A review of cancer related surrogate outcomes used for PBAC decision making

The Cancer Surrogate Report for the PBAC
July 2023

Prepared by

Kah-Ling Sia, Leonie Picton, Esa Chen, Karen Yong, Dennis Petrie, Peter Ghijben, Gopisankar Mohanannair Geethadevi, Kaung Mon Winn, Anna Crothers and Jing Jing Li

Monash University Evaluation Group

Centre for Health Economics

Monash Business School

Monash University

Contents

Executive summary

Surrogate outcomes, such as time to cancer progression, can be early predictors for the long-term efficacy and effectiveness of new treatments. They have become increasingly important in cancer medicine research and clinical practice due to contamination of randomised controlled trials (RCTs) from subsequent treatments after the primary outcome is reached and increasing pressures for early access, registration and subsidy of new effective medicines. While evidence on surrogate outcomes can help predict the likely impact on more final outcomes, such as overall survival (OS), they may also provide poor predictions if treatments compress morbidity rather than delay or prevent mortality. When validation frameworks are applied, few surrogate outcomes used in clinical trials have been shown to be strong predictors of OS (Gyawali et al., 2020; Haslam et al., 2019; Savina et al., 2018; Walia et al., 2022). Decision makers are often faced with the difficult task of making recommendations about the long-term effectiveness and cost effectiveness of new treatments based on limited evidence.

The objective of the Cancer Surrogate Report is to provide a narrative review to describe how surrogate outcomes have informed past PBAC decisions for cancer treatments and to explore whether high-quality evidence on the validity of these surrogate outcomes already exists. A final aim of this research is to identify areas where further research, such as systematic reviews and/or meta-analyses, or observational studies based on registry and/or administrative data, are required to provide better evidence for future health technology assessment (HTA) decision making using surrogate outcomes. We conducted a detailed examination of PBAC decisions in cancer treatments in the last 10 years (January 2012 to May 2022) where a surrogate outcome was the key clinical evidence relied on by the PBAC in its recommendations. PBS public summary documents (PSDs) from the PBS website were used to identify relevant submissions and resubmissions of cancer drugs considered by the PBAC. Submissions or re-submissions were included for detailed data extraction if a surrogate outcome was used in lieu of OS or if the PSD clearly states that a surrogate outcome for another clinically meaningful outcome was relied upon for the clinical claim or PBAC decision. A literature review was also conducted to identify recently published, high-quality meta-analyses assessing the validity of surrogate outcomes in cancer therapies. Based on the evidence from the literature, a qualitative assessment was provided of emerging contemporary evidence on the validity of surrogate outcomes in cancer and, in particular, the strength of the relationship to final outcomes and identify areas of further research.

A total of 1691 PSDs for submissions/resubmissions to the PBAC were identified, of which 1206 were excluded as they were for either, non-cancer treatments or were supportive treatments in the cancer population (i.e., treatments for pain, nausea or adverse events). Out of the 498 cancer submissions/resubmissions considered between January 2012 and May 2022, reporting in the PSDs suggest that the PBAC's decision had at least partially relied on a surrogate outcome for 357 (72%) submissions. Of these submissions, given 110 also provided evidence of a statistically significant impact on OS in addition to the surrogate outcome, this meant for 247 (50%) submissions, the PBAC's decision was primarily based on the surrogate outcomes and thus met the eligibility criteria for detailed narrative review. Table ES1 summarises the relative use of surrogate outcomes by the most prevalent cancer indications and the PBAC recommendations for the included submissions/resubmissions.

Table ES1: Frequency of surrogate outcomes as main clinical evidence in PBAC submissions by cancer type and the PBAC's recommendations for submissions/resubmissions that primarily relied on surrogate outcomes

Cancer type	Total Submissions N	n (%)	Recommended, n (%)	Deferred, n (%)	Not recommended, n (%)	Not recommended due to immature OS^, n (%)
Lung cancer	66[a]	30 (45.5)	16 (53.3)	6 (20.0)	8 (26.7)	3 (37.5)
Blood cancer	152	83 (54.6)	34 (41.0)	10 (12.0)	39 (47.0)	23 (59.0)
Skin cancer	49[b]	28 (57.1)	14 (50.0)	3 (10.7)	11 (39.3)	6 (54.5)
Breast cancer	48[c]	31 (64.6)	12 (38.7)	3 (9.7)	17 (54.8)	12 (70.6)
Other cancers*	183	75 (41.0)	33 (44.0)	9 (12.0)	33 (44.0)	23 (69.7)
All cancers	**498**	**247 (49.6)**	**109 (44.1)**	**30 (12.1)**	**108 (43.7)**	**67 (62.0)**

N=number; OS=overall survival; PSD=Public Summary Document

^ Of the submissions for cancer that relied on a surrogate but were not recommended, the total number (%) not recommended due to inadequate or insufficient data on overall survival to support the clinical claim

* Other cancers were pooled together and include cancer types with less than 15 submissions/resubmissions

a Exclude 2 submissions for lung cancer and other cancers

b Exclude 1 submission for skin cancer and lung cancer

c Exclude six submissions for breast cancer and gastro-intestinal cancer

Of the 247 submissions/resubmissions for cancer that relied on a surrogate, 44% (109) received a positive recommendation and 44% (108) were not recommended by the PBAC. Of the submissions/resubmissions not recommended, 62% (67) were not recommended due to inadequate or insufficient data on overall survival to support the clinical claim. Breast cancer had the highest proportion of submissions that relied on surrogate outcomes (65%) and of those submissions with surrogate outcomes were also the most likely to be not recommended (55%).

Figure ES1 summarises the 247 PSDs considered by the PBAC for each year between 2012 and 2022 by their broad cancer types. Within the 247 included PBAC decisions, the PBAC had considered 91 drugs for 22 broad cancer indications.

Figure ES1: Submissions/resubmission included (N=247) in the detailed surrogate review by broad cancer types and considered by the PBAC between 2012-2022.
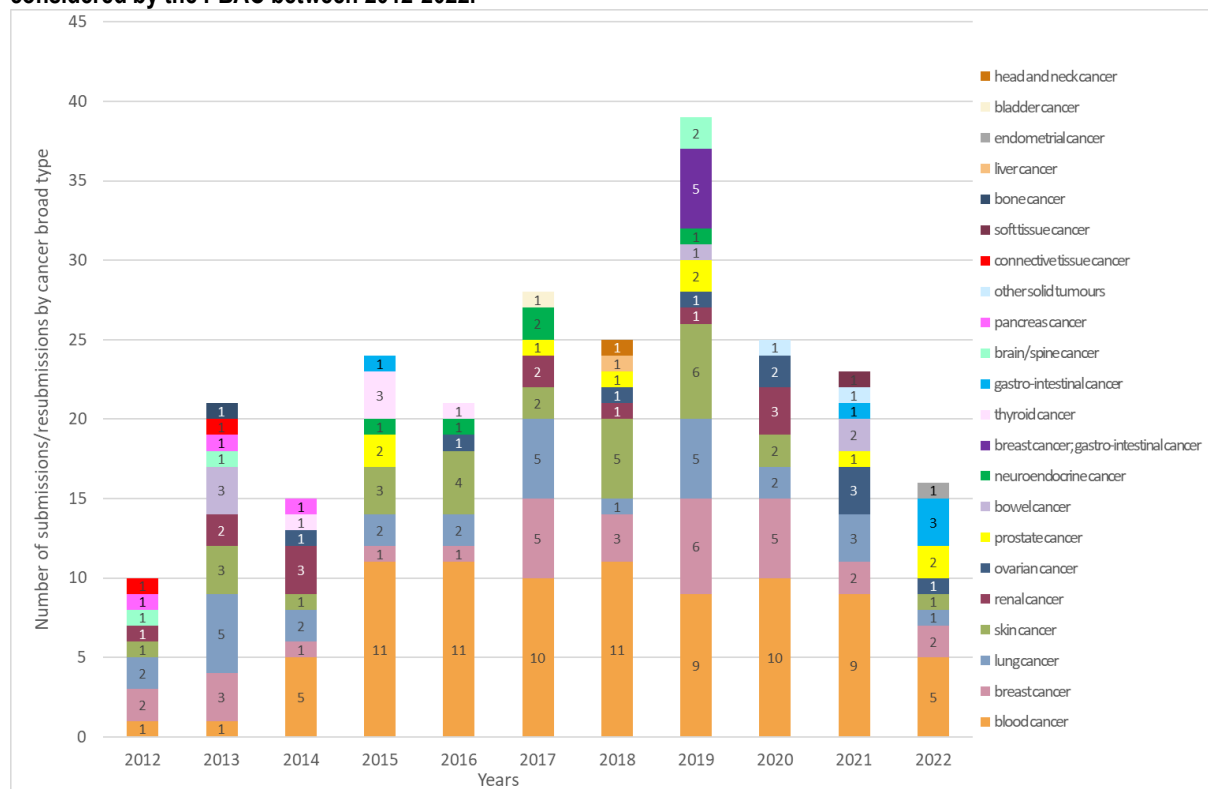
Source: Compiled during the review.

Table ES2 summarises the main surrogate outcomes reported in the submissions/resubmissions and relied on by the PBAC for its decision-making by cancer type. Progression-free survival (PFS) was the most commonly reported surrogate outcome, used in 76% (187/247) of submissions/resubmissions, and measured as either a primary or secondary outcome in the clinical trial(s) supporting the submission. Other commonly used surrogates include measures of response (i.e. overall/objective response rate), clinical/complete response (, in blood cancer), relapse/recurrence free survival (RFS, in skin cancer) and invasive disease-free survival (iDFS, in breast cancer). There are challenges relating to the consistency of evaluating surrogate measures in clinical trials. For example, the terms overall response rate (RR) and objective RR are often used interchangeably, however, objective RR is more commonly used for evaluating treatment effect on solid tumours and overall RR is more commonly used in blood cancers. Surrogate outcome definitions and how they are measured vary across trial protocols, which makes it difficult to compare outcomes across trials and elicit a reliable estimate of treatment effect. Standardised guidelines for measuring outcomes in clinical trials, such as the Response Evaluation Criteria in Solid Tumours (RECIST), are available for many cancers, however they are not universally adopted in clinical trials. Criteria are also updated periodically, making comparisons to older clinical trials more difficult. PBAC has had to consider the potential for bias in surrogate measurements when deliberating on the evidence presented in submissions.

Table ES2: Main surrogate outcomes relied on in PBAC submissions/resubmissions by cancer type

Cancer type (number of submissions)	BCR	Biomarker	Clearance of solar keratosis	Complete Remission	Clinical/Complete Response	pCR	Cytogenic response	DCR	DFS	DMFS	DOR	DRR	iDFS	EFS	Incidence of cancer	MFS	MRD	Objective RR	Overall RR	PFS	POF	response	RFS	Symptom progression	TTNT	TTP	VGPR
Blood (83)				4	10		2				3			2	1		1	10	14	63			3		1	1	4
Breast (31)						1^			4				6	1	1				2	21	1					1	
Lung (30)								1			1							13		27							
Skin (28)			3							3		1						6	2	14			7				
Renal (13)																		1		13							
Ovarian (10)																				10							
Prostate (9)		3§														4				2					2	1	1
Bowel (7)																		1		6							
Gastrointestinal (5)																		2		5							
Gastrointestinal + Breast~ (5)	1					3^					1								4	2						1	
Neuro-endocrine (5)																				5						1	
Thyroid (5)																				5							
Brain/spine (4)																		1		2		2				2	
Pancreatic (3)																				3							
Connective tissue (2)																				2							
Other solid tumours (2)																				2							
Bladder (1)																				1							
Bone (1)																							1				
Endometrial (1)																1				1							
Head & neck (1)																				1							
Liver (1)																				1							
Soft tissue (1)																				1							

BCR=breast conservation rate; C$_{trough}$=trough concentration; DCR=disease control rate; DFS=disease free survival; DMFS=distant metastasis free survival; DOR=duration of response; DRR=durable response rate; EFS=event free survival; iDFS=invasive disease-free survival; MFS=metastasis-free survival; MRD=minimal residual disease; pCR=pathological complete response; POF=premature ovarian failure; RFS=relapse/recurrence free survival; RR=response rate; TTNT=time to next treatment; TTP=time to progression; VGPR=very good partial response; §biomarkers for prostate cancer=serum testosterone and prostate specific antigen levels; ^includes primary breast tumour (bpCR) and total response (tpCR); ~submissions for biosimilar products

Of the submissions/resubmissions that primarily relied on surrogates, 65% (161) presented OS data based on interim trial results. In some instances, the pivotal trials supporting the submission have since published their final OS results. Table ES3 presents the comparison of the published final OS results versus the interim OS trial results relied on in the submissions for PBAC decisions. The final OS results were generally consistent with the interim OS results.

Table ES3: Comparison of the published final OS results versus interim OS results relied on in the submission

Cancer type	Trials^ with interim OS results at the time of submission	Statistical significance consistent[a] - HR consistent[d] (number of trials^)	Statistical significance consistent[a] - HR improved[d] (number of trials^)	Statistical significance consistent[a] - HR worsened (number of trials^)	Not consistent on statistical significance[a,b] (number of trials^)
Lung cancer	11	2	1	1	0
Blood cancer	38	8	0	1	3[c]
Skin cancer	9	5	0	0	1
Breast cancer	15	6	0	1	1
Other cancers*	28	8	0	1	2
All cancers	**101**	**29**	**1**	**4**	**7**

cf=compared to; HR=hazard ratio; OS=overall survival

^ the number of trials by drug-indication pairs (exclude resubmissions relying on the same trials).

* Includes broad cancer areas with less than 15 submissions/resubmissions, namely: bladder, bone, bowel, brain/spine, breast and gastrointestinal, connective tissue, endometrial, gastrointestinal, head and neck, liver, neuroendocrine, ovarian, pancreas, prostate, renal, soft tissue, solid tumours and thyroid.

a excluding trials for which the HR for OS were not reported in the PSD and trials for which this variable was not applicable (e.g. single arm studies used as the key clinical evidence)

b for our sample OS in final analysis reached statistical significance

c this included the CALGB trial (lenalidomide6) which had significant HR in interim analysis but had non significant HR in final analysis.

d HR consistent = HR within 0.1 difference compared to interim; HR improved = HR reduced >0.1 compared to interim; HR worsened = HR increased >0.1 compared to interim

In submissions/resubmissions that presented modelled economic evaluations (i.e., a cost-effectiveness or cost-utility analysis), the surrogate outcome from the clinical evaluations were used in 85% (140/165) of cases. Therefore, the cost-effectiveness was likely heavily dependent on the assumed relationship between the surrogate outcome and OS or other clinically meaningful future outcomes.

Common themes identified in the narrative review included the immaturity of data presented in the clinical claim, the lack of validation of surrogate outcomes, the quality of the data, including potential bias in measurement of outcomes, confounding from crossover to the active arm in clinical trials and sample size. Comments by PBAC relating to these themes were similar across cancer types.

Validation studies analysing the correlation between surrogate endpoints and OS have presented mixed results, with most studies showing low to moderate correlation between PFS or objective/overall RR and OS. For example, based on correlation, studies have reported PFS to be a validated surrogate for OS in diffuse large B cell lymphoma and non-Hodgkin Lymphoma but inconclusive across many other cancers (e.g., bowel, ovarian, prostate). The validation studies had used different criteria for assessing the correlation between surrogate and OS. Some studies report correlation at the patient level, but many only at the trial level for either aggregate measures or treatment effect. There was also no consistent threshold accepted as sufficient for validation of a surrogate, some have used $R^2 \geq 0.6$, which may not be robust enough for HTA where the surrogate outcome is used to predict OS beyond the available trial data. For some meta-analyses the disease stage and type of therapy were also unclear. Validation studies of surrogate outcomes using a consistent framework tailored to the needs of HTA is thus lacking.

We used PBAC meeting documentation (primarily PSDs) as our primary evidence. This was highly relevant as a stocktake of PBAC's prior decisions relying on surrogates, but was limited as a source of information to provide an in depth understanding of the validity of the surrogate outcomes. In particular, we found it difficult to characterise surrogate outcomes using our intended evidence framework (level of evidence, strength of association and quantification of the expected effect on the

patient centred outcome) based on information published in the PSDs. Further research using more extensive meeting documentation (including full submission documentation) may be required in order to assess how the validation of putative surrogate endpoints were addressed by the PBAC, perhaps focusing on specific surrogates or cancer types. We also found limited use and mention of the PBAC's guidance on surrogate outcomes (i.e., Appendix 5 of the PBAC guidelines) in the meeting documentation. This may indicate lack of focus in previous submissions on assessing the surrogate evidence or that a more simplified reporting framework is needed. This should be explored in further research.

Our review further highlights the need to weigh up the strength of available evidence on the validity of surrogate used across the cancer types and quantify how well the surrogate predicts final clinical outcome and hence the level of uncertainty around the treatment benefit and cost-effectiveness associated with the PBAC decisions. For example, more systematic reviews and meta-regressions of evidence between surrogate and final outcome is needed (Gyawali et al., 2020; Walia et al., 2022), starting with a priority cancer type. Validation may also depend on treatment class and cancer subtypes which may add further complexity. In addition, administrative data on medication use linked with death records for patients treated before and after the listing of new treatment could be utilised.

In cases where the PBAC had to make decisions relying on immature OS, this report highlighted the high availability of additional OS data from pivotal trials post the PBAC's decision. Another area of fruitful research would be to provide an in-depth analysis and comparison of the anticipated OS as modelled in the submissions versus actual observed based on final trial data for recommended submissions. Data from clinical registries capturing both surrogate and final outcomes and linking administrative datasets (e.g., PBS, MBS and National Death Index data) may be useful additional sources of evidence in this regard, particularly as they capture the benefits and costs of downstream treatments missing in RCT data.

The PBAC may also wish to understand the relationship between surrogate and final outcomes for specific case studies, for example, the anticancer activity of newer immunotherapy agents programmed cell death protein 1 (PD-1) and programmed death-ligand 1 (PD-L1) inhibitors may be different to other cancer therapies in that they may take longer to show an effect, and due to the immune response, may also induce a 'pseudoprogression' which impacts the measurement of true disease progression. A comparison of final trial data with the evidence presented in the submissions would aid in closing the long-term evidence gap with this class of drug. Cancer treatments may also have different treatment effects in different cancers. Understanding a treatment's real-world efficacy may deepen our understanding on how best to estimate its effect on OS using different surrogates. Different therapies in the same drug class may also have different long-term outcomes, even if they exhibit comparable early evidence at the time of the PBAC recommendations. In this case linked administrative data (e.g., PBS, MBS and National Death Index) could be useful to confirm the estimated final outcomes in the submissions.

Introduction

Surrogate outcomes, such as time to cancer progression, can be early predictors for the long-term efficacy and effectiveness of new treatments. They have become increasingly important in cancer medicine research and clinical practice due to contamination of randomised controlled trials from subsequent treatments after the primary outcome is reached and increasing pressures for early access, registration and subsidy of new effective medicines. While evidence on surrogate outcomes can help predict the likely impact on more final outcomes, such as overall survival (OS), they may also provide poor predictions if treatments compress morbidity rather than delay or prevent mortality. When validation frameworks are applied, few surrogate outcomes used in clinical trials have been shown to be strong predictors of overall survival (Gyawali et al., 2020; Haslam et al., 2019; Savina et al., 2018; Walia et al., 2022). Decision makers are often faced with the difficult task of making recommendations about the long-term effectiveness and cost effectiveness of new treatments based on limited evidence.

Current literature from international HTA agencies, including those in Europe, Canada, and the United States' Food and Drug Administration (FDA), show that whilst surrogate endpoints enable shorter clinical trials and therefore more expedient approval, reliance on such endpoints can be problematic if they fail to fully capture the complete risk-benefit profile of the health technology in terms of survival and toxicity (Grigore et al., 2020; Gyawali et al., 2020; Johnson et al., 2011; Walia et al., 2022). Reviews conducted on surrogate endpoints used in FDA applications have been shown to frequently overestimate treatment benefits and underestimate toxicity after long-term data is gathered (Gyawali et al., 2020; Walia et al., 2022). Medicines funded based on surrogate outcomes therefore may never translate into the estimated long-term benefits (i.e., longer survival and/or better quality of life) that was assumed and relied on in decision making.

As surrogate endpoints in medicine applications have become common to inform licensing arrangements, HTA agencies are under increasing pressure to also utilise such evidence in their recommendations (Kemp & Prasad, 2017). Many HTA agencies recommend surrogate endpoints only be used when they are strongly correlated with survival and quality of life (Buyse et al., 2010; Ciani et al., 2017). However, acquiring such information (particularly for rare cancers or cancer subtypes) can increase patients' delay to access medicines. In turn, the available treatment options are greatly reduced as patients wait longer to access medicines, which can limit their expected survival and overall quality of life.

In Australia, the PBAC has guidance for submissions using surrogate outcomes in [Appendix 5 of its Guidelines](). The surrogate outcome(s) chosen should be patient-relevant, have a current and consistent relationship with clinically relevant patient outcomes (e.g., OS), and be consistent with the health states defined in the natural history of the disease or condition. The submission should include justification for its use of surrogate outcomes, including any biological basis for using the surrogate measure, how the change in the surrogate measure is expected to alter clinically relevant patient outcomes, and whether anything is expected to change the relationship between the surrogate measure and the patient outcome. The surrogate outcome(s) should be well defined with units of measurement, the measurement tool/criteria, evidence supporting the reliability of the measurement tool/criteria, the variability across observers or different measurement tools, the measurement of the comparative treatment effect (e.g., odds ratio, standardised mean difference), and a minimal clinically important difference (MCID).

The objective of this report is to provide a narrative review to describe how surrogate outcomes have informed past PBAC decisions for cancer treatments and to explore whether high-quality evidence on

the validity of these surrogate outcomes already exists. A final aim of this research is to identify areas where further research, such as systematic reviews and/or meta-analyses, or observational studies based on registry and/or administrative data, are required to provide better evidence for future HTA decision making using surrogate outcomes. We conducted a detailed examination of PBAC decisions for cancer treatments in the last 10 years (January 2012 to May 2022) where a surrogate outcome was the key clinical evidence relied on by the PBAC in its recommendations. A literature review was also conducted to identify recently published, high-quality meta-analyses assessing the validity of surrogate outcomes in cancer therapies. Based on the evidence from the literature, we provide a qualitative assessment of emerging contemporary evidence on the validity of surrogate outcomes in cancer and, in particular, the strength of the relationship to final outcomes and identify areas of further research.

Method

We used the PBS PSDs from the [PBS website](#) to identify relevant submissions and resubmissions of cancer drugs considered by the PBAC between January 2012 and May 2022. PSDs were included if the submission or resubmission was for a cancer treatment (i.e., chemotherapy or anti-cancer treatment). Supportive treatments for cancer were excluded. Submissions or resubmissions were included for detailed data extraction if a surrogate outcome was used in lieu of OS (a common patient relevant clinical outcome in cancer) or if the PSD clearly states that a surrogate outcome for another clinically meaningful outcomes was relied upon for the clinical claim or PBAC decision. Table 1 summarises the key inclusion and exclusion criteria for identifying relevant submissions of cancer treatments with a PSD to be included in the narrative review.

Table 1: Eligibility criteria for submissions or resubmissions of cancer therapies in the review

Inclusion criteria	Exclusion criteria
- Submission or resubmission with a PSD available, that was considered by the PBAC between January 2012 and May 2022. - Submission or resubmission for a chemotherapy or other anti-cancer treatment. - Submission or resubmission where a surrogate outcome was used for OS or the PSD clearly states that a surrogate outcome for another clinically meaningful outcome was relied upon for the clinical claim or PBAC decision.	- Submission or resubmission for a supportive treatment in a cancer population (e.g., treatments for pain, nausea or adverse events associated with anti-cancer treatments). - The OS (or the claimed clinically meaningful endpoint) was statistically significantly different between treatment and comparator.

PBAC=Pharmaceutical Benefits Advisory Committee; PSD=public summary document; OS=overall survival

A Microsoft Excel data extraction form was created to record relevant data characteristics from all the included PSDs. The variables collected include:

- cancer type
- oncology medicine class
- surrogate measure used
- whether the surrogate outcome was the primary, secondary, or other outcome in key trials
- surrogate effect size (including confidence interval)
- other reported clinical outcomes
- whether the surrogate(s) were used as inputs in the economic mode,
- PBAC recommendation
- PBAC advice on the potential validity of the surrogate or the surrogate's MCID
- any documented comments in the PSDs on the link between surrogate and clinical outcomes relied on by the PBAC, including the level of evidence, strength of association and quantification of the expected effect on the final outcome.

Oncology medicine class for each medication-indication pair was additionally informed using the United States government [National Cancer Institute classifications](#).

Each PSD was assessed for additional descriptive statistics including submission type (e.g., whether it was a cost minimisation, cost utility, or cost effectiveness type submission) and comparator used in the trial to assist with the further understanding of the submissions. Any instances where the PBAC had considered the OS data presented alongside the surrogate outcomes to be immature (and when the OS reported was from a trial interim analysis), were also identified. In these instances, the final trial data was sought from the literature based on the National Clinical Trial number (Zarin et al., 2011). Given issues associated with loss of randomisation, attrition and patient selection, the search was

limited to results from the Phase II/III clinical trials and did not cover any extended follow ups that many clinical trials may have.

All primary data collected from PSDs were quality assured for accuracy, consistency and completeness by an experienced PBAC evaluator.

Descriptive analysis was conducted to understand the proportion of submissions where PBAC has had to rely on surrogate measures rather than clinical outcomes, the proportion of positive PBAC recommendations that relied on surrogate measures and the proportion not recommended due to inadequate or insufficient data on OS to support the clinical claim. To better understand where surrogate outcomes have been used this was also broken down by indication.

The detailed narrative review of the included decisions was conducted to describe and discuss how surrogate outcomes have informed past PBAC decisions for cancer treatments. We summarise any reported links between surrogate and clinical outcomes in the meeting documentation, including the level of evidence, strength of association and quantification of the expected effect on the final outcomes and how the evidence was viewed and relied on by the PBAC in its recommendations. We demonstrate these with quotes extracted from the meeting documentation, focusing on quotes from PSDs due to the commercial-in-confidence nature of other meeting documentation.

The narrative was grouped by cancer type and by oncology medicine class where there was sufficient sample size to provide this further breakdown.

The narrative review prioritised submissions where the surrogate measured was a primary or secondary outcome in the clinical trial evidence. The review also prioritised submissions where there was a superiority claim on the clinical benefit, i.e., rather than cost minimisation submissions where the validity of the surrogate was likely already at least partially accepted for the comparator treatment, although data from all submissions fitting the eligibility criteria were reviewed.

Content analysis was used to summarise any observable patterns and trends in the use of surrogate measures across the indications to identify important similarities or discrepancies in the use of surrogate measures between PBAC submissions for the same indication. NVivo software (released in March 2020) was used for the content analysis.

Recent high-quality meta-analyses were identified that assessed the validity of surrogate outcomes in cancer therapies. The search was conducted on MEDLINE, EMBASE, PubMed and Google Scholar for articles assessing the validity of surrogate endpoints in cancer research. Search terms included (cancer or neoplasms or oncology) and ((surrogate and validation) or (surrogate and correlation)) and (survival or mortality). The articles were then screened for quality and relevance and their reference lists were scanned to identify any further relevant articles. The included studies were grouped by cancer type and surrogate endpoint. A qualitative comparison of the information was provided on the correlation between the surrogate measure effect size and the effect size for the clinical outcomes and the assumed effect of the change in the surrogate measure on the clinical outcome considered or accepted by the PBAC. The report also identified areas where further detailed systematic review and meta-analyses may be needed to assist future decision making.

Results

A total of 1691 PSDs for submissions/resubmissions to the PBAC were identified between January 2012 and May 2022, of which 1206 were excluded as they were for either, non-cancer treatments or were supportive treatments in the cancer population (i.e., treatments for pain, nausea or adverse events). Out of the 498 cancer submissions/resubmissions considered between January 2012 and May 2022, reporting in the PSDs suggest that the PBAC's decision had at least partially relied on a surrogate outcome for 357 (72%) submissions. Of these submissions, given 110 also provided evidence of a statistically significant impact on OS in addition to the surrogate outcome, this meant for 247 (50%) submissions, the PBAC's decision was primarily based on the surrogate outcomes and thus met the eligibility criteria for further detailed narrative review. Figure 1 presents the PRISMA flow diagram of the relevant submissions of cancer drugs with a PSD identified for inclusion in the detailed narrative review. Where a submission was for multiple cancers, it was reviewed for each of the indicated cancer types.

Figure 1: PRISMA flow diagram
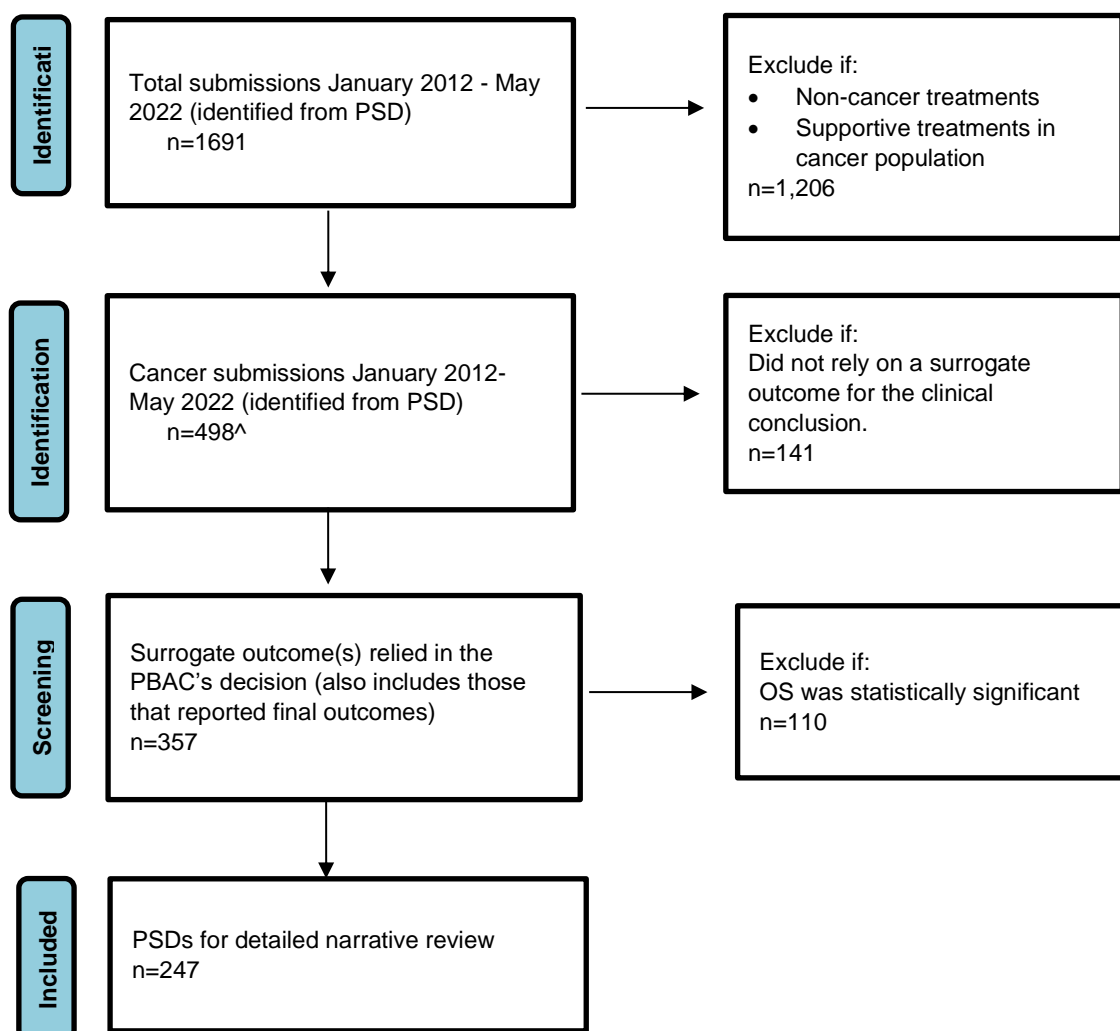
Source: complied during the review based on publication by Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. Doi: 10.1136/bmj.n71
OS=overall survival; PBAC=Pharmaceutical Benefits Advisory Committee; PSD=public summary documents; n=number
^ Total submission-indication pairs i.e. included 11 PSDs of submissions for two indications.

Characteristics of the included PBAC decisions

Within the 247 included PBAC decisions, the PBAC had considered 91 drugs for 22 broad cancer indications between January 2012 and May 2022. Since 2014, blood cancer was the predominant indication where surrogate outcomes were being heavily relied upon by the PBAC and included treatments for leukaemia (12%, 29/247 PSDs), lymphoma (14%, 35/247 PSDs) and myeloma (8%, 19/247 PSDs). The other most common cancer indications considered included lung cancer (12%, 30/247 PSDs), skin cancer (11%, 28/247 PSDs) and breast cancer (13%, 31/247 PSDs).

Figure 2 summarises the 247 PSDs considered by the PBAC for each year between 2012 and 2022 in the narrative review by their broad cancer types.

Figure 2: Submissions/resubmissions included (N=247) in the detailed surrogate review by broad cancer types and considered by the PBAC between 2012-2022.
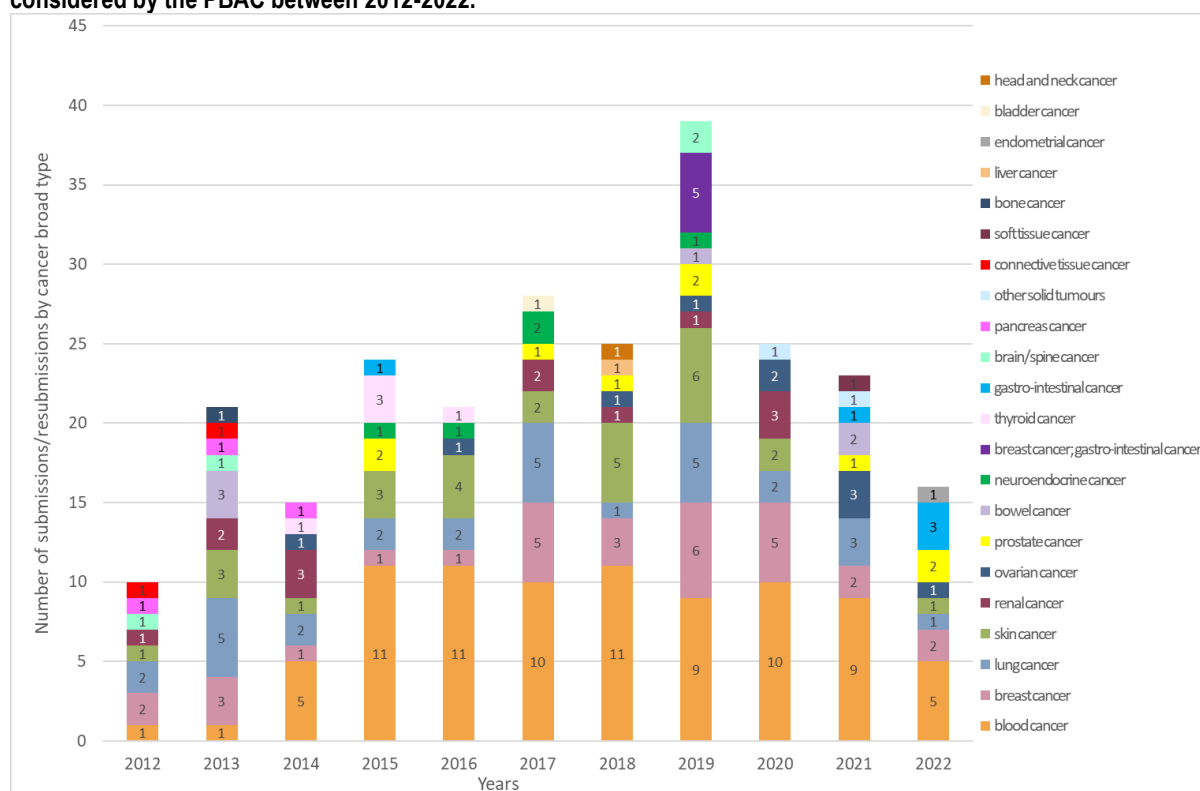
Source: Compiled during the review.

Table 2 summarises the relative use of surrogate outcomes by the most prevalent cancer indications and the PBAC recommendations for the included submissions/resubmissions.

Table 2: Frequency of surrogate outcomes as main clinical evidence in PBAC submissions/resubmissions by cancer type and the PBAC's recommendations for submissions/resubmissions that primarily relied on surrogate outcomes.

Cancer type	Total Submissions N	n (%)	Recommended, n (%)	Deferred, n (%)	Not recommended, n (%)	Not recommended due to immature OS^, n (%)
Lung cancer	66[a]	30 (45.5)	16 (53.3)	6 (20.0)	8 (26.7)	3 (37.5)
Blood cancer	152	83 (54.6)	34 (41.0)	10 (12.0)	39 (47.0)	23 (59.0)
Skin cancer	49[b]	28 (57.1)	14 (50.0)	3 (10.7)	11 (39.3)	6 (54.5)
Breast cancer	48[c]	31 (64.6)	12 (38.7)	2 (6.5)	17 (54.8)	12 (70.6)
Other cancers*	183	75 (41.0)	33 (44.0)	9 (12.0)	33 (44.0)	23 (69.7)
All cancers	**498**	**247 (49.6)**	**109 (44.1)**	**30 (12.1)**	**108 (43.7)**	**67 (62.0)**

N=number; OS=overall survival; PSD=Public Summary Document

Of the 247 submissions/resubmissions for cancer that relied on a surrogate, 44% (109) received a positive recommendation and 44% (108) were not recommended by the PBAC. The PBAC recommended listing with a risk sharing arrangement (RSA) on 13% (33/247) of occasions and with managed access on 2% (4/247) of occasions. Of the submissions/resubmissions not recommended, 62% (67) were not recommended due to inadequate or insufficient data on OS to support the clinical claim. Breast cancer had the highest proportion of submissions that relied on surrogate outcomes (65%) and of those submissions with surrogate outcomes were also the most likely to be not recommended (55%).

Table A1 (Attachment 1) summarises the overall characteristics of the cancer submissions/resubmissions (N=247) that relied on surrogate outcomes for the clinical claim or PBAC decision. These are summaries of data collected from PBAC decision documentation for the project based on the review objectives.

The most common oncology medicine class was targeted therapies at 79%. 45% (110/247) of PSDs were for resubmissions. Progression-free survival (PFS) was the most common surrogate outcome, used in 76% (187/247) of submissions/resubmissions, and measured as either a primary or secondary outcome in the clinical trial(s) supporting the submissions. Objective response rate (RR) and overall RR were used in 14% (35/247) and 9% (22/247) of submissions/resubmissions, respectively. These terms are often used interchangeably, however, we reported them separately because objective RR is more commonly used for evaluating treatment effect on solid tumours and overall RR is more commonly used in blood cancers. Additionally, objective RR was included in the FDA guidance document 'Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics Guidance for Industry,' whereas overall RR was not (U.S. Department of Health and Human Services Food and Drug Administration, 2018).

Table **3** summarises the main surrogate outcomes reported in the submissions/resubmissions and relied on by the PBAC for its decision making by cancer type.

Table 3: Main surrogate outcomes relied on in PBAC submissions/resubmissions by cancer type

Cancer type (number of submissions)	BCR	Biomarker	Clearance of solar keratosis	Complete Remission	Clinical/Complete Response	pCR	Cytogenic response	DCR	DFS	DMFS	DOR	DRR	iDFS	EFS	Incidence of cancer	MFS	MRD	Objective RR	Overall RR	PFS	POF	response	RFS	Symptom progression	TTNT	TTP	VGPR
Blood (83)				4	10		2				3			2	1		1	10	14	63			3		1	1	4
Breast (31)						1^			4			6	1	1					2	21	1					1	
Lung (30)								1			1							13		27							
Skin (28)			3							3	1							6	2	14			7				
Renal (13)																		1		13							
Ovarian (10)																				10							
Prostate (9)		3§														4				2				2	1	1	
Bowel (7)																		1		6							
Gastrointestinal (5)																		2		5							
Gastrointestinal + Breast~ (5)	1					3^					1								4	2						1	
Neuro-endocrine (5)																				5						1	
Thyroid (5)																				5							
Brain/spine (4)																		1		2		2				2	
Pancreatic (3)																				3							
Connective tissue (2)																				2							
Other solid tumours (2)																				2							
Bladder (1)																				1							
Bone (1)																								1			
Endometrial (1)																		1		1							
Head & neck (1)																				1							
Liver (1)																				1							
Soft tissue (1)																				1							

BCR=breast conservation rate; C$_{trough}$=trough concentration; DCR=disease control rate; DFS=disease free survival; DMFS=distant metastasis free survival; DOR=duration of response; DRR=durable response rate; EFS=event free survival; iDFS=invasive disease-free survival; MFS=metastasis-free survival; MRD=minimal residual disease; pCR=pathological complete response; POF=premature ovarian failure; RFS=relapse/recurrence free survival; RR=response rate; TTNT=time to next treatment; TTP=time to progression; VGPR=very good partial response; §biomarkers for prostate cancer=serum testosterone and prostate specific antigen levels; ^includes primary breast tumour (bpCR) and total response (tpCR); ~submissions for biosimilar products

Of the submissions/resubmissions that relied on surrogates, 65% (161) presented OS data based on interim trial results. In some instances, the pivotal trials supporting the submission have since published their final OS results. Table 4 presents the comparison of the published final OS results versus the interim OS trial results relied on in the submissions for PBAC decisions. The final OS results were generally consistent with the interim OS results. Attachment 4 presents further detail of the trials that now have published final OS results compared to the interim OS results relied on in the submissions.

Table 4: Comparison of the published final OS results versus interim OS results relied on in the submission

Cancer type	Trials^ with interim OS results at the time of submission	Statistical significance consistent[a] - HR consistent[d] (number of trials^)	Statistical significance consistent[a] - HR improved[d] (number of trials^)	Statistical significance consistent[a] - HR worsened (number of trials^)	Not consistent on statistical significance[a,b] (number of trials^)
Lung cancer	11	2	1	1	0
Blood cancer	38	8	0	1	3[c]
Skin cancer	9	5	0	0	1
Breast cancer	15	6	0	1	1
Other cancers*	28	8	0	1	2
All cancers	**101**	**29**	**1**	**4**	**7**

cf=compared to; HR=hazard ratio; OS=overall survival

^ the number of trials by drug-indication pairs (exclude resubmissions relying on the same trials).

* Includes broad cancer areas with less than 15 submissions/resubmissions, namely: bladder, bone, bowel, brain/spine, breast and gastrointestinal, connective tissue, endometrial, gastrointestinal, head and neck, liver, neuroendocrine, ovarian, pancreas, prostate, renal, soft tissue, solid tumours and thyroid.

a excluding trials for which the HR for OS were not reported in the PSD and trials for which this variable was not applicable (e.g. single arm studies used as the key clinical evidence)

b for our sample OS in final analysis reached statistical significance

c this included the CALGB trial (lenalidomide6) which had significant HR in interim analysis but had non significant HR in final analysis.

d HR consistent = HR within 0.1 difference compared to interim; HR improved = HR reduced >0.1 compared to interim; HR worsened = HR increased >0.1 compared to interim

Two thirds of the submissions/resubmissions (66%, 162/247) claimed superior clinical efficacy, and most presented a cost-utility analysis to the nominated main comparator (59%, 145/247). In submissions/resubmissions that presented modelled economic evaluations (i.e., a cost-effectiveness or cost-utility analysis), the surrogate outcome from the clinical evaluations were used in 85% (140/165) of cases. Therefore, the cost-effectiveness was likely heavily dependent on the assumed relationship between the surrogate outcome and OS or other clinically meaningful future outcomes.

Lung cancer

Table A2 and Table A3 (Attachment 1) summarise the characteristics of the lung cancer submissions/resubmissions (N=30) that relied on surrogate outcomes for the clinical claim or PBAC decision.

All 30 submissions were for targeted therapies and included 16 drugs or 21 drug-indication pairs (i.e., 5 drugs had submitted to the PBAC for multiple indications in lung cancer). 43% (13/30) of PSDs were for resubmissions. The PBAC had considered four surrogate measures in lung cancer during the study period, PFS (90%, 27/30), objective RR (43%, 13/30), duration of response (3%, 1/30) and disease control (3%, 1/30), measured as either a primary or secondary outcome in the clinical trials supporting the submission. Half of the PSDs (n=15) claimed superior clinical efficacy and presented a cost-utility analysis to the nominated main comparator. In these submissions, the surrogate outcome from the clinical evaluations was used in the modelled economic evaluations.

The PBAC considered the clinical effectiveness of the treatment (relative to comparator) to be clinically significant in 73% (11/15) of the submissions that claimed superiority. The main comparators across the pivotal trials supporting all the submissions/resubmissions included active comparators (57%, 17/30), placebo (17%, 5/30) or none (23%, 7/30) given it was a single arm study.

Of the 30 submissions, 53% (16/30) received a positive PBAC recommendation, with 27% (8/30) not recommended and 20% (6/30) deferred. The PBAC recommended listing with an RSA on 20% (6/30)

of occasions and with managed access on 3% (1/30) of occasions. Of the submissions/resubmissions with a positive PBAC recommendation, 44% (7/16) presented immature OS data (i.e. from an interim trial analysis). In all instances, the pivotal trials have since published their final OS results.

Of the 14 submissions/resubmissions that relied on a surrogate but were not recommended or were deferred by the PBAC, the main reason (64%) was related to both clinical and economic uncertainty. Further, in 37% (3/8) of submissions/resubmissions that were not recommended by the PBAC, the PBAC had documented in the PSDs that the OS data were immature (see Table 2).

Of all the submissions that relied on surrogate measures, the PBAC had documented comments about the surrogate and OS relationship in the PSDs in 30% (9/30) of the cases.

Narrative review of past PBAC decisions using surrogates in lung cancer
Progression-free survival (PFS)
PFS was the most commonly used surrogate in PBAC submissions for lung cancer treatments. The FDA defines PFS as 'the time from randomisation until objective tumour progression or death, whichever occurs first' (U.S. Department of Health and Human Services Food and Drug Administration, 2018). There are several factors which may introduce bias in PFS measurement in RCTs, such as the definition and assessment criteria for progression in the protocol, whether progression is objectively assessed, and protocols which allow for crossover to the active arm of the trial. Definitions of PFS may differ across RCTs, which may make evaluation of indirect comparisons between trials more challenging as the results could be biased in either direction. Examples of PBAC's documented views on PFS measurement in PSDs are included in Text Box 1.

Interpretation of disease progression may also differ between investigators and independent review committees (IRCs). Progression assessed by an IRC is preferred to investigator assessed progression as the risk of bias is lower. PBAC considers the potential for bias in PFS measurements when deliberating on the evidence presented in submissions. Additionally, trial protocols are not always clear on the processes for maintaining objectivity for IRCs, such as blinding. The PBAC have previously noted discrepancies between investigator and IRC assessed PFS in trials, particularly in single-arm or unblinded studies where the investigator is aware of patient assignment.

Text Box 1: Comments relating to measurement of PFS in PBAC decisions for lung cancer

> *'The PBAC agreed with ESC that there were significant limitations with the data presented in the indirect comparisons including… lack of trial exchangeability, due to varying eligibility criteria and PFS measurement….'.* (paragraph 7.2, afatinib, PSD, November 2015 PBAC meeting) [not recommended]
>
> *'PFS was the primary outcome in both ALTA-1L and ALEX, primarily assessed by the IRC in ALTA-1L and by investigators in ALEX, and conversely in either study. The base case PFS outcome used in the submission was from the IRC assessment from both trials, and that the results assessed by either the IRC or investigators within each of the trials were consistent, suggesting a low chance of bias between the methods.'* (paragraph 6.14, brigatinib, PSD, November 2019 PBAC meeting) [recommended]
>
> *'The PBAC considered the claim of superior treatment effect of osimertinib versus erlotinib/gefitinib in terms of PFS was reasonable based on the evidence presented from the FLAURA trial. ….. PBAC noted the estimated proportion of patients alive and progression-free (investigator-assessed, Response Evaluation Criteria In Solid Tumours (RECIST)-defined) was 50.9% in the osimertinib arm and 24.4% in the erlotinib/gefitinib arm at 18 months. The median PFS was 8.7 months longer for osimertinib (18.9 months) when compared with the median PFS of the comparator arm (10.2 months). First-line osimertinib was also associated with a statistically significant reduction in the risk of disease progression or death compared with erlotinib and gefitinib, with a hazard ratio (HR) of 0.46 (95% CI: 0.37, 0.57;*

In establishing clinical benefit for a new therapy, an advantage of PFS is that it is not confounded by the treatment effects of any post-progression therapies, unlike OS (Hashim et al., 2018). The potential for patients to benefit from subsequent treatments however may be relevant for a funding decision and a reason to act cautiously on the use of PFS as a surrogate, as the likely magnitude of benefit on OS for the intervention is likely to be much lower than that suggested by PFS differences (e.g., real-world benefits may be reduced costs from subsequent treatments rather than increased OS). There is also potential for selection bias using PFS, particularly in the timing and selection of patients who switch to the active treatment arm in crossover trials if protocols for crossover allow for investigator discretion, and this may bias PFS in favour of the active treatment. Given trials whose primary outcome is a surrogate outcome often include crossover to active treatment for those randomised to placebo, any future OS outcomes are likely impacted by treatment switching. Even when OS is adjusted for subsequent therapies, the treatment effect measurement may be uncertain (Fiteni et al., 2017). A recent systematic review analysed PFS and objective response rate in trials that allowed treatment switching and those that did not, and found that a significant treatment effect threshold for median PFS of 4.2 months predicted OS benefit (Hashim et al., 2018). These figures are slightly higher than those reported in a 2006 study (Johnson et al., 2006) referenced in the PSD for the crizotinib submission in November 2013, which noted extensive crossover from the comparator arm to crizotinib. Examples of the PBAC's documented views on data from crossover trials are included in Text Box 2.

Text Box 2: Comments relating to PFS and OS data from crossover trials in PBAC decisions for lung cancer

platinum chemotherapy in relation to PFS, the uncertainty in crossover adjustment of the OS data remained that key confounding factor. As such, the magnitude of the overall survival benefit of osimertinib treatment compared with chemotherapy remained uncertain.' (paragraph 7.7, osimertinib, PSD, July 2018 PBAC meeting) [deferred]

Objective Response Rate

Objective response rate (RR) was the second most frequently used surrogate endpoint in PBAC submissions for lung cancer therapies. Objective RR is a direct measure of the antitumor activity of the treatment which can be used in single arm studies (U.S. Department of Health and Human Services Food and Drug Administration, 2018). Definitions of objective RR vary, but it is usually the sum of partial responses and complete responses. The FDA guidelines recommend that response criteria should be clearly defined in the trial protocol, and be assessed using standardised criteria such as those defined in the Response Evaluation Criteria In Solid Tumours (RECIST) guidelines for measuring tumour size using radiographic imaging (Eisenhauer et al., 2009). One of the disadvantages of objective RR is that doesn't measure length of response like PFS does (Hashim et al., 2018). Examples of the PBAC's documented views on objective RR data in PSDs are summarised in Text Box 3.

Text Box 3: Comments relating to Objective RR in the PBAC's decisions for lung cancer

'The primary outcome of Study 1001 was ORR [Objective RR] by IRC. The discordance rate between IRC and investigator assessment for ORR for the pooled cohort EXP2:EXP5 was 17.8% and 25.9% for IC-ORR. Differences in response based on assessment method are observed in most of the response categories, which compromise the robustness of the data and were larger in the assessment of CNS metastases for lorlatinib treated patients… The PBAC was satisfied that intracranial activity of lorlatinib provides clinical benefit in patients with CNS metastases, based on the primary outcomes (ORR and IC-ORR) of Study 1001, despite the lack of comparative efficacy data for lorlatinib.' (paragraph 6.2 and 7.2, lorlatinib, PSD, November 2019 PBAC meeting) [recommended]

'The PBAC noted the ORR [Objective RR] (the primary outcome in the study) was 45%, median PFS was 8.9 months and median OS was 17.6 months for the combined liquid and tissue biopsy group. The PBAC noted the ORR was 46%, median PFS 11.0 months and median OS 20.4 months in the tissue biopsy group which together suggest the results from the overall sample may be conservative for the Australian treated population as testing will be based on tissue biopsy'. (paragraph 7.5, tepotinib, PSD, November 2021 PBAC meeting) [recommended]

Expediate literature review on the validity of surrogate endpoints used in clinical trials of therapies for lung cancer

Use of surrogate endpoints in clinical trials assists in obtaining efficacy data for new treatments faster, and in obtaining marketing approvals sooner, allowing patients to access effective therapies more quickly (Haslam et al., 2019). However, since OS is considered the most meaningful clinical endpoint, surrogate endpoints should be validated to show that they do predict the treatment effect on the patient-relevant clinical endpoint, OS. A summary of findings from recently published high quality literature on the validity of surrogate endpoints in lung cancer trials is presented in Table 5.

Table 5: Descriptive summary of literature on the validity of surrogate endpoints (Lung cancer)

Review article^ (year)	Type	# trials included	Main Surrogate	Summary of findings
Walia[10] (2022)	Review of FDA validation studies	15 (9 Lung)	Overall RR, PFS	No strong correlations in lung cancer
Haslam[9] (2019)	Systematic review of trial-level meta-analyses	78 (11 Lung)	Multiple	High correlation for PFS and OS and DFS and OS in adjuvant setting; Low-medium correlation in metastatic setting
Zhao[39] (2019)	Systematic review and meta-analysis	50	PFS and 1-year milestone survival	PFS is a possible surrogate in immunotherapy trials; 1-year milestone survival strongly correlated with OS in second-line NSCLC trials
Li[37] (2019)	Systematic review and meta-analysis of ICI RCTs in advanced NSCLC	9	CR, Objective RR, PFS	PFS was significantly correlated with OS; CR and Objective RR not correlated with OS
Hashim[35] (2018)	Systematic review and meta-analysis	146	Objective RR, PFS	Aimed to assess the impact of crossover and post progression treatments on surrogate/OS relationship. Calculated significant treatment effect required in trials for an expected significant OS benefit – Objective RR (41%) and median PFS (4.2 months)
Fiteni[40] (2017)	Systematic review and meta-analysis	20	PFS, DFS	PFS and DFS validated as surrogates for OS in operable NSCLC adjuvant trials and locally advanced NSCLC radiotherapy trials. All other correlations were low.
Ito[38] (2019)	Systematic review and meta-analysis of trials using ICI in people with high PD-L1 expression	7	Objective RR, PFS	Objective RR and HR(PFS) are strongly correlated with HR(OS) in ICI trials and could be predictors of survival in ICI NSCLC trials. Objective RR and PFS may be useful to predict OS in trials selecting participants with high PD-L1 expression.
Shameer[36] (2021)	Meta-analysis	81	PFS	Low-moderate correlations.
Belin[2] (2020)	Systematic review	91 (19 Lung)	PFS	PFS was validated in 4/19 included lung cancer studies, and partially validated in a further 1 study.

DFS=disease-free survival; FDA=The United States Food and Drug Administration; CR=complete response; HR=hazard ratio; ICI=immune checkpoint inhibitors; NSCLC=non-small cell lung cancer; OR=odds ratio; OS=overall survival; PD-L1=programmed death ligand 1; PFS=progression-free survival; RCT=randomised controlled trial; RR=response rate

^Literature review citations are listed in Attachment 2

Overall, the validation studies found that commonly used surrogates such as PFS and objective RR have not shown sufficient correlation with OS to be considered valid indicators of OS in most scenarios. In many submissions to PBAC for lung cancer treatments, clinical trials were ongoing and final OS data were not available. For many trials, particularly those in early stage disease, life expectancy can be many years, and the cohort is unlikely to reach median OS during follow-up. PFS and objective RR have shown sufficient correlation with OS to be considered valid indicators of OS in operable, locally advanced NSCLC (Fiteni et al., 2017). However, effective treatments in metastatic disease may also impact OS such that median OS may take years to reach. Examples of comments on the link between the surrogate endpoint and the clinical endpoint when OS data are immature are in Text Box 4.

Text Box 4: The PBAC's comments on validity of PFS in PBAC decisions for lung cancer

'The PBAC considered the claim of superior effectiveness in terms of improvement in PFS versus pemetrexed was supported. However, the PBAC did not consider PFS alone to be an adequately informative outcome for patients with advanced NSCLC.' (paragraph 9.2, crizotinib, PSD, November 2013 PBAC meeting) [deferred]

'The PBAC noted the improvement in progression free survival (PFS) associated with first-line treatment with osimertinib compared with first-line treatment with erlotinib or gefitinib. However, the magnitude of benefit in overall survival (OS) was uncertain, as the data provided were still immature.' (paragraph 7.1, osimertinib, PSD, July 2019 PBAC meeting) [not recommended]

'….published literature on NSCLC has not supported PFS as a surrogate for quality of life or OS, the two most important clinical outcomes for patients with advanced cancer.' (paragraph 6.15, afatinib, PSD, July 2015 PBAC meeting) [not recommended]

Decision making for lung cancer submissions when OS data are immature

Reporting of OS is often delayed in RCTs, particularly for cancers of long survivorship. Increasingly drug licensing decisions are based on shorter term outcomes such as PFS. This poses a major challenge for reimbursement agencies who must evaluate the treatments over the long term, including estimating patient survival. The PBAC have recommended listing of new treatments when OS benefit was uncertain due to immaturity of data. Comments in the PSDs on the reasoning behind the recommendations point to: clinical need, particularly for patient groups who have very few other treatment choices or where the new treatment offers a safer option; small patient populations; benefits in other patient-relevant outcomes such as quality of life; alignment with clinical practice guidelines; and pragmatic decisions based on the available evidence and cost-effectiveness against therapies with similar modes of action. Examples of stated reasons supporting the PBAC's decision to recommend the new treatments despite uncertain OS benefits are summarised in Text Box 5.

Text Box 5: The PBAC's stated reasons for recommending submissions with immature OS data in lung cancer

'The PBAC considered that, at the reduced price, afatinib could be considered to be cost-effective in comparison with platinum-based doublet chemotherapy, based on afatinib's superiority in terms of progression free survival and quality of life, and different toxicity profile, despite the evidence showing no additional overall survival benefit for first-line afatinib over chemotherapy in patients with NSCLC who are EGFR mutation positive.' (paragraph 12.23, afatinib, PSD, July 2013 PBAC meeting) [recommended]

'The PBAC considered that prioritising the access to TKIs to patients with NSCLC and activating EGFR mutations would position this class of agents in the treatment algorithm where they would deliver a net benefit to patients. The PBAC noted that providing access to TKIs in first-line for EGFR mutation positive NSCLC would also be consistent with clinical practice guidelines and the consensus of the EGFR/TKI stakeholder meeting.' (paragraph 12.22, afatinib, PSD, July 2013 PBAC meeting) [recommended] AND (paragraph 12.25, erlotinib, PSD, July 2013 PBAC meeting) [recommended]

'In making this recommendation, the PBAC noted the relatively small population of patients with ALK-positive NSCLC and the clinical need for additional targeted therapies with different safety profiles than currently available treatments for this condition.' (paragraph 7.2, alectinib, PSD, July 2017 PBAC meeting) [recommended]

'The treatment is expected to address a high and urgent unmet clinical need as there are currently no immunotherapies listed on the PBS for the first-line treatment of NSCLC patients with TPS<50%.' (paragraph 7.16, pembrilizumab, PSD, July 2019 PBAC meeting) [recommended]

'Given the small number of patients expected to be eligible for crizotinib and the clinical need for a more effective alternative than pemetrexed, and noting the difficult consequences for patients following the

Blood cancer

Table A4 and Table A5 (Attachment 1) summarise data collected from PBAC decision documentation (PSDs) on the characteristics of the blood cancer submissions/resubmissions (N=83) that relied on surrogate outcomes for the clinical claim or PBAC decision. Of these, 35% (29/83) were in the broad subcategory of leukemia, 42% (35/83) in lymphoma, and 23% (19/83) in myeloma. In total, there were 29 drugs or 50 drug-indication pairs (i.e., 14 drugs had submitted to the PBAC for multiple indications). The most commonly reported surrogate outcomes were PFS (76%, 63/83 PSDs) and overall response rate (17%, 14/83 PSDs), measured as either a primary or secondary outcome in the clinical trials supporting the submission.

Of the 83 submissions, 42% received a positive PBAC recommendation, with 47% (39/83) not recommended and 11% (9/38) deferred. Of all the submissions that relied on surrogate measures, the PBAC had documented views in PSDs on the surrogate and OS relationship in 22% (18/83) of the cases.

Leukemia

Of the 29 included submissions for leukemia, the majority were for targeted therapies (93%, 27/83 PSDs). 48% (14/29) of PSDs were for resubmissions. There were 72% (21/29) of submissions/resubmissions that included a claim of superior clinical efficacy and 67% (14/21) of these used the surrogate outcome in the model (1 cost-effectiveness and 13 cost-utility analyses).

The PBAC considered the clinical effectiveness of the treatment (relative to comparator) to be clinically significant in 90% (19/21) of the PSDs that claimed superiority. The main comparators across the pivotal trials supporting all the submissions/resubmissions included active comparators (52%, 15/29), placebo (14%, 4/29) or none (31%, 10/29) given it was a single arm study.

Of the 29 submissions for leukemia, 52% (15/29) received a positive PBAC recommendation, with 31% (9/29) not recommended and 17% (5/29) deferred. The PBAC recommended listing with an RSA on 7 (24%) occasions and with managed access on 1 occasion. Of the submissions/resubmissions with a positive PBAC recommendation, 47% (7/15) presented immature OS data (i.e. from an interim trial analysis).

Of the 14 submissions/resubmissions that relied on a surrogate but were not recommended or were deferred by the PBAC, the main reason (57%, 8/14) was related to both clinical and economic uncertainty. Further, in 89% (8/9) of submissions/resubmissions that were not recommended by the PBAC, the PBAC had documented in the PSDs that the OS data reviewed by the committee was immature.

Lymphoma

Of the 35 included submissions for lymphoma, the majority of submissions were for targeted therapies (74%, 26/35 PSDs). 43% (15/35) of PSDs were for resubmissions. There were 91% (32/35) of submissions/resubmissions that claimed superior clinical efficacy and 78% (25/32) of these used the surrogate outcome in the model (2 cost-effectiveness and 23 cost-utility analyses).

The PBAC considered the clinical effectiveness of the treatment (relative to comparator) to be clinically significant in 47% (15/32) of the PSDs that claimed superiority. The main comparators across the

pivotal trials supporting all the submissions/resubmissions included active comparators (43%, 15/35), placebo (6%, 2/35) or none (51%, 18/35) given it was a single arm study.

Of the 35 submissions, 40% (14/35) received a positive PBAC recommendation, with 49% (17/35) not recommended and 11% (4/35) deferred. The PBAC recommended listing with an RSA on 6 (17%) occasions and there were no recommendations for managed access. Of the submissions/resubmissions with a positive PBAC recommendation, 71% (10/14) presented immature OS data (i.e. from an interim trial analysis).

Of the 21 submissions/resubmissions that relied on a surrogate but were not recommended or were deferred by the PBAC, the main reason (76%, 16/21) was related to both clinical and economic uncertainty. Further, in 47% (8/17) of submissions/resubmissions that were not recommended by the PBAC, the PBAC had documented in the PSDs that the OS data reviewed by the committee was immature.

Myeloma

Of the 19 included submission for myeloma treatments, the majority were for targeted therapies (58%, N=11 PSDs). 42% (8/19) of PSDs were for resubmissions. There were 53% (10/19) of submissions/resubmissions that claimed superior clinical efficacy. Three of those made two clinical claims (non-inferior and superior) for different comparators or different treatment settings. Overall, 70% (7/10) of these used the surrogate outcome in the model (cost-utility analyses (9/10) and/or cost-effectiveness analyses (1/10)).

The PBAC considered the clinical effectiveness of the treatment (relative to comparator) to be clinically significant in 60% (6/10) of the PSDs that claimed superiority. The main comparators across the pivotal trials supporting all the submissions/resubmissions included active comparators (26%, 5/19) or placebo (79%, 15/19).

Of the 19 submissions, 32% (6/19) received a positive PBAC recommendation, with 68% (13/19) not recommended. The PBAC recommended listing with an RSA on 1 (5%) occasion and there were no recommendations for managed access. Of the submissions/resubmissions with a positive PBAC recommendation, one third (2/6) presented immature OS data.

Of the 13 submissions/resubmissions that relied on a surrogate but were not recommended by the PBAC, the main reason (46%, 6/13) was related to both clinical and economic uncertainty. Further, in 54% (7/13) of submissions/resubmissions that were not recommended by the PBAC, the PBAC had documented in the PSDs that the OS data reviewed by the committee was immature.

Narrative review of past PBAC decisions using surrogates in blood cancers
Leukemia

Improving survival is the most clinically meaningful and patient-relevant endpoint in treating acute myeloid leukemia (AML). The death rate from AML is highest in the first 3 years, after which the death rate plateaus and some patients with remissions lasting past 3 years may never relapse (Medeiros, 2018). Clinical trials investigating new induction therapies require follow-up of at least 3 years in order for sufficient number of OS events to evaluate the impact of the treatment on OS. Hence, surrogate endpoints are used to allow earlier assessment of therapies.

PFS was the most common surrogate used in the submissions for therapies to treat leukemia. Other surrogate endpoints relied on for evidence of efficacy included complete remission, event-free survival (EFS), transfusion independence, cytogenic response and objective RR, either alone or in conjunction with PFS. Complete remission has been considered an acceptable surrogate for patient-

relevant outcomes because clinical trials have shown that patients achieving complete remission gain a survival benefit, however, achieving complete remission does not always result in improved OS (Medeiros, 2018). Clinical trial protocols do not always use the standard definition of complete remission, making comparison of trial data more complicated. EFS is useful in AML clinical trials as it is not confounded by the impact of subsequent treatments, and can reduce the length of follow-up in clinical trials because events usually occur in the first year (Medeiros, 2018).

Minimal residual disease

Minimal residual disease (MRD) and bridge to transplant are emerging surrogates in AML clinical trials (Estey et al., 2016), however MRD testing methods are still developing. PBAC considered MRD in 1 submission for a leukemia treatment. Text Box 6 gives an example of how PBAC considered the MRD evidence.

Text Box 6: Comments on MRD as an emerging surrogate endpoint in the PBAC's decisions for leukemia

'The PBAC noted [that] the key study (BLAST) recruited patients with an MRD level of ≥ 10^{-3}, while the threshold proposed in the restriction was ≥ 10^{-4}. The PBAC noted that a threshold of ≥ 10^{-4} is currently used in clinical practice to define MRD, but noted this was based on a consensus of clinical opinion rather than a strong evidence base. Notwithstanding this, the PBAC considered that an MRD level of ≥ 10^{-4} was appropriate for inclusion in the blinatumomab restriction. The PBAC also considered that, as testing technology improves over time, the level of residual disease detected would decrease which may further lessen the applicability of the study data. Thus, the PBAC considered that the level of MRD should be explicitly stated in the PBS restriction.' (paragraph 7.7, blinatumomab, PSD, July 2018 PBAC meeting) [not recommended]

Clinical trial data

Small sample sizes and the typically older age group of AML patients means that trials may not have sufficient power for statistically significant results. Crossover to subsequent treatments is also a confounding factor for survival data in leukemia clinical trials. The PBAC's documented views on submissions where the evidence presented is not statistically significant are summarised in Text Box 7.

Text Box 7: Comments relating to patient population, differences in trial sample sizes, eligibility criteria, study design and duration of follow-up in the PBAC's decisions for leukemia

'The PBAC noted that the trial data demonstrated no additional benefit in terms of OS for venetoclax + obinutuzumab compared to chlorambucil + obinutuzumab. The PBAC considered that OS was not an informative outcome in CLL due to downstream effective treatments and the older mean age at diagnosis (70 years).' (paragraph 7.7, venetoclax, PSD, March 2020 PBAC meeting) [not recommended]

'The results from the MAIC demonstrated that acalabrutinib and ibrutinib were not statistically different in terms of progression free survival (HR = 0.72; 95% CI: 0.33, 1.60) or overall survival (HR = 0.92; 95% CI: 0.38, 2.27). The PBAC considered that the results of the MAICs were uncertain due to the low effective sample size in the acalabrutinib arm after matching (N=44, reduced from 132), differences in the duration of follow-up between the trials and differences between the trials in eligibility criteria.' (paragraph 7.7, acalabrutinib, PSD, March 2020 PBAC meeting) [recommended]

'The PBAC recalled that it had previously considered the key trial AFLA-0701, and various post hoc subgroup analyses by cytogenetic risk, and although it had considered the clinical claim of superior effectiveness was reasonable in terms of event-free survival, the claim was not adequately supported in terms of overall survival, and there were also applicability concerns (para 7.9 and 6.52, gemtuzumab, PSD, March 2021 PBAC meeting). The PBAC noted new clinical evidence presented …. The PBAC considered that the new evidence alone did not address its uncertainties with respect to the clinical

claim. However, it also recognised that no further clinical data was expected and that the cure claim itself was not implausible, although the magnitude of the benefit had been poorly supported.' (paragraph 7.6, gemtuzumab, PSD, November 2021 PBAC meeting) [recommended]

'A range of other factors have been identified by the PBAC as influential in its review of statistical adjustment methods in the context of cross-over, including: the degree of cross-over ….; whether there is a sufficient number of patients to inform the adjustment methods….; the extent to which cross-over was triggered by progression events, and if so, how these were assessed using evaluation of symptomatic or asymptomatic events; … and whether there is any corroborating evidence provided to support the use of progression-free survival as a surrogate measure for overall survival in this specific condition.' (paragraph 6.11, ibrutinib, PSD, November 2015 PBAC meeting) [deferred]

Decision making for leukemia submissions when OS data are immature

Submissions continue to present evidence based on immature survival data. PBAC has recommended submissions based on interim data where there is a high clinical need, and where there is sufficient evidence of efficacy based on the surrogate measure. PBAC has also commented on reviewing updates for OS when new data is released. Examples of PBAC comments regarding immature OS data are in Text Box 8.

Text Box 8: Comments relating to immature OS data and data updates in the PBAC's decisions for leukemia

'The PBAC recalled that, based on ELEVATE-TN, it had previously considered that a claim of superior comparative effectiveness between acalabrutinib and chlorambucil + obinutuzumab was reasonable, although the magnitude of the benefit had been uncertain at that time due to the immaturity of the data. The PBAC was more confident this claim was reasonable in relation to PFS, noting that the more mature data in the resubmission with follow-up over approximately four years showed consistency of effect. At the same time, it remained of the view that it would be inappropriate to model an overall survival gain in the economic evaluation as this data remained immature and there may not be a difference over the longer term given subsequent lines of effective therapy are available.' (paragraph 7.8, acalabrutinib, PSD, November 2021 PBAC meeting) [not recommended]

'The preliminary results from the TOWER trial presented in the PSCR provide reassurances that evidence that is more robust will be forthcoming in the foreseeable future to better inform the clinical effectiveness and cost effectiveness. The PBAC has proposed a plan to review this evidence as soon as it becomes available to ensure patients receiving medicines on the PBS are being treated according to the best available evidence and that the cost of the treatment remains justified in terms of acceptable cost-effectiveness. Should the modelled extent of benefits not be realised, the Committee has recommended measures to minimise the risk of unjustified health care expenditure.' (paragraph 7.6, blinatumomab, July 2016 PBAC meeting) [recommended]

'The PBAC recalled it previously considered that while blinatumomab is effective in eliminating MRD and is associated with durable relapse-free survival, it remained unclear whether blinatumomab would lead to long-term gains in overall survival given the lack [of] reliable comparative data and the relative immaturity of the data from the BLAST study. However, the PBAC considered the updated data from the BLAST study (see paragraph 4.5) reinforced that treatment with blinatumomab may be associated with an overall survival advantage noting that the plateau in overall survival after 48 months indicated in the overall survival data available at the March 2019 consideration was maintained.' (paragraph 5.3, blinatumomab, PDS, July 2019 PBAC meeting) [recommended]

Stem cell transplant in AML

Achieving stem cell transplant (STC) is a relevant outcome for induction therapies. However, SCT may confound long term follow-up data and make it difficult to determine the effect of the induction therapy. Achieving SCT does not necessarily improve OS. There have been conflicting views expressed in PSDs relating to which endpoint is most relevant to consider for HTA decisions. Examples of comments made by PBAC on SCT as an outcome, and as a confounder for OS are in Text Box 9.

Text Box 9: Comments relating to SCT from the PBAC's decisions in leukemia

'The PBAC considered that the preliminary TOWER trial results provided by the sponsor in the PSCR indicated the likely superiority in efficacy of blinatumomab over standard care chemotherapy. However, the PBAC disagreed with the ESC that the most clinically relevant outcome is transplant rate and post-transplant survival, instead considering that overall survival was the most relevant outcome to inform decision-making.' (paragraph 7.7, blinatumomab, PSD, July 2016 PBAC meeting) [recommended]

'The ESC noted that inotuzumab was associated with a higher rate of HSCT than blinatumomab and considered that this was a clinically meaningful outcome that likely contributed to the overall survival and event free survival outcomes achieved with inotuzumab.' (paragraph 6.43, inotuzumab, PSD, November 2018 PBAC meeting) [recommended]

'The PBAC had also previously noted that overall survival was potentially confounded as 22.1% of patients in the control arm subsequently received gemtuzumab, and due to the use of salvage therapies and HSCT. The PBAC had further noted that more patients in the control arm (39.0%) underwent HSCT compared to in the gemtuzumab arm (23.7%) (para 7.8, gemtuzumab PSD, March 2021 PBAC meeting).' (paragraph 6.26, gemtuzumab, PSD, November 2021 PBAC meeting) [recommended]

Validity of surrogate endpoints used in clinical trials of therapies for leukemia

The rapid literature review for recent evidence identified 3 articles on the correlation between surrogate endpoints and OS in leukemia. These are summarised in Table 6.

Table 6: Summary of validation studies for surrogate endpoints used in leukemia clinical trials

Review article^ (year)	Type	# trials included	Main Surrogate	Findings
Assouline[3] (2022)	Systematic review	31 (13 AML)	EFS	EFS is highly correlated with OS at trial-level in newly diagnosed AML.
Walia[10] (2022)	Systematic review of FDA approvals	15 (1 AML)	Multiple	In AML, high correlation between EFS and OS.
Belin[2] (2020)	Methodological systematic review	91 (1 CLL)	PFS	Possible surrogate.

AML=acute myeloid leukemia; EFS=event free survival; OS=overall survival; PFS=progression-free survival
^Literature review citations are listed in Attachment 2

Lymphoma

Many lymphomas can be cured, and the indolent subtypes of Non-Hodgkin Lymphoma (NHL) are often considered chronic diseases. Being diagnosed with indolent NHL does not necessarily reduce life expectancy (Non-Hodgkin Lymphoma (NHL) - Lymphoma Australia). Because modern treatments are improving patient survival, clinical trials are taking longer to reach median PFS, and lengthy follow-up times are necessary for median OS results. Surrogate endpoints based on response, which can be measured at earlier time points are becoming more widely adopted in clinical trials (Mangal, Salem, Li, et al., 2018). However, the most common surrogate endpoint relied on in submissions for therapies to treat lymphoma was PFS, followed by objective RR and overall RR. Complete response and very good partial response (VGPR) were also used. The PBAC had noted the value consumers place on being progression-free. Examples of comments referring to impact of therapies on quality of life are in Text Box 10.

Text Box 10: Comments relating to quality of life from the PBAC's decisions in lymphoma

'Representatives of the PBAC met with Lymphoma Australia prior to the PBAC meeting, and reported the following key points to the PBAC in relation to the agenda items for CLL and indolent NHL:
- Consumers place high importance on having access to the best available treatments. Where cure is not possible, the eventual goal would be to enable indolent lymphomas and CLL to be treated as chronic diseases. Ultimately patients may die of conditions unrelated to their lymphoma.
- Patients may relapse multiple times in the course of the disease, and will be treated on relapse. As PBS subsidy may influence the choice of treatment, subsidising the most clinically effective treatments is critical to ensure the best value for the taxpayer.
- Patients may be diagnosed at a young age and live for years after diagnosis, and therefore place a high value on PFS. Patients who are well during the progression free period can resume day-to day functions including participating in the workforce and family life. In this context, the decision for the patient rests on a balance of the PFS gained against the quality of life impacts of drug toxicity. The psychological impact of patients' fear of relapse can have a highly detrimental effect on their quality of life.
- With regard to bendamustine, Lymphoma Australia noted that bendamustine has been available overseas for many years, but noted that this was the first application for PBAC consideration of the drug in Australia.' (paragraph 6.3, bendamustine, PSD, March 2015 PBAC meeting) [deferred]

'The PBAC noted and welcomed this input. PBAC recognises that a drug may be useful even when it does not provide a survival advantage, but does provide quality of life benefits. In terms of using PFS to value the benefits of a drug, PBAC recalled that some of the most informative submissions seen to date have presented economic models that incorporate the impacts on quality of life when patients are in a PFS state, capturing the fact that PFS is not a homogenous state. It was noted that exploring how patients could provide more input to rigorous measurement of Quality of Life would be valuable in future consumer submissions.' (paragraph 6.4, bendamustine, PSD, March 2015 PBAC meeting) [deferred]

'The PBAC considered that the restriction should not limit use to patient's whose treatment has a curative intent, noting that this would be ambiguous and would preclude use in some patients who would derive improvements in quality of life.' (paragraph 7.6, brentuximab, PSD, March 2015 PBAC meeting) [not recommended]

'The PBAC noted that the clinical claim for obinutuzumab over rituximab was made based on superior investigator-assessed PFS. The PBAC reiterated that it considered PFS to be a potentially important outcome in indolent diseases such as follicular lymphoma. However, for this submission the PBAC was concerned that the modest gain in PFS over rituximab may be offset by increases in serious AEs. In addition, the PBAC noted that no difference in OS or health-related quality of life measures was demonstrated between treatment arms. Hence, the PBAC considered the clinical claim of superior effectiveness to be inadequately supported.' (paragraph 7.6, obinutuzumab, PSD, November 2017 PBAC meeting) [not recommended]

Even though relying on surrogate endpoints as evidence of efficacy is common in submissions, not all surrogate endpoints have been independently validated to show a strong correlation with the clinical endpoint. PBAC guidelines set out requirements for submissions using surrogate endpoints, however, these are rarely referred to in PSDs. Even so, PSDs do show PBAC concerns regarding the validity of the surrogate endpoints in some indications. Text Box 11 shows examples where the PBAC had raised surrogate validity concerns.

Text Box 11: Comments relating to validity of surrogate endpoints from the PBAC's decisions in lymphoma

'The submission noted that PFS and OS endpoints from ASPEN were not mature at the time of submission, thus disease response rates were presented as meaningful surrogates of effectiveness. However, as the submission did not address the requirements outlined in the PBAC Guidelines for translating comparative treatment effects of proposed surrogate measures to target clinical outcomes, this claim was unsupported (Appendix 5, PBAC Guidelines v5.0, 2016).' (paragraph 6.16, zanubrutinib, PSD, July 2021 PBAC meeting) [not recommended]

'The PBAC considered that PFS was difficult to interpret and compare across the trials due to differences in the patient populations and outcome measurement. The PBAC also considered that time to next treatment (TTNT), adjusted for the time therapy ceased, may be a more appropriate outcome for considering the duration of treatment response.' (paragraph 6.14, brentuximab, PSD, July 2018 PBAC meeting) [not recommended]

'Not only does this implicitly assume that PFS is a valid surrogate for OS for patients with R/R PMBCL receiving immunotherapy, but it also assumes that the same relationship between PFS and OS holds for both immunotherapy and chemotherapy. There is increasing evidence to suggest that PFS is not a reliable surrogate for OS in many oncology settings, especially for immunotherapies[7,8,9]. Therefore, the validity of the PFS data generated using this approach was highly uncertain.' (paragraph 6.49. pembrolizumab, PSD, March 2020 PBAC meeting) [recommended]

[7] Buyse M, Burzykowski T, Saad ED. The search for surrogate endpoints for immunotherapy trials. Annals of translational medicine. 2018; 6(11):231.

[8] Haslam A, Hey SP, Gill J, Prasad V. A systematic review of trial-level meta-analyses measuring the strength of association between surrogate end-points and overall survival in oncology. Eur J Cancer. 2019 Jan; 106:196-211.

[9] Kumar S, Rajkumar SV. Surrogate endpoints in randomised controlled trials: a reality check. Lancet. 2019 Jul 27; 394(10195):281-3.

'For the economic evaluation, a surrogate relationship was assumed to apply between response in CTCL and the median durations of PFS and OS from the registry. This link between the surrogate outcome (response) and the final outcome (survival) was not justified in the resubmission. The application of this link from response to survival may not be appropriate in CTCL since, unlike in other NHL subtypes, response criteria for Mycosis Fungoides / Sézary Syndrome(MF/SS) have not been demonstrated to correlate with prognosis for survival.' (paragraph 6.26, vorinostat, PSD, November 2016 PBAC meeting) [deferred]

'The PBAC noted that limited data were available for overall survival and progression free survival, and that the majority of the data for efficacy was for the assessment of clinical response and was mostly in the salvage patient group. The PBAC considered that data on survival and/or potential quality of life impacts of the palliative group of patients would have been useful in estimating the benefit of brentuximab vedotin in this population.' (paragraph 7.10, brentuximab, PSD, November 2016 PBAC meeting) [recommended]

'The PBAC agreed with the ESC that the superiority claim was likely reasonable in terms of response outcomes, noting its advice that the magnitude of benefit was uncertain, and that response outcomes are unlikely to translate into survival outcomes. Nonetheless, the PBAC was certain of a response benefit, and given the substantial value of this outcome to patients in terms of improving quality of life and given lack of treatments available on the PBS, considered that zanubrutinib provided a high added therapeutic value in the treatment of WM.' (paragraph 7.6, zanubrutinib, PSD, July 2020 PBAC meeting) [not recommended]

Decision making for lymphoma submissions when OS data are immature

Because long-term follow-up is necessary to reach median PFS and OS, PBAC had commented on the maturity and quality of evidence presented for decision making. In rare cancers, clinical trial sample sizes can be small and not powered to show statistically significant differences in treatment effect. Examples of relevant PBAC comments are in Text Box 12.

Text Box 12: Comments relating to data maturity and quality of evidence from the PBAC's decisions in lymphoma

> *'The committee noted that no overall survival benefit was demonstrated in the StiL trial, but that median overall survival was not reached in either group at the time of analysis (median follow up of 45 months).'* (paragraph 7.9, bendamustine, PSD, March 2015 PBAC meeting) [deferred]
>
> *'In reviewing the resubmission, the ESC considered that any difference in OS between treatments was unlikely to be observed in the clinical trial setting given patients with WM survive for a relatively long time due to the indolent nature of WM and given the rarity of WM. The ESC acknowledged that additional long-term trial data was unlikely to be forthcoming for WM.'* (paragraph 6.16, zanubrutinib, PSD, March 2022 PBAC meeting) [recommended]
>
> *'The PBAC noted that the end of follow-up results for PFS and OS from the BRIGHT trial were anticipated to be reported in July 2017 and reiterated that it would wish to see and review these data when released.'* (paragraph 7.3, bendamustine, PSD, July 2015 PBAC meeting) [recommended]
>
> *'The PBAC considered that the data presented in the submission were of poor quality with a high risk of bias. The submission was based on a naïve comparison of datasets with small sample sizes and sparse event data.'* (paragraph 7.8, brentuximab, PSD, March 2015 PBAC meeting) [not recommended]

Use of surrogates and OS projections to inform the economic model

Economic models routinely use surrogate endpoints such as PFS to project long-term outcomes. In diseases where life expectancy is long, the use of surrogate endpoint data collected at early time points may add uncertainty to economic models. Examples of comments relating to surrogate endpoints in economic models are in Text Box 13.

Text Box 13: Comments relating to surrogate endpoints informing economic models from the PBAC's decisions in lymphoma

> *'For the economic evaluation, a surrogate relationship was assumed to apply between response in PTCL and the median durations of PFS and OS from the registry. This translation from a surrogate outcome (response) to a final outcome (survival) was not sufficiently substantiated in the submission.'* (paragraph 6.20, romidepsin, PSD, November 2016 PBAC meeting) [not recommended]
>
> *'…the significant uncertainty created as a result of extrapolating PFS from immature data, which likely biased the estimates in favour of obinutuzumab. The PBAC considered that the fitted parametric curve for the obinutuzumab plus bendamustine arm was unlikely to reliably predict PFS post cessation of obinutuzumab maintenance and in the trial approximately 18% of patients in the obinutuzumab plus bendamustine arm were still receiving obinutuzumab maintenance at the time of the most recent data analysis. The PBAC considered that it would have been more appropriate to allow for earlier convergence of the PFS curves, consistent with the waning of the obinutuzumab maintenance effect; the submission used a 15-year time horizon, however, the PBAC considered that a 10-year time horizon was more appropriate given the relative poor prognosis of patients with follicular lymphoma that is [refractory] to rituximab….'* (paragraph 7.7, obinutuzumab, PSD, November 2016 PBAC meeting) [not recommended]
>
> *'Overall, the PBAC considered that the underlying lack of evidence to support the claim of superior efficacy resulted in unreliable inputs and optimistic extrapolations, leading to an economic model that did not provide a plausible estimate of the cost effectiveness of pralatrexate.'* (paragraph 7.9, pralatrexate, PSD, July 2017 PBAC meeting) [not recommended]

Validity of surrogate endpoints used in clinical trials of therapies for lymphoma

A summary of the recent literature on surrogate endpoints used in lymphoma clinical trials is in Table 7.

Table 7: Summary of validation studies for surrogate endpoints used in lymphoma clinical trials

Review article^ (year)	Type	# trials included	Main Surrogate	Findings
Assouline[3] (2022)	Systematic review	31 (18 Lymphoma)	EFS	EFS is strong surrogate for OS in DLBCL treated with immunochemotherapy at patient- and trial-level.
Haslam[9] (2019)	Systematic Review	78 (3 lymphoma)	pCR, EFS, PFS,	High correlation between PFS and 2-year OS in DLBCL. Low correlations with pCr and medium correlations with EFS/PFS in NHL.
Belin[2] (2020)	Methodological systematic review	91 (3 lymphoma)	PFS	Validated in first-line DLBCL and NHL. Not validated in FL post-autologous transplant.
Mangal[5] (2018)	Database review	73	CR, Overall RR as predictors of PFS	ORR and CR correlated with mPFS in NHL, including FL.

CR=complete response; DLBCL=diffuse large B-cell lymphoma; EFS=event free survival; FL=follicular lymphoma; IMiD=immunomodulatory imide drug; mPFS=median progression-free survival; MRD=minimal residual disease; NHL=non-Hodgkin lymphoma; OS=overall survival; pCR=partial complete response; PFS=progression-free survival; RR=response rate
^Literature review citations are listed in Attachment 2

Multiple Myeloma

Multiple Myeloma (MM) is a cancer of plasma cells characterised by remissions and relapses. The 5-year survival rate for MM is 55% (American Cancer Society, Survival Rates for Multiple Myeloma). Median PFS of 3-4 years have been reported in clinical trials, and it is not uncommon for patients to live with MM for 15-20 years (Avet-Loiseau et al., 2020).

Newly Diagnosed Multiple Myeloma

In newly diagnosed multiple myeloma (NDMM), median OS is often not reached by the end of trial follow-up, and subsequent therapies may also improve OS, confounding OS data during long-term follow-up (Daniele et al., 2022). Because it is not practical to wait many years to reach median OS in this setting, it is common for PFS to be the primary endpoint of clinical trials, with response measures being used as surrogates for PFS rather than OS (Mangal, Salem, Menon, et al., 2018). The submissions for NDMM during the study period included a total of 3 PSDs for 2 drugs. The submission for bortezomib relied on Very Good Partial Response (VGPR) for its clinical claim for induction therapy in NDMM. The submissions for lenalidomide were for maintenance therapy and relied on PFS for their clinical claim.

Very Good Partial Response (VGPR)

In 2006, the International Myeloma Working Group (IMWG) met to reform practice and proposed 'uniform response criteria' to help make comparisons between MM treatments more precise (Durie et al., 2006). One outcome of this meeting was to update the definition for the response outcome VGPR, which identifies patients whose response to treatment is such that they might experience outcomes similar to those in complete response. The IMWG defined VGPR as 'serum and urine M-component detectable by immunofixation but not on electrophoresis or 90% or greater reduction in serum M-component plus urine M-component <100 mg per 24 h'. The latter definition is less susceptible to observer bias and the IMWG claim it is easier to use.

Text Box 14 includes comments from PSDs on surrogate endpoints in the NDMM setting.

Text Box 14: Comments on surrogate endpoint use in NDMM in the PBAC's decisions

'The PBAC noted that the indirect comparison did not show any statistically significant differences in very good partial response (VGPR) rates for either the VcD …. versus TD …. or VcAD …. versus TAD …. comparisons, indirect OR (95%CI): 0.91 (0.36, 2.26) and 1.59 (0.91, 2.78), respectively). However, the PBAC noted that VGPR has been previously accepted as a measure of effectiveness of induction chemotherapy in the setting of stem cell transplants (SCTs) when the submission for thalidomide for newly diagnosed multiple myeloma patients was considered in 2009. From the previous thalidomide submission, the PBAC noted that a post-hoc analysis of eight year survival data conducted as part of the Barlogie trial (Barlogie et al (2008)) demonstrated a possible improvement in overall survival. The overall eight year survival estimates were 56% for the thalidomide group compared with 45% in the control group (p=0.09).
The PBAC noted that the current bortezomib submission did not provide supportive data regarding long term benefit. However, based on the indirect comparison of VGPR and the fact that PBAC has regarded bortezomib as equivalent to thalidomide for other multiple myeloma indications, the PBAC considered that bortezomib is likely to be non-inferior to thalidomide. The PBAC acknowledged it is very difficult to isolate the impact of one drug in an induction regimen given that multiple myeloma is treated with multiple other drugs over a relatively long period.' (paragraph 12.2, bortezomib, PSD, March 2012 PBAC meeting) [recommended]

'The PBAC considered that it was implausible that lenalidomide would increase OS compared with thalidomide, but have no impact on PFS, especially in the context of maintenance therapy where the aim of treatment is to extend the period of time in the progression-free state.' (paragraph 7.10, lenalidomide, PSD, March 2018 PBAC meeting) [not recommended]

Relapsed or Refractory Multiple Myeloma

Most patients with MM will relapse (Nathwani et al., 2022). Newer treatments such as lenalidomide and bortezomib have improved survival times for patients with relapsed or refractory multiple myeloma (RRMM) to 2.5 years from relapse (Kumar et al., 2008). In submissions for therapies to treat RRMM, PFS was the most common surrogate endpoint. Response outcomes were also presented, mostly with PFS. Even though the survival outlook in RRMM is not as distant as in NDMM, submissions are presenting clinical trials with immature OS data. Text Box 15 shows examples of comments from PSDs relating to the link between PFS and OS in RRMM.

Text Box 15: Comments relating to PFS and OS in RRMM in the PBAC's decisions

'…Further, it was unclear whether PFS was a good surrogate for OS in multiple myeloma, particularly given it is a relapsing condition.' (paragraph 6.10, carfilzomib, PSD, November 2016 PBAC meeting) [not recommended]

'The PBAC considered that the economic analysis provided by the resubmission was favourable to DBd treatment and that the overall survival gains with DBd treatment were large in comparison to that observed in the trial and hence uncertain. The PBAC noted that the treatment landscape in multiple myeloma is rapidly changing and expected survival rates for patients have improved. However, the PBAC noted that, using the second-line subgroup population, """"% ("""""%) of DBd patients were modelled to be alive at 20 years (15 years) and considered that this was clinically implausible. The PBAC also considered that the persistent treatment effect (i.e. the curves did not converge) over the 20 year (15 year) time horizon was not adequately supported by the data.' (paragraph 7.11, daratumumab, PSD, march 2019 PBAC meeting) [not recommended]

'The PBAC noted that the hazard ratio for overall survival (OS) was statistically significant for Cd versus Ld from ENDEAVOR (HR = 0.76; 95% CI: 0.63, 0.92), but was not statistically significant for ILd versus Ld from TOURMALINE (HR = 0.87; 95% CI: 0.64, 1.18). However, the PBAC considered that the data from the TOURMALINE trial were too immature to assess efficacy in terms of OS.' (paragraph 7.8, ixazomib, PSD, November 2020 PBAC meeting) [not recommended]

> *'The PBAC noted that although PBd demonstrated an improvement compared to Bd in terms of progression free survival (HR = 0.61; 95% CI: 0.49, 0.77) in the OPTIMISMM trial, PBd provided no statistically significant improvement in terms of overall survival (HR = 0.91; 95% CI: 0.70, 1.18).'* (paragraph 7.6, pomalidomide, PSD, July 2019 PBAC meeting) [not recommended]
>
> *'In regard to OS, the PBAC noted that despite the longer follow-up, the survival data presented in the resubmission remained immature with an event rate of 35% and 39% for SBd and Bd arms, respectively. The difference in OS based on the updated data cut-off was not statistically significant between the two trial arms (HR = 0.88; 95% CI: 0.63, 1.22), a result that was consistent with the February 2020 data cut-off as presented in the July 2021 submission (HR = 0.84; 95% CI: 0.57, 1.23). The PBAC considered although this result may be in part due to the impact of crossover within the Bd treatment arm to SBd treatment, the impact of SBd on OS remained uncertain. In contrast, the PBAC noted that for Cd, the clinical evidence (ENDEAVOR) demonstrated a significant improvement in OS for Cd compared with Bd (HR = 0.76; 95% CI: 0.63, 0.92).'* (paragraph 7.9, selinexor, PSD, March 2022 PBAC meeting) [not recommended]
>
> *'The PBAC also noted that although ASPIRE had 20 months more follow-up than ENDEAVOR, the data from both trials were immature'.* (paragraph 7.5, carfilmozib, PSD, November 2016 PBAC meeting) [not recommended]

Validity of surrogate outcomes used in clinical trials of therapies for MM

The literature review identified 4 articles validating response outcomes against PFS as the clinical endpoint (Avet-Loiseau et al., 2020; Daniele et al., 2022; Mangal, Salem, Menon, et al., 2018; Munshi et al., 2020). In some submissions, clinical trial data is being presented where median PFS has not yet been reached in the active treatment arm, and measures of response to treatment are being relied upon for the clinical claim. VGPR is a surrogate measure with a clear definition which has been validated to predict median PFS in MM (Mangal, Salem, Menon, et al., 2018). MRD has been identified as a possible surrogate for PFS (Avet-Loiseau et al., 2020; Daniele et al., 2022) and OS (Munshi et al., 2020), however none of the submissions relied on MRD for their clinical claim.

Results of meta-analyses validating PFS as a surrogate for OS were mixed, ranging from no correlation, to medium and high correlation (Belin et al., 2020; Haslam et al., 2019). Belin et al. validated PFS as a surrogate for OS in a meta-analysis of chemotherapy and targeted therapy trials in MM, however, Haslam et al. used a different validation method on the same data and found only medium correlation. Table 8 summarises the validation studies for surrogate endpoints used in MM clinical trials.

Table 8: Summary of validation studies for surrogate outcomes used in multiple myeloma clinical trials

Review article^ (year)	Type	# trials included	Main Surrogate	Findings
Avet-Loiseau[8] (2020)	Systematic review and meta-analysis	6	MRD	MRD negativity possible surrogate for PFS (Prentice criteria) in NDMM
Daniele[4] (2022)	Systematic review	75	CR, MRD, Overall RR	Overall RR and CR significantly correlated with PFS MRD and sCR (stringent CR) possible surrogates in NDMM
Mangal[6] (2018)	Database analyses	102	CB, CR, DC, Overall RR, VGPR	VGPR was superior to CB, Overall RR or CR in predicting median PFS. mPFS was longer with IMiDs in RRMM trials
Belin[3] (2020)	Methodological Systematic review	91 (2 MM)	PFS	PFS was validated as a surrogate for OS in 1 of 2 included studies
Haslam[9] (2019)	Systematic review and meta-analysis	78 (1 MM)	PFS	PFS(HR) and PFS (logHR) had medium correlation with OS
Munshi[11] (2020)	Systematic review and Meta-analysis	93	MRD	MRD negativity has potential to predict PFS and OS

CB=clinical benefit; CR=complete response; DC=disease control; IMiD=immunomodulatory imide drug; mPFS=median progression-free survival; MM=multiple myeloma; MRD=minimal residual disease; NDMM=newly diagnosed multiple myeloma; OS=overall survival; PFS=progression-free survival; RR=response rate; RRMM=relapsed or refractory multiple myeloma; VGPR=very good partial response
^Literature review citations are listed in Attachment 2

Skin cancer

Table A6 and Table A7 (Attachment 1) summarise the characteristics of the skin cancer submissions/resubmissions (N=28) that relied on surrogate outcomes for the clinical claim or PBAC decision.

In total, there were 12 drugs or 16 drug-indication pairs (i.e., 2 drugs had submitted to the PBAC for multiple indications in skin cancer). Half (14/28) of the PSDs were for resubmissions. Most of the skin cancer submissions were for targeted therapies (86%, 24/28 PSDs). The PBAC had considered 7 surrogate measures in skin cancer during the study period. The most frequently used surrogates were PFS (50%, 14/28), recurrence-free survival (25%, 7/28), and objective RR (21%, 6/28), measured as either a primary or secondary outcome in the clinical trials supporting the submission. 61% (17/28) of the PSDs claimed superior clinical efficacy and most presented a cost-utility analysis to the nominated main comparator (16/17 PSDs). Of note, one resubmission made two claims for different populations, one non-inferiority claim supported by a cost-minimisation analysis and one superior claim supported by a cost-utility analysis. Surrogate outcome(s) from the clinical evaluations were used in the modelled economic evaluations in 83% (15/18) of submissions/resubmissions.

The PBAC considered the clinical effectiveness of the treatment (relative to comparator) to be clinically significant in 56% (10/18) of the submissions/resubmissions that claimed superiority. The main comparators across the pivotal trials supporting all the submissions/resubmissions included active comparators (39%, 11/28), placebo (36%, 10/28), none (21%, 6/28) given it was a single arm study, or same drug at a different dose or schedule (4%, 1/28).

Of the 28 submissions, 50% (14/28) received a positive PBAC recommendation, with 39% (11/28) not recommended and 11% (3/28) deferred. The PBAC recommended listing with an RSA on 14% (4/28) of occasions and with managed access on one occasion (4%) (pembrolizumab, March 2015 PBAC meeting). Of the submissions/resubmissions with a positive PBAC recommendation, 71% (10/14) presented immature OS data. For 4 (29%) of these submissions/resubmissions, the pivotal trials have since published their final OS results.

Of the 14 submissions/resubmissions that relied on a surrogate but were not recommended or were deferred by the PBAC, the main reason (71%) was related to both clinical and economic uncertainty. Further, 55% (6/11) of submissions/resubmissions that were not recommended by the PBAC, the PBAC had documented in the PSDs that the OS data reviewed by the committee was immature to support the clinical claim (see Table 2).

Of all the submissions that relied on surrogate measures, the PBAC had documented comments about the surrogate and OS relationship in the PSDs in 36% (10/28) of the cases.

Narrative review of past PBAC decisions using surrogates in skin cancer

The majority of submissions for skin cancer were for melanoma treatments (19/28). There were 3 submissions for 2 drugs (vismodegib and sonidegib) to treat basal cell carcinoma (BCC); 5 submissions for 2 drugs (cemiplimab and ingenol) to treat or prevent squamous cell carcinoma (SCC) and 1 submission for a treatment (avelumab) for Merkel cell carcinoma (MCC).

Measurement of surrogate endpoints

How and when surrogate endpoints are measured continues to evolve. The Response Evaluation Criteria in Solid Tumors (RECIST) criteria is the standard for measuring response in solid tumours (Eisenhauer et al., 2009), though how the criteria are applied may vary between clinical trials. Lack of uniformity across clinical trials has been identified as a barrier for data comparison, and the International Neoadjuvant Melanoma Consortium has published recommendations for clinical trial design to address this (Amaria et al., 2019). The anticancer activity of newer immunotherapy agents programmed cell death protein 1 (PD-1) and programmed death-ligand 1 (PD-L1) inhibitors is different to other cancer therapies in that they may take longer to show an effect, and due to the immune response, may also induce a 'pseudoprogesssion' which impacts the measurement of true disease progression (Nie et al., 2020). Examples of PBAC comments relating to measurement of surrogate outcomes are in Text Box 16.

Text Box 16: Comments relating to measurement of PFS and RR in the PBAC's decisions for skin cancer

> *'The PBAC also noted that progression in the trials was determined by meeting RECIST criteria based largely on repeated images taken of the tumours, which may not be as relevant or meaningful an outcome to the patient as progression which manifests with symptoms. The PBAC noted that adjustment was made in the assessment of PFS by the trials for the phenomenon of "pseudo-progression" (where the immunotherapy may induce an early increase in the size of tumours), but could not be confident that this adjustment was long enough for all patients, nor how such an inadequacy might bias the comparisons across these immunotherapies.'* (paragraph 7.4, nivolumab plus ipilimumab, PSD, November 2015 PBAC meeting) [not recommended]
>
> *'The PBAC considered that the statistically significant increase in overall response rates (from 12% with ipilimumab to 33% with pembrolizumab) and the statistically significant prolongation in median progression-free survival based on the RECIST criteria (from 2.8 months with ipilimumab to 4.1 months with pembrolizumab) were likely to be clinically meaningful. Changing the definition and assessor of progression events increased the difference in median progression-free survival (from 3.3 months with ipilimumab to 7.2 months with pembrolizumab). However, the PBAC noted that overall response rates for pembrolizumab were substantially less than the 90% response rate which was strongly reported by consumers who met with PBAC representatives prior to the meeting. The PBAC was unable to determine how the discrepancy between the consumers' perception and the measured benefit had arisen. The PBAC noted that there might be some responder bias in these perceptions (those who experienced or were aware of instances of poor outcomes would be less likely to advocate for the medicine), and that the selected results of subgroups of KN-001 published in two peer-reviewed journal articles (Robert et al, 2014 and Hamid et al, 2013) had conveyed a more favourable impression of pembrolizumab than the data submitted from KN-001 or KN-006 as the basis for the PBAC consideration. The PBAC noted again that early results of studies are often very favourable and that*

over time the true impact of a medicine is revealed to be less impressive.' (paragraph 7.12, pembrolizumab, PSD, March 2015 PBAC meeting) [recommended]

Recurrence-free survival

While PFS (15/28) and objective or overall RR (8/28) were the most common surrogate measures relied on in submissions for skin cancer treatments, recurrence-free survival (RFS) data was presented in 7/28 submissions.

The broad definition of RFS, also known as relapse-free survival, is time from randomisation to recurrence or death (Suciu et al., 2017), however, like other surrogate endpoints, clinical trial protocols vary in the definition and measurement of RFS. PBAC has commented on the measurement of RFS and the validity of RFS as a surrogate for OS in skin cancer submissions, noting that the validity of RFS as a surrogate may differ depending on disease stage and the therapeutic class under investigation in the clinical trial. Examples of comments by ESC and PBAC on the validity of RFS are in Text Box 17.

Text Box 17: The ESC's and PBAC's comments on the measurement and validity of RFS in PBAC decisions for skin cancer

'The ESC noted that the definition of RFS in CA209238 was the time between the date of randomisation and the date of first recurrence (local, regional or distant metastasis), new primary melanoma, or death (whatever the cause), whichever occurs first; and in CA184029 was the time between the date of randomisation and the date of first recurrence (local, regional or distant metastasis) or death (whatever the cause), whichever occurs first. The ESC noted the use of a composite endpoint consisting of different health outcomes and different definitions of RFS across the two trials potentially affects the transitivity of the trials used in the indirect comparison.' (paragraph 7.4, nivolumab, PSD, July 2018 PBAC meeting) [not recommended]

'The ESC considered that: the suitability of RFS as a surrogate for overall survival in the assessment of pembrolizumab as adjuvant therapy for resectable Stage III melanoma has not been established; and; only overall survival data will capture the total impact of listing pembrolizumab as adjuvant therapy for the overall treatment of the target patients. The ESC advised [that] further evidence was required to establish the surrogacy relationship, if any, between RFS and overall survival with PD-1 inhibitor therapy, which may also need to be cancer specific.' (paragraph 6.10, Pembrolizumab, PSD, November 2018 PBAC meeting) [not recommended]

'No overall survival (OS) data were presented in the resubmission. The resubmission reproduced similar arguments as in the previous submission, regarding RFS as a surrogate for OS (although the RFS to OS surrogate relationship was no longer applied in the economic model). OS remains the most clinically appropriate endpoint that would capture the "overall clinical benefit" associated with adjuvant nivolumab therapy in the completely resected curative setting. The ESC advised that DMFS [distant metastases free survival] may be more closely related to OS than RFS, although subject to similar levels of uncertainty in the absence of OS data.' (paragraph 6.10, nivolumab, PSD, march 2019 PBAC meeting) [not recommended]

Other surrogate endpoints presented in submissions to PBAC for skin cancer treatments

Clearance of solar keratosis, durable response rate and distant metastases-free survival (DMFS) were also presented as evidence to support the clinical claims made in submissions for skin cancer treatments. Clearance of solar keratosis was presented as a surrogate for prevention of squamous cell carcinoma (SCC) in 3 submissions for the topical treatment ingenol. Durable response rate is a relatively new surrogate endpoint used in immunotherapy clinical trials for cancers including melanoma and a standard definition has not yet been adopted (Borcoman et al., 2019). The trial included in the PBAC submission defined durable response rate as 'objective response lasting continuously ≥ 6 months' (Andtbacka et al., 2015). DMFS is a measure of how well the treatment prevents metastases and is defined as the 'time from randomisation to the development of any distant

metastases or death' (Amabile et al., 2021). Like other surrogate endpoints, clinical trials have measured DMFS in different ways, which adds complexity to interpreting the results (Mo et al., 2023). PBAC comments relating to these emerging surrogate endpoints are included in Text Box 18.

Text Box 18: Comments on validity of less commonly used surrogate endpoints in the PBAC's decisions for skin cancers

'The PBAC considered that convincing data were not presented to quantify the reduction in risk of SCC that would be attributed to solar keratosis clearance. The PBAC considered that such data would be essential to establish solar keratosis clearance as a surrogate outcome for reduction in progression to SCC. The PBAC therefore considered that it was not possible to quantify the clinical benefit in terms of reduced SCCs that would accrue from clearance of solar keratoses.' (paragraph 12.4, ingenol, PSD, November 2013 PBAC meeting) [not recommended]

'The PBAC considered that durable response rate, the primary endpoint in the key trial (OPTiM), was not a validated endpoint, as it was clinician assessed, and allowed patients to be counted as having durable response despite having disease relapse or disease progression after 6 months.' (paragraph 7.5, talimogene, PSD, July 2016 PBAC meeting) [not recommended]

'The ESC advised that DMFS may be more closely related to OS than RFS, although subject to similar levels of uncertainty in the absence of OS data.' (paragraph 6.10, nivolumab, PSD, March 2019 PBAC meeting) [not recommended]

Line of therapy

Therapies may have different efficacy in different stages of disease, when used in combination, or if given as neoadjuvant or adjuvant therapy. Examples of PBAC comments on the efficacy of treatments when used in different settings are included in Text Box 19.

Text Box 19: The PBAC's comments relating to evidence of effectiveness in different lines of treatment and in patients in different stages of disease in PBAC decisions for skin cancer

'The PBAC considered that the claim of superior effectiveness compared with chemotherapy was reasonable. In the second line setting, the PBAC noted PFS at 1 year with avelumab was 30% compared with 0% for standard of care, and OS at 6 months was 70% compared with approximately 30%, and at 12 months was 50% compared with 0%. The PBAC also noted for responders, an extended duration of benefit with the median duration of response not reached in the analysis based on a ≥24 months of follow-up. In the first line setting, the PBAC noted that JAVELIN Merkel 200 study part B only had 30 patients and the data was less mature. However, the results were favourable towards avelumab relative to standard of care (Study Obs001 and Iyer 2016) demonstrating improved ORR (50%-51% vs. 29%-55%), a higher CR rate (16%-18% vs. 13%-14%), improved duration of response (11.3 months vs 2.8-6.7 months), and improved 6-month OS (83% vs. 67%). Additionally, the PBAC noted that the durable response rate (DRR) was higher in the first-line cohort (37.4%) than in the second-line and later cohort (29.1%) of the JAVELIN Merkel 200 study, and this is consistent with the claim that avelumab may provide greater benefit in the first-line compared with second-line and later setting.' (paragraph 7.10, avelumab, PSD, July 2018 PBAC meeting) [recommended]

'The PBAC agreed with the ESC in that there was genuine clinical equipoise amongst clinicians as to the optimal sequence of therapies in BRAF mutant metastatic melanoma. The PBAC noted that two international RCTs comparing the effects on overall survival of BRAF-targeted therapy and immunotherapy (DREAMseq and SECOMBIT) were in progress and requested that the results of these trials be made available to the PBAC once they are completed.' (paragraph 7.8, nivolumab and ipilimumab, PSD, November 2019 PBAC meeting) [recommended]

'The PBAC considered that it was not appropriate to rely solely on data from the STEVIE trial to inform the vismodegib response rate, as that study reported a higher response compared to the other studies, had the lowest proportion of participants with metastatic disease, and used investigator rather than independent response assessment. A less optimistic interpretation of the total data, giving more weight to response rates from the other studies, independent assessment, and response rates in metastatic

disease, would suggest a vismodegib response rate substantially less than 50%.' (paragraph 5.5, visodegib, PSD, November 2016 PBAC meeting) [Outcome: Advice provided]

Decision making for skin cancer submissions when OS data are immature

Comments made by the PBAC relating to immature data in skin cancer submissions are similar to those PBAC had made for lung and blood cancer submissions. Examples of PBAC comments on immature data are included in Text Box 20.

Text Box 20: The PBAC's comments relating immature OS data in submissions for skin cancer.

'The PBAC considered that the evidence presented in the submission was immature, particularly in the first line setting, and the magnitude of benefit was uncertain. However, the Committee also considered that it was unlikely that phase III data would be available in the near future. The PBAC noted that there is emerging phase II data for anti-PD-1 drugs such as nivolumab and pembrolizumab that show this class of drugs are effective in the treatment of MCC.' (paragraph 7.9, avelumab, PSD, July 2018 PBAC meeting) [recommended]

'The PBAC considered that the sponsor needs to contribute to the cost of the proposed post market data collection, and also needs to provide updated overall survival data to the PBAC, when it is available, from the BREAK-3 randomised trial.' (paragraph 6.4, dabrafenib, PSD, July 2013 PBAC meeting) [recommended]

'A data update from CA238, nivolumab versus ipilimumab, provided RFS and distant metastases free survival (DMFS) data at a minimum of 36 months follow-up (compared to 24 months in the March 2019 submission), and a 7-year update was available for CA029 (ipilimumab versus placebo, compared to 5-years in the March 2019 submission). The PBAC noted that treatment effect for nivolumab versus ipilimumab was maintained in terms of RFS and DMFS at 36 months follow-up. The PBAC noted that no updated OS data was provided for CA238. The PBC noted that the treatment effect for ipilimumab versus placebo was maintained in terms of OS at 7 years of follow-up.' (paragraph 5.7, nivolumab, PSD, July 2019 PBAC meeting) [deferred]

'However, the PBAC considered that interpretation of these results were limited by the single arm design of Study 1423 and Study 1540 and small sample sizes. The PBAC also agreed with the ESC that the overall survival (OS) data for cemiplimab were immature thus further limiting interpretation.' (paragraph 6.19, cemiplimab, PSD, November 2020 PBAC meeting) [not recommended]

'The PBAC noted the updated clinical data provided from the September 2016 database lock of CA209-067. This data reaffirmed the PBAC's previous conclusion that NIVO+IPI demonstrated an improvement in PFS over nivolumab monotherapy. Nonetheless, the OS data was still immature and showed no statistically significant effect for NIVO+IPI beyond that of nivolumab monotherapy, and there was no evidence of improved quality of life.' (paragraph 7.5, nivolumab and ipilimumab, PSD, March 2017 PBAC meeting) [not recommended]

Expediate literature review on the validity of surrogate endpoints used in clinical trials of therapies for skin cancer

A review of published validation studies identified 4 articles assessing surrogate endpoints used in skin cancer trials and 3 larger meta-analyses which included multiple cancer types. While none of the meta-analyses found definitive evidence that any of the surrogates accurately predicted OS, there was evidence that PFS and RFS are associated with OS in trials of immune checkpoint inhibitors, and this may be an area for future research. Table 9 presents a summary of the main findings on surrogates used in skin cancer trials.

Table 9: Descriptive summary of literature on the validity of surrogate outcomes used in melanoma clinical trials

Review article^ (year)	Type	# trials included	Main Surrogate	Findings
Belin[2] (2020)	Meta-analysis	91 (2 melanoma)	PFS	Possible for Chemotherapy/targeted therapy Not validated for Immunotherapy
Haslam[9] (2019)	Systematic review of meta-analyses	78 (2 melanoma)	PFS, RFS	RFS had low surrogacy strength in the adjuvant setting (when all data were included*); PFS had high surrogacy strength in the metastatic setting
Walia[10] (2022)	Systematic review of validation studies by FDA	15 (3 melanoma)	Overall RR, PFS	Low correlation
Coart[55] (2020)	Meta-analysis of IPD	2	RFS	In adjuvant therapy with immune checkpoint inhibitors: moderate association between RFS and OS at trial level; strong patient-level association
Nie[56] (2020)	Meta-analysis	8	DCR, Objective RR, PFS	PFS is possible surrogate for OS in anti-PD-1/PD-L1 trials in advanced melanoma. ORR and DCR not strongly correlated with OS
Sheth[57] (2020)	Analyses of trial data submitted to FDA for treatment of advanced melanoma	13	Objective RR, PFS	Weak correlations between ORR or PFS and OS; Responders in either arm (experimental or control) had better PFS and OS than non-responders
Suciu[54] (2017)	Meta-analysis of IPD	13	RFS	Possible surrogate at patient level in adjuvant interferon trials in Resected Stage II-III melanoma

FDA=United States Food and Drug Administration; IPD=individual patient data; DCR=disease control rate; PFS=progression-free survival; RFS=relapse-free survival; RR=response rate; *meta-analysis of Suciu (2017)
^ Literature review citations are listed in Attachment 2

Breast cancer

Table A8 and Table A9 (Attachment 1) summarise the characteristics of the breast cancer submissions/resubmissions (N=31) that relied on surrogate outcomes for the clinical claim or PBAC decision.

In total, there were 15 drugs or 22 drug-indication pairs (i.e., 6 drugs had submitted to the PBAC for multiple indications in lung cancer). 35% (11/31) PSDs were for resubmissions. Most of the breast cancer submissions were for targeted therapies (90%, 28/31 PSDs). The PBAC had considered 14 surrogates in breast cancer during the study period. The most frequently used surrogates were PFS (68%, 21/31), invasive disease-free survival (19%, 6/31), and disease/event free survival (16%, 5/31), measured as either a primary or secondary outcome in the clinical trials supporting the submission. Most of the PSDs (81%, 25/31) claimed superior clinical efficacy and presented a cost-utility analysis to the nominated main comparator (77%, 24/31). In both cases, 2 PSDs had an additional clinical claim of non-inferiority and presented a cost-minimisation analysis to support that claim. In all submissions with a modelled economic evaluation, the surrogate outcome from the clinical evaluations was used in the model.

The PBAC considered the clinical effectiveness of the treatment (relative to comparator) to be clinically significant in 48% (12/25) of the submissions/resubmissions that claimed superiority. The main comparators across the pivotal trials supporting all the submissions/resubmissions included placebo (77%, 24/31), active comparators (13%, 4/31), none (6%, 2/31) given it was a single arm study, or same drug at a different dose or schedule (3%, 1/31).

Of the 31 submissions, 39% (12/31) received a positive PBAC recommendation, with 55% (17/31) not recommended and 6% (2/31) deferred. The PBAC recommended listing with an RSA on 13% (4/31) of occasions and there were no recommendations to list with managed access. Of the submissions/resubmissions with a positive PBAC recommendation, 75% (9/12) presented immature

OS data. For 7 (58%) of these submissions/resubmissions, the pivotal trials have since published their final OS results.

Of the 19 submissions/resubmissions that relied on a surrogate but were not recommended or were deferred by the PBAC, the main reason (84%) was related to both clinical and economic uncertainty. Further, for 71% (12/17) of submissions/resubmissions that were not recommended by the PBAC, the PBAC had documented that the OS data was immature to support the clinical claim (see Table 2).

Narrative review of past PBAC decisions using surrogates in breast cancer

While PFS continued to be the main surrogate outcome presented as evidence of efficacy in PBAC submissions (20/31), invasive disease-free survival (iDFS) is an emerging surrogate endpoint first used in PBAC submissions in 2019. There was one submission which relied on pathological complete response rate of the primary breast tumour (bpCR) for a neoadjuvant therapy. The comments by PBAC relating to PFS and immature OS, measurement of progression in clinical trials, and the use of surrogates creating more uncertainty in economic models in submissions for breast cancer were similar to comments made for lung, blood and skin cancer submissions.

Invasive disease-free survival

The DATECAN initiative made recommendations for surrogate endpoint definitions in breast cancer through a consensus process in 2015 (Gourgou-Bourgade et al., 2015). The expert panel considered iDFS more relevant than DFS in breast cancer. A recent systematic review of DFS, which included trials with varying definitions of DFS, some of which were consistent with iDFS, concluded that DFS could be used as a surrogate for OS in human epidermal growth factor receptor 2 positive (HER2+) early breast cancer (Saad et al., 2019). Examples of comments by PBAC relating to iDFS in submissions are included in Text Box 21.

Text Box 21: Comments relating to iDFS in the PBAC's decisions for breast cancer

'The PBAC considered that it may be appropriate to restrict use to HR+ patients given the greater improvement in iDFS in this subgroup, however the ExteNET trial was not powered to detect if HR status is a treatment effect modifier.' (paragraph 7.2, neratinib, PSD, March 2019 PBAC meeting) [not recommended]

'Overall, the PBAC considered the difference in iDFS to be small and uncertain given the potential for a high risk of bias due to protocol amendments, the reliance on a subgroup of the ExteNET trial, and potential applicability issues with the trial relating to underrepresentation of node negative patients and the prior neoadjuvant and adjuvant treatments used. Further, the PBAC noted without overall survival data the long term benefits of neratinib therapy are unknown.' (paragraph 7.11, neratinib, PSD, March 2019 PBAC meeting) [not recommended]

'The PBAC was satisfied that abemaciclib in combination with ET was both statistically and clinically superior to the nominated comparator in improving IDFS and DRFS. While acknowledging that the IDFS benefit appeared modest and was smaller compared to agents seen previously, the PBAC considered a 3.5% absolute difference may be clinically meaningful in the adjuvant setting where the goal is cure. The PBAC considered that IDFS being employed as a surrogate for OS was uncertain but generally plausible. However, the PBAC noted the relationship between IDFS and OS is uncertain for abemaciclib, given the OS data were immature and no difference in OS was observed at the most recent data cutoff.' (paragraph 7.8, abemaciclib, PSD, March 2022 PBAC meeting) [not recommended]

'The primary outcome of the KATHERINE trial was iDFS. The PBAC considered that the claim that T-DM1 was superior to trastuzumab in terms of comparative effectiveness was reasonable with respect to iDFS. The PBAC considered the 50% reduction in the risk of recurrence or death in the T-DM1 treatment group, which was statistically significant in favour of T-DM1 with HR=0.50 (95% CI: 0.39, 0.64; p=<0.0001), was clinically meaningful. The PBAC also noted the iDFS benefits of T-DM1 were

demonstrated across different subgroups, regardless of ER status, nodal status, or whether neoadjuvant therapy comprised HER2 monotherapy (trastuzumab alone) versus doublet HER2 therapy (trastuzumab + additional HER2-directed agents).

The PBAC noted that OS, which was a secondary outcome, was not statistically significant (HR = 0.70 (95% CI: 0.47, 1.05)), but considered this was likely due to the immaturity of the trial data with respect to this outcome (6.6% of patients across both arms had died at the data-cut reported in the submission). The PBAC considered that a gain in OS is plausible given the strong iDFS results reported in the well-conducted trial and considered there was a moderate level of certainty around the OS benefits. However, the PBAC considered any OS gain was of uncertain magnitude given the lack of statistically significant OS results.' (paragraph 7.7-7.8, trastuzumab-emtansine, PSD, November 2019 PBAC meeting) [recommended]

Pathological complete response

Another emerging surrogate, pathological complete response (pCR), is yet to show strong correlation with OS (Conforti et al., 2021; Haslam et al., 2019; Savina et al., 2018; Walia et al., 2022), with the exception of an association with improved EFS and OS in triple negative breast cancer (TNBC) and HER2+ breast cancer in neoadjuvant trials (Liu et al., 2021). PBAC considered one submission which presented pCR results as the main clinical evidence. PBAC comments on pCR are summarised in Text Box 22.

Text Box 22: Comments relating to pCR in the PBAC's decisions for breast cancer

'The PBAC noted that the primary endpoint in NEOSPHERE was pathological complete response in the breast (bpCR) with total response (tpCR) assessed post-hoc. In PEONY the primary endpoint was tpCR. The PBAC considered tpCR to be more clinically relevant than bpCR. The PBAC noted, although meta-analyses of neoadjuvant trial data have indicated this outcome informs prognosis, that tpCR has not been definitively demonstrated to be a surrogate endpoint for disease free or overall survival. The PBAC noted tpCR increased by approximately 18% with the addition of neoadjuvant pertuzumab in both NEOSPHERE (17.8%; 95% CI 4.6%, 31.0%, p=0.008) and PEONY (17.5%, 95% CI 6.9%, 28.0%, p=0.001).

The PBAC noted this difference exceeded the minimal clinically important difference (MCID) defined in the submission of 15%. However, the PBAC considered the MCID of 15% was not adequately justified, noting the role of tpCR as a surrogate measure for patient relevant outcomes was unclear, especially with the availability of T-DM1 in the adjuvant setting for patients without a tpCR.' (paragraph 7.4-7.5, pertuzumab, PSD, March 2020 PBAC meeting) [not recommended]

Place in therapy

Neoadjuvant, preoperative therapies are aimed at better rates of breast conservation and downstaging primary tumours (Gion et al., 2021). With the increase in research and use of therapies in the neoadjuvant setting, the clinical place of therapies in adjuvant settings is changing. PBAC comments on submissions received at a time of change in clinical practice are included in Text Box 23.

Text Box 23: The PBAC's comments on use of therapies in the neoadjuvant and adjuvant settings in PBAC decisions for breast cancer

'The PBAC considered the clinical place in therapy for pertuzumab for HER2 positive, lymph node positive eBC is unclear. The PBAC noted that there is a move towards neoadjuvant therapy in patients with high-risk HER2+ eBC. This was supported by an Australian breast cancer consumer group which highlighted that pertuzumab's place in therapy is changing and its use in the neoadjuvant setting is becoming more widely recommended for women with HER2 positive eBC. The PBAC noted that a neoadjuvant approach allows assessment of response to therapy at the time of surgery. A recent large meta-analysis of eBC trials with neoadjuvant chemotherapy reported the prognostic importance of pathological complete response (pCR), which correlates with event free survival and OS. The PBAC

also noted that the results from a trial assessing T-DM1 in patients with residual invasive disease after completing neoadjuvant chemotherapy + trastuzumab support a change in the treatment pathway and that T-DM1 may become an alternative treatment to pertuzumab in the adjuvant setting.' (paragraph 7.2, pertuzumab, PSD, March 2019 PBAC meeting) [not recommended]

'The PBAC noted that there was no improvement in PFS or DFS demonstrated in the NEOSPHERE trial (PFS HR=0.69, 95% CI 0.34, 1.40; DFS HR=0.60, 95% CI 0.28, 1.27), and that the trial was not powered to assess these outcomes. The PBAC further noted that in the KATHERINE trial patients with residual disease derived a similar benefit from adjuvant T-DM1 regardless of whether they received pertuzumab in addition to trastuzumab in the neoadjuvant setting, and that the DFS at 3 years was high (88%) and similar to that observed with neoadjuvant pertuzumab (92% at 3 years). The PBAC considered that with the current treatment algorithm, it is unclear whether patients who achieve pCR post neoadjuvant therapy have improved DFS or OS over patients with residual disease who receive adjuvant T-DM1. The PBAC also considered it is unclear whether patients who do not achieve a pCR and go on to receive adjuvant TDM1 derive any benefit from neoadjuvant pertuzumab.' (paragraph 7.7, pertuzumab, PSD, March 2020 PBAC meeting) [not recommended]

Decision making for breast cancer submissions when OS data are immature

Comments made by the PBAC relating to immature data in breast cancer submissions (Text Box 24) are similar to those PBAC had made for lung, blood and skin cancer submissions. PBAC has also noted that for some treatments in the early breast cancer setting, an OS benefit may not be evident for many years.

Text Box 24: The PBAC's comments on immature OS data presented in submissions for breast cancer.

'The PBAC noted the iDFS event-free rates for node-positive patients were 91.99% vs. 90.15% at 3 years (difference of 1.84%) and 89.88% vs. 86.68% at 4 years (difference of 3.2%), for Ptz+T+Chemo versus T+Chemo, respectively. The PBAC maintained that the magnitude of absolute benefit was small. The improvement in OS was not statistically significant in the ITT population (P=0.4673) or in the lymph node positive subgroup (P=not reported). The PSCR argued that the benefits in the curative adjuvant setting occur over a long time horizon, with trastuzumab trials not showing a statistically significant improvement in OS at 3 and 4 years but showing a statistically significant improvement after 10-11 years. The PBAC agreed with the ESC's view that an OS benefit for pertuzumab may be seen over time, as was seen for trastuzumab, however this benefit has not yet been demonstrated.' (paragraph 7.6, pertuzumab, PSD, March 2019 PBAC meeting) [not recommended]

'The PBAC noted that final overall survival results from the PALOMA-2 trial are expected to be available in 2020 and advised that if listed, the sponsor should provide these results to the PBAC.' (paragraph 6.20, palbociclib, PSD, March 2018 PBAC meeting) [recommended]

Expediate literature review on the validity of surrogate endpoints used in clinical trials of therapies for breast cancer

The literature review identified 8 relevant breast cancer specific meta-analyses on surrogate endpoints and 4 broader meta-analyses which included breast cancer. Two focused on early breast cancer, one on advanced and two on metastatic breast cancer. The FDA accepts durable objective response rates, PFS, disease-free survival (DFS), event-free survival (EFS) and pathological complete response (pCR) in applications for treatment of breast cancer. EFS is used in the neoadjuvant setting while DFS is used in the adjuvant setting (Gyawali et al., 2020). Results of validation studies for PFS have been conflicting, with Gyawali et al finding PFS more strongly correlated in second line therapies, but Forsythe et al finding stronger correlation in the first line setting (Forsythe et al., 2018). No study found strong evidence of pCR as a predictor for OS, however, 2 studies found that DFS was correlated with OS in HER2+ breast cancer. Table 10 summarises the key findings from the review papers.

Table 10: Summary of validation studies for surrogate endpoints used in breast cancer clinical trials.

Review article^ (year)	Type	# trials/ studies included	Main Surrogate	Findings
Belin[2] (2020)	Methodological review of meta-analyses	91 (13 BC)	PFS	PFS validated in 3/13 studies and possibly validated in a further 1 study
Haslam[9] (2019)	Systematic review of meta-analyses	78 (15 BC)	pCR, PFS	All: Low-medium surrogacy strength across neoadjuvant, metastatic and adjuvant settings.
Walia[10] (2022)	Systematic review of validation studies by FDA	15 (2 BC)	CBR, pCR	Low correlation with OS
Savina[15] (2018)	Critical review of meta-analyses	53 (12 BC)	pCR, DFS, PFS, TTP	Low-medium correlation with OS
Saad[17] (2019)	Systematic review and meta-analysis of DFS in adjuvant HER2+ EBC trials	8	DFS	DFS associated with OS a trial-level
Lux[18] (2019)	Systematic review and meta-analysis of HR+ HER2− MBC	16	PFS	Possible surrogate: if the upper confidence limit of $HR_{PFS} < 0.6$ (STE) there may be a statistically significant effect on OS
Li[19] (2018)	Meta-analysis of ABC RCTs	37	PFS, TTP	All: Moderate correlation with OS
Hirai[20] (2020)	Meta-analysis and correlation analysis of PFS in TNBC	14	PFS	PFS strongly correlated with OS in advanced or metastatic TNBC
Guarneri[21] (2022)	Meta-analysis of RCTs of neoadjuvant therapy in HER2+ EBC	4	pCR, RFS	Patients who achieved pCR had improved RFS and OS; the association was stronger in the HR− subgroup
Forsythe[22] (2018)	Systematic review and meta-analysis of RCTs in HR+, HER2− MBC	40	PFS, TTP	Significant association between PFS/TTP and OS; Stronger correlation in trials including only HR+ patients; STE for OS benefit was 5-6 months of incremental PFS/TTP
Liu[23] (2021)	Meta-analysis of RCTs in neoadjuvant setting	25	pCR	pCR associated with improved EFS and OS in TNBC and HER2+ BC N.B. The strength of correlation was not tested.
Gyawali[24] (2020)	Systematic review of studies for surrogates accepted by FDA for breast cancer	13	pCR, DFS, EFS, Objective RR, PFS	DFS strongly correlated with OS in HER2 positive trials; EFS yet to be validated; pCR, ORR and PFS not strongly correlated with OS, but PFS possible surrogate in 2nd line therapies, not 1st line
Conforti[25] (2021)	Systematic review and meta-analysis of pCR in neoadjuvant RCTs in EBC	54	pCR	Weak association between pCR and DFS or OS.

ABC=advanced breast cancer; BC=breast cancer; CBR=clinical benefit rate; pCR=pathological complete response; DFS=disease free survival; EBC=early breast cancer; EFS=event free survival; FDA=United States Food and Drug Administration; HER2=human epidermal growth factor receptor 2; HR= hormone receptor; HR_{PFS}=Hazard ratio for PFS; MBC=metastatic breast cancer; PFS=progression-free survival; RCT=randomised controlled trials; RFS=recurrence free survival; RR=response rate; STE=surrogate threshold effect; TNBC=triple negative breast cancer; TTP=time to progression
^ Literature review citations are listed in Attachment 2

Broad cancer types with less than 15 submissions/resubmissions

Table A10 and Table A11 (Attachment 1) summarise the characteristics of broad cancer types with less than 15 submissions/resubmissions (N=75), that relied on surrogate outcomes for the clinical claim or PBAC decision. These are summaries of data collected from PBAC decision documentation for the project based on the review objectives. Broad cancer types with more than 5 PSDs are described individually.

Common themes identified in the narrative review for lung, blood, skin and breast cancer included the immaturity of data presented in the clinical claim, the lack of validation of surrogate outcomes, the quality of the data, including potential bias in measurement of outcomes, confounding from crossover to the active arm in clinical trials and sample size. Comments by PBAC relating to these themes were similar across cancer types. For the remaining cancer types with < 15 submissions, the narrative review focused on surrogate endpoints unique to that cancer type and any new emergent themes.

There were no new emergent themes in the remaining submissions, which predominately relied on PFS or response outcomes for their clinical claim, however, there were two additional surrogate measures used to support submissions for prostate cancer treatments.

Renal cancer

In total, there were 13 submissions/resubmissions for 7 drugs or 7 drug-indication pairs. 62% (8/13) of PSDs were for resubmissions. The PBAC had previously considered two surrogates in renal cancer: PFS (100%), and objective response rate (8%, 1/13), measured as either a primary or secondary outcomes in the clinical trials supporting the submission. Most of the submissions/resubmissions (62%, 8/13) claimed superior clinical efficacy. There were 6 submissions/resubmissions (46%) that presented a cost-utility analysis and 2 that presented cost-consequence analyses (15%) to the nominated main comparator. In all submissions that presented modelled economic evaluations, the surrogate outcome from the clinical evaluations was used in model.

The PBAC considered the clinical effectiveness of the treatment (relative to comparator) to be clinically significant in 1/8 (13%) submissions that claimed superiority. The main comparators across the pivotal trials supporting all the submissions/resubmissions were placebo (39%, 5/13) or active comparators (62%, 8/13).

Of the 13 submissions, 46% (6/13) received a positive PBAC recommendation, with 54% (7/13) not recommended and none deferred. The PBAC recommended listing with an RSA on 23% (3/13) of occasions and there were no recommendations to list with managed access. Of the submissions/resubmissions with a positive PBAC recommendation, 33% (2/6) presented immature OS data. In both instances, the pivotal trials have since published their final OS results.

Of the 7 submissions/resubmissions that relied on a surrogate but were not recommended by the PBAC, the main reason (71%) was related to clinical uncertainty. Further, 71% (5/7) of submissions/resubmissions that were not recommended by the PBAC, the PBAC had documented that the OS data reviewed by the committee was immature to support the clinical claim.

Table 11 summarises the key findings of the literature review for surrogate outcomes used in renal cancer clinical trials.

Table 11: Summary of validation studies for surrogate endpoints with OS used in renal cancer clinical trials

Review article^ (year)	Type	# trials/ studies included	Main Surrogate	Findings
Belin[2] (2020)	Methodological review of meta-analyses	91 (4 renal)	PFS	PFS validated in 1/4 studies
Haslam[9] (2019)	Systematic review of meta-analyses	78 (6 RCC)	DFS, PFS	Low correlation of DFS and PFS with OS in adjuvant setting; Mixed results in metastatic setting; Medium correlation in immunotherapy trials
Walia[10] (2022)	Systematic review of validation studies by FDA	15 (1 RCC + 3 including RCC)	PFS	Low-medium correlation
Savina[15] (2018)	Critical review of meta-analyses	53 (5 renal)	DFS, PFS	Medium correlation
Harshman[51] (2018)	Meta-analysis of trials in localised RCC	13	DFS	Moderate correlation with OS
Bria[52] (2015)	Benchmarking analysis of trials of targeted therapy or immunotherapy for advanced RCC	19	PFS	PFS was correlated with OS in 1st line treatments of advanced RCC
Zhang[53] (2019)	Meta-analysis RCTS that used immunotherapies	11 (1 RCC)	Objective RR, PFS	Low correlation

DFS=disease free survival; PFS=progression-free survival; RCC=renal cell carcinoma; RCT=randomised controlled trials; RR=response rate;
^ Literature review citations are listed in Attachment 2

Ovarian cancer

In total, there were 10 submissions/resubmissions for 3 drugs or 8 drug-indication pairs (i.e., 2 drugs had submitted to the PBAC for multiple indications in ovarian cancer). 30% (3/10) of PSDs were for resubmissions. The PBAC had considered PFS as a surrogate in ovarian cancer (10/10 PSDs), measured as a primary outcome in the clinical trials supporting the submission. Most of the PSDs (70%, 7/10) claimed superior clinical efficacy, and most presented a cost-utility analysis (70%, 7/10) to the nominated main comparator. In both cases, one resubmission had an additional clinical claim of non-inferiority and presented a cost-minimisation analysis to support that claim. In all submissions that presented modelled economic evaluations, the surrogate outcome from the clinical evaluations was used in model.

The PBAC considered the clinical effectiveness of the treatment (relative to comparator) to be clinically significant in 3/7 (43%) of the submissions/resubmissions that claimed superiority. The main comparators across the pivotal trials supporting all the submissions/resubmissions were placebo (90%, 9/10) or active comparators (10%, 1/10).

Of the 10 submissions, 40% (4/10) received a positive PBAC recommendation, with 50% (5/10) not recommended and none deferred. There were no recommendations to list with an RSA or managed access. All the submissions/resubmissions with a positive PBAC recommendation presented immature OS data. In all instances, the pivotal trials are yet to publish their final OS results.

Of the 6 submissions/resubmissions that relied on a surrogate but were not recommended by the PBAC, the main reason (50%) was related to clinical uncertainty. Further, for 80% (4/5) of submissions/resubmissions that were not recommended by the PBAC, the PBAC had documented that the OS data was immature to support the clinical claim.

Table 12 summarises the key findings of the literature review for surrogate outcomes used in ovarian cancer clinical trials.

Table 12: Summary of validation studies for surrogate endpoints used in ovarian cancer clinical trials.

Review article^ (year)	Type	# trials/ studies included	Main Surrogate	Findings
Belin[2] (2020)	Methodological review of meta-analyses	91 (2 ovarian)	PFS	Not validated
Haslam[9] (2019)	Systematic review of meta-analyses	78 (3 ovarian)	PFS	Mixed
Savina[15] (2018)	Critical review of meta-analyses	53 (1 ovarian)	PFS	Medium correlation
Sjoquist[42] (2018)	Meta-analysis of trials on 1st line treatment for early OC	26	PFS	Moderate correlation
Siddiqui[43] (2017)	Systematic review and meta-analysis of RCTs in advanced, recurrent OC	39	Objective RR	Objective RR is possible surrogate for OS in 2nd and further line therapy
Paoletti[44] (2020)	Systematic review and Meta-analysis of PFS in 1st line ovarian cancer trials	17	PFS	High correlation at individual level but low correlation at trial level
Shimokawa[45] (2018)	Systematic review and meta-analysis of 2nd/3rd line chemotherapy trials in advanced or recurrent EOC	22	PFS, PPS	PFS was moderately associated with OS; PPS was strongly associated with OS

OC=ovarian cancer; OS=overall survival; PFS=progression-free survival; PPS=post-progression survival; RR=response rate
^ Literature review citations are listed in Attachment 2

Prostate cancer

In total, there were 9 submissions/resubmissions for 5 drugs or 7 drug-indication pairs (i.e., 2 drugs had submitted to the PBAC for multiple indications in prostate cancer). 33% (3/9) of PSDs were for resubmissions. The PBAC had previously considered six surrogates in prostate cancer. The most frequently used surrogates were metastasis-free survival (44%, 4/9), serum testosterone (33%, 3/9), and PFS (22%, 2/9), measured as a primary outcome in the clinical trials supporting the submission. Around half of the PSDs (56%, 5/9) claimed superior clinical efficacy, and presented a cost-utility analysis (56%, 5/9) to the nominated main comparator. In 80% (4/5) of submissions, the surrogate outcome from the clinical evaluations was used in the modelled economic evaluations.

The PBAC considered the clinical effectiveness of the treatment (relative to comparator) to be clinically significant in 60% (3/5) of the submissions/resubmissions that claimed superiority. The main comparators across the pivotal trials supporting the submissions/resubmissions were placebo (67%, 6/9) or active comparators (22%, 2/9).

Of the 9 submissions, 33% (3/9) received a positive PBAC recommendation, with 56% (5/9) not recommended and one deferred (11%). There were no recommendations to list with an RSA or managed access. Of the submissions/resubmissions with a positive PBAC recommendation, 33% (1/3) presented immature OS data. For one submission/resubmission, the pivotal trial has published their final OS results.

Of the 6 submissions/resubmissions that relied on a surrogate but were not recommended or were deferred by the PBAC, the main reason (67%) was related to clinical uncertainty. Further, for 40% (2/5) of submissions/resubmissions that were not recommended by the PBAC, the PBAC had documented that the OS data was immature to support the clinical claim.

Surrogate measures used in prostate cancer

Men diagnosed with prostate cancer may live with the disease for a long time, with 78% of men surviving 10 years from diagnosis (Cancer Research UK). Therefore, surrogate outcomes using biomarkers or measures of disease progression are common endpoints in prostate cancer clinical trials. Of the 9 submissions for prostate cancer, 2 relied on metastasis-free survival (MFS), with a

further 2 submissions including MFS, and 3 relied on biomarkers (e.g. serum testosterone levels or prostate specific antigen). MFS is defined as time to the occurrence of distant metastases or death and was validated as a surrogate for OS in radiotherapy trials of localised prostate cancer in 2017 (Xie et al., 2017). A recent meta-analysis also considered validation of MFS, biochemical failure and local failure in trials of surgery and hormone therapies in addition to radiotherapy trials in localised prostate cancer and concluded that MFS was a valid surrogate endpoint for OS due to strong correlation with OS (R^2= 0·78 95% CI: 0·59–0·89) (Gharzai et al., 2021). This study also found that biomarker-based endpoints, such as biochemical failure, did not have sufficient correlation with OS to be considered valid surrogate measures in prostate cancer. Comments made by the PBAC about use of MFS and biomarker surrogates in the clinical evidence and economic models for prostate cancer submissions are included in Text Box 25.

Text Box 25: PBAC comments on use of surrogate measures in prostate cancer submissions

'The PBAC advised that any resubmission would be a major submission and would require an economic model that does not assume a direct surrogate relationship between MFS and OS, regardless of the ratio used. The PBAC considered that the efficacy parameters of MFS/PFS, time to symptomatic progression and time to chemotherapy would offer a method of assessing the clinical benefit and cost-effectiveness of apalutamide.' (paragraph 7.13, apalutamide, PSD, November 2018 PBAC meeting) [not recommended]

'The PBAC noted that the economic model used investigator-assessed MFS in the base case, rather than blinded independent central review (BICR)-assessed MFS. The PBAC considered that there was a greater potential for detection bias with investigator-assessed MFS. Further, the PBAC noted the choice of investigator-assessed MFS was not conservative, which the PBAC considered to be an issue in the context of the large gain in MFS that was modelled based on immature data (with only 26% and 58% of patients progressing to metastases in the apalutamide and placebo arms, respectively). Overall, the PBAC considered that, in this case, BICR-assessed MFS was the more appropriate measure for use in the economic model.' (paragraph 7.10, apalutamide, PSD, July 2019 PBAC meeting) [not recommended]

'… the PBAC noted the study had limitations including the trial was based on surrogate endpoints (serum testosterone and PSA levels) which do not translate to improved overall survival.' (paragraph 7.5, abiraterone and methylprednisolone, PSD, March 2022 PBAC meeting) [not recommended]

Table 13 summarises the key findings of the literature review for surrogate outcomes used in prostate cancer clinical trials.

Table 13: Summary of validation studies for surrogate endpoints used in prostate cancer clinical trials

Review article^ (year)	Type	# trials/ studies included	Main Surrogate	Findings
Xie[60] (2017)	Systematic review of meta-analyses	24 (DFS) 19 (MFS)	DFS, MFS	DFS and MFS both valid surrogates for OS; MFS more strongly correlated with OS than DFS
Gharzai[50] (2021)	Meta-analysis of trials for localised prostate cancer	75	Biochemical/ local failure, PFS, MFS	PFS moderately correlated with OS; MFS strongly correlated with OS

DFS=disease free survival; MFS=metastasis free survival; OS=overall survival; PFS=progression-free survival; RR=response rate
^ Literature review citations are listed in Attachment 2

Bowel cancer

In total, there were 6 submissions/resubmissions for 3 drugs or 5 drug-indication pairs (i.e., 2 drugs had submitted to the PBAC for multiple indications in bowel cancer). 67% (4/6) of PSDs were for resubmissions. The PBAC had considered two surrogates in renal cancer: PFS (100%), and objective response rate (17%, 1/6), measured as either a primary or secondary outcome in the clinical trials supporting the submission. Half of the PSDs claimed superior clinical efficacy, and half claimed non-

inferior clinical efficacy. In 33% (2/6) of submissions that presented modelled economic evaluations, the surrogate outcome from the clinical evaluations was used in the model.

The PBAC considered the clinical effectiveness of the treatment (relative to comparator) to be clinically significant in one (20%, 1/5) of the submissions/resubmissions that claimed superiority. The main comparator across the pivotal trials supporting all the submissions/resubmissions was placebo (67%, 4/6), while one resubmission (17%) used an active comparator and one submission (17%) had no comparator.

Of the 6 submissions, 50% received a positive PBAC recommendation, with 50% not recommended. The PBAC recommended listing with an RSA on one (17%) occasion and there were no recommendations to list with managed access. Of the submissions/resubmissions with a positive PBAC recommendation, 67% (2/3) presented immature OS data. In all instances, the pivotal trials have since published their final OS results.

Of the 3 submissions/resubmissions that relied on a surrogate but were not recommended or were deferred by the PBAC, the main reason (67%) was related to clinical uncertainty. Further, for 67% (2/3) of submissions/resubmissions that were not recommended by the PBAC, the PBAC had documented that the OS data was immature to support the clinical claim.

Table 14 summarises the key findings of the literature review for surrogate outcomes used in bowel cancer clinical trials.

Table 14: Summary of validation studies for surrogate endpoints used in bowel cancer clinical trials

Review article[^] (year)	Type	# trials/ studies included	Main Surrogate	Findings
Belin[2] (2020)	Methodological review of meta-analyses	91 (15 colorectal)	PFS	PFS validated as surrogate for OS in 8/15 studies; with partial validation in a further 1
Haslam[9] (2019)	Systematic review of meta-analyses	78 (19 colorectal)	DFS, PFS, Overall RR, TTP	Mixed for all surrogates, ranging from low to high correlation
Savina[15] (2018)	Critical review of meta-analyses	53 (20 colorectal)	DFS, PFS	Medium-High correlation between DFS and OS in adjuvant fluoropyrimidine trials
Burzykowski[12] (2019)	Pooled analysis of RCTs using tumour size-based endpoints in metastatic colorectal cancer	20	Time to nadir, depth of nadir	Not acceptable surrogates in 1st line treatment
Ecker[13] (2022)	Meta-analysis of institutional cohort with resected colorectal liver metastases	2983 individuals	RFS	Minimal correlation between RFS and OS
Cremolini[14] (2017)	Systematic review and meta-analysis of 2nd line trials of targeted therapies in metastatic colorectal cancer	20	Objective RR, PFS	Objective RR showed poor correlation with OS; PFS showed moderate correlation with OS

DFS=disease free survival; OS=overall survival; PFS=progression-free survival; RCT=randomized controlled trial; RFS=recurrence free survival; RR=response rate
[^] Literature review citations are listed in Attachment 2

Other cancers

There were 37 PSDs for broad cancer areas with less than 5 submissions/resubmissions each. These were for the following cancers types: bladder; bone; brain/spine; breast and gastrointestinal; connective tissue; endometrial; gastrointestinal; head and neck; liver; neuroendocrine; pancreas; soft tissue; thyroid; and solid tumours.

In total, there were 22 drugs or 25 drug-indication pairs (i.e., 4 drugs had submitted to the PBAC for multiple indications in bowel cancer). 46% (17/37) of PSDs were for resubmissions. The PBAC had previously considered eight surrogates in these broad cancer types. The most frequently used surrogates were PFS (84%, 31/37) and objective RR (11%, 4/37), measured as a primary outcome in the clinical trials supporting the submission. 70% (26/37) of the PSDs claimed superior clinical efficacy, and 30% (11/37) claimed non-inferior clinical efficacy. Most presented a cost-utility analysis (57%, 21/37) to the nominated main comparator. In 88% (21/24) of submissions that presented a modelled economic evaluation, the surrogate outcome from the clinical evaluations was used in the model

The PBAC considered the clinical effectiveness of the treatment (relative to comparator) to be clinically significant in 25/26 (96%) of the submissions/resubmissions that claimed superiority. The main comparator across the pivotal trials supporting all the submissions/resubmissions was placebo (62%, 23/37), but same drug with different dose or schedule (14%, 5/37), active comparator (8%, 3/37), and no comparator (16%, 6/37) were also used.

Of the 37 submissions, 46% (17/37) received a positive PBAC recommendation, with 35% (13/37) not recommended and 19% (7/37) deferred. The PBAC recommended listing with an RSA on one (3%) occasion and recommended to list with managed access on one occasion (3%). Of the submissions/resubmissions with a positive PBAC recommendation, 18% (3/17) presented immature OS data. Of these, there were two that have final OS results available. Overall, for 65% (11/17) of submissions/resubmissions, the pivotal trials have since published their final OS results.

Of the 20 submissions/resubmissions that relied on a surrogate but were not recommended or were deferred by the PBAC, the main reasons were related to both clinical and cost uncertainty (40%) or cost uncertainty (40%). Further, for 77% (10/13) of submissions/resubmissions that were not recommended by the PBAC, the PBAC had documented that the OS data to support the clinical claim were immature.

Narrative review of past PBAC decisions using surrogates in other cancers
There were no new emergent themes in the remaining submissions.

Expediate literature review on the validity of surrogate endpoints used in clinical trials of therapies for other cancers
The findings of the most relevant articles measuring the association between surrogate and clinical outcomes identified in the literature review for these cancer types are summarised in Table A12 (Attachment 3). No validation studies were identified for surrogate outcomes used in thyroid, endometrial or bone cancer clinical trials. The results of the validation studies for surrogate outcomes in cancers with ≤5 submissions were similar to those of the cancers with higher numbers of submissions, in that it was generally inconclusive. There was often overlap in the RCTs included in systematic reviews and meta-analyses and sometimes these reviews reported conflicting findings. For example, in neuroendocrine cancer, 2 validation studies were found which assessed the correlation between PFS and OS using the same clinical trial data in different ways. One study concluded PFS was a suitable surrogate for OS (Imaoka et al); the other did not (Belin et al.).

Literature review

The literature search identified 59 relevant, recently published articles assessing a total of 17 surrogate endpoints in 17 different cancer types. These are summarised in Table 15.

Table 15: Surrogate endpoint validation studies by cancer type and surrogates identified from PSDs

Cancer type	Number of PSDs	Surrogate validation studies~ n[ref]	Surrogate endpoint(s)^	Validation of surrogate
Bladder	1	1 [1]	1-year survival, Overall RR, PFS	Not validated
Blood	83	10 [2-11]	EFS, PFS, MRD, VGPR	Inconclusive. Some studies assessed surrogate endpoints for PFS. EFS possible surrogate for OS in AML trials and DLBCL immuno-chemotherapy trials; PFS possible surrogate for OS in DLBCL and NHL trials; MRD may predict PFS in MM
Bone	1	0		
Bowel	7	6 [2, 9, 12-15]	DFS, Overall/ Objective RR, PFS, RFS, Time to nadir, TTP	Inconclusive
Brain/spine	4	2 [2, 16]	PFS	PFS possible surrogate for OS in glioblastoma trials
Breast	31	11 [2, 15, 17-25]	pCR, DFS, PFS, RFS, TTP	Inconclusive. Mixed results for pCR and DFS: DFS strongly correlated with OS in HER2+ trials; PFS surrogacy strength varied with setting and patient population
Endometrial	1	0		
Gastro-intestinal	5	5 [2, 9, 26-28]	DFS, EFS, PFS, TTP	Inconclusive
Head and neck	1	4 [2, 9, 29, 30]	DFS, EFS, PFS	Inconclusive
Liver	1	5 [9, 31-34]	Objective response, Overall RR, PFS, TTP,	Inconclusive
Lung	30	9 [2, 9, 10, 35-40]	CR, Objective RR, PFS, TTD	Inconclusive. PFS possible surrogate in ICI trials; PFS and DFS possible surrogates in adjuvant setting; Varied results for other NSCLC settings
Neuroendocrine	5	2 [2, 41]	PFS	Inconclusive*
Ovarian	10	7 [2, 9, 15, 42-45]	DCR, Objective RR, PFS	Inconclusive
Pancreatic	3	6 [2, 9, 46-49]	DCR, DFS, PFS,	Inconclusive
Prostate	9	2 [9, 50]	Biochemical failure, MFS, PFS	MFS possible surrogate for OS. PFS inconclusive. Biochemical failure not validated.
Renal	13	7 [2, 9, 10, 15, 51-53]	Objective RR, PFS, TTD	Inconclusive
Skin	28	7 [2, 9, 10, 54-57]	DCR, Objective/ Overall RR, PFS, RFS	Inconclusive. PFS possible surrogate for OS in chemotherapy and targeted therapy trials, and in anti-PD-1/PD-L1 trials; RFS possible in adjuvant trials
Soft tissue	1	3 [2, 15, 58]	EFS, PFS	Inconclusive. EFS had good correlation
Solid tumours	2	2 [2, 59]	PFS	Inconclusive
Thyroid	5	0		

AML=acute myeloid leukemia; CR=complete response; DCR=disease control rate; DFS=disease free survival; DLBCL=diffuse large B cell lymphoma; EFS=event free survival; HER2=human epidermal growth factor receptor 2; ICI=immune checkpoint inhibitor; MFS=metastasis free survival; MM=multiple myeloma; MRD=minimal residual disease; NHL=non-Hodgkin lymphoma; NSCLC=non-small cell lung cancer; pCR= pathological complete response; PD-1=programmed cell death protein 1; PD-L1=programmed death-ligand 1; PFS=progression free survival; RFS=recurrence free survival; RR=response rate; TTD=time to treatment discontinuation; TTP=time to progression; VGPR=very good partial response; *2 studies used different validation criteria on the same data
~ Studies validating surrogates for >1 cancer type are included in the count for each cancer type
^ References from this table are listed in Attachment 2

The articles were published between 2017 and 2022, and are systematic reviews or meta-analyses based on trial-level or individual patient-level data. Some of the articles specifically targeted surrogate endpoints used to support successful applications to the FDA (Alabaku et al., 2022; Gyawali et al., 2020; Walia et al., 2022).

Validation methods

Methods used to validate the surrogate endpoints varied between review articles. Some studies applied a validation framework to the trial data, such as the German Institute of Quality and Efficiency in Heath Care (IQWiG) guidelines, which assesses the reliability of the data, in addition to the statistical analyses, to inform conclusions on the validity of the surrogate. Other study protocols used statistical methods alone to assess the correlation between the surrogate and clinical endpoints. The thresholds accepted as sufficient for validation of a surrogate were inconsistent across the studies. In some instances, a surrogate found to be valid for a particular medication-indication pair in one meta-analysis did not meet the validation criteria of a different meta-analysis, even though data from the same clinical trial were used.

Discussion

Submissions continue to present evidence based on immature survival data. This is because surrogate endpoints are now primary outcomes in many clinical trials. The use of surrogate endpoints allows for smaller cohort sizes and shorter durations of follow-up, which reduces the cost of drug development and gives patients access to effective treatments earlier than would be the case if trials continue until sufficient events had occurred to show a significant OS benefit. For many trials, particularly those in early-stage disease, life expectancy can be many years, and the cohort is unlikely to reach median OS during follow-up. In rare cancers, the evidence may be limited to single arm studies, and even if RCTs are available, the sample sizes are often small and not powered to show statistically significant differences in treatment effect. This means the PBAC, like other HTA authorities around the world, is required to make decisions based on intermediate endpoints with no certainty that such benefits would translate into the most patient relevant outcome, increased survival. Economic models routinely use surrogate endpoints such as PFS to project long-term outcomes. In diseases where life expectancy is long, the use of surrogate endpoint data collected at early time points may add uncertainty to the economic models the PBAC relies on for cost-effectiveness assessments.

The PBAC had recommended submissions based on interim data where there is a high clinical need, and where there is sufficient evidence of efficacy based on the surrogate measure. There are advantages and disadvantages of using surrogate outcomes. The PBAC had noted the value consumers place on being progression-free, which may improve quality of life even if it may not improve life expectancy. In establishing clinical benefit for a new therapy, an advantage of PFS is that it is not confounded by the treatment effects of downstream treatments. The potential for patients to benefit from subsequent treatments however may be relevant for a funding decision and a reason to act cautiously on the use of PFS as a surrogate, as the likely magnitude of benefit on OS for the intervention is likely to be much lower than that suggested by PFS differences. Real-world benefits may be reduced costs from subsequent treatments rather than improved survival. Response

outcomes, such as complete response and objective response rate, can be measured at early timepoints in clinical trials, but they do not measure length of response like PFS does.

There are challenges relating to the consistency of evaluating surrogate measures in clinical trials. Surrogate outcome definitions and how they are measured vary across trial protocols, which makes it difficult to compare outcomes across trials and elicit a reliable estimate of treatment effect. Expert panels such as the RECIST Working Group (Eisenhauer et al., 2009) and the International Myeloma Working Group (Durie et al., 2006) have attempted to standardise response criteria, however these criteria have not been uniformly adopted in clinical trial protocols. The PBAC has had to consider the potential for bias in surrogate measurements when deliberating on the evidence presented in submissions.

The surrogate outcome most used in PBAC submissions for cancer therapies was PFS, which was used in submissions for all but one cancer type (i.e., bone cancer). Time to progression, which is similar to PFS but does not include death events, was used in 5 cancer types. Composite response outcomes, overall and/or objective response rate, were used in submissions for 10 cancer types. Specific response measures, such as clinical or complete response, were also presented in submissions for several cancer types, particularly blood cancer submissions, whereas the emerging surrogate outcomes, minimal residual disease and very good partial response, were less relied on in submissions to PBAC.

The narrative review found that saturation was reached on the themes of surrogate outcome measurement and the links between the most common surrogates, PFS and objective/overall RR, and OS after analysing the PSDs for lung, blood, skin and breast cancer submissions. The remaining cancers with <15 submissions during the study period were analysed with a focus on surrogate endpoints unique to that cancer type and any new themes. There were no new emergent themes in the remaining submissions, which predominately relied on PFS or response outcomes for their clinical claim.

Validation studies analysing the correlation between surrogate endpoints and OS have presented mixed results, with most studies showing low to moderate correlation between PFS or objective/overall RR and OS. For example, based on correlation, studies have reported PFS to be a validated surrogate for OS in diffuse large B cell lymphoma and non-Hodgkin Lymphoma but inconclusive across many other cancers (e.g., bowel, ovarian, prostate). The validation studies had used different criteria for assessing the correlation between surrogate and OS. With some reporting correlation at the patient level, but many only at the trial level for either aggregate measures or treatment effect. There was also no consistent threshold accepted as sufficient for validation of a surrogate, although some have used $R^2 \geq 0.6$, which may not be robust enough for HTA where the surrogate outcome is used to predict OS beyond the available trial data. For some meta-analyses the disease stage and type of therapy were also unclear. More validation studies of surrogate outcomes tailored to the needs of HTA using a more consistent framework is thus needed.

We used PBAC meeting documentation (primarily PSDs) as our primary evidence. This was highly relevant as a stocktake of PBAC's prior decisions relying on surrogates, but was limited as a source of information to provide an in-depth understanding of the validity of the surrogate outcomes. We found it difficult to characterise surrogate outcomes using our intended evidence framework (level of evidence, strength of association and quantification of the expected effect on the patient centred outcome) based on information published in the PSDs. Further research using more extensive meeting documentation (including full submission documentation) may be required to assess how the validation of putative surrogate endpoints were addressed by the PBAC, perhaps focusing on specific surrogates or cancer types. We also found limited use and mention of the PBAC's guidance on

surrogate outcomes (i.e., Appendix 5 of the PBAC guidelines) in the meeting documentation. This may indicate lack of focus in previous submissions on assessing the surrogate evidence or that a more simplified reporting framework is needed. This should be explored in further research.

Our research also highlighted the need for additional research to provide HTA committees the evidence to make decisions about cancer treatments whenever a surrogate is relied on in the clinical claim. For example, research in cardiovascular disease has meant that blood pressure control or blood lipid levels are now validated and widely accepted as surrogates for more final clinical outcomes in cardiovascular disease. To gather evidence, more systematic reviews and meta-regressions of evidence between surrogate and final outcome is needed (Gyawali et al., 2020; Walia et al., 2022), starting with a priority cancer type. Validation may also depend on treatment class and cancer subtypes which may add further complexity. In addition, administrative data on medication use linked with death records for patients treated before and after the listing of a new treatment could be utilised.

In cases where the PBAC had to make decisions relying on immature OS, this report highlighted the high availability of additional OS data from pivotal trials post the PBAC's decision. Another area of fruitful research would be to provide an in-depth analysis and comparison of the anticipated OS as modelled in the submissions versus actual observed survival based on final trial data for recommended submissions. Data from clinical registries capturing both surrogate and final outcomes and linking to administrative datasets (e.g., PBS, MBS and National Death Index data) may be useful additional sources of evidence in this regard, particularly as they capture the benefits and costs of downstream treatments missing in RCT data.

The PBAC may also wish to understand the relationship between surrogate and final outcomes for specific case studies, for example, the anticancer activity of newer immunotherapy agents programmed cell death protein 1 (PD-1) and programmed death-ligand 1 (PD-L1) inhibitors may be different to other cancer therapies in that they may take longer to show an effect, and due to the immune response, may also induce a 'pseudoprogression' which impacts the measurement of true disease progression. Within the included data there were submissions for 7 PD-1/PDL-1 inhibitors (nivolumab, pembrolizumab, darvulamab, avelumab, atezolizumab, cemiplimab and dostarlimab) in 9 broad cancer types (blood, lung, skin, breast, gastric, bowel, endometrial, head and neck, and renal). A comparison of final trial data with the evidence presented in the submissions would aid in closing the long-term evidence gap with this class of drug.

Cancer treatments may also have different treatment effects in different cancers. For example, pembrolizumab had relied on a surrogate outcome in submissions for 7 broad cancer types (blood, lung, skin, breast, gastric, bowel, and head and neck cancer) and was subsequently PBS listed for 6 indications (melanoma, Hodgkin Lymphoma, B-cell lymphoma, non-small cell lung cancer, urothelial cancer, colorectal cancer, and squamous cell carcinoma of the oral cavity, pharynx or larynx). Understanding its real-world efficacy may deepen our understanding on how best to estimate its effect on OS using different surrogates. Different therapies in the same drug class may also have different long-term outcomes, even if they exhibit comparable early evidence at the time of the PBAC recommendations. The epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors gefitinib, erlotinib and afatinib, were all recommended for EGFR mutation-positive non-small cell lung cancer after the July 2013 PBAC meeting. One of the reasons given for recommending these therapies was because they had been included in the clinical practice guidelines for treating EGFR mutation-positive non-small cell lung cancer. In this case linked administrative data (e.g., PBS, MBS and National Death Index) could be useful to confirm the estimated final outcomes in the submissions.

Conclusion

Surrogate outcomes are now commonly presented as evidence to support clinical claims in submissions to the PBAC. While there is some promising evidence of validation in a small number of drug classes, lines of therapy and cancer types, there is very little consensus on the reliability of surrogate outcomes such as PFS to predict OS. Further research into which surrogates are valid in which circumstances is necessary to assist HTA committees to make funding decisions/recommendations. Highlighting the need for better evidence and requiring submissions to the PBAC to provide better justification to support the use of a surrogate outcome in the clinical claim and/or modelled economic evaluation will likely assist the PBAC to make consistent decisions in this area, although the current guidance in Appendix 5 of the PBAC Guidelines may require streamlining and simplification to improve uptake.

Attachment 1

Table A1: All cancers: Descriptive statistics of cancer PSDs that relied on surrogates (N=247)

Variables	N (%)
Total cancer submissions that relied on surrogates (number of PSDs)	247 (100)
Total drugs	91 (36.8)
Drug-indication pairs	161 (65.2)
Oncology medicine class	
chemotherapy	10 (4.0)
hormone therapy	5 (2.0)
targeted therapy	194 (78.5)
immunotherapy	29 (11.7)
other	9 (3.6)
First submission	137 (55.5)
Recommendation01	
not recommended	108 (43.7)
recommended	110 (44.5)
deferred	29 (11.7)
Recommendation	
do not list	108 (43.7)
list at price	64 (25.9)
list with Criteria	30 (12.1)
list similar in class	4 (1.6)
list at lower price	12 (4.9)
defer	29 (11.2)
PBAC recommend list with RSA	33 (13.4)
PBAC recommend list with managed access	4 (1.6)
If not recommended or deferred, reason for decision:	137 (55.5)
clinical uncertainty	29 (21.2)
cost uncertainty	28 (20.4)
both	78 (56.9)
other	2 (1.3)
Appropriate nominated comparator	220 (89.1)
PBAC agreed there is unmet need for drug	116 (47.0)
Drug is for end-of-life	2 (0.8)
TGA - orphan	18 (7.3)
Drug under conditional marketing authorisation	2 (0.8)
Clinical claim for efficacy	
non-inferior	75 (30.4)
superior	162 (65.6)
non -inferior and superior	*8 (3.2)*
inferior	0
not reported or not applicable	2 (0.8)
Main comparator in pivotal trial	
no comparator	51 (20.6)
active comparator	80 (32.4)
placebo comparator	105 (42.5)
active and placebo comparators	*2 (0.8)*
same drug, different dose or schedule	9 (3.6)
PBAC consider trial comparator appropriate:	
no	70 (28.3)
yes	156 (63.2)
not reported	9 (3.6)
not sure	12 (4.9)
Submission's clinical claim rely on final clinical endpoint	149 (60.3)
Primary outcome of pivotal trial based on clinical scale	0
Surrogate	
Objective response rate	35 (14.2)
Overall response rate	22 (8.9)

Variables	N (%)
PFS	187 (75.7)
PBAC judgement about surrogate outcome's validity	
not valid	72 (29.1)
valid	68 (27.5)
possible	74 (30.0)
not mentioned	33 (13.4)
Surrogate level	
primary outcome	229 (92.7)
secondary outcome	14 (5.7)
other	2 (0.8)
not reported	2 (0.8)
OS data presented	205 (83.0)
OS maturity	
mature	29 (11.7)
immature^	143 (57.9)
not reported	46 (18.6)
not applicable	29 (11.7)
OS based on interim results	161 (65.2)
Final trial OS results available	161 (65.2)
PBAC considered effect (relative to comparator) was clinically significant	106 (42.9)
Effect size was statistically significant	97 (39.3)
Significant benefit in subgroup	93 (37.7)
Cost-effectiveness type	
cost-minimisation	66 (26.7)
cost-effectiveness (e.g. cost per responder)	12 (4.9)
cost-utility	145 (58.7)
*cost-minimisation and cost-effectiveness**	*1 (0.4)*
*cost-minimisation and cost-utility**	*7 (2.8)*
cost-consequence (e.g. cost analysis)	3 (1.2)
none presented	13 (5.3)
Surrogate in the model	140 (84.8)
OS in the model	N=165
similar to trial data presented in submission	36 (21.8)
not similar to trial data presented in submission	95 (57.6)
not applicable	3 (1.8)
not reported	30 (18.2)
OS converged in the model	N=165
not converged	27 (16.4)
converged	30 (18.2)
not applicable	3 (1.8)
not reported	104 (63.0)
PBS restriction narrower than TGA indication	177 (71.7)

Source: compiled during the review

OS=overall survival; PBAC=Pharmaceutical Benefits Advisory Committee; PBS=Pharmaceutical Benefits Scheme; PFS=progression free survival; PSD=Public Summary Document; RSA=risk sharing arrangement; TGA=Therapeutic Goods Administration.

^ Considered 'immature' if the PSD stated that the data were immature.

* The submission made multiple claims, each claim supported by a separate model.

Table A2: Lung cancer: Descriptive statistics of the lung cancer PSDs that relied on surrogates (N=30)

Variables	N (%)
Total lung cancer submissions that relied on surrogates (number of PSDs)	30 (100.0)
Total drugs	16 (53.3)
Drug-indication pairs	21 (70)
Oncology medicine class	
targeted therapy	30 (100.0)
First submission	17 (56.7)
Recommendation01	

Variables	N (%)
not recommended	8 (26.7)
recommended	16 (53.3)
deferred	6 (20.0)
Recommendation	
do not list	8 (26.7)
list at price	10 (33.3)
list with Criteria	6 (20.0)
list similar in class	0
list at lower price	0
defer	6 (20.0)
PBAC recommend list with RSA	6 (20.0)
PBAC recommend list with managed access	1 (3.3)
If not recommended or deferred, reason for decision:	N=14
clinical uncertainty	4 (28.6)
cost uncertainty	1 (7.1)
both	9 (64.3)
Appropriate nominated comparator	28 (93.3)
PBAC agreed there is unmet need for drug	9 (30.0)
Drug is for end-of-life	0
TGA - orphan	2 (6.7)
Drug under conditional marketing authorisation	1 (3.3)
Clinical claim for efficacy	
non-inferior	15 (50.0)
superior	15 (50.0)
inferior	0
Main comparator in pivotal trial	
no comparator	7 (23.3)
active comparator	17 (56.7)
placebo comparator	5 (16.7)
same drug, different dose or schedule	1 (3.3)
PBAC consider trial comparator appropriate:	
no	10 (33.3)
yes	20 (66.7)
not reported	0
not sure	0
Submission's clinical claim rely on final clinical endpoint	23 (76.7)
Primary outcome of pivotal trial based on clinical scale	0
Surrogate	30 (100.0)
Objective response rate	13 (43.3)
PFS	27 (90.0)
duration of response	1 (3.3)
disease control	1 (3.3)
PBAC judgement about surrogate outcome's validity	
not valid	2 (6.7)
valid	15 (50.0)
possible	12 (40.0)
not mentioned	1 (3.3)
Surrogate level	
primary outcome	27 (90.0)
secondary outcome	3 (10.0)
OS data presented	28 (93.3)
OS maturity	
mature	7 (23.3)
immature^	13 (43.3)
not reported	8 (26.7)
not applicable	2 (6.7)
OS based on interim results	15 (50.0)
Final trial OS results available	28 (93.3)

Variables	N (%)
PBAC considered effect (relative to comparator) was clinically significant	11 (73.3)[a]
Effect size was statistically significant	18 (60.0)
Significant benefit in subgroup	9 (30.0)
Cost-effectiveness type	
cost-minimisation	15 (50.0)
cost-effectiveness (e.g. cost per responder)	0
cost-utility	15 (50.0)
cost-consequence (e.g. cost analysis)	0
none presented	0
Surrogate in the model	15 (50.0)
OS in the model	N=15
similar to trial data presented in submission	6 (40.0)
not similar to trial data presented in submission	6 (40.0)
not reported	3 (20.0)
OS converged in the model	N=15
not converged	2 (13.3)
converged	2 (13.3)
not reported	11 (73.3)
PBS restriction narrower than TGA indication	25 (83.3)

Source: compiled during the review

OS=overall survival; PBAC=Pharmaceutical Benefits Advisory Committee; PBS=Pharmaceutical Benefits Scheme; PFS=progression free survival; PSD=Public Summary Document; RSA=risk sharing arrangement; TGA=Therapeutic Goods Administration.

^ Considered 'immature' if the PSD stated that the data were immature.

a Denominator=15 (submissions/resubmissions that claimed superiority)

Table A3: Lung cancer by recommendation: Descriptive statistics of the lung cancer PSDs that relied on surrogates (N=30) by PBAC recommendation

Lung cancer PSDs that relied on surrogates	Not recommended N=8	Recommended N=16	Deferred N=6
If not recommended or deferred, reason for decision:			
clinical uncertainty	4 (50.0%)	0	0
cost uncertainty	0	0	1 (16.7%)
both	4 (50.0%)	0	5 (83.3%)
NA	0	16 (100%)	0
OS maturity:			
mature	2 (25.0%)	4 (25.0%)	1 (16.7%)
immature^	2 (25.0%)	7 (43.8%)	4 (66.7%)
not reported	4 (50.0%)	3 (18.8%)	1 (16.7%)
not applicable	0	2 (12.5%)	0
OS based on interim results	2 (25.0%)	8 (50%)	5 (83.3%)
Final trial OS results available	8 (100%)	14 (87.5%)	6 (100%)
Final trial OS results available by OS maturity in the submission:			
mature OS	*2*	*4*	*1*
- Final trial OS results available	*2*	*4*	*1*
immature OS	*2*	*7*	*4*
- Final trial OS results available	*2*	*7*	*4*

Source: compiled during the review

OS=overall survival; PBAC=Pharmaceutical Benefits Advisory Committee; PSD=Public Summary Document

^ Considered 'immature' if the PSD stated that the data were immature.

Table A4: Blood cancer: Descriptive statistics of the blood cancer PSDs that relied on surrogates (N=83)

Variables	Leukemia N (%)	Lymphoma N (%)	Myeloma N (%)	Total N (%)
Total blood cancer submissions that relied on surrogates (number of PSDs)	29 (100.0)	35 (100.0)	19 (100.0)	83 (100.0)
Total drugs	11 (37.9)	14 (40.0)	9 (47.4)	29 (34.9)
Drug-indication pairs	18 (62.1)	22 (62.9)	12 (63.2)	50 (60.2)
Oncology medicine class				
chemotherapy	0	6 (17.1)	0	6 (7.2)
targeted therapy	27 (93.1)	26 (74.3)	11 (57.9)	64 (77.1)
immunotherapy	2 (6.9)	3 (8.6)	6 (31.6)	11 (13.3)
other	0	0	2 (10.5)	2 (2.4)
First submission	15 (51.7)	20 (57.1)	11 (57.9)	46 (55.4)
Recommendation01				
not recommended	9 (31.0)	17 (48.6)	13 (68.4)	39 (47.0)
recommended	15 (51.7)	14 (40.0)	6 (31.6)	35 (42.2)
deferred	5 (17.2)	4 (11.4)	0	9 (10.8)
Recommendation				
do not list	9 (31.0)	17 (48.6)	13 (68.4)	39 (47.0)
list at price	6 (20.7)	8 (22.9)	5 (26.3)	19 (22.9)
list with Criteria	8 (27.6)	5 (14.3)	1 (5.3)	14 (16.9)
list similar in class	1 (3.4)	0	0	1 (1.2)
list at lower price	0	1 (2.9)	0	1 (1.2)
defer	5 (17.2)	4 (11.4)	0	9 (10.8)
PBAC recommend list with RSA	7 (24.1)	6 (17.1)	1 (5.3)	14 (16.9)
PBAC recommend list with managed access	1 (3.4)	0	0	1 (1.2)
If not recommended or deferred, reason for decision:	14 (48.3)	21 (60.0)	13 (68.4)	48 (57.8)
clinical uncertainty	0	0	3 (23.1)	3 (6.3)
cost uncertainty	6 (42.9)	5 (23.8)	4 (30.8)	15 (31.3)
both	8 (57.1)	16 (76.2)	6 (46.2)	30 (62.5)
Appropriate nominated comparator	23 (79.3)	33 (94.3)	14 (73.7)	70 (84.3)
PBAC agreed there is unmet need for drug	10 (34.5)	24 (68.6)	6 (31.6)	40 (48.2)
Drug is for end-of-life	0	1 (2.9)	1 (5.3)	2 (2.4)
TGA - orphan	2 (6.9)	1 (2.9)	0	3 (3.6)
Drug under conditional marketing authorisation	0	0	0	0
Clinical claim for efficacy				
non-inferior	7 (24.1)	3 (8.6)	9 (47.4)	19 (22.9)
superior	20 (70.0)	32 (91.4)	7 (36.8)	59 (71.1)
non-inferior and superior	*1 (3.4)*	*0*	*3 (15.8)*	*4 (4.8)*
inferior	0	0	0	0
not applicable	1 (3.4%)	0	0	1 (1.2%)
Main comparator in pivotal trial				
no comparator	10 (34.5)	18 (51.4)	0	28 (33.7)
active comparator	15 (51.7)	15 (42.9)	3 (15.8)	33 (39.8)
placebo comparator	4 (13.8)	2 (5.7)	13 (68.4)	19 (22.9)
active and placebo comparators	*0*	*0*	*2 (10.5)*	*2 (2.4)*
same drug, different dose or schedule	0	0	1 (5.3)	1 (1.2)
PBAC consider trial comparator appropriate:				
no	13 (44.8)	13 (37.1)	3 (15.8)	29 (34.9)
yes	13 (44.8)	15 (42.9)	14 (73.7)	42 (50.6)
not reported	0	2 (5.7)	1 (5.3)	3 (3.6)
not sure	3 (10.3)	5 (14.3)	1 (5.3)	9 (10.8)
Submission's clinical claim rely on final clinical endpoint	19 (65.5)	28 (80.0)	16 (84.2)	63 (75.9)
Primary outcome of pivotal trial based on clinical scale	0	0	0	0
Surrogate				
complete remission	4 (13.8%)	0	0	4 (4.8%)
clinical/complete response	2 (6.9%)	7 (20.0%)	1 (5.3%)	10 (12.0%)
cytogenic response	2 (6.9%)	0	0	2 (2.4%)
DOR	0	3 (8.6%)	0	3 (3.6%)

Variables	Leukemia N (%)	Lymphoma N (%)	Myeloma N (%)	Total N (%)
EFS	2 (6.9%)	0	0	2 (2.4%)
incidence of cancer	1 (3.4%)	0	0	1 (1.2%)
MRD	1 (3.4%)	0	0	1 (1.2%)
objective response rate	1 (3.4%)	9 (25.7%)	0	10 (12.0%)
overall response rate	2 (6.9%)	9 (25.7%)	3 (15.8%)	14 (16.9%)
PFS	17 (58.6%)	29 (82.9%)	17 (89.5%)	63 (75.9%)
RFS	3 (10.3%)	0	0	3 (3.5%)
TTNT	1 (3.4%)	0	0	1 (1.2%)
TTP	0	1 (2.9%)	0	1 (1.2%)
VGPR	0	2 (5.7%)	2 (10.5%)	4 (4.8%)
PBAC judgement about surrogate outcome's validity				
not valid	3 (10.3)	14 (40.0)	1 (5.3)	18 (21.7)
valid	5 (17.2)	1 (2.9)	4 (21.1)	10 (12.0)
possible	13 (44.8)	16 (45.7)	9 (47.4)	38 (45.8)
not mentioned	8 (27.6)	4 (11.4)	5 (26.3)	17 (20.5)
Surrogate level				
primary outcome	26 (89.7)	32 (91.4)	17 (89.5)	75 (90.4)
secondary outcome	3 (10.3)	3 (8.6)	1 (5.3)	7 (8.4)
other	0	0	1 (5.3)	1 (1.2)
OS data presented	21 (72.4)	29 (82.9)	17 (89.5)	67 (80.7)
OS maturity				
mature	2 (6.9)	4 (11.4)	3 (15.8)	9 (10.8)
immature^	19 (65.5)	21 (60.0)	9 (47.4)	49 (59.0)
not reported	5 (17.2)	7 (20.0)	6 (31.6)	18 (21.7)
not applicable	3 (10.3)	3 (8.6)	1 (5.3)	7 (8.4)
OS based on interim results	20 (69.0)	26 (74.3)	15 (78.9)	61 (73.5)
Final trial OS results available	18 (62.1)	23 (65.7)	13 (68.4)	54 (65.1)
PBAC considered effect (relative to comparator) was clinically significant[a]	19 (90.5)	15 (46.9)	6 (60.0)	40 (63.5)
Effect size was statistically significant	16 (55.2)	12 (34.3)	3 (15.8)	31 (37.3)
Significant benefit in subgroup	14 (48.3)	19 (54.3)	8 (42.1)	41 (49.4)
Cost-effectiveness type				
cost-minimisation	6 (20.7)	3 (8.6)	8 (42.1)	17 (20.5)
cost-effectiveness (e.g. cost per responder)	3 (10.3)	4 (11.4)	0	7 (8.4)
cost-utility	17 (58.6)	28 (80.0)	7 (36.8)	52 (62.7)
cost-consequence (e.g. cost analysis)	0	0	0	0
cost-minimisation and cost-effectiveness	0	0	1 (5.3)	1 (1.2)
cost-minimisation and cost-utility	1 (3.4)	0	2 (10.5)	3 (3.6)
none presented*	2 (6.9)	0	1 (5.3)	3 (3.6)
Surrogate in the model	14 (48.3)	25 (71.4)	7 (36.8)	46 (55.4)
OS in the model	N=21	N=32	N=10	N=63
similar to trial data presented in submission	2 (9.5)	5 (15.6)	3 (30.0)	10 (15.9)
not similar to trial data presented in submission	15 (71.4)	20 (62.5)	4 (40.0)	39 (61.9)
not applicable	0	3 (9.4)	0	3 (4.8)
not reported	4 (19.0)	4 (12.5)	3 (30.0)	11 (17.5)
OS converged in the model	N=21	N=32	N=10	N=63
not converged	3 (14.3)	8 (25.0)	5 (50.0)	16 (25.4)
converged	9 (42.9)	7 (21.9)	0	16 (25.4)
not applicable	0	3 (9.4)	0	3 (4.8)
not reported	9 (42.9)	14 (43.8)	5 (50.0)	28 (44.4)
PBS restriction narrower than TGA indication	26 (89.7)	17 (48.6)	13 (68.4)	56 (67.5)

Source: compiled during the review

DOR=duration of response; EFS=event free survival; MRD=minimal residual disease; OS=overall survival; PBAC=Pharmaceutical Benefits Advisory Committee; PBS=Pharmaceutical Benefits Scheme; PFS=progression free survival; PSD=Public Summary Document; RFS=relapse/recurrence free survival; RSA=risk sharing arrangement; TGA=Therapeutic Goods Administration; TTNT=time to next treatment; VGPR=very good partial response.

^ Considered 'immature' if the PSD stated that the data were immature.

* The resubmission did not present the economic evaluation, however the cost-utility analysis was presented in the previous submission

a Denominator is the number of submissions/resubmissions that made a claim of superiority

Table A5: Blood cancer by PBAC recommendation: Descriptive statistics of the blood cancer PSDs that relied on surrogates (N=83) by PBAC recommendation

Blood cancer PSDs that relied on surrogates	Not recommended N=39	Recommended N=35	Deferred N=9
If not recommended or deferred, reason for decision:			
clinical uncertainty	2 (5.1%)	0	0
cost uncertainty	9 (23.1%)	0	6 (66.7%)
both	27 (69.2%)	0	3 (33.3%)
NA	0	35 (100%)	0
OS maturity:			
mature	4 (10.3%)	5 (14.3%)	0
immature^	23 (59.0%)	19 (54.3%)	7 (77.8%)
not reported	10 (25.6%)	8 (22.9%)	0
not applicable	2 (5.1%)	3 (8.6%)	2 (22.2%)
OS based on interim results	30 (76.9%)	24 (68.6%)	7 (77.8%)
Final trial OS results available	26 (66.7%)	21 (60.0%)	7 (77.8%)
Final trial OS results available by OS maturity in the submission:			
mature OS	*4*	*5*	*0*
- Final trial OS results available	*4*	*5*	*0*
immature OS	*23*	*19*	*7*
- Final trial OS results available	*13*	*10*	*7*

Source: compiled during the review

OS=overall survival; PBAC=Pharmaceutical Benefits Advisory Committee; PSD=Public Summary Document

^ Considered 'immature' if the PSD stated that the data were immature.

Table A6: Skin cancer: Descriptive statistics of the skin cancer PSDs that relied on surrogates (N=28)

Variables	N (%)
Total skin cancer submissions that relied on surrogates (number of PSDs)	28 (100)
Total drugs	12 (42.9)
Drug-indication pairs	16 (57.1)
Oncology medicine class	
targeted therapy	24 (85.6)
immunotherapy	1 (3.6)
other	3 (10.7)
First submission	14 (50.0)
Recommendation01	
not recommended	11 (39.3)
recommended	14 (50.0)
deferred	3 (10.7)
Recommendation	
do not list	11 (39.3)
list at price	10 (35.7)
list with Criteria	2 (7.1)
list similar in class	0
list at lower price	2 (7.1)
defer	3 (10.7)
PBAC recommend list with RSA	4 (14.3)
PBAC recommend list with managed access	1 (3.6)
If not recommended or deferred, reason for decision:	14 (50.0)
clinical uncertainty	2 (14.3)
cost uncertainty	1 (7.1)
both	10 (71.4)
other	1 (7.1)a
Appropriate nominated comparator	26 (92.9)

Variables	N (%)
PBAC agreed there is unmet need for drug	16 (57.1)
Drug is for end-of-life	0
TGA - orphan	0
Drug under conditional marketing authorisation	1 (3.6)
Clinical claim for efficacy	
non-inferior	10 (35.7)
superior	17 (60.7)
non-inferior and superior	*1 (3.6)[b]*
inferior	0
Main comparator in pivotal trial	
no comparator	6 (21.4)
active comparator	11 (39.3)
placebo comparator	10 (35.7)
same drug, different dose or schedule	1 (3.6)
PBAC consider trial comparator appropriate:	
no	5 (17.9)
yes	16 (57.1)
not reported	4 (14.3)
not sure	3 (10.7)
Submission's clinical claim rely on final clinical endpoint	22 (78.6)
Primary outcome of pivotal trial based on clinical scale	0
Surrogate	
clearance of solar keratosis	3 (10.7)
distant metastases free survival	3 (10.7)
durable response rate	1 (3.6)
objective response rate	6 (21.4)
overall response rate	2 (7.1)
PFS	14 (50.0)
RFS	7 (25.0)
PBAC judgement about surrogate outcome's validity	
not valid	13 (46.4)
valid	1 (3.6)
possible	3 (10.7)
not mentioned	11 (39.3)
Surrogate level	
primary outcome	26 (92.9)
not reported	2 (7.1)
OS data presented	17 (60.7)
OS maturity	
mature	0
immature[c]	19 (67.9)
not reported	2 (7.1)
not applicable	7 (25.0)
OS based on interim results	18 (64.3)
Final trial OS results available[d]	6 (21.4)
PBAC considered effect (relative to comparator) was clinically significant	10 (55.6)^
Effect size was statistically significant	6 (21.4)
Significant benefit in subgroup	1 (3.6)
Cost-effectiveness type	
cost-minimisation	8 (28.6)
cost-effectiveness (e.g. cost per responder)	2 (7.1)[e]
cost-utility	16 (57.1)
cost-consequence (e.g. cost analysis)	1 (3.6)
cost-minimisation and cost-utility	*1 (3.6)[b]*
none presented	0
Surrogate in the model	15 (83.3)[f]
OS in the model	N=18
similar to trial data presented in submission	0

Variables	N (%)
not similar to trial data presented in submission	13 (46.4)[f]
not reported	5 (27.8)[f]
OS converged in the model	N=18
not converged	1 (5.6)[f]
converged	6 (33.3)[f]
not reported	11 (61.1)[f]
PBS restriction narrower than TGA indication	16 (57.1)

Source: compiled during the review

OS=overall survival; PBAC=Pharmaceutical Benefits Advisory Committee; PBS=Pharmaceutical Benefits Scheme; PFS=progression free survival; PSD=Public Summary Document; RFS=relapse/recurrence free survival; RSA=risk sharing arrangement; TGA=Therapeutic Goods Administration

a The PBAC deferred a recommendation of a drug noting that it was unable to proceed until the listing of another drug was agreed.

b The resubmission made two claims, each in a different population: one claim of non-inferiority (supported by a cost-minimisation analysis) and one claim of superiority (supported by a cost-utility analysis)

c Considered 'immature' if the PSD stated that the data were immature.

d Final trial OS results available for all trials relied upon in the submission

e Two cost-effectiveness analyses presented for non-inferiority claims

f Denominator=18 (cost-effectiveness and cost-utility analyses). The resubmission that presented the cost-minimisation and cost-utility analyses did not include information about the model.

^ Denominator=18 (submissions/resubmissions that made a claim or superiority)

Table A7: Skin cancer by PBAC recommendation: Descriptive statistics of the skin cancer PSDs that relied on surrogates (N=28) by PBAC recommendation

Skin cancer PSDs that relied on surrogates	Not recommended N=11	Recommended N=14	Deferred N=3
If not recommended or deferred, reason for decision:			
clinical uncertainty	1 (9.1%)	0	1 (33.3%)
cost uncertainty	0	0	1 (33.3%)
both	10 (90.9%)	0	0
other	0	0	1 (33.3%)
NA	0	14 (100%)	0
OS maturity:			
mature	0	0	0
immature^	6 (54.5%)	10 (71.4%)	3 (100%)
not reported	1 (9.1%)	1 (7.1%)	0
not applicable	4 (36.4%)	3 (21.4%)	0
OS based on interim results	6 (54.5%)	9 (64.3%)	3 (100%)
Final trial OS results available			
Final trial OS results available by OS maturity in the submission:			
mature OS	*0*	*0*	*0*
- Final trial OS results available	*0*	*0*	*0*
immature OS	*6*	*10*	*3*
- Final trial OS results available	*0*	*4*	*1*

Source: compiled during the review

OS=overall survival; PBAC=Pharmaceutical Benefits Advisory Committee; PSD=Public Summary Document

^ Considered 'immature' if the PSD stated that the data were immature.

Table A8: Breast cancer: Descriptive statistics of the breast cancer PSDs that relied on surrogates (N=31)

Variables	N (%)
Total breast cancer submissions that relied on surrogates (number of PSDs)	31 (100)
Total drugs	15 (48.4)
Drug-indication pairs	22 (71.0)
Oncology medicine class	
targeted therapy	28 (90.3)
hormone therapy	2 (6.5)

Variables	N (%)
chemotherapy	1 (3.2)
First submission	20 (64.5)
Recommendation01	
not recommended	17 (54.8)
recommended	12 (38.7)
deferred	2 (6.5)
Recommendation	
do not list	17 (54.8)
list at price	6 (19.4)
list with Criteria	2 (6.5)
list similar in class	0
list at lower price	4 (12.9)
defer	2 (6.5)
PBAC recommend list with RSA	4 (12.9)
PBAC recommend list with managed access	0
If not recommended or deferred, reason for decision:	19 (61.3)
clinical uncertainty	3 (15.8)
cost uncertainty	0
both	16 (84.2)
Appropriate nominated comparator	26 (83.9)
PBAC agreed there is unmet need for drug	12 (38.7)
Drug is for end-of-life	0
TGA - orphan	1 (3.2)
Drug under conditional marketing authorisation	0
Clinical claim for efficacy	
non-inferior	5 (16.1)
superior	23 (74.2)
non-inferior and superior	2 (6.5)[b]
inferior	0
not reported	1 (3.2)
Main comparator in pivotal trial	
no comparator	2 (6.5)
active comparator	4 (12.9)
placebo comparator	24 (77.4)
same drug, different dose or schedule	1 (3.2)
PBAC consider trial comparator appropriate:	
no	6 (19.4)
yes	23 (74.2)
not reported	2 (6.5)
not sure	0
Submission's clinical claim rely on final clinical endpoint	26 (83.9)
Primary outcome of pivotal trial based on clinical scale	0
Surrogate	
Pathological complete response	1 (3.2)
DFS/EFS^	5 (16.1)
Invasive disease-free survival	6 (19.4)
Incidence of breast cancer	1 (3.2)
Overall response rate	2 (6.5)
PFS	21 (67.7)
Premature ovarian failure	1 (3.2)
Time to progression	1 (3.2)
PBAC judgement about surrogate outcome's validity	
not valid	13 (41.9)
valid	7 (22.6)
possible	8 (25.8)
not mentioned	3 (9.7)
Surrogate level	
primary outcome	30 (96.8)

Variables	N (%)
other	1 (3.2)
OS data presented	25 (80.6)
OS maturity	
mature	0
immature[a]	22 (71.0)
not reported	3 (9.7)
not applicable	6 (19.4)
OS based on interim results	23 (74.2)
Final trial OS results available	19 (61.3)
PBAC considered effect (relative to comparator) was clinically significant	12 (48.0)[c]
Effect size was statistically significant	14 (45.2)
Significant benefit in subgroup	11 (35.5)
Cost-effectiveness type	
cost-minimisation	5 (16.1)
cost-effectiveness (e.g. cost per responder)	0
cost-utility	22 (71.0)
cost-consequence (e.g. cost analysis)	0
cost-minimisation and cost-utility	*2 (6.5)[b]*
none presented	2 (6.5)
Surrogate in the model	24 (100)[d]
OS in the model	N=24
similar to trial data presented in submission	4 (16.7)
not similar to trial data presented in submission	15 (62.5)
not reported	5 (20.8)
OS converged in the model	N=24
not converged	1 (4.2)
converged	2 (8.3)
not reported	21 (87.5)
PBS restriction narrower than TGA indication	18 (58.1)

Source: compiled during the review

DFS=disease-free survival; EFS=event-free survival; OS=overall survival; PBAC=Pharmaceutical Benefits Advisory Committee; PBS=Pharmaceutical Benefits Scheme; PFS=progression free survival; PSD=Public Summary Document; RSA=risk sharing arrangement; TGA=Therapeutic Goods

^ Includes distant disease-free survival, disease-free survival including ductal carcinoma in situ), and event-free/disease-free survival

[a] Considered 'immature' if the PSD stated that the data were immature.

[b] The resubmission made two claims, each in a different population: one claim of non-inferiority (supported by a cost-minimisation analysis) and one claim of superiority (supported by a cost-utility analysis)

[c] Denominator=25 (submissions/resubmissions that claimed superiority)

[d] Denominator=24 (cost-utility analysis and the submission that presented cost-utility analysis with a cost-minimisation). One PSD with a superior claim did not present a model.

Table A9: Breast cancer by PBAC recommendation: Descriptive statistics of the breast cancer PSDs that relied on surrogates (N=31) by PBAC recommendation

Breast cancer PSDs that relied on surrogates	Not recommended N=17	Recommended N=12	Deferred N=2
If not recommended or deferred, reason for decision:			
clinical uncertainty	2 (11.8%)	0	1 (50%)
cost uncertainty	0	0	0
both	15 (88.2%)	0	1 (50%)
NA	0	12 (100%)	0
OS maturity:			
mature	0	0	0
immature^	12 (70.6%)	9 (75.0%)	1 (50%)
not reported	2 (11.8%)	1 (8.3%)	0
not applicable	3 (17.6%)	2 (16.7%)	1 (50%)
OS based on interim results	12 (70.6%)	10 (83.3%)	1 (50%)
Final trial OS results available	10 (58.8%)	7 (58.3%)	2 (100%)
Final trial OS results available by OS maturity in the submission:			
mature OS	*0*	*0*	*0*
- Final trial OS results available	*0*	*0*	*0*
immature OS	*12*	*9*	*1*
- Final trial OS results available	*9*	*6*	*1*

Source: compiled during the review

OS=overall survival; PBAC=Pharmaceutical Benefits Advisory Committee; PSD=Public Summary Document

^ Considered 'immature' if the PSD stated that the data were immature.

Table A10: Other cancers: Descriptive statistics of broad cancer types with less than 15 PSDs that relied on surrogates (N=75)

Variables	Renal N (%)	Ovarian N (%)	Prostate N (%)	Bowel N (%)	Other N (%)*
Total submissions that relied on surrogates (number of PSDs)	13 (100)	10 (100)	9 (100)	6 (100)	37 (100)
Total drugs	7 (53.8)	3 (30)	5 (55.6)	3 (50)	22 (59.5)
Drug-indication pairs	7 (53.8)	8 (80)	7 (77.8)	5 (83.3)	25 (67.6)
Oncology medicine class					
chemotherapy	0	0	0	0	3 (8.1)
hormone therapy	0	0	2 (22.2)	0	1 (2.7)
targeted therapy	13 (100)	10 (100)	7 (77.8)	6 (100)	33 (89.2)
immunotherapy	0	0	0	0	0
other	0	0	0	0	0
First submission	5 (38.5)	7 (70)	6 (66.7)	2 (33.3)	20 (54.1)
Recommendation01					
not recommended	7 (53.8)	5 (50.0)	5 (55.6)	3 (50.0)	13 (35.1)
recommended	6 (46.2)	4 (40.0)	3 (33.3)	3 (50.0)	17 (45.9)
deferred	0	1 (10.0)	1 (11.1)	0	7 (18.9)
Recommendation					
do not list	7 (53.8)	5 (50.0)	5 (55.6)	3 (50.0)	13 (35.1)
list at price	1 (7.7)	1 (10.0)	2 (22.2)	1 (16.7)	14 (37.8)
list with Criteria	3 (23.1)	1 (10.0)	0	1 (16.7)	1 (2.7)
list similar in class	1 (7.7)	0	1 (11.1)	0	1 (2.7)
list at lower price	1 (7.7)	2 (20.0)	0	1 (16.7)	1 (2.7)
defer	0	1 (10.0)	1 (11.1)	0	7 (18.9)
PBAC recommend list with RSA	3 (23.1)	0	0	1 (16.7)	1 (2.7)
PBAC recommend list with managed access	0	0	0	0	1 (2.7)
If not recommended or deferred, reason for decision:	7 (53.8)	6 (60.0)	6 (66.7)	3 (50.0)	20 (54.1)
clinical uncertainty	5 (71.4)	3 (50.0)	4 (66.7)	2 (66.7)	3 (15.0)
cost uncertainty	0	2 (33.3)	1 (16.7)	0	8 (40.0)

Variables	Renal N (%)	Ovarian N (%)	Prostate N (%)	Bowel N (%)	Other N (%)*
both	2 (28.6)	1 (16.7)	1 (16.7)	1 (33.3)	8 (40.0)
other	0	0	0	0	1 (5.0)[a]
Appropriate nominated comparator	11 (84.6)	10 (100)	9 (100)	5 (83.3)	35 (94.6)
PBAC agreed there is unmet need for drug	6 (46.2)	6 (60.0)	2 (22.2)	2 (33.3)	23 (62.2)
Drug is for end-of-life	0	0	0	0	0
TGA - orphan	1 (7.7)	1 (10.0%)	0	0	10 (27.0)
Drug under conditional marketing authorisation	0	0	0	0	0
Clinical claim for efficacy					
non-inferior	5 (38.5)	3 (30.0)	4 (44.4)	3 (50.0)	11 (29.7)
superior	8 (61.5)	6 (60.0)	5 (55.6)	3 (50.0)	26 (70.3)
non-inferior and superior	*0*	*1 (10%)*	*0*	*0*	*0*
inferior	0	0	0	0	0
Main comparator in pivotal trial					
no comparator	0	0	1 (11.1)	1 (16.7)	6 (16.2)
active comparator	8 (61.5)	1 (10.0)	2 (22.2)	1 (16.7)	3 (8.1)
placebo comparator	5 (38.5)	9 (90.0)	6 (66.7)	4 (66.7)	23 (62.2)
same drug, different dose or schedule	0	0	0	0	5 (13.5)
PBAC consider trial comparator appropriate:					
no	4 (30.8)	2 (20.0)	4 (44.4)	4 (66.7)	6 (16.2)
yes	9 (69.2)	8 (80.0)	5 (55.6)	2 (33.3)	31 (83.8)
not reported	0	0	0	0	0
not sure	0	0	0	0	0
Submission's clinical claim rely on final clinical endpoint	3 (23.1)	0	0	1 (16.7)	11 (29.7)
Primary outcome of pivotal trial based on clinical scale	0	0	0	0	0
Surrogate					
Pathological complete response	0	0	0	0	3 (8.1)
Metastasis-free survival	0	0	4 (44.4)	0	0
Objective response rate	1 (7.7)	0	0	1 (16.7)	4 (10.8)
Overall response rate	0	0	0	0	3 (8.1)
PFS	13 (100)	10 (100)	2 (22.2)	6 (100)	31 (83.8)
Response[b]	0	0	0	0	3 (8.1)
Serum testosterone	0	0	3 (33.3)	0	0
Symptom progression	0	0	2 (22.2)	0	0
Time to next treatment	0	0	1 (11.1)	0	0
Time to progression	0	0	1 (11.1)	0	4 (10.8)
PBAC judgement about surrogate outcome's validity					
not valid	8 (65.1)	2 (20.0)	2 (22.2)	1 (16.7)	13 (35.1)
valid	4 (30.8)	6 (60.0)	5 (55.6)	5 (83.3)	15 (40.5)
possible	1 (7.7)	2 (20.0)	2 (22.2)	0	8 (21.6)
not mentioned	0	0	0	0	1 (2.7)
Surrogate level					
primary outcome	12 (92.3)	10 (100)	7 (77.8)	6 (100)	36 (97.3)
secondary outcome	1 (7.7)	0	2 (22.2)	0	1 (2.7)
OS data presented	13 (100)	10 (100)	6 (66.7)	6 (100)	33 (89.2)
OS maturity					
mature	2 (15.4)	1 (10.0)	3 (33.3)	1 (16.7)	6 (16.2)
immature^	7 (53.8)	9 (90.0)	3 (33.3)	4 (66.7)	17 (45.9)
not reported	4 (30.8%)	0	0	1 (16.7)	10 (27.0)
not applicable	0	0	3 (33.3)	0	4 (10.8)
OS based on interim results	9 (69.2)	9 (90.0)	3 (33.3)	4 (66.7)	19 (51.4)
Final trial OS results available	12 (92.3)	4 (40.0)	7 (77.8)	6 (100)	25 (67.6)
PBAC considered effect (relative to comparator) was clinically significant[c]	1 (12.5)	3 (42.8)	3 (60.0)	1 (20.0)	25 (96.2)
Effect size was statistically significant	1 (7.7)	6 (60.0)	3 (33.3)	2 (33.3)	16 (43.2)
Significant benefit in subgroup	9 (69.2)	8 (80.0)	0	3 (50)	11 (29.7)

Variables	Renal N (%)	Ovarian N (%)	Prostate N (%)	Bowel N (%)	Other N (%)*
Cost-effectiveness type					
cost-minimisation	5 (38.5)	2 (20.0)	4 (44.4)	3 (50.0)	7 (18.9)
cost-effectiveness (e.g. cost per responder)	0	0	0	0	3 (8.1)
cost-utility	6 (46.2)	6 (60.0)	5 (55.6)	2 (33.3)	21 (56.8)
cost-consequence (e.g. cost analysis)	2 (15.4)[d]	0	0	0	0
cost-minimisation and cost-utility	0	*1 (10.0)*	0	0	0
none presented	0	1 (10.0)[g]	0	1 (16.7)[e]	6 (16.2)[f]
Surrogate in the model[j]	6 (100)	7 (100)	4 (80.0)	2 (100%)	21 (87.5)
OS in the model	N=6	N=7	N=5	N=2	N=24
similar to trial data presented in submission	6 (100)	4 (57.1)	2 (40.0)	0	4 (16.7)
not similar to trial data presented in submission	0	3 (42.9)	2 (40.0)	2 (100)	15 (62.5)
not reported	0	0	1 (20.0)	0	5 (20.8)
OS converged in the model	N=6	N=7	N=5	N=2	N=24
not converged	1 (16.7)	3 (42.9)	0	0	3 (12.5)
converged	1 (16.7)	1 (14.3)	0	0	2 (8.3)
not reported	4 (66.7)	3 (42.9)	5 (100)	2 (100)	19 (79.2)
PBS restriction narrower than TGA indication	13 (100)	8 (80.0)	8 (88.9)	5 (83.3)	28 (75.7)

Source: compiled during the review

OS=overall survival; PBAC=Pharmaceutical Benefits Advisory Committee; PBS=Pharmaceutical Benefits Scheme; PFS=progression free survival; PSD=Public Summary Document; RSA=risk sharing arrangement; TGA=Therapeutic Goods Administration

* Includes broad cancer areas with less than 5 submissions/resubmissions, namely: bladder, bone, brain/spine, breast and gastrointestinal, connective tissue, endometrial, gastrointestinal, head and neck, liver, neuroendocrine, pancreas, soft tissue, solid tumours and thyroid.

^ Considered 'immature' if the PSD stated that the data were immature.

[a] Pending MSAC advice on codependent testing

[b] Includes pathological complete response, pathological complete response of the breast only, SEGA response rate, tumour response

[c] Denominator is the number of submissions/resubmissions that claimed clinical superiority.

[d] Two superior claims presented cost-consequence analyses

[e] The resubmission did not present the economic evaluation, however the cost-utility analysis was presented in the previous submission

[f] Two resubmissions with economic evaluation presented in previous submissions, four submission of biosimilar brands for listed indications.

[g] Minor submission requesting a change to the listed patient population, no economic analyses were presented as the submission assumed that the cost-effectiveness of treating the requested population was similar to the listed population

[j] Denominator=24 (cost-utility analysis and the submission that presented cost-utility analysis with a cost-minimisation). One PSD with a superior claim did not present a model.

Table A11: Other cancers by PBAC recommendation: Descriptive statistics of broad cancer types with less than 15 PSDs that relied on surrogates (N=75) by PBAC recommendation

Other cancer PSDs that relied on surrogates*	Not recommended N=33	Recommended N=33	Deferred N=9
If not recommended or deferred, reason for decision:			
clinical uncertainty	17 (51.5%)	0	0
cost uncertainty	5 (15.2%)	0	6 (66.7%)
both	11 (33.3%)	0	2 (22.2%)
NA	0	33 (100%)	1 (11.1%)#
OS maturity:			
mature	6 (18.2%)	6 (18.2%)	1 (11.1%)
immature^	23 (69.7%)	12 (36.4%)	5 (55.6%)
not reported	2 (6.1%)	11 (33.3%)	2 (22.2%)
not applicable	2 (6.1%)	4 (12.1%)	1 (11.1%)
OS based on interim results	23 (69.7%)	15 (45.5%)	6 (66.7%)
Final trial OS results available	28 (84.8%)	21 (63.6%)	7 (77.8%)
Final trial OS results available by OS maturity in the submission:			
mature OS	*6*	*6*	*1*
- Final trial OS results available	*6*	*6*	*1*
immature OS	*23*	*12*	*5*
- Final trial OS results available	*18*	*6*	*3*

Source: compiled during the review

OS=overall survival; PBAC=Pharmaceutical Benefits Advisory Committee; PSD=Public Summary Document

* Includes broad cancer areas with less than 15 submissions/resubmissions, namely: bladder, bone, bowel, brain/spine, breast and gastrointestinal, connective tissue, endometrial, gastrointestinal, head and neck, liver, neuroendocrine, ovarian, pancreas, prostate, renal, soft tissue, solid tumours and thyroid.

^ Considered 'immature' if the PSD stated that the data were immature.

One submission was deferred pending MSAC advice on codependent testing.

Attachment 2

Surrogate validation studies from Table 10

1.   Abdel-Rahman, O., *Surrogate end points for overall survival in trials of PD-(L)1 inhibitors for urinary cancers: a systematic review.* Immunotherapy, 2018. **10**(2): p. 139-148.
2.   Belin, L., et al., *Progression-free survival as a surrogate for overall survival in oncology trials: a methodological systematic review.* British Journal of Cancer, 2020. **122**(11): p. 1707-1714.
3.   Assouline, S., et al., *Validity of event-free survival as a surrogate endpoint in haematological malignancy: Review of the literature and health technology assessments.* Critical Reviews in Oncology/Hematology, 2022. **175**: p. 103711.
4.   Daniele, P., et al., *Response rates and minimal residual disease outcomes as potential surrogates for progression-free survival in newly diagnosed multiple myeloma.* PLoS ONE, 2022. **17(5 May) (no pagination)**.
5.   Mangal, N., et al., *Relationship between response rates and median progression-free survival in non-Hodgkin's lymphoma: A meta-analysis of published clinical trials.* Hematological Oncology, 2018. **36(1)**: p. 37-43.
6.   Mangal, N., et al., *Use of depth of response to predict progression-free survival in relapsed or refractory multiple myeloma: Evaluation of results from 102 clinical trials.* Hematological Oncology, 2018. **36(3)**: p. 547-553.
7.   Estey, E., et al., *New drug approvals in acute myeloid leukemia: what's the best end point?* Leukemia, 2016. **30**(3): p. 521-525.
8.   Avet-Loiseau, H., et al., *Minimal Residual Disease Status as a Surrogate Endpoint for Progression-free Survival in Newly Diagnosed Multiple Myeloma Studies: A Meta-analysis.* Clinical Lymphoma, Myeloma and Leukemia, 2020. **20(1)**: p. e30-e37.
9.   Haslam, A., et al., *A systematic review of trial-level meta-analyses measuring the strength of association between surrogate end-points and overall survival in oncology.* Eur J Cancer, 2019. **106**: p. 196-211.
10.   Walia, A., A. Haslam, and V. Prasad, *FDA validation of surrogate endpoints in oncology: 2005-2022.* J Cancer Policy, 2022. **34**: p. 100364.
11.   Munshi, N.C., et al., *A large meta-analysis establishes the role of MRD negativity in long-term survival outcomes in patients with multiple myeloma.* Blood Advances, 2020. **4**(23): p. 5988-5999.
12.   Burzykowski, T., et al., *Evaluation of Continuous Tumor-Size-Based End Points as Surrogates for Overall Survival in Randomized Clinical Trials in Metastatic Colorectal Cancer.* JAMA Network Open, 2019. **(no pagination)**.
13.   Ecker, B.L., et al., *Recurrence-free survival versus overall survival as a primary endpoint for studies of resected colorectal liver metastasis: a retrospective study and meta-analysis.* Lancet Oncol, 2022. **23**(10): p. 1332-1342.
14.   Cremolini, C., et al., *Surrogate Endpoints in Second-Line Trials of Targeted Agents in Metastatic Colorectal Cancer: A Literature-Based Systematic Review and Meta-Analysis.* Cancer Res Treat, 2017. **49**(3): p. 834-845.
15.   Savina, M., et al., *Meta-analyses evaluating surrogate endpoints for overall survival in cancer randomized trials: A critical review.* Crit Rev Oncol Hematol, 2018. **123**: p. 21-41.
16.   Suh, C.H., et al., *Woptimized image-based surrogate endpoints in targeted therapies for glioblastoma: A systematic review and meta-analysis of phase III randomized controlled trials.* Korean Journal of Radiology, 2020. **21(4)**: p. 471-482.
17.   Saad, E.D., et al., *Disease-free survival as a surrogate for overall survival in patients with HER2-positive, early breast cancer in trials of adjuvant trastuzumab for up to 1 year: a systematic review and meta-analysis.* The Lancet Oncology, 2019. **20(3)**: p. 361-370.

18. Lux, M.P., et al., *Surrogate threshold effect based on a meta-analysis for the predictive value of progression-free survival for overall survival in hormone receptor-positive, HER2-negative metastatic breast cancer.* Breast Cancer Res Treat, 2019. **176**(3): p. 495-506.

19. Li, L. and Z. Pan, *Progression-Free Survival and Time to Progression as Real Surrogate End Points for Overall Survival in Advanced Breast Cancer: A Meta-Analysis of 37 Trials.* Clinical Breast Cancer, 2018. **18(1)**: p. 63-70.

20. Hirai, T., et al., *Meta-analyses on progression-free survival as a surrogate endpoint for overall survival in triple-negative breast cancer.* Breast Cancer Research and Treatment, 2020. **181(1)**: p. 189-198.

21. Guarneri, V., et al., *Survival after neoadjuvant therapy with trastuzumab-lapatinib and chemotherapy in patients with HER2-positive early breast cancer: a meta-analysis of randomized trials.* ESMO Open, 2022. **7**(2): p. 100433.

22. Forsythe, A., et al., *Progression-free survival/time to progression as a potential surrogate for overall survival in HR+, HER2- metastatic breast cancer.* Breast cancer (Dove Medical Press), 2018. **10**: p. 69-78.

23. Liu, H., et al., *Pathologic Complete Response and Its Impact on Breast Cancer Recurrence and Patient's Survival after Neoadjuvant Therapy: A Comprehensive Meta-Analysis.* Comput Math Methods Med, 2021. **2021**: p. 7545091.

24. Gyawali, B., S.P. Hey, and A.S. Kesselheim, *Evaluating the evidence behind the surrogate measures included in the FDA's table of surrogate endpoints as supporting approval of cancer drugs.* EClinicalMedicine, 2020. **21**: p. 100332.

25. Conforti, F., et al., *Evaluation of pathological complete response as surrogate endpoint in neoadjuvant randomised clinical trials of early stage breast cancer: systematic review and meta-analysis.* Bmj, 2021. **375**: p. e066381.

26. Liu, H., et al., *Evaluation of Event-Free Survival Surrogating Overall Survival as the Endpoint in Neoadjuvant Clinical Trials of Gastroesophageal Adenocarcinoma.* Front Oncol, 2022. **12**: p. 835389.

27. Kataoka, K., et al., *Surrogacy of progression-free survival (PFS) for overall survival (OS) in esophageal cancer trials with preoperative therapy: Literature-based meta-analysis.* Eur J Surg Oncol, 2017. **43**(10): p. 1956-1961.

28. Ajani, J.A., et al., *Disease-free survival as a surrogate endpoint for overall survival in adults with resectable esophageal or gastroesophageal junction cancer: A correlation meta-analysis.* European Journal of Cancer, 2022. **170**: p. 119-130.

29. Kumarasamy, C., et al., *Prognostic value of miRNAs in head and neck cancers: A comprehensive systematic and meta-analysis.* Cells, 2019. **8(8) (no pagination)**.

30. Black, C.M., et al., *Correlation Between Early Time-to-Event Outcomes and Overall Survival in Patients With Locally Advanced Head and Neck Squamous Cell Carcinoma Receiving Definitive Chemoradiation Therapy: Systematic Review and Meta-Analysis.* Frontiers in Oncology, 2022. **12 (no pagination)**.

31. Llovet, J.M., R. Montal, and A. Villanueva, *Randomized trials and endpoints in advanced HCC: Role of PFS as a surrogate of survival.* Journal of Hepatology, 2019. **70**(6): p. 1262-1277.

32. Kudo, M., et al., *Objective Response Predicts Survival in Advanced Hepatocellular Carcinoma Treated with Systemic Therapies.* Clinical Cancer Research, 2022. **28(16)**: p. 3443-3451.

33. Celsa, C., et al., *Are radiological endpoints surrogate outcomes of overall survival in hepatocellular carcinoma treated with transarterial chemoembolization?* Liver International, 2021. **41(5)**: p. 1105-1116.

34. Cabibbo, G., et al., *Progression-free survival early assessment is a robust surrogate endpoint of overall survival in immunotherapy trials of hepatocellular carcinoma.* Cancers, 2021. **13(1)**: p. 1-14.

35. Hashim, M., et al., *Do Surrogate Endpoints Better Correlate with Overall Survival in Studies That Did Not Allow for Crossover or Reported Balanced Postprogression Treatments? An Application in Advanced Non–Small Cell Lung Cancer.* Value in Health, 2018. **21**(1): p. 9-17.

36. Shameer, K., et al., *Correlation Between Early Endpoints and Overall Survival in Non-Small-Cell Lung Cancer: A Trial-Level Meta-Analysis.* Front Oncol, 2021. **11**: p. 672916.

37. Li, J., et al., *Complete response associated with immune checkpoint inhibitors in advanced non-small-cell lung cancer: a meta-analysis of nine randomized controlled trials.* Cancer Management and Research, 2019. **11**: p. 1623-1629.

38. Ito, K., et al., *The impact of high PD-L1 expression on the surrogate endpoints and clinical outcomes of anti-PD-1/PD-L1 antibodies in non-small cell lung cancer.* Lung Cancer, 2019. **128**: p. 113-119.

39. Zhao, S., et al., *Progression-free survival and one-year milestone survival as surrogates for overall survival in previously treated advanced non-small cell lung cancer.* International Journal of Cancer, 2019. **144(11)**: p. 2854-2866.

40. Fiteni, F., V. Westeel, and F. Bonnetain, *Surrogate endpoints for overall survival in lung cancer trials: a review.* Expert Rev Anticancer Ther, 2017. **17**(5): p. 447-454.

41. Imaoka, H., et al., *Progression-free survival as a surrogate endpoint in advanced neuroendocrine neoplasms.* Endocrine-Related Cancer, 2017. **24(9)**: p. 475-483.

42. Sjoquist, K.M., et al., *Progression-free survival as a surrogate endpoint for overall survival in modern ovarian cancer trials: a meta-analysis.* Therapeutic Advances in Medical Oncology, 2018. **10**(no pagination).

43. Siddiqui, M.K., et al., *Objective response rate is a possible surrogate endpoint for survival in patients with advanced, recurrent ovarian cancer.* Gynecologic Oncology, 2017. **146(1)**: p. 44-51.

44. Paoletti, X., et al., *Assessment of Progression-Free Survival as a Surrogate End Point of Overall Survival in First-Line Treatment of Ovarian Cancer: A Systematic Review and Meta-analysis.* JAMA Network Open, 2020. **3(1) (no pagination)**.

45. Shimokawa, M., et al., *Overall survival and post-progression survival are potent endpoint in phase III trials of second/third-line chemotherapy for advanced or recurrent epithelial ovarian cancer.* J Cancer, 2018. **9**(5): p. 872-879.

46. Petrelli, F., et al., *Disease-free survival is not a surrogate endpoint for overall survival in adjuvant trials of pancreatic cancer: a systematic review of randomized trials.* HPB (Oxford), 2017. **19**(11): p. 944-950.

47. Ricci, C., et al., *Disease-free survival as a measure of overall survival in resected pancreatic endocrine neoplasms.* Endocrine-Related Cancer, 2020. **27(5)**: p. 275-283.

48. Nie, R.C., et al., *Disease-free survival as a surrogate endpoint for overall survival in adjuvant trials of pancreatic cancer: a meta-analysis of 20 randomized controlled trials.* BMC Cancer, 2020. **20**(1): p. 421.

49. Makris, E.A., et al., *Surrogate End Points for Overall Survival in Metastatic, Locally Advanced, or Unresectable Pancreatic Cancer: A Systematic Review and Meta-Analysis of 24 Randomized Controlled Trials.* Annals of Surgical Oncology, 2017. **24(8)**: p. 2371-2378.

50. Gharzai, L.A., et al., *Intermediate clinical endpoints for surrogacy in localised prostate cancer: an aggregate meta-analysis.* The Lancet Oncology, 2021. **22(3)**: p. 402-410.

51. Harshman, L.C., et al., *Evaluation of disease-free survival as an intermediate metric of overall survival in patients with localized renal cell carcinoma: A trial-level meta-analysis.* Cancer, 2018. **124(5)**: p. 925-933.

52. Bria, E., et al., *Progression-free survival as primary endpoint in randomized clinical trials of targeted agents for advanced renal cell carcinoma. Correlation with overall survival, benchmarking and power analysis.* Crit Rev Oncol Hematol, 2015. **93**(1): p. 50-9.

53. Zhang, J., et al., *Endpoint surrogacy in oncological randomized controlled trials with immunotherapies: a systematic review of trial-level and arm-level meta-analyses.* Ann Transl Med, 2019. **7**(11): p. 244.

54. Suciu, S., et al., *Relapse-Free Survival as a Surrogate for Overall Survival in the Evaluation of Stage II–III Melanoma Adjuvant Therapy.* JNCI: Journal of the National Cancer Institute, 2017. **110**(1): p. 87-96.

55. Coart, E., et al., *Evaluating the potential of relapse-free survival as a surrogate for overall survival in the adjuvant therapy of melanoma with checkpoint inhibitors.* European Journal of Cancer, 2020. **137**: p. 171-174.

56. Nie, R.-C., et al., *Surrogate endpoints for overall survival in anti-programmed death-1 and anti-programmed death ligand 1 trials of advanced melanoma.* Therapeutic Advances in Medical Oncology, 2020. **12**: p. 1758835920929583.

57. Sheth, M. and J. Ko, *Exploring the relationship between Overall Survival (OS), Progression Free Survival (PFS) and Objective Response Rate (ORR) in patients with advanced melanoma.* Cancer Treatment and Research Communications, 2021. **26**: p. 100272.

58. Tanaka, K., et al., *Surrogacy of intermediate endpoints for overall survival in randomized controlled trials of first-line treatment for advanced soft tissue sarcoma in the pre- and post-pazopanib era: A meta-analytic evaluation.* BMC Cancer, 2019. **19(1) (no pagination)**.

59. Alabaku, O., et al., *Trends in endpoint use in pivotal trials and efficacy for US Food and Drug Administration–approved solid tumor therapies, 1995-2021.* Journal of Managed Care & Specialty Pharmacy, 2022. **28**(11): p. 1219-1223.

60. Xie, W., et al., *Metastasis-Free Survival Is a Strong Surrogate of Overall Survival in Localized Prostate Cancer*. Journal of Clinical Oncology, 2017. 35(27): 3097-3104.

Attachment 3

Table A12: Summary of validation studies for surrogate endpoints used in clinical trials for cancers with ≤5 PBAC submissions.

Review article^ (year)	Type	# trials/ studies included	Main Surrogate	Findings
Bladder				
Abdel-Rahman[1] (2017)	Systematic review of trials of PD-L1 inhibitors for urinary cancers	13	Overall RR, PFS	Weak correlation with OS
Bone cancer				
No validation studies identified				
Brain/Spine				
Belin[2] (2020)	Methodological review of meta-analyses	91 (1 glioblastoma)	PFS	PFS validated a trial level as a surrogate for OS
Suh[16] (2020)	Systematic review and meta-analysis in targeted therapies for glioblastoma	23	PFS	PFS strongly correlated with OS
Connective/ soft tissue				
Belin[2] (2020)	Methodological review of meta-analyses	91 (3 soft tissue sarcoma)	PFS	No validation
Savina[15] (2018)	Critical review of meta-analyses	53 (1 soft tissue sarcoma)	PFS, RR	Medium correlation of PFS and RR with OS in metastatic disease
Tanaka[58] (2019)	Meta-analysis of surrogate endpoints in RCTs of 1st line treatment for ASTS	27	PFS	Modest correlation between PFS and OS
Endometrial				
No validation studies identified				
Gastro-intestinal				
Belin[2] (2020)	Methodological review of meta-analyses	91 (5 gastric)	PFS	PFS validated as a surrogate for OS in 2/5 at trial level with a further 2/5 in favour of PFS as surrogate
Savina[15] (2018)	Critical review of meta-analyses	53 (3 gastric)	DFS, PFS	High association between DFS and OS in older publications of adjuvant chemotherapy trials; Medium to poor correlation of PFS and OS in chemotherapy trials
Liu[26] (2022)	Systematic review of neoadjuvant RCTs for gastroesophageal adenocarcinoma	8	EFS	Strong correlation between EFS and OS at trial level
Kataoka[27] (2017)	Systematic review of preoperative therapy in resectable esophageal cancer	10	PFS	Not correlated
Ajani[28] (2022)	Systematic review of neoadjuvant or preoperative therapy in resectable esophageal or gastroesophageal cancer	26	DFS, PFS	$HR_{DFS/PFS}$ was strongly correlated with HR_{OS}
Head and neck				
Belin[2] (2020)	Methodological review of meta-analyses	91 (4 head and neck)	PFS	PFS validated in 2/4 studies

Review article^ (year)	Type	# trials/ studies included	Main Surrogate	Findings
Haslam[9] (2019)	Systematic review of meta-analyses	78 (2 head and neck)	EFS	High correlation with OS in adjuvant setting; Medium correlation in neoadjuvant setting
Kumarasamy[29] (2019)	Systematic review and meta-analysis of head and neck cancer trials	34	Biomarker (microRNA)	Potential predictor for DFS and OS
Black[30] (2022)	Systematic review and meta-analysis of chemoradiation trials of advanced head and neck SCC	31	EFS	EFS strongly correlated with OS
Liver				
Haslam[9] (2019)	Systematic review of meta-analyses	78 (1 hepatocellular)	TTP	Low-medium surrogacy strength
Llovet[31] (2019)	Systematic review of advanced HCC RCTs	21	PFS	A $HR_{PFS} \leq 0.6$ was predictive of treatment effect on OS
Kudo[32] (2022)	Systematic review and meta-analysis of RCTs of systemic therapies in advanced HCC	34	Objective Response (using RECIST or mRECIST)	Modest correlation with OS. Possible use in Phase II trials for 'proof of concept' but not supported in Phase III trials.
Celsa[33] (2021)	Systematic review and meta-analysis of TACE for unresectable HCC	13	TTP	Moderate correlation between TTP and OS
Cabibbo[34] (2021)	Systematic review and meta-analysis of immunotherapy trials in advanced HCC	49 (11 ICI, 38 MKI)	PFS	Correlation varied depending on treatment and evaluation time point. In ICI trials, PFS (Q1-PFS and 12-month PFS-RMST) was 'robust' surrogate for OS
Neuroendocrine				
Belin[2] (2020)	Methodological review of meta-analyses	91 (1 neuro-endocrine)	PFS	No validation
Imaoka[41] (2017)	Systematic review of advanced neuroendocrine cancer trials	20	PFS	PFS was significantly correlated with OS
Pancreas				
Belin[2] (2020)	Methodological review of meta-analyses	91 (5 pancreas)	PFS	PFS validated in 2/5 studies, with a further 2/5 in favouring PFS as a surrogate
Haslam[9] (2019)	Systematic review of meta-analyses	78 (4 pancreas)	DCR, DFS, PFS/TTP, RR	DCR and DFS medium strength surrogates; PFS varied from low-high strength; RR low strength
Petrelli[46] (2017)	Systematic review and meta-analysis of adjuvant RCTs	12	DFS	Weak correlation between DFS and OS in adjuvant therapy for resected pancreatic cancer
Ricci[47] (2020)	Systematic review of RCTs and Monte Carlo simulation	6	DFS	Strong correlation between DFS and OS

Review article[^] (year)	Type	# trials/ studies included	Main Surrogate	Findings
Nie[48] (2020)	Systematic review and meta-analysis in adjuvant non-metastatic pancreatic cancer	20	DFS	Strong correlation between DFS and OS
Makris[49] (2017)	Systematic review and meta-analysis of RCTs in metastatic, locally advanced or unresectable pancreatic cancer	24	DCR, PFS, RR	Strong correlation between PFS and OS in 1st line chemotherapy trials
Solid tumours				
Belin[2] (2020)	Methodological review of meta-analyses	91 (1 solid tumours)	PFS	Not validated
Alabaku[59] (2022)	Review of trends in endpoint use in FDA approved solid tumour therapies 1995-2021	251 cancer-drug indication pairs (6.4% 'other')	PFS	PFS increased by 3 months since 1995, but OS has not changed
Thyroid				
No recent, relevant studies identified				

CBR=clinical benefit rate; pCR=pathological complete response; DCR=disease control rate; DFS=disease free survival; EBC=early breast cancer; EFS=event free survival; EOC=epithelial ovarian cancer; FDA=United States Food and Drug Administration; HER2=human epidermal growth factor receptor 2; HR= hormone receptor; $HR_{DFS/PFS/OS}$=hazard ratio; ICI=immune checkpoint inhibitor; MBC=metastatic breast cancer; MKI=multikinase inhibitors; OC=ovarian cancer; ORR=durable objective response rate; PD-L1=programmed death-ligand 1; PFS=progression-free survival; PPS=post-progression survival; RCC=renal cell carcinoma; RCT=randomised controlled trials; RR=response rate; TTP=time to progression

[^] Literature review citations are listed in Attachment 2

Attachment 4

Table A13: Trials that now have published final OS results versus interim OS results relied on in the submission

Drug name	Indication	PBAC meeting date	Trial name	NCT #	Interim OS, HR	Interim OS, 95%CI HR	Final OS, HR	Final OS, 95%CI HR	Final OS, aligned code	Consistency results	PBAC Rec
Lung cancer											
Afatinib	first-line, locally advanced or metastatic NSCLC	1/07/2013	LUX Lung 3	NCT00949650	0.9	0.66, 1.25	0.88	0.66, 1.17	1	consistent[a]	Rec
Crizotinib	ALK +ve, advanced NSCLC	1/11/2013	A8081007	NCT00932893	1.0	0.68, 1.54	0.85	0.66, 1.10	2	HR improved[b]	Defer
Osimertinib	first-line treatment, locally advanced or metastatic, EGFR +ve, NSCLC	1/07/2019	FLAURA	NCT02296125	0.63	0.45, 0.88[f]	0.80	0.64, 1.00	3	HR worsened[c]	Not Rec
Lorlatinib	Stage IIIB or Stage IV NSCLC	1/12/2021	CROWN	NCT03052608	0.72	0.41, 1.25	0.72	0.41, 1.25	1	consistent[a]	Rec
Blood cancer											
Bendamustine	NHL: previously untreated indolent stage III-IV NHL; MCL	1/03/2015	StiL	NCT00991211	0.70	0.48, 1.04	0.82	0.58, 1.15	3	HR worsened[c]	Defer
Obinutuzumab	NHL subtype: rituximab-refractory, follicular lymphoma	1/11/2016	GADOLIN	NCT01059630	0.72	0.48, 1.08	0.77	0.57, 1.03	1	consistent[a]	Not Rec
Obinutuzumab	NHL subtype: previously untreated advanced, follicular lymphoma	1/11/2017	GALLIUM	NCT01332968	0.82	0.54, 1.22	0.86	0.63, 1.18	1	consistent[a]	Not Rec
Obinutuzumab	CLL: unfit elderly patients with comorbidities	1/03/2015	CLL11	NCT01010061	0.70	0.47, 1.02	0.76	0.60, 0.97	4	not consistent[d]	Rec
Ibrutinib	CLL/SLL: first-line treatment CLL or SLL in unfit patients	1/11/2017	RESONATE-2	NCT01722487	0.44	0.21, 0.92[g]	0.45	0.28, 0.74	1	consistent[a]	Not Rec
Carfilzomib	MM: relapsed / refractory	1/11/2016	ENDEAVOR	NCT01568866	0.79	0.58, 1.08	0.76	0.63, 0.92	4	not consistent[d]	Not Rec
Lenalidomide	MM: maintenance therapy, newly diagnosed, undergone ASCT	1/03/2018	CALGB	NCT00114101	0.61	0.46, 0.80[h]	0.52	0.26, 1.02	4	not consistent[d,e]	Not Rec
Plitidepsin	MM: relapsed / refractory	1/07/2019	ADMYRE	NCT01102426	0.8	0.60, 1.11	0.80	0.60, 1.11	1	consistent[a]	Not Rec
Elotuzumab	MM: relapsed / refractory	1/11/2020	ELOQUENT-2	NCT01239797	0.77	0.61, 0.97#	0.82	0.68, 1.00	1	consistent[a]	Not Rec
Ixazomib	MM: relapsed / refractory	1/11/2020	TOURMALINE MM-1	NCT01564537	0.87	0.64,1.18	0.94	0.78, 1.13	1	consistent[a]	Not Rec
Daratumumab	MM: second-line	1/07/2021	COLUMBA	NCT03277105	0.91	0.66, 1.25	0.92	0.72, 1.18	1	consistent[a]	Rec

Drug name	Indication	PBAC meeting date	Trial name	NCT #	Interim OS, HR	Interim OS, 95%CI HR	Final OS, HR	Final OS, 95%CI HR	Final OS, aligned code	Consistency results	PBAC Rec
Gemtuzumab ozogamicin	AML: previously untreated, de novo CD33 +ve	1/03/2021	ALFA-0701	NCT00927498	0.81	0.60, 1.10	0.81	0.60, 1.09	1	consistent[a]	Not Rec
Skin cancer											
Dabrafenib	BRAF V600 mutation +ve advanced or metastatic melanoma	1/03/2013	BREAK-3	NCT01227889	0.75	0.44, 1.29	0.82	0.57, 1.18	1	consistent[a]	Defer
Pembrolizumab	unresectable stage III or stage IV melanoma	1/03/2015	KN-006	NCT01866319	0.6	0.43, 0.84[j]	0.68	0.53, 0.87	1	consistent[a]	Rec
Nivolumab plus ipilimumab	unresectable stage III or stage IV melanoma	1/11/2015	CA209-069	NCT01927419	0.73	0.39, 1.36	0.74	0.43, 1.26	1	consistent[a]	Not Rec
Cobimetinib with vemurafenib	BRAF V600 mutation +ve unresectable or metastatic melanoma	1/03/2016	coBRIM	NCT01689519	0.7	0.55, 0.90#	0.80	0.64, 0.99	1	consistent[a]	Rec
Talimogene laherparepvec	unresectable stage III or stage IV melanoma	1/07/2016	OPTiM	NCT00769704	0.79	0.62, 1.00	0.79	0.62, 1.00	1	consistent[a]	Not Rec
Pembrolizumab	unresectable stage III or stage IV melanoma, V600 BRAF mutation	1/03/2020	KEYNOTE-006	NCT01866319	0.7	0.44, 1.11	*	*	4	not consistent[d]	Rec
Breast cancer											
Everolimus	HRe+/HER2-, advanced	1/03/2013	BOLERO-2	NCT00863655	0.77	0.57, 1.04	0.89	0.73, 1.10	3	HR worsened[c]	Not Rec
Trastuzumab emtansine	HER2+, metastatic	1/07/2013	EMILIA	NCT00829166	0.68	0.55, 0.85[k]	0.75	0.64, 0.88	1	consistent[a]	Not Rec
Palbociclib	HRe+/HER2-, advanced or metastatic	1/03/2017	PALOMA-1	NCT00721409	0.81	0.49, 1.35	0.90	0.62, 1.29	1	consistent[a]	Not Rec
Ribociclib	first-line treatment, HRe+/HER2-, advanced	1/07/2017	MONALEESA-2	NCT01958021	0.746	0.52, 1.08	0.76	0.63, 0.93	4	not consistent[d]	Not Rec
Talazoparib	gBRCAm HER2- locally advanced inoperable or metastatic	1/11/2019	EMBRACA	NCT01945775	0.76	0.55, 1.06	0.85	0.67, 1.073	1	consistent[a]	Not Rec
Atezolizumab	unresectable locally advanced or metastatic TNBC	1/03/2020	IMpassion130	NCT02425891	0.87	0.75, 1.02	0.87	0.75, 1.02	1	consistent[a]	Not Rec
Ribociclib	first-line treatment, not been previously treated with an aromatase inhibitor, advanced	1/07/2020	MONALEESA-3	NCT02422615	0.72	0.57, 0.92#	0.73	0.59, 0.90	1	consistent[a]	Defer

Drug name	Indication	PBAC meeting date	Trial name	NCT #	Interim OS, HR	Interim OS, 95%CI HR	Final OS, HR	Final OS, 95%CI HR	Final OS, aligned code	Consistency results	PBAC Rec
Abemaciclib	HRe+/HER2-, locally advanced or metastatic	1/03/2021	MONARCH-2	NCT02107703	0.76	0.61, 0.95#	0.76	0.61, 0.95	1	consistent[a]	Rec
Renal cancer											
Lenvatinib	Stage IV clear cell variant	1/11/2017	Study 205	NCT01136733	0.55	0.30, 1.01	0.51	0.30, 0.88	4	not consistent[d]	Not Rec
Cabozantinib	Stage IV (unresectable) clear cell variant	1/12/2017	METEOR	NCT01865747	0.67	0.53, 0.83#	0.66	0.53, 0.83	1	consistent[a]	Rec
Cabozantinib	first-line treatment of Stage IV clear cell variant	1/03/2019	CABOSUN	NCT01835158	0.8	0.53, 1.21	0.8	0.50, 1.26	1	consistent[a]	Not Rec
Ovarian cancer											
Niraparib	ovarian, fallopian tube or primary peritoneal	1/03/2021	NOVA	NCT01847274	1.1	0.83, 1.46	1.1	0.83, 1.46	1	consistent[a]	Not Rec
Prostate cancer											
Apalutamide	non-metastatic castration-resistant	1/11/2018	SPARTAN	NCT01946204	0.7	0.47, 1.04	0.78	0.64, 0.96	4	not consistent[d]	Not Rec
Bowel cancer											
Panitumumab	metastatic; first-line	1/03/2013	PRIME	NCT00339183	0.88	0.73, 1.06	0.88	0.73, 1.06	1	consistent[a]	Rec
Panitumumab	metastatic; later-line	1/11/2013	ASPECCT	NCT01001377	0.97	0.84, 1.11	0.94	0.82, 1.07	1	consistent[a]	Rec
Pembrolizumab	first-line treatment of dMMR metastatic	1/03/2021	KN177	NCT02563002	0.77	0.54, 1.09	0.74	0.53, 1.03	1	consistent[a]	Rec
Gastro-intestinal cancer											
Ripretinib	gastrointestinal stromal tumour	1/03/2021	INVICTUS	NCT03353753	0.36	0.21, 0.62[m]	0.36	0.21, 0.62	1	consistent[a]	Not Rec
Pancreas cancer											
Everolimus	pancreatic neuroendocrine tumour	1/11/2012	RADIANT-3	NCT00789828	0.89	0.64, 1.23	0.94	0.7, 1.2	1	consistent[a]	Not Rec
Thyroid cancer											
Sorafenib	locally advanced / metastatic	1/07/2014	DECISION	NCT00984282	0.8	0.54, 1.19	0.92	0.71, 1.21	3	HR worsened[c]	Not Rec

ALK=anaplastic lymphoma kinase; AML=Acute myeloid leukemia; ASCT=autologous stem cell transplantation; cf=compared with; CLL=chronic lymphocytic leukaemia; dMMR=mismatch repair deficient; EGFR=epidermal growth factor receptor; gBRCAm=germline breast cancer susceptibility gene mutated; HER=human epidermal growth factor receptor; HR=hazard ratio; HRe=hormone-receptor; MCL=mantle cell lymphoma; MM=multiple myeloma; NHL=Non-Hodgkin lymphoma; NSCLC=non-small cell lung cancer; OS=overall survival, SLL=small lymphocytic leukemia; TNBC=triple-negative breast cancer.

* *Exact HR for the BRAF mutation positive subgroup was not reported in the publication but the HR was statistically significant (Figure 3 of Schachter et al 2017).*

The submission presented indirect treatment comparisons versus active comparators, for which the indirect HRs were not statistically significant.

a **Consistent**: statistical significance consistent AND HR consistent (HR within 0.1 difference versus interim results).

b **HR improved**: statistical significance consistent; HR improved cf interim results (HR reduced >0.1).

c **HR worsened**: statistical significance consistent; HR worsened cf interim results (HR increased >0.1).

d **Not consistent** on statistical significance (for our sample this would mean the OS in final analysis reached statistical significance).

e Significant HR in interim analysis but had non significant HR in final analysis.

f Unadjusted interim OS analysis (p=0.0068) did not reach formal statistical significance required at interim analysis P<0.0015 (determined by the O'Brien-Fleming approach).

g Indirect comparison of ibrutinib vs obinutuzumab showed no significant OS advantage using data from RESONATE-2 (at May 2015 cut-off, OS HR = 0.16 (0.05, 0.56)). However, the submission provided updated OS for RESONATE-2 (February 2016 data cut) for ibrutinib vs chlorambucil (HR=0.44; 95% CI: 0.21, 0.92). The PBAC noted the small number of events at the May 2015 cut-off, and that at the February 2016 data cut the relative OS gain had reduced (from HR 0.16 to 0.44). The PBAC noted there was no statistical comparison (i.e. indirect comparisons) of survival with the nominated comparators using the updated results, and considered that any survival advantage over the comparators will be reduced from that estimated using the first data cut.

h The PBAC noted a non-significant improvement in OS in a network meta-analysis, although the CALGB trial reported a significant improvement in OS.

j OS data was immature and no statistically significant difference was observed at the trial's pre-specified level (0.00002) for an interim analysis. The PBAC expected that the HR would tend towards the null with additional follow-up.

k The PBAC considered that the results from the indirect analysis versus comparator (trastuzumab + chemotherapy) included the null and had wide CIs.

m Due to the hierarchal test testing procedure of the endpoints, OS could not be formally tested for statistical significance because the objective response was not significant; the nominal p-value displayed is based on 2-sided stratified log-rank test.

References

Alabaku, O., Laffey, T. N., Suh, K., & Li, M. (2022). Trends in endpoint use in pivotal trials and efficacy for US Food and Drug Administration–approved solid tumor therapies, 1995-2021. *Journal of Managed Care & Specialty Pharmacy*, *28*(11), 1219-1223. https://doi.org/10.18553/jmcp.2022.28.11.1219

Amabile, S., Roccuzzo, G., Pala, V., Tonella, L., Rubatto, M., Merli, M., Fava, P., Ribero, S., Fierro, M. T., Queirolo, P., & Quaglino, P. (2021). Clinical Significance of Distant Metastasis-Free Survival (DMFS) in Melanoma: A Narrative Review from Adjuvant Clinical Trials. *J Clin Med*, *10*(23). https://doi.org/10.3390/jcm10235475

Amaria, R. N., Menzies, A. M., Burton, E. M., Scolyer, R. A., Tetzlaff, M. T., Antdbacka, R., Ariyan, C., Bassett, R., Carter, B., Daud, A., Faries, M., Fecher, L. A., Flaherty, K. T., Gershenwald, J. E., Hamid, O., Hong, A., Kirkwood, J. M., Lo, S., Margolin, K., . . . Long, G. V. (2019). Neoadjuvant systemic therapy in melanoma: recommendations of the International Neoadjuvant Melanoma Consortium. *Lancet Oncol*, *20*(7), e378-e389. https://doi.org/10.1016/s1470-2045(19)30332-8

Andtbacka, R. H., Kaufman, H. L., Collichio, F., Amatruda, T., Senzer, N., Chesney, J., Delman, K. A., Spitler, L. E., Puzanov, I., Agarwala, S. S., Milhem, M., Cranmer, L., Curti, B., Lewis, K., Ross, M., Guthrie, T., Linette, G. P., Daniels, G. A., Harrington, K., . . . Coffin, R. S. (2015). Talimogene Laherparepvec Improves Durable Response Rate in Patients With Advanced Melanoma. *J Clin Oncol*, *33*(25), 2780-2788. https://doi.org/10.1200/jco.2014.58.3377

Avet-Loiseau, H., Ludwig, H., Landgren, O., Paiva, B., Morris, C., Yang, H., Zhou, K., Ro, S., & Mateos, M. V. (2020). Minimal Residual Disease Status as a Surrogate Endpoint for Progression-free Survival in Newly Diagnosed Multiple Myeloma Studies: A Meta-analysis. *Clinical Lymphoma, Myeloma and Leukemia*, *20(1)*, e30-e37. https://doi.org/https://dx.doi.org/10.1016/j.clml.2019.09.622

Belin, L., Tan, A., De Rycke, Y., & Dechartres, A. (2020). Progression-free survival as a surrogate for overall survival in oncology trials: a methodological systematic review. *British Journal of Cancer*, *122*(11), 1707-1714. https://doi.org/10.1038/s41416-020-0805-y

Borcoman, E., Kanjanapan, Y., Champiat, S., Kato, S., Servois, V., Kurzrock, R., Goel, S., Bedard, P., & Le Tourneau, C. (2019). Novel patterns of response under immunotherapy. *Annals of Oncology*, *30*(3), 385-396. https://doi.org/10.1093/annonc/mdz003

Buyse, M., Sargent, D. J., Grothey, A., Matheson, A., & de Gramont, A. (2010). Biomarkers and surrogate end points--the challenge of statistical validation. *Nat Rev Clin Oncol*, *7*(6), 309-317. https://doi.org/10.1038/nrclinonc.2010.43

Ciani, O., Buyse, M., Drummond, M., Rasi, G., Saad, E. D., & Taylor, R. S. (2017). Time to Review the Role of Surrogate End Points in Health Policy: State of the Art and the Way Forward. *Value in Health*, *20*(3), 487-495. https://doi.org/https://doi.org/10.1016/j.jval.2016.10.011

Conforti, F., Pala, L., Sala, I., Oriecuia, C., De Pas, T., Specchia, C., Graffeo, R., Pagan, E., Queirolo, P., Pennacchioli, E., Colleoni, M., Viale, G., Bagnardi, V., & Gelber, R. D. (2021). Evaluation of pathological complete response as surrogate endpoint in neoadjuvant randomised clinical trials of early stage breast cancer: systematic review and meta-analysis. *BMJ*, *375*, e066381. https://doi.org/10.1136/bmj-2021-066381

Daniele, P., Mamolo, C., Cappelleri, J. C., Bell, T., Neuhof, A., Tremblay, G., Musat, M., & Forsythe, A. (2022). Response rates and minimal residual disease outcomes as potential surrogates for progression-free survival in newly diagnosed multiple myeloma. *PLoS ONE*, *17(5 May) (no pagination)*, Article e0267979. https://doi.org/https://dx.doi.org/10.1371/journal.pone.0267979

Durie, B. G. M., Harousseau, J. L., Miguel, J. S., Bladé, J., Barlogie, B., Anderson, K., Gertz, M., Dimopoulos, M., Westin, J., Sonneveld, P., Ludwig, H., Gahrton, G., Beksac, M., Crowley, J., Belch, A., Boccadaro, M., Turesson, I., Joshua, D., Vesole, D., . . . on behalf of the International

Myeloma Working, G. (2006). International uniform response criteria for multiple myeloma. *Leukemia*, *20*(9), 1467-1473. https://doi.org/10.1038/sj.leu.2404284

Eisenhauer, E. A., Therasse, P., Bogaerts, J., Schwartz, L. H., Sargent, D., Ford, R., Dancey, J., Arbuck, S., Gwyther, S., Mooney, M., Rubinstein, L., Shankar, L., Dodd, L., Kaplan, R., Lacombe, D., & Verweij, J. (2009). New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*, *45*(2), 228-247. https://doi.org/10.1016/j.ejca.2008.10.026

Estey, E., Othus, M., Lee, S. J., Appelbaum, F. R., & Gale, R. P. (2016). New drug approvals in acute myeloid leukemia: what's the best end point? *Leukemia*, *30*(3), 521-525. https://doi.org/10.1038/leu.2015.262

Fiteni, F., Westeel, V., & Bonnetain, F. (2017). Surrogate endpoints for overall survival in lung cancer trials: a review. *Expert Rev Anticancer Ther*, *17*(5), 447-454. https://doi.org/10.1080/14737140.2017.1316196

Forsythe, A., Chandiwana, D., Barth, J., Thabane, M., Baeck, J., & Tremblay, G. (2018). Progression-free survival/time to progression as a potential surrogate for overall survival in HR+, HER2- metastatic breast cancer. *Breast cancer (Dove Medical Press)*, *10*, 69-78. https://doi.org/10.2147/BCTT.S162841

Gharzai, L. A., Jiang, R., Wallington, D., Jones, G., Birer, S., Jairath, N., Jaworski, E. M., McFarlane, M. R., Mahal, B. A., Nguyen, P. L., Sandler, H., Morgan, T. M., Reichert, Z. R., Alumkal, J. J., Mehra, R., Kishan, A. U., Fizazi, K., Halabi, S., Schaeffer, E. M., . . . Spratt, D. E. (2021). Intermediate clinical endpoints for surrogacy in localised prostate cancer: an aggregate meta-analysis. *The Lancet Oncology*, *22(3)*, 402-410. https://doi.org/https://dx.doi.org/10.1016/S1470-2045%2820%2930730-0

Gion, M., Pérez-García, J. M., Llombart-Cussac, A., Sampayo-Cordero, M., Cortés, J., & Malfettone, A. (2021). Surrogate endpoints for early-stage breast cancer: a review of the state of the art, controversies, and future prospects. *Ther Adv Med Oncol*, *13*, 17588359211059587. https://doi.org/10.1177/17588359211059587

Gourgou-Bourgade, S., Cameron, D., Poortmans, P., Asselain, B., Azria, D., Cardoso, F., A'Hern, R., Bliss, J., Bogaerts, J., Bonnefoi, H., Brain, E., Cardoso, M. J., Chibaudel, B., Coleman, R., Cufer, T., Dal Lago, L., Dalenc, F., De Azambuja, E., Debled, M., . . . Dabakuyo-Yonli, T. S. (2015). Guidelines for time-to-event end point definitions in breast cancer trials: results of the DATECAN initiative (Definition for the Assessment of Time-to-event Endpoints in CANcer trials)^{†}. *Annals of Oncology*, *26*(5), 873-879. https://doi.org/10.1093/annonc/mdv106

Grigore, B., Ciani, O., Dams, F., Federici, C., de Groot, S., Möllenkamp, M., Rabbe, S., Shatrov, K., Zemplenyi, A., & Taylor, R. S. (2020). Surrogate Endpoints in Health Technology Assessment: An International Review of Methodological Guidelines. *PharmacoEconomics*, *38*(10), 1055-1070. https://doi.org/10.1007/s40273-020-00935-1

Gyawali, B., Hey, S. P., & Kesselheim, A. S. (2020). Evaluating the evidence behind the surrogate measures included in the FDA's table of surrogate endpoints as supporting approval of cancer drugs. *EClinicalMedicine*, *21*, 100332. https://doi.org/10.1016/j.eclinm.2020.100332

Hashim, M., Pfeiffer, B. M., Bartsch, R., Postma, M., & Heeg, B. (2018). Do Surrogate Endpoints Better Correlate with Overall Survival in Studies That Did Not Allow for Crossover or Reported Balanced Postprogression Treatments? An Application in Advanced Non–Small Cell Lung Cancer. *Value in Health*, *21*(1), 9-17. https://doi.org/https://doi.org/10.1016/j.jval.2017.07.011

Haslam, A., Hey, S. P., Gill, J., & Prasad, V. (2019). A systematic review of trial-level meta-analyses measuring the strength of association between surrogate end-points and overall survival in oncology. *Eur J Cancer*, *106*, 196-211. https://doi.org/10.1016/j.ejca.2018.11.012

Hernandez-Villafuerte, K., Fischer, A., & Latimer, N. (2018). CHALLENGES AND METHODOLOGIES IN USING PROGRESSION FREE SURVIVAL AS A SURROGATE FOR OVERALL SURVIVAL IN

ONCOLOGY. *Int J Technol Assess Health Care*, *34*(3), 300-316. https://doi.org/10.1017/S0266462318000338

Johnson, J. R., Ning, Y.-M., Farrell, A., Justice, R., Keegan, P., & Pazdur, R. (2011). Accelerated Approval of Oncology Products: The Food and Drug Administration Experience. *JNCI: Journal of the National Cancer Institute*, *103*(8), 636-644. https://doi.org/10.1093/jnci/djr062

Kemp, R., & Prasad, V. (2017). Surrogate endpoints in oncology: when are they acceptable for regulatory and clinical decisions, and are they currently overused? *BMC Med*, *15*(1), 134. https://doi.org/10.1186/s12916-017-0902-9

Kumar, S. K., Rajkumar, S. V., Dispenzieri, A., Lacy, M. Q., Hayman, S. R., Buadi, F. K., Zeldenrust, S. R., Dingli, D., Russell, S. J., Lust, J. A., Greipp, P. R., Kyle, R. A., & Gertz, M. A. (2008). Improved survival in multiple myeloma and the impact of novel therapies. *Blood*, *111*(5), 2516-2520. https://doi.org/10.1182/blood-2007-10-116129

Liu, H., Lv, L., Gao, H., & Cheng, M. (2021). Pathologic Complete Response and Its Impact on Breast Cancer Recurrence and Patient's Survival after Neoadjuvant Therapy: A Comprehensive Meta-Analysis. *Comput Math Methods Med*, *2021*, 7545091. https://doi.org/10.1155/2021/7545091

Mangal, N., Salem, A. H., Li, M., Menon, R., & Freise, K. J. (2018). Relationship between response rates and median progression-free survival in non-Hodgkin's lymphoma: A meta-analysis of published clinical trials. *Hematological Oncology*, *36(1)*, 37-43. https://doi.org/https://dx.doi.org/10.1002/hon.2463

Mangal, N., Salem, A. H., Menon, R. M., & Freise, K. J. (2018). Use of depth of response to predict progression-free survival in relapsed or refractory multiple myeloma: Evaluation of results from 102 clinical trials. *Hematological Oncology*, *36(3)*, 547-553. https://doi.org/https://dx.doi.org/10.1002/hon.2514

Medeiros, B. C. (2018). Interpretation of clinical endpoints in trials of acute myeloid leukemia. *Leukemia Research*, *68*, 32-39. https://doi.org/https://doi.org/10.1016/j.leukres.2018.02.002

Mo, D.-C., Huang, J.-F., Luo, P.-H., Chen, L., & Huang, S.-X. (2023). Distant metastasis-free survival with adjuvant pembrolizumab for resected stage IIB or IIC melanoma. *The Lancet Oncology*, *24*(1), e7. https://doi.org/https://doi.org/10.1016/S1470-2045(22)00703-3

Munshi, N. C., Avet-Loiseau, H., Anderson, K. C., Neri, P., Paiva, B., Samur, M., Dimopoulos, M., Kulakova, M., Lam, A., Hashim, M., He, J., Heeg, B., Ukropec, J., Vermeulen, J., Cote, S., & Bahlis, N. (2020). A large meta-analysis establishes the role of MRD negativity in long-term survival outcomes in patients with multiple myeloma. *Blood Advances*, *4*(23), 5988-5999. https://doi.org/10.1182/bloodadvances.2020002827

Nathwani, N., Bell, J., Cherepanov, D., Sowell, F. G., Shah, R., McCarrier, K., & Hari, P. (2022). Patient perspectives on symptoms, health-related quality of life, and treatment experience associated with relapsed/refractory multiple myeloma. *Support Care Cancer*, *30*(7), 5859-5869. https://doi.org/10.1007/s00520-022-06979-7

Nie, R.-C., Yuan, S.-Q., Wang, Y., Zou, X.-B., Chen, S., Li, S.-M., Duan, J.-L., Zhou, J., Chen, G.-M., Luo, T.-Q., Zhou, Z.-W., & Li, Y.-F. (2020). Surrogate endpoints for overall survival in anti-programmed death-1 and anti-programmed death ligand 1 trials of advanced melanoma. *Therapeutic Advances in Medical Oncology*, *12*, 1758835920929583. https://doi.org/10.1177/1758835920929583

Saad, E. D., Squifflet, P., Burzykowski, T., Quinaux, E., Delaloge, S., Mavroudis, D., Perez, E., Piccart-Gebhart, M., Schneider, B. P., Slamon, D., Wolmark, N., & Buyse, M. (2019). Disease-free survival as a surrogate for overall survival in patients with HER2-positive, early breast cancer in trials of adjuvant trastuzumab for up to 1 year: a systematic review and meta-analysis. *The Lancet Oncology*, *20(3)*, 361-370. https://doi.org/https://dx.doi.org/10.1016/S1470-2045%2818%2930750-2

Savina, M., Gourgou, S., Italiano, A., Dinart, D., Rondeau, V., Penel, N., Mathoulin-Pelissier, S., & Bellera, C. (2018). Meta-analyses evaluating surrogate endpoints for overall survival in cancer

randomized trials: A critical review. *Crit Rev Oncol Hematol*, *123*, 21-41. https://doi.org/10.1016/j.critrevonc.2017.11.014

Suciu, S., Eggermont, A. M. M., Lorigan, P., Kirkwood, J. M., Markovic, S. N., Garbe, C., Cameron, D., Kotapati, S., Chen, T.-T., Wheatley, K., Ives, N., de Schaetzen, G., Efendi, A., & Buyse, M. (2017). Relapse-Free Survival as a Surrogate for Overall Survival in the Evaluation of Stage II–III Melanoma Adjuvant Therapy. *JNCI: Journal of the National Cancer Institute*, *110*(1), 87-96. https://doi.org/10.1093/jnci/djx133

U.S. Department of Health and Human Services Food and Drug Administration. (2018). *Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics Guidance for Industry*. Silver Spring, MD, USA

Walia, A., Haslam, A., & Prasad, V. (2022). FDA validation of surrogate endpoints in oncology: 2005-2022. *J Cancer Policy*, *34*, 100364. https://doi.org/10.1016/j.jcpo.2022.100364

Xie, W., Regan, M. M., Buyse, M., Halabi, S., Kantoff, P. W., Sartor, O., Soule, H., Clarke, N. W., Collette, L., Dignam, J. J., Fizazi, K., Paruleker, W. R., Sandler, H. M., Sydes, M. R., Tombal, B., Williams, S. G., & Sweeney, C. J. (2017). Metastasis-Free Survival Is a Strong Surrogate of Overall Survival in Localized Prostate Cancer. *J Clin Oncol*, *35*(27), 3097-3104. https://doi.org/10.1200/jco.2017.73.9987

Zarin, D. A., Tse, T., Williams, R. J., Califf, R. M., & Ide, N. C. (2011). The ClinicalTrials.gov results database--update and key issues. *N Engl J Med*, *364*(9), 852-860. https://doi.org/10.1056/NEJMsa1012065