recurrent cohort also Date of recurrence	Age, included as a spline (AgeAtDiagnosis from CaNISC) For the recurrent cohort, AgeAtDiagnosis (first diagnosis), DateofDiagnosis (first diagnosis) & Date of recurrence
Grade	GRADE from NCRD	FirstGradeOfDifferentiationPathology from CaNISC
Charlson co-morbidity score	The Charlson co-morbidity score is a commonly used scoring system for medical comorbidities, consisting of a grouped score calculated based on the absence (0) and presence (≥1) of 14 pre-specified medical conditions (**Appendix 5**). The score was calculated using information on secondary diagnoses (ICD-10 codes) recorded in HES APC (England) / PEDW (Wales) within the 24-month period prior to a person's diagnosis. For analysis, the CCI is grouped into three categories: • **0** none of the 14 pre-specified comorbidities. • **1** only 1 of the 14 pre-specified comorbidities. • **2+** 2 or more of the 14 pre-specified comorbidities	

SCARF Index[B]	Using DIAG_01-DIAG20 from HES APC	Using Diagnosis01-12 from PEDW
Diagnosis year	Year extracted from diagnosisdatebest (NCRD)	Year extracted from DIAGNOSIS_DATE (Cohort data)
ER Status (Invasive only)	Using ER_STATUS; ER_SCORE from NCRD	Using ERStatus from CaNISC
HER2 Status (Invasive only)	Using HER2_STATUS from NCRD	Using Her2Status and Her2FishStatus from CaNISC.
Source of Metastases	Whether metastasis was identified via NCRD or HES (Applicable to *de novo* cohort only)	

7.3 Handling of missing data

For the risk-adjustment, missing values within the risk adjustment model were imputed with multiple imputation using chained equations, creating ten datasets and pooling model estimates using Rubin's Rules. The imputation models included all the variables in the risk adjustment model.

8. Outlier Process

The NAoMe will not be implementing an outlier process for the findings of this report due to two key limitations: a) the *de novo* population is too small when indicators are produced at a trust level and b) the recurrent population is incomplete, and we do not have information to know how representative this sample is.

Appendix 1: Routine data sources

Overview of the data sources used for the SotN Report.

Country	Data source	Content
England	Cancer registry (NCRD and RCRD)	Data on all aspects of the cancer registration including information from hospital pathology systems.
England	COSD	Cancer Outcomes and Services dataset (COSD) items, are submitted routinely by service providers via multidisciplinary team (MDT) electronic data collection systems to the National Cancer Data Repository (NCDR) on a monthly basis.
England	SACT	Systemic Anti-Cancer Therapy (SACT) data contains information on chemotherapy dates, regimen(s) and dose(s).
England	RTDS	Radiotherapy dataset (RTDS) contains information on radiotherapy treatment including dates, prescription region and dose.
England	HES	Hospital Episode Statistics (HES) is the administrative database of all NHS hospital admissions in England; records were supplied by NHS Digital to NCRAS.
England	PCPD	Primary Care Prescription Database (PCPD) contains information on the use of endocrine therapy.
Wales	CaNISC	Cancer Network Information System Cymru (Canisc) contains data on all aspects of the cancer registration including investigations. (OLD SYSTEM)
Wales	CDF	Clinical Dataset Form (CDF) contains data on all aspects of the cancer registration including investigations (NEW SYSTEM)
Wales	PEDW	Patient Episode Database for Wales (PEDW) is the administrative database of all NHS hospital admissions in Wales.
Wales	RTH	Radiotherapy data (RTH) contains information on radiotherapy treatment.
England & Wales	ONS	Office for National Statistics (ONS) death data including date of death and cause of death.

Appendix 2: Charlson co-morbidity score

Reference:

Armitage JN, van der Meulen JH. Identifying co-morbidity in surgical patients using administrative data with the Royal College of Surgeons Charlson Score. Br J Surg 2010;97:772-81. doi https://doi.org/10.1002/bjs.6930

Pre-specified conditions included in the assignment of Charlson co-morbidity score.

Charlson co-morbidity score Conditions
Myocardial infarction
Dementia
Diabetes mellitus
Metastatic solid tumour
Congestive cardiac failure
Chronic pulmonary disease
Hemiplegia or paraplegia
AIDS/HIV infection
Peripheral vascular disease
Rheumatological disease
Renal disease
Cerebrovascular disease
Liver disease
Any malignancy

Note: AIDS/HIV diagnoses cannot be identified in HES APC data because of legal requirements for NHS trusts to remove patient identifiers from legally restricted records, including those containing diagnoses of HIV/AIDS. These diagnoses are also not found in linked PEDW data.

Appendix 3: WHO Performance Status

Reference:

Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. American Journal of Clinical Oncology. 1982;5(6):649-56

WHO Performance Status	**Definition**
0	Fully active, able to carry on all pre-disease performance without restriction.
1	Restricted in physically strenuous activity but ambulatory & able to carry out work of a light or sedentary nature.
2	Ambulatory and capable of all self-care but unable to carry out any work activities. Up & about more than 50% of waking hours.
3	Capable of only limited self-care, confined to bed or chair more than 50% of waking hours.
4	Completely disabled. Cannot carry on any self-care. Totally confined to bed or chair.
5	Dead.

Appendix 4: Secondary Care Administrative Records Frailty (SCARF) Index

Reference:

Jauhari Y, Gannon MR, Dodwell D, et al. Construction of the secondary care administrative records frailty (SCARF) index and validation on older women with operable invasive breast cancer in England and Wales: a cohort study. BMJ Open 2020;10:e035395. doi: 10.1136/bmjopen-2019-035395

Deficit			
Activity limitation	Diabetic complications	Hypotension	Requirement for care
Anaemia	Falls	Ischaemic heart disease	Respiratory disease
Arthritis	Foot problems	Incontinence	Skin ulcer
Cardiac arrhythmias	Fragility fracture	Neurodegenerative disorders	Sleep disturbance
Cerebrovascular disease	Hearing impairment	Nutritional Problems	Social vulnerability
Chronic kidney disease	Heart failure	Osteoporosis	Thyroid disease
Cognitive and mental health problems	Heart valve disease	Peptic ulcer	Urinary system disease
Diabetes	Hypertension	Peripheral vascular disease	Visual impairment