Australian Government
Australian Institute of
Health and Welfare

Metastatic breast cancer - first national estimates

Methodology paper

AIHW

Contents

Aim of this project

The aim of this project is to estimate the prevalence of metastatic breast cancer (MBC) in Australia and in each state and territory at the end of 2024.

The purpose of this paper is to describe the data and methods used, along with assumptions and data limitations. This provides transparency and a basis for improving future estimates.

Data available to estimate metastatic breast cancer prevalence

National (Australian Cancer Database (ACD)) cancer data only records primary diagnosis of cancer and has details of primary stage of breast cancer for two out of the eight Australian jurisdictions (Victoria and Tasmania); i.e. stage data for other jurisdictions has not been provided to the ACD, and no information on progression of disease from a less advanced stage at diagnosis to metastatic disease is available in the ACD.

The Cancer and Treatment Linked analysis asset (CaT-Link) is a recently created national enduring linkage of selected variables from the following databases: ACD, national death index (NDI), Medicare Benefits Scheme (MBS) and Pharmaceutical Benefits Scheme (PBS).

CaT-Link includes many of the same variables as the ACD (excluding stage data) and can potentially be used to identify cases that have received MBS or PBS services that are specifically for metastatic breast cancer (MBC). Such cases will be a subset of all cases of MBC.

Recently, Cancer Institute New South Wales (CINSW) developed a protocol for identifying MBC cases using linked cancer registry and numerous different sources of treatment data, which was based on earlier work by Lord et al 2019, Smith et al 2025, and Lord et al 2022. This was used to count the number of prevalent cases of MBC in NSW. Cancer Council Victoria (CCV) and Cancer Alliance Queensland (CAQ) then adapted this protocol as closely as possible to available linked data so as to count cases in their states, with all three then comparing methodologies and results.

The methods outlined in this paper, describe how prevalence of MBC cases that have received MBC-specific MBS or PBS services (i.e. a subset of all Australian MBC cases) have been counted, and then how these counts were ’scaled-up’, so as to estimate total MBC prevalence, using findings from the CINSW, CCV and CAQ studies.

CaT-Link is anticipated to be superseded when data from the ACD for all states and territories are included as part of the cancer module of the National Health Data Hub (https://www.aihw.gov.au/reports-data/nhdh) which is anticipated to occur within the next year, at which time additional treatment data sources (e.g. hospital admissions data) would also become available to help improve the completeness and certainty around national estimates of MBC.

This improved linked data capacity would also greatly enhance the national ability to better understand cancer diagnosis, treatment, costs and outcomes.

Until that time, the results presented here should be considered preliminary experimental estimates and interpreted cautiously.

The method – overview

The aim of this work is to estimate the prevalence of metastatic breast cancer (MBC) in Australia and in each state and territory at the end of 2024.

Although breast cancer can be diagnosed in males, over 99% of cases are diagnosed in females. This work estimates MBC in Australian women, while a national estimate is also provided for men.

The available data from CaT-Link is capable of counting MBC-specific MBS or PBS treatments to 2023 (for cases diagnosed with cancer between 1982 and 2019) but does not contain information about any other form of treatment (e.g. hospitalisation)) and is capable of counting people diagnosed with breast cancer to the end of 2019. Information on stage at which cancer is originally diagnosed is not available, and the fact of death is available to the end of 2022 (although incomplete counts are also available for 2023). In addition, changes in the PBS, especially around 2015, create a substantial break in series which restricts the ability to project into the future, to using only the data available for 2016 to 2019. However, the data that is available is considered virtually complete and covers the whole of Australia.

As outlined in the preceding paragraph, the challenges are considerable. However, recent work in this space by CINSW, CCV and CAQ provides a means of overcoming these challenges.

CINSW, CCV and CAQ have used a methodology developed by CINSW, based on earlier work by Lord et al 2022, as the basis to count prevalent MBC cases in their states for two periods, for breast cancer cases diagnosed:

1. since 2008 and who were still alive at the end of 2020 (a period for which they all had data and could therefore check the comparability of the estimates produced for all three states) and
2. from as early as possible (in some cases as early as 1972) and who were still alive at the end of 2020, 2021 or 2022, respectively, projected to the end of 2024.

In addition, CCV was able to access MBS and PBS data for some of 2008-2020 and share information on the prevalent cases that could be identified through MBS or PBS data, the number who had only been identified through MBS or PBS data, and the number who had received other (non-MBS/PBS) treatment. This, and the ability to compare prevalence of MBS/PBS-treated MBC cases in these states with prevalent MBC cases counted by CINSW, CCV and CAQ, provided a mechanism to use the available national MBS and PBS data to estimate the number of prevalent MBC cases in Australia at the end of 2024.

This study:

1. makes use of the work of the CINSW, CCV and CAQ in calibrating a common approach to estimating MBC prevalence at the end of 2020 (with recruitment of cases only from 2008), to develop sufficient understanding to develop a method to estimate national MBC prevalence at the end of 2024.
2. Uses the Longitudinal Cohort Method (developed during this study) to estimate MBS/PBS-treated MBC prevalence at the end of successive years to 2024, for people identified as having MBS/PBS-treated MBC in each year from 2002 to 2024.
3. Uses these (MBS/PBS-treated MBC prevalence counts for VIC and QLD from 2 above) to adjust (CCV+CAQ MBC prevalence at end of 2024) for potentially missing MBS/PBS-treated cases.

4. Uses the ratio of (CINSW+CCV+CAQ (MBS/PBS adjusted)) prevalent MBC cases at the end of 2024 from 3 above)/(CINSW+VIC+QLD prevalent MBS/PBS-treated MBC cases at end of 2022), applied to the number of prevalent MBS/PBS-treated MBC cases at the end of 2022 in Australia and the other states and territories, to estimate MBC prevalence for these populations at the end of 2024.

In addition, MIAMOD methodology, relying on breast cancer incidence and mortality data alone, was used in parallel with the method described above, to estimate MBC prevalence using a completely independent method. These estimates can be compared to the LCM estimates to assess how similar the two sets of estimates are and to investigate any differences.

The work falls into the following three segments:

1. Development of an understanding of the issue and its complexities, and then development of successive methods that are considered to be reliable enough to calculate an estimate of MBC prevalence at the end of 2020, for cases recruited from 2008. This is the period for which CINSW, CCV, CAQ and AIHW had the least differences between the data available to them, when the state-specific data available to AIHW was at its richest, and where there was maximum opportunity to understand enough to develop a working model.
2. Development and application of the best available method to estimate MBC prevalence at the end of 2024.
3. Application of MIAMOD methodology to independently calculate MBC prevalence at the end of 2024.

Background, and factors affecting development of the methodology

CINSW, CCV and CAQ calibrated a common approach to estimating MBC prevalence at the end of 2020 (with recruitment of cases only from 2008).

Sharing of methods by CINSW and some summary Victoria and Queensland data tables by CCV and CAQ, allowed AIHW to develop insights into how the available data could be used to estimate national prevalence.

The logical available approach was to scale up AIHW's complete MBS/PBS-treated prevalence count based on the ratio of CINSW, CCV and CAQ state estimates of MBC to AIHW's equivalent MBS/PBS-treated MBC prevalence counts for these states.

However, several challenges had to be overcome before this was possible.

1. AIHW was only able to count cases diagnosed with MBC up to the end of 2019, who were still alive at the end of 2022 (because these were the periods for which the relevant data were available). A method therefore needed to be developed to reliably estimate the number of cases diagnosed in 2020 and still alive at the end of 2020 (for comparison with CINSW, CCV and CAQ counts for this period), as well as the number of cases diagnosed in 2020 to 2024, and the number of cases diagnosed in previous years who were still alive at the end of 2024.
2. CCV and CAQ had, respectively, restricted and no access to MBS and PBS data, and so their MBC prevalence counts are assumed to be under-estimated to some extent. Based on data provided by CCV, a method was developed to estimate the approximate number of missing cases for CCV and CAQ.

3. Treatments available for people with MBC changed abruptly in July 2015, with the effective transfer of the Herceptin program to the PBS. This resulted in a very large spike in the number of prevalent MBS/PBS-treated MBC cases that year. The upshot is that there is no clear trend in the number of prevalent cases of MBS/PBS-treated MBC who were originally diagnosed with MBC from 2015. This required the development of a method to estimate prevalent MBS/PBS-treated MBC cases diagnosed each year from 2020 to 2024, as well as the prevalence in 2023 and 2024 of the MBS/PBS-treated cases that had been diagnosed each year from 2002. In solving this problem, the Longitudinal Cohort Method (LCM) was developed. The LCM also addresses the problems described in 1 and 2 above.

These issues and underlying concepts are discussed from page 9.

Estimation of MBC prevalence at the end of 2024

Having standardised their approaches as much as possible, CINSW, CCV and CAQ estimated cases prevalent at the end of 2024. However, each state was able to 'recruit' cases initially diagnosed with breast cancer starting from different years (NSW starting from 1972, Qld from 1982 and Victoria from 2008 (but then adjusted to account for missing cases diagnosed with breast cancer between 1982 and 2007). This likely results in the NSW estimate being a little more complete, but it also increases the possibility of more linkage errors occurring resulting in a greater number of deceased cases incorrectly being counted as alive and cases that have moved overseas as still prevalent in Australia. In addition, the treatment data used by each state to identify cases treated for metastatic disease, while broadly similar, are state-specific, and span different periods, which could result in relatively greater, or fewer, counts of MBC. However, in each case these are the best and most complete counts that are currently available.

So as to overcome the data issues mentioned previously, particularly around projecting estimates 5 years into the future (for 2020 to 2024) where trends are not clearly visible, a number of successive methods were developed, assessed and then discarded. Ultimately, lessons learned during this process resulted in the development of the Longitudinal Cohort Method (LCM) which appears to yield plausible projected estimates of MBS/PBS-treated MBC prevalence to 2024.

For diagnosis years and censor years in which counts were not possible because of insufficient data, the LCM was used to estimate the number of cases of MBS/PBS-treated MBC diagnosed in each year from 2003 to 2024 prevalent at the end of each year 2016 to 2024.

The MBS/PBS-treated MBC cases in VIC and QLD that are prevalent at the end of 2024 (from the previous paragraph), were then used to estimate the number that were potentially uncounted by CCV and CAQ because of, respectively, limited or no access to MBS or PBS data, which were then added to the CCV and CAQ estimate of MBC prevalence at the end of 2024.

Australian and jurisdictional MBC prevalence at end of 2024 was then calculated as

$A \times B/C$

Where:

- A is the MBS/PBS-treated MBC prevalence at the end of 2022 for Australia or any state or territory,
- B is the MBS/PBS adjusted MBC prevalence at the end of 2024 for NSW, Victoria and Queensland combined, and

- C is the MBS/PBS-treated MBC prevalence at the end of 2022 for NSW, Victoria and Queensland combined.

In addition, adjustment has been made for perceived linkage errors/migration overseas that would inflate counts and cross-border issues that would diminish counts. However, in practice, this was found to make little difference to the final estimate of MBC prevalence.

Assumptions around the incidence and survival trends for MBS/PBS-treated cases from 2016-19 to 2024 make little noticeable difference to the estimates of MBC prevalence at the end of 2024.

Cases of metastasised breast cancer that have not yet been diagnosed (and treated) as MBC are not included in this prevalence estimate.

Estimation of national MBC prevalence is described from page 20.

MIAMOD methodology to estimate MBC prevalence at the end of 2024

The MIAMOD method, and the value of the parameters (e.g. estimates of survival) used in its application, are described from page 37 of the detailed methods section.

Background and factors affecting development of the methodology

This task is challenging for the following reasons:

- no cancer stage data is available in CaT-Link;
- the only treatment data available in CaT-Link are based on MBS and PBS data. No hospital data is yet available, nor are the other various jurisdiction-specific databases that have been used by CINSW, CCV and CAQ to compile what are considered to be relatively complete counts.
- A break in series for MBS/PBS-treated MBC prevalence between 2014 and 2015 complicates trend identification.
- The latest censor year for which complete counts (as opposed to estimates) of MBS/PBS-treated MBC prevalence are possible, is 2019.

For the purposes of this project, the following is assumed or known:

- All cases of diagnosed breast cancer are included on state and territory cancer registries, and in the ACD (noting that cancer registry data are known to be 'virtually complete').
- CINSW, CCV and CAQ counts are complete, noting that all jurisdictions are aware of the potential for under- or over-counting and have been transparent with their methods. With them, we are aware of some areas where adjustment can be made to account for statistical uncertainty.
- The probability that MBC is treated with MBS and PBS co-funded treatments will be the same in all states and territories.
- Incidence and prevalence of MBS/PBS-treated MBC do not change smoothly over time between 2002 and 2019 but are influenced by a substantial break in series between 2014 and 2015 associated with the transition of function from the Herceptin program. Trends from 2015 to 2019 are not clear. However, it is assumed that from 2016-2019 to 2024, MBC incidence increases at 2.4% each year, a similar rate to that for breast cancer in Australia, and stage 4 de novo breast cancer in Victoria.
- In the ACD (but not in CaT-Link) cancer stage at primary diagnosis is known for Victorian breast cancer cases from 2006 to 2021. This has been useful in allowing a count of de novo MBC incidence and prevalence, and has been useful, along with breast cancer incidence data from Cancer data in Australia (AIHW 2025) in estimating the likely direction and rate of MBC incidence trends (as above).
- The MBC prevalence estimates produced by CINSW, CCV, CAQ and AIHW all rely on a different mix of datasets that cover different periods of time. While there is a great deal of commonality between the types of datasets used (e.g. state-specific variants of similar datasets), and the periods covered, completeness of prevalence counts for some states may be relatively higher or lower than others. During the current phase of national data development, this is considered to be unavoidable.
- The earliest count of MBS/PBS-treated MBC that it is possible to produce using CaT-Link data will be for cases diagnosed in 2002, because CaT-Link does not include stage at diagnosis data, or other treatment data (e.g. hospitals data), and 2002 is the earliest year for which MBS and PBS are available in CaT-Link. CINSW, CCV and

CAQ each have at least some stage at diagnosis data and have some treatment data (with which to identify MBC cases) preceding 2002. As a consequence, CINSW, CCV and CAQ may be able to count prevalent cases starting from an earlier period (compared with the AIHW using the methods described in this paper). However, based on the methods used here, this should not affect AIHW estimates of national prevalence.

- Counting of MBC prevalent cases was undertaken without access to MBS and PBS data:
    - by CCV for all MBC cases diagnosed before 2018, and
    - by CAQ for all MBC cases diagnosed in all years.
- In the three years for which CCV was able to identify MBC cases using MBS or PBS data (2018 to 2020), 47% of all prevalent cases diagnosed since 2008 were identifiable using MBS or PBS data, and 16% of all cases were *only* identifiable through the use of MBS or PBS data (with one third (33%) of MBS/PBS-treated prevalent cases being identifiable only by virtue of the fact that they had received these services). Note that, in the years 2018, 2019 and 2020, respectively, 16%, 17% and 15% of prevalent cases were identified only through MBS/PBS data (demonstrating internal consistency across time).
- CINSW indicated that for the same period, 14% of all cases were *only* identifiable through the use of MBS or PBS data (with one third (34%) of MBS/PBS-treated prevalent cases being identifiable only by virtue of the fact that they had received these services). However, in the years 2018, 2019 and 2020, respectively, 11%, 14% and 17% of prevalent cases were identified only through MBS/PBS data (which does not demonstrate internal consistency across time and possibly suggests a trend).
- It is assumed that 33% of prevalent MBS/PBS-treated MBC cases in Victoria and Queensland will only be identifiable by virtue of the fact that they have accessed those services. So as to estimate potentially missed counts, one third of prevalent MBS/PBS-treated MBC cases diagnosed in years for which CCV and CAQ did not have MBS or PBS data in their linkages were add to their 2024 prevalence counts. This may slightly over-count cases if this tendency diminishes over time (as suggested by the CINSW data in the previous point).
- CINSW had access to MBS data from March 2003 and PBS data from January 2002.
- It is assumed that MBC cases accessing the Herceptin program prior to 2015 (when the program was rolled into the PBS) will be fully counted amongst the prevalent cases newly ‘diagnosed’ in 2015. It is possible that lack of access to Herceptin data could result in some undercounting of prevalence, although the impact on prevalence would likely be relatively small by 2024 due to attrition over the years.
- Linkage will not be perfect in all cases because it is ‘probabilistic’ (based on the probability of the available personal identifiers belonging to the same person). Review of MBS/PBS-treated MBC cases (from CaT-Link) that are ‘prevalent’ (apparently alive) at the end of 2019 shows that around 4.9% of cases are likely to in fact be deceased or overseas (having no MBS or PBS activity at all in 2019), and that the number of cases flagged as deceased, but still accessing MBS or PBS from 2020 (and therefore assumed to be alive) is equivalent to 1.7% of prevalence at the end of 2019. In probabilistic data linkage, there are known to be both ‘false positives’ (e.g., records that link to a death record but are actually alive) and ‘false negatives’ (e.g., records that don’t link to a death record but are actually dead). In theory, the proportions of these should be approximately even (roughly cancel each other out).

Therefore, at least some of the difference between these two percentages (4.9% - 1.7% = 3.2%) could be people who, since being identified as having MBC have moved (permanently) or travelled (temporarily during the relevant periods) overseas. At present, the impact of linkage errors on CINSW, CCV and CAQ estimates is unknown to this study. For the purposes of this work, linkage errors are assumed to be the same for CINSW, CCV and CAQ as they are for AIHW.

Identifying and counting cases of MBS/PBS-treated MBC

In CaT-Link, the only relevant variables available to identify a diagnosis of metastatic breast cancer (MBC) are:

- date of breast cancer diagnosis,
- The date of service and fact of service for all MBS item numbers,
- The date of service and fact of service for all PBS item numbers,
- state of residence at the time of breast cancer diagnosis, at the censor date (e.g. end of 2020 or end of 2022 for 2024 projections), and state in which the MBC was first identified using all the available data.

Identification of an MBC case

A case of MBC is identified where a person has been diagnosed with breast cancer, and then accesses, for the first time, one of the MBS or PBS services specified in Appendix A. The access could be within 3 months of the breast cancer diagnosis date (in which case it is considered a de novo or primary case of MBC), or it could be months or years later than this (a recurrent, or progressed, case of MBC).

Because, in CaT-Link, the date of service can precede the date of diagnosis, there is some uncertainty around the reliability of actual date that an event is reported. This issue needs to be thoroughly investigated. In the meantime, the date of diagnosis is considered to be the calendar month of diagnosis, and the date of MBS or PBS service provision is considered to be the calendar month in which this service was provided.

An MBC case is considered to be de novo if the MBS or PBS service is provided in the same month as the breast cancer diagnosis, or in any of the three months following the month of breast cancer diagnosis. This means that, for de novo cases, the average period between diagnosis and service is 3 months, with the minimum being 2 months and the maximum being 4 months. Where the MBS or PBS service is received at any time after the third month following the diagnosis month, the case is considered to be recurrent. Where a recurrent case has been identified, the date of MBC diagnosis is considered to be the date of the first MBC-specific MBS or PBS service provision.

It is considered to be prudent to continue this treatment of time until confidence in the quality of date variables improves.

The programming definition of de novo and recurrent may differ slightly between CINSW, CCV, CAQ and AIHW.

Jurisdiction to which a case is allocated

The jurisdiction allocated to each MBC case can be:

1. the person's state of residence at the time of MBC diagnosis,
2. the person's usual residence at each censor date.
3. the state registry in which the primary case of breast cancer was registered.

It would be possible to report prevalence for all three of these definitions, but it would be possible to report incidence for only 1 and 3 above, because usual residence at censor date will not be available for cases who have died.

Other considerations are that:

- MBC prevalence estimates calculated by CINSW, CCV and CAQ are assumed to reflect the registry that registered the primary breast cancer diagnosis. As these state counts are the basis by which MBS/PBS-treated MBC prevalence counts were inflated to estimates of total MBC prevalence, alignment as far as possible is important when deriving inflation factors.
- Reporting by state of usual residence at the time of MBC diagnosis is important in that it allows comparison of both incidence (the cases originally diagnosed) and prevalence (the cases still alive at censor date (2020 or 2022)).
- Reporting MBC prevalence by state of usual residence at the censor date would be, assumably, the most useful for the purposes of planning. People can migrate between states, and this definition would provide the most up-to-date estimate of prevalence in each jurisdiction.

State of usual residence at each censor date would be as at the end of 2019 (the latest year for which cancer diagnosis data was available in CaT-Link) for the 2008-2020 prevalence estimates, and 2022 (the latest year for which MBS/PBS data was available in CaT-Link) for 2024 prevalence estimates.

Prevalence counts using each definition have been extracted from CaT-Link and compared. No meaningful difference in count was found between these definitions, consequently state of usual residence at the time of MBC diagnosis is used.

MBS and PBS items used to identify MBC

MBS and PBS item numbers used in this analysis, the conditions developed by CINSW for their use, and the programming logic applied by AIHW are detailed in Appendix A.

Exclusion of some cases accessing all MBS and specific PBS services

The MBS items relate to radiation services that specify that the service is for a secondary site (i.e. for a cancer that has metastasised). However, if the patient has been diagnosed with more than one primary cancer (e.g. breast cancer and colorectal cancer), it can sometimes be unclear for which cancer these services were intended.

Because the stage at diagnosis variable is not included in CaT-Link, it is not possible to identify (and exclude from the count of people with MBC) people diagnosed with breast cancer as well as another type of cancer which has been diagnosed de novo, at metastatic stage.

The available options are to:

- include all such cases,
- exclude all such cases, or

- otherwise attempt to identify cases for which the MBS service is unlikely to be for the other cancer type and include these only.

The same issue (lack of stage data and the options available) applies to inclusion or exclusion of cases treated with PBS items that relate to Gemcitabine, Vinorelbine and Megestrol.

Preliminary analyses show the total prevalence count of MBS/PBS-treated MBC (2008-2019) at the end of 2019 was:

- 5,013 if all prevalent cases that have also been diagnosed with another cancer type, and which had been identified using these MBS and specific PBS items, are included, and
- 4,579 if all prevalent cases that have also been diagnosed with another cancer type, and which had been identified using these MBS and specific PBS items, are excluded.

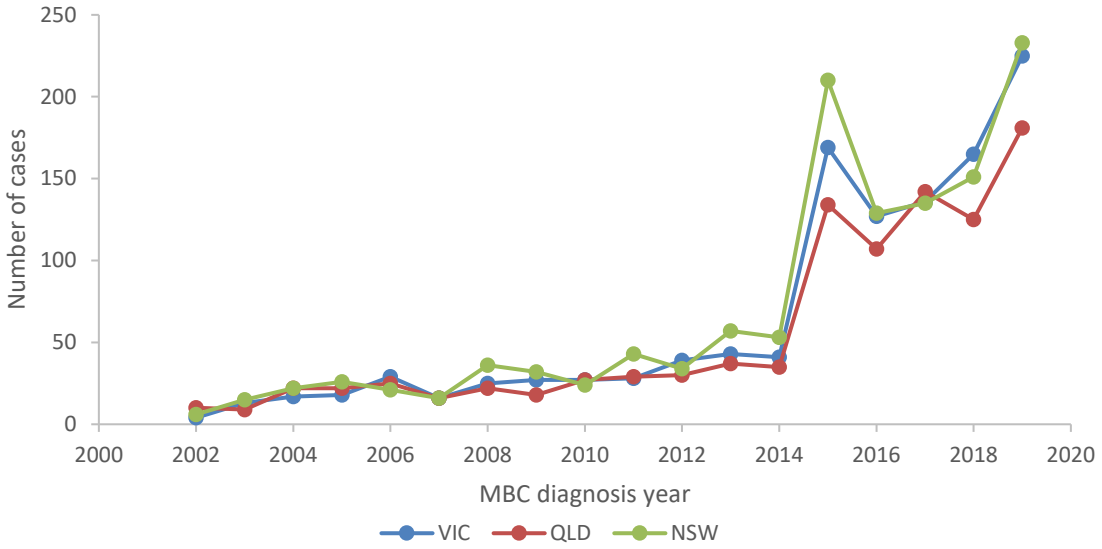
The difference between the two is relatively small (<10% of the smaller count) but relevant.

A partial solution, in the absence of stage data, is to apply a logical condition. If the other cancer type was diagnosed after the specific MBS or PBS items had been accessed, then the service is unlikely to have related to the other cancer, and so such cases could be assumed to be MBC. Conversely, if the other cancer is diagnosed before the MBS or specific PBS service had been accessed, then it would not be possible to know whether the service was to treat metastatic breast cancer or the metastatic form of the other cancer, and the case would be excluded (if a conservative approach was preferred). Such an approach (excluding cases where the other cancer is diagnosed before the first MBS or specific PBS service) yields a national MBS/PBS-treated MBC prevalence of 4,715 compared with 4,579 and 5,013 detailed above. In the absence of access to stage data, this approach has been taken.

The Herceptin program

Figure 1 shows MBS/PBS-treated MBC prevalence at end of 2019, by year of MBC diagnosis.

Figure 1: MBS/PBS-treated MBC prevalent cases at the end of 2019, by year of MBC diagnosis, NSW, Victoria and Queensland



Notes:

Source: AIHW analysis of CaT-Link data

As the cessation of the Herceptin program for treating metastatic breast cancer and the uptake of this function in the PBS occurred around this time, the break in series for the relevant MBS/PBS services between 2014 and 2015 is assumed to be due to this change. A review of prevalence for MBC cases treated by the full range of individual MBC-specific PBS items confirms this assumption.

The Herceptin program

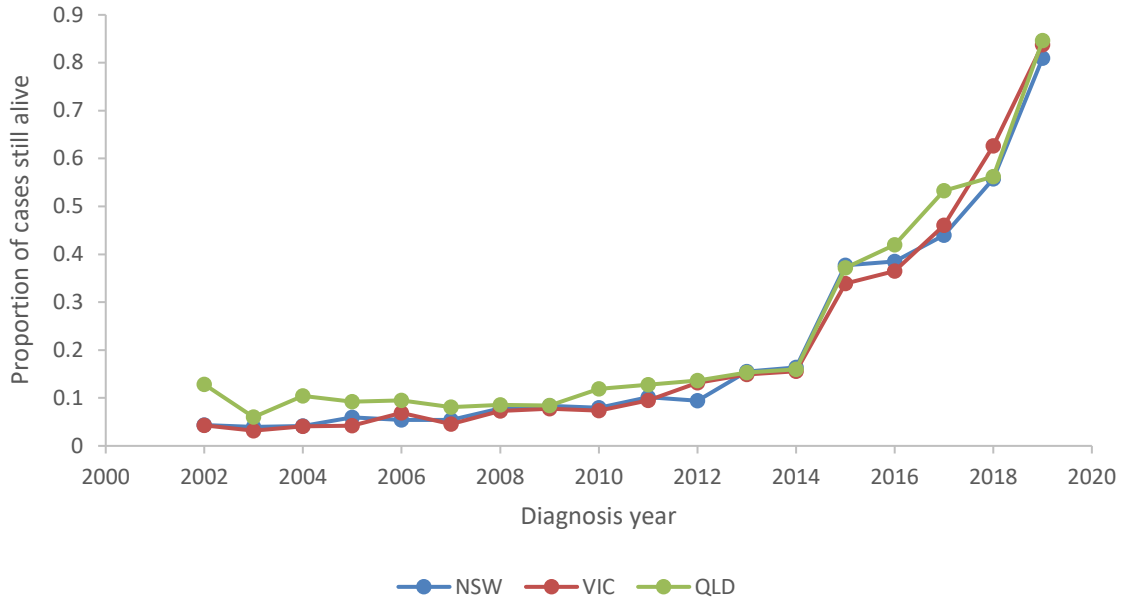
Prior to 1 July 2015, non-PBS-subsidised Herceptin was subsidised by the Australian Government free-of-charge to eligible patients for late-stage metastatic breast cancer through the special Herceptin programme that was administered by the Department of Human Services outside the PBS.

From December 2001, Herceptin was subsidised for early stage and locally advanced HER2 positive breast cancer through the Pharmaceutical Benefits Scheme (PBS).

From 1 July 2015, the Herceptin PBS listing was extended to include treatment for Metastatic (Stage IV – also called secondary or advanced) HER2 positive breast cancer (PBS 2015).

The percentage of MBS/PBS-treated cases diagnosed between 2015 and 2019 and still alive at the end of 2019, follow a recognisable survival curve (Figure 2), suggesting that survival probability expectations and therefore MBS/PBS treatment in this period is reasonably consistent. However, survival probabilities prior to 2015 are substantially lower, and clearly from a different series (i.e. without Herceptin).

Figure 2: Proportion of MBS/PBS-treated MBC cases still alive at end of 2019, by year of MBC diagnosis, for MBC cases diagnosed in NSW, Victoria and Queensland



Source: AIHW analysis of CaT-Link data

If all cases accessing the Herceptin program transitioned to the PBS, then the number of prevalent cases previously treated with Herceptin (i.e. actually diagnosed in previous years) is assumed to be completely counted amongst the additional number of prevalent cases first visible ('apparently' diagnosed with MBC) in 2015.

The prevalence of MBS/PBS-treated MBC cases in 2016 is several times greater than in 2014 (even after considering the effect of MBC-related deaths in this period), which is assumed to reflect the very substantial role played by the Herceptin program in previous years.

CINSW linkage includes MBS and PBS data from at least 2003, CCV linkage includes MBS and PBS data from 2018, and CAQ linkage did not have access to MBS and PBS data.

The proposed method for adjusting CCV and CAQ MBC prevalence estimates is to assume, based on limited data shared by CINSW and CCV, that in all the years from 2002 to 2020, a third of all prevalent MBS/PBS-treated MBC cases diagnosed in each year will be unable to be counted where MBS and PBS data are not included in the data linkage.

Accuracy of probabilistic data linkage

Probabilistic data linkage is a method used to link records from different datasets by using mathematical probabilities to calculate the likelihood that a pair of records refer to the same person or entity, even when unique identifiers are not available or cannot be used.

Logical tests were applied to CaT-Link data to assess the opportunity for linkage errors to affect the AIHW counts of MBS/PBS-treated MBC prevalence through the counting of apparently:

- live cases who were actually deceased ('false negatives'), and

- deceased cases who were actually still alive ('false positives').

Preliminary analyses show that, at the end of 2019, 337 out of 6,896 prevalent MBS/PBS-treated MBC cases (4.9%) had not accessed any MBS or PBS service throughout 2019 (Table 1). These cases are likely to not be prevalent cases (i.e., living in Australia) and are therefore more likely to be deceased or to have moved, or be travelling, overseas.

Table 1: Percentage of MBS/PBS-treated MBC prevalent cases with linkage error, by jurisdiction, at end of 2019

State of usual residence at MBC diagnosis	Percentage of apparently live MBC patients with no MBS/PBS service for 1 year before 31/12/2019	Percentage of apparently deceased MBC patients at end of 2019 with MBS or PBS activity in 2020-22 (as a percentage of MBS/PBS-treated prevalence at end of 2019)
NSW	4.4	1.8
VIC	4.9	2.1
QLD	6.9	1.4
WA	4.6	1.1
SA	1.8	1.4
TAS	5.7	1.4
ACT	7.6	1.9
NT	17.2	4.3
Australia	**4.9**	**1.7**

Source: AIHW analysis of CaT-Link data

Another potential issue is cases flagged as deceased, but who are likely to actually be alive. Of the 22,245 MBC cases diagnosed between 2002 and 2019 and who were flagged as deceased by the end of 2019, 117 had MBS/PBS service activity in the period 2020 to 2023 and were therefore likely to have been living in Australia. This is equivalent to 1.7% of the 6,896 prevalent cases at the end of 2019.

While around a third of all prevalent MBS/PBS-treated MBC cases were born overseas, around two thirds of prevalent MBS/PBS-treated MBC cases, without any MBS or PBS activity in the past year, were born overseas. This provides some support for the hypothesis that (at least) some of these people may have been overseas during the relevant periods and may help explain why there are more false negatives than false positives (because with probabilistic linkage there is usually a similar percentage of false negatives and false positives).

On balance (from the discussion above), it appears that prevalence is more likely to be overcounted than undercounted as a result of potential data linkage errors and travel or relocation overseas.

It is unclear what level of linkage error was experienced by CINSW, CCV and CAQ when counting prevalent cases but, for the purposes of this work, error is assumed to be similar in each jurisdiction to the error for each state in CaT-Link (Table 1).

As a further check, assessment of the 'fact of death' flag was assessed within the ACD. The ACD (a national linkage of Australian cancer registration data and the national death index) has been continuously managed over the years with the aim of minimising such linkage errors and is therefore considered to be relatively error free (noting that there will still be some errors). When deaths appear to have been missed in the registry data received by the

AIHW, the state or territory registry is advised of any cases that, according to the NDI, are deceased. Registries can then update their data.

Cross-border issues

Because of the state-specific nature of data available to CINSW, CCV and CAQ, some cases will be missed.

Identification of MBC cases requires an initial diagnosis of breast cancer, followed by treatment by at least one service that relates specifically to metastatic breast cancer.

Cross-border issues can result in a state having access to fact of breast cancer diagnosis, but not the fact of MBC treatment, and vice versa.

If a person is diagnosed with early-stage breast cancer in one state/territory, then moves to another state/territory and subsequently develops and is treated for metastatic breast cancer, this case may not be counted by either the state in which the breast cancer was originally diagnosed or by the state in which the MBC is subsequently treated.

A similar situation exists where a person living with MBC in one state accesses all of their treatment in another state/territory, i.e. while the person has not migrated to another state, they are accessing their specialist MBC treatment in another state.

Both situations can result in missed counts. Unless at least one treatment occurs in the state in which a person was originally diagnosed with breast cancer, then individual states may have difficulty counting these, and the case may remain uncounted by both the state where the breast cancer was diagnosed and the state where the MBC is being treated.

This problem does not exist in the data available in CaT-Link, because data is available for the whole of the Australian population (i.e. there are no cross-border issues). However, in CaT-Link it is possible to count the number of cases who were diagnosed with breast cancer while living in one state, and who were then living in another state when they commenced MBS/PBS treatment for MBC. These cases may not be counted by states, and it is assumed that, because the same principles apply, a similar proportion of cases that access hospital or radiotherapy services will also be affected.

Nationally, 3.5% of prevalent MBS/PBS-treated cases are affected, but this varies from state to state (see Table 2 below).

In CaT-Link, 2.1% of prevalent MBS/PBS-treated MBC cases diagnosed between 2015 and 2019 access all of their MBC-specific MBS and PBS services in another state, with substantial variation between states indicating the potential for miscounting (Table 2).

It is unclear whether hospital and radiotherapy clinic services are similarly affected.

MBS and PBS data in the linked data available to CCV and CINSW could have potentially been allocated on the basis of the state in which cases live at the time of treatment, the state where cases live at some particular point in time, the state in which health service providers and pharmacies are located/registered, or all of these definitions.

Table 2: Percentage of prevalent MBS/PBS-treated MBC cases at the end of 2019, in each state, that were originally diagnosed with breast cancer in another state or were accessing all MBS/PBS treatments in another state

State of residence in which MBC treatment commenced	Percentage of prevalent cases for which breast cancer was originally diagnosed in another state	Percentage of all prevalent cases diagnosed 2015-2019, accessing all MBC-specific MBS and PBS services in another state
NSW	2.8	3.9
VIC	2.4	1.3
QLD	5.2	1.5
WA	2.4	0.4
SA	2.2	0.4
TAS	5.7	1.2
ACT	9.1	3.4
NT	13.8	15.2
Australia	**3.6**	**2.1**

Source: AIHW analysis of CaT-Link data.

The basis for allocation of MBS and PBS data to the linkage available to CCV is unclear.

MBS and PBS data in the linkage available to CINSW included MBS and PBS activity for all NSW residents diagnosed with primary breast cancer and registered with the NSW cancer registry; consequently cases diagnosed with breast cancer in NSW, who have moved to another state before being diagnosed with MBC and treated with MBC-specific MBS or PBS service will be included in CINSW counts, but similar cases that have been diagnosed with breast cancer in another state and then moved to NSW and subsequently diagnosed with MBC will not. However, people who were diagnosed with BC in another state but treated in NSW hospitals or diagnosed with BC in NSW but treated in hospitals in another state (except for ACT), would not be visible in CINSW counts.

CAQ did not have access to MBS and PBS data in the linkage available to them so this issue, in relation to MBS and PBS data, doesn't apply. However, like CINSW, CAQ would not be able to count cases which were diagnosed with BC in another state but treated in Queensland hospitals, or those diagnosed with BC in Queensland but then treated in hospitals in another state.

Because of current uncertainty about the basis for allocation of MBS and PBS data to states, this second type of cross border issue is acknowledged as potential, but no attempt has been made to adjust counts; consequently, prevalence counts may be slightly underestimated due to these factors.

Only with complete linked data for the whole of Australia would it be possible to be sure that all cases were counted.

Data sources used by CINSW, CCV, CAQ and AIHW to estimate prevalent cases at end of 2024

Table 3 below describes the range of data used by CINSW, CCV, CAQ and AIHW to estimate MBC prevalence in Australia at the end of 2024.

All other things being equal, the earlier the recruitment of people diagnosed with breast cancer, and the wider the range of data sources used to identify whether a person was treated for metastatic breast cancer, the more complete the prevalence count is likely to be.

CINSW had access to the deepest and broadest data with which to count prevalent cases, which was then projected from the end of 2020 to the end of 2024.

CCV was only able to identify cases for which the initial breast cancer diagnosis was from 2008, and they had access to MBS and PBS data only from 2018. CCV adjusted their counts to allow for the fact that recruitment started only from 2008.

CAQ was able to recruit breast cancer cases from 1982, and available treatment data extended back more than 15 years. It is assumed that some cases will be missed because of lack of access to MBS and PBS data which can be estimated by AIHW in this work.

MBS and PBS data was the only treatment data to which AIHW had access, but breast cancer recruitment was from as early as 1982, and MBC identification was possible from 2002. However, 2019 was the last year in which it was possible to identify new cases of MBC.

The opportunity for all contributors to count 'all' cases differs between organisations. So as to better calibrate approaches, the four organisations calculated and compared MBC prevalence estimates to the end of 2020, with recruitment of breast cancer cases from 2008.

CCV and CAQ were able to independently calculate similar rates for Victoria and Queensland, with NSW rates being a little higher (which is not unexpected given their access to more data). With access only to MBS and PBS data, AIHW was only able to compare MBS/PBS-treated counts.

Potentially, a trade-off exists between more complete data (a larger array of datasets and a deeper recruitment period) and linkage (as well as coding) error. More complete data will tend to improve prevalence estimates, but it also has the potential to increase the number of cases with linkage error.

Table 3: Data sources used by CINSW, CCV and CAQ and AIHW to estimate MBC prevalence at end of 2024

Database/type	NSW	VIC	QLD	AIHW
State Cancer Registry data	1972-2020	2008-2021	1982-2022	1982-2019
NSW Cancer Registry Episodes	2011-2020			
State Hospital admissions data	2001-2020	2007-2021	2001-2024	
ACT Hospital admissions data	2004-2020			
State Radiotherapy course data	2009-2020	2011-2021	2007-2024	
NSW Outpatient Systemic Therapies Oncology Dataset	2013-2020			
MBS	2003-2020	2018-2021		2002-2022
PBS	2002-2020	2018-2021		2002-2022
National Death Index	1980-2020	2008-2021	1982-2024*	1982-2022

Notes: *Excluding NDI data for 2023 2024

Source: pers com Cancer Institute NSW, Luc te Marvelde, Nathan Dunn, Brett Davis

Estimation of MBC prevalence at the end of 2024, using the Longitudinal Cohort Method

The previous section discussed background issues and factors affecting development of the methodology used to estimate national prevalence at the end of 2024 (five years beyond the most recently available cancer data in CaT-Link). Successive methods were developed and then successively discarded until the Longitudinal Cohort method (LCM) was devised.

This section describes the LCM and its application to estimating MBC prevalence for Australian men and women at the end of December 2024.

The Longitudinal Cohort Method

A number of different methods have been developed and considered for the estimation of national MBS/PBS-treated MBC prevalence at end of 2024. All have either methodological issues, rely on unstable data, or require assumptions.

However, in building and testing successive methods, what appears to be a transparent, reliable and replicable method (the LCM) was developed.

The LCM is used to estimate the ‘unknown’ values (pink cells) in Table 4, which then provides the counts required to adjust CCV and CAQ MBC prevalence estimates, which in turn are used to inflate the MBS/PBS-treated MBC prevalence counts for Australia and the other jurisdictions so as to report MBC prevalence for these populations.

This section describes how the estimates in the pink cells in Table 4 are calculated, and then how these estimates are used to estimate national and jurisdictional MBC prevalence at the end of 2024.

Using the LCM to estimate MBS/PBS-treated MBC prevalence at the end of 2024

Table 4 describes MBS/PBS-treated MBC prevalence censored at the end of successive years (2016 to 2022, and projected to 2024), by year of MBC diagnosis. The pink cells have been populated using the LCM, which is described in this section.

Table 4: MBS/PBS-treated MBC prevalence for Australian women, for successive censorship years to 2024, by year of MBC diagnosis

	Counts and estimates of prevalence at end of each censorship year								
MBC diagnosis year	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**	**2024**
2002	36	33	32	31	30	30	28	26	24
2003	84	80	79	74	71	66	60	56	52
2004	103	96	93	89	87	82	79	72	67
2005	110	101	92	88	86	82	80	77	70
2006	108	100	97	94	89	88	86	84	81
2007	108	97	90	84	77	73	65	64	62
2008	149	134	129	123	113	110	106	94	92
2009	153	133	123	113	105	102	95	92	82
2010	215	181	159	134	119	109	101	94	91
2011	241	210	179	158	137	130	123	114	106
2012	322	257	212	190	157	146	139	132	122
2013	436	361	286	243	212	189	168	160	151
2014	579	421	347	263	214	187	170	151	144
2015	1,665	1,331	1097	941	834	743	681	542	516
2016	1,588	1,168	929	757	643	560	498	444	405
2017		1,561	1,118	881	728	606	530	488	435
2018			1,424	1,010	764	634	560	498	450
2019				1,576	1,158	944	818	611	538
2020					1,605	1,162	917	766	631
2021						1,677	1,214	958	800
2022							1,751	1,267	1,001
2023								1,829	1,324
2024									1,911
All years	5,897	6,264	6,486	6,849	7,229	7,720	8,269	8,618	9,154

Notes:

Breast cancer recruited from 1982, MBS/PBS services used to identify metastatic disease since 2002.

Estimates in pink cells are derived from incidence and apparent survival based on counts in the other cells. In this table, estimates assume that incidence and survival increase by, respectively, 2.4% and 2.0% per year from 2020.

Source: AIHW analysis of CaT-Link Data.

The numbers in the uncoloured cells are actual prevalence counts extracted from CaT-Link.

The pink highlighted cells have been populated using the trends apparent for previous years to estimate the prevalence counts for 2023 and 2024, while accounting for the break in series after 2014.

Across rows, prevalence counts in successive years will be a function of survival at the time. Down columns, prevalence counts are contributed by women diagnosed in successive years,

which will be a function of both the rate at which new cases are diagnosed and the survival rate.

Finally, while changes in the number of cases treated with MBS or PBS will be strongly affected by the nature of (and the rules around access to) MBS and PBS services, the underlying direction of changes in MBC incidence and prevalence should be reflective of known trends in both national breast cancer and Victorian stage 4 (de novo) breast cancer.

The Longitudinal Cohort Method potentially provides a reliable pathway to estimating future MBS/PBS-treated MBC prevalence for all jurisdictions, with a minimum of intermediary steps.

Background and assumptions for this method

The break in series for MBC-specific MBS/PBS services between 2014 and 2015, the spike in incidence and prevalence in 2015, and the lack of incidence data beyond 2019, all combine to make prediction of incidence and prevalence beyond 2019 challenging.

Changes in incidence prior to 2015 relate to cases that accessed MBS or PBS but not the Herceptin program; however, from 2015 onwards, changes in incidence relate to cases that accessed MBS and PBS that had taken on the role of the Herceptin program. Therefore, no single trend can be ascribed for the entire period.

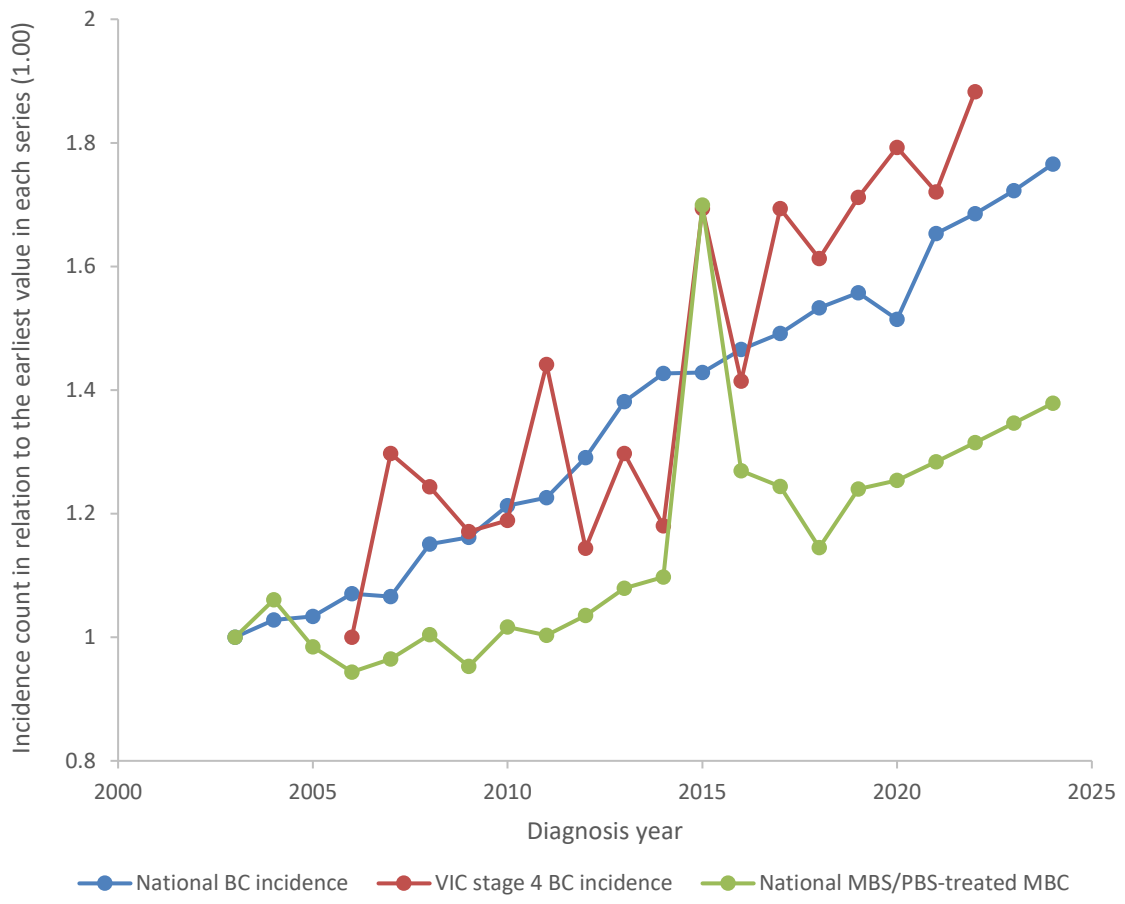
The spike in apparent incidence in 2015 is assumed to be associated with the transition from the Herceptin program to the PBS of previously uncounted MBC cases and is not considered to be part of the series to 2019.

Any trends for incidence and prevalence beyond 2019 will need to rely on data in the four years 2016, 2017, 2018 and 2019 only.

In the absence of a clear understanding of the incidence trend beyond 2019, it has been assumed that incidence of MBS/PBS-treated MBC cases will increase at a similar rate to that of breast cancer in Australia and stage 4 breast cancer in Victoria (two series for which data are available).

Figure 3 compares Australian breast cancer incidence, Victorian stage 4 breast cancer incidence, and Australian MBS/PBS-treated MBC incidence between 2003 and 2024, equivalised (so that all incidences are relative to the incidence at the start of each series).

Figure 3: Comparison of numbers of new cases of Australian MBS/PBS-treated MBC (including projections to 2024), Australian breast cancer, and Victorian stage 4 breast cancer, equivalised to 1.00 at the start of each series, 2003 to 2024



Note: National breast cancer (BC) incidence for 2021 to 2024 are projections from CdiA. MBS/PBS-treated MBC incidence for 2020 to 2024 are projections based on the percentage annual increase experienced by the other series, applied to the average MBS/PBS-treated MBC incidence for 2016-2019.

Source: AIHW analysis of CaT-Link and ACD data, and CdiA.

Based on regression lines through national breast cancer incidence from 2003 to 2019 (thereby excluding 2020, which is considered to be atypical due to the pandemic, and due to its position at the end of a series able to exert substantial influence on a regression line), Australian breast cancer incidence was increasing by an average of 2.3% per annum around 2020.

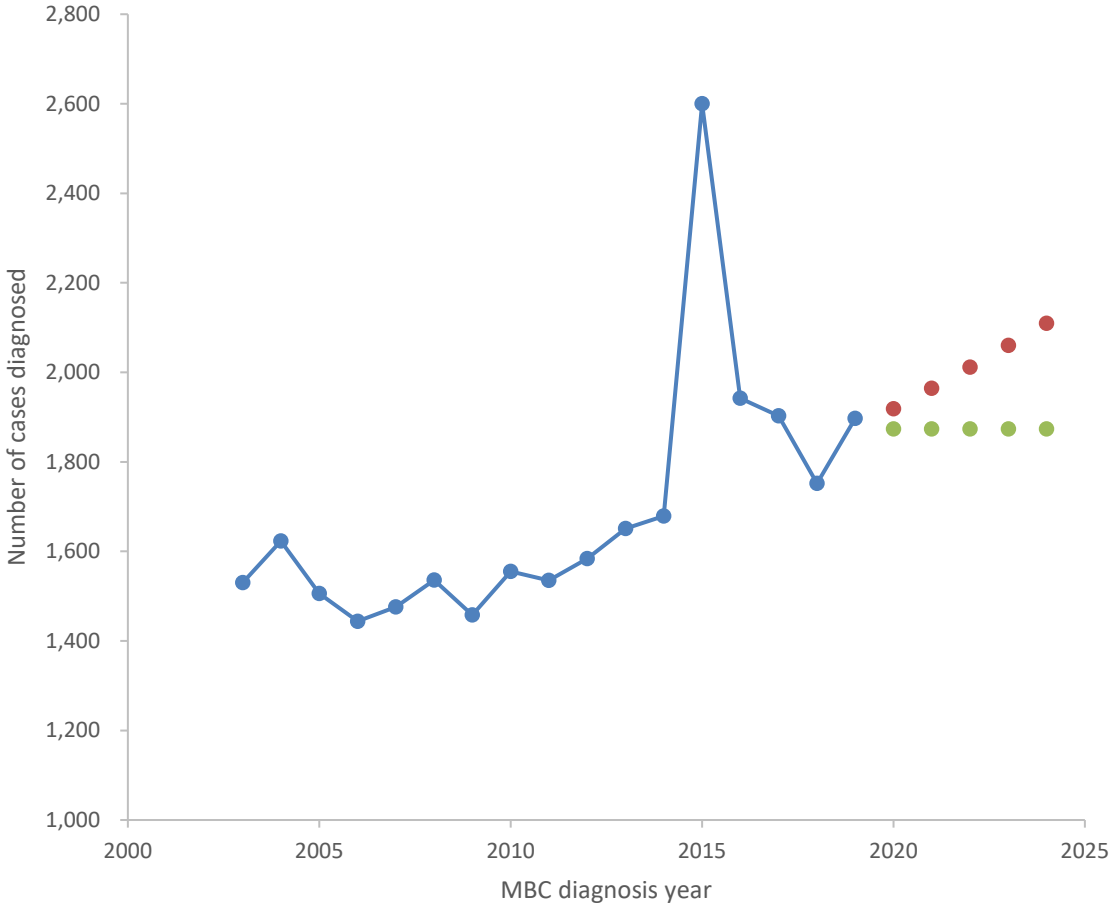
Based on regression lines through Victorian stage 4 breast cancer incidence for 2006 to 2022 (no reduction was observed during 2020, and additional data points were available to 2022), Victorian stage 4 incidence was increasing by 2.5% per annum around 2020.

In the absence of information to the contrary, Australian MBS/PBS-treated MBC incidence has been assumed to be increasing at a similar rate (2.4%) each year between 2020 and 2024 for this work.

MBS/PBS-treated MBC incidence is plotted in Figure 4, along with two proposed projections to 2024. Both red and green projections are based on the average MBC incidence between 2016 and 2019. The green series assumes no change in average MBC incidence between

2016 and 2024. The red series assumes that MBC incidence increases by 2.4% each year from the average incidence during the period 2016-2019.

Figure 4: Incidence of MBS/PBS-treated MBC, by diagnosis year, 2003 to 2024, Australia, with two series of projections to 2024



Notes:

Blue series describes actual MBS/PBS-treated MBC incidence counts to 2019.

Green series is the average incidence for the period 2016 to 2019 projected forward as potential incidence counts in 2020 to 2024.

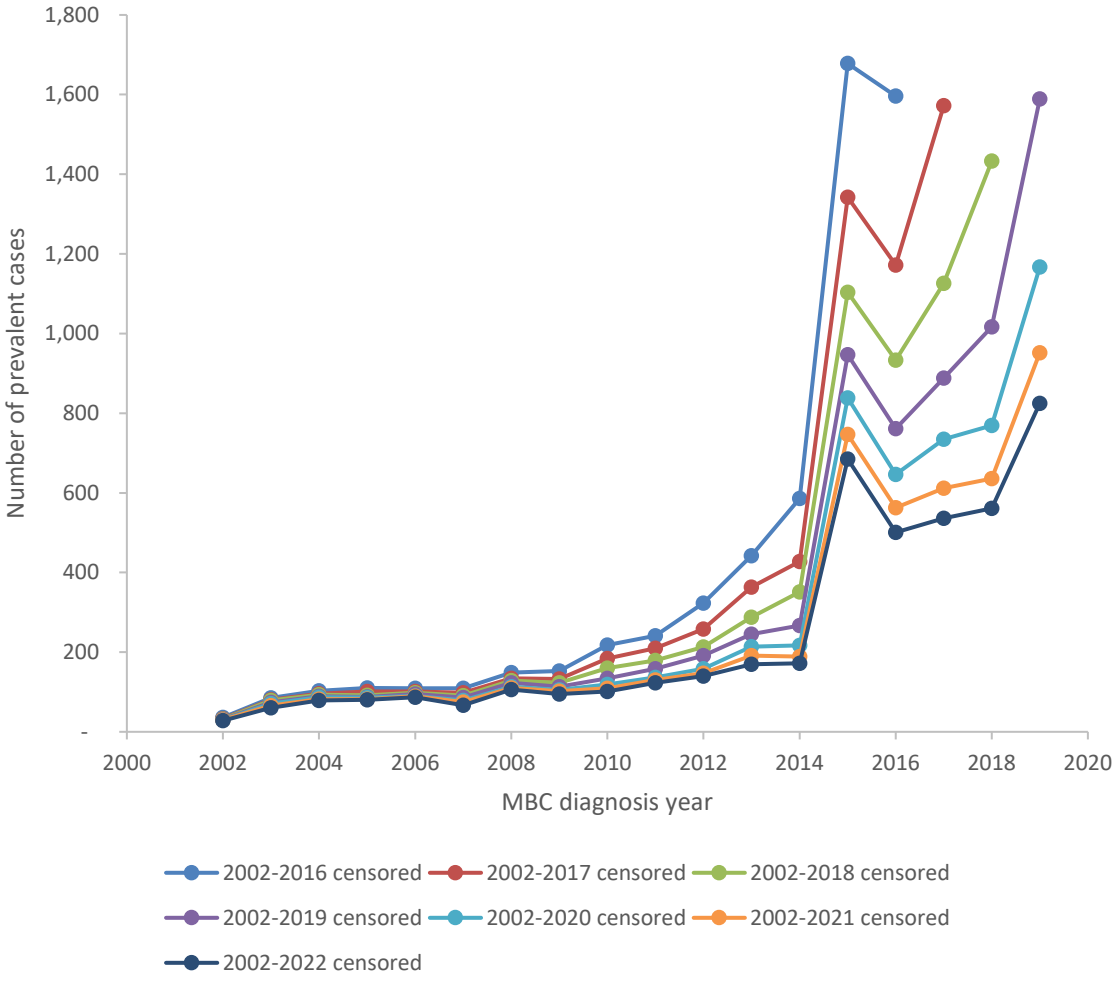
Red series is the unweighted average of the annual growth rate in the Australian breast cancer incidence count (2.3%) and the annual growth rate in the Victorian stage 4 breast cancer incidence count (2.5%), (the average taken as 2.4% pa), applied to the average incidence for the period 2016 to 2019, projected forward as potential incidence counts in 2020 to 2024.

Source: AIHW analysis of CaT-Link and CdiA data.

While, on balance, the red series appears more likely, the green series is also possible. Consequently, where MBC incidence is used to estimate future prevalence, both series are used, providing some (high and low) guard rails around estimates.

Figure 5 shows MBS/PBS-treated MBC prevalence censored at 2002 through to 2019 (the most recent year for which MBS/PBS-treated prevalence for people diagnosed in all previous years to 2002). While it is relatively straightforward to predict near-future prevalence based on the smooth growth in historic prevalence (e.g. for Australian breast cancer and for Victorian stage 4 breast cancer), Figure 5 demonstrates the difficulty of reliably estimating future MBS/PBS-treated MBC prevalence based on historic prevalence.

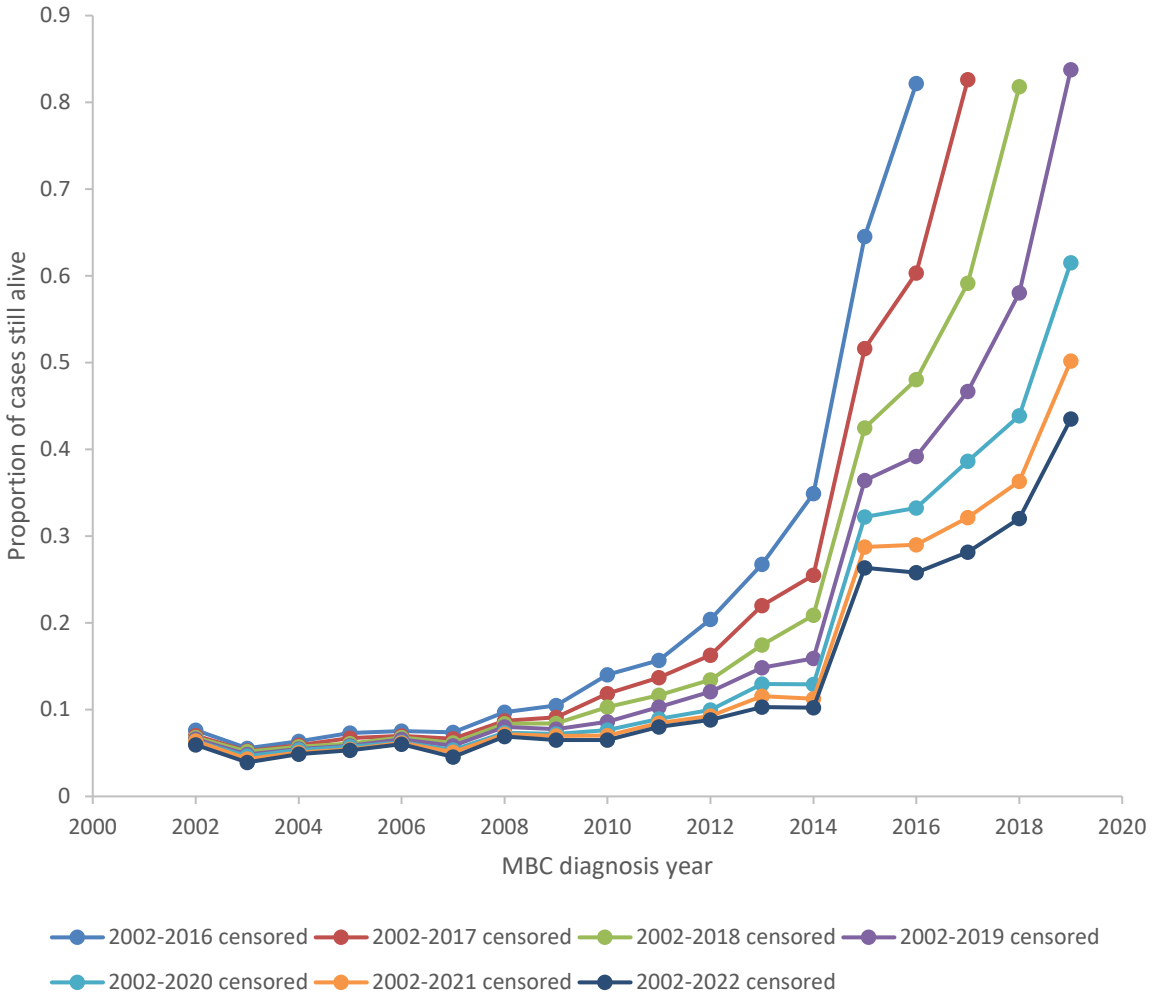
Figure 5: Prevalence of MBS/PBS-treated MBC cases at the end of each censorship year, by year of MBC diagnosis, 2002 to 2019



Source: AIHW analysis of CaT-Link data.

Figure 6 shows the percentage of MBS/PBS-treated cases diagnosed in each year, still alive at the end of successive censorship years. The break in series between 2014 and 2015 is obvious. While there is some variation from censor year to censor year, the survival experiences from 2016 to 2019 have been reasonably consistent.

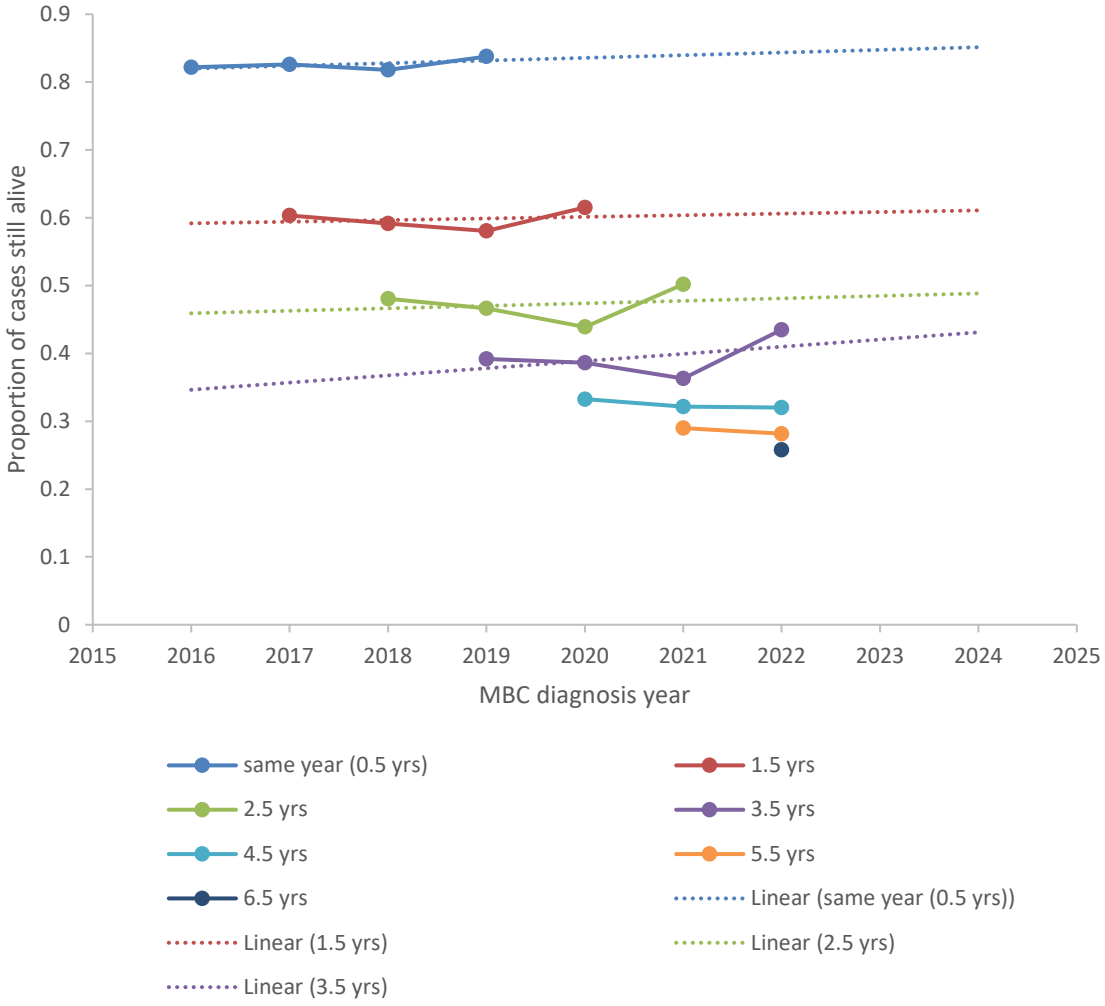
Figure 6: Proportion of incident MBS/PBS-treated MBC cases, still alive at end of each censorship year, by year of MBC diagnosis, 2002 to 2019



Source: AIHW analysis of CaT-Link data.

Figure 7 represents survival for MBC cases censored at 2019 to 2022 described in Figure 6, grouping survival by time from MBC diagnosis to censor date. Each series is based on a maximum of only four data points, so allocation of a trend is challenging, but it is reassuring that all trends based on no fewer than four data points have positive slope.

Figure 7: Proportion of incident MBS/PBS-treated MBC cases, still alive at the end of each censorship year, by year at diagnosis, by average time since MBC diagnosis



Source: AIHW analysis of CaT-Link data.

Progress in MBC treatment methods over the years would make it more likely that survival is increasing, but this is difficult to demonstrate (for MBS/PBS-treated cases) based on only four data points. Because it is not possible to describe a trend, survival for this period could be assumed to be constant.

For the purposes of estimating prevalence, survival for the period 2016 to 2024 could be calculated as the average survival for each of the periods in Figure 7 above. Consequently, the percentage of incident cases that were treated with MBC-specific MBS/PBS services and still alive would be assumed to be:

- 82.6% beyond the end of the year of MBC diagnosis
- 59.8% beyond the end of the first year after MBC diagnosis
- 47.2% beyond the end of the third year after MBC diagnosis

- 39.4% beyond the end of the fourth year after MBC diagnosis
- 32.5% beyond the end of the fifth year after MBC diagnosis
- 28.6% beyond the end of the sixth year after MBC diagnosis, and
- 25.8% beyond the end of the seventh year after MBC diagnosis.

It is plausible that survival is slightly higher in each successive censor year, but, from the available data, this is not clear or measurable.

Because it is more likely that survival is increasing over time, it is presumed that using these constant values for survival will slightly underestimate prevalence at the end of 2024. Alternatively, survival could be assumed to increase nominally by (2% per annum, for example) and the relative size of resultant prevalences compared as a form of sensitivity analysis.

In practice, MBC prevalence estimate calculated assuming that survival is constant (with the values described in the dot points above) has been found to be very similar to that calculated on the assumption that survival increases at 2% per annum from the average survival in 2016-2019.

The method for estimating unknown MBS/PBS-treated MBC prevalence

Necessarily, the methods used to populate pink cells will not be uniform for all of Table 4.

Estimation of 2020 to 2024 prevalence for years of diagnosis 2020 to 2024

Assuming incidence increases at 2.4% each year and that survival is constant, for 2020 diagnosed cases:

- the estimate for the 2020 censored prevalence count for cases diagnosed during 2020 is 82.6% of the estimated number diagnosed in that year (1.024 times the average number of cases diagnosed with MBS/PBS-treated MBC in 2016 to 2019)
- the estimate for the 2021 censored count for cases diagnosed in 2020 is 59.8% of the estimated number of MBS/PBS-treated MBC cases diagnosed in 2020 (see paragraph above)
- the estimate for subsequent censor years is the appropriate survival percentage (see dot points in previous section) of the estimated number of MBS/PBS-treated MBC cases diagnosed in 2020
- and so on.

Assuming incidence increases at 2.4% each year and that survival is constant, for 2021 diagnosed cases:

- the estimate for the 2021 censored prevalence count for cases diagnosed during 2021 is 82.6% of the estimated number of MBS/PBS-treated MBC cases diagnosed in 2021 (with that estimated number being 1.024 times 1.024 times the average number of prevalent cases diagnosed with MBC in 2016 to 2019).
- and then continue as for 2020 diagnosed cases above.

Assuming incidence increases at 2.4% each year and that survival is constant, for 2022 diagnosed cases:

- the estimate for the 2022 censored count for cases diagnosed during 2022 is 82.6% of the estimated number of MBS/PBS-treated MBC cases diagnosed in 2022 (1.024

times 1.024 times 1.024 times the average number of prevalent cases diagnosed with MBC in 2016 to 2019)

- and then continue as for 2020 diagnosed cases above.

Estimation for 2023 and 2024 uses a repetition of the same simple procedure.

If the assumption is that survival is not constant, but that it increases at 2% per annum on average, survival estimates above are multiplied by 1.02 for 2020, by 1.02 times 1.02 for 2021, and so on.

If the assumption is that incidence remains static, then the incidence of new cases of MBS/PBS-treated MBC diagnosed in each year from 2020 to 2024 is equal to the average number of cases diagnosed with MBS/PBS-treated MBC in 2016 to 2019.

Estimation of 2023 and 2024 prevalence for years of diagnosis 2015 to 2019

Estimation uses the same technique for censor years 2023 and 2024 as above, except the number of diagnosed cases is known, and doesn't need to be estimated.

Estimation of 2023 and 2024 prevalence for years of diagnosis 2002 to 2014

The number of prevalent cases of MBC in 2023, that were diagnosed in 2014, is calculated as the number diagnosed in 2014 who were prevalent at the end of 2022, multiplied by the ratio (for those diagnosed the year before (2013)) of the 2021-censored prevalence to the 2022 prevalence. This assumes that the relativity between the 2023 and 2022 prevalences for people diagnosed in 2014, will be the same as the relativity between the 2022 and 2021 prevalences for people diagnosed the year before.

Similarly, the relativity between the 2024 and (estimated above) 2023 prevalence for cases diagnosed in 2014, is assumed to be the same as the relativity between the 2022 and 2021 prevalences for people diagnosed two years before (2012).

This process is repeated for all years back to 2002 with the exception that, when there are no more previous years for which relativities can be calculated, the relativity for the earliest year is used. Consequently, the 2023 prevalence for people diagnosed in 2003 and in 2002, and the 2024 prevalence for people diagnosed in 2002 are all calculated using the ratio of the number of 2022 prevalent cases to 2021 prevalent cases diagnosed in 2002.

Estimation of total prevalence censored at the end of each year to 2024

The total prevalent cases at the end of each censor year is equal to the sum of the prevalence counts in each column of the table.

Use of the estimates in Table 4

In addition to documenting the number of MBC cases still alive at the end of 2024 that have accessed MBC-specific MBS or PBS services, the purpose of Table 4 is to be the first step in estimating the number of MBS and PBS services that may not have been counted by CCV and CAQ due to, respectively, limited or no access to MBS and PBS data in the data-linkages available to them.

Estimation of MBC prevalence using LCM estimates

Steps involve:

1. Use of the LCM to estimate MBS/PBS-treated MBC prevalence at the end of 2024 for Australia. This is achieved through using the LCM to populate the pink cells in Table 4. Use of national estimates to calculate incidence trends and survival will result in

less volatile estimates for jurisdictions but assumes that these relationships are identical across jurisdictions.

2. Using the same survival rates and assumptions about incidence and survival trends as for Australia, use the LCM to estimate prevalence of cases diagnosed with MBS/PBS-treated MBC in each year in NSW, VIC and QLD, who are still alive at the end of each year 2016 to 2024.
3. The number of cases that would not have been detectable by CCV and CAQ would then be estimated using the previously described rules (allocation of one third of the MBS/PBS-treated MBC prevalence count for any years in which MBS and PBS data were not included in the data-linkage used for counting cases). These estimated missing cases would then be added to the 2024 MBC prevalence estimates calculated by CCV and CAQ to yield the AIHW MBS/PBS-adjusted 2024 MBC prevalence for these states.
4. Calculate the ratio of aggregated NSW, VIC and QLD adjusted 2024 MBC prevalence (from 3 above) to the aggregated NSW, VIC and QLD 2022 prevalence of MBS/PBS-treated MBC (from the 2022 column of Table 4 (or from Table 5 below) – note that the 2022 prevalences will be actual counts and will therefore exclude prevalence of cases diagnosed in 2020, 2021 and 2022). Apply this ratio to the 2022 censored prevalence of MBS/PBS-treated MBC for each jurisdiction and for Australia (from Table 5 below) to yield the estimated MBC prevalence at the end of 2024.

Table 5: Prevalence at end of 2022, of MBS/PBS-treated MBC in Australian women, by residential jurisdiction at the time of MBC diagnosis, by MBC diagnosis year 2002 to 2019

MBC diagnosis year	State or territory of residence at the time of MBC diagnosis									
	NSW	VIC	QLD	WA	SA	TAS	ACT	NT	Unknown	Total
2002	6	n.p.	10	n.p.	6	n.p.	n.p.	n.p.	n.p.	28
2003	15	n.p.	9	7	13	n.p.	n.p.	n.p.	n.p.	60
2004	22	17	22	n.p.	7	5	n.p.	n.p.	n.p.	79
2005	27	18	22	n.p.	11	n.p.	n.p.	n.p.	n.p.	80
2006	21	29	24	n.p.	6	n.p.	n.p.	n.p.	n.p.	86
2007	16	16	15	6	5	n.p.	n.p.	n.p.	n.p.	65
2008	36	25	22	8	11	n.p.	n.p.	n.p.	n.p.	106
2009	31	28	18	n.p.	5	n.p.	5	n.p.	n.p.	95
2010	24	27	27	n.p.	8	n.p.	5	n.p.	n.p.	101
2011	43	28	29	8	10	n.p.	n.p.	n.p.	n.p.	123
2012	33	39	30	15	14	5	n.p.	n.p.	n.p.	139
2013	55	43	37	10	10	n.p.	8	n.p.	n.p.	168
2014	52	40	35	17	19	n.p.	n.p.	n.p.	n.p.	170
2015	207	169	133	82	63	7	12	n.p.	n.p.	681
2016	128	126	107	60	44	18	7	n.p.	n.p.	498
2017	133	134	142	57	42	7	8	n.p.	n.p.	530
2018	151	165	125	56	34	20	5	n.p.	n.p.	560
2019	233	223	180	79	60	21	15	n.p.	n.p.	818
Total	**1,233**	**1,144**	**987**	**419**	**368**	**108**	**79**	**39**	**10**	**4,387**

Note: Counts less than 5 (and counts that could be used to back-calculate counts less than 5) have been suppressed (n.p.).

Source: AIHW analysis of CaT-Link Data.

Assumptions and limitations

Development of national health data continues. While individual data sources are rich, connectivity between data sets and between jurisdictions remains limited.

Until such time as Australia's national health data capacity has sufficiently matured, it will not be possible to accurately count all cases of MBC (or indeed report on a wide range of other important issues) with certainty.

In this work, estimates have been calculated from what data was available at the time, specifically:

1. counts by CINSW, CCV and CAQ, of MBC prevalence in NSW, VIC and QLD, based on a wide range of treatment data sources.
2. counts by AIHW of MBC prevalence for MBC cases whose treatment included MBS or PBS subsidised services for all of Australia and for each of the states and territories.

National estimates have then been calculated by inflating the national count in 2 above, by the ratio of the aggregated NSW, VIC and QLD counts in 1 above divided by the aggregated NSW, VIC and QLD counts in 2 above.

This process is not perfect, but it provides a means of calculating a meaningful and representative estimate of the likely number of affected people.

There are a number of assumptions involved in this work, and there are several limitations, including the opportunity for over- or under-counting. These are outlined below, while their influence on the final calculation of MBC prevalence estimates is explained in the following section.

This work relies on numerous assumptions which are mentioned throughout this paper. In summary, it has been assumed that:

1. all cases of breast cancer are registered and have been included in this work because the ACD is known to be virtually complete.
2. CINSW, CCV, CAQ and AIHW counts are largely complete after any adjustment for limited access to MBS and PBS data, and for potential linkage and cross border issues. In practice, some counts may be less complete than others.
3. the probability that MBC is treated with MBS and PBS co-funded treatments is the same in all states and territories. This seems highly probable, notwithstanding the opportunity for sub-state regional differences. Jurisdictions with highly decentralised populations or with special arrangements may differ. Estimates of MBC prevalence for jurisdictions with lower levels of access to MBS and PBS could conceivably be lower than in reality. At a national level, the impact is assumed to be limited.
4. In the absence of clear trends, national MBS/PBS-treated MBC incidence is assumed to increase annually by 2.4% p.a. between 2016-2019 and 2024, in line with trends for breast cancer in Australia and stage 4 breast cancer in Victoria. Trends in states and territories are assumed to be the same as for Australia. Changing the assumption to a static incidence over this period was found to have little impact on MBC prevalence estimates.
5. In the absence of clear trends, MBS/PBS-treated MBC survival is assumed to either remain steady between 2016 and 2024 or increase annually by 2% p.a. from the average survival in 2016-2019, through to 2024. An assumption that survival has increased over time seems more plausible, but the size of that increase (in the

short-term) is unclear. Whichever assumption is adopted has little impact on the final MBC prevalence counts.

6. Adjustment of CCV and CAQ MBC prevalence estimates to reflect cases that are likely to have been missed because of limited access to MBS and PBS data, assumes that a third of MBS/PBS-treated cases alive at the end of 2024 will not have received any other MBC-specific treatment. This percentage is based on data from CCV and CINSW for 2018 to 2020. It is assumed that this pattern also applies to Queensland, and to the adjoining periods 2015-2017 and 2021-2024. Without adjustment, cases will be missed.
7. In CINSW data, the percentage of prevalent cases (diagnosed with breast cancer from 2008) at the end of 2020 that could be identified using MBS and PBS data was 14% versus Victoria's 16%, potentially as a consequence of CINSW greater opportunity to identify cases using more available data sources. Additionally, while this percentage remained relatively consistent over time for Victoria, it was lower in earlier years (11% in 2018 versus 16% in 2020) for CINSW, potentially demonstrating a greater opportunity for cases to be identified using other data sources the longer a case was receiving treatment. Estimation of potentially missed cases in Victoria and Queensland, assumes that a third of all prevalent cases treated with MBS or PBS will be missing from CCV and CAQ counts for these states. It is possible that this could overcount cases, but the effect on overall counts is assumed to be relatively small.
8. The Herceptin program was rolled into the PBS in 2015. AIHW, CINSW, CCV and CAQ had no access to these data. This deficit has been treated by assuming that, prior to 2015, the functionality of the Herceptin program plus the PBS is the same, from 2015, as for cases treated with PBS alone. All prevalent cases that were being treated under the Herceptin program are assumed to have been counted in 2015 for the first time when they transitioned to the PBS.
9. Linkage of deaths data with the other datasets is based on probabilistic techniques and so will never be 100% accurate. In CaT-Link, around five percent (4.9%) of cases prevalent at the end of 2019 appear to have no MBS or PBS activity of any kind during 2019, indicating that they are likely to be deceased (linkage error) or to have been overseas (not a linkage error, but no longer prevalent (or countable) in Australia). In addition, a number of MBC cases flagged as deceased at the end of 2019, were receiving MBS or PBS services in the following years and were likely alive (this corresponds to a number equivalent to 1.7% of MBS/PBS-treated cases prevalent at the end of 2019). In the absence of information to the contrary, this linkage error is assumed to also apply to CINSW, CCV and CAQ counts of total MBC prevalence, on the assumption that the causes of linkage error (and being out of the country) will be similar in all four linked data sets.
10. Identification of a case is dependent on knowing that a person has been diagnosed with breast cancer, and then that they are treated with an MBC-specific treatment. If a person is diagnosed with breast cancer in one state, then moves to another state, where they are later diagnosed with and treated for MBC, it is probable that they will not be visible as having MBC by either state. Because CaT-Link is a national data collection, the same problem does not apply to MBS/PBS-treated cases identified by AIHW. Nationally, 3.5% of MBS/PBS-treated prevalent cases would not be visible by states (with the exception of CINSW), and it is assumed that the same applies to cases who access other (e.g. hospital) treatments (including for CINSW).

11. It is assumed that MBC prevalent cases counted by CINSW, CCV and CAQ are cases that were living in NSW, Victoria and Queensland at the time of their MBC diagnosis.
12. In this work, state has been defined as the state of residence at the time of MBC diagnosis. It is known that state of residence at the time of MBS/PBS-treated MBC diagnosis is to all practical purposes, the same as the state of residence at the end of 2024.
13. The treatments that have been used to identify breast cancer cases that have developed metastatic disease were conservatively selected. Importance was placed on minimising the opportunity for non-metastatic cases being counted as metastatic. Treatments that could be used to treat either breast cancer at metastatic stage or breast cancer at an earlier stage have been excluded. Consequently, the true prevalence of MBC is considered more likely to be higher than reported here (rather than lower, noting that is also considered possible). However, CINSW was also able to establish that less than 2% of deceased cases of breast cancer did not access any of the MBC-specific treatments they had used to count live cases of MBC (You et al (submitted)). This is not to be interpreted that more than 98% of the prevalent cohort have been counted because the survival probabilities for the prevalence cohort (under one set of treatment regimens) could differ from those for deceased cases (who may be more likely to receive other treatments); but it is encouraging, and suggests that a large proportion of prevalent cases are likely to be counted using these specifications.
14. The number of cases counted will depend, all other things being equal, on the number of datasets available and the length of time over which people diagnosed with breast cancer can be recruited. CINSW, CCV and CAQ have each estimated MBC prevalence using the most complete data available to them at the time. Data available to CINSW was extensive in scope and over time. CCV was restricted to recruiting from 2008 onwards (but has adjusted so as to compensate) and had some (but not all) MBS and PBS data. CAQ recruits far back in time but did not have access to MBS or PBS data. Due to the relatively large size of their populations, prevalence counts by these three states are critical for estimating national prevalence. It is assumed, even with adjustment of counts to reflect likely uncounted MBS/PBS-treated cases, that national prevalence could still be unavoidably under or over-estimated to some degree.
15. In the MIAMOD methodology, all breast cancer deaths are assumed to be MBC deaths and survival is assumed to be constant for the study period from 2008 to 2022. Assuming that survival does indeed increase over time (which is considered likely), MIAMOD is considered likely to slightly under-estimate prevalence at the end of 2022 compared with the other method used in this work.
16. From a mathematical perspective, the MIAMOD methodology estimates of MBC prevalence at the end of 2022 and 2024 should be similar.

Impact of key assumptions on estimated MBC prevalence and rationale for choice of final model

The size of the MBC prevalence estimate is strongly determined by the CINSW, CCV and CAQ estimates for NSW, Victoria and Queensland. The MBC prevalence estimates for each of these states have their own strengths and weakness, with NSW potentially being the most

complete due to having more data sources and years of data available. Using NSW alone as the basis for scaling up the national MBS/PBS estimates would be one potential approach. This would yield a national MBC prevalence estimate of around 24,000.

However, a potentially safer approach is to use the combined estimates for NSW, Victoria and Queensland, yielding a lower, but more stable and potentially more geographically or nationally representative estimate of 20,800. The use of an 'average' approach to scaling up will inevitably mean that estimates for NSW, Victoria and Queensland in this report will be, respectively, lower than CINSW counts for NSW and higher than CCV counts for Victoria and CAQ counts for Queensland (both of which have also been adjusted for potentially missed cases that were treated with MBS and PBS).

The net effect of attempting to account for linkage errors and cross border issues has negligible effect on the calculated number of prevalent cases.

Assuming that MBS/PBS-treated MBC incidence remains either stable or increases at 2.4% pa after 2016-2019 has negligible effect on the 2024 MBC prevalence estimate.

Assuming that MBS/PBS-treated MBC survival remains either stable or increases at 2% pa after 2016-2019 also has negligible effect on the 2024 MBC prevalence estimate.

Adjusting the CCV and CAQ estimates for VIC and QLD, by adding to each, an estimate of the number of prevalent MBC cases likely missed due to limited or no access to MBS and PBS data, added around 1,000 cases to the national 2024 MBC prevalence count.

Based on all the available information, it is considered that the most reliable estimate:

- relies on CINSW, CCV and CAQ state estimates of MBC prevalence (in aggregate) to scale up and account for cases missed through treatment data available to the AIHW being currently restricted to MBS and PBS alone.
- Involves adjusting VIC and QLD MBC prevalence counts upwards, based on MBS/PBS counts from CaT-Link to account for cases missed through limited (CCV) or no (CAQ) access to MBS and PBS data.
- Though it makes little difference to the final estimate, incidence is assumed to increase at 2.4% p.a., and survival is assumed to increase at 2% p.a., for the period from 2016-2019 to 2024.
- Adjusts AIHW MBS/PBS-treated MBC prevalence counts for all states and territories, and CINSW, CCV and CAQ MBC prevalence estimates for their states down to account for the difference between counted cases that are likely deceased or overseas, and a smaller number of cases flagged as deceased that are likely alive (i.e. to account for linkage errors and cases that could be overseas).
- Adjusts CINSW (partially), CCV and CAQ MBC prevalence estimates for their states up to account for cross border issues related to cases originally diagnosed with breast cancer out of state.

The prevalence estimate calculated using MIAMOD, is a useful addition. The MIAMOD method calculates prevalence using a completely different and independent method. In this work, the MIAMOD estimate of 19,200 prevalent cases at the end of 2024 adds confidence to the other estimate, given that MIAMOD, relying on average survival across the entire study period (2008-2022), will tend to yield a lower prevalence than the other method for which the survival experience of prevalent cases will be more reflective of more recent times.

Trends

The increase in prevalence in Table 4 assumes increasing incidence and survival, is influenced by a greater opportunity to count cases in later years, and is potentially influenced by the change to the Herceptin program.

Notwithstanding, prevalence is expected to increase over the years.

Prevalence is dependent on the number of new cases being added each year (i.e. on incidence of new cases), and on survival (which influences the number lost each year due to death).

The prevalence of MBC will remain static if the number of new cases each year equals the number of deaths each year.

While it is unknown at this stage, it is assumed that the number of new cases of MBC is increasing at about the same rate as for breast cancer generally, and stage 4 breast cancer in Victoria (between 2.3% and 2.5% annually). Although it is not yet clear by how much, survival is assumed to be improving as treatments progressively improve. The logical conclusion is that the prevalence of MBC is currently increasing, either because of an increase in the number of new cases each year, or because of better survival (a smaller percentage of cases dying each year), or both.

However, until national data linkages are more comprehensive, it will be difficult to be entirely sure about the exact nature and size of any change over the years, and the degree to which this annual increase is slowing or speeding up.

Undiagnosed MBC

This work counts only those cases that have been diagnosed with MBC and for whom MBC-specific treatment has commenced. It does not include those cases whose breast cancer has metastasised but have not yet been diagnosed with MBC (or who have been diagnosed but treatment has not yet commenced).

The concept of present but undiagnosed MBC is beyond scope for this work, but the concept may have relevance to policy.

Approximately 2,000 prevalent cases of MBS/PBS-treated MBC are estimated to have been diagnosed during 2024, indicating that there were altogether very approximately, around 4,000 new prevalent cases of MBC diagnosed in that year, still alive by the end of that year (on the basis that for every prevalent case of MBS/PBS-treated MBC there is (very approximately) another prevalent MBC case that has received treatment that did not involve MBS or PBS services).

The number of additional undiagnosed cases of MBC prevalent at the end of 2024, will depend on the average time taken for newly metastasised breast cancer to become diagnosed and then treated. For example, if the average time between breast cancer becoming metastatic and then subsequently diagnosed (and treatment started) is one year, then it may be reasonable to assume that, as well as those cases that have been diagnosed and treated, there could be (very approximately) an additional 4,000 cases with the disease, that had yet to be diagnosed with MBC and treatment started.

The need for strategic national data development

Estimates produced using the methods described in this paper should be considered preliminary experimental estimates, indicative, and interpreted cautiously for the reasons expressed.

Development of national linked cancer data capacity in the AIHW National Health Data Hub will provide a progressively greater capacity to estimate MBC prevalence, with progressively greater confidence.

Such capacity is also expected to enable, for the first time, answers to a large number of other currently unanswerable questions about cancer, health treatment and systems, efficiency and effectiveness, and health system expenditure.

MIAMOD methodology to estimate MBC prevalence at the end of 2024

The MIAMOD methodology was developed by De Angelis et al (1994) and was previously used by Clements et al (2012) to estimate 2004 MBC prevalence in Australia.

MIAMOD (Mortality Incidence Analysis MODel) is a technique for estimating prevalence from death information, when incidence data is limited (e.g., when estimating metastatic breast cancer incidence).

Use is made of the fact that (virtually) all breast cancer deaths were, at some point in time, instances of metastatic breast cancer, because it is very unlikely for people to die from breast cancer that isn't metastatic (Palmieri et al 2022).

The Australian Cancer Database contains some observations of de novo MBC for 2006-20; from which it is possible to estimate survival probabilities (for de novo cases and (with modelling) for recurrent cases) by follow-up year since diagnosis and age groups 0-49, 50-69, 70+.

Using least squares regression, these survival probabilities are used to back-cast incidence counts up to 10 years before the year of death (after which the survival probability becomes negligible).

Imputed MBC incidence and observed breast cancer deaths can then be used to estimate prevalence at the end of a specific year.

Application of the MIAMOD technique to estimate MBC prevalence

At the time, the MIAMOD software was not available and so was programmed in SAS based on the method provided by De Angelis et al 1994, and information on its application in this context provided by Clements et al 2012.

The MIAMOD software has since been made available by Professor Roberta DeAngelis and the l'Istituto Superiore di Sanità.

Data and assumptions

All invasive breast cancer deaths are assumed to be metastatic breast cancer (MBC) deaths. Under this assumption, breast cancer death counts, disaggregated by age group and time-period, and MBC survival probabilities, are used to estimate MBC incidence (by age group and time-period).

MBC survival probabilities are assumed constant with respect to time of diagnosis. That is, survival is a function of diagnosis age group and the number of follow-up years since diagnosis. Survival probabilities for more than 10 years follow-up have been found to be negligibly small, and so, are set to zero in this particular application of the MIAMOD technique.

We observe breast cancer death counts for the 11 age groups 0-39, 40-44, 45-49, …, 80-84, 85+ and 5-year time periods. The time-periods used when estimating prevalence for the end of 2022 are 2008-12, 2013-17 and 2018-22. Hence, the dataset consists of 33 points generated by the cross-classification of age group and time-period.

Age group 0-39 consists of the eight youngest 5-year age groups and so will be referred to as age group 8; the remaining 10 age groups will be referred to as x = 9, 10, …, 18. Time periods 2008-12, 2013-17, 2018-22 will be labelled t=1, 2, 3.

The modelling and estimation of incidence

Let

x = age group (x = 8, 9, 10, …, 18)

t = 5-year time-period (t = 1, 2, 3)

$D(x,t) =$ invasive breast cancer death count for age group x and time-period t

$P(x,u) =$ MBC survival probability for u follow-up years in age group x

$\overline{P(x)} =$ median MBC survival for age group x (over 5 years) = $\big(P(x,2) + P(x,3)\big)/2$

For age group x and time-period t, the MBC incidence count $I(x,t)$ is modelled by

$$I(x,t) = C + Ax + Bt, \qquad (1)$$

where C, A, B are parameters to be estimated.

Given equation (1), it follows that the ratio $D(x,t)/\big(1-\overline{P(x)}\big)$ is equal to

$$\Big((C + Ax + Bt) + \big(C + A(x-1) + B(t-1)\big)P(x-1,5) + \big(C + A(x-2) + B(t-2)P(x-2,10)\big)\Big)$$

This model can be rewritten as

$$C\big(1 + P(x-1,5) + P(x-2,10)\big) + A\big(x + (x-1)P(x-1,5) + (x-2)P(x-2,10)\big) + B\big(t + (t-1)P(x-1,5) + (t-2)P(x-2,10)\big)$$

The 33 cross-classifications of age group (x) and time-period (t), each provide a data point of the form $\big(x, t, D(x,t)/\big(1-\overline{P(x)}\big)\big)$ – parameters C, A, B are estimated by applying ordinary least squares regression to the model, for $D(x,t)/\big(1-\overline{P(x)}\big)$, described immediately above. These parameter estimates will be denoted $\hat{C}$, $\hat{A}$, $\hat{B}$.

The prevalence estimates

For age group x, the estimated incidence counts for 2018-22, 2013-17, 2008-12 are

$$\hat{I}(x,3) = \hat{C} + \hat{A}x + 3\hat{B}$$
$$\hat{I}(x,2) = \hat{C} + \hat{A}x + 2\hat{B}$$
$$\hat{I}(x,1) = \hat{C} + \hat{A}x + \hat{B}$$

Of the $\hat{I}(x,3)$ patients MBC diagnosed in 2018-22, the number estimated to survive until the end of 2022, $S(x,2022)$, is given by

$$S(x,2022) = \hat{I}(x,3)\overline{P(x)}$$

Similarly, for the earlier two time periods, the corresponding estimated survivor counts are

$$S(x, 2017) = \hat{I}(x, 2)\overline{P(x)}$$

and

$$S(x, 2012) = \hat{I}(x, 1)\overline{P(x)}$$

At the end of 2022, the prevalent cases in age group *x*, where *x* = 10,11, …, 18, will be comprised of the following individuals.

- Patients in diagnosis age group *x*-2 diagnosed in 2008-12, and alive at the end of 2012, who survive another 10 years:
- Patients in diagnosis age group *x*-1 diagnosed in 2013-17, and alive at the end of 2017, who survive another 5 years.
- Patients in diagnosis age group x diagnosed in 2018-22 who are alive at the end of 2022.

Hence, the prevalence count is estimated by

$$S(x, 2022) + S(x - 1{,}2017)P(x - 1{,}5) + S(x - 2{,}2012)P(x - 2{,}10)$$

The corresponding prevalence estimates for age groups 9 and 8 are

$$S(9{,}2022) + S(8{,}2017)P(8{,}5) \text{ and } \mathrm{S}(8{,}2022)$$

Appendix A: Coding used to identify MBS/PBS-treated MBC

Table A1: MBS and PBS services used to identify de novo and recurrent metastatic breast cancer

Data source	Criteria	Description
MBS	Criterion 1	First radiation service with MBS item 15227, 15242, 15257, 15272, which are items that specify 'secondary site' and were available from May 2003. Since diagnosis information is not available in the MBS data, the MBS items were not included for the following people who had also been diagnosed with another primary cancer (other than breast cancer) since the radiation services could be aimed at treating another primary cancer: 1. People who were diagnosed with and died from another primary cancer other than breast cancer 2. People diagnosed with another primary cancer before this service was accessed (see note).
PBS	Criterion 2	First PBS dispensing of anti-neoplastic medicines using the ATC codes and PBS item numbers in Supplementary Table A2. Gemcitabine, Vinorelbine, and Megestrol can be used to treat metastatic breast cancer as well as other cancer types. However, PBS item numbers specific to metastatic breast cancer could not be found. Since diagnosis information is not available in the PBS data, these drugs were not included for the following people who had also been diagnosed with another primary cancer (other than breast cancer) since the drugs could be aimed at treating another primary cancer: 1. People who were diagnosed with and died from another primary cancer other than breast cancer 2. People diagnosed with another primary cancer before this service was accessed (see note).

Note: The criteria used by AIHW are an adaptation of the original CINSW method for identifying MBC cases using MBS and PBS data that were shared with CCV, CAQ and AIHW. The original criteria excluded cases which were diagnosed with another primary cancer 'with distant extent who are still alive'. While information about distant extent may have been available to CINSW, CCV and CAQ, it was not available to AIHW. Consequently, the second parts of criteria 1 and 2 above were modified to at least partially address the issue as detailed in the paragraph below.

Source: Pers com Cancer Institute of NSW, February 2025, based on You et al (submitted).

Criteria 1.2 and 2.2 have been adjusted because, in CaT-Link, it is not possible to identify stage at primary diagnosis. A partial solution, in the absence of stage data, is to apply a logical condition. If the other cancer type was diagnosed after the MBS or specific PBS items had been accessed, then the service is unlikely to have related to the other cancer, and so such cases could be assumed to be MBC. Conversely, if the other cancer is diagnosed before the MBS or specific PBS service had been accessed, then it would not be possible to know whether the service was to treat metastatic breast cancer or the metastatic form of the

other cancer, and the case would be excluded (if a conservative approach was preferred). Such an approach (excluding cases where the other cancer is diagnosed before the first MBS or specific PBS service) yields a national MBS/PBS-treated MBC prevalence of 4,715 compared with 4,579 and 5,013 if all affected cases are, respectively, excluded or included.

Table A2: ATC and PBS item numbers used to count PBS-treated MBC

ATC code	PBS item number	Description
L01BC05		Gemcitabine
L01CA04		Vinorelbine
L02AB01		Megestrol
L01FD01	10383L, 10391X,10401K, 10402L, 10798H, 10803N, 10811B, 10817H	Trastuzumab for Metastatic (Stage IV) HER2 positive breast cancer
L01FD02	10267J, 10308M,10333W, 10334X	Pertuzumab for Metastatic (Stage IV) HER2 positive breast cancer
L01XX41	10140Q 10144X	Eribulin for Metastatic breast cancer
L02AB02	14038W, 2728N	Medroxyprogesterone for advancer breast cancer
L01FD04	13713R, 13718B	Trastuzumab deruxtecan for Metastatic (Stage IV) HER2 positive breast cancer
L01EH01	11251E, 9148L	Lapatinib for Metastatic (Stage IV) HER2 positive breast cancer
L01FD03	10281D, 10282E	Trastuzumab emtansine for Metastatic (Stage IV) HER2 positive breast cancer

Source: Pers com Cancer Institute of NSW, February 2025, based on You et al (submitted).

Programming logic for counting MBC prevalence for MBS/PBS-treated cases in CaT-Link

Identifying in-scope invasive breast cancer patients

The programming steps, in the following two paragraphs, describe how invasive breast cancer patients were identified and partitioned into two datasets, ACD_EXCLUDE and ACD_OTHER_TIME. The former consists of patients whose only cancer diagnosis is breast cancer, ACD_OTHER_TIME consists of breast cancer patients who also have a cancer diagnosis at another site.

- Subset the ACD Extract in CaT-Link to all patients with an ICD10='C50' record and diagnosis date from 1982 to 2019 (which is the entire scope of the Extract); their unique personal identifier (AIHW_PPN) and their year and month of invasive breast cancer diagnosis is collected. This dataset was named ACD_MASTER.
- This master dataset is merged (using AIHW_PPN) with the ACD Extract sub-set to patients with ICD10 **not** equal to 'C50'. Patients belonging only to the master dataset are assigned to ACD_EXCLUDE, patients belonging to both datasets are assigned to ACD_OTHER_TIME. For patients in ACD_OTHER_TIME we also collect the year and month of their other cancer diagnoses.

MBC cases identified through MBS radiation services

The pertinent MBS Item Codes are 15227, 15242, 15257, 15272. The MBS Extract in CaT-Link was used, which has service information for 2002-2022, sub-set to these Item Codes – in creating this MBS subset, the scope was restricted to services up to and including 2019 (to

align with the final year of the ACD Extract) and collected patient unique identifier (AIHW_PPN) and year and month of service. This dataset has been combined with ACD_EXCLUDE and ACD_OTHER_TIME, in separate merging processes, to identify MBS-treated cases of MBC; details of these merging processes are as follows:

- ACD_EXCLUDE is merged with the MBS Extract subset (using AIHW_PPN), all patients that can be linked are classified as MBC diagnoses.
- ACD_OTHER_TIME is merged with the MBS Extract subset (using AIHW_PPN), a linked patient is classified as an MBC diagnosis if and only if they have a service occurring on or after their invasive breast cancer diagnosis and strictly before their other cancer diagnosis.

MBC cases identified through PBS services Gemcitabine, Vinorelbine, Megestrol

Gemcitabine, Vinorelbine, Megestrol have been defined in terms of PBS codes according to the following Item Code sets.

Gemcitabine = ('01134Q','01144F','01145G','04439P','05586B','05587C',

'05588D','05843M','05844N','05845P','05852B','05936K','05937L','07246J','08049P',

'08050Q','09401T','09402W','09414L','09463C').

Vinorelbine = ('04620E','05992J','05993K','07263G','08280T','08281W','09009E','09010F').

Megestrol = ('02731R','02734X').

The PBS Extract in CaT-Link, which has service information for 2002-2022, was sub-set to these Item Codes – in creating this PBS subset the scope was restricted to services up to 2019 and collected patient unique identifier (AIHW_PPN) and year and month of service. This PBS Extract subset, named PBS_SUBSET_1, has been combined with ACD_EXCLUDE and ACD_OTHER_TIME, in separate merging processes, to identify PBS-treated cases of MBC; details of these merging processes are given in the following two paragraphs.

- ACD_EXCLUDE is merged with PBS_SUBSET_1 (using AIHW_PPN), all patients that can be linked are classified as MBC diagnoses.
- ACD_OTHER_TIME is merged with PBS_SUBSET_1 (using AIHW_PPN), a linked patient is classified as an MBC diagnosis if and only if they have a service occurring on or after their invasive breast cancer diagnosis and strictly before their other cancer diagnosis.

MBC cases identified through other PBS services

The relevant PBS item codes for this part of the MBC identification procedure are contained in the last 7 rows of Table A2 above.

The PBS Extract in CaT-Link was sub-set to these specific Item Codes – in creating this PBS subset the scope was restricted to services up to 2019 and collected patient unique identifier (AIHW_PPN) and year and month of service. This PBS Extract subset, named PBS_SUBSET_2, has been merged with ACD_MASTER (using AIHW_PPN) - all patients that could be linked were classified as MBC cases.

The counting of MBC incidence and prevalence

MBC cases identified through MBS, Gemcitabine, Vinorelbine, Megestrol or other PBS services (as described above) were combined into a single SAS dataset. Incidence for 2002-

19 was obtained by sub-setting this combined dataset to the first occurrence of each AIHW_PPN (the unique patient identifier) and noting the resulting record count.

The corresponding prevalence count at the end of a censor year (such as 2019 or 2020) was then obtained by merging this incidence dataset with the National Death Index (NDI) Extract in CaT-Link (using AIHW_PPN). The prevalence is then the count of patients for whom there is either no corresponding NDI record, or the corresponding record specifies a DEATH_YEAR strictly after the censor year.

Acknowledgements

This methods paper was written by Andrew Phillips and Brett Davis, under the supervision of Justin Harvey and Amy Young.

Direction and advice were provided by members of the Australian Cancer Data Alliance (ACDA) Metastatic Breast Cancer Working Group, in close association with Breast Cancer Network Australia (BCNA).

The work builds on methods for counting cases of metastatic breast cancer (MBC) developed by Cancer Institute NSW (CINSW) and relies heavily on recent estimates of MBC prevalence, as well as some more detailed counts, produced by CINSW, Cancer Council Victoria (CCV) and Cancer Alliance Queensland (CAQ).

In particular, Hui You, Nicola Creighton and Richard Walton of the CINSW, Luc te Marvelde of the CCV and Nathan Dunn and Danica Cossio of the CAQ provided expert advice on MBC estimation in NSW, Victoria and Queensland.

Mark Short provided advice particular to the Australian Cancer Database (ACD).

National cancer registration data (the ACD) would not be possible without ongoing contribution of essential data from all members of the Australasian Association of Cancer Registries.

Professor Roberta De Angelis of the l'Istituto Superiore di Sanità kindly provided access to MIAMOD software.

Abbreviations

AACR: Australasian Association of Cancer Registries
ACD: Australian Cancer Database
ACDA: Australian Cancer Data Alliance
AIHW: Australian Institute of Health and Welfare
ATC: Anatomical Therapeutic Chemical classification system
BCNA: Breast Cancer Network Australia
CAQ: Cancer Alliance Queensland
CdiA: Cancer data in Australia
CaT-Link: Cancer and Treatment Linked analysis asset
CINSW: Cancer Institute NSW
CCV: Cancer Council Victoria
MBC: Metastatic breast cancer
MBS: Medicare Benefits Scheme
MIAMOD: Mortality Incidence Approach MODel
NCDF: National Cancer Data Framework
NDI: National Death Index
NHDH: National Health Data Hub
PBS: Pharmaceutical Benefits Scheme
WHO: World Health Organization

NSW: New South Wales
VIC: Victoria
QLD: Queensland
SA: South Australia
WA: Western Australia
TAS: Tasmania
ACT: Australian Capital Territory
NT: Northern Territory

Glossary

Anatomical Therapeutic Chemical classification system (ATC): The ATC (ATC/DDD) is a drug classification system, in which drugs are divided into different groups according to the organ or system on which they act and their chemical, pharmacological and therapeutic properties. The ATC is maintained by the WHO Collaborating Centre for Drug Statistics in Oslo.

Age-standardisation: a method of removing the influence of age when comparing populations with different age structures. This is usually necessary because the rates of many diseases vary strongly (usually increasing) with increasing age. The age structures of the different populations are converted to the same 'standard' structure; then the disease rates that would have occurred with that structure are calculated and compared.

Cancer: refers to a large range of diseases in which some of the body's cells become defective, begin to multiply out of control, can invade and damage the area around them, and can also spread to other parts of the body to cause further damage.

Cancer incidence: the number of new cancers diagnosed during a specified time period (usually one year).

Cancer mortality: the number of deaths occurring during a specified time period (usually one year) for which the underlying cause of death is cancer.

Chemotherapy: The use of drugs (chemicals) to prevent or treat disease, with the term being applied for treatment of cancer rather than for other uses.

Crude rate: The number of events in a given period divided by the size of the population at risk in a specified time period. death due to cancer: A death where the underlying cause is indicated as cancer.

Death due to cancer: A death where the underlying cause is indicated as cancer.

De novo: The first occurrence of cancer in the body.

Diagnosis: The process of identifying cancer based on its signs and symptoms. A definitive diagnosis of cancer can only be made by a pathologist.

Incidence: The number of new cases (of an illness or event, and so on) in a given period.

MIAMOD: Mortality Incidence Approach MODel, a statistical method used to estimate disease prevalence based on incidence and deaths data alone.

Medicare Benefits Scheme: The MBS is part of Australia's public health insurance scheme. Through the Scheme, the Australian Government subsidises the costs of a wide range of health services.

Metastasis: See Secondary site cancer.

Metastatic breast cancer: Breast cancer that has spread to other parts of the body.

Mortality due to cancer: The number of deaths that occurred during a specified period (usually a year) for which the underlying cause of death was recorded as cancer.

Pathology: The study of disease processes. A specialist in this field is called a pathologist. Sub-specialised diagnostic activities with relevance for cancer are histopathology/histology (microscopic examination and description of tissue) and haematopathology (the microscopic examination and description of blood and lymph).

Pharmaceutical Benefits Scheme: The PBS is a national, government-funded scheme that subsidises the cost of a wide range of pharmaceutical drugs for all Australians. The Schedule of Pharmaceutical Benefits (schedule) lists all the medicinal products available under the PBS and explains the uses for which they can be subsidised.

Prevalence: the number of people alive with a prior diagnosis of cancer at a given time.

Primary cancer: A tumour that is at the site where it first formed (see also secondary site cancer).

Probabilistic data linkage: a method used to link records from different datasets by using mathematical probabilities to calculate the likelihood that a pair of records refer to the same person or entity, even when unique identifiers are not available or cannot be used.

Projection: Longer-term extrapolation of recent trend data using unknown parameters such as expected future populations.

Radiotherapy: The use of high-energy radiation from x-rays, gamma rays, neutrons, protons, and other sources to kill cancer cells and shrink tumours. The radiation source may be applied externally, or internally.

Recurrence: Cancer that has returned (recurred) after a period of remission. The cancer may recur at the primary site, or elsewhere in the body, as a secondary tumour.

Relative survival: the ratio of observed survival of a group of persons with cancer to expected survival of those in the corresponding general population after a specific interval (such as 1, 3 or 5 years) following diagnosis.

Risk factor: Any factor that represents a greater risk of a health disorder or other unwanted condition or event. Some risk factors are regarded as causes of disease, others are not necessarily so. Along with their opposites, namely protective factors, risk factors are known as ‘determinants’.

Screening: Testing or examination of asymptomatic individuals for a specific cancer. The screening process may be indiscriminate, opportunistic (during a routine health check) or systematic.

Secondary site cancer: A tumour that originated from a cancer elsewhere in the body. Also referred to as a metastasis.

Stage: The extent of a cancer in the body. Staging is usually based on the size of the tumour, whether lymph nodes contain cancer, and whether cancer has spread from the original site to other parts of the body.

Survival: A general term indicating the probability of being alive for a given amount of time after a particular event, such as diagnosis of cancer.

Targeted (molecular-based) treatments: Improved understanding of the molecular characteristics of tumours, and the genetic causes, have led to targeted, precision or personalised treatments for some cancers. An example is the use of Herceptin to treat HER-2 type breast cancers (those that have a mutation of the *HER-2* gene).

References

AIHW (2025a) *Cancer data in Australia*, AIHW, Australian Government, accessed 16 October 2025.

BCNA (2025a). Groundbreaking data sees Victorians living with metastatic breast cancer counted for first time. BCNA, accessed 24 October 2025.

BCNA (2025b) Landmark step for women with metastatic breast cancer in Queensland. BCNA, accessed 24 October 2025.

Clements MS, Roder DM, Xue Qui Yu, Sam Egger, Dianne L O'Connell (2012) Estimating prevalence of metastatic breast cancer: a means of filling a data gap. *Cancer Causes Control*, 23, 1625-1634.

De Angelis G, De Angelis R, Frova L, Verdecchia A (1994). MIAMOD: a computer package to estimate chronic disease morbidity using mortality and survival data. *Computer methods and programs in biomedicine*, 44, 99-107.

Lord SJ, Kiely BE, Pearson SA, et al 2019. Metastatic breast cancer incidence, site and survival in Australia, 2001–2016: a population-based health record linkage study protocol. *BMJ Open* 2019; 9: e026414.

Lord SJ, Daniels B, Kiely BE, et. al. 2022. Long term risk of distant metastasis in women with non-metastatic breast cancer and survival after metastasis detection: A population-based linked health records study. Med J Aust. 2022; 217: pp. 402-9.

Palmieri, C., Owide, J., & Fryer, K. (2022). Estimated Prevalence of Metastatic Breast Cancer in England, 2016-2021. JAMA network open, 5(12), e2248069. https://doi.org/10.1001/jamanetworkopen.2022.48069.

PBS 2015. Transitioning of Herceptin subsidy to the Pharmaceutical Benefits Scheme (PBS). Canberra: Department of Health, Disability and Ageing. Viewed 2 November 2021, herceptin-faq-1-july-2015.pdf.

Smith AL, Yu XQ, O'Connell DL, et. al 2025. Metastatic breast cancer prevalence in New South Wales, Australia, in 2016: A health record linkage study. Asia Pac J Clin Oncol. 2025; 21: pp. 407-14.

You, Creighton, Walton, Lord, Smith, Kiely, Roder, Rushton, O'Brien. Estimating the number of people living with metastatic breast cancer: a population-based retrospective cohort study using linked health data. (submitted).

Related publications

The following publications may be of interest:

Cancer data in Australia provides detailed statistics on incidence, mortality, prevalence survival, for all major cancers in Australia since 1982.

BreastScreen Australia monitoring report 2025 provides statistics on Australia's national breast screening program.

Cancer topic page at AIHW, provides a portal to Australian cancer data generally as well as links to each of the state and territory cancer registries which also report data for their jurisdictions.