

***Anti-Müllerian
hormone (AMH)
testing for female
patients who will or
have received
gonadotoxic
treatment***

January 2018

MSAC application no. 1434

Assessment report

© Commonwealth of Australia 2018

ISSN (Online) 1443-7139

Internet site <http://www.msac.gov.au/>.

This work is copyright. You may download, display, print and reproduce this material in unaltered form only (retaining this notice) for your personal, non-commercial use or use within your organisation. Apart from any use as permitted under the Copyright Act 1968, all other rights are reserved. Requests and inquiries concerning reproduction and rights should be addressed to Commonwealth Copyright Administration, Attorney-General's Department, Robert Garran Offices, National Circuit, Barton ACT 2600 or posted at <http://www.ag.gov.au/>.

Electronic copies of the report can be obtained from the Medical Service Advisory Committee's Internet site at <http://www.msac.gov.au/>.

Enquiries about the content of the report should be emailed to hta@health.gov.au.

The technical information in this document is used by the Medical Services Advisory Committee (MSAC) to inform its deliberations. MSAC is an independent committee which has been established to provide advice to the Minister for Health on the strength of evidence available on new and existing medical technologies and procedures in terms of their safety, effectiveness and cost-effectiveness. This advice will help to inform government decisions about which medical services should attract funding under Medicare.

MSAC's advice does not necessarily reflect the views of all individuals who participated in the MSAC evaluation.

This report was prepared by Ms Sharon Kessels, Dr Ruchi Mittal, Dr Judy Morona, Ms Skye Newton, Ms Camille Schubert, Ms Jo Milverton and Prof Tracy Merlin from Adelaide Health Technology Assessment (AHTA). Clinical advice was provided by Dr Antoinette Anazodo, A/Prof Catharyn Stern, Prof William Ledger and Dr David Molloy. The report was commissioned by the Australian Government Department of Health. The report was edited by Mr Ben Ellery.

The suggested citation for this document is:

Kessels S, Mittal R, Morona J, Newton S, Schubert C, Milverton J, Merlin T. (2018). Anti-Müllerian hormone (AMH) testing for female patients who will or have received gonadotoxic treatment. MSAC application 1434, Assessment Report. Commonwealth of Australia, Canberra, ACT.



CONTENTS

Contents	iii
Tables	viii
Figures	xiii
Executive summary	1
Anti-Müllerian hormone testing for female patients preceding or following gonadotoxic treatment	1
Alignment with agreed PICO Confirmation	2
Proposed medical service	2
Proposal for public funding	3
Population	3
Comparator details	3
Clinical management algorithm(s)	4
Clinical claim	4
Approach taken to the evidence assessment	4
Characteristics of the evidence base	5
Results	6
Summary of findings	11
Translation issues	11
Economic evaluation	11
Estimated extent of use and financial implications	12
Consumer impact summary	13
Other relevant considerations	14
Acronyms and abbreviations	15
Section A Context	18
A1 Items in the agreed PICO Confirmation	18
A2 Proposed medical service	19
Fertility and the ovarian follicle pool	19
Ovarian failure and AMH testing	19
A3 Proposal for public funding	20
A4 Proposed population	21
Patients prior to, or after completion of gonadotoxic treatment	21
Why gonadotoxic treatment is a concern	23
Estimated incidence	24
Uptake of AMH testing in Australia	24

A5	Comparator details	25
	Clinical management algorithms	28
A6	Clinical Claim	32
A7	Summary of the PICO.....	32
	Direct evidence.....	33
	Linked evidence	35
A8	Consumer impact statement.....	39
Section B	Clinical evaluation	41
B1	Direct evidence.....	42
B1.1	Literature sources and search strategies.....	42
B1.2	Results of literature search	42
	Appraisal of the evidence.....	43
B1.3	Risk of bias assessment.....	44
B1.4	Characteristics of the evidence base	44
B1.5	Outcome measures and analysis.....	44
B1.6	Results of the systematic literature review	45
	Is it safe?	45
	Is it effective?	45
B2	Linked evidence approach	46
B2.1	Basis for linked evidence.....	46
B2.2	Steps for linked analysis.....	46
B3	Analytical validity	47
B3.1	Reference standard	47
B3.2	Literature sources and search strategies.....	47
B3.3	Results of literature search	47
B3.4	Risk of bias assessment.....	48
B3.5	Characteristics of the evidence base	49
B3.6	Outcome measures and analysis.....	50
	Analytical validity.....	50
	Receiver operator characteristic (ROC) analysis.....	51
	Correlation	52
B3.7	Results of the systematic literature review	53
	Is it accurate?	53
B3.8	Extended assessment of reliability evidence	56
B3.9	Concordance analysis	56

	Overall per cent agreement	56
	Correlation between tests.....	57
B3.10	Interpretation of evidence on analytical validity.....	58
	Analytical validity of AMH testing compared with a clinical reference standard	58
	Comparing the accuracy of different AMH tests	59
	Accuracy of AMH testing compared with AFC.....	60
	Accuracy of AMH testing compared with FSH testing	60
	Accuracy of AMH testing compared with E2 testing	61
	Accuracy of AMH testing compared with inhibin B testing	61
B4	Clinical validity.....	62
B4.1	Measures of clinical validity	62
B4.1.1	Reference standard.....	62
B4.1.2	Risk of bias assessment.....	62
B4.1.3	Characteristics of the evidence base	63
B4.1.4	Outcome measures and analysis	67
B4.1.5	Results of the systematic literature review.....	68
	Prognostic and predictive value of AMH testing	68
	Is it accurate? - Usefulness of AMH at varying risks of ovarian failure (clinical validity)	
	85
B5	Clinical utility	94
B5.1	Impact of AMH testing on clinical management (therapeutic efficacy)	94
B5.1.1	Risk of bias assessment.....	94
B5.1.2	Characteristics of the evidence base	94
B5.1.3	Outcome measures and analysis	94
B5.1.4	Results of the systematic literature review.....	95
	Does AMH testing impact clinical management?.....	95
B5.2	Therapeutic effectiveness.....	97
B5.2.1	Risk of bias assessment.....	97
B5.2.2	Characteristics of the evidence base	98
B5.2.3	Outcome measures and analysis	98
B5.2.4	Results of the systematic literature review.....	98
	Does the change in management improve health outcomes?	98
B6	Impact of repeat testing / monitoring	103
B7	Extended assessment of comparative harms.....	105
B8	Interpretation of the clinical evidence.....	106

Section C	Translation issues	110
Section D	Economic evaluation	111
	D.1. Overview	111
	D.2. Populations and settings.....	111
	D.3. Structure and rationale of the economic evaluation.....	112
	Literature review	112
	D.4. Inputs to the economic evaluation.....	112
	Test costs and associated costs	112
	D.5. Results of the economic evaluation	113
	D.6. Sensitivity analyses.....	113
Section E	Financial implications	115
	E.1. Justification of the selection of sources of data	115
	E.2. Use and costs of AMH testing	115
	Estimated use of AMH testing	115
	Estimated costs of AMH testing	119
	E.3. Changes in Use and cost of other medical services.....	120
	E.4. Financial implications for the MBS	120
	E.5. Financial implications for government health budgets	120
	E.6. Identification, estimation and reduction of uncertainty	120
	Alternative referral and uptake rates	120
	Reduced use in younger and older women.....	121
	Leakage	121
	AMH as a replacement test.....	121
Section F	Other relevant considerations	123
Appendix A	Clinical experts and assessment group	124
	Clinical experts	124
	Assessment group	124
Appendix B	Search strategies	125
	Bibliographic databases	125
	Search terms	128
Appendix C	Studies included in the systematic review.....	130
Appendix D	Evidence profile tables	152
Appendix E	Excluded studies.....	159
	Full text could not be retrieved.....	159

No extractable data for analytical validity section or only predictive/prognostic data	159
Incorrect population	159
Incorrect intervention	159
No comparator	160
Incorrect outcomes.....	160
Excluded from predictive/prognostic section	160
Not predictive/prognostic.....	160
Data not extractable	161
Incorrect population or study population too small (n<10).....	162
Appendix F Quality appraisal of prognostic evidence.....	163
Appendix G Correlation data	164
References	168

TABLES

Table 1	Proposed MBS item descriptor	3
Table 2	Key features of the included linked evidence.....	5
Table 3	Summary of the economic evaluation	12
Table 4	Costs associated with testing ovarian reserve, and incremental cost (per patient).....	12
Table 5	Estimated costs of AMH testing, 2018–19 to 2022–23	13
Table 6	Proposed MBS item descriptor	21
Table 7	Risk of amenorrhea in women treated with chemotherapy and radiotherapy, adapted from Loren et al. 2013 (ASCO guidelines).....	22
Table 8	Incidence of cancer in Australian females.....	24
Table 9	Estimated number of new cases, by type of cancer most relevant for fertility preservation for 2017, females aged 0 to 44	24
Table 10	Relevant MBS item for comparator tests	27
Table 11	Explanatory notes for comparator MBS item number.....	27
Table 12	Relevant MBS item for ultrasound used for antral follicle count	28
Table 13	PICO criteria and research questions for direct evidence in populations 1, 2 and 3	33
Table 14	PICO criteria and research questions for direct evidence in populations 4, 5 and 6	34
Table 15	PICO criteria and research question for the analytical validity of AMH testing in populations 1 to 3	35
Table 16	PICO criteria and research question for determining prognostic value of AMH in population 1, 2 and 3	35
Table 17	PICO criteria and research question for determining predictive value of AMH in population 1, 2 and 3	36
Table 18	PICO criteria and research question to determine the impact of the AMH test on patient management in population 1, 2 and 3	36
Table 19	PICO criteria and research question to determine the effectiveness of change in management in population 1, 2 and 3	37
Table 20	PICO criteria and research question for analytical validity of AMH testing in populations 4, 5 and 6	38

Table 21	PICO criteria and research question to determine the impact of the AMH test on patient management in population 4, 5 and 6	38
Table 22	PICO criteria and research question to determine the effectiveness of change in management in population 1, 2 and 3	39
Table 23	Key features of the included evidence comparing intervention with comparator against reference standard	47
Table 24	Studies reporting on the correlation between the AMH test and other tests.....	48
Table 25	Modified QUADAS-2 risk of bias results for analytical validity and concordance studies	48
Table 26	Modified QUADAS-2 risk of bias results for studies reporting correlations between AMH and other tests	49
Table 27	Modified QUADAS-2 risk of bias results for studies reporting the correlation between different AMH tests	49
Table 28	Summary of the AMH tests used in the included studies	50
Table 29	Analytical validity data extraction.....	51
Table 30	Concordance data extraction	52
Table 31	Analytical validity of the EIA AMH/MIS assay and pico-AMH ELISA against the clinical reference standard of resumption of menses	54
Table 32	AUC values from ROC analysis conducted by Su et al. (2011)	55
Table 33	Analytical validity of ovarian reserve markers using thresholds determined by ROC curve analysis	55
Table 34	Spearman’s rank correlation between two AMH assays.....	56
Table 35	2 × 2 concordance data from Miyoshi at al. (2013)	57
Table 36	Summary of the correlation between the AMH test and other tests used to determine ovarian function.....	57
Table 37	Modified QUADAS-2 risk of bias results for clinical validity studies	63
Table 38	Key features of the included evidence on prognostic or predictive value of AMH testing, ordered by study size	64
Table 39	Key features of the included evidence on clinical validity of AMH testing, ordered by study size	67
Table 40	Analytical validity data extraction.....	67

Table 41	Baseline AMH levels in breast cancer patients with CRA at follow-up compared with baseline AMH levels in patients without CRA	70
Table 42	Baseline FSH levels in breast cancer patients with CRA at follow-up compared with baseline FSH levels in patients without CRA	71
Table 43	Baseline E2 levels in breast cancer patients with CRA at follow-up compared with baseline E2 levels in patients without CRA	72
Table 44	Baseline inhibin B levels in breast cancer patients with CRA at follow-up compared with baseline inhibin B levels in patients without CRA	72
Table 45	Baseline AFC count in breast cancer patients with CRA at follow-up compared with baseline AFC count in patients without CRA.....	72
Table 46	Age at baseline in breast cancer patients with CRA at follow-up compared with age at baseline in patients without CRA at follow-up	72
Table 47	Multivariate regression analysis for prediction of ovarian function in breast cancer patients.....	76
Table 48	AUC values from ROC analysis conducted by Chai et al. (2014)	79
Table 49	Predictive value of AMH predicting pregnancy outcomes in women undergoing gonadotoxic treatment, ordered by study size	81
Table 50	Association between number of in vitro matured oocytes cryopreserved and AFC and serum AMH levels, univariate linear regression.....	83
Table 51	Threshold values of AFC and serum AMH levels for obtaining ≤ 2 or ≥ 8 , 10 or 15 mature oocytes frozen.....	84
Table 52	Correlation between oocytes retrieved and AMH, AFC, FSH, inhibin B and/or E2 values	85
Table 53	Clinical validity of the AMH test against the clinical reference standard of resumption of menses.....	87
Table 54	Clinical validity of the AMH test against the clinical reference standard of resumption of menses.....	90
Table 55	Clinical validity of ovarian reserve markers using thresholds determines by ROC curve analysis.....	93
Table 56	Controlled ovarian hyperstimulation protocol influenced by AMH levels, in St Mary's hospital, Manchester (from 2008)	96
Table 57	IVF parameters and outcomes in the nomogram group (individualised FSH dosage) compared with a control group (standardised FSH dosage)	99

Table 58	Pre-embryology clinical outcomes in a conventional dosage protocol compared with an AMH-tailored protocol	100
Table 59	Outcomes after individualised follitropin delta use compared to standard follitropin alpha use.....	101
Table 60	Summary of findings table.....	107
Table 61	Classification of the comparative effectiveness and safety of the proposed therapeutic medical service compared with its main comparator and guide to the suitable type of economic evaluation.....	111
Table 62	Summary of the economic evaluation	112
Table 63	Various costs associated with ovarian reserve testing	113
Table 64	Costs associated with testing ovarian reserve, and incremental cost per patient	113
Table 65	Scenario analyses: AMH is used as replacement test to either AFC, or FSH+E2.....	114
Table 66	Data sources used in the financial analysis	115
Table 67	Incident cases of cancer (females 0–44 years) estimated to undergo gonadotoxic treatment (2017).....	116
Table 68	Estimate of number of AMH tests performed, 2018–19 to 2022–23	118
Table 69	Proposed or advertised fees ^a for AMH testing	119
Table 70	Estimated costs of AMH testing, 2018–19 to 2022–23.....	119
Table 71	Financial implications of listing the AMH test on the MBS, sensitivity analysis varying referral and uptake rates	120
Table 72	Financial implications of MBS listing the AMH test: scenario analysis of leakage to other indications	121
Table 73	Scenario analyses, potential cost-offsets per test to MBS expenditure, if AMH is a replacement test	121
Table 74	Scenario analyses, potential cost-offsets to MBS expenditure, if AMH is a replacement test for AFC.....	122
Table 75	Scenario analyses, potential cost-offsets to MBS, if AMH is a replacement test for FSH+E2	122
Table 76	Bibliographic databases.....	125
Table 77	Additional sources of literature	125
Table 78	Specialty websites.....	126

Table 79	HTA websites	126
Table 80	Suggested search terms for AMH testing to measure ovarian reserve to determine the requirement for cryopreservation of ovarian tissue or oocytes prior to or following gonadotoxic treatment, PubMed.....	128
Table 81	Suggested search terms for AMH testing to measure ovarian reserve to determine the requirement for cryopreservation of ovarian tissue or oocytes prior to or following gonadotoxic treatment, Embase.....	129
Table 82	Profiles of studies on analytical validity studies providing 2x2 or ROC data included in the systematic literature review	130
Table 83	Profiles of studies on diagnostic concordance between different AMH tests included in the systematic literature review.....	132
Table 84	Profiles of studies on diagnostic concordance between AMH and other tests included in the systematic literature review.....	132
Table 85	Prognostic evidence on ovarian function / amenorrhea / menses	137
Table 86	Predictive evidence: AMH levels predicting ovarian response.....	146
Table 87	Prognostic evidence: AMH levels predicting pregnancy	148
Table 88	Evidence profile table for the analytical validity of AMH testing for girls or women undergoing gonadotoxic treatment	152
Table 89	Evidence profile table for the prognostic and predictive evidence of AMH testing for women undergoing gonadotoxic treatment.....	155
Table 90	Evidence profile table for the clinical validity of AMH testing for girls or women undergoing gonadotoxic treatment	158
Table 91	Quality appraisal for studies included in predictive and prognostic evidence section of B4, measured using QUIPS tool	163
Table 92	Correlation of AMH and AFC	164
Table 93	Correlation between AMH and ovarian biopsy (primordial follicle count)	164
Table 94	Correlation between AMH and FSH levels	165
Table 95	Correlation between AMH and E2 levels	166
Table 96	Correlation between AMH and inhibin B levels.....	167

FIGURES

Figure 1	Changes in hormone levels and AFC with decreasing ovarian reserve approaching menopause and post-menopause.	19
Figure 2	Clinical management algorithm for assessment of ovarian reserve and cryopreservation of oocytes or ovarian tissue. Pathway showing cryopreservation of oocytes is current practice in post-pubertal cases. Pathway showing cryopreservation of ovarian tissue is for the proposed service. The relevant population for both pathways includes females aged 0 to 45 years, pre-gonadotoxic treatment.....	29
Figure 3	Current clinical management algorithm for the assessment of ovarian reserve and post-pubertal cryopreservation oocytes for females aged 0 to 45 years, following completion of gonadotoxic treatment.....	30
Figure 4	Proposed clinical management algorithm for the assessment of ovarian reserve and cryopreservation of oocytes or ovarian tissue for females aged 0 to 45 years, following completion of gonadotoxic treatment	31
Figure 5	Summary of the process used to identify and select studies for the assessment	43
Figure 6	PPV values with increasing prevalence of treatment-induced ovarian failure from 20% to 80%, AMH measured prior to treatment	87
Figure 7	NPV values with increasing prevalence of treatment-induced ovarian failure from 20% to 80%, AMH measured prior to treatment	88
Figure 8	PPV values with increasing prevalence of treatment-induced ovarian failure from 20% to 80% (AMH measured post-treatment).....	91
Figure 9	NPV values with increasing prevalence of treatment-induced ovarian failure from 20% to 80% (AMH measured post-treatment).....	91

EXECUTIVE SUMMARY

Main issues for MSAC consideration

- No direct evidence was identified to determine the safety and incremental effectiveness of anti-Müllerian hormone (AMH) testing **in addition** to other standard tests, or compared to other standard tests alone, in patients prior to or following completion of gonadotoxic therapy.
- The test is generally considered to be safe.
- Different AMH assays correlated highly with each other, although they differed greatly in sensitivity.
- AMH testing does provide some incremental information for predicting ovarian functioning. However, the relationship between AMH testing and the most clinically relevant outcome of pregnancy or a live birth was not significant.
- No studies were identified which reported on how the prognostic information is being used in the target population, i.e. no evidence on how AMH results impact on the management of women undergoing gonadotoxic treatment.
- In women undergoing IVF, but not the target population, there is evidence that AMH values may influence the starting dosage of recombinant follicle-stimulating hormone or human menopausal gonadotropin during ovarian hyperstimulation for the retrieval of oocytes. However, the studies showed a lack of standardisation regarding how the AMH score was used or should be used in this broader population.
- Given the lack of incremental outcome data, only a cost-analysis could be undertaken for the economic evaluation.
- The major challenge for clinicians attempting to interpret AMH values for use in clinical care is the lack of standardisation. It is recommended that clinicians should always use the same laboratory to avoid problems with result interpretation. It is also critical to understand how that laboratory calibrates their clinical thresholds to ensure accurate interpretation of the result. A good quality assurance program would be essential for implementation and development of standardised clinically relevant thresholds for AMH testing in Australia.

ANTI-MÜLLERIAN HORMONE TESTING FOR FEMALE PATIENTS PRECEDING OR FOLLOWING GONADOTOXIC TREATMENT

This contracted assessment examines the evidence to support the listing of anti-Müllerian hormone (AMH) testing on the Medicare Benefits Schedule (MBS). The service would be exclusively used in women who have had or will have gonadotoxic treatment. The applicant has claimed that the successful listing of the technology in the target population may lead to better assessments of

ovarian function prior to cancer treatment, enable better prediction of the return of reproductive function following gonadotoxic treatment, and/or improve decision-making regarding the need for fertility preservation following or prior to gonadotoxic treatment.

ALIGNMENT WITH AGREED PICO CONFIRMATION

This contracted assessment of AMH testing addresses most of the PICO¹ elements that were pre-specified in the PICO Confirmation that was approved the PICO Advisory Sub-Committee (PASC). Due to lack of evidence in some of the steps of the linked analysis, a modified approach has been presented in the results section. As evidence was limited, especially regarding young girls, data could not be separated into subgroups. Where possible, studies were divided in a pre-treatment AMH testing group and a post-treatment AMH testing group.

PROPOSED MEDICAL SERVICE

Ovarian failure is associated with the absence of primordial follicles in the ovarian cortex, leading to a lack of maturing follicles, which is linked to menopause (Amir et al. 2010). AMH is a glycoprotein believed to represent non-cyclical, continuous primordial follicle growth and is an indirect measure of the resting ovarian follicle pool, i.e. the ovarian reserve (Bozza et al. 2014). AMH is present in blood and can be measured from birth until menopause.

In women undergoing gonadotoxic therapy or surgery, the primordial follicles in the ovaries can be damaged, which may lead to ovarian failure, infertility and early menopause. Usually AMH levels drop during gonadotoxic treatment, with the possibility of some recovery after finishing treatment. This is dependent on age, dose and type of treatment (Bozza et al. 2014).

A series of biochemical and ultrasonographic tests have been developed as indirect markers to measure ovarian reserve, such as AMH, follicle-stimulating hormone (FSH), estradiol (E2), inhibin B, antral follicle count (AFC), and measurement of ovarian volume, as ovarian reserve cannot be measured directly.

Different commercial assays are used in clinical practice to measure serum AMH levels. Scores of different assays are not necessarily comparable, so when interpreting AMH values, clinicians should be aware of which AMH assay was used.

AMH tests are currently paid for out-of-pocket by the patient.

¹ Population, Intervention/Investigation/Index test, Comparator, Outcomes

PROPOSAL FOR PUBLIC FUNDING

The scheduled fee for an AMH test as proposed by the applicant is \$100, which is higher than the amount currently charged by providers offering the service in the private sector. The proposed item descriptor is shown below in Table 1.

Table 1 Proposed MBS item descriptor

Category 6 - Pathology services
MBS [item number]
Anti-Müllerian Hormone (AMH) for female patients who will or have received gonadotoxic treatment.
Fee: \$100
Explanatory notes: <ul style="list-style-type: none">- Diagnosis requiring treatment with gonadotoxic therapy.- Service to be provided by a suitably trained and accredited fertility specialist.- Female patient aged 0–45 years of age will eligible for this service.- Maximum of one AMH test in female prior to initial or relapse treatment with gonadotoxic treatment for malignant or non-malignant disease.- Assessment of ovarian reserve with a maximum of one AMH test per year following treatment in patients who have had gonadotoxic treatment to assess the gonadotoxic effects of treatment, to assess pubertal delay, to assess ovarian failure, to assess the need for fertility preservation following treatment and to assess the need for assisted reproductive treatment for family planning.

POPULATION

The proposed population is female pre-menopausal patients who will be having, or have had, gonadotoxic treatment. Gonadotoxic treatment includes any treatment which is associated with a risk of ovarian damage or sterility. This includes treatment for malignancy, e.g. chemotherapy, irradiation, as well as treatment for precancerous or benign conditions, e.g. pelvic surgery.

In Australia, the estimated number of new cases of cancer to be diagnosed in 2017 among women aged under 45 years is 6,520. It was estimated that 22% of women indicated for AMH testing have a non-malignant condition. Information on estimated uptake of AMH testing in Australia among females seeking specialist fertility with non-malignant conditions requiring gonadotoxic treatment was unavailable. However, the available evidence suggests that uptake of AMH testing among female oncology patients up to 44 years, currently about 50%, would increase to 90% if the service were made available without out-of-pocket costs to patients.

COMPARATOR DETAILS

AMH testing would be done in addition to the current standard tests for measuring ovarian reserve. Ovarian reserve tests include both biochemical tests and ovary ultrasound imaging. The PICO Confirmation listed basal FSH, E2 and inhibin B measurement and AFC as valid comparators. The measurement of inhibin B is currently not listed on the MBS, however FSH and E2 tests are listed under MBS number 66695. AFC using ultrasound is listed under MBS item number 55065.

CLINICAL MANAGEMENT ALGORITHM(S)

Figure 2 shows the current and proposed clinical management algorithm for females undergoing testing prior to gonadotoxic treatment. The current and proposed clinical management algorithms for females tested after undergoing gonadotoxic treatment are shown in Figure 3 and Figure 4, respectively. In most cases, AMH will be measured in addition to ovarian reserve tests already done in clinical practice (mainly FSH, E2 and AFC).

CLINICAL CLAIM

The final PICO Confirmation did not state a clear clinical claim. Based on the available information in the PICO Confirmations, the evidence and the public consultation feedback, the assessment group identified several uses for AMH testing.

Prior to gonadotoxic treatment, an AMH test is used to provide prognostic information about risk of infertility, thereby informing the decision whether to undergo fertility preservation. Or, in other words, whether fertility is likely to be sufficient after gonadotoxic therapy such that fertility preservation is unnecessary, versus a scenario requiring that patients are counselled that they are likely to have a low number of oocytes, and that the chance of conception is low. Pre-gonadotoxic treatment, an AMH test may also be used to predict the response to superovulation, thereby enabling the yield of good quality oocytes to be maximised while reducing the risk of ovarian hyperstimulation by adjusting starting doses of rFSH.

Following gonadotoxic treatment, an AMH test is used to monitor the level of ovarian function. Monitoring ovarian function may inform whether patients are candidates for natural conception or may need assisted reproductive therapy (ART), using previously cryopreserved ovarian tissue, oocytes or embryos to conceive. Other claimed benefits of monitoring ovarian function include the ability to guide HRT based on information on ovarian reserve (i.e. is there premature ovarian insufficiency), to diagnose ovarian failure, and to assist with counselling and supportive care of early menopause.

APPROACH TAKEN TO THE EVIDENCE ASSESSMENT

A systematic review of published and unpublished literature was undertaken. The medical literature was searched on the 6th of June 2017 to identify relevant studies and systematic reviews published during the period 1990 to June 2017. Searches were conducted of the databases and sources as per Appendix B. Attempts were also made to source unpublished or grey literature.

Studies were selected by a single reviewer with a second reviewer assessing 10% of the most relevant citations, as determined by the algorithm within Rayyan. Appraisal of the evidence was conducted in four stages: (1) appraisal of the risk of bias within individual studies included in the review; (2) appraisal of the precision, size of effect and clinical importance of the results reported in the evidence base as they relate to the pre-specified primary outcomes for this assessment; (3)

rating the overall quality of the evidence per outcome, across studies, based on the study limitations (risk of bias), imprecision, inconsistency of results, indirectness of evidence, and the likelihood of publication bias, and; (4) integration of this evidence across outcomes for conclusions about the net clinical benefit of the test in the context of Australian clinical practice.

No studies meeting the PICO criteria for direct evidence were identified. Therefore, a linked evidence approach was used to evaluate the evidence.

CHARACTERISTICS OF THE EVIDENCE BASE

The number of studies included in each step of the linked evidence is shown in Table 2. Most identified studies enrolled patients with malignancies, with the most common form of cancer being breast cancer. Some studies enrolled patients with endometriosis undergoing surgery.

No evidence was found in the target population on therapeutic efficacy or therapeutic effectiveness. The evidence presented in these sections did not match the proposed MBS populations and was therefore not included in Table 2. Most studies did not evaluate the incremental value of AMH testing in addition to tests currently available in current clinical practice. Due to the indirectness of most results, most of the evidence was considered very low quality using the GRADE system.

Table 2 Key features of the included linked evidence

Type of evidence	Description	Evidence base
Analytical validity (Section B3)	Two studies compared different AMH assays (correlation). Two studies reported the accuracy of AMH compared to other tests with ovarian failure as the reference standard. Twelve studies reported concordance / correlation data of the AMH test results with one or more comparator tests.	k=14 n=914
Prognostic and predictive evidence (Section B4.1.5)	Eleven studies reported on the prognostic value of AMH testing in predicting ovarian failure in the target population. Six studies reported on the prognostic value of AMH testing in predicting pregnancy or live births in the target population. Four studies reported on the predictive value of AMH testing in predicting ovarian response to hyperstimulation in the target population.	k=21 n=1,760
Clinical validity (Section B4.1.5)	Three studies reported on the clinical validity of AMH testing prior to gonadotoxic treatment in predicting the resumption of menses after treatment. Four studies reported on the clinical validity of AMH testing after gonadotoxic treatment in determining the resumption of menses after treatment.	k=7 n=310
Therapeutic efficacy (Section B5.1)	No evidence was found on AMH informing change in management in women undergoing gonadotoxic treatment.	k=0 n=0
Therapeutic effectiveness (Section B5.2)	No evidence was found to determine how change in management due to AMH testing impacts health outcomes in women undergoing gonadotoxic treatment.	k=0 n=0

^a Reference standard available

^b Reference standard not available

Due to evidence gaps in the last two steps of the linked analysis (Therapeutic efficacy and therapeutic effectiveness) there is uncertainty when estimating incremental effectiveness of AMH testing.

RESULTS

On the basis of the benefits and harms reported in the evidence base it is suggested that, relative to other ovarian reserve tests alone, the AMH test and associated interventions has **non-inferior safety and uncertain incremental effectiveness**.

Safety

Test adverse events

No studies meeting the PICO criteria regarding safety of AMH testing were identified. However, the test is done through a routine blood test and this is generally considered safe.

Effectiveness

Direct effectiveness

No direct evidence was identified to determine the effectiveness of AMH testing in addition to other standard tests, compared to other standard tests alone, in patients prior to or following completion of gonadotoxic therapy.

Effectiveness from linked evidence

1. Analytical validity

The information on the analytical validity of the AMH test focuses on whether the test results are reliable, and how they correlate with other tests which are currently in clinical practice. Studies which used menstrual status as the reference standard found that the accuracy of AMH testing varied substantially between different assays, and that about one in three women who resumed menstruation after treatment had undetectable AMH, i.e. a negative result.

One study reported on the diagnostic accuracy of the pico-AMH ELISA (enzyme-linked immunosorbent assay test and the EIA AMH/MIS assay compared to menstrual status, i.e. the absence or presence of menses, as a clinical reference standard. The sensitivity of the pico-AMH ELISA was significantly higher (71% compared to 11%) than the older EIA AMH/MIS assay (with detectable AMH as the cut-off). The large difference in sensitivity between the two AMH tests is likely due to the 40-fold difference in the level of AMH detectable by these tests with the pico-AMH ELISA test being more sensitive.

Two studies compared the AMH Gen II ELISA test with the Ansh Labs AMH ELISA, Ultrasensitive AMH ELISA, and pico-AMH ELISA, as well as the Ultrasensitive AMH ELISA with the pico-AMH ELISA. There was a high degree of correlation between these four tests. However, there is concern that although

different AMH assays are highly correlated, the results are not generalisable. One study showed that while each laboratory showed good reproducibility when using a single test, the between-laboratory variability, even using the same assay, was 40%.

The major challenge for clinicians attempting to interpret AMH values for use in clinical care is the lack of standardisation. It is recommended that clinicians should always use the same laboratory to avoid problems with result interpretation. It is also critical to understand how that laboratory calibrates their clinical thresholds to ensure accurate interpretation of the result. A good quality assurance program would be essential for implementation and development of standardised clinically relevant thresholds for AMH testing in Australia.

ROC analysis comparing AMH serum levels and AFC with the menstrual status of women after chemotherapy found that the area under the curve (AUC) for the first generation AMH assay, the Diagnostics Systems Lab (DSL) ACTIVE® AMH/MIS ELISA and AFC were 0.71 and 0.82, respectively. This means the DSL ACTIVE® AMH/MIS ELISA test showed moderate performance in detecting the menstrual status of women, compared with a good performance for AFC, the reference standard. The AUCs for the comparators FSH and inhibin B were 0.72 (moderate performance) and 0.63 (poor performance), respectively.

2. Prognostic and predictive value

Prognostic value of AMH in predicting ovarian function - Pre-treatment AMH testing

AMH testing was considered to be mainly of prognostic and predictive value. Eleven studies were identified which considered whether AMH values could predict ovarian function at follow-up in the target population. These studies included mostly women with breast cancer. Six out of seven studies that reported mean or median baseline AMH values stated that the group of women with chemotherapy-related amenorrhea (CRA) at follow-up had statistically significantly lower baseline AMH levels, i.e. levels measured before treatment. Follow-up was six months to five years after enrolment or treatment. Higher mean age and lower AFC count were also associated with CRA at follow-up (in 4/4 studies and 2/2 studies, respectively). The other comparator tests were less convincing (FSH was different in 2/4 studies; E2 in 1/2 studies; inhibin B in 1/3 studies).

Univariate analysis alone is insufficient to determine the incremental prognostic value of AMH testing. AMH levels decrease with age and correlate with other hormone tests. In the multivariate analyses, AMH remained a significant predictor for continuation or resumption of ovarian function in four out of five studies. The largest study reported that women with an AMH above 0.7 ng/mL had three times greater probability of a shorter time to ovarian recovery than women with an AMH under 0.7 ng/mL over a median of 163 days. FSH (≤ 10 IU/L) and age (< 40 years) were also predictive for shorter time to ovarian recovery, with hazard ratios of 4.7 (95%CI 1.3, 16.8) and 3.39 (95%CI 1.74, 6.60), respectively.

Two studies were included that presented a ROC curve to assess the accuracy of AMH for predicting ovarian function and to determine the optimal cut-off point for AMH values. Both studies reported an AUC for AMH of >0.8 which indicates good test performance. One study showed AMH values ≥ 3.32 ng/mL were protective for the occurrence of oligomenorrhea or amenorrhea after treatment, with a sensitivity of 85% and a specificity of 75% (n=52). The cut-off for AFC was <13 follicles to predict oligo- or amenorrhea, having a higher risk at six-month follow-up with a sensitivity of 83% and a specificity of 62%. In the study by Anderson et al., a classification mosaic chart including age (38.6 years as a cut-off) and AMH score (3.8 pmol/L as a cut-off) gave a 98.2% and a specificity of 80.0% in predicting amenorrhea after two years for breast cancer patients (n=75).

The lack of standardisation makes it hard to interpret AMH values. Determining standardised clinically relevant thresholds for AMH testing in Australia may improve the prognostic value of the test in clinical practice.

Prognostic value of AMH in predicting ovarian function - Post-treatment AMH testing

There is limited evidence available to determine the prognostic value of post-treatment AMH testing in the patient population (k=2). The incremental value of AMH testing could not be determined. However, the limited evidence does indicate that a woman with detectable AMH, measured by pico-AMH ELISA post-chemotherapy, is very likely to have ongoing menses for at least three years, whereas a woman with undetectable AMH is likely to continue to have amenorrhoea due to the lack of ovarian reserve.

Prognostic value of AMH in predicting pregnancy

Most studies (5/6) did not find a predictive relationship between AMH levels and pregnancy rate. All women underwent a gonadotoxic treatment for endometriosis, breast cancer or lymphoma, and wanted to fall pregnant. Follow-up was >1 year in all studies. Five studies reported AMH thresholds and showed how many women achieved pregnancy when divided into groups based on AMH score. Only one small study found a lower pregnancy rate in the group with low AMH levels.

These studies showed that the difference between those with low and high AMH is not large enough to inform clinical decisions. Even women with undetectable and critically low AMH values had given birth during follow-up in multiple studies. This raised the hypothesis that the relationship between low AMH levels and reproductive outcomes may differ in cancer survivors compared to women from the general population (Hamre et al. 2012). However the study populations may have been too small to find a relationship between pregnancy and AMH levels.

Predictive value of AMH in predicting response to ovarian stimulation

One study including 340 cancer patients investigated the association between the number of in vitro matured oocytes cryopreserved and AFC and AMH levels using ROC curve analysis. In this study, the

AMH test had moderate (0.7 to 0.8) to good (0.8 to 0.9) test performance at predicting oocyte yield. AFC ultrasound had a good (0.8 to 0.9) test performance.

Three studies looked at the correlation between oocytes retrieved and AMH, AFC, inhibin B and/or E2 values. A moderate linear relationship between AMH and oocyte yield was found. Moderate relationships were also found for AFC and inhibin B with oocyte yield in one study. An inverse correlation was found for FSH and age with oocyte yield.

An association was found between AMH levels and ovarian response to hyperstimulation in women undergoing gonadotoxic treatment. However, there was insufficient evidence to determine whether the AMH test had an incremental predictive value in addition to other standard tests to determine response to ovarian stimulation. When looking at the broader population of otherwise healthy women undergoing in vitro fertilisation (IVF) or intracytoplasmic sperm injection (ICSI), Broer et al. stated that combining the results of AFC and AMH did not improve prediction of poor response. This supports the findings by the Practice Committee of the American Society for Reproductive Medicine in 2015, which stated that combined ovarian reserve test models do not consistently improve predictive ability over that of single ovarian reserve tests, and that there is insufficient evidence to indicate that a combination of tests are more useful than each test alone in ovarian reserve testing in a broad population. Tobler et al. (2015) stated that the predictive value of AMH testing on the outcomes of ART has yet to be definitively established, and the exact role of the use of AMH testing for fertility treatments still needs to be determined. There is a need for AMH cut-off values related to patient relevant outcomes, e.g. pregnancy, oocyte yield, and return of menstruation, as evidence on these outcomes is still lacking. However, it is considered a first-line test of ovarian reserve in Australia and is reported to have direct value in the management of IVF cycles in the general population (Tobler et al. 2015).

3. Clinical validity

The evidence that was included in the analytical validity and the prognostic section and provided sensitivity/specificity or 2 × 2 data was included in the clinical validity section to determine how the test would perform in women undergoing gonadotoxic treatment of different levels of gonadotoxicity. This gives an indication of the usefulness of the test when it is used before or after treatments with low risk (20–30%), intermediate risk (40–70%) and high risk (70–80%) of ovarian failure.

Three studies reported on the accuracy of AMH testing prior to gonadotoxic treatment in predicting the resumption of menses after treatment. First generation test EIA AMH/MIS ASSAY and DSL ACTIVE® MIS/AMH ELISA showed some lack of utility for AMH measured both prior to and after treatment compared to resumption of menses. A positive pre-treatment or post-treatment test result, above a certain threshold, only offered useful information to women having treatment associated with intermediate to high risk of ovarian failure. At least 76% of women with a ≥50% risk

of ovarian failure receiving a positive pre-treatment AMH result and 79 to 94% of women with a positive post-treatment AMH test result who underwent high-risk treatment actually resumed menses. On the other hand, a negative pre- or post-treatment test result only offers useful information to women having low-risk treatment, where the negative predictive value (NPV) indicated that 88 to 95% and 71 to 88% of women with a negative pre- and post-treatment test result would not have resumption of menses, respectively. Of women having treatment with a high risk of ovarian failure, 44 to 68% and 21 to 45% with a negative pre- or post-treatment test result would have amenorrhea at follow-up (after treatment), respectively. This is equivalent to chance and not clinically useful.

AMH Gen II ELISA and the pico-AMH ELISA tests performed much better. In women undergoing treatment with a low risk of ovarian failure, 63 to 74% with a detectable baseline AMH Gen II ELISA test result would have resumption of menses, and 94 to 96% of women having treatment with a high risk of ovarian failure who have a detectable AMH level would have resumption of menses after treatment. The NPV indicated that 81 to 99% of women undergoing treatment with a low risk of ovarian failure and an undetectable AMH level would not have resumption of their menses (Henry et al. 2014). The pico-AMH ELISA showed that around 90% of women who had a positive post-treatment test result and underwent high-risk treatment would have resumption of menses, with only 47 to 64% having a positive test in women undergoing low risk treatment. A negative test result was only meaningful in the low risk treatment group, with 87 to 97% of women with a negative post-treatment AMH result having amenorrhea. Therefore, the positive test result was only clinically useful at higher prevalence rates of ovarian failure, whereas the negative test result was only useful in the group at low risk of ovarian failure.

4. Therapeutic efficacy (change in management)

The research questions on clinical utility could not be answered, due to lack of evidence on the impact of AMH on change in clinical management among women undergoing gonadotoxic treatment. The only evidence of impact of AMH testing on change in clinical management was found in a broader population, i.e. healthy women undergoing IVF.

In women undergoing IVF, there is evidence that AMH values may influence the starting dosage of recombinant FSH (rFSH) or hMG during ovarian hyperstimulation for the retrieval of oocytes. However, the studies showed a lack of standardisation or guidelines on how the AMH score was used or should be used.

5. Therapeutic effectiveness (health benefit from change in management)

No evidence was found to determine how the change in management due to AMH testing impacts health-related outcomes in women undergoing gonadotoxic treatment, the target population. Due to lack of evidence, a non-systematic search was performed to investigate whether individualised

dosage of rFSH/hMG based on AMH results lead to better health outcomes in women undergoing IVF/ICSI.

Three studies were identified which investigated whether an individualised starting dosage based on AMH levels impacted IVF outcomes, and all three studies measured a higher rate of optimal oocyte yield in the individualised dosage group. However, only one retrospective study showed a difference in pregnancy or birth rates between groups, and it was not possible to determine which factors were contributing to these outcomes. This was the case due to the before- and after design of the study and the introduction of a number of organisational and procedural changes alongside the introduction of AMH-tailored dosing. The two clinical trials did not find a difference in pregnancy or birth rates between standard dosage and individualised dosage informed by AMH.

Even though there was insufficient effectiveness evidence regarding pregnancy and birth outcomes, two out of the three studies indicated that there was a significant reduction in the incidence of ovarian hyperstimulation syndrome (OHSS) and/or preventive interventions for OHSS using the AMH-tailored dosage protocol, compared to a standardised dosage in a population of healthy women undergoing IVF. One of these trials compared a standard dose of follitropin alpha with an AMH and bodyweight-tailored dose of follitropin delta (Nyboe Andersen et al. 2017). It is unknown whether the outcome differences between groups in this trial were due to the individualised dosage or due to the different follitropin used.

Notably, the studies discussed in the clinical utility section did not meet the PICO criteria and mostly excluded patients at the extremes of ovarian reserve. The nomogram that was used in two of the studies to determine dosage only gives an input for women aged 25 to 40 years old and may not be very useful if the AMH levels are extremely low. This nomogram has not been validated in some of the proposed MBS populations (women <25 years old and/or with low or undetectable AMH), and therefore the generalisability of this evidence is unknown.

SUMMARY OF FINDINGS

A summary of the findings and the quality of the evidence is presented in Table 60.

TRANSLATION ISSUES

No clinical evidence for the incremental benefit of AMH as an additional test over the current practice was identified, therefore effectiveness was not modelled, and no translation studies were necessary.

ECONOMIC EVALUATION

On the basis of the conclusion of the clinical evaluation that relative to other ovarian reserve tests alone, additional AMH testing has **non-inferior safety and uncertain incremental effectiveness**,

economic modelling of outcomes was not appropriate. Therefore a cost-analysis was conducted for the economic evaluation.

A summary of the key characteristics of the economic evaluation is provided in Table 3.

Table 3 Summary of the economic evaluation

Perspective	Australian health care
Comparator	AFC, FSH and E2 tests
Type of economic evaluation	Cost-analysis
Sources of evidence	Systematic review
Outcomes	No health outcomes, cost per patient estimated only
Methods used to generate results	Investigative pathway cost-comparison
Software packages used	Microsoft Excel 2013

AFC = antral follicle count; E2 = estradiol; FSH = follicle-stimulating hormone

The overall costs and incremental costs as calculated for the proposed use of AMH and the comparator are shown in Table 4. The use of the proposed AMH test would be expected to increase the cost of current practice by \$100 per patient, if listed at the proposed fee.

Table 4 Costs associated with testing ovarian reserve, and incremental cost (per patient)

	AMH + current practice^a	Current practice
Specialist consultations for referral and review	\$128.55	\$128.55
Test costs	\$266.74	\$166.74
Total cost per patient	\$395.29	\$295.29
Incremental cost per patient		\$100.00

AMH = anti-Müllerian hormone; AFC = antral follicle count; E2 = estradiol; FSH = follicle-stimulating hormone

^a AFC and FSH+E2 measurements are considered as current practice for ovarian reserve testing and the intervention includes all these tests and AMH.

The base case analysis assumes AMH is used as an additional test to the current practice (AFC and FSH+E2). However, alternative scenarios are costed where AMH replaces one of the existing tests used to estimate ovarian reserve. If AMH were to replace AFC in the current practice it would result in a cost saving of \$17 per patient compared to existing practice. However, if AMH were used to replace FSH+E2, i.e. the proposed intervention being use of AMH and AFC, it would result in a net cost increase.

ESTIMATED EXTENT OF USE AND FINANCIAL IMPLICATIONS

An epidemiological approach has been used to estimate the eligible population. Uptake rates, and the use of multiple tests per patient are also incorporated into the overall estimates of the financial implications of the proposed MBS funding of AMH testing to inform fertility management in female patients preceding or following gonadotoxic treatment.

The applicant has proposed the scheduled fee for an AMH test is \$100. This is higher than the fee currently charged by many providers offering the service to privately funded patients, ranging from \$55 to \$98.

Table 5 summarises the estimated costs of AMH testing to the MBS, i.e. government expenditure, and patients, i.e. out-of-pocket expenditure. The base case analysis estimates that AMH testing will cost approximately \$570,000 to the MBS each year.

Table 5 Estimated costs of AMH testing, 2018–19 to 2022–23

	2018–2019	2019–2020	2020–2021	2021–2022	2022–2023
Projected number of AMH tests	6,682	6,706	6,730	6,754	6,779
Cost of AMH to the MBS	\$567,961	\$570,005	\$572,057	\$574,117	\$576,184
Cost of AMH to the patients	\$100,228	\$100,589	\$100,951	\$101,315	\$101,679
Total cost of AMH test	\$668,189	\$670,595	\$673,009	\$675,432	\$677,863

AMH = anti-Müllerian hormone; MBS = Medicare Benefits Schedule

Sensitivity analyses around referral and uptake rates, restricted population size, potential leakage and proposed intervention as a replacement test are presented in Section E.6. Lower referral/uptake rates and restricting the target population to women aged 25 to 39 years decreases estimated costs to the MBS. Cost impact to the MBS due to potential leakage is very high.

CONSUMER IMPACT SUMMARY

During the public consultation period, before the PICO Confirmation for application 1434 was finalised, 78 responses were received. The responses were mainly supportive of AMH testing. A number of advantages and disadvantages, or issues that would need to be addressed, were emphasised during the public consultation period.

The benefits suggested via public consultation were that:

- AMH testing may provide vital information to young patients with regard to their fertility status, allowing them to make informed decisions about potential fertility interventions
- AMH tests can measure the risk for premature ovarian failure, enabling possibility of early intervention (hormone replacement therapy or oocyte cryopreservation)
- AMH testing could potentially reduce costs by allowing targeted fertility interventions
- listing AMH testing on the MBS would lead to more equitable fertility monitoring
- listing AMH testing on the MBS could lead to reduced anxiety and improved psychological wellbeing via better understanding around potential fertility function in the future and in relation to medical costs.

The disadvantages, or issues suggested for address, as determined by public consultation were that:

- AMH levels don't predict the quality of remaining oocytes
- AMH test results may cause psychological distress for the patient (negative/undetectable results)
- pre-test counselling would be critical to minimise the negative psychological impact of an unexpected or negative result.

OTHER RELEVANT CONSIDERATIONS

In women attending a fertility clinic in Ireland, a population not representative of the target population, the awareness of the clinical relevance of AMH testing was low. Women with low AMH levels reported feelings of devastation, isolation and loss of femininity and purpose. Girls and women receiving a low test result could become unnecessarily anxious about their fertility, and as the results of the test are hard to interpret, receiving proper counselling is important. Regardless of the test results, the women indicated that having information about their ovarian reserve was important and that it impacted their decision-making. However, the current priorities of women undergoing gonadotoxic treatment may be different from women undergoing fertility treatments, and no evidence was found on the psychological impact of AMH testing on the target population.

ACRONYMS AND ABBREVIATIONS

AFC	antral follicle count
AHTA	Adelaide Health Technology Assessment
AI	aromatase inhibitor
AIHW	Australian Institute of Health and Welfare
AMH	anti-Müllerian hormone
ANOVA	analysis of variance
ART	assisted reproductive technology
ARTG	Australian Register of Therapeutic Goods
ASCO	American Society of Clinical Oncology
AUC	area under the curve
BMT	bone marrow transplant
CI	confidence interval
CRA	chemotherapy-related amenorrhea
CoV	coefficient(s) of variation
E2	estradiol
EIA	enzyme immunoassay
ELISA	enzyme-linked immunosorbent assay
FSH	follicle-stimulating hormone
FP	false positive
FPR	false-positive rate
GnRH	gonadotropin-releasing hormone
HCT	haemopoietic cell transplant
HESP	Health Expert Standing Panel
HL	Hodgkin lymphoma
hMG	human menopausal gonadotropin, also called menotropin

HR	hazard ratio
HRT	hormone replacement therapy
HRQoL	health-related quality of life
HSCT	haematopoietic stem cell transplant
HTA	health technology assessment
ICER	incremental cost-effectiveness ratio
ICSI	intracytoplasmic sperm injection
IVF	in vitro fertilisation
LH	luteinising hormone
LoD	limit of detection
LoQ	limit of quantification
LR	likelihood ratio
MBS	Medicare Benefits Schedule
MD	mean difference
MII	mature metaphase II (oocyte)
MIS	Müllerian-inhibiting substance
MSAC	Medical Services Advisory Committee
NA	not applicable
NHMRC	National Health and Medical Research Council
NPV	negative predictive value
NR	not reported
NS	not significant
OHSS	ovarian hyperstimulation syndrome
OR	odds ratio
OTC	ovarian tissue cryopreservation
PASC	PICO Confirmation Advisory Sub-Committee of the MSAC
PICO	Population, Intervention/Investigation/Index test, Comparator(s), Outcome(s)

PPV	positive predictive value
QALY	quality-adjusted life year
QUIPS	Quality in Prognosis Studies
RCT	randomised controlled trial
rFSH	recombinant follicle-stimulating hormone
ROC	receiver operator characteristic
RS	reference standard
SD	standard deviation
SROC	summary receiving operating characteristic
TBI	total body irradiation
TGA	Therapeutic Goods Administration
TP	true positive
TPR	true-positive rate

This is a contracted assessment of anti-Müllerian hormone (AMH) testing for female patients preceding or following to gonadotoxic treatment to assess the need for fertility preservation or to determine the return of reproductive function following treatment. This assessment is intended for the Medical Services Advisory Committee (MSAC). MSAC evaluates new and existing health technologies and procedures for which funding is sought under the Medicare Benefits Schedule (MBS) in terms of their safety, effectiveness and cost-effectiveness, while taking into account other issues such as access and equity. MSAC adopts an evidence-based approach to assessment based on reviews of the scientific literature and other information sources including clinical expertise.

Adelaide Health Technology Assessment (AHTA) has been commissioned by the Australian Government Department of Health to conduct a systematic literature review and economic evaluation of AMH testing for female patients undergoing gonadotoxic treatment. This assessment has been undertaken in order to inform MSAC's decision whether the proposed medical service should be publicly funded.

Appendix A lists the people involved in the development of this assessment report, including the applicants and clinical experts.

The proposed use of AMH testing in Australian clinical practice was outlined in a Protocol, now a PICO Confirmation, presented to and accepted by the former Protocol Advisory Sub-Committee. This committee is now known as the PICO Confirmation Advisory Sub-Committee (PASC). The PICO Confirmation was released for public comment in June 2016 and ratified out-of-session in February 2017.

A1 ITEMS IN THE AGREED PICO CONFIRMATION

This contracted assessment of AMH testing addresses most of the PICO elements that were pre-specified in the PICO Confirmation approved by PASC. The PICO criteria were amended to fit a linked analysis approach.

Due to lack of evidence in some steps of the linked analysis, an alternative approach has been presented in the results section. The final PICO Confirmation approved by PASC divided the population between females aged 0 to 14 years, 15 to 25 years and 26 to 45 years. As evidence was limited (especially regarding young girls), data could not be separated in subgroups. Where possible, studies were divided in a pre-treatment AMH testing group and a post-treatment AMH testing group.

FERTILITY AND THE OVARIAN FOLLICLE POOL

When follicles are recruited from the ovarian primordial follicle pool and move into the growing follicle pool they start to express AMH and inhibin B. AMH is considered an indirect index for the number of follicles in the resting primordial pool (Bozza et al. 2014). In post-pubertal girls and women, a limited number of follicles are selected from the growing follicle pool in every menstrual cycle, influenced by FSH. This is called cyclic recruitment. One follicle is then selected from the growing follicle pool and ovulates under the influence of luteinising hormone (LH) (Visser et al. 2012). The selected ovarian follicles which do not ovulate during the cycle will degenerate in a process called atresia. FSH levels need to increase to a critical threshold concentration to save follicles from atresia and to allow FSH-dependent selection of a limited number of follicles. Throughout reproductive life, the number of primordial and growing follicles will decline. Subsequently, serum inhibin B and estradiol (E2) concentrations decline, which leads to a rise in FSH (and LH) levels in menopause (Figure 1).

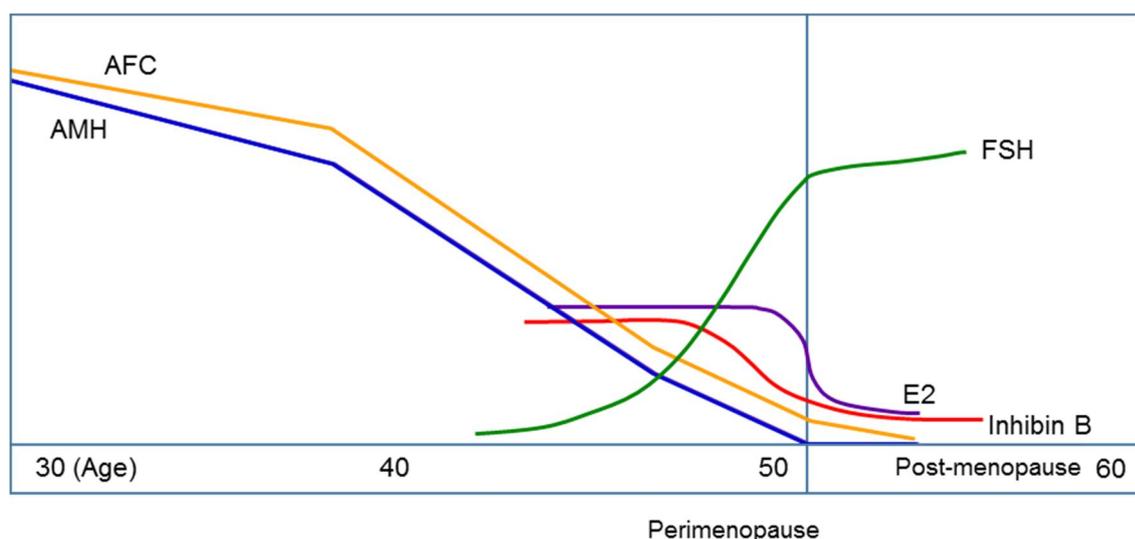


Figure 1 Changes in hormone levels and AFC with decreasing ovarian reserve approaching menopause and post-menopause.

The blue area depicts ovarian reserve. The lines for AMH, FSH, E2 and inhibin B reflect the change in serum levels of these hormones between ages 30 and 60 years. The line for AFC reflects the decrease in follicle counts between ages 30 and 60 years. AFC = antral follicle count, AMH = anti-Müllerian hormone, E2 = estradiol, FSH = follicle-stimulating hormone

OVARIAN FAILURE AND AMH TESTING

Ovarian failure is associated with the absence of primordial follicles in the ovarian cortex, resulting in an absence of maturing follicles, which is linked to menopause (Amir et al. 2010). AMH is a glycoprotein of the transforming growth factor β family, believed to represent the non-cyclical, continuous primordial follicle growth. This means it can be considered an indirect quantifier of the resting ovarian follicle pool, i.e. the ovarian reserve (Bozza et al. 2014).

AMH is produced by granulosa cells of primary, preantral and small antral follicles, and is subsequently released in the circulation (Bozza et al. 2014; Iwase et al. 2014). AMH levels can be measured from birth until menopause. The hormone protects primordial follicles by slowing their rapid recruitment, preventing premature follicle pool depletion (Bozza et al. 2014).

In women undergoing gonadotoxic therapy or surgery, damage can occur in the primordial follicles in the ovaries, which can lead to ovarian failure, infertility and the early onset of menopause. Usually AMH levels drop during gonadotoxic treatment, with the possibility of partial recovery after finishing treatment. However, the level of recovery is dependent on age, type of treatment and dosage (Bozza et al. 2014) (see Section A4).

Measuring AMH and other hormones

Ovarian reserve cannot be measured directly. Therefore, a series of markers and ultrasonographic tests have been adopted as indirect markers. These include AMH, FSH, estradiol (E2), inhibin B, AFC and measurement of ovarian volume. Measuring AMH is gaining popularity as AMH is very sensitive to changes with advancing age and it is relatively consistent throughout the menstrual cycle (Iwase et al. 2014). FSH, inhibin B and E2 show cyclical fluctuations.

Different commercial assays are used in clinical practice to measure serum AMH. Prior to 2011, two AMH assays were available: DSL and Immunotech (IOT) (Tobler et al. 2015). These assays use different primary antibodies against AMH and different standards; therefore crude values reported can differ substantially (Nelson & La Marca 2011). A second generation assay is now available (Beckman Coulter AMH Gen II). Interpreting AMH values requires awareness that scoring is not necessarily comparable across the different assays; therefore, it is important to know which AMH assay has generated the score.

A3 PROPOSAL FOR PUBLIC FUNDING

The proposed MBS item descriptor was presented in the PICO Confirmation submitted by the applicant and is summarised in Table 6. The applicant has proposed that the scheduled fee for an AMH test should be \$100. This is higher than the fee currently charged by many providers offering the service to privately funded patients.²

² See websites (access date: 2 November 2017): IVF Australia (<https://www.ivf.com.au/ovarian-reserve-amh-test/>), \$80; Fertility North (<http://www.fertilitynorth.com.au/amh-test/>), \$55; Repromed (<http://repromed.com.au/what-to-expect/preliminary-investigations/amh-blood-test/>), \$98; Clinpath Pathology ([http://www.clinipathpathology.com.au/media/96085/anti-mullerian%20hormone%20\(amh\).pdf](http://www.clinipathpathology.com.au/media/96085/anti-mullerian%20hormone%20(amh).pdf)), \$60; i-screen (<https://www.i-screen.com.au/app/register/amh-test>), \$85.

Table 6 Proposed MBS item descriptor

Category 6 - Pathology services
MBS [item number]
Anti-Müllerian hormone (AMH) for female patients who will or have received gonadotoxic treatment.
Fee: \$100
Explanatory notes: Diagnosis requiring treatment with gonadotoxic therapy Female patient aged 0–45 years of age will eligible for this service. Maximum of one AMH test in female prior to initial or relapse treatment with gonadotoxic treatment for malignant or non-malignant disease. Assessment of ovarian reserve with a maximum of one AMH test per year following treatment in patients who have had gonadotoxic treatment to assess the gonadotoxic effects of treatment, to assess pubertal delay, to assess ovarian failure, to assess the need for fertility preservation following treatment and to assess the need for assisted reproductive treatment for family planning.

A4 PROPOSED POPULATION

PATIENTS PRIOR TO, OR AFTER COMPLETION OF GONADOTOXIC TREATMENT

The proposed population includes female patients of reproductive age who will be having, or have had, gonadotoxic treatment. Gonadotoxic treatment includes any treatment which is associated with a risk of ovarian damage or sterility. This includes treatment for malignancy (i.e. chemotherapy, irradiation), and treatment for precancerous or benign conditions (pelvic surgery). Ovarian function is key for induction of puberty, for fertility and for timing of menopause.

Examples of malignant disorders in women of reproductive age include:

- breast cancer
- haematological malignancies (e.g. Hodgkin’s and non-Hodgkin’s lymphoma, leukaemia)
- ovarian cancer
- cervical cancer

Examples of non-malignant disorders in women of reproductive age (requiring gonadotoxic treatment) include:

- endometriosis
- ovarian cysts
- auto-immune diseases (e.g. systemic lupus erythematosus, rheumatoid arthritis)

The nature of the treatment often determines the degree of damage to the ovaries, and it is often difficult to give an accurate fertility prognosis before the start of the treatment (Anderson & Wallace 2013). The gonadotoxicity of combination chemotherapy treatments varies according to the specific agents used, their cumulative doses, the protocol used, and the reproductive potential of the patient

at the time of treatment. High-dose alkylating agents and ionising radiation have well-recognised gonadotoxicity, which leads to infertility in a high proportion of patients (Roberts et al. 2015). Quantifying the gonadotoxic effects of each chemotherapy regimen is difficult and poorly studied to date. The impact of these treatments, such as ovarian failure, infertility and early menopause, increases with age. The American Society of Clinical Oncology (ASCO) guidelines have estimated the risk of amenorrhea in women treated with chemotherapy and radiotherapy (Table 7) (Loren et al. 2013).

Table 7 Risk of amenorrhea in women treated with chemotherapy and radiotherapy, adapted from Loren et al. 2013 (ASCO guidelines)

Degree of risk	Treatment protocol	Patient and dose factors	Common uses
High risk of amenorrhea (>70%)	Any alkylating agent (e.g., busulfan, carmustine, cyclophosphamide, ifosfamide, lomustine, melphalan, procarbazine) + total body irradiation	-	Conditioning for HSCT for leukaemias, lymphomas, myelomas, Ewing's sarcoma, neuroblastoma, choriocarcinoma
	Any alkylating agent + pelvic radiation	-	Ovarian cancer, sarcomas
High risk of amenorrhea (>70%)	Total cyclophosphamide	5 g/m ² in women age >40 7.5 g/m ² in women and girls age <20	Multiple cancers: breast cancer, non-Hodgkin lymphoma, conditioning for haematopoietic stem cell transplantation
	Protocols containing procarbazine: MOPP, BEACOPP	>3 cycles >6 cycles	Hodgkin lymphoma
	Protocols containing temozolomide or BCNU + cranial radiation	-	Brain tumour
	Whole abdominal or pelvic radiation doses	>6 Gy in adult women >10 Gy in post-pubertal girls >15 Gy in pre-pubertal girls	Wilms tumour, neuroblastoma, sarcomas, Hodgkin lymphoma, ovarian
	Total Body Irradiation (TBI) doses	-	Haematopoietic stem cell transplantation
	Cranial radiation	>40 Gy	Brain tumour
Intermediate risk of amenorrhea (~30-70%)	Total cyclophosphamide	5 g/m ² in women age 30-40	Multiple cancers, breast cancer
	AC for breast cancer	x4 + Paclitaxel or Docetaxel in women age <40	Breast cancer
	Monoclonal Antibodies (e.g., Bevacizumab (Avastin))	-	Colon cancer, non-small-cell lung cancer, head and neck cancer, breast cancer
	FOLFOX4	-	Colon cancer
	Protocols containing cisplatin	-	Cervical cancer
	Abdominal/pelvic radiation	10-15 Gy in pre-pubertal girls, 5-10 Gy in post-pubertal girls	Wilms tumour, neuroblastoma, spinal tumours, brain tumours, relapsed acute lymphoblastic

Degree of risk	Treatment protocol	Patient and dose factors	Common uses
			leukaemia or non-Hodgkin lymphoma
Lower risk of amenorrhea (<30%)	Protocols containing non-alkylating agents or lower levels of alkylating agents (e.g. ABVD, CHOP, COP; multi-agent therapies for leukaemia)	-	Hodgkin lymphoma, non-Hodgkin lymphoma, leukaemia
	Protocols for breast cancer containing cyclophosphamide (e.g. CMF, CEF, or CAF)	Women <30 years old	Breast cancer
	Anthracycline + cytarabine	-	Acute myeloid leukaemia
Very low / no risk of amenorrhea (negligible)	Multi-agent therapies using vincristine	-	Leukaemia, Lymphoma, Breast cancer, Lung cancer
	Radioactive iodine	-	Thyroid cancer
Unknown risk of amenorrhea	Monoclonal Antibodies (e.g., Cetuximab (Erbix), Trastuzumab (Herceptin))	-	Colon cancer, Non-small-cell lung cancer, head and neck cancer, breast cancer
	Tyrosine kinase inhibitors (e.g., Erlotinib (Tarceva), Imatinib (Gleevec))	-	Non-small-cell lung cancer, pancreatic cancer, chronic myeloid leukaemia, gastrointestinal stromal tumours

ABVD = adriamycin, bleomycin, vinblastine, dacarbazine; AC = adriamycin, cyclophosphamide; BEACOPP = bleomycin, etoposide, adriamycin, cyclophosphamide, oncovin, procarbazine, prednisone; CAF = cyclophosphamide, doxorubicine, 5-fluorouracil; CEF = cyclophosphamide, epirubicin, 5-fluorouracil; CHOP = cyclophosphamide, hydroxydaunorubicin, oncovin, prednisone/prednisolone; CMF = cyclophosphamide, methotrexate, 5-fluorouracil; COP = cyclophosphamide, vincristine, prednisone; FOLFOX = folinic acid, fluorouracil, oxaliplatin; HSCT = haematopoietic stem cell transplant; MOPP = mustargen, oncovin, procarbazine, prednisone; TBI = total body irradiation

Source: (Loren et al. 2013)

WHY GONADOTOXIC TREATMENT IS A CONCERN

Some women undergoing a gonadotoxic treatment will be able to reproduce naturally. However, gonadotoxic treatments, e.g. alkylating agents, can have two distinct effects on ovarian function, leading to infertility. The first effect is from immediate damage to the growing follicle population, characterised by amenorrhea during or immediately after treatment. Depending on the extent of primordial follicle loss due to this treatment, premature ovarian insufficiency and continuation of amenorrhea may occur at a later date. If a sufficient pool of follicles remains after treatment, the population of growing follicles will be replenished and menses can resume (Anderson & Wallace 2013). When there is only a partial loss of primordial follicles during treatment, premature ovarian failure may only manifest after years or decades. It can be difficult to predict the risk of infertility for individuals prior to the commencement of treatment.

For patients diagnosed with cancer, the main concern is initially long-term survival. However, with more and more patients surviving cancer, the loss of fertility due to gonadotoxic cancer therapies becomes an important issue (Munoz et al. 2016). Cancer survivors with an irregular menstrual function have been found to have lower quality of life scores than those with regular cycles

(Kondapalli et al. 2014). Likewise, patients with debilitating non-malignant conditions requiring gonadotoxic treatments face similar fertility issues. Gonadotoxic treatment in females of any age can lead to subsequent infertility, as the ovaries are susceptible to damage before, during, and after puberty.

ESTIMATED INCIDENCE

In Australia, the estimated number of new cancer cases diagnosed in 2017 among women aged less than 45 years old is 6,520 (Table 8). The AIHW provides a breakdown by common cancer types associated with ovarian failure and fertility preservation, and therefore relevant to AMH testing, for females aged 0 to 44 (Table 9). This approach does not capture rare tumours which may be treated with gonadotoxic treatment and does not provide the proportion of cases with the different tumour types who would be treated with gonadotoxic treatment.

Table 8 Incidence of cancer in Australian females

Population group	Age	Estimated new cases in 2017
Paediatric	0–14 years	322
Adolescent / young adult	15–24 years	432
Adults	25–44 years	5,766
Australian population	0–44 years	6,520

Source: AIHW Cancer in Australia 2017 Supplementary tables Chapter 3 Incidence of Cancer

Table 9 Estimated number of new cases, by type of cancer most relevant for fertility preservation for 2017, females aged 0 to 44

Tumour type	Estimated new cases in 2017
Acute myeloid leukaemia	86
Brain	167
Breast cancer	1864
Cervix	426
Hodgkin's lymphoma	177
Non-Hodgkin's lymphoma	197
Other soft tissue	79
Ovary	164
Uterus	126
Total	3286

Source: AIHW Cancer in Australia 2017 Supplementary tables Chapter 3 Incidence of Cancer

UPTAKE OF AMH TESTING IN AUSTRALIA

The estimated use of AMH testing is described in detail in section E.2 of this report. It was suggested that approximately 50% of women of reproductive age who are diagnosed with a malignant condition will undergo gonadotoxic treatment affecting their fertility and be eligible for AMH testing. Table 67 shows the number of cases of cancer estimated to undergo gonadotoxic treatment in 2017 (50% of total cancer incidence of 6,520 is 3,260).

There is a lack of evidence on the proportion of patients referred to fertility specialists who are undergoing gonadotoxic treatment. One recently published systematic review reported referral rates between 14% and 67% for the period 2012 to 2016 (Logan et al. 2017).

Australian data on the proportion of women with non-malignant conditions undergoing gonadotoxic treatment and AMH testing could not be identified. A recent meta-analysis conducted to determine cohort epidemiological characteristics and success rates of autologous ovarian tissue transplantation reported that approximately 78% of the women undergoing fertility preservation due to gonadotoxic treatment had malignant conditions and the other 22% had non-malignant indications (Pacheco & Oktay 2017).

Data provided by clinical experts for one Victorian centre indicated that nearly 53% of the referred and counselled oncology patients aged 0 to 44 years underwent AMH testing at least once during the 2011 to 2017 period. Of these patients, 86% had an AMH test before or within three months of their first consultation. It is therefore likely that they were tested prior treatment. The MBS item proposes a maximum of one AMH test per patient per year, however not all women in the target population will undergo routine annual AMH tests after treatment. Table 68 estimates the current and expected number of AMH tests annually over the next five years, if the service were to be listed on the MBS.

A5 COMPARATOR DETAILS

The comparator is usually defined as the current practice most likely to be replaced or added to by the proposed medical service. AMH testing would be done in addition to the current standard tests to measure ovarian reserve. Ovarian reserve testing can include both biochemical tests and ultrasound imaging of the ovaries (American Society for Reproductive Medicine Practice Committee 2015).

The PICO Confirmation reported the following tests as comparators:

- basal FSH measurement
- estradiol (E2) measurement
- inhibin B measurement
- AFC, if post-pubertal.

FSH is secreted by the pituitary gland in the brain and stimulates the onset of new follicular growth and an increase in estradiol (E2) concentrations (American Society for Reproductive Medicine Practice Committee 2015). Basal serum FSH concentrations increase around the third day of the menstrual cycle with advancing reproductive age. To predict a woman's ovarian reserve, FSH is

usually measured on the third day of the menstrual cycle. Reliability of FSH measurements is limited due to high inter- and intra-cycle variability, and the absence of a consistent cut-off point for abnormal levels. However, FSH is widely used as a measure of ovarian reserve, despite these limitations. High FSH values have been correlated with, but do not necessarily predict, poor ovarian stimulation and failure to conceive (American Society for Reproductive Medicine Practice Committee 2015).

Basal E2 level measurements aid the interpretation of the FSH measurements. An early increase of E2 is a characteristic of reproductive aging, and can return an otherwise elevated FSH level to the normal range. Therefore, normal FSH levels concurrent with an increased E2 level at the second, third or fourth day of the menstrual cycle, may be associated with diminished ovarian reserve, poor response and lower pregnancy rates. Basal E2 measurements have poor inter- and intra-cycle reliability, and should not be used in the absence of other tests to determine diminished ovarian (American Society for Reproductive Medicine Practice Committee 2015).

Inhibin B is a heterodimeric glycoprotein produced by the granulosa cells in the ovary. It suppresses synthesis and secretion of FSH. Inhibin B levels are highest in the late follicular phase and luteal phase of the menstrual cycle (Chada et al. 2003). The levels of this hormone rise with gonadotropin-releasing hormone (GnRH) and FSH stimulation, and therefore will show very high variability within the menstrual cycle and between menstrual cycles (American Society for Reproductive Medicine Practice Committee 2015). It is therefore a poor measure of ovarian reserve. Most studies reported in a 2015 review showed that inhibin B does not discriminate between pregnancy and failure to conceive (American Society for Reproductive Medicine Practice Committee 2015).

Antral follicles can be measured using transvaginal ultrasonography during the early follicular phase. Most studies define antral follicles as those measuring two to ten millimetres in diameter in the greatest two-dimensional plane, although some studies define antral follicles as those measured three to eight millimetres. AFC has been identified as a predictor of ovarian response in ART, where low AFC is considered to be three to six antral follicles and is associated with poor response to ovarian stimulation. However, AFC seems unable to reliably predict the ability to conceive (American Society for Reproductive Medicine Practice Committee 2015). AFC has low inter-cycle variability and low to moderate inter-observer variability (Hsu et al. 2011). Inter- and intra-observer variability may be limiting, especially in centres with less experience or lower quality equipment (American Society for Reproductive Medicine Practice Committee 2015).

The MBS item descriptors for the relevant comparators, including explanatory notes, are summarised in Table 10 to Table 12. Inhibin B measurements are currently not listed on the MBS, however FSH and E2 are listed under MBS item 66695. AFC is done using ultrasound under MBS item 55065.

Table 10 Relevant MBS item for comparator tests

<p>Category 6 - Pathology Services</p>
<p>MBS 66695</p> <p>Quantitation in blood or urine of hormones and hormone binding proteins - ACTH, aldosterone, androstenedione, C-peptide, calcitonin, cortisol, DHEAS, 11-deoxycortisol, dihydrotestosterone, FSH, gastrin, glucagon, growth hormone, hydroxyprogesterone, insulin, LH, estradiol, oestrone, progesterone, prolactin, PTH, renin, sex hormone binding globulin, somatomedin C(IGF-1), free or total testosterone, urine steroid fraction or fractions, vasoactive intestinal peptide - 1 test</p> <p>(Item is subject to rule 6) Fee: \$30.50 Benefit: 75% = \$22.90 85% = \$25.95 (See para TN.1.4 of explanatory notes to this Category)</p>
<p>MBS 66698</p> <p>2 tests described in item 66695</p> <p>(Item is subject to rule 6) Fee: \$43.70 Benefit: 75% = \$32.80 85% = \$37.15 (See para TN.1.4 of explanatory notes to this Category)</p>
<p>MBS 66701</p> <p>3 tests described in item 66695</p> <p>(Item is subject to rule 6) Fee: \$56.90 Benefit: 75% = \$42.70 85% = \$48.40</p> <p>(See para TN.1.4 of explanatory notes to this Category)</p>
<p>MBS 66704</p> <p>4 tests described in item 66695</p> <p>(Item is subject to rule 6) Fee: \$70.15 Benefit: 75% = \$52.65 85% = \$59.65</p> <p>(See para TN.1.4 of explanatory notes to this Category)</p>
<p>MBS 66707</p> <p>5 or more tests described in item 66695</p> <p>(Item is subject to rule 6) Fee: \$83.35 Benefit: 75% = \$62.55 85% = \$70.85</p> <p>(See para TN.1.4 of explanatory notes to this Category)</p>

Table 11 Explanatory notes for comparator MBS item number

<p>Category 3 - Therapeutic procedures</p>
<p>TN. 1.4. Assisted Reproductive Technology ART Services - (Items 13200 to 13221)</p> <p>Medicare benefits are not payable in respect of ANY other item in the Medicare Benefits Schedule (including Pathology and Diagnostic Imaging) in lieu of or in connection with items 13200 - 13221. Specifically, Medicare benefits are not payable for these items in association with items 104, 105, 14203, 14206, 35637, pathology tests or diagnostic imaging. A treatment cycle that is a series of treatments for the purposes of ART services is defined as beginning either on the day on which treatment by superovulatory drugs is commenced or on the first day of the patient's menstrual cycle, and ending either; not more than 30 days later, or if a service mentioned in item 13212, 13215 or 13321 is provided in connection with the series of treatments-on the day after the day on which the last of those services is provided.</p> <p>The date of service in respect of treatment covered by Items 13200, 13201, 13203, 13206, 13209 and 13218 is DEEMED to be the FIRST DAY of the treatment cycle.</p> <p>Items 13200, 13201, 13202 and 13203 are linked to the supply of hormones under the Section 100 (National Health Act) arrangements. Providers must notify the Department of Human Services of Medicare card numbers of patients using</p>

Category 3 - Therapeutic procedures

hormones under this program, and hormones are only supplied for patients claiming one of these four items. Medicare benefits are not payable for assisted reproductive services rendered in conjunction with surrogacy arrangements where surrogacy is defined as 'an arrangement whereby a woman agrees to become pregnant and to bear a child for another person or persons to whom she will transfer guardianship and custodial rights at or shortly after birth'. NOTE: Items 14203 and 14206 are not payable for artificial insemination.

Table 12 Relevant MBS item for ultrasound used for antral follicle count

Category 5 - Diagnostic Imaging Services

MBS 55065

PELVIS, ultrasound scan of, by any or all approaches, where:

- (a) the patient is referred by a medical practitioner; and
- (b) the service is not associated with a service to which an item in Subgroup 2, or 3, applies; and
- (c) the referring practitioner is not a member of a group of practitioners of which the providing practitioner is a member; and
- (d) the service is not solely a transrectal ultrasonic examination of the prostate gland, bladder base and urethra, or any of those organs; and
- (e) the service is not performed with item 55014, 55017, 55036 or 55038 on the same patient within 24 hours (R)(K)

Bulk bill incentive

Fee: \$98.25 Benefit: 75% = \$73.70 85% = \$83.55

(See para IN.0.19 of explanatory notes to this Category)

CLINICAL MANAGEMENT ALGORITHMS

Figure 2 shows the current and proposed clinical management algorithm for females prior to gonadotoxic treatment. The proposed intervention is shown in red. The current and proposed clinical management algorithms for females following the completion of gonadotoxic treatment are shown in Figure 3 and Figure 4, respectively. In most cases, the AMH test will be done in addition to current ovarian reserve tests, i.e. FSH, E2 and AFC.

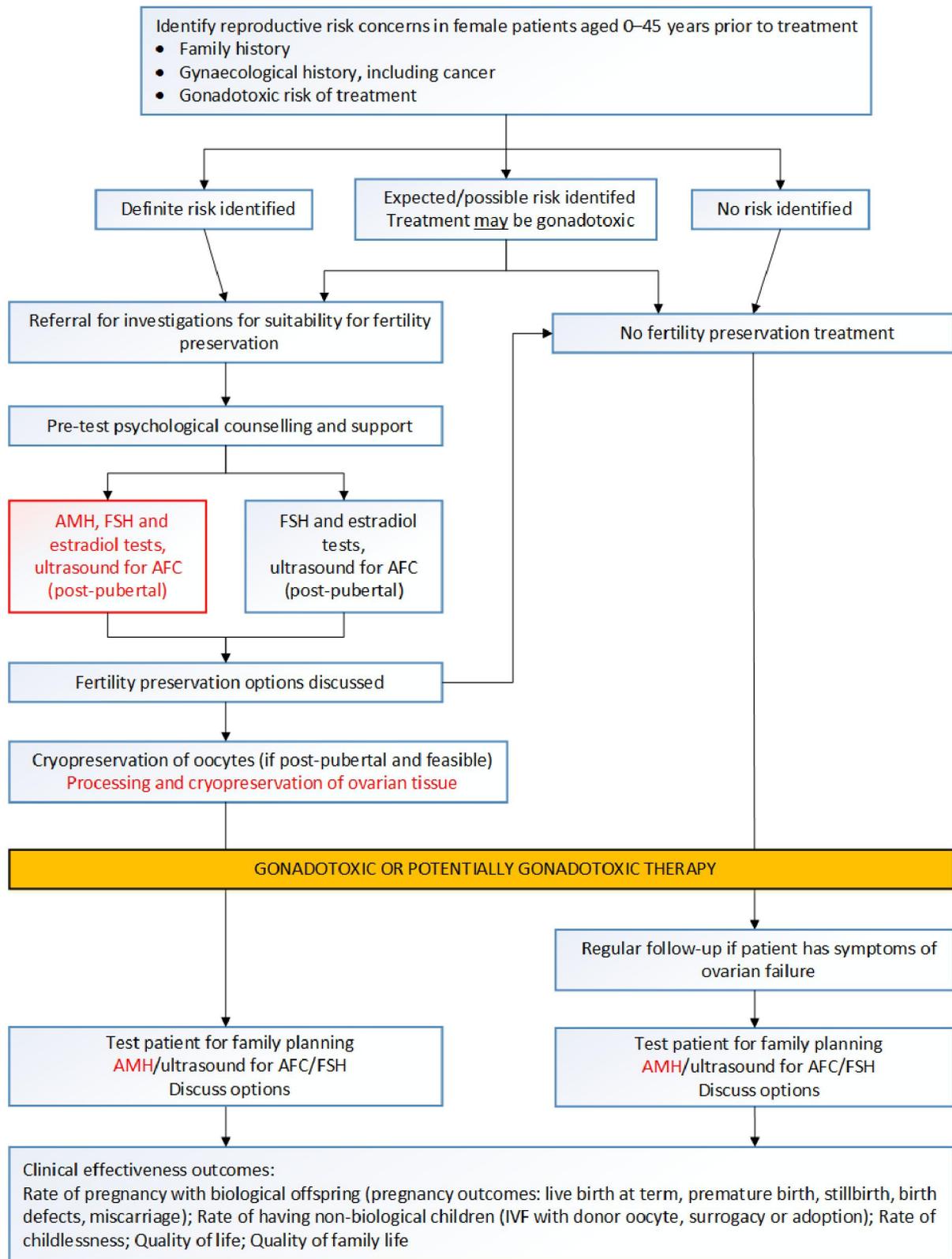


Figure 2 Clinical management algorithm for assessment of ovarian reserve and cryopreservation of oocytes or ovarian tissue. Pathway showing cryopreservation of oocytes is current practice in post-pubertal cases. Pathway showing cryopreservation of ovarian tissue is for the proposed service. The relevant population for both pathways includes females aged 0 to 45 years, pre-gonadotoxic treatment

AFC = antral follicle count, AMH = anti-Müllerian hormone, FSH = follicle-stimulating hormone
Source: Adapted from algorithm 4 from the final protocol for MSAC application 1434.

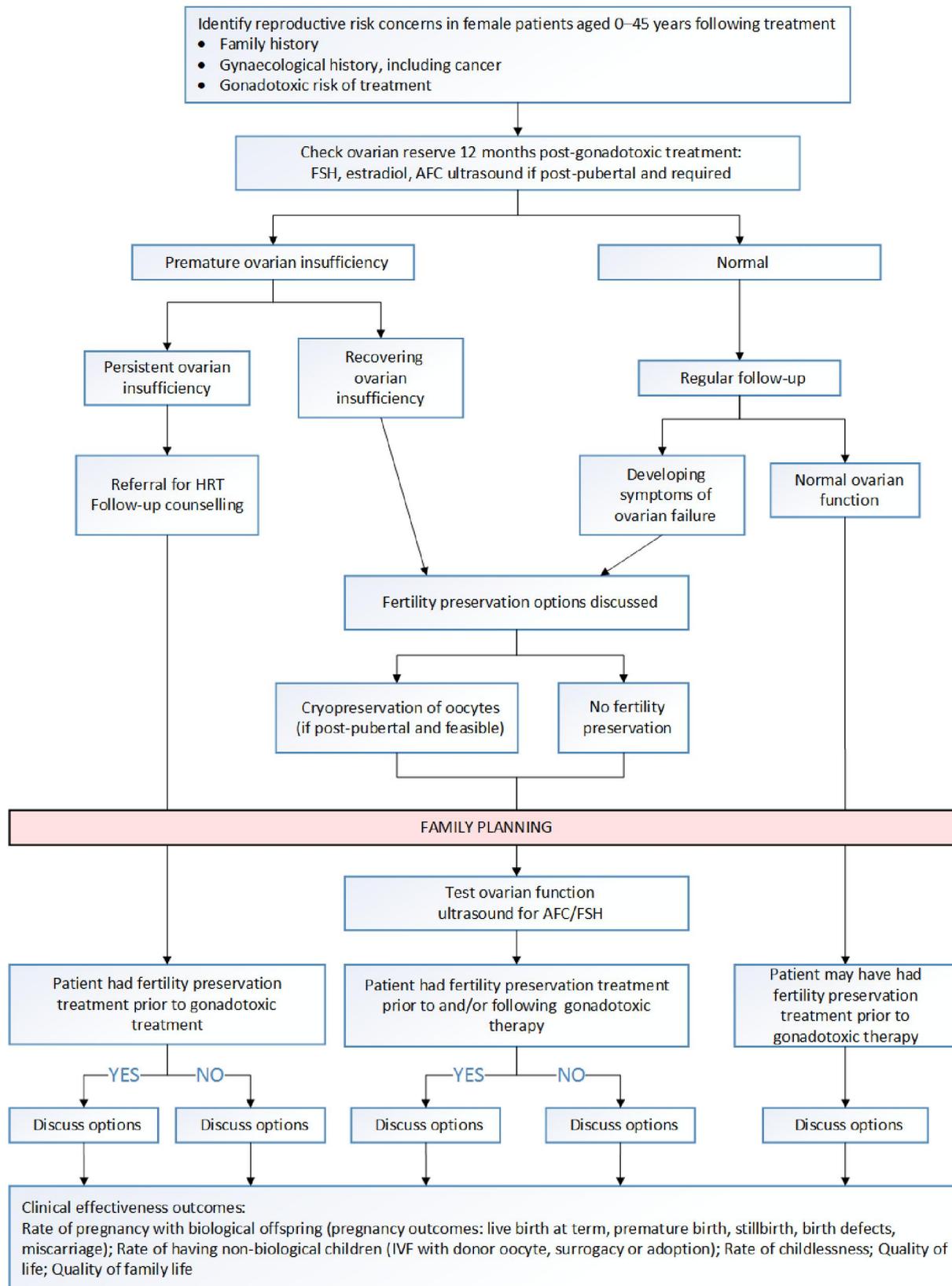


Figure 3 Current clinical management algorithm for the assessment of ovarian reserve and post-pubertal cryopreservation oocytes for females aged 0 to 45 years, following completion of gonadotoxic treatment

AFC = antral follicle count, AMH = Anti-Müllerian Hormone, E2 = estradiol, FSH = follicle-stimulating hormone, HRT = hormone replacement therapy

Source: Adapted from algorithm 4 from the final protocol for MSAC application 1434.

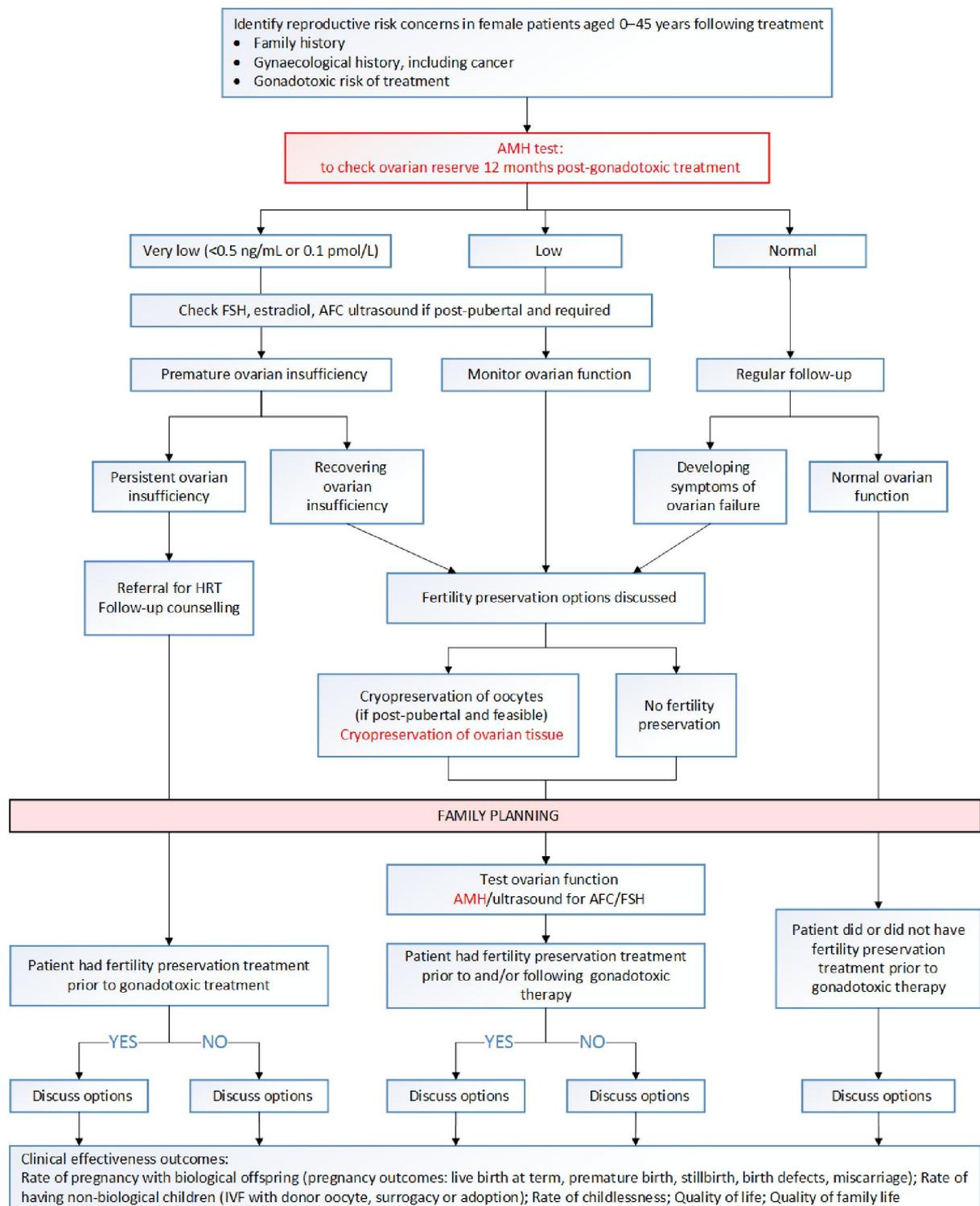


Figure 4 Proposed clinical management algorithm for the assessment of ovarian reserve and cryopreservation of oocytes or ovarian tissue for females aged 0 to 45 years, following completion of gonadotoxic treatment

The proposed test (AMH) is shown in red, current practice included the cryopreservation of oocytes. The proposed intervention for cryopreservation of ovarian tissue (MSAC application 1435 Part B) is also highlighted in red.

AFC = antral follicle count, AMH = anti-Müllerian hormone, FSH = follicle-stimulating hormone, HRT = hormone replacement therapy

Source: Adapted from algorithm 4 from the final protocol for MSAC application 1434.

A6 CLINICAL CLAIM

The early consultation PICO Confirmation indicated AMH testing would be used to: (1) assess ovarian function prior to cancer treatment; (2) estimate the return of reproductive function following gonadotoxic treatment; (3) assess the need for fertility preservation prior to or following gonadotoxic treatment, or to recommend the start of ART in patients planning to start a family. The final PICO Confirmation did not make a clear clinical claim. Based on the available information in the PICO Confirmation documents, the evidence, and the consultation feedback, the assessment group identified that AMH testing is used for several purposes.

Prior to gonadotoxic treatment, an AMH test is used to provide prognostic information about risk of infertility, thereby informing the decision whether to undergo fertility preservation. Or, in other words, whether fertility is likely to be sufficient after gonadotoxic therapy such that fertility preservation is unnecessary, versus a scenario requiring that patients are counselled that they are likely to have a low number of oocytes, and that the chance of conception is low. Pre-gonadotoxic treatment, an AMH test may also be used to predict the response to superovulation, thereby enabling the yield of good quality oocytes to be maximised while reducing the risk of ovarian hyperstimulation by adjusting starting doses of rFSH.

Following gonadotoxic treatment, an AMH test is used to monitor the level of ovarian function. Monitoring ovarian function may inform whether patients are candidates for natural conception or may need ART, using previously cryopreserved ovarian tissue, oocytes or embryos to conceive. Other claimed benefits of monitoring ovarian function include the ability to guide HRT based on information on ovarian reserve (i.e. is there premature ovarian insufficient), to diagnose ovarian failure, and to assist with counselling and supportive care of early menopause.

A7 SUMMARY OF THE PICO

A guiding framework for the PICO is determined by MSAC for each assessment using a PICO Confirmation. The PICO Confirmation is a document that describes the current clinical practice and reflects the likely future practice with the proposed medical service.

The PICO that were pre-specified to guide the systematic literature review for direct evidence are presented in Table 13 and Table 14. Table 13 lists the PICO for the population of patients undergoing AMH testing prior to gonadotoxic treatment and Table 14 shows the PICO for the population of patients having AMH testing post-gonadotoxic treatment. The PICO criteria for direct and linked evidence for the female population are divided into three population subgroups according to age.

DIRECT EVIDENCE

Table 13 PICO criteria and research questions for direct evidence in populations 1, 2 and 3

Population	<ol style="list-style-type: none"> 1. Paediatric female patients aged 0–14 years prior to receiving gonadotoxic treatment 2. Adolescent/young adult female patients aged 15–25 years prior to receiving gonadotoxic treatment 3. Adult female patients aged 26–45 years prior to receiving gonadotoxic treatment
Intervention	<ol style="list-style-type: none"> a) AMH test in addition to other standard tests, to measure ovarian reserve to determine if cryopreservation of ovarian tissue, oocytes or embryos for fertility preservation is required, and predict chance of success of fertility preservation b) AMH test in addition to other standard tests, to predict response to superovulation and determine the level of hormones used for superovulation
Comparators	<ol style="list-style-type: none"> a) Other standard tests (FSH, E2, AFC ultrasound if post-pubertal) to measure ovarian reserve to determine if cryopreservation of ovarian tissue, oocytes or embryos for fertility preservation is required, and predict chance of success of fertility preservation b) Other standard tests (FSH, E2, AFC ultrasound if post-pubertal) to predict response to superovulation and determine the level of hormones used
Outcomes	<p><u>Safety:</u></p> <ul style="list-style-type: none"> Adverse events associated with the tests Adverse events associated with superovulation, such as OHSS Adverse events associated with procurement of ovarian tissue or oocytes Anxiety about viability of frozen tissue or oocytes Anxiety about disposal of excess or unwanted frozen tissue or oocytes <p><u>Effectiveness:</u></p> <ul style="list-style-type: none"> Rate of pregnancy with biological offspring Pregnancy outcomes (live birth at term, premature birth, stillbirth, birth defects, miscarriage) Rate of having non-biological children (IVF with donor oocyte, surrogacy or adoption) Rate of childlessness Quality of life Quality of relationship and family life Yield of good quality oocytes (intermediate effectiveness outcome) <p><u>Cost-effectiveness</u></p> <ul style="list-style-type: none"> Cost Cost per QALY or cost per live birth ICER
Study design	Randomised trials, cohort studies, case series, or systematic reviews of these study designs
Search period	1990–June 2017
Language	Studies in languages other than English will only be translated if they represent a higher level of evidence than that available in the English language evidence base
<ol style="list-style-type: none"> a. What is the safety, effectiveness, and cost-effectiveness of having an AMH test to measure ovarian reserve in addition to other standard tests, versus other standard tests alone, for determining the need for cryopreservation of ovarian tissue, oocytes or embryos, and predicting the success of fertility preservation, prior to receiving gonadotoxic treatment? b. What is the safety, effectiveness, and cost-effectiveness of having an AMH test in addition to other standard tests, versus other standard tests alone, to predict response to superovulation and determine the level of hormones used, prior to receiving gonadotoxic treatment? 	

AFC = antral follicle count; AMH = Anti-Müllerian hormone; E2 = estradiol, FSH = follicle-stimulating hormone; ICER = incremental cost-effectiveness ratio; IVF = in vitro fertilisation, OHSS = ovarian hyperstimulation syndrome; QALY = quality-adjusted life year

Table 14 PICO criteria and research questions for direct evidence in populations 4, 5 and 6

Population	<ol style="list-style-type: none"> 1. Paediatric female patients aged 0–14 years following completion of gonadotoxic treatment 2. Adolescent/young adult female patients aged 15–25 years following completion of gonadotoxic treatment 3. Adult female patients aged 26–45 years following completion of gonadotoxic treatment
Intervention	<p>AMH test in addition to other standard tests, to assess the level of ovarian function for:</p> <ol style="list-style-type: none"> a) helping inform whether patients need to use ART to conceive, using cryopreserved ovarian tissue, oocytes or embryos, or should try natural conception b) providing information about decline in ovarian reserve (premature ovarian insufficiency), and whether patients should undergo fertility preservation, if not already performed prior to treatment c) monitoring whether pre-pubertal patients have delayed puberty, in order to allow treatment d) diagnose ovarian failure and assist with counselling and supportive care of early menopause
Comparators	Other standard tests (FSH, E2, AFC ultrasound if post-pubertal) to assess level of ovarian function
Outcomes	<p><u>Safety</u></p> <ul style="list-style-type: none"> Adverse events associated with procurement of ovarian tissue or oocytes Adverse events associated with the test(s) Anxiety about viability of frozen tissue/oocytes Anxiety about disposal of excess or unwanted frozen tissue or oocytes <p><u>Effectiveness</u></p> <ul style="list-style-type: none"> Rate of pregnancy with biological offspring Pregnancy outcomes (live birth at term, premature birth, stillbirth, birth defects, miscarriage) Rate of having non-biological children (IVF with donor oocyte, surrogacy or adoption) Rate of childlessness Quality of life Quality of relationship and family life Psychosocial or health outcomes associated with pubertal timing <p><u>Cost-effectiveness</u></p> <ul style="list-style-type: none"> Cost Cost per QALY or cost per live birth ICER
Study design	Randomised trials, cohort studies, case series, or systematic reviews of these study designs
Search period	1990–June 2017
Language	Studies in languages other than English will only be translated if they represent a higher level of evidence than that available in the English language evidence base
<p>What is the safety, effectiveness, and cost-effectiveness of having an AMH test in addition to other standard tests to measure ovarian function, compared to other standard tests alone in female patients following completion of gonadotoxic treatment?</p>	

AFC = antral follicle count; AMH = anti-Müllerian hormone; ART = assisted reproductive technology; FSH = follicle-stimulating hormone; ICER = incremental cost-effectiveness ratio; IVF = in vitro fertilisation; OHSS = ovarian hyperstimulation syndrome; QALY = quality-adjusted life year

LINKED EVIDENCE

The PICO that were pre-specified to guide the systematic literature review for a linked evidence approach, are presented in Table 15 to Table 19 outlines the criteria to be used for assessing the impact of a change in fertility preservation options. Assessment of the effectiveness of cryopreservation of ovarian tissue was assessed in MSAC 1435 (Part B: females). It was stated in the research protocol that if other changes in management were identified, such as differential use of hormone treatment, then the health impact of these would be assessed.

Table 19, which include the PICO for the linked analysis in the population prior to receiving gonadotoxic treatment, and Table 20 to Table 22, which show the PICO for the population following completion of gonadotoxic treatment.

Table 15 PICO criteria and research question for the analytical validity of AMH testing in populations 1 to 3

Population	<ol style="list-style-type: none"> 1. Paediatric female patients aged 0–14 years prior to receiving gonadotoxic treatment 2. Adolescent/young adult female patients aged 15–25 years prior to receiving gonadotoxic treatment 3. Adult female patients aged 26–45 years prior to receiving gonadotoxic treatment
Intervention	AMH test
Comparator	Other standard tests (FSH, E2) Other AMH test (repeated test, different laboratory, different assay etc).
Evidentiary standard	AFC ultrasound if post-pubertal
Outcomes	Concordance (per cent positive agreement), unsatisfactory or uninterpretable test results, reliability (intra-observer or intra-instrument variability/agreement)
Search period	1990–June 2017
Language	Studies in languages other than English will only be translated if they represent a higher level of evidence than that available in the English language evidence base
<p>What is the concordance of AMH, FSH and E2 in female patients prior to receiving gonadotoxic treatment?</p> <p>What is the concordance between AMH and AFC ultrasound, compared to FSH and E2 in female patients prior to receiving gonadotoxic treatment?</p> <p>What is the reliability of AMH testing, compared to FSH and E2 in female patients prior to receiving gonadotoxic treatment?</p>	

AFC = antral follicle count; AMH = anti-Müllerian hormone; E2 = estradiol; FSH = follicular stimulating hormone; SROC = summary receiving operating characteristic

Table 16 PICO criteria and research question for determining prognostic value of AMH in population 1, 2 and 3

Population	<ol style="list-style-type: none"> 1. Paediatric female patients aged 0–14 years prior to receiving gonadotoxic treatment 2. Adolescent/young adult female patients aged 15–25 years prior to receiving gonadotoxic treatment 3. Adult female patients aged 26–45 years prior to receiving gonadotoxic treatment
Intervention	AMH test in addition to other standard tests, to measure ovarian reserve
Comparators	Other standard tests (FSH, E2, AFC ultrasound if post-pubertal), to measure ovarian reserve
Outcomes	Rate of pregnancy achieved naturally, rate of pregnancy achieved through ART (with biological offspring), rate of childlessness, quality of life/family life
Study design	Prospective cohort study, all or none, analysis of prognostic factors among persons in a single

Population	<ol style="list-style-type: none"> 1. Paediatric female patients aged 0–14 years prior to receiving gonadotoxic treatment 2. Adolescent/young adult female patients aged 15–25 years prior to receiving gonadotoxic treatment 3. Adult female patients aged 26–45 years prior to receiving gonadotoxic treatment
	arm of a randomised controlled trial, retrospective cohort study, case series, or systematic reviews of these study designs
Search period	1990–June 2017
Language	Studies in languages other than English will only be translated if they represent a higher level of evidence than that available in the English language evidence base
Will the extra information generated as a result of the AMH test be of additional prognostic value in female patients prior to receiving gonadotoxic treatment, compared to FSH, E2 and/or AFC ultrasound alone?	

AFC = antral follicle count; ART = assisted reproductive technology; AMH = anti-Müllerian hormone; E2 = estradiol; FSH = follicular stimulating hormone

Table 17 PICO criteria and research question for determining predictive value of AMH in population 1, 2 and 3

Population	<ol style="list-style-type: none"> 1. Paediatric female patients aged under 14 years (if post-pubertal) prior to receiving gonadotoxic treatment 2. Adolescent/young adult female patients aged 15–25 years prior to receiving gonadotoxic treatment 3. Adult female patients aged 26–45 years prior to receiving gonadotoxic treatment
Intervention	AMH test in addition to other standard tests, to predict response to superovulation
Comparators	Other standard tests (FSH, E2, AFC ultrasound if post-pubertal), to predict response to superovulation
Outcomes	Yield of good quality oocytes, rate of OHSS
Study design	Prospective cohort study, all or none, analysis of prognostic factors among persons in a single arm of a randomised controlled trial, retrospective cohort study, case series, or systematic reviews of these study designs
Search period	1990–June 2017
Language	Studies in languages other than English will only be translated if they represent a higher level of evidence than that available in the English language evidence base
Does the addition of AMH testing to standard tests allow better prediction of response to superovulation than standard tests alone?	

AFC = antral follicle count; AMH = anti-Müllerian hormone; E2 = estradiol; FSH = follicular stimulating hormone; OHSS = ovarian hyper stimulation syndrome

Table 18 PICO criteria and research question to determine the impact of the AMH test on patient management in population 1, 2 and 3

Population	<ol style="list-style-type: none"> 1. Paediatric female patients aged 0–14 years prior to receiving gonadotoxic treatment 2. Adolescent/young adult female patients aged 15–25 years prior to receiving gonadotoxic treatment 3. Adult female patients aged 26–45 years prior to receiving gonadotoxic treatment
Intervention	<ol style="list-style-type: none"> a) AMH test in addition to other standard tests, to measure ovarian reserve to determine if cryopreservation of ovarian tissue, oocytes or embryos for fertility preservation is required, and predict chance of success of fertility preservation b) AMH test in addition to other standard tests, to predict response to superovulation and determine the level of hormones used
Comparators	a) Other standard tests (FSH, E2, AFC ultrasound if post-pubertal) to measure ovarian reserve to determine if cryopreservation of ovarian tissue, oocytes or embryos for fertility

Population	<ol style="list-style-type: none"> 1. Paediatric female patients aged 0–14 years prior to receiving gonadotoxic treatment 2. Adolescent/young adult female patients aged 15–25 years prior to receiving gonadotoxic treatment 3. Adult female patients aged 26–45 years prior to receiving gonadotoxic treatment
	<p>preservation is required, and predict chance of success of fertility preservation</p> <p>b) Other standard tests (FSH, E2, AFC ultrasound if post-pubertal) to predict response to superovulation and determine the level of hormones used</p>
Outcomes	<p>Change in fertility preservation pathway (cryopreservation of oocytes and/or ovarian tissue undertaken or avoided), patient compliance, time from testing to gonadotoxic treatment</p> <p>Levels of hormones used for superovulation</p>
Study design	Randomised trials, cohort studies, case series or systematic reviews of these study designs
Search period	1990–June 2017
Language	Studies in languages other than English will only be translated if they represent a higher level of evidence than that available in the English language evidence base
Does the addition of the AMH test lead to a change in management in female patients prior to receiving gonadotoxic treatment, compared to FSH, E2 and/or AFC ultrasound alone?	

AFC = antral follicle count; AMH = anti-Müllerian hormone; E2 = estradiol; FSH = follicular stimulating hormone

Table 19 outlines the criteria to be used for assessing the impact of a change in fertility preservation options. Assessment of the effectiveness of cryopreservation of ovarian tissue was assessed in MSAC 1435 (Part B: females). It was stated in the research protocol that if other changes in management were identified, such as differential use of hormone treatment, then the health impact of these would be assessed.

Table 19 PICO criteria and research question to determine the effectiveness of change in management in population 1, 2 and 3

Population	<ol style="list-style-type: none"> 1. Paediatric female patients aged 0–14 years prior to receiving gonadotoxic treatment 2. Adolescent/young adult female patients aged 15–25 years prior to receiving gonadotoxic treatment 3. Adult female patients aged 26–45 years prior to receiving gonadotoxic treatment
Intervention	<ol style="list-style-type: none"> a) Cryopreservation of ovarian tissue for fertility preservation^a b) Cryopreservation of oocytes for fertility preservation
Comparators	<ol style="list-style-type: none"> a) No cryopreservation of ovarian tissue for fertility preservation^a b) No cryopreservation of oocytes for fertility preservation
Outcomes	Rate of pregnancy with biological offspring, rate of childlessness, quality of life/family life
Study design	Randomised trials, cohort studies, or systematic reviews of these study designs
Search period	1990–June 2017
Language	Studies in languages other than English will only be translated if they represent a higher level of evidence than that available in the English language evidence base
Does cryopreservation of ovarian tissue, oocytes or embryos prior to receiving gonadotoxic treatment lead to better family outcomes in female patients compared to no cryopreservation?	

AFC = antral follicle count; AMH = anti-Müllerian hormone; FSH = follicular stimulating hormone

^a The assessment of the effectiveness and safety of cryopreservation of ovarian tissue, prior to receiving gonadotoxic treatment, in female patients aged 0–45 years, compared to no cryopreservation, will be included in MSAC Assessment 1435, and summarised briefly in Assessment 1434.

Table 20 PICO criteria and research question for analytical validity of AMH testing in populations 4, 5 and 6

Population	4. Paediatric female patients aged 0–14 years following completion of gonadotoxic treatment 5. Adolescent/young adult female patients aged 15–25 years following completion of gonadotoxic treatment 6. Adult female patients aged 26–45 years following completion of gonadotoxic treatment
Intervention	AMH test
Comparator	a. Other standard tests (FSH, E2) b. Other AMH test (repeated test, different laboratory, different assay, etc)
Evidentiary standard	AFC ultrasound if post-pubertal
Outcomes	Concordance (per cent positive agreement), unsatisfactory or uninterpretable test results, reliability (intra-observer or intra-instrument variability/agreement)
Search period	1990–June 2017
Language	Studies in languages other than English will only be translated if they represent a higher level of evidence than that available in the English language evidence base
<p>What is the concordance of AFC, AMH, FSH and E2 in female patients following completion of gonadotoxic treatment?</p> <p>What is the concordance between AMH and AFC ultrasound, compared to FSH and E2 in female patients following completion of gonadotoxic treatment?</p> <p>What is the reliability of AMH testing, compared to FSH and E2 in female patients following completion of gonadotoxic treatment?</p>	

AFC = antral follicle count; AMH = anti-Müllerian hormone; E2 = estradiol; FSH = follicular stimulating hormone; SROC = summary receiving operating characteristic

Table 21 PICO criteria and research question to determine the impact of the AMH test on patient management in population 4, 5 and 6

Population	4. Paediatric female patients aged 0–14 years following completion of gonadotoxic treatment 5. Adolescent/young adult female patients aged 15–25 years following completion of gonadotoxic treatment 6. Adult female patients aged 26–45 years following completion of gonadotoxic treatment
Intervention	AMH test in addition to other standard tests
Comparators	Other standard tests (FSH, E2, AFC ultrasound if post-pubertal)
Outcomes	Change in fertility preservation pathway or uptake (cryopreservation of oocytes and/or ovarian tissue undertaken or avoided) Change in treatment e.g. hormone therapy for delayed puberty or premature ovarian insufficiency
Study design	Randomised trials, cohort studies, case series or systematic reviews of these study designs
Search period	1990–June 2017
Language	Studies in languages other than English will only be translated if they represent a higher level of evidence than that available in the English language evidence base
<p>Is there a change in management after AMH testing in female patients following completion of gonadotoxic treatment, compared to FSH, E2 and/or AFC ultrasound if post-pubertal?</p>	

AFC = antral follicle count; AMH = anti-Müllerian hormone; E2 = estradiol; FSH = follicular stimulating hormone

Table 22 PICO criteria and research question to determine the effectiveness of change in management in population 1, 2 and 3

Population	1. Paediatric female patients aged 0–14 years following completion of gonadotoxic treatment 2. Adolescent/young adult female patients aged 15–25 years following completion of gonadotoxic treatment 3. Adult female patients aged 26–45 years following completion of gonadotoxic treatment
Intervention	a) Cryopreservation of ovarian tissue for fertility preservation ^a b) Cryopreservation of oocytes for fertility preservation
Comparators	a) No cryopreservation of oocytes for fertility preservation b) No cryopreservation of ovarian tissue for fertility preservation
Outcomes	Rate of pregnancy with biological offspring, rate of childlessness, quality of life/family life
Study design	Randomised trials, cohort studies, or systematic reviews of these study designs
Search period	1990–June 2017
Language	Studies in languages other than English will only be translated if they represent a higher level of evidence than that available in the English language evidence base
Does cryopreservation of ovarian tissue or oocytes following completion of gonadotoxic treatment lead to better family outcomes in female patients compared to no cryopreservation?	

AFC = antral follicle count; AMH = anti-Müllerian hormone; FSH = follicular stimulating hormone

^a Assessed as part of MSAC assessment 1435 (Part B: females).

A8 CONSUMER IMPACT STATEMENT

Public consultation was sought prior to finalising the PICO Confirmation for application 1434. Of the 78 responses received, most were supportive of AMH testing. A number of advantages and disadvantages, or issues that would need to be addressed, were emphasised during the public consultation period.

The benefits suggested via public consultation were that:

- AMH testing is seen as a way to provide vital information to young patients with regard to their fertility status, allowing them to make informed decisions about potential fertility interventions
- AMH tests can measure the risk for premature ovarian failure and this makes the implementation of early interventions possible, such as HRT or oocyte cryopreservation
- AMH testing could potentially reduce costs by allowing targeted fertility interventions
- listing AMH testing on the MBS would lead to equitable fertility monitoring for all patients, not only the people who can afford it
- listing AMH testing on the MBS could lead to reduced anxiety and improved psychological wellbeing due to improved understanding and less uncertainty around potential fertility function in the future and in relation to medical costs.

The disadvantages, or issues suggested for address, as determined by public consultation were that:

- AMH levels don't predict the quality of the oocytes in reserve
- AMH test results may cause psychological distress for the patient (if levels are low or undetectable)
- pre-test counselling would be critical to minimise the negative psychological impact of an unexpected or negative result.

In addition, one clinician responded that AMH results are difficult to interpret in pre-pubertal children. There is more evidence required to prove that there is a correlation between AMH levels and fertility reserve in this population. It was stated that for "paediatric patients AMH is important in long-term follow-up to assess return of reproductive function and assess need for intervention. It may be useful pre-treatment in selective cases but is not required routinely."

SECTION B

CLINICAL EVALUATION

Determination of the clinical effectiveness of an investigative medical service requires either direct or linked evidence. Direct evidence for AMH testing would high-quality comparative studies evaluating the use of AMH testing and any subsequent change in management compared to other standard tests alone (FSH, E2, AFC ultrasound if post-pubertal). Direct evidence investigates the direct impact of the test on health outcomes. Randomised controlled trials provide the highest quality evidence for this comparison. In the absence of direct evidence, linked evidence may be used. A linked evidence approach provides a combined analysis of the following outcomes:

- analytical validity - how reliable is the test, and how well does the test correlate with other measures of ovarian reserve, such as AFC ultrasound? (Section B3)
- clinical validity (prognostic value) - can the test result predict the ability to conceive? (Section B4)
- clinical validity (predictive value) - can the test result predict response to ovarian stimulation? (Section B4)
- impact on clinical decision-making - does the test result in a change in treatment decisions/fertility preservation decisions? (Section B5)
- clinical utility - does the change in fertility preservation method/decision or patient management change their health or fertility outcomes? (Section B5)

As there was no direct evidence to assess AMH testing, an evidence analysis was undertaken using a linked approach for this assessment.

B1 DIRECT EVIDENCE

B1.1 LITERATURE SOURCES AND SEARCH STRATEGIES

The medical literature was searched on the 6 June 2017 to identify relevant studies and systematic reviews published during the period January 1990 to June 2017. The search included databases and sources described in Appendix B. Attempts were also made to source unpublished or grey literature (Appendix B). Search terms are described in Table 80 and Table 81.

A single set of searches was performed to identify evidence for MSAC assessment 1434 (AMH testing) and 1435 part B (Ovarian cryopreservation) as both assessments are linked and include the same patient population. It was initially proposed that MSAC assessment 1435 part B would be part of the final step of the linked analysis for MSAC assessment 1434.

It was proposed that if evidence was identified regarding how AMH produces a change in patient management, unrelated to fertility preservation, i.e. change in treatment for delayed puberty or premature ovarian insufficiency, then targeted searches would be performed to assess the health impact of these management strategies. As no change in management data in the target population were identified, no additional systematic searches were performed.

B1.2 RESULTS OF LITERATURE SEARCH

A PRISMA flowchart (Figure 5) provides a graphic depiction of literature search results and application of the study selection criteria (Liberati et al., 2009).

Studies were selected by a single reviewer with a second reviewer assessing 10% of the most relevant citations, as determined by an algorithm within Rayyan citation management software. Where there was doubt about study inclusion, a third independent reviewer was consulted.

Studies that could not be retrieved, or that met the inclusion criteria but had data that were insufficient or of inadequate quality for inclusion, are listed as Excluded Studies in Appendix E.

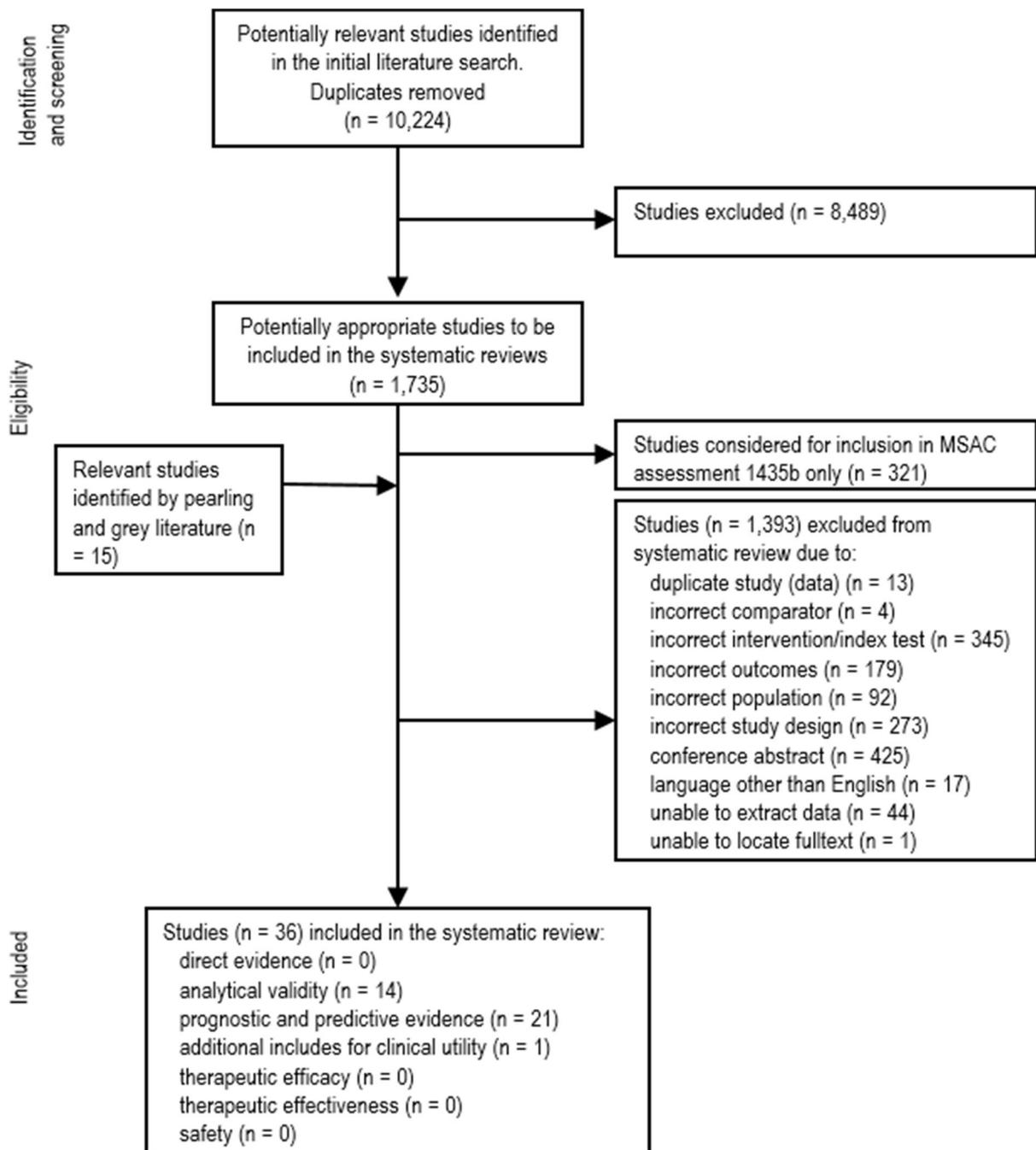


Figure 5 Summary of the process used to identify and select studies for the assessment

A profile of each included study is given in Appendix C. This study profile provides details of the authors, study identification, publication year, study design and quality, i.e. level of evidence and risk of bias, study location, length of follow-up of patients, study population characteristics, description of the test and associated interventions, if relevant, description of the comparator, if available, description of the reference standard or evidentiary standard, and the relevant outcomes assessed.

APPRAISAL OF THE EVIDENCE

Appraisal of the evidence was conducted in four stages:

Stage 1: Risk of bias appraisal for individual studies, or systematic reviews, included in the review. Some risk of bias items were assessed for the study as a whole, while others were assessed at the outcome level. See subsections B1.3, B3.3, B4.1.2, B5.1.1.

Stage 2: Appraisal of the precision, size of effect and clinical importance of the results reported in the evidence base as they relate to the pre-specified primary outcomes for this assessment. See subsections B1.6, B3.6, B4.1.5, B5.1.4, B5.2.4.

Stage 3: Rating the overall quality of the evidence per outcome, across studies, based on the study limitations (risk of bias), imprecision, inconsistency of results, indirectness of evidence, and the likelihood of publication bias. See Evidence profile tables, Appendix D.

Stage 4: Evidence integration across outcomes for conclusions about the net clinical benefit of the test and associated interventions in the context of Australian clinical practice. See Section B.8.

B1.3 RISK OF BIAS ASSESSMENT

No studies meeting the PICO criteria for direct evidence were identified.

B1.4 CHARACTERISTICS OF THE EVIDENCE BASE

No studies meeting the PICO criteria for direct evidence were identified.

B1.5 OUTCOME MEASURES AND ANALYSIS

No studies meeting the PICO criteria for direct evidence were identified.

B1.6 RESULTS OF THE SYSTEMATIC LITERATURE REVIEW

IS IT SAFE?

Summary

What is the safety of having an AMH test to measure ovarian reserve in addition to other standard tests, versus other standard tests alone, for determining the need for cryopreservation of ovarian tissue, oocytes or embryos, and predicting the success of fertility preservation, prior to receiving gonadotoxic treatment?

No studies on the safety of AMH testing were identified.

What is the safety of having an AMH test in addition to other standard tests, versus other standard tests alone, to predict response to superovulation and determine the level of hormones used, prior to receiving gonadotoxic treatment?

No studies on the safety of AMH testing were identified.

What is the safety of having an AMH test in addition to other standard tests to measure ovarian function compared to other standard tests alone in female patients following completion of gonadotoxic treatment?

No studies on the safety of AMH testing were identified.

No studies meeting the PICO criteria for direct evidence were identified.

IS IT EFFECTIVE?

Summary

What is the effectiveness of having an AMH test to measure ovarian reserve in addition to other standard tests, versus other standard tests alone, for determining the need for cryopreservation of ovarian tissue, oocytes or embryos, and predicting the success of fertility preservation, prior to receiving gonadotoxic treatment?

No studies meeting the PICO criteria for direct evidence were identified.

What is the effectiveness of having an AMH test in addition to other standard tests, versus other standard tests alone, to predict response to superovulation and determine the level of hormones used, prior to receiving gonadotoxic treatment?

No studies meeting the PICO criteria for direct evidence were identified.

What is the effectiveness of having an AMH test in addition to other standard tests to measure ovarian function compared to other standard tests alone in female patients following completion of gonadotoxic treatment?

No studies meeting the PICO criteria for direct evidence were identified.

No studies meeting the PICO criteria for direct evidence were identified.

B2 LINKED EVIDENCE APPROACH

B2.1 BASIS FOR LINKED EVIDENCE

No direct evidence was identified; thus, a linked evidence approach was undertaken.

B2.2 STEPS FOR LINKED ANALYSIS

To construct a linked evidence analysis, the different evidentiary requirements considered are:

- the prognostic performance and clinical validity, where relevant, of the investigative medical service (Sections B3 and B4)
- the clinical utility of the investigative medical service, i.e. the impact of test results on patient management, the contribution and clinical importance of false negatives versus false positives, and the direct impact of each therapeutic model service option on health outcomes (Section B5)
- the relative safety of performing the investigative service, both the immediate safety issues of directly performing the test and the flow on safety issues that arise as a result of conducting the investigative service (section B7).

Conclusions linking the different steps of the linked evidence approach can be found in section B8.

B3 ANALYTICAL VALIDITY

This section presents the evidence directly comparing AMH tests with one or more of the comparator tests, to determine the concordance and reliability of the tests, i.e. analytical validity. Studies investigating whether the tests predict a certain outcome in the future, with the reference standard measured after a certain follow-up period, are discussed in section B4.1.5. This is the prognostic and predictive value of AMH testing. These outcomes included ovarian function, oocyte yield for cryopreservation, and pregnancy.

B3.1 REFERENCE STANDARD

There is no true reference standard for measurement of ovarian reserve. The evidentiary standard as per the PICO for determining ovarian reserve in post-pubertal women before or after treatment with gonadotoxic treatments is AFC via transvaginal ultrasound. However, the included literature often used a clinical reference standard for ovarian failure, defined as the absence of menses for at least one year. Studies using the clinical reference standard were also included.

B3.2 LITERATURE SOURCES AND SEARCH STRATEGIES

See Section B.1 for details.

B3.3 RESULTS OF LITERATURE SEARCH

Two studies were identified that provided sensitivity and specificity data for AMH tests, with CRA as the clinical reference standard. Decanter et al. compared the ability of the AMH and AFC tests to diagnose resumption of menses in women who underwent chemotherapy for either breast cancer or haematological malignancies (Decanter et al. 2014). Su and colleagues investigated the ability of AMH, AFC, FSH and inhibin B tests to diagnose CRA in post-chemotherapy breast cancer survivors (Su et al. 2011). The characteristics of these two analytical validity studies are summarised in Table 23. A full profile of each included study is also included in Appendix C.

Table 23 Key features of the included evidence comparing intervention with comparator against reference standard

Trial/study	N	Level of evidence	Risk of bias	Patient population	Key outcome(s)	Result used in meta-analysis
Decanter et al. (2014)	58	III-2	Low	Women who had had chemotherapy	Sensitivity/ Specificity	Not used
Su et al. (2011)	55	III-2	Low	Female post-chemotherapy breast cancer survivors	Sensitivity/ Specificity	Not used

III-2 = a comparison with non-blinded reference standard

Miyoshi et al (2013) compared the ability of the AMH test to predict high FSH levels in women who had survived childhood cancer. This was the only study to provide the concordance of two tests to identify a clinically relevant outcome. Eleven studies reported on the correlation between AMH levels and the AFC or other hormones (Table 24).

Additionally, two studies reported on the correlation between different AMH tests (Table 24).

Table 24 Studies reporting on the correlation between the AMH test and other tests

Comparison	Studies reporting correlation
Two different AMH tests	de Souza et al. (2015), Su et al. (2014)
AMH vs AFC	Biacchiardi et al. (2011), Lee et al. (2011), Nielsen et al. (2013), Paradisi et al. (2016), Partridge et al. (2010)
AMH vs density of primordial follicles	Fabbri et al. (2014)
AMH vs FSH	Decanter et al. (2014), Fabbri et al. (2014), Kim et al. (2016), Lee et al. (2011), Lutchman Singh et al. (2007), Paradisi et al. (2016), van Beek et al. (2007)
AMH vs inhibin B	Beneventi et al. (2014), Kim et al. (2016), Lee et al. (2011), Lutchman Singh et al. (2007), Paradisi et al. (2016), van Beek et al. (2007)
AMH vs E2	Fabbri et al. (2014), Kim et al. (2016), Lutchman Singh et al. (2007), Paradisi et al. (2016)

AFC = antral follicle count; AMH = anti-Müllerian hormone; E2 = estradiol; FSH = follicle-stimulating hormone

B3.4 RISK OF BIAS ASSESSMENT

The Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool was used to assess the risk of bias (Whiting et al. 2011). It was modified to include a comparator for the studies assessing diagnostic accuracy data and the concordance study. The results are summarised in Table 25. Su et al. (2011) included FSH and inhibin B as comparators to AMH. Decanter et al. (2014) compared AMH tests to the clinical reference standard, CRA. According to QUADAS-2, the risk of bias associated with the AMH index test and the comparators is unclear. This is because the studies were not blinded, though because these tests are immunoassays, mostly ELISAs, and reading the result is automated, the risk of bias is considered low. However, there may be applicability issues when using these results for a linked evidence analysis. These studies used clinical reference standards that were valid clinical outcomes for interventional studies, but they were not included in the PICO for analytical validity studies. The concordance study did not include a reference standard.

Table 25 Modified QUADAS-2 risk of bias results for analytical validity and concordance studies

Study	Risk of bias				Applicability				
	Patient selection	Index test	Comparator	Reference standard	Flow and timing	Patient selection	Index test	Comparator	Reference standard
Decanter et al. (2014)	😊	?	?	😊	😊	😊	😊	😊	⊗
Su et al. (2011)	😊	?	😊	?	😊	😊	😊	😊	⊗
Miyoshi et al. (2013)	😊	?	😊	NA	😊	😊	😊	😊	NA

😊 = low risk; ⊗ = high risk; ? = unclear risk; NA = not applicable

Table 26 and Table 27 summarise the risk of bias for patient selection and for each test conducted in the concordance studies. Again, the risk of bias associated with the biochemical tests (AMH, FSH, E2 and inhibin B) is unclear due to the lack of blinding in all but the study by de Souza et al. (2015). However, as these tests are immunoassays, mostly ELISAs, and the reading of the result is automated, the risk of bias is low. The studies with a high risk of bias for the FSH test, inhibin B test or both these tests did not provide details about the testing.

Table 26 Modified QUADAS-2 risk of bias results for studies reporting correlations between AMH and other tests

Study	Patient selection	AMH test	AFC test	Ovarian biopsy	FSH test	E2 test	Inhibin B test
Beneventi et al. (2014)	😊	?	ND	ND	ND	ND	?
Biacchiardi et al. (2011)	😊	?	😊	ND	ND	ND	ND
Decanter et al. (2014)	😊	😊	ND	ND	⊗	ND	ND
Fabbri et al. (2014)	😊	?	ND	😊	?	?	ND
Kim et al. (2016)	😊	😊	ND	ND	?	?	?
Lee et al. (2011)	😊	?	😊	ND	⊗	ND	⊗
Lutchman Singh et al. (2007)	😊	😊	ND	ND	?	?	?
Nielsen et al. (2013)	😊	?	😊	ND	ND	ND	ND
Paradisi et al. (2016)	😊	?	😊	ND	?	?	?
Partridge et al. (2010)	😊	?	😊	ND	ND	ND	ND
van Beek et al. (2007)	😊	😊	ND	ND	?	ND	?

😊 = low risk; ⊗ = high risk; ? = unclear risk; ND = not done

Table 27 Modified QUADAS-2 risk of bias results for studies reporting the correlation between different AMH tests

Study	Patient selection	AMH Gen II ELISA	Ansh Labs AMH ELISA	Ultrasensitive AMH ELISA	Pico-AMH ELISA
de Souza et al. (2015)	😊	😊	😊	ND	ND
Su et al. (2014)	😊	😊	ND	?	?

😊 = low risk; ⊗ = high risk; ? = unclear risk; ND = not done

B3.5 CHARACTERISTICS OF THE EVIDENCE BASE

Appendix C summarises the study characteristics for individual studies included in the evidence base.

All except two studies enrolled patients with cancer, the most common form being early breast cancer (Stages I–III). Five studies enrolled patients with non-gynaecological malignancies (Fabbri et al. 2014; Paradisi et al. 2016), early breast cancer (Lee et al. 2011; Su, HC et al. 2014) and ovarian endometriomas (Biacchiardi et al. 2011) before treatment. Five studies enrolled patients with early breast cancer (Kim, HA et al. 2016; Partridge et al. 2010; Su et al. 2011), polymyositis (de Souza et al. 2015) and either breast cancer or lymphoma (Decanter et al. 2014). Four studies enrolled patients who had survived childhood cancers (Beneventi et al. 2014; Miyoshi et al. 2013; Nielsen et al. 2013; van Beek et al. 2007), and one study enrolled women with early breast cancer both before and after treatment (Lutchman Singh et al. 2007). Thus, the included populations all met the PICO criteria.

The studies used a variety of different AMH tests, which are summarised in Table 28. These tests all had different precision parameters, including limit of detection and inter-assay coefficient(s) of variation (CoV). The comparator tests also varied between studies (see Appendix C for details). However, the AFC test, which was considered as the evidentiary standard, was conducted in a similar fashion in all studies that included this test.

Table 28 Summary of the AMH tests used in the included studies

AMH Assay	Manufacturer	Inter-assay CoV	LoD
EIA AMH/MIS ELISA	Beckman Coulter, Immunotech, Marseille, France	14.2%, 13.0% and 12.6 % at 1.37, 2.61 and 5.20 ng/mL	3 pmol/L, 0.42 ng/mL
DSL ACTIVE® MIS/AMH ELISA	Diagnostic Systems Laboratories, Webster, TX	8.0%, 4.8% and 6.7% at 0.164, 0.917 and 4.527 ng/mL	0.63 pmol/L, 0.09 ng/mL
AMH Gen II ELISA	Beckman Coulter, Brea, CA	5.6% and 4.5% at 4.42 and 14.0 ng/mL	0.57 pmol/L, 0.08 ng/mL
USCN AMH ELISA	USCN Life Science Inc, Houston, TX	<12% over 0.062–5.0 ng/mL range	0.2 pmol/L, 0.03 ng/mL
Ultrasensitive AMH/MIS ELISA	Ansh Labs, Webster, TX	4.6%, 4.8% and 2.0% at 0.346, 0.715 and 1.853 ng/mL	0.16 pmol/L, 0.023 ng/mL
Automated Access AMH test	Beckman Coulter	3.04-5.76% over 0.16-10.0 ng/mL range	0.14 pmol/L, 0.02 ng/mL
Elecsys AMH assay (Automated Cobas test)	Roche Diagnostics, Rotkreuz, Switzerland (headquarters)	<2.2% across 0.95–5.19 ng/mL range	0.07 pmol/L, 0.01 ng/mL
Pico-AMH ELISA	Ansh Labs, Webster, TX	4.5%, 2.3% and 3.8% at 0.023, 0.087 and 0.373 ng/mL	0.0086 pmol/L, 0.0012 ng/mL

AMH = anti-Müllerian hormone; CoV = coefficient(s) of variation; EIA = enzyme Immunoassay; ELISA = enzyme-linked immunosorbent assay; LoD = limit of detection; MIS = Müllerian-inhibiting substance
Note: 1 ng/mL AMH is equivalent to 7.14 pmol/L

B3.6 OUTCOME MEASURES AND ANALYSIS

The outcomes measured in the included studies, along with the statistical methods used to analyse the results, are included in Appendix C.

ANALYTICAL VALIDITY

To assess the analytical validity of the proposed test, studies were only included if they provided data that could be extracted into a classic 2 × 2 table, in which the results of the index test or the comparator were cross-classified against the results of the reference standard³ using Bayes' Theorem (Table 29). This means that for each study reporting outcomes for analytical validity a threshold needs to be applied, such that AMH levels above the threshold are considered positive,

³ Armitage, P, Berry, G & Matthews, JNS 2002, *Statistical methods in medical research*, fourth edn, Blackwell Science, Oxford.

Deeks, JJ 2001, 'Systematic reviews of evaluations of diagnostic and screening tests', in M Egger, G Davey Smith & DG Altman (eds), *Systematic Reviews in Healthcare: Meta-Analysis in Context*, second edn, BMJ Publishing Group, London, pp. 248–282.

while those below the threshold are considered negative. In some studies a detectable level of AMH was considered positive.

Table 29 Analytical validity data extraction

		Clinical reference standard		
		<i>Have outcome</i>	<i>Do not have outcome</i>	
Index test	<i>Test +</i>	true positive	false positive	Total test positive
Or comparator	<i>Test -</i>	false negative	true negative	Total test negative
		Total with outcome	Total without outcome	

Test sensitivity was calculated as the proportion of people with the clinical outcome, determined by the reference standard, who were test positive:

Sensitivity (true-positive rate) = number with true-positive result / total with clinical outcome

Test specificity was calculated as the proportion of people without the clinical outcome, determined by the reference standard, who were test negative:

Specificity (true-negative rate) = number with true-negative result / total without clinical outcome

The 95%CI was calculated by exact binomial methods.

RECEIVER OPERATOR CHARACTERISTIC (ROC) ANALYSIS

For a given diagnostic test, the true-positive rate (TPR) against false-positive rate (FPR) can be plotted on a ROC curve, where:

TPR = number with true-positive result / total with clinical outcome = sensitivity

FPR = number with false-positive result / total without clinical outcome = 1 - specificity

The ROC space comprises all possible combinations of the TPR and the FPR, and the position of a point in the ROC space shows the trade-off between sensitivity and specificity. The AUC is the average sensitivity given that all values of specificity are equally likely. The AUC serves as a global measure of test performance. Given two randomly chosen people, one with and one without a disease or condition, the AUC can be interpreted as the probability that the diagnostic test will rank suspicion of disease higher in the person with the disease compared to the disease-free individual.

In this assessment, the AUC is interpreted according to the following test performance cut-offs:

- >0.9 = very good
- 0.8–0.9 = good
- 0.7–0.8 = moderate
- <0.7 = poor

CORRELATION

To assess the concordance between the AMH test and the comparator tests, data were extracted into a classic 2 × 2 table (Table 30).

Table 30 Concordance data extraction

		Test 1		
		Test 1 +	Test 1 -	
Test 2	Test 2 +	Positive with both tests (a)	Negative with test 1 Positive with test 2 (b)	Total test 2 positive
	Test 2 -	Positive with test 1 Negative with test 2 (c)	Negative with both tests (d)	Total test 2 negative
		Total test 1 positive	Total test 1 negative	

The overall, positive and negative per cent agreement between the tests were calculated. Overall per cent agreement is defined as the number of people who are either positive or negative for both tests divided by the total number of people. Positive per cent agreement is defined as the number of people who are positive for both tests divided by the number who are positive for either test. Negative per cent agreement is defined as the number of people who are negative for both tests divided by the number who are negative for either test.

$$\text{Overall per cent agreement} = (a + d) / (a + b + c + d)$$

$$\text{Positive per cent agreement} = a / (a + b + c)$$

$$\text{Negative per cent agreement} = d / (b + c + d)$$

The correlation studies used two different correlation coefficients, either Spearman's rank correlation coefficient (ρ) or Pearson's bivariate correlation coefficient (r).

Spearman's rank correlation coefficient measures the strength and direction of association between two ranked variables, and ranges between zero (no tendency for Y to increase or decrease when X increases) and ± 1 (when X and Y are perfectly monotonically related). The sign of the Spearman correlation indicates the direction of association between X and Y:

- if Y tends to increase when X increases, ρ is positive
- if Y tends to decrease when X increases, ρ is negative.

Pearson's bivariate correlation coefficient is a measure of the linear correlation between two variables. It has a value between +1 and -1, where:

- 1 is total positive linear correlation
- 0 is no linear correlation
- -1 is total negative linear correlation.

IS IT ACCURATE?**Summary****Analytical validity of AMH tests compared with a clinical reference standard**

Decanter et al. (2014) reported that the sensitivity of the EIA AMH/MIS assay was only 11% (95%CI 2, 28), compared with the pico-AMH ELISA of 71% (95%CI 51, 87). This can be explained by the 40-fold difference in the level of AMH detectable by the two AMH tests, and shows that if the presence of detectable AMH in the serum is used to determine ovarian recovery in women soon after completion of chemotherapy, a sensitive AMH test is needed to detect very low AMH levels. Even then, about one in three samples of women who will resume menstruation had a false-negative result in this study (undetectable AMH). The specificity for both tests was 93% at the 3-month follow-up, suggesting that undetectable AMH was a strong indicator for non-resumption of menses. Only 7% of samples of women who had not resumed menstruation were false positive (i.e. detectable AMH levels).

ROC analysis of the ability of the various hormones to discriminate between women with and without menses

Su et al. (2011) measured AFC and serum levels of AMH, FSH and inhibin B to identify CRA in 56 female, post-chemotherapy late reproductive-aged breast cancer survivors. The study used ROC analysis and found that only the AFC test had an AUC above 0.8, indicating good accuracy. However, they reported that the combination of the AFC test with the AMH or FSH tests improved the accuracy compared to the AFC test alone.

Su et al. (2011) also calculated the sensitivity and specificity of the tests using the threshold values determined by ROC curve analysis. AFC was more specific (89% versus 60–64%) than the AMH, FSH and inhibin B tests. Thus, there were less women having a false-positive result indicating CRA with AFC (11%) than with the three hormone tests (36–40%). However, the sensitivity of the AFC, AMH and FSH tests did not vary greatly (76–79%). 21–24% of women who had CRA would have had a false-negative result.

What is the concordance of AMH, FSH and E2 in female patients prior to receiving or following completion of gonadotoxic treatment?

Eleven studies reported correlations between AMH and FSH. Two studies found AMH serum levels to be negatively correlated with FSH, and another found a negative correlation between detectable/undetectable AMH and FSH levels. This was expected, as per Figure 1. Five studies found no significant correlation between AMH and FSH serum levels.

Three studies found a positive correlation between AMH and E2 levels, as expected according to Figure 1. Studies that included blood samples from the early follicular phase found a negative or no correlation between the two tests. The reason for these discrepancies is unclear.

What is the concordance between AMH and AFC ultrasound, compared to FSH and E2 in female patients prior to receiving or following completion of gonadotoxic treatment?

All included studies (k=8) found a positive correlation between AFC and AMH serum levels. Two studies showed a stronger correlation between AFC and AMH, compared to AFC and FSH or E2.

What is the reliability of AMH testing, compared to FSH and E2 in female patients prior to receiving or following completion of gonadotoxic treatment?

No studies were identified investigating the reliability of AMH compared to FSH and E2. However, high intra- and inter-cycle variability has been observed for FSH and E2 levels (American Society for Reproductive Medicine Practice Committee 2015).

Analytical validity of two AMH tests compared with a clinical reference standard

The study by Decanter et al. (2014) compared the sensitivity and specificity of two AMH tests (EIA AMH/MIS assay and pico-AMH ELISA) to detect ovarian recovery by the resumption of menses during the ovarian recovery period in 30 women with either early breast cancer or lymphoma, 3 to 24 months after the end of chemotherapy. The distribution of the serum samples was as follows: 9 were taken at 3-months post-chemotherapy, 13 at 6-months, 11 at 9-months, 18 at 12-months, 2 at 18-months, and five at 24-months.

For this study, results under the first calibrator value were expressed as undetectable. This value was 3.0 pmol/L (0.42 ng/mL) for the EIA AMH/MIS assay and 0.07 pmol/L (0.01 ng/mL) for the pico-AMH ELISA. Samples were selected from women according to the menstrual status, such that they constituted two equally sized groups at the time of sampling. The accuracy results at the time of sampling, including 28 samples from women who had resumption of menses, are summarised in Table 31.

Table 31 Analytical validity of the EIA AMH/MIS assay and pico-AMH ELISA against the clinical reference standard of resumption of menses

Study ID	Result	EIA AMH/MIS assay [95%CI]	Pico-AMH ELISA [95%CI]	Difference
Decanter et al. (2014) Undetectable AMH versus RS at time of AMH	Sensitivity	11% [2, 28]	71% [51, 87]	60%
	Specificity	93% [78, 99]	83% [65, 94]	-7%

AMH = anti-Müllerian hormone; CI = confidence interval; EIA = enzyme immunoassay; MIS = Müllerian-inhibiting substance

There was a 40-fold difference in the level of detectable AMH between the two AMH tests. Thus, it is not surprising that more samples had undetectable levels with the less sensitive EIA AMH/MIS assay (53 samples) compared with the pico-AMH ELISA (33 samples). The sensitivity of the EIA AMH/MIS assay was only 11% (95%CI 2, 28), compared with the pico-AMH ELISA of 71% (95%CI 51, 87). Thus, if the presence of detectable AMH in the serum is used to determine ovarian recovery as defined by the return of menses in women who have completed chemotherapy, a sensitive AMH test is needed

to detect very low AMH levels in the early recovery period. Even then, about 30% of samples from women who will resume menstruation were falsely negative (undetectable AMH). However, undetectable AMH was a strong indicator for non-resumption of menses for both tests.

ROC analysis of the ability of the various hormones to discriminate between women with and without menses

Su et al. (2011) conducted a retrospective study and used Poisson regression methods to model the optimum thresholds for AFC as measured by transvaginal ultrasound, and serum levels of AMH, FSH and inhibin B to identify CRA in 56 female, post-chemotherapy late reproductive-aged breast cancer survivors of at least 1-year duration. CRA was defined as at least 12 months of amenorrhea occurring after start of chemotherapy. Thirty-four women (60.7%) met criteria for CRA. ROC curves were generated for each model as shown in Table 32. This study used an older first generation AMH test, the DSL ACTIVE® AMH/MIS ELISA, which had a lower limit of detection for AMH of 25 pg/mL, according to the manufacturer. This limit coincided with the optimum threshold to distinguish women with and without CRA. Only the AFC test had an AUC above 0.8, indicating good test performance in identifying women with CRA. The tests for AMH and FSH showed moderate test performance, but inhibin B had poor test performance in identifying women with CRA. The combination of the AFC test with the AMH or FSH tests improved the performance of the AFC test alone.

Table 32 AUC values from ROC analysis conducted by Su et al. (2011)

Single-variable models	AUC	Two-variable models	AUC
AFC (total 2–10 mm) <1	0.82	AFC <1 + AMH ≤25 pg/mL	0.87
AMH ≤25 pg/mL	0.71	AFC <1 + FSH ≥40IU/L	0.87
FSH ≥40IU/L	0.72	AMH ≤25 pg/mL + FSH ≥40IU/L	0.74
Inhibin B ≤5 pg/mL	0.63	-	-

AFC = antral follicle count; AMH = anti-Müllerian hormone; AUC = area under the curve; FSH = follicle-stimulating hormone; E2 = estradiol; ROC = receiver operator characteristic

Su et al. (2011) also calculated the sensitivity and specificity of the tests using the threshold values determined by ROC curve analysis (Table 33). AFC was more specific (89% versus 60–64%) than the AMH, FSH and inhibin B tests in distinguishing late reproductive-aged women with and without CRA, indicating less women had a false-positive result indicating CRA with AFC (11%) than with the three hormone tests (36–40%). However, the sensitivity of the AFC, AMH and FSH tests did not vary greatly (76–79%), but 21 to 24% of women who had CRA would have had a false-negative result.

Table 33 Analytical validity of ovarian reserve markers using thresholds determined by ROC curve analysis

	AFC (total 2–10 mm) <1	AMH ≤25 pg/mL	FSH ≥40IU/L	Inhibin B ≤5pg/mL
Specificity	89%	60%	64%	64%
Sensitivity	79%	76%	78%	54%

AFC = antral follicle count; AMH = anti-Müllerian hormone; FSH = follicle-stimulating hormone; NPV = negative predictive value; PPV = positive predictive value; ROC = receiver operator characteristic

Source: Su et al. (2011)

B3.8 EXTENDED ASSESSMENT OF RELIABILITY EVIDENCE

Two studies reported on the correlation between different AMH tests, which were calculated using Spearman’s rank correlation coefficient test. Su et al. (2014) compared the AMH Gen II ELISA, Ultrasensitive AMH ELISA and pico-AMH ELISA tests, and de Souza et al. (2015) compared the AMH Gen II ELISA and the Ansh Labs AMH ELISA tests. The correlations between the tests are shown in Table 34. The median correlation for all four comparisons is 0.942 (range 0.92–0.99). Thus, there is a high degree of correlation between these four tests.

Su et al. (2014) reported a systematic bias between the assays that should be taken into account when converting AMH values between assays. They reported that a measurement of 1.0 ng/mL by the AMH Gen II ELISA would correspond to 1.94 ng/mL when using the Ultrasensitive AMH ELISA, and 1.77 ng/mL when using the pico-AMH ELISA.

Table 34 Spearman’s rank correlation between two AMH assays

Study, comparison	Population	Correlation
de Souza et al. (2015) AMH Gen II ELISA versus Ansh Labs AMH ELISA	8 female polymyositis patients being treated or previously treated with immunosuppressive agents and corticosteroids	rho=0.964; p<0.0001
Su et al. (2014) AMH Gen II ELISA versus Ultrasensitive AMH ELISA	90 newly diagnosed breast cancer patients prior to cancer treatment	rho=0.92; p<0.001
Su et al. (2014) AMH Gen II ELISA versus pico-AMH ELISA	90 newly diagnosed breast cancer patients prior to cancer treatment	rho=0.92; p<0.001
Su et al. (2014) Ultrasensitive AMH ELISA versus pico-AMH ELISA	90 newly diagnosed breast cancer patients prior to cancer treatment	rho=0.99; p<0.001

AMH = anti-Müllerian hormone; ELISA = enzyme-linked immunosorbent assay

B3.9 CONCORDANCE ANALYSIS

OVERALL PER CENT AGREEMENT

Miyoshi et al. (2013) provided 2 × 2 data on the concordance of the EIA AMH/MIS ELISA and the Access FSH chemiluminescent EIA tests to detect patients with low AMH and high FSH values in 53 female childhood cancer survivors (Table 35). Sixteen women had high FSH levels and 28 had low AMH levels (i.e. <5 pmol/L). The overall, positive and negative per cent agreements between low AMH and high FSH levels were 77.4%, 57.1% and 67.6%, respectively.

Table 35 2 × 2 concordance data from Miyoshi et al. (2013)

	High FSH	Low-normal FSH	
Low AMH	16	12	28
Normal AMH	0	25	25
	16	37	53

Overall per cent agreement = 41/53 = 77.4%
Positive per cent agreement = 16/28 = 57.1%
Negative per cent agreement = 25/37 = 67.6%

AMH = anti-Müllerian hormone; FSH = follicle-stimulating hormone

CORRELATION BETWEEN TESTS

Eleven studies reported on the correlation between AMH testing and other tests providing information for determining ovarian reserve using either Spearman's rank correlation test or Pearson's bivariate correlation test. The results from individual studies for each comparison are shown in Appendix G and a summary of the findings is provided in Table 36.

Table 36 Summary of the correlation between the AMH test and other tests used to determine ovarian function

Comparison	Correlation test	Median correlation (range)
AMH vs AFC	Spearman's rank correlation	Positive: 0.62 (0.44–0.83), k=4 Pre-cancer: 0.48 (0.44–0.52), k=2 Post-cancer: 0.72, k=1 Post-childhood cancer: 0.83, k=1
	Pearson's bivariate correlation	Pre-endo surgery: 0.842 (k=1)
AMH vs density of primordial follicles	Pearson's bivariate correlation	Positive pre-cancer: 0.23 (k=1)
AMH vs FSH	Spearman's rank correlation	Negative: -0.50 (-0.52, -0.47; k=2), NS (k=3) Pre-cancer: -0.47, k=1, NS, k=2 Post-cancer: NS, k=1 Post-childhood cancer: -0.52, k=1
	Pearson's bivariate correlation	Serum AMH vs serum FSH levels: NS, k=3 Detectable AMH vs serum FSH: -0.55, k=1
AMH vs E2	Spearman's rank correlation	Positive: 0.30 (0.15–0.44), k=2 Negative: -0.64, k=1 Pre-cancer -0.64 and 0.15, k=2
	Pearson's bivariate correlation	Positive post-cancer: 0.44, k=1 Pre- and post-cancer: NS (k=1)
AMH vs inhibin B	Spearman's rank correlation	Positive: 0.72 (0.44–0.87), k=5 Pre-cancer: 0.32 (0.22–0.42), k=2 Post-cancer(4.8 years): 0.52, k=1 Post-childhood cancer: 0.44 (0.43–0.44), k=2
	Pearson's bivariate correlation	Positive pre- and post-cancer: 0.84, k=1

AFC = antral follicle count; AMH = anti-Müllerian hormone; FSH = follicle-stimulating hormone; E2 = estradiol; k = number of studies

All included studies found AMH serum levels to be positively correlated with AFC, density of primordial follicles, and inhibin B serum levels. When the serum levels of AMH and inhibin B are compared with AFC and ovarian reserve in Figure 1, they all trend downwards as women approach menopause. Thus, a positive correlation is to be expected and confirms the ability of the AMH test to detect decreased AMH serum levels relative to the decreases detected by the other tests.

Two studies found AMH serum levels to be negatively correlated with FSH serum levels, and another study that did not find a significant correlation between AMH and FSH serum levels, found a negative correlation between detectable/undetectable AMH and FSH serum levels. From Figure 1, it can be seen that as AMH levels decrease prior to menopause, FSH levels increase, especially after 40 years of age. Thus, a negative correlation would be expected. However, five studies found no significant correlation between AMH and FSH serum levels. The reasons for these discrepancies are unclear. Of the four studies that used early follicular phase blood samples for testing, two showed a correlation and two showed no significant correlation. Of the three studies that used blood samples from any stage of the menstrual cycle, or did not report the stage at which samples were taken, one showed a correlation between detectable/undetectable AMH and FSH serum levels.

Three studies found a positive correlation between AMH and E2 serum levels, as expected in accordance with Figure 1. Both these studies used blood samples taken at any time during the menstrual cycle. The two studies that compared AMH and E2 serum levels from early follicular phase blood samples found either a negative correlation or no significant correlation. Again, the reason for these discrepancies is unclear.

Two studies reported both the concordance between AMH and AFC, and the concordance between FSH or E2 and AFC (Lee et al. 2011; Paradisi et al. 2016). Both studies used the Spearman rank correlation test and showed a stronger correlation between AMH and AFC ($\rho=0.44$; $p<0.005$ and $\rho=0.52$; $p<0.001$, respectively), compared to FSH and AFC (ρ non-significant and $\rho=0.02$; $p=0.776$, respectively). One study also compared E2 with AFC, yielding a ρ value of 0.15 ($p=0.040$), compared to 0.52 ($p<0.001$) for AMH compared to AFC.

B3.10 INTERPRETATION OF EVIDENCE ON ANALYTICAL VALIDITY

The information on the analytical validity of the AMH test focuses on whether the test results are reliable, and how they correlate with other tests which are currently in clinical practice.

This is a slightly different concept to prognostic value or clinical validity, which is whether AMH testing can predict who will have ovarian failure or other clinically relevant outcomes such as the ability to conceive in the future, including a time element.

Studies which used the presence or absence of menses as the reference standard found that the accuracy of AMH testing varied largely between different assays, and that about 30% of women who will resume menstruation had a false-negative result (undetectable AMH).

ANALYTICAL VALIDITY OF AMH TESTING COMPARED WITH A CLINICAL REFERENCE STANDARD

One study reported on the analytical validity of the pico-AMH ELISA test and the EIA AMH/MIS assay compared to menstrual status as a clinical reference standard. The sensitivity of the pico-AMH ELISA was significantly higher (71% compared to 11%) than the older EIA AMH/MIS assay. The vast

difference in sensitivity between the two AMH tests is likely due to the 40-fold difference in the level of AMH detectable by these tests with the pico-AMH ELISA test being more sensitive.

Many women had very low AMH levels in the early ovarian recovery period. Thus, if the presence of detectable AMH in the serum is used to determine ovarian recovery in women soon after completion of chemotherapy, a sensitive AMH test is needed to detect these very low AMH levels. Even using a sensitive test, about 30% of women who will resume menstruation had a false-negative result (undetectable AMH).

COMPARING THE ACCURACY OF DIFFERENT AMH TESTS

Two studies compared the AMH Gen II ELISA test with the Ansh Labs AMH ELISA, Ultrasensitive AMH ELISA, and pico-AMH ELISA, as well as the Ultrasensitive AMH ELISA with the pico-AMH ELISA. There was a high degree of correlation between these four tests. One study reported a systematic bias between the assays that should be taken into account when converting AMH values between assays, noting a measurement of 1.0 ng/mL by the AMH Gen II ELISA would correspond to:

- 1.94 ng/mL when using the Ultrasensitive AMH ELISA
- 1.77 ng/mL when using the pico-AMH ELISA.

However, there is concern that although different AMH assays are highly correlated, the results are not generalisable. The standard curves of different AMH assays are not always parallel, and no universally applicable conversion factor exists to standardise results. Furthermore, the thresholds developed and reported for one commercial AMH assay are not generalisable to other commercial assays.

One study showed that while each laboratory showed good reproducibility when using a single test, the between-laboratory variability, even using the same assay, was 40%. Variability of results between laboratories can also be due to the method used to calibrate clinical thresholds, which may differ according to the patient population and clinical outcomes used and laboratory specimen handling procedures, which may affect the serum AMH values.

The major challenge for clinicians attempting to interpret AMH values for use in clinical care is the lack of standardisation. It is recommended that clinicians should always use the same laboratory to avoid problems interpreting results. It is also critical to understand how that laboratory calibrates their clinical thresholds to ensure accurate interpretation. A rigorous quality assurance program would be essential for implementation and development of standardised clinically relevant thresholds for AMH testing in Australia.

ACCURACY OF AMH TESTING COMPARED WITH AFC

ROC analysis comparing AMH serum levels and AFC with the menstrual status of women after chemotherapy found that the AUC for the first generation AMH assay, the DSL ACTIVE® AMH/MIS ELISA and AFC were 0.71 and 0.82, respectively. This means the DSL ACTIVE® AMH/MIS ELISA test showed moderate performance in detecting the menstrual status of women, compared with good performance for AFC. When the AMH and AFC results were combined the AUC increased to 0.87. The sensitivity of the AFC and AMH tests did not vary greatly (79 and 76%), but 21% and 24% of women who had CRA would have had a false-negative result (high AMH and/or AFC), respectively.

Less women without CRA had a false-positive result indicating CRA with AFC (low count; 11%) than with AMH (low serum level; 40%).

Among the included correlation studies, five studies found AMH serum levels to be positively correlated with AFC, while none failed to report a significant correlation between AMH serum levels and AFC.

ACCURACY OF AMH TESTING COMPARED WITH FSH TESTING

Miyoshi et al. (2013) found that the overall, positive and negative per cent agreements between low AMH serum levels measured by the EIA AMH/MIS ELISA test and high FSH serum levels were 77.4%, 57.1% and 67.6%, respectively. Thus, approximately 23% of patients would have discordant AMH and FSH results.

The AUC for the DSL ACTIVE® AMH/MIS ELISA first generation test and FSH were 0.71 and 0.72, respectively. This means both the DSL ACTIVE® AMH/MIS ELISA and FSH test showed moderate performance in detecting the menstrual status of women. When the AMH and FSH results were combined the AUC increased to 0.74, indicating moderate performance. The sensitivity and specificity of the AFC and AMH tests did not vary greatly, but 22 to 24% of women who had CRA would have had a false-negative result (high AMH and/or AFC), and 36 to 40% would have had a false-positive result indicating CRA (low AMH and/or AFC).

Among the included correlation studies, two found AMH serum levels were negatively correlated with FSH serum levels, one study that did not find a significant correlation between AMH and FSH serum levels, found a negative correlation between detectable/undetectable AMH and FSH serum levels, and five studies found no significant correlation between AMH and FSH serum levels.

The reasons for these discrepancies are unclear. They did not coincide with timing of the blood sample used for testing. Of the four studies that used early follicular phase blood samples for testing, two showed a correlation and two showed no significant correlation. Of the three studies that used blood samples from any stage of the menstrual cycle, or did report the stage at which sampling occurred, one showed a correlation between detectable/undetectable AMH and FSH serum levels.

ACCURACY OF AMH TESTING COMPARED WITH E2 TESTING

Among the included correlations studies, two found AMH serum levels to be positively correlated with E2 serum levels, one found AMH serum levels to be negatively correlated with E2 serum levels, and one study did not find a significant correlation between AMH and E2 serum levels.

The reasons for these discrepancies may be due to the timing of the blood sample used for testing.

Of the two studies that used early follicular phase blood samples for testing, one showed a negative correlation and the other showed no significant correlation. Of the two studies that used blood samples from any stage of the menstrual cycle, or did report the stage at which sampling occurred, both showed a positive correlation between AMH and E2 serum levels.

ACCURACY OF AMH TESTING COMPARED WITH INHIBIN B TESTING

The AUC for DSL ACTIVE® AMH/MIS ELISA and inhibin B were 0.71 and 0.63, respectively. This means the DSL ACTIVE® AMH/MIS ELISA test showed moderate performance in detecting the menstrual status of women compared with a poor performance for inhibin B. Less women with CRA had a false-negative result with AMH (high serum level; 24%) than with inhibin B (high serum level; 46%). The specificity of the AMH and inhibin B tests did not vary greatly, but 36 to 40% would have had a false-positive result indicating CRA (low AMH and/or inhibin B serum levels).

Among the included correlations studies, six found AMH serum levels to be positively correlated with inhibin B serum levels and none failed to show a significant correlation between AMH and inhibin B serum levels.

B4 CLINICAL VALIDITY

B4.1 MEASURES OF CLINICAL VALIDITY

The clinical validity of AMH testing is about how well it predicts clinical outcomes. The results are presented in two parts: one presents the prognostic and predictive value of AMH testing, with ovarian function, response to ovarian stimulation and pregnancy as outcomes, and the other part presents how the test would perform in predicting ovarian function in women undergoing treatments with different levels of gonadotoxicity.

The clinical validity of a test often depends on the prevalence or pre-test probability of the target condition or outcome of interest, i.e. ovarian function. The key measures used are the PPVs and NPVs, which are the probabilities of premature ovarian failure or the continuation or resumption of menses in a tested individual. These measures are heavily dependent on the prevalence of disease in the study population, in this case, the risk of ovarian failure. As the risk of ovarian failure depends on a range of factors, e.g. age, type of treatment, the second part of B4.1.5 (Is it accurate? - Usefulness of AMH at varying risks of ovarian failure) included eligible evidence from the prognostic and analytical validity evidence to determine how the test would perform in women undergoing gonadotoxic treatments of different gonadotoxicity levels, effectively different levels of disease prevalence. This should give an indication of the usefulness of the test when it is used before or after gonadotoxic treatments associated with low risk (20–30%), intermediate risk (40–70%) and high risk (70–80%) of ovarian failure.

B4.1.1 REFERENCE STANDARD

The clinical outcomes serve as the reference standard. In the prognostic section outcomes were ovarian function, measured as the presence or absence of amenorrhea, achieving pregnancy, live birth or both, and the response to controlled hyperstimulation, i.e. oocyte yield. The second section aimed at determining the usefulness of AMH determining ovarian function, with the presence or absence of a menstrual cycle as the reference standard.

B4.1.2 RISK OF BIAS ASSESSMENT

The studies included in the prognostic evidence section were assessed for risk of bias using the Quality in Prognosis Studies (QUIPS) Risk of Bias Assessment Instrument for Prognostic Factor Studies (Hayden et al. 2013). The table showing overall risk of bias and risk of bias for the different domains per study is shown in Appendix F.

Of the seven studies included in the second part second part of B4.1.5 (Is it accurate? - Usefulness of AMH at varying risks of ovarian failure), two studies were included for evidence on analytical validity

(Decanter et al. 2014; Su et al. 2011) and assessed by the QUADAS-2 checklist in Section B3.3. Four of the seven studies were included in the prognostic evidence section and initially assessed by the QUIPS tool (Appendix F) (Anders et al. 2008; Chai et al. 2014; D'Avila et al. 2015; Henry et al. 2014). The study by Jantke et al. (2012) was the only study which did not meet the criteria for inclusion in either the analytical validity or the prognostic section, given that no comparator test was applicable and the AMH levels and outcomes were measured at the same time, and therefore the test was not predictive or prognostic. However, it provided 2 x 2 data which means it was eligible to be included in for clinical validity. The risk of bias of these seven studies was reassessed with the QUADAS-2 tool for clinical validity, as shown in Table 37. The risk of bias assessment of the studies by Su et al. 2011 and Decanter et al. 2014 remained largely the same as in Section B3.3, except for the applicability. A clinical outcome is applicable for the clinical validity section, and no comparators have been included in the clinical validity section.

Table 37 Modified QUADAS-2 risk of bias results for clinical validity studies

Study	Risk of bias						Applicability		
	Patient selection	Index test	Comparator	Reference standard	Flow and timing	Patient selection	Index test	Comparator	Reference standard
Anders et al. (2008)	😊	?	?	😊	😊	😊	😊	NA	😊
Chai et al. (2014)	😊	?	?	😊	😊	😊	😊	NA	😊
D'avila et al. (2015)	😊	?	?	😊	😊	😊	😊	NA	😊
Decanter et al. (2014)	😊	?	?	😊	😊	😊	😊	NA	😊
Henry et al. (2014)	😊	?	?	😊	😊	😊	😊	NA	😊
Jantke et al. (2012)	😞	?	NA	?	😞	😊	😊	NA	😊
Su et al. (2011)	😊	?	😊	?	😊	😊	😊	NA	😊

😊 = low risk; 😞 = high risk; ? = unclear risk; NA = not applicable

B4.1.3 CHARACTERISTICS OF THE EVIDENCE BASE

Profiles of the individual studies included in the evidence base are provided in Appendix C.

Characteristics for studies included for the prognostic and predictive value of AMH testing are summarised in Table 38. The characteristics for studies included for clinical validity, i.e. the usefulness of AMH testing, are shown in

Table 39.

Table 38 Key features of the included evidence on prognostic or predictive value of AMH testing, ordered by study size

Trial/study	N	Level of evidence	Risk of bias	Patient population	Analysis / key outcome(s)
Sonigo et al. (2016)	340	IV	Low	Women breast cancer, haematological malignancies, other diseases	Predicting ovarian response Association between number of in vitro matured oocytes cryopreserved and AFC and serum AMH levels Sensitivity and specificity
Dezellus et al. (2017)	249	IV	Low	Women with breast cancer	Predicting ovarian function Baseline serum hormone values, grouped by menstrual status at follow-up
Stochino-Loi et al. (2017)	180	IV	Moderate	Women with stage 3 and 4 endometriosis	Predicting pregnancy Association of pre-surgery AMH with occurrence of pregnancy and pregnancy outcomes
Hamy et al. (2016)	134	IV	Moderate	Women with breast cancer	Predicting pregnancy Association of baseline AMH and end-of-chemotherapy AMH with occurrence of pregnancy
Su et al. (2010)	127	IV	Low	Female breast cancer survivors	Predicting ovarian function Median hormone values in patients with CRA vs hormone values in patients with no CRA on second follow-up
Ruddy et al. (2014)	124	II	Moderate	Women with breast cancer	Predicting ovarian function Baseline serum hormone values, grouped by menstrual status at follow-up Multivariate analysis, predicting ovarian function
Su et al. (2014)	109	II	Low	Women with breast cancer	Predicting ovarian function Pre-chemotherapy levels of AMH, FSH, inhibin B and E2, and unadjusted HRs of return of ovarian function Multivariate analysis predicting ovarian function
Iwase et al. (2016)	58	IV	Moderate	Women with endometriosis	Predicting pregnancy Serum hormone values at baseline and follow-up, grouped by pregnancy status at follow-up
Anderson et al. (2013)	55	II	Low	Women with early breast cancer	Predicting ovarian function Baseline serum hormone values, grouped by menstrual status at follow-up Multivariate analysis predicting ovarian function ROC curve
D'avila et al.	47	IV	Moderate	Women with breast	Predicting ovarian function

Trial/study	N	Level of evidence	Risk of bias	Patient population	Analysis / key outcome(s)
(2015)				cancer	Baseline serum hormone values, grouped by menstrual status at follow-up ROC curve
Chai et al. (2014)	42	IV	Low	Women with early breast cancer	Predicting ovarian function Sensitivity and specificity ROC curve
Lee et al. (2011)	41	IV	Low	Women with breast cancer	Predicting ovarian response Multivariate regression analysis
Manno et al. (2016)	38	IV	Moderate	Women with breast cancer, HL, NHL, or other types of cancer	Predicting ovarian response Correlation of hormone levels with oocytes retrieved
Ozaki et al. (2016)	35	IV	Moderate	Women with endometriosis	Predicting pregnancy Association of post-surgery AMH with occurrence of spontaneous pregnancy
Lind et al. (2016)	34	IV	Moderate	Women with benign ovarian cysts	Predicting pregnancy Association of post-surgery AMH with occurrence of pregnancy and live birth
Anderson & Cameron (2011)	33	II	Low	Women with early breast cancer	Predicting ovarian function Baseline median serum hormone values, grouped by menstrual status at follow-up Multivariate analysis ROC curve
Henry et al. (2014)	27	II	Low	Women with breast cancer	Predicting ovarian function Multivariate analysis, predicting ovarian function
Takae et al. (2015)	27	IV	Moderate	Women with breast cancer	Predicting ovarian response Correlation of hormone levels with oocytes retrieved
Yu et al. (2010)	26	IV	Moderate	Women with early breast cancer	Predicting ovarian function Baseline serum hormone values, grouped by menstrual status at follow-up
Anders et al. (2008)	22	IV	Moderate	Women with early breast cancer	Predicting ovarian function Baseline median serum hormone values, grouped by menstrual status at follow-up
Pup et al. (2014)	12	IV	Moderate	Women with lymphoma undergoing HSCT	Predicting pregnancy Association of pre-treatment AMH with occurrence of pregnancy

AFC = antral follicle count; AMH = anti-Müllerian hormone; CRA=chemotherapy-related amenorrhea; E2 = estradiol; FSH = follicle-stimulating hormone; HL = Hodgkin's lymphoma; HR = hazard ratio; NHL = non-Hodgkin's lymphoma; HSCT = haematopoietic stem cell transplant; ROC = receiver operator characteristic

I = systematic review of level II studies;

II = a prospective cohort study

III-1 = all or none

III-2 = analysis of prognostic factors among persons in a single arm of a randomised controlled trial

III-3 = a retrospective cohort study

IV = case series, or cohort study of persons at different stages of disease

Because evidence for the incremental value of AMH as an addition to standard testing to predict ovarian function is lacking in most studies, it was deemed inappropriate to assign these studies a level II grading as per the NHMRC prognostic levels of evidence. Instead, these studies were graded as level IV, given their incremental prognostic value could not be determined.

Table 39 Key features of the included evidence on clinical validity of AMH testing, ordered by study size

Trial/study	N	Level of evidence ^a	Risk of bias	Patient population	Analysis / key outcome(s)
Su et al. (2011)	109	III-2	Low	Women with breast cancer	Sensitivity, specificity, LR+, LR-
Jantke et al. (2012)	86	III-2	Medium	Childhood cancer survivors (leukaemia, Hodgkin's, non-Hodgkin's, lymphoma)	Sensitivity, specificity, LR+, LR-
D'avila et al. (2015)	47	III-2	Low	Women with breast cancer	Sensitivity, specificity, LR+, LR-
Chai et al. (2014)	42	III-2	Low	Women with early breast cancer	Sensitivity, specificity, LR+, LR-
Decanter et al. (2014)	30	III-2	Low	Women with early breast cancer or lymphoma	Sensitivity, specificity, LR+, LR-
Henry et al. (2014)	27	III-2	Low	Women with breast cancer	Sensitivity, specificity, LR+, LR-
Anders et al. (2008)	22	III-2	Low	Women with early breast cancer	Sensitivity, specificity, LR+, LR-

LR+ = positive likelihood ratio; LR- = negative likelihood ratio

^a Diagnostic accuracy levels of evidence, as per NHMRC hierarchy of evidence

B4.1.4 OUTCOME MEASURES AND ANALYSIS

The outcomes measured in the included studies, along with the statistical methods used to analyse the results are provided in Appendix C. To assess the accuracy of the proposed test, studies were only included if they provided data that could be extracted into a classic 2 × 2 table (Table 40), in which the results of the index test were cross-classified against the results of the reference standard⁴ using Bayes' Theorem.

Table 40 Analytical validity data extraction

		Reference standard		
		<i>Disease +</i>	<i>Disease -</i>	
Index test	<i>Test +</i>	true positive	false positive	Total test positive
Or comparator	<i>Test -</i>	false negative	true negative	Total test negative
		Total with disease	Total without disease	

The positive likelihood ratio (LR+) is the probability of a person who has the disease, in this case ovarian failure, testing positive divided by the probability of a person who does not have the disease testing positive. The LR+ is calculated as follows:

⁴ Armitage, P, Berry, G & Matthews, JNS 2002, *Statistical methods in medical research*, fourth edn, Blackwell Science, Oxford.

Deeks, JJ 2001, 'Systematic reviews of evaluations of diagnostic and screening tests', in M Egger, G Davey Smith & DG Altman (eds), *Systematic Reviews in Healthcare: Meta-Analysis in Context*, second edn, BMJ Publishing Group, London, pp. 248–282.

$$LR+ = TP \text{ rate} / FP \text{ rate} = a / (a + c) / [b / (b + d)] = \text{sensitivity} / (1 - \text{specificity})$$

The negative likelihood ratio (LR-) is the probability of a person who has the disease, in this case ovarian failure, testing negative divided by the probability of a person who does not have the disease testing negative. The LR- is calculated as follows:

$$LR- = FN \text{ rate} / TN \text{ rate} = c / (a + c) / [d / (b + d)] = (1 - \text{sensitivity}) / \text{specificity}$$

A LR of 1 means that the test does not provide any useful diagnostic information, whereas LR+ >5 and LR- <0.2 can suggest strong diagnostic ability (MSAC 2005).

Negative and positive predictive values (NPV, PPV) were also reported. They are defined as follows:

PPV (proportion of positive results that are true positives) = true positives / true + false positives

NPV (proportion of negative results that are true negatives) = true negatives / true + false negatives

B4.1.5 RESULTS OF THE SYSTEMATIC LITERATURE REVIEW

PROGNOSTIC AND PREDICTIVE VALUE OF AMH TESTING

Summary

Will the extra information generated as a result of the AMH test be of additional prognostic value in female patients prior to receiving gonadotoxic treatment, compared to FSH, E2 and/or AFC ultrasound alone?

AMH testing was considered to be mainly of prognostic and predictive value. Eleven studies were identified on using AMH to predict ovarian function at follow-up in women undergoing gonadotoxic treatment. Mostly women with breast cancer were included. Six of the seven studies reporting mean or median baseline AMH values showed significantly lower baseline AMH levels in women with CRA, with a follow-up of 6 months to 5 years after enrolment or treatment. Higher mean age and lower AFC count were also associated with CRA at follow-up (in 4/4 studies and 2/2 studies, respectively). The other comparator test values were only associated with CRA in about half the studies.

AMH remained a significant predictor for the continuation or resumption of ovarian function in the multivariate analysis in 80% of studies (4/5). The largest study reported that over a median of 163 days, women with an AMH above 0.7 ng/mL have a threefold greater probability of shorter time to ovarian recovery compared to women with an AMH level <0.7 ng/mL. FSH (≤ 10 IU/L) and age (<40 years) were also predictive for shorter time to ovarian recovery, with HRs of 4.7 (95%CI 1.3, 16.8) and 3.39 (95%CI 1.74, 6.60), respectively.

ROC curve analysis of the accuracy of AMH for predicting ovarian function was included in two studies to determine the optimal cut-off point for AMH values. An AUC for AMH of >0.8 was observed in both studies which indicates good test performance. One study showed AMH values ≥ 3.32 ng/mL were protective for the occurrence of oligomenorrhea or amenorrhea after treatment, with a sensitivity of 85% and a specificity of 75%

(n=52). AFC had a sensitivity of 83% and a specificity of 62%, with a cut-off of <13 follicles to predict oligomenorrhea or amenorrhea at the 6-month follow-up. One study including breast cancer patients presented a classification mosaic chart including cut-offs for age and AMH score set at 38.6 years and 3.8 pmol/L, respectively, and showed a sensitivity and specificity of 98.2% and 80.0%, respectively, in predicting amenorrhea after two years (n=75). The lack of standardisation complicates the interpretation of the AMH values and therefore determining standardised clinically relevant thresholds for AMH testing in Australia may improve the prognostic value of the test in clinical practice.

The prognostic value of post-treatment AMH testing in the patient population is hard to determine due to limited evidence (k=2). No evidence was identified on the incremental value of AMH testing. However, the limited evidence available indicated that a woman with detectable AMH post-chemotherapy, measured by pico-AMH ELISA, is likely to have ongoing menses for at least three years, whereas a woman with undetectable AMH, and therefore lack of ovarian reserve, is likely to continue to have amenorrhoea.

A predictive relationship between AMH levels and pregnancy rate was only identified in one out of six studies. Only one small study found a lower pregnancy rate in the group with low AMH levels.

Does the addition of AMH testing to standard tests allow better prediction of response to superovulation than standard tests alone?

One study showed the association between the number of in vitro matured oocytes cryopreserved and AFC and AMH levels in cancer patients by presenting a ROC curve (n=340). The AMH test showed moderate (0.7–0.8) to good (0.8–0.9) test performance at predicting oocyte yield. AFC ultrasound had good (0.8–0.9) test performance.

Three studies investigated the correlation between oocytes retrieved and AMH, AFC, inhibin B and/or E2 values. A moderate linear relationship between AMH and oocyte yield was found. One study also showed moderate relationships between AFC and inhibin B with oocyte yield. FSH and age showed an inverse correlation with oocyte yield.

An association was found between AMH levels and ovarian response to hyperstimulation in women undergoing gonadotoxic treatment. However, it could not be determined whether the AMH test in addition to other standard tests had an incremental predictive value in determining response to ovarian stimulation.

The information generated as a result of the AMH test is mainly of prognostic and predictive value.

Studies on the prognostic value of AMH were included based on the PICO criteria in Table 16 and studies investigating the predictive value of AMH were included based on the PICO criteria in Table 17.

AMH test results were used to predict ovarian function or treatment related amenorrhea, ovarian response to controlled ovarian hyperstimulation, and the chance of spontaneous pregnancy.

Prognostic value of AMH - predicting ovarian function

Univariate analysis in women tested before receiving gonadotoxic treatment

Eleven studies with extractable data on AMH predicting ovarian function in women undergoing gonadotoxic treatment were included. This information is expected to be useful for influencing whether women undergo fertility preservation prior to treatment.

Seven of these studies presented pre-treatment AMH values in breast cancer patients with CRA compared to baseline AMH values in patients without CRA (Table 41). All but one of these studies (6/7, 85%) reported a statistically significant lower baseline AMH in women who had CRA at follow-up, compared to women who continued menstruation. Follow-up was six months to five years after enrolment or the start of treatment, depending on the study. The small study that did not find a significantly lower baseline AMH in the group with CRA reported that only one patient had an AMH level above the lower normal range of 0.05 ng/mL at one-year follow-up. However, in the same time period, 15 patients had resumed menstruation (Yu et al. 2010).

The large study by Dezellus et al. (2017) reported that the prevalence of CRA in breast cancer patients was 93.4% after chemotherapy, and this decreased to 82.2% after six months (n=152 patients with amenorrhea) and was 64.3% after 12 months follow-up (n=119 patients). Mean baseline AMH levels were significantly different between groups at six months' follow-up, however this was no longer the case after 12 months.

Table 41 Baseline AMH levels in breast cancer patients with CRA at follow-up compared with baseline AMH levels in patients without CRA

Study	Population	Follow-up	Baseline AMH measurement	CRA group (AMH in ng/mL)	No CRA group (AMH in ng/mL)	p-value
Anders et al. (2008)	N=21 women with early breast cancer age 18–55 years	1 year	Median (range)	0.16 (0.006–1.5)	1.09 (0.64–3.8)	0.02 ^a
Anderson & Cameron (2011)	N=33 women with early breast cancer	5 years	Mean ± SD	0.7 ± 0.1	2.5 ± 0.4	<0.0001 ^b
Anderson et al. (2013)	N=55 at 1-year, N=46 at 2-year follow-up Pre-menopausal women with early breast cancer	1 year 2 years	Mean ± SD	6.6 ± 1.5 4.0 ± 0.9	16.6 ± 4.8 17.2 ± 5.1	0.01 ^b <0.0001 ^b
D'Avila et al. (2015)	N=47 women with breast cancer	6 months	Median and interquartile range	1.31 (0.72–2.89) ^c	5.34 (2.71–8.15)	p<0.001 ^d
Dezellus et al. (2017)	N=249 women with breast cancer	6 months	Mean ± SD	3.42 ± 2.59	5.33 ± 3.72	0.0027 ^d (NS after longer follow-up)
Ruddy et al. (2014)	N=124 at 1-year, N=100 at 18-month follow-up Pre-menopausal women with breast cancer	1 year 18 months	Median	0.08 (n=102) 0.06 (n=81)	1.3 (n=22) 1.25 (n=19)	<0.0001 ^e <0.0001 ^e

Study	Population	Follow-up	Baseline AMH measurement	CRA group (AMH in ng/mL)	No CRA group (AMH in ng/mL)	p-value
Yu et al. (2010)	N=26 women with breast cancer	1 year	Median	0.97 (n=11)	0.98 (n=15)	NS ^b

AMH = anti-Müllerian hormone; CRA = chemotherapy-related amenorrhea; NS = not significant; SD = standard deviation

^a Wilcoxon two-sample test

^b Student t-test

^c The median value of patients with oligomenorrhea or amenorrhea at follow-up

^d Mann-Whitney U-test

^e Wilcoxon rank sum test

The studies also included one or more comparator tests and/or age, to determine whether baseline test values and/or age could predict CRA at follow-up. The results are for FSH (Table 42), E2 (Table 43), inhibin B (CRA = chemotherapy-related amenorrhea; E2 = estradiol; NS = not significant; SD = standard deviation

Table 44) and AFC (CRA = chemotherapy-related amenorrhea; NS = not significant; SD = standard deviation

Table 45) measurements, and age at baseline (AFC = antral follicle count; CRA = chemotherapy-related amenorrhea; SD = standard deviation

Table 46). Of the four studies measuring baseline FSH in addition to AMH, two studies (total n=54) identified significantly higher baseline FSH levels in women with CRA at follow-up of either one or five years. Two other studies (total n=81) found no relationship between baseline FSH levels and CRA at 1-year follow-up.

Of the two studies measuring E2 levels at baseline, only one study (n=21) found a significantly higher baseline E2 level in women with a menstrual cycle at follow-up.

Of three studies including inhibin B in the analyses, one study found that baseline inhibin B was significantly lower in women with CRA at follow-up (n=21). The other two studies (total n=88) did not find a significant difference.

When looking at the two studies reporting AFC values, both studies showed significantly lower AFC counts in patients with amenorrhea at follow-up. Furthermore, higher age was significantly associated with CRA at follow-up in all studies (4/4).

Table 42 Baseline FSH levels in breast cancer patients with CRA at follow-up compared with baseline FSH levels in patients without CRA

Study	Population	Follow-up	Baseline FSH measurement	CRA (FSH in IU/L)	No CRA (FSH in IU/L)	p-value
Anders et al. (2008)	N=21 women with early breast cancer age 18–55 years	12 months	Median (range)	2.1 (0–25.3)	0 (0–1.1)	0.16
Anderson & Cameron. (2011)	N=33 women with early breast cancer	5 years	Mean ± SD	13.1 ± 1.8	5.2 ± 0.7	0.0008
Anderson et al. (2013)	N=55 at 1-year, N=46 at 2-year follow-up Pre-menopausal women with early breast cancer	1 year 2 years	Mean ± SD	4.9 ± 0.6 5.6 ± 0.8	3.6 ± 0.9 3.3 ± 0.5	NS NS
Yu et al.	N=26 women with breast	1 year	Median	4.5 (n=11)	6.8 (n=15)	NS

Study	Population	Follow-up	Baseline FSH measurement	CRA (FSH in IU/L)	No CRA (FSH in IU/L)	p-value
(2010)	cancer					

CRA = chemotherapy-related amenorrhea; FSH = follicle-stimulating hormone; NS = not significant; SD = standard deviation

Table 43 Baseline E2 levels in breast cancer patients with CRA at follow-up compared with baseline E2 levels in patients without CRA

Study	Population	Follow-up	Baseline E2 measurement	CRA	No CRA	p-value
Anders et al. (2008)	N=21 women with early breast cancer age 18–55 years)	12 months	Median (range)	83.3 (31.3–383.4) pg/mL	112.0 (41.8–160.4) pg/mL	0.96
Anderson & Cameron. (2011)	N=33 Women with early breast cancer	5 years	Mean ± SD	265 ± 19 pmol/L	292 ± 31 pmol/L	NS

CRA = chemotherapy-related amenorrhea; E2 = estradiol; NS = not significant; SD = standard deviation

Table 44 Baseline inhibin B levels in breast cancer patients with CRA at follow-up compared with baseline inhibin B levels in patients without CRA

Study	Population	Follow-up	Baseline inhibin B measurement	CRA (inhibin B in pg/mL)	No CRA (inhibin B in pg/mL)	p-value
Anders et al. (2008)	N=21 women with early breast cancer (18-55 years)	12 months	Median (range)	33.2 (0–187.6)	108.8 (28.6–425.2)	0.03
Anderson & Cameron. (2011)	N=33 women with early breast cancer	5 years	Mean ± SD	51.8 ± 7.9	68.9 ± 11.6	NS
Anderson et al. (2013)	N=55 at 1-year, N=46 at 2-year follow-up Pre-menopausal women with early breast cancer	1 year 2 years	Mean ± SD	37.6 ± 5.8 34.2 ± 6.2	32.4 ± 12.0 38.1 ± 14.6	NS NS

CRA = chemotherapy-related amenorrhea; NS = not significant; SD = standard deviation

Table 45 Baseline AFC count in breast cancer patients with CRA at follow-up compared with baseline AFC count in patients without CRA

Study	Population	Follow-up	AFC measurement (Baseline count)	CRA	No CRA	p-value
Anderson & Cameron. (2011)	N=33 women with early breast cancer	5 years	Mean ± SD	8.7 ± 1.2	19.8 ± 3.0	0.0004
D'Avila et al. (2015)	N=47 women with breast cancer	6 months	Median and interquartile range	9 (7.75–12) ^b	13.5 (11–16)	p<0.001

AFC = antral follicle count; CRA = chemotherapy-related amenorrhea; SD = standard deviation

Table 46 Age at baseline in breast cancer patients with CRA at follow-up compared with age at baseline in patients without CRA at follow-up

Study	Population	Follow-up	Age (years) at baseline	CRA	No CRA	p-value
-------	------------	-----------	-------------------------	-----	--------	---------

Anderson et al. (2013)	N=55 at 1 year, N=46 at 2 year follow-up Pre-menopausal women with early breast cancer	1 year 2 years	Mean ± SD	43.3 ± 0.7 43.9 ± 0.8	37.9 ± 0.8 37.9 ± 2.0	0.03 0.004
D'avila et al. (2015)	N=47 women with breast cancer	6 months	Median and interquartile range	36.5 ± 3.8	32.9 ± 3.5	0.02
Dezellus et al. (2017)	N=249 women with breast cancer	6 months	Mean ± SD	35.3 ± 3.7	33.6 ± 3.7	0.0137
Ruddy et al. (2014)	N=124 at 1 year, N=100 at 18 month follow-up Pre-menopausal women with breast cancer	1 year 18 months	Median	46 (n=102) 46 (n=81)	36.5 (n=22) 37 (n=19)	<0.0001 <0.0001

CRA = chemotherapy-related amenorrhea; SD = standard deviation

Multivariate analysis in women tested before undergoing gonadotoxic treatment

Univariate analysis alone is insufficient to determine the incremental prognostic value of AMH testing, as the different variables are not independent. AMH values decrease with age and correlate with other hormone tests (e.g. FSH, E2; see Section B3). Five studies conducted a multivariate regression analysis to adjust for other variables, such as age, treatment, and/or other hormone tests (Anderson & Cameron 2011; Anderson et al. 2013; Henry et al. 2014; Ruddy et al. 2014; Su, HC et al. 2014). The results are shown in

Table 47.

Table 47 Multivariate regression analysis for prediction of ovarian function in breast cancer patients

Study	Population	Follow-up	Variables included / adjusted for	Outcomes	p-value
Anderson & Cameron. (2011)*	N=33 women with early breast cancer	4–5 years	Pre-chemotherapy age, AMH, and FSH	Only AMH was a significant predictor for ongoing menses AMH OR=13.0 (95%CI 2.5, 66.7)	0.02
Anderson et al. (2013)*	N=46 pre-menopausal women with early breast cancer	2 years	Age, pre-treatment AMH, FSH and inhibin B	Only AMH was a significant predictor for amenorrhea AMH OR=0.013 (95%CI 0.001, 0.227)	0.005
Henry et al. (2014)*	N=29 pre- and peri-menopausal women with newly diagnosed breast cancer	18 months (average 13.6 months)	Age at enrolment, baseline AMH	No factors remained statistically significant in the multivariate analysis	NS
Ruddy et al. (2014)*	N=124 at 1-year, N=100 at 18-month follow-up Pre-menopausal women with breast cancer	1 year 18 months	Baseline AMH, age, race, whether or not the patient received bevacizumab, and use of tamoxifen	Age was predictor for 12 month CRA AMH OR=0.83 (95%CI 0.58, 1.20) Age OR=1.20^a (95%CI 1.10, 1.33) AMH and age were predictors for 18 month CRA AMH OR=0.41^b (95%CI 0.18, 0.95) Age OR=1.18^c (95%CI 1.04, 1.34)	<u>12 months follow-up</u> AMH: 0.32 Age: 0.00003 <u>18 months follow-up</u> AMH: 0.04 Age: 0.008
Su et al. (2014)**	N=109 women with breast cancer treated with chemotherapy and at least 3 months of amenorrhea	Median 163 days (range 4–1009 days)	Age, body size, race, chemotherapy regimen, tamoxifen exposure	Predictive for shorter time to return of ovarian function: Baseline AMH levels >0.7 ng/mL: HR=2.9 (95%CI 1.5, 5.6) FSH levels ≤10 IU/L: HR=4.7 (95%CI 1.3, 16.8) Age <40 years: HR=3.39 (95%CI 1.74, 6.60)	AMH: 0.002 FSH: 0.018 Age: <0.001

AMH = anti-Müllerian hormone; CRA = chemotherapy-related amenorrhea; FSH = follicle-stimulating hormone; HR = hazard ratio; NS = not significant; OR = odds ratio

^a For every 1 year increase in age, there was a 20% increase in the odds of developing CRA at 12 months

^b For every 1 ng/mL increase in AMH, there was a 59% decrease in the odds of developing CRA

^c For every 1 year increase in age, there was a 18% increase in the odds of developing CRA at 18 months

*Multivariate logistic regression analysis

**Multivariable Cox regression analysis

In the multivariate analyses, AMH remained a significant predictor for continued or return of ovarian function in four out of five studies. Only the study with the smallest study population did not find significant factors for predicting ovarian function (Henry et al. 2014). In the two largest studies, higher age was also a significant predictor for CRA and longer time to return of ovarian function (Ruddy et al. 2014; Su, HC et al. 2014). Ruddy et al. (2014) reported that for every one-year increase in age, there was a 20% increase in the odds of developing CRA at the 12-month follow-up, and an 18% increase in the odds of having CRA at the 18-month follow-up. AMH was only predictive at 18 months' follow-up in this study, with a 59% decrease in the odds of having CRA at 18 months of follow-up for every 1 ng/mL increase in AMH level. These data suggest there is likely to be an incremental prognostic effect, if it remains significant after controlling for age and FSH.

Su et al. (2014) reported a hazard ratio (HR) of 2.9 (95%CI 1.5, 5.6) when baseline levels of AMH were higher than 0.7 ng/mL for shorter time to return of ovarian function. This means that women with an AMH above 0.7 ng/mL have three times greater probability of a shorter time to ovarian recovery compared to women with an AMH under 0.7 ng/mL, over a median of 163 days. FSH (≤ 10 IU/L) and age (< 40 years) were also predictive for shorter time to ovarian recovery, with HRs of 4.7 (95%CI 1.3, 16.8) and 3.39 (95%CI 1.74, 6.60), respectively.

Receiver operating characteristic curve analyses

Three studies conducted a ROC curve analysis⁵ to assess the accuracy of AMH for predicting ovarian function and to determine the optimal cut-off point for AMH values (Anderson & Cameron 2011; Anderson et al. 2013; D'Avila et al. 2015). As the data used in the ROC curve from the studies by Anderson et al. (2013) and Anderson and Cameron (2011) were combined and fully captured in the study by Anderson et al. (2013), the results from the 2011 study are not considered in this report.

Using a combined dataset, Anderson et al. (2013) included 75 women with breast cancer. The AUC for predicting ongoing menses at 2-year follow-up was 0.90 (95%CI 0.82, 0.97) and 0.88 (95%CI 0.78, 0.97) for AMH and age, respectively. Based on this, a classification mosaic chart was derived, with age and AMH as predictor variables. An AMH value below 3.8 pmol/L predicted amenorrhea, whereas an AMH value above 20.3 pmol/L predicted ongoing menses. When baseline AMH values fall between these threshold values, an age threshold of 38.6 years was defined above which amenorrhea was predicted. This model had a sensitivity of 98.2% and a specificity of 80.0% in predicting amenorrhea after two years (n=75).

The 2015 study by D'Avila et al. provided a ROC curve for predicting amenorrhea or oligomenorrhea six months after chemotherapy treatment for breast cancer. AMH was slightly better at predicting these outcomes than AFC (AUC for AMH = 0.86; AUC for AFC = 0.81). The study showed AMH values

⁵ Section B3.5 provides information on ROC curve analysis.

≥3.32 ng/mL were protective against the occurrence of oligomenorrhea or amenorrhea after treatment, with a sensitivity of 85% and a specificity of 75% (n=52). The cut-off for AFC was <13 follicles to predict oligomenorrhea or amenorrhea, having a higher risk at 6-month follow-up with a sensitivity of 83% and a specificity of 62%.

Thresholds

Different thresholds for baseline AMH were used to predict ovarian function. Anders et al. (2008) reported an AMH threshold of 0.7 ng/mL, with 50% (3/6) and 92% (11/12) of women who had a value above and below the threshold affected by CRA after one year of follow-up, respectively. The relative risk for this outcome was 0.83. Su et al. (2014) used the same threshold (Multivariate analysis in women tested before undergoing gonadotoxic treatment), in addition to a threshold of 0.17 ng/mL, whereas one other study also used a low threshold of 0.16 ng/mL (Henry et al. 2014). One study (D'Avila et al. 2015) used a much higher AMH threshold (3.32 ng/mL) for predicting CRA.

AMH tests in women after gonadotoxic treatment - prediction of ovarian function

Two studies that met the inclusion criteria included women who had the AMH test after gonadotoxic treatment (Chai et al. 2014; Su 2010).

Su et al. (2010) included 127 female post-chemotherapy breast cancer survivors, of which 111 women provided data on menstrual status at follow-up (2 were deceased, 1 declined and 13 were not able to be reached). AMH was measured one to four years after treatment and follow-up was at two to seven years after treatment.

Return of menses occurred in nine women who had CRA when AMH was measured and had been measured at follow-up. It was reported that AMH levels did not differ between women with CRA reversal compared to women with continued amenorrhea (p=0.92). Similarly, no differences were observed between groups for inhibin B levels (p=0.27) and FSH levels (p=0.73). However, women with CRA reversal were younger on average (mean age 41.7 years, range 38.4–44.8 years vs 47.3 years, range 40.3–56 years]; p<0.001) and received more dose-dense therapy (risk ratio 6.4; p=0.03) than women with continued amenorrhea.

Four women experienced CRA between AMH measurement and the follow-up visit. Compared to women without CRA, the average AMH value in these four women was significantly lower (25.2 pg/mL, range <25–233.5 pg/mL vs 179.4 pg/mL, range 96.2–334.1 pg/mL; p=0.03). FSH was also higher in women with CRA (48.1 IU/L, range 13.3–173.7 IU/L vs 17.4 IU/L, range 12.2–24.7 IU/L; p=0.04). It was reported that women in which CRA occurred had a similar age to those with continuation of menstruation.

The study by Chai et al. (2014) looked at the ability of the pico-AMH ELISA to detect AMH in the 1-year post-chemotherapy serum of 42 women diagnosed with early stage breast cancer who had or had not resumed menstruation three years after chemotherapy. In this study, detection of AMH by

the pico-AMH ELISA predicted future menses in the validation cohort of 42 women with a sensitivity of 91% (95%CI 59%, 100%) and 9% false-negative rate, but was less accurate in predicting amenorrhoea, with a specificity of 74% (95%CI 55%, 88%) and 26% FPR).

Chai et al. (2014) measured the AMH, FSH, E2 and inhibin B serum levels two years post-chemotherapy in 39 women diagnosed with early breast cancer. These women were subdivided into three groups according to their menstrual pattern over the follow-up period (2–5 years): ongoing menses (n=10), transient amenorrhoea (n=4), and amenorrhoea (n=25). The authors then used ROC curve analysis to determine the ability of the various hormones to discriminate between the 35 women with and without menses (Table 48). When the four women with transient amenorrhoea were included the ROC analysis, the AUC varied little for AMH and FSH, but reduced considerably for E2 and inhibin B. This was due to the marked variability of the levels of these hormones in these four women over the 3-year testing period. The AUC for AMH was above 0.9, indicating very good test performance to predict CRA in women post-chemotherapy. For FSH and E2, the AUC was above 0.8, indicating good predictive test performance. For inhibin B the AUC was below 0.7, indicating poor test performance. As the AMH test performs better at predicting ovarian function than the other tests, it is likely that it provides incremental prognostic information.

Table 48 AUC values from ROC analysis conducted by Chai et al. (2014)

AUC [95%CI] and/or p-value for:	35 women with ongoing menses and amenorrhoea	All 39 women
AMH	0.99 [0.97, 1.01]; <0.001	0.97; <0.05
FSH	0.86 [0.73, 0.98]; 0.001	0.87; <0.05
E2	0.93 [CI 0.83, 1.03]; <0.001	0.86; <0.05
Inhibin B	0.74 [0.54, 0.94]; 0.03	0.65; >0.05 (NS)

AMH = anti-Müllerian hormone; AUC = area under the curve; CI = confidence interval; E2 = estradiol; FSH = follicle-stimulating hormone; NS = not significant; ROC = receiver operator characteristic

Chai et al. (2014) determined that the optimum threshold according to the ROC analysis for the pico-AMH ELISA was 16.1 pg/mL, giving a sensitivity of 96% and specificity of 90% in predicting CRA. The authors concluded that this value was close enough to the limit of detection of the assay. Thus, the AMH results can be interpreted to indicate that a woman with detectable AMH post-treatment is very likely to have ongoing menses for at least three years whereas a woman with undetectable AMH will continue to have amenorrhoea due the lack of ovarian reserve.

One study was identified on AMH testing in childhood cancer survivors to assess ovarian reserve after gonadotoxic treatment (Lunsford et al. 2014). The study was not an official include as the results were not prognostic or predictive, but it did report that all patients with delayed pubertal development (n=10) had an AMH <1ng/mL, and the majority of these patients (90%) had an undetectable AMH value. It was reported that AMH levels were statistically significantly lower in the patients with delayed puberty compared with those with normal pubertal progression (p<0.001).

Prognostic value of AMH - predicting spontaneous pregnancy

Six studies reporting on whether AMH was able to predict pregnancy were identified. Five studies reported AMH thresholds and showed how many women achieved pregnancy when divided into groups based on AMH score (see Table 49).

All women underwent gonadotoxic treatment for endometriosis, breast cancer or lymphoma and wanted to become pregnant. Follow-up was at least one year in all studies as shown in Table 49. Only a small study by Ozaki et al. (n=35) reported a higher pregnancy rate in patients with a high AMH, compared to AMH <1.1 ng/mL. This study included women with endometriosis who underwent laparoscopic cystectomy. The other four studies did not report a statistically significant pregnancy rate difference in women with higher AMH values, compared to women with lower AMH values. This suggests that although AMH may assist in predicting which women will have amenorrhea or menses, this relationship is not strong enough to predict ability to conceive.

The recently published and largest study by Stochino-Loi et al. (n=180) included women with endometriosis undergoing ovarian endometrioma ablation. In women with a low baseline AMH score (<2 ng/mL) 34 pregnancies were reported in 46 women (73.9%). In women with a high baseline AMH score (≥ 2 ng/mL) 100 pregnancies were reported in 134 women (74.6%). The probability of pregnancy within 12 months was 50% and 65% in the women with a low and high AMH value, respectively (p=0.19). The estimation of the probability of the rate of live births (HR) using a Cox multivariate model in women with AMH level <2 ng/mL was 0.98 (95%CI 0.6, 1.5) at 36 months after surgery, compared to women with AMH level ≥ 2 ng/mL. This means the impact of preoperative AMH level on the probability of live birth was not statistically significant after adjusting for women's age, antecedents of ovarian cystectomy, ablation of ovarian endometriomas, documented preoperative infertility, and colorectal surgery for endometriosis (Stochino-Loi et al. 2017).

The other larger study by Hamy et al. (n=134) included women with breast cancer. Of the women who became pregnant during the follow-up of this study, 65% had undetectable AMH levels (11/17), compared to 70% among women who did not achieve pregnancy. This was confirmed by a study in lymphoma cancer survivors. This study was not officially included in the systematic review due to the data not being prognostic, i.e. AMH was measured only at the end of the follow-up period. However it stated that nearly a quarter (4/15) of women with critically low AMH values at the end of the study (≤ 0.3 ng/mL) had given birth during the follow-up period, and it reported that another three women with an AMH value ≤ 0.3 ng/mL became pregnant shortly after the follow-up survey and AMH measurement (Hamre et al. 2012). The authors stated that the relation between low AMH levels and reproductive outcomes may differ in cancer survivors compared to older women from the general population (Hamre et al. 2012).

Table 49 Predictive value of AMH predicting pregnancy outcomes in women undergoing gonadotoxic treatment, ordered by study size

Study	Population	Follow-up	Threshold AMH	Outcomes low AMH	Outcomes normal AMH	Outcomes high AMH	p-value
Stochino-Loi et al. (2017)	N=180 endometriosis patients undergoing ovarian endometrioma ablation Mean age: 30.5 ± 4.1 years	AMH was measured preoperatively Follow-up: >1 year Questionnaires completed at 1, 3, and 5 years after surgery	Low: <2 ng/mL (n=46) High: ≥2 ng/mL (n=134)	Pregnancy: 34/46 (73.9%) Through ART: 14/34 (41.2%) Spontaneous: 20/34 (58.8%)		Pregnancy: 100/134 (74.6%) Through ART: 46/100 (46%) Spontaneous: 54/100 (54%)	0.53
				Probability of pregnancy 12 month follow-up: 50% [95%CI 34, 69%] 24 month follow-up: 77% [95%CI 61, 90%] 36 month follow-up: 83% [95%CI 68, 94%]		Probability of pregnancy 12 month follow-up: 65% [95%CI 55, 75%] 24 month follow-up: 77% [95%CI 68, 86%] 36 month follow-up: 83% [95%CI 75, 90%]	0.39
			Low: <1 ng/mL (n=22) Normal: 1-1.9 ng/mL (n=24) High: ≥2 ng/mL (n=134)	Pregnancy outcomes Ongoing (>12wk): 4/34 (11.8%) Delivery: 25/34 (73.5%) Miscarriage: 4/34 (11.8%) Ectopic pregnancy: 1/34 (2.9%)		Pregnancy outcomes Ongoing (>12wk): 21/100 (21%) Delivery: 72/100 (72%) Miscarriage: 7/100 (7%) Ectopic pregnancy: 0/100 (0%)	0.19
				Pregnancy: 15/22 (68.2%) Through ART: 5/15 (33.3%) Spontaneous: 10/15 (66.7%)		Pregnancy outcomes Ongoing (>12wk): 0/15 (0%) Delivery: 13/15 (86.7%) Miscarriage: 1/15 (6.7%) Ectopic pregnancy: 0/15 (0)	Pregnancy outcomes Ongoing (>12wk): 21/100 (21%) Delivery: 72/100 (72%) Miscarriage: 7/100 (7%) Ectopic pregnancy: 0/100 (0%)
				Pregnancy: 19/24 (79.2%) Through ART: 9/19 (47.4%) Spontaneous: 10/19 (52.6%)	Pregnancy: 100/134 (74.6%) Through ART: 46/100 (46%) Spontaneous: 54/100 (54%)	0.69	
				Pregnancy outcomes Ongoing (>12wk): 4/19 (21%) Delivery: 13/19 (68.4%) Miscarriage: 2/19 (10.5%) Ectopic pregnancy: 1/19 (4.2%)	Pregnancy outcomes Ongoing (>12wk): 21/100 (21%) Delivery: 72/100 (72%) Miscarriage: 7/100 (7%) Ectopic pregnancy: 0/100 (0%)	0.63	
						0.07	

Study	Population	Follow-up	Threshold AMH	Outcomes low AMH	Outcomes normal AMH	Outcomes high AMH	p-value
Hamy et al. (2016) Moderate risk of bias	N=134 women with breast cancer Median age (range): 35.5 (26–43)	AMH was measured on the first day of chemotherapy Follow-up: Median 59 months (range 11–104)	Very low: <0.75 ng/mL Low: 0.75–1.5 ng/mL High: >1.5–2.5 ng/mL Very high: >2.5 ng/mL	<u>Pregnancy:</u> Very low AMH (reference class) HR=1	<u>Pregnancy:</u> Low: HR =0.34 [95%CI 0.04, 3.29] High: HR=1.55 [95%CI 0.35, 6.95]	<u>Pregnancy:</u> Very high AMH: HR=2.32 [95%CI 0.60, 8.9]	NS
Ozaki et al. (2016) Moderate risk of bias	N=35 endometriosis patients aiming for pregnancy after laparoscopic cystectomy Mean age: 33.5 ± 4.9	AMH was measured 6 months post-surgery Follow-up after AMH: 18 months	Low: <1.1 ng/mL Normal/high: >1.1 ng/mL	<u>Spontaneous pregnancy:</u> 14.3%	-	<u>Spontaneous pregnancy:</u> 59.2%	0.04
Lind et al. (2016) Moderate risk of bias	N=34 women with benign ovarian cysts undergoing surgery and AMH tests, attempting to conceive Mean age at surgery: 30.4 ± 5.9	AMH was measured 6 months post-surgery Follow-up after AMH: 18 months	Low: <1 ng/mL Normal: 1–3.5 ng/mL High: >3.5 ng/mL	<u>Pregnancy:</u> 2/4 (50%) <u>Live birth:</u> 0/4 (0%)	<u>Pregnancy:</u> 11/23 (48%) <u>Live birth:</u> 8/23 (35%)	<u>Pregnancy:</u> 4/7 (57%) <u>Live birth:</u> 3/7 (43%)	0.911 ^a 0.312 ^a
Pup et al. (2014) Moderate risk of bias	N=12 lymphoma patients undergoing HSCT Mean age at diagnosis: 26 (range 18–37)	AMH was measured pre-treatment Follow-up: NR (years)	Low: ≤0.16 ng/mL Normal/high: >0.16 ng/mL	<u>Pregnancy:</u> 1/5 (20%)	-	<u>Pregnancy:</u> 2/7 (28.6%)	1.00

AMH = anti-Müllerian hormone; ART = assisted reproductive technology; CI = confidence interval; HR = hazard ratio; HSCT = haematopoietic stem cell transplant; NR = not reported; NS = not significant

^a Pearson Chi-squared test

Iwase et al. also reported results on pregnancy in endometriosis patients who underwent laparoscopic cystectomy and had AMH results. They stated that the AMH level measured one year post-operatively was significantly higher in women who achieved pregnancy (3.44 ± 1.78 ng/mL; n=17) compared to those who did not get pregnant (2.17 ± 2.24 ; n=24) ($p=0.049$). However, no difference in mean AMH values were observed when it was measured directly after surgery ($p=0.122$) or one month post-operatively ($p=0.682$).

Taniguchi et al. did not report pregnancy nor birth rates and therefore did not meet inclusion criteria for this report. However, they stated that in women with endometriosis the post-operative decline in AMH levels after one year of follow-up was significantly lower in patients achieving spontaneous pregnancy, compared to women requiring infertility treatment. Median (25th, 75th percentile) AMH values were 0.34 ng/mL (0.24, 0.40) versus 0.48 ng/mL (0.36, 0.60), respectively ($p<0.05$; n=40) (Taniguchi et al. 2016).

Although there is a trend showing that those with a higher AMH level are more likely to have a pregnancy, the difference between those with low and high AMH is too small to inform clinical decisions. No mention was made regarding the use of fertility preservation, so it can only be assumed that ART was undertaken with gametes collected post-treatment (e.g. in the study by Stochino-Loi et al.).

Predictive value of AMH - predicting response to controlled hyperstimulation

The large study by Sonigo et al. (2016) evaluated the association between the number of in vitro matured oocytes cryopreserved and AFC and AMH levels in 340 cancer patients. The majority of the study population consisted of breast cancer patients (n=300). Also included were 14 women with haematological malignancies and 26 women with other forms of cancer. In the study, 301 women opted for oocyte cryopreservation, 39 chose embryo freezing and 47 women decided to undergo ovarian tissue cryopreservation in combination with in vitro maturation. Odds ratios and the AUCs for number of oocytes frozen are shown in Table 50. The AUCs reported would mean that AMH would have moderate (0.7–0.8) to good (0.8–0.9) ability to predict oocyte yields above or below a threshold, and AFC would have good (0.8–0.9) accuracy to predict oocyte yield.

Sonigo et al. also reported the sensitivity and specificity of AMH and AFC at different thresholds for predicting the number of mature oocytes frozen (Table 51).

Table 50 Association between number of in vitro matured oocytes cryopreserved and AFC and serum AMH levels, univariate linear regression

Oocyte yield	AMH (ng/mL)	AFC (n)
	OR [95%CI]	
≤2 oocytes	0.63 [0.53, 0.76] ^a	0.85 [0.81, 0.89] ^a
≥8 oocytes	1.33 [1.21, 1.46] ^a	1.09 [1.06, 1.12] ^a
≥10 oocytes	1.30 [1.19, 1.42] ^a	1.09 [1.06, 1.12] ^a

Oocyte yield	AMH (ng/mL)	AFC (n)
≥15 oocytes	1.17 [1.05, 1.31] ^b	1.07 [1.03, 1.1] ^a
	AUC [95%CI]	
≤2 oocytes	0.79 [0.73, 0.86]	0.81 [0.76, 0.86]
≥8 oocytes	0.80 [0.74, 0.86]	0.80 [0.74, 0.84]
≥10 oocytes	0.81 [0.75, 0.88]	0.82 [0.76, 0.88]
≥15 oocytes	0.79 [0.65, 0.92]	0.90 [0.84, 0.97]

AFC = antral follicle count, AMH = anti-Müllerian hormone, CI = confidence interval, AUC = area under the curve, OR = odds ratio

^a p<0.0001

^b p=0.006

Source: (Sonigo et al. 2016)

Table 51 Threshold values of AFC and serum AMH levels for obtaining ≤2 or ≥8, 10 or 15 mature oocytes frozen

	Threshold (≥)	Threshold of oocytes frozen	Sensitivity [95%CI]	Specificity [95%CI]
AMH (ng/mL)	3.9	≥15	0.89 [0.52, 1.00]	0.61 [0.55, 0.67]
	3.7	≥10	0.84 [0.68, 0.94]	0.64 [0.57, 0.7]
	3.5	≥8	0.82 [0.71, 0.91]	0.63 [0.56, 0.70]
	3.0	≤2	0.77 [0.65, 0.86]	0.80 [0.63, 0.77]
AFC (n)	28	≥15	0.90 [0.55, 1.00]	0.78 [0.73, 0.82]
	20	≥10	0.88 [0.76, 0.96]	0.62 [0.56, 0.68]
	19	≥8	0.82 [0.72, 0.89]	0.63 [0.56, 0.69]
	19	≤2	0.83 [0.74, 0.90]	0.61 [0.55, 0.68]

AFC = antral follicle count, AMH = anti-Müllerian hormone, CI = confidence interval

Source: (Sonigo et al. 2016)

Three smaller studies were included which investigated the correlation between AMH score and number of oocytes retrieved during controlled hyperstimulation in women undergoing gonadotoxic treatment (Table 52). Lee et al. (2011) performed a multiple regression analysis to adjust for AMH, age, FSH, inhibin B and AFC which were correlated with the number of oocytes retrieved in the univariate analysis. The correlation between number of oocytes and AMH, AFC and age was $r=0.71$ (all $p<0.05$). The correlation between mature, metaphase II oocytes retrieved and AMH and AFC was $r=0.64$ (all $p<0.05$). An AMH cut-off of 1.2 ng/mL was used to determine if yield was higher with a higher AMH score. In the group with low AMH (≤ 1.2 ng/mL; $n=18$) the mean oocyte yield was 11.3 ± 9.7 , whereas in the group with high AMH (>1.2 ng/mL; $n=23$) the mean oocyte yield was 19.7 ± 8.8 ($p<0.01$). This was 6.8 ± 5.6 and 12.7 ± 5.8 for mature, metaphase II oocytes, respectively ($p<0.01$). With >1.2 ng/mL used as an AMH threshold for high AMH score and >4 mature, metaphase II oocytes retrieved as a successful oocyte yield, the sensitivity and specificity of AMH for predicting successful oocyte yield would be 68.9% (95%CI 50.0, 83.9) and 87.5% (95%CI 47.4, 99.7), respectively (Lee et al. 2011).

The two other studies did not report whether the correlations were statistically significant; however the correlations between AMH value and oocytes retrieved are between 0.45 and 0.60 in all three studies, which indicates a moderate linear relationship. Takae et al. (2015) also reported an inverse moderate correlation with age and the number of oocytes retrieved ($r = -0.48$).

Table 52 Correlation between oocytes retrieved and AMH, AFC, FSH, inhibin B and/or E2 values

Study	Population	Correlation between:	AMH	AFC	FSH	Inhibin B	Peak E2 (pg/mL)
Lee et al. (2011)	N=41 women with stage ≤3 breast cancer	No. of oocytes retrieved: No. of MII retrieved: No. of embryos cryopreserved:	(logAMH) 0.454** 0.583** 0.508**	0.658** 0.605** 0.588**	-0.375* -0.360* NS	0.385* 0.434* 0.595**	-
Manno et al. (2016)	N=38 patients with breast cancer, HL, NHL, or other types of cancer	(Pearson r) No. of oocytes retrieved: No. of oocytes vitrified:	0.46 0.39	0.16 0.26	- -	- -	0.35 0.40
Takae et al. (2015)	N=27 breast cancer patients	(Spearman correlation coefficient by log rank test) No. of oocytes retrieved	0.60	-	-	-	-

AFC = antral follicle count; AMH = anti-Müllerian hormone; E2 = estradiol; FSH = follicle-stimulating hormone; MII = mature, metaphase II oocyte

*p<0.005 (2-tailed)

**p<0.001 (2-tailed)

IS IT ACCURATE? - USEFULNESS OF AMH AT VARYING RISKS OF OVARIAN FAILURE (CLINICAL VALIDITY)

Summary

The accuracy of AMH testing prior to gonadotoxic treatment in predicting the resumption of menses after treatment was reported in three studies. First generation AMH tests showed limited utility for AMH measured both prior to and after treatment for predicting the resumption of menses. A positive test result, i.e. above a certain threshold, only offers useful information for women undergoing intermediate to high-risk treatment. At least 76% of women with a ≥50% risk of ovarian failure received a positive pre-treatment AMH result and 79 to 94% of women undergoing high-risk treatment with a positive post-treatment AMH test result resumed menses. On the other hand, a negative AMH test result only offered useful information to women having treatment with low risk of amenorrhea. Here, the NPV indicated that 88 to 95% and 71 to 88% of women with a negative pre- and post-treatment AMH test result would not have resumption of menses, respectively. Among women having treatment with a high risk of ovarian failure, 44 to 68% and 21 to 45% with a negative pre- or post-treatment test result would have amenorrhea at follow-up, respectively. This is equivalent to chance and not clinically useful.

Second generation AMH tests performed much better. In women with a low risk of ovarian failure, up to three quarters of women with a detectable baseline AMH Gen II ELISA test result would have resumption of menses, and around 95% of women with a detectable AMH level undergoing a treatment that poses high risk of ovarian failure would have resumption of menses after treatment. The NPV showed that 81 to 99% of women with an undetectable AMH level and a low risk of ovarian failure would continue to have amenorrhea. The pico-AMH ELISA showed that around 90% of women undergoing high-risk treatment who had a positive post-treatment AMH test result would have resumption of menses. Conversely, only 47 to 64% of women with a positive pico-AMH ELISA result and undergoing low-risk treatment would resume menses. In the low-risk treatment group, only a negative result was useful, with 87 to 97% of women having amenorrhea with a negative AMH result measured after treatment. In conclusion, the positive AMH test result was only clinically useful at higher prevalence rates of ovarian failure, i.e. in scenarios of high-risk treatment, whereas the negative test result was only useful in the group at low risk of ovarian failure, i.e. in scenarios of low-risk treatment.

The evidence that was included in the analytical validity and the prognostic section and provided sensitivity/specificity or 2 × 2 data was included here to determine how the test would perform in women undergoing gonadotoxic treatment of different levels of gonadotoxicity. This gives an indication of the usefulness of the test when it is used before or after treatments posing low risk (20–30%), intermediate risk (40–70%) and high risk (70–80%) of ovarian failure.

Clinical validity of AMH testing measured prior to gonadotoxic treatment compared with a clinical reference standard to predict resumption of menses after treatment

Three studies reported on the accuracy of AMH testing prior to gonadotoxic treatment in predicting the resumption of menses after treatment and provided sensitivity/specificity or 2 × 2 data. The study by Anders et al. (2008) compared the sensitivity and specificity of the DSL ACTIVE® MIS/AMH ELISA measured prior to gonadotoxic treatment for early stage breast cancer in 27 women to detect the resumption of menses within one year of treatment. The pre-chemotherapy median AMH levels among women aged less than 35 years was significantly higher compared with those aged 35 years or older (2.72 versus 0.47 ng/mL, $p < 0.0001$). The authors reported that women with pre-treatment levels below the median AMH value were more likely to experience CRA than those who were above the median.

The study by D'Avila et al. (2015) considered the sensitivity and specificity of the EIA AMH/MIS ELISA in measuring AMH levels prior to gonadotoxic treatment for predicting post-treatment resumption of menstruation in 52 women diagnosed with breast cancer. The authors reported that women with AMH values < 3.32 ng/mL prior to gonadotoxic treatment were more likely to develop oligomenorrhea or amenorrhea after completion of their treatment (Table 53). The study by Henry et al. (2014) investigated the sensitivity and specificity of the AMH Gen II ELISA to detect AMH in 27 women with stage I to stage III breast cancer, prior to receiving neoadjuvant or adjuvant chemotherapy, compared with a reference standard of resumed menstruation (Table 53).

As the Australian prevalence rate of women resuming menstruation after gonadotoxic therapy cannot be estimated accurately, and depends on a range of factors including type of treatment (Section A, Table 7), the PPV and NPV were calculated for women undergoing gonadotoxic treatments with low risk (20–30%), intermediate risk (40–70%) and high risk (70–80%) of ovarian failure (Figure 6 and Figure 7). The LR+ and LR-, which are not dependent on the prevalence rate, were also calculated (Table 53).

Table 53 Clinical validity of the AMH test against the clinical reference standard of resumption of menses

Study ID	Result	EIA AMH/MIS assay [95%CI]	DSL ACTIVE® AMH/MIS ELISA	AMH Gen II ELISA
Anders et al. (2008) AMH levels below the median prior to treatment versus RS post-treatment	Sensitivity Specificity LR+ LR-	NA	75% [19, 99] 79% [49, 95] 3.50 [1.11, 11.07] 0.32 [0.06, 1.78]	NA
D'Avila et al (2015) AMH levels <3.32 ng/mL prior to treatment versus RS post-treatment	Sensitivity Specificity LR+ LR-	85% 75% 3.40 0.20	NA	NA
Henry et al. (2014) Undetectable AMH levels prior to treatment versus RS post-treatment	Sensitivity Specificity LR+ LR-	NA	NA	95% [74, 100] 86% [42, 100] 6.63 [1.08, 40.84] 0.06 [0.01, 0.42]

AMH = anti-Müllerian hormone; CI = confidence interval; EIA = enzyme Immunoassay; LR+ = positive likelihood ratio; LR- = negative likelihood ratio; NA = not applicable; MIS = Müllerian-inhibiting substance; RS = reference standard

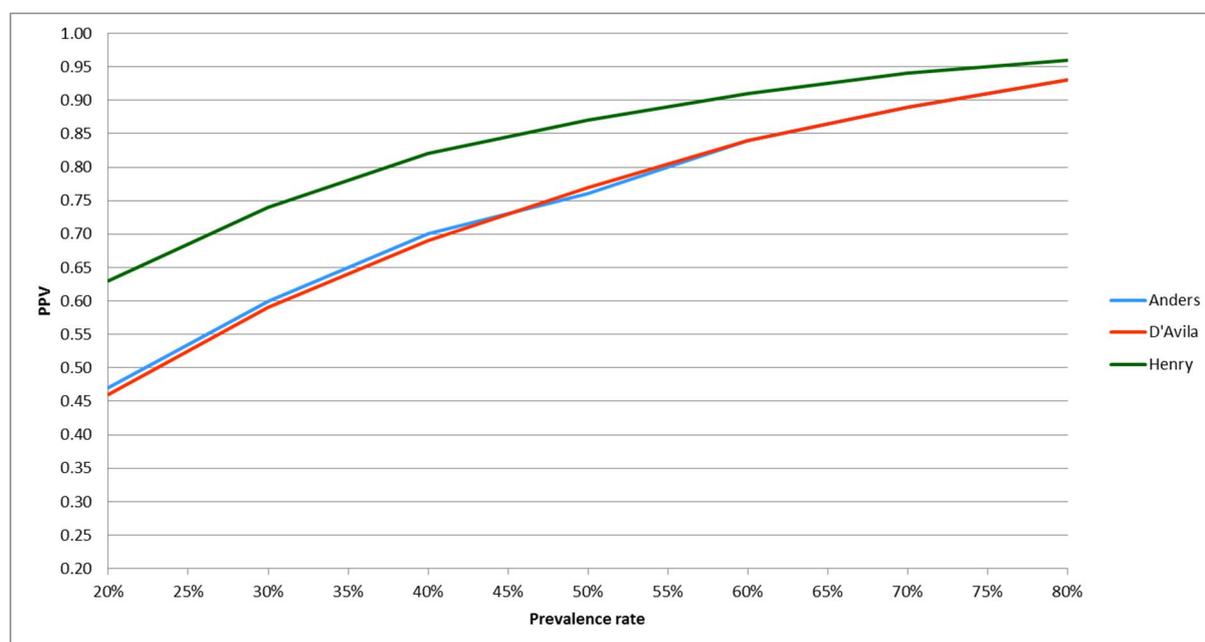


Figure 6 PPV values with increasing prevalence of treatment-induced ovarian failure from 20% to 80%, AMH measured prior to treatment

Anders = PPV for DSL ACTIVE® AMH/MIS ELISA from Anders et al. (2008); D'Avila = PPV for EIA AMH/MIS assay from D'Avila et al. (2015); Henry = PPV for AMH Gen II ELISA from Henry et al. (2014)

Women who have treatment associated with a low risk (20–30%) of ovarian failure, tested using the generation 1 DSL ACTIVE® AMH/MIS ELISA or EIA AMH/MIS assays, have a PPV of 0.46 to 0.60 indicating that approximately half of the women with a positive test result would actually have resumption of menses. This is equivalent to chance and is not clinically useful. Conversely, of the women having treatment with a high risk (70–80%) of ovarian failure, 89 to 93% with a positive DSL ACTIVE® AMH/MIS ELISA or EIA AMH/MIS assay result would have resumption of menses. Thus, a positive test result only offers useful information to women having intermediate to high-risk

treatment. In this scenario, at least 76% of women with a $\geq 50\%$ risk and a positive result would actually resume menses and thus be true positives (Figure 6).

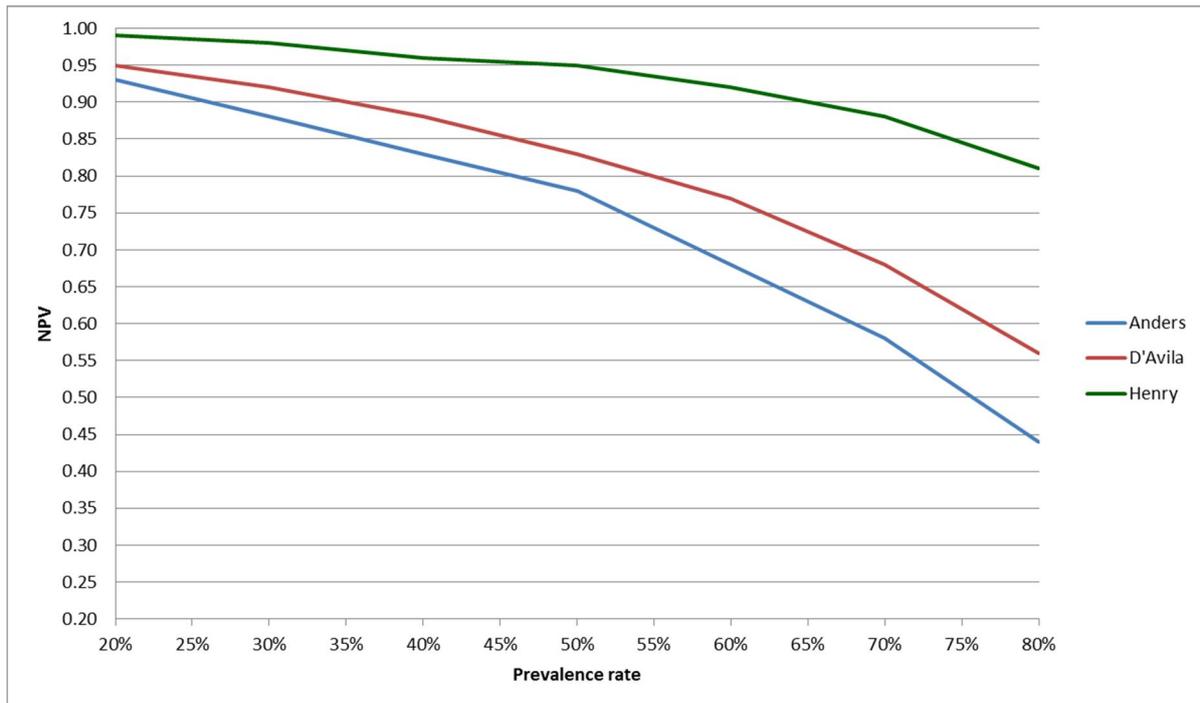


Figure 7 NPV values with increasing prevalence of treatment-induced ovarian failure from 20% to 80%, AMH measured prior to treatment

Anders = NPV for DSL ACTIVE® AMH/MIS ELISA from Anders et al. (2008); D'Avila = NPV for EIA AMH/MIS assay from D'Avila et al. (2015); Henry = NPV for AMH Gen II ELISA from Henry et al. (2014)

On the other hand, a negative test result only offers useful information to women having low-risk treatment, where the NPV indicated that 88 to 95% of women with a negative test result would not have resumption of menses. Among women having treatment with a high risk of ovarian failure, 44 to 68% with a negative test result would not have resumption of menses (Figure 7). This is equivalent to chance and not clinically useful.

The point estimates for the LR+ and LR- calculated for the DSL ACTIVE® AMH/MIS ELISA and EIA AMH/MIS assays, confirm the conclusions reached due to the PPV and NPV values. The LR+ point estimate and 95% CIs for the DSL ACTIVE® AMH/MIS ELISA indicate that a positive test result is 3.5 times more likely, or at least as likely with 95% confidence, to come from women who have resumed menstruation, compared to those women who have not. The LR- point estimate indicates that a negative test result is three to five times more likely to come from women who have not resumed menstruation than from those who have. However, the 95% CIs for the LR- from the DSL ACTIVE® AMH/MIS ELISA indicate that a positive test result is just as likely to come from a woman with resumption of menses as from a woman without menses. Thus, these assays may provide no additional information on the likelihood of resuming and/or continuing menstruation after treatment. The 95% CIs could not be derived for the EIA AMH/MIS assay.

The AMH Gen II ELISA was more accurate. Among the women who have treatment associated with a low risk of ovarian failure, the PPV indicates that 63 to 74% with a positive test result would have resumption of menses. Conversely, 94 to 96% of women having treatment with a high risk of ovarian failure who have a positive test result would have resumption of menses (Figure 6). Therefore, the test is clinically useful at most prevalence rates. A negative test result provided useful information for all women who had treatment associated with a risk of ovarian failure. The NPV indicated that 81 to 99% of women with a negative test result would not resume menses (Figure 7).

The LR+ and LR- estimates and 95% CIs also indicate that the AMH Gen II ELISA is more accurate than the first generation assay, and provides some useful information. The LR+ point estimates and 95% CIs indicate that a positive test result from the pico-AMH ELISA is 7-fold more likely, or at least as likely with 95% confidence, to come from women who have resumed menstruation compared to those who have not. Similarly, the LR- indicates that a negative test result is more than ten times more likely, or at least two times more likely with 95% confidence, to have come from women who have not resumed menstruation compared to those who have. Thus, the LR+ and LR- indicate an increase in confidence that the AMH Gen II ELISA test result is correct.

Clinical validity of AMH testing compared with a clinical reference standard when measured after gonadotoxic treatment

The study by Decanter et al. (2014) compared the sensitivity and specificity of the EIA AMH/MIS assay and pico-AMH ELISA to detect ovarian reserve with the resumption of menses during the ovarian recovery period in 30 women with either early breast cancer or lymphoma, three to 24 months after the end of chemotherapy. The distribution of the serum samples was as follows: nine were taken three months after chemotherapy, 13 at six months, 11 at nine months, 18 at 12 months, two at 18 months, and five at 24 months. The study by Chai et al. (2014) also looked at the ability of the pico-AMH ELISA to detect AMH in one-year post-chemotherapy serum of 42 women with early stage breast cancer who had or had not resumed menstruation three years after chemotherapy.

Su et al. (2011) used Poisson regression methods to model the optimum thresholds for serum levels of AMH using an older first generation AMH test, the DSL ACTIVE® AMH/MIS ELISA, which had a lower limit of detection for AMH of 25 pg/mL. Using ROC curve analysis and resumption of menses at least one year post-treatment as the reference standard, the authors calculated the sensitivity and specificity of the AMH test to detect AMH in 56 late reproductive-aged breast cancer patients to be 60% and 76%, respectively.

The study by Jantke et al. (2012) investigated the ability of the AMH test to detect AMH in 86 women who had resumed menstruation compared with those who had not after receiving gonadotoxic treatment up to 14 years previously.

The PPV and NPV were calculated for women undergoing gonadotoxic treatments with low risk (20–30%), intermediate risk (40–70%) and high risk (70–80%) of ovarian failure. The LR+ and LR- are shown in Table 54.

Table 54 Clinical validity of the AMH test against the clinical reference standard of resumption of menses

Study ID	Result	EIA AMH/MIS assay [95%CI]	DSL ACTIVE® AMH/MIS ELISA	Pico-AMH ELISA [95%CI]	AMH measured by diagnostic laboratory
Decanter et al. (2014) Undetectable AMH versus RS at time of AMH	Sensitivity Specificity LR+ LR-	11% [2, 28] 93% [78, 99] 1.61 [0.29, 8.92] 0.96 [0.82, 1.12]	NA	71% [51, 67] 83% [65, 94] 4.29 [1.86, 9.87] 0.34 [0.19, 0.63]	NA
Chai et al (2014) Undetectable AMH at 1-year post-chemotherapy versus RS at 3-years	Sensitivity Specificity LR+ LR-	NA	NA	91% (59, 100] 74% [55, 88] 3.52 [1.88, 6.58] 0.12 [0.02, 0.8]	NA
Jantke et al. (2012) Undetectable AMH versus RS at time of AMH	Sensitivity Specificity LR+ LR-	NA	NA	NA	86% [65, 97] 100% [48, 100] infinity 0.14 [0.05, 0.39]
Su et al. (2011) Undetectable AMH versus RS at time of AMH	Sensitivity Specificity LR+ LR-	NA	60% 76% 2.50 0.53	NA	NA

AMH = anti-Müllerian hormone; CI = confidence interval; EIA = enzyme immunoassay; LR+ = positive likelihood ratio; LR- = negative likelihood ratio; NA = not applicable; MIS = Müllerian-inhibiting substance; RS = reference standard

Women who have treatment associated with a low risk (20–30%) of ovarian failure, tested using the first generation DSL ACTIVE® AMH/MIS ELISA or EIA AMH/MIS assays, have a PPV of 0.28 to 0.52, indicating that only 28 to 52% of the women with a positive test result would actually have resumption of menses (Figure 8). This is no better than chance. Conversely, of the women having treatment with a high risk (70–80%) of ovarian failure, 79 to 91% with a positive EIA AMH/MIS assay result would have resumption of menses. Thus, a positive test result only offers useful information for women having high-risk treatment, with at least 79% of those women receiving a positive result resuming menses, being true positives.

On the other hand, a negative test result only offers useful information to women having low-risk treatment, where the NPV indicated that 71 to 88% of women with a negative test result would not have resumption of menses (Figure 9). Among women having treatment with a high risk of ovarian failure, 21 to 45% with a negative test result would not have resumption of menses. This is no better than chance and is not clinically useful.

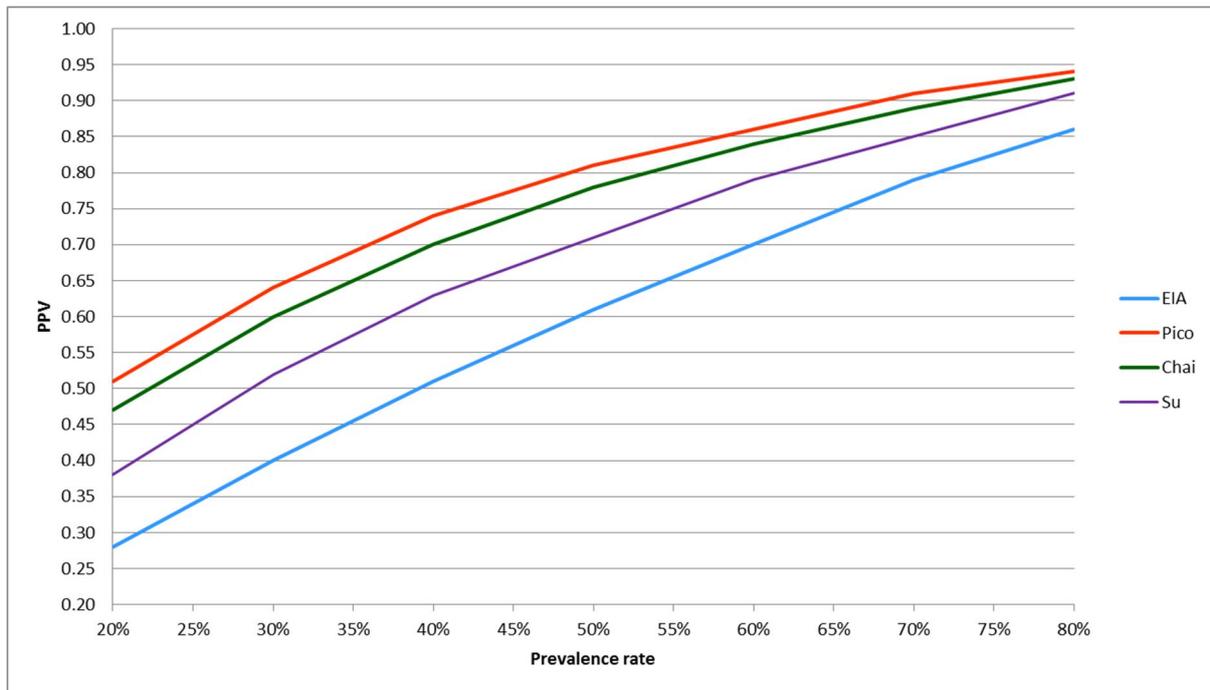


Figure 8 PPV values with increasing prevalence of treatment-induced ovarian failure from 20% to 80% (AMH measured post-treatment)

EIA = PPV for EIA AMH/MIS assay from Decanter et al. (2014); Pico = PPV for pico-AMH ELISA from Decanter et al. (2014); Chai = PPV for pico-AMH ELISA from Chai et al. (2014)

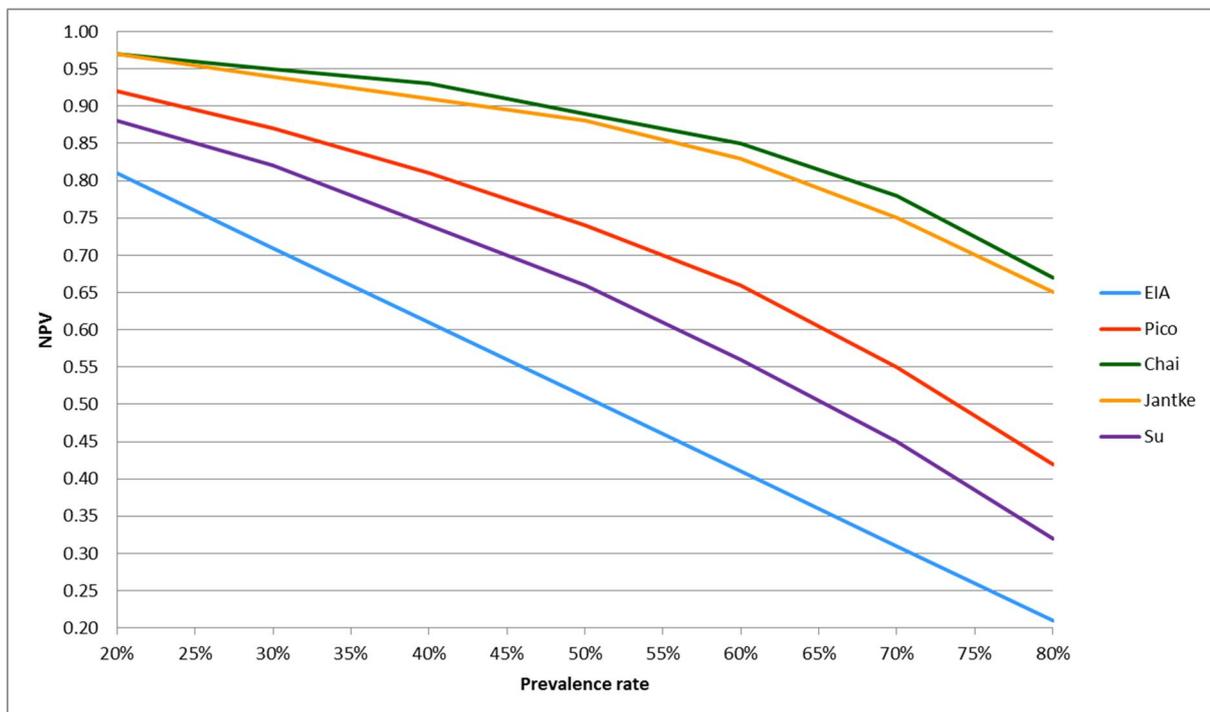


Figure 9 NPV values with increasing prevalence of treatment-induced ovarian failure from 20% to 80% (AMH measured post-treatment)

EIA = NPV for EIA AMH/MIS assay from Decanter et al. (2014); Pico = NPV for pico-AMH ELISA from Decanter et al. (2014); Chai = NPV for pico-AMH ELISA from Chai et al. (2014); Jantke = NPV for AMH measured in a diagnostic laboratory from Jantke et al. (2012)

The point estimates for the LR+ and LR- calculated for the DSL ACTIVE® AMH/MIS ELISA and EIA AMH/MIS assays confirm the conclusions based on the PPV and NPV values. However, the 95% CIs for the LR+ from EIA AMH/MIS assay indicate that a positive test result is just as likely to come from

a woman who has resumed menses as a woman without menses. Likewise, a negative test result is equally likely to come from a woman who has resumed menses as a woman without menses. Thus, the EIA AMH/MIS assay provides no additional information about ovarian reserve of these women, nor the likelihood of future menstruation status. The 95% CIs could not be derived for the DSL ACTIVE® AMH/MIS ELISA. The LR+ point estimate indicates that a positive test result is 2.5 times more likely to come from women who have resumed menstruation than from those who have not, whereas the LR- point estimate indicates that a negative test result is approximately twice as likely to come from women who have not resumed menstruation versus those who have.

The pico-AMH ELISA test was more accurate. Among the women who have treatment associated with a low risk of ovarian failure, the PPV indicates that 47 to 64% with a positive test result would have resumption of menses; however, this is only slightly better than chance. Conversely, 89 to 94% of women having treatment with a high risk of ovarian failure who have a positive test result would have resumption of menses. Therefore, the test is only clinically useful at higher prevalence rates (Figure 8).

A negative test result provided useful information for women who had treatment associated with a low risk of ovarian failure, where the NPV indicated that 87 to 97% with a negative test result would not resume menses (Figure 9). However, the clinical value of a negative test result for women having treatment with a high risk of ovarian failure differed between the two studies, with 42 to 55% of patients not resuming their menses in the study by Decanter et al. (2014) (equal to chance) and 67 to 78% not resuming their menses in the study by Chai et al. (2014).

The LR+ and LR- estimates and 95% CIs also indicate that the pico-AMH ELISA is more accurate than the EIA AMH/MIS assay, and provides some useful information. The LR+ point estimates and 95% CIs indicate that a positive test result from the pico-AMH ELISA is four times more likely to come from women who have resumed menstruation than from those who have not, or at least two times as likely with 95% confidence. Similarly, the LR- indicates that a negative test result is about three to 10 times, or 1.2-times with 95% confidence, more likely to come from women who have not resumed menstruation than from those who have. Thus, the LR+ and LR- indicate a small increase in confidence that the pico-AMH ELISA test result is correct.

In the study by Jantke et al. (2012), which used a diagnostic laboratory to measure AMH levels, the specificity of the AMH test compared to resumption of menstruation as the reference standard was 100%. Therefore, there were no false-positive patients and the PPV was 100% for all prevalence values and was not plotted. The NPV values were very similar to those for the pico-AMH test used in the study by Chai et al. (2014). The LR- indicates that a negative test result is about seven times more likely to come from women who have not resumed menstruation than from those who have, or at least 2.5-times more likely with 95% confidence. Thus, LR- indicates an increase in confidence that the diagnostic AMH test result is correct.

ROC analysis of the ability of the various hormones to discriminate between women with and without menses, after gonadotoxic treatment

Su et al. (2011) calculated the clinical validity of the AFC, AMH, FSH and inhibin B tests to distinguish between late reproductive-aged women with and without CRA using the threshold values determined by ROC curve analysis (Table 55). The AMH test used was the first generation DSL ACTIVE® AMH/MIS ELISA test.

When the PPV and NPV were compared at the study prevalence rate of CRA, approximately 60%, AFC outperformed the AMH, FSH and inhibin B tests (Table 55). The PPV values indicated that nine out of ten women who tested positive using AFC actually had CRA compared with eight out of ten for AMH, FSH and inhibin B. The NPV values showed that seven out of ten women predicted to have menses according to the AFC result actually did not have CRA, compared with six out of ten for AMH and FSH and four out of ten for inhibin B.

Table 55 Clinical validity of ovarian reserve markers using thresholds determined by ROC curve analysis

	AFC (total 2–10 mm) <1	AMH ≤25 pg/mL	FSH ≥40IU/L	Inhibin B ≤5pg/mL
PPV	90%	79%	81%	76%
NPV	73%	56%	59%	41%

AFC = antral follicle count; AMH = anti-Müllerian hormone; FSH = follicle-stimulating hormone; NPV = negative predictive value; PPV = positive predictive value; ROC = receiver operator characteristic
Source: Su et al. (2011)

B5 CLINICAL UTILITY

Clinical utility refers to how likely the test is to significantly impact on patient management and health outcomes.

No evidence of the impact of AMH on change in clinical management was identified in the eligible population, i.e. women undergoing gonadotoxic treatment. Because the test has only been part of clinical practice for a couple of years and evidence on clinical utility of the test in women undergoing gonadotoxic treatment is not yet available, it was decided to estimate the potential of AMH testing to change clinical management by looking at a broader population. To do this, a non-systematic search was conducted in a broad population, i.e. healthy women undergoing IVF treatment, to identify a possible change in patient management or decision-making influenced by AMH testing (therapeutic efficacy).

B5.1 IMPACT OF AMH TESTING ON CLINICAL MANAGEMENT (THERAPEUTIC EFFICACY)

B5.1.1 RISK OF BIAS ASSESSMENT

As no evidence was identified on clinical utility in the eligible patient population, risk of bias was not assessed for clinical utility.

B5.1.2 CHARACTERISTICS OF THE EVIDENCE BASE

The evidence presented in this section of the report did not match the proposed MBS populations. A non-systematic search was performed for existing systematic reviews, high-quality, recently published studies and international guidelines in a broader population of women undergoing IVF treatment to identify a possible change in patient management or decision-making influenced by AMH testing and associated health outcomes.

Two recently published randomised controlled trials (RCTs) and two retrospective studies were identified on whether AMH results change ovarian hyperstimulation protocols in the IVF population. These studies were included to help determine whether AMH impacts the starting dose of gonadotrophins in IVF or intracytoplasmic sperm injection (ICSI) cycles (i.e. therapeutic efficacy) and whether this leads to better health outcomes (i.e. therapeutic effectiveness).

B5.1.3 OUTCOME MEASURES AND ANALYSIS

Evidence for the impact of AMH on clinical management was reported primarily as dosage change to the controlled ovarian hyperstimulation protocol, influenced by AMH levels.

B5.1.4 RESULTS OF THE SYSTEMATIC LITERATURE REVIEW

DOES AMH TESTING IMPACT CLINICAL MANAGEMENT?

Summary

Does the addition of the AMH test lead to a change in management of female patients prior to receiving gonadotoxic treatment, compared to FSH, E2 and/or AFC ultrasound alone?

The research questions on clinical utility could not be answered, due to lack of evidence on the impact of AMH on change in clinical management among women undergoing gonadotoxic treatment.

Does AMH testing lead to a change in management of female patients following completion of gonadotoxic treatment, compared to FSH, E2 and/or AFC ultrasound if post-pubertal?

The research questions on clinical utility could not be answered, due to lack of evidence on the impact of AMH on change in clinical management among women undergoing gonadotoxic treatment.

The only evidence of impact of AMH testing on change in clinical management was found in a broader population, i.e. healthy women undergoing IVF. In women undergoing IVF, there is evidence that AMH values may influence the starting dosage of recombinant FSH or hMG during ovarian hyperstimulation for the retrieval of oocytes. However, the studies showed a lack of standardisation or guidelines regarding how the AMH score was used or should be used.

The original research questions determined a priori to determine the impact of the AMH test on patient management were: *“Does the addition of the AMH test lead to a change in management in female patients prior to receiving gonadotoxic treatment, compared to FSH, E2 and/or AFC ultrasound alone?”* and *“Is there a change in management after AMH testing in female patients following completion of gonadotoxic treatment, compared to FSH, estradiol and/or AFC ultrasound, if post-pubertal?”* Due to lack of evidence, these research questions could not be answered. We therefore aimed to show how AMH impacts management in women who did not undergo gonadotoxic treatment.

Section B4 provided material on how AMH predicts prognosis in regards to ovarian function more accurately than FSH, and E2. It is hypothesised that this information may be used to influence fertility preservation decisions. However, no data were identified showing that this is the case.

There is evidence that AMH testing can help predict how many oocytes will be retrieved through ovarian hyperstimulation, although it was not clear the extent to which AMH provided incremental information over FSH. If it could be shown that AMH testing provides superior information to FSH, then it is hypothesised that this information could be used to try and alter the dosage of gonadotrophins, to avoid too few or too many oocytes being retrieved.

AMH testing influencing starting dosage of gonadotrophins in IVF/ICSI cycles (not in target population undergoing gonadotoxic treatment)

A retrospective study by Yates et al (2011), reported on the introduction of AMH-tailored stimulation protocols in an IVF clinic in Manchester, UK, in September 2008. After the introduction of AMH testing in the IVF clinic, basal AMH levels were measured in all women within three months prior to controlled ovarian hyperstimulation. Women who had AMH levels above 48.5 pmol/L had further tests to exclude granulosa cell tumours or polycystic ovary syndrome and underwent counselling. Women with an acceptable AMH level were stratified and underwent different treatments (see Table 56). Those with higher AMH levels received lower doses of gonadotrophins and vice versa. Before AMH-tailored protocols, dosages were determined by FSH levels (10.0 IU/L threshold) and age (threshold 35 years).

Table 56 Controlled ovarian hyperstimulation protocol influenced by AMH levels, in St Mary's hospital, Manchester (from 2008)

AMH level	Controlled ovarian hyperstimulation protocol based on AMH
<2.2 pmol/L	Exclude, counsel and offer alternative assisted reproductive treatment
2.2–15.6 pmol/L	300 IU hMG + GnRH antagonist from day 6 of stimulation
15.7–28.6 pmol/L	200 IU recombinant FSH or 225 IU hMG in Long down regulation protocol + GnRH agonist
>28.6 pmol/L	150 IU hMG + GnRH antagonist from day 6 of stimulation

AMH = anti-Müllerian hormone; FSH = follicle-stimulating hormone; hMG = human menopausal gonadotrophin; GnRH = gonadotrophin releasing hormone

Source: Yates et al. (2011)

A retrospective study by Papaleo et al. (2016) reported the treatment protocol used in two centres in Italy (n=398). They stated the standard starting dose of recombinant FSH or hMG was determined based on age, body weight, AFC, AMH and FSH, ranging from 100 to 225 IU per day. The exact starting dosage was selected by the treating physician and determined by their personal experience (Papaleo et al. 2016). The doses were then adjusted according to ovarian response, observed by pelvic ultrasound on day 5 or day 6 of the stimulation cycle. The two centres displayed different controlled ovarian stimulation protocols and women were prescribed different starting doses, despite the women's characteristics being similar between centres. This is likely due to the absence of standard guidelines and appropriate patient-based tailoring of treatment, i.e. dosage partly depends on the clinician's experience. Even though this study did not report a change in management due to a standardised protocol based on AMH levels, it reported who would have received a different starting dose if a standardised nomogram based on AMH, day 3 serum FSH level, and age had been used. In this model, AMH is the leading predictor, explaining most of the model variation, followed by serum FSH and female age. Results showed that almost 90% (22/25) of women who had a suboptimal oocyte yield and starting dose would have received a higher starting dose if the nomogram had been used. Conversely, 49/398 patients (12.3%) had an oocyte yield above the target. Of these women, 26 of 49 (53%) would have received a lower starting dose if the nomogram had been followed. However, it is notable that the study design was retrospective, and if is not possible to determine if the increased or decreased dosage of FSH in women with a

suboptimal response would lead to an optimal response. Furthermore, it is not known whether the nomogram can be used for women and girls undergoing gonadotoxic treatment, as they are often younger than 25 years old, and their AMH may be low. The nomogram currently only gives an input for women aged 25 to 40 years, and there is no evidence of the validity of this method in the proposed MBS populations.

B5.2 THERAPEUTIC EFFECTIVENESS

The original research questions determined a priori to determine the effectiveness of change in management were:

- *“Does cryopreservation of ovarian tissue, oocytes or embryos prior to receiving gonadotoxic treatment lead to better family outcomes in female patients compared to no cryopreservation?”*
- *“Does cryopreservation of ovarian tissue or oocytes following completion of gonadotoxic treatment lead to better family outcomes in female patients compared to no cryopreservation?”*

These questions were based on the assumption that the evidence would show that AMH levels would influence fertility decision-making in women undergoing gonadotoxic treatment. Due to the absence of evidence to show any impact on patient management due to AMH results in the target patient population, we aimed to determine how AMH levels influence decision-making in a broad population (see B5.1.4). The change in management evidence showed that AMH results may influence the starting dosage of gonadotrophins or FSH in women undergoing controlled ovarian hyperstimulation for the retrieval of oocytes. Therefore in this section we aimed to answer the following question:

- *“Do individualised starting dosages of gonadotrophins or FSH based on AMH levels lead to better health outcomes, i.e. a higher percentage of women with an optimal number of retrieved oocytes, higher pregnancy rates and less adverse events, in women undergoing IVF/ICSI?”*

B5.2.1 RISK OF BIAS ASSESSMENT

The studies discussed in this section did not meet the PICO criteria and were therefore not assessed for risk of bias.

B5.2.2 CHARACTERISTICS OF THE EVIDENCE BASE

Three of the studies on therapeutic effectiveness are discussed below: two RCTs and one retrospective cohort study. These studies did not match the proposed MBS populations and therefore did not meet the PICO criteria (see B5.1.2).

B5.2.3 OUTCOME MEASURES AND ANALYSIS

The key outcomes reported in the therapeutic effectiveness section were the mean number of retrieved oocytes, fertilisation, implantation, pregnancy and birth rates, and the incidence of OHSS. Groups with standard FSH starting dosage were compared with groups with an individualised starting dosage, partly influenced by AMH test results.

B5.2.4 RESULTS OF THE SYSTEMATIC LITERATURE REVIEW

DOES THE CHANGE IN MANAGEMENT IMPROVE HEALTH OUTCOMES?

Summary

Does cryopreservation of ovarian tissue, oocytes or embryos prior to receiving gonadotoxic treatment lead to better family outcomes in female patients compared to no cryopreservation?

This question was not answered here due to lack of evidence indicating a change in management.

Does cryopreservation of ovarian tissue or oocytes following completion of gonadotoxic treatment lead to better family outcomes in female patients compared to no cryopreservation?

This question was not answered here due to lack of evidence indicating a change in management.

No evidence was found to determine how the change in management due to AMH testing impacts health-related outcomes in women undergoing gonadotoxic treatment (the target population).

Do individualised starting dosages of gonadotrophins or FSH based on AMH levels lead to better health outcomes, i.e. a higher percentage of women with an optimal number of retrieved oocytes, higher pregnancy rates and less adverse events, in women undergoing IVF/ICSI?

Three studies were identified on whether an individualised starting dosage based on AMH levels impacted IVF outcomes, and a higher rate of optimal oocyte yield in the individualised dosage group was identified in all three studies. However, a difference in pregnancy or birth rates between groups was only identified in one study, and due to the before and after design of the study, and the introduction of a number of organisational and procedural changes alongside the introduction of AMH-tailored dosing, it was not known which factors were contributing to these outcomes. The two clinical trials were not able to find a difference in pregnancy or birth rates between standard FSH dosage and AMH informed individualised dosage.

Although individualised dosage did not have an impact on live birth rates, two out of three studies showed a significant reduction in the incidence of OHSS and/or preventive interventions for OHSS using the AMH-tailored dosage protocol compared to a standardised dosage in a population of healthy women undergoing IVF. One of these trials compared a standard dose of follitropin alpha with an AMH and bodyweight-tailored dose of follitropin

delta. It is unknown whether the outcome differences between groups in this trial were due to the individualised dosage or the different follitropin used.

Notably, the studies discussed in the clinical utility section mostly excluded patients at the extremes of ovarian reserve, and did not meet the PICO criteria. The generalisability of this evidence is unknown.

It should be noted that the studies discussed in the clinical utility section mostly excluded patients at the ‘extremes’ of ovarian reserve, and did not meet the PICO criteria. The generalisability of this evidence is unknown.

An RCT by Allegra et al. (2017) was published last year to investigate the performance of the nomogram used by Papaleo et al. (2016) (see section B5.1.4) in selecting the most appropriate FSH starting dose in IVF/ICSI cycles (Allegra et al. 2017). They included 194 otherwise healthy women undergoing IVF/ICSI cycles, and randomised them to receive a standard starting dose of rFSH based on their age (150 IU if ≤35 years and 225 IU if >35 years) or on the basis of their ovarian reserve and age, by using the nomogram including age, day 3 serum FSH levels and AMH. Three women dropped out due to personal reasons (n=2) or spontaneous pregnancy (n=1). Women were aged between 18 and 40 years, had serum AMH concentrations between 1.0 and 4.0 ng/mL and normal menstrual cycles. Women with endometriosis, previous ovarian surgery and any known metabolic or endocrinological disease were excluded from this study, i.e. those who would be relevant to the gonadotoxic treatment group were specifically excluded. The most prevalent cause of infertility was of male origin in both groups (55/99; 56% and 52/92; 57% for the control group and nomogram group, respectively). The results are shown in Table 57. The number of growing follicles (≥11 mm), number of large follicles (≥17 mm), treatment duration, number of retrieved and mature oocytes were not significantly different between groups. Furthermore, patients in the two groups showed comparable fertilisation, cleavage, implantation and clinical pregnancy rates. The percentage of women with at least one cryopreserved embryo was not significantly different between groups (27/83; 28.9% vs 17/84; 20.2% for the nomogram and the control group, respectively). There were no cases of moderate to severe OHSS.

Table 57 IVF parameters and outcomes in the nomogram group (individualised FSH dosage) compared with a control group (standardised FSH dosage)

Parameter / outcome	Control group (n=99)	Nomogram group (n=92)	p-value
Mean starting dose of rFSH, mean IU ± SD	182.6 ± 37.4	201.1 ± 28.4	0.001
Women with dose adjustment, n (%)	72 (73)	56 (61)	0.01
Women with optimal (8-14) retrieved oocytes, n (%)	42 (42.4)	58 (63.0)	0.0037
Women with <8 retrieved oocytes, n (%)	40 (40.4)	24 (26.1)	0.040
Women with >14 retrieved oocytes, n (%)	17 (17.2)	10 (10.9)	NS
Women with at least one cryopreserved embryo, n (%)	17/84 (20.2)	24/83 (28.9)	NS
Clinical pregnancy rate per embryo transfer, n (%)	32 (41.0)	29 (39.7)	NS
Clinical pregnancy rate per started cycle, n (%)	32 (32.3)	29 (31.5)	NS

Source: (Allegra et al. 2017)

rFSH = recombinant follicle-stimulating hormone; NS = not significant; SD = standard deviation; IU = international units

Clinical pregnancy rate and live birth rate are considered the most important outcomes, whereas number of oocytes retrieved is considered an intermediate outcome. In the trial by Allegra et al. the clinical pregnancy rate and number of cryopreserved embryos was similar in both groups, however it was stated the study was not designed to detect any difference that was not the primary outcome, i.e. the rate of women with an appropriate response. Furthermore, it is not known whether the nomogram is usable for women and girls undergoing gonadotoxic treatment, as they are often younger than 25 years old, and their AMH may be low. The nomogram currently only gives an input for women aged 25 to 40 years and there is no evidence of the validity of this method in the proposed MBS populations. Patients at the extremes of ovarian reserve were excluded from the trial, and it is these women particularly who may benefit from personalised treatment (Allegra et al. 2017).

The retrospective study by Yates et al. (2011) reported outcomes of the AMH-tailored dosage protocol compared with the standardised dosage before AMH, in addition to the change in management evidence presented in section B5.1.4. The pre-embryology outcomes of this study are shown in Table 58. There was a significant reduction in the incidence of OHSS using the new AMH-tailored protocol. Furthermore, a higher number of women in the AMH group underwent embryo transfer compared with the conventional group (87.5% vs 78.9%; $p=0.002$), the overall pregnancy rate per cycle improved from 17.9% to 27.7% ($p=0.002$) and the live birth rate increased from 15.9% to 23.9% ($p=0.007$). However, at the same time of the introduction of the AMH-tailored protocols, there was a change in embryological culture media used during the fertilisation and preimplantation of oocytes. Alongside this change, there were a number of organisational and protocol changes in the laboratory used in the study (Yates, Roberts & Nardo 2012). Therefore, given the before and after design of the study, it is not possible to determine which factors were contributing to the different outcomes. Some of the observed effects, i.e. the post-embryology outcomes, may not be solely related to the AMH-tailored dosage protocol, and are therefore not presented. Yates et al. (2012) did show that an AMH-tailored protocol may decrease the incidence of OHSS in women undergoing IVF treatment.

Table 58 Pre-embryology clinical outcomes in a conventional dosage protocol compared with an AMH-tailored protocol

Clinical outcomes	Conventional protocol (n=346)	AMH-tailored protocol (n=423)	Unadjusted p-value (fisher's exact test)	Adjusted p-value (logistic regression) ^a
Cancelled cycle due to poor response, n (%)	14 (4.0)	14 (3.3%)	0.7	0.57
Cancelled cycle due to elective freeze all, n (%)	0 (0.0)	3 (0.7%)	0.26	0.066
Cancelled cycle due to other reasons, n (%)	4 (1.2)	4 (0.9%)	1	0.8
Mean number \pm SD of oocytes retrieved	12.4 \pm 7.8	10.6 \pm 6.9	0.001 ^b	0.007 ^b
OHSS leading to cycle	24 (6.9)	10 (2.3%)	0.002	0.004

Clinical outcomes	Conventional protocol (n=346)	AMH-tailored protocol (n=423)	Unadjusted p-value (fisher's exact test)	Adjusted p-value (logistic regression) ^a
cancellation and/or freeze all, n (%)				
OHSS leading to hospital admission, n (%)	10 (2.9)	5 (1.2%)	0.12	0.15

AMH = anti-Müllerian hormone; OHSS = ovarian hyperstimulation syndrome; SD = standard deviation

^a adjusted for age, previous pregnancy, male factor, unexplained cause, and for post-transfer end-points, ICSI and number of embryos transferred.

^b Mann-Whitney U-test for unadjusted and ordinary regression analysis for adjusted.

Source: (Yates et al. 2011).

The ESTHER-1 trial suggested that AMH testing can be used to individualise the dosage of a new rFSH (follitropin delta). The rFSH follitropin delta is currently being assessed by the Pharmaceutical Benefits Advisory Committee for reimbursement in Australia, and if listed, AMH levels will be used to inform the dosage of this hormone. The ESTHER-1 study was a randomised, multicentre trial which compared the efficacy and safety of follitropin delta with individualised dosing based on AMH and body weight, with conventional follitropin alfa dosing for ovarian stimulation in women undergoing IVF (Nyboe Andersen et al. 2017). The RCT was conducted at 37 different locations in 11 different countries (Belgium, Brazil, Canada, Czech Republic, Denmark, France, Italy, Poland, Russia, Spain, and United Kingdom). Women in the intervention group were randomised to a fixed subcutaneous dose of follitropin delta, determined by their serum AMH level at screening using the automated Elecsys AMH Immunoassay by Roche Diagnostics International and body weight (AMH <15 pmol/L: 12 µg; AMH ≥15 pmol/L: 0.10–0.19 µg/kg; the maximum daily dose was 12 µg). Women randomised to follitropin alpha received a daily subcutaneous dose of 150 IU (11 µg) in the first five days, not determined by ovarian reserve tests or other factors. After five days the dose could be adjusted based on follicular response, with 450 IU set as the maximum dose. All pregnancies were followed until four weeks after live birth. The main baseline characteristics were similar in both intervention groups. The main outcomes per group are shown in Table 59. No significant differences were observed in pregnancy, live birth or implantation outcomes between intervention groups. There were no significant differences between treatment groups in terms of oocytes retrieved. However, among women with an AMH level of <15 pmol/L, i.e. potential hypo-responders, individualised follitropin delta was associated with more oocytes (8.0 compared to 7.0; p=0.004). In women with an AMH of ≥15 pmol/L, individualised follitropin delta was associated with a lower oocyte yield (11.6 compared to 13.3; p=0.002). The individualised dosage group also had fewer women requiring OHSS preventive measures.

Table 59 Outcomes after individualised follitropin delta use compared to standard follitropin alpha use

Outcome per started cycle	Follitropin delta dose based on AMH and body weight (n=665)	Standard follitropin alpha dose (n=661)	Difference [95%CI] or p-value
Ongoing pregnancy (one viable fetus 10-11 weeks after transfer), n (%)	204 (30.7)	209 (31.6)	-0.9% [-5.9%, 4.1%]
Ongoing implantation, number of viable	206/585 (35.2)	209/584 (35.8)	-0.6% [-6.1, 4.8%]

Outcome per started cycle	Follitropin delta dose based on AMH and body weight (n=665)	Standard follitropin alpha dose (n=661)	Difference [95%CI] or p-value
fetuses 10-11 weeks after transfer divided by number of blastocysts transferred (%)			
Women with live birth, n (%)	198 (29.8)	203 (30.7)	-0.9% [-5.8, 4.0%]
Women with live neonate at 4 weeks after birth, n (%)	198 (29.8)	201 (30.4)	-0.6% [-5.5%, 4.3%]
Extreme ovarian response (<4 or ≥15 oocytes), n (%)	169 (26.6)	201 (31.3)	0.001
Women with AMH <15 pmol/L, n	280	290	
<i>Mean number oocytes retrieved ± SD</i>	<i>8.0 ± 4.3</i>	<i>7.0 ± 3.9</i>	<i>0.004</i>
<i>Poor responders (<4 oocytes), n (%)</i>	<i>33 (11.8)</i>	<i>52 (17.9)</i>	<i>0.039</i>
Women with AMH ≥15 pmol/L, n	355	353	
<i>Mean number oocytes retrieved ± SD</i>	<i>11.6 ± 5.9</i>	<i>13.3 ± 6.9</i>	<i>0.002</i>
<i>Excessive responders (≥15 oocytes), n (%)</i>	<i>99 (27.9)</i>	<i>124 (35.1)</i>	<i>0.038</i>
Safety outcomes			
Preventive interventions, n (%)	15 (2.3)	30 (4.5)	0.005
Early OHSS (any grade), n (%)	17 (2.6)	20 (3.0)	0.291
All OHSS (any grade), n (%)	23 (3.5)	32 (4.8)	0.238
Hospitalisation due to OHSS, n (%)	2 (0.3)	6 (0.9)	0.108

AMH = anti-Müllerian hormone; CI = confidence interval; OHSS = ovarian hyperstimulation syndrome
Source: (Nyboe Andersen et al. 2017)

This evidence corresponds with the recently published study by the OPTIMIST study group (van Tilborg et al. 2017). In this study FSH dosage was based on AFC counts in women with infertility problems undergoing IFV/ICSI, and an individualised FSH dose did not result in better live birth rates (n=1,515). They did observe a reduction of the rate of mild and moderate OHSS with individualised dosing. When a post-hoc analysis was done with AMH levels as the ovarian reserve test, using the statistical method of standardisation to correct for non-concordant test results, comparable results were observed.

B6 IMPACT OF REPEAT TESTING / MONITORING

The proposed MBS item descriptor indicates that repeat AMH testing will be performed for patients who have had gonadotoxic treatment, “to assess the gonadotoxic effects of the treatment, to assess pubertal delay, to assess ovarian failure, to assess the need for fertility preservation following treatment, and to assess the need for assisted reproductive treatment for family planning”.

Studies have been identified in which AMH is used to monitor the decline or recovery of AMH during and after gonadotoxic treatment (Decanter et al. 2010; Dillon et al. 2013; Everhov et al. 2014; Hamy et al. 2014; Kim, YJ, Cha & Kim 2017; Somigliana et al. 2012; van der Kooi et al. 2017). The aim of these studies was to determine the gonadotoxic effects of certain treatments.

Decanter et al. (2010), Hamy et al. (2014) and Dillon et al. (2013) aimed to investigate the effects of chemotherapy on the ovarian follicles and measures of ovarian reserve, with repeated AMH measurements which were performed prior to, during and after treatment in cancer patients. Dillon et al. and Hamy et al. reported an acute impaired ovarian reserve during treatment and a degree of recovery post-treatment. Decanter reported similar results, with a strong decrease in AMH concentrations after the start of chemotherapy. Post-treatment AMH concentrations increased by the third month of follow-up, with the degree of recovery dependent on the toxicity of the chemotherapy drugs used. The study by Everhov et al. (2014) followed pre-menopausal women with cervical cancer, investigating the effects of surgery and or chemoradiation on AMH levels. In women undergoing radical hysterectomy, pelvis lymphadenectomy and salpingo-oophorectomy and/or chemoradiation, levels of AMH were undetectable in all women (n=23) after treatment. A 45% reduction in AMH levels was seen in the women undergoing radical hysterectomy and pelvic lymphadenectomy with ovarian preservation (n=9).

The study by van der Kooi et al. (2017) investigated whether the decline in AMH levels in childhood cancer survivors differed from that observed in a healthy normal population. In this investigation of longitudinal changes in ovarian function over time, median AMH levels in cancer survivors were below the 50th percentile in a healthy population, at 5 years after cessation of treatment and at a second follow-up visit (median 3.2 years after the first visit). In women with a sustained ovarian function, the decline in AMH levels were similar to that in the normal, healthy population.

One systematic review was identified on serum AMH level modification after surgical excision of ovarian endometriomas (Somigliana et al. 2012). Nine of 11 studies showed a significant decrease in serum AMH levels after surgery. The study by Kim et al. (2017) also measured preoperative and post-operative AMH levels to assess the effect of cystectomy on ovarian reserve in patients with endometrioma and other benign cysts (n=75). A significant decline in AMH levels was observed in the group with endometriosis when comparing pre- and post-operative AMH levels, similar to the systematic review by Somigliana et al.

In conclusion, the evidence identified a decline in AMH levels after gonadotoxic treatments, with the possibility of recovery of AMH levels over time based on level of toxicity, the woman's age and other factors. However, no evidence was identified on how the AMH test results obtained after completion of gonadotoxic treatment were used in clinical practice. Evidence is lacking on how repeat testing or monitoring AMH levels changes patient management.

B7 EXTENDED ASSESSMENT OF COMPARATIVE HARMS

For an add-on investigation, any additional harms are those associated with the add-on test itself. As AMH testing can be undertaken as part of a routine a blood test, no additional harms due to the addition of AMH testing are expected.

There are theoretical harms which may result downstream if the results of an AMH test are used to guide decisions or treatments which are more harmful than management in the absence of information provided by AMH testing. For example, if a low AMH test result means that a patient is more likely to undergo fertility preservation procedures such as ovarian tissue, oocyte or embryo cryopreservation, the patient may delay treatment by a short interval. If it were found that this interval negatively affected their health, then this could be a downstream harm of AMH. However, there was no evidence of AMH influencing decisions or management of patients prior to or following gonadotoxic treatment. In the broader population of those undergoing IVF for male or female infertility, use of individualised dosing, incorporating AMH levels to determine use of gonadotropins, was associated with a reduced risk of OHSS than the group with standardised dosing. It is therefore hypothesised that use of AMH testing may result in superior safety than not using AMH testing, if this information is used to influence dosing of gonadotropins.

B8 INTERPRETATION OF THE CLINICAL EVIDENCE

It is important to classify the therapeutic profile of the proposed investigative test and associated interventions in relation to the main comparators, i.e. whether it is therapeutically superior, inferior or equivalent to the comparators.

The evidence profile summarised in Table 60 suggests that **relative to other ovarian reserve tests alone, the AMH test and associated interventions have non-inferior safety and uncertain incremental effectiveness.**

No direct evidence was identified to determine the effectiveness of AMH testing in addition to other standard tests, compared to other standard tests alone, in patients prior to or following completion of gonadotoxic therapy. No studies meeting the PICO criteria regarding safety of AMH testing were identified. However, the test is done through a routine blood test and this is generally considered safe.

AMH testing was proposed by the applicants to be an additional investigation to those already used in clinical practice for women who are at risk of premature ovarian failure due to gonadotoxic treatment. Therefore, in order to demonstrate the benefit of AMH testing, evidence of incremental benefit over and above the tests already being performed is required. AMH testing prior to gonadotoxic treatment was found to be a predictor of amenorrhea or menses after gonadotoxic treatment, even when other prognostic factors such as age and FSH were controlled for in multivariate analyses. This suggests that AMH does provide some additional useful information for predicting ovarian functioning. However, the relationship between AMH testing and the most clinically relevant outcome of a live birth was not significant. No studies were identified which actually reported on how the prognostic information is being used in the population undergoing gonadotoxic treatment.

AMH levels do significantly relate to the number of oocytes which are able to be retrieved for IVF. In the broader population, there is evidence that this information can be used to prevent cases of OHSS, although the study which demonstrated this was confounded by a change in FSH, and the method used to determine dosage.

Studies on the analytical validity of the AMH test suggest AMH levels have a moderate to strong positive correlation with AFC and inhibin B levels. This is not surprising, as these counts and hormone levels trend downwards as women approach menopause (Figure 1). Different AMH assays correlated highly with each other, although they differed greatly in how sensitive they were. This leads to concerns that different thresholds will need to be determined for each assay, as the interpretation of scores will differ.

Table 60 Summary of findings table

Section in report	Aim / outcomes	Participants (studies)	Quality of evidence	Results	Interpretation	Quality of evidence using GRADE
B1 Direct evidence	Direct effectiveness and safety of AMH testing and comparative tests	K=0 N=0	NA	No evidence was identified on the direct effectiveness or harms of AMH testing in women undergoing gonadotoxic treatment.	No conclusions can be made on direct effectiveness. No safety concerns were raised. AMH is measured through a routine blood test and this is generally considered safe.	NA
B3 Analytical validity	Accuracy of AMH in diagnosing ovarian failure (compared to other test(s))	K=2 N=113	Risk of bias: 0 Inconsistency: 0 Indirectness: -1 Imprecision: 0 Other considerations: 0	One study reported that sensitivity of AFC, AMH and FSH tests did not vary greatly however AFC was more specific than AMH, FSH and inhibin B. One study reported a 40-fold difference in the level of detectable AMH between two different AMH tests.	Analytical validity can be highly impacted by the type of assay/test used. No conclusions can be drawn on the incremental diagnostic value of AMH testing in diagnosing ovarian failure.	⊕○○○ Very low quality
	Concordance / correlation of AMH with comparator tests	K=12 N=703	Risk of bias: 0 Inconsistency: -1 Indirectness: -1 Imprecision: 0 Other considerations: 0	AMH levels are positively correlated with AFC, density of primordial follicles, and inhibin B serum levels. The correlations between AMH and FSH or E2 were inconsistent.	Correlations were observed between AMH and the comparator tests.	⊕○○○ Very low quality
	Correlation between AMH assays	K=2 N=98	Risk of bias: 0 Inconsistency: 0 Indirectness: -1 Imprecision: 0 Other considerations: 0	There was a high degree of correlation between the different AMH tests ($\rho > 0.9$).	Different AMH tests are highly correlated, but there may be systemic biases when converting AMH values between assays.	⊕○○○ Very low quality
B4 Prognostic and predictive value	Prognostic value: AMH predicting ovarian failure	K=11 N=861	Risk of bias: 0 Inconsistency: 0 Indirectness: 0 Imprecision: 0 Other considerations: 0	In most multivariate analyses, baseline AMH levels remained a predictor for ovarian function. AMH testing post-treatment also performs well in predicting CRA.	An incremental prognostic value of AMH testing in predicting ovarian function was observed. AMH testing may be a predictor for CRA at follow-up.	⊕⊕○○ Low quality

Section in report	Aim / outcomes	Participants (studies)	Quality of evidence	Results	Interpretation	Quality of evidence using GRADE
	Prognostic value: AMH predicting pregnancy / live births	K=6 N=453	Risk of bias: -1 Inconsistency: 0 Indirectness: -1 Imprecision: -1 Other considerations: 0	Most studies did not find a significant relationship between AMH levels and pregnancy rate.	No conclusions can be drawn on the incremental prognostic value of AMH testing in predicting pregnancy or live birth rate.	⊕⊖⊖⊖ Very low quality
	Predictive value: AMH predicting ovarian response	K=4 N=446	Risk of bias: 0 Inconsistency: 0 Indirectness: -1 Imprecision: 0 Other considerations: 0	One study presented that AMH had moderate (AUC 0.7 to 0.8) to good (0.8 to 0.9) test performance at predicting oocyte yield. AFC ultrasound had a good (0.8 to 0.9) test performance. The other three studies found a moderate relationship between AMH and oocyte yield.	An association was found between AMH levels and ovarian response to hyperstimulation in women undergoing gonadotoxic treatment. However, no conclusions can be drawn on the incremental predictive value of AMH testing in determining response to ovarian stimulation.	⊕⊖⊖⊖ Very low quality
B4 Clinical validity	Accuracy and usefulness of AMH	K=7 N=310	Risk of bias: 0 Inconsistency: 0 Indirectness: -1 Imprecision: 0 Other considerations: 0	First generation AMH tests showed lack of utility AMH measured both prior to and after treatment compared to resumption of menses / CRA. Second generation tests perform better, however the positive test result was only clinically useful in women at high risk of ovarian failure, whereas the negative test result was only useful in the group at low risk of ovarian failure.	Regarding the second generation AMH tests and lab tests, a positive AMH result (detectable AMH) only gives an accurate prediction of ovarian function in women at high risk of ovarian failure. A negative test result (undetectable AMH) is only able to predict amenorrhea in women at low risk of ovarian failure. First generation AMH tests showed lack of utility in predicting ovarian function.	⊕⊖⊖⊖ Very low quality
B5 Clinical utility	Change in clinical management due to AMH test results (therapeutic efficacy)	K=0 N=0	NA	No evidence was found on AMH informing change in management in women undergoing gonadotoxic treatment. In a broader population, AMH values may influence the starting dose of hormones for ovarian hyperstimulation.	The research questions on clinical utility could not be answered, due to lack of evidence on the impact of AMH on change in clinical management in women undergoing gonadotoxic treatment.	NA

Section in report	Aim / outcomes	Participants (studies)	Quality of evidence	Results	Interpretation	Quality of evidence using GRADE
	Impact of change in management on health outcomes (Therapeutic effectiveness)	K=0 N=0	NA	No evidence was found to determine how change in management due to AMH testing impacts health outcomes in women undergoing gonadotoxic treatment. The only evidence identified was conducted in women undergoing IVF, and suggested that individualised dosage based on AMH level may lead to a decrease in OHSS in a broad population.	The research questions on clinical utility could not be answered, due to lack of evidence in the targeted patient population. The generalisability of the available evidence in a broad population to the target population is unknown.	NA

AFC = antral follicle count; AMH = anti-Müllerian hormone; AUC = area under the curve; CRA = chemotherapy-related amenorrhea; E2 = estradiol; FSH = follicle-stimulating hormone; IVF = in vitro fertilisation NA = not available

SECTION C

TRANSLATION ISSUES

AMH is proposed as an additional test to the current ovarian reserve tests AFC, FSH and E2. No clinical evidence was found showing the incremental benefit of this additional test over the current practice. Therefore, only a cost-comparison is provided in Section D and there are no translation studies.

SECTION D

ECONOMIC EVALUATION

D.1. OVERVIEW

The clinical evaluation suggested that, relative to other ovarian reserve tests alone, the AMH test and associated interventions have **non-inferior safety and uncertain incremental effectiveness**. Table 61 sets out the framework that was used to classify the clinical evidence in Section B, so that a decision could be made about the type of economic analysis to undertake in this section.

Table 61 Classification of the comparative effectiveness and safety of the proposed therapeutic medical service compared with its main comparator and guide to the suitable type of economic evaluation

Comparative safety		Comparative effectiveness		
	Inferior	Uncertain ^a	Non-inferior ^b	Superior
Inferior	Health forgone: need other supportive factors	Health forgone possible: need other supportive factors	Health forgone: need other supportive factors	? Likely CUA
Uncertain ^a	Health forgone possible: need other supportive factors	?	?	? Likely CEA/CUA
Non-inferior ^b	Health forgone: need other supportive factors	?	CMA	CEA/CUA
Superior	? Likely CUA	? Likely CEA/CUA	CEA/CUA	CEA/CUA

CEA=cost-effectiveness analysis; CMA=cost-minimisation analysis; CUA=cost-utility analysis

? = reflect uncertainties and any identified health trade-offs in the economic evaluation, as a minimum in a cost-consequences analysis.

^a 'Uncertainty' covers concepts such as inadequate minimisation of important sources of bias, lack of statistical significance in an underpowered trial, detecting clinically unimportant therapeutic differences, inconsistent results across trials, and trade-offs within the comparative effectiveness and/or the comparative safety considerations.

^b An adequate assessment of 'non-inferiority' is the preferred basis for demonstrating equivalence

Given the lack of any incremental outcome data, only a cost-analysis could be undertaken for the economic evaluation.

D.2. POPULATIONS AND SETTINGS

The proposed population is female patients aged 0 to 44 years who will be having, or have had, gonadotoxic treatment. This includes treatment for malignancy, as well as for precancerous or benign conditions.

AMH testing would be done in addition to the current standard tests, predominantly AFC and FSH combined with E2 (FSH+E2), to measure ovarian reserve. AMH is a pathology service and is expected that most claims for this test will be as an outpatient service.

D.3. STRUCTURE AND RATIONALE OF THE ECONOMIC EVALUATION

The clinical assessment does not provide adequate evidence to identify nor quantify the incremental effectiveness of AMH test as an additional test to the current practice. Therefore no outcomes are analysed in this economic evaluation. A cost-analysis of AMH testing as an addition to current practice compared to current practice is presented. The costs include those related to testing, including patient co-payments, and the cost of referrals for testing, where applicable.

A summary of the key characteristics of the economic evaluation is provided in Table 62.

Table 62 Summary of the economic evaluation

Perspective	Australian health care
Comparator	AFC, FSH and E2 tests
Type of economic evaluation	Cost-analysis
Sources of evidence	Systematic review
Outcomes	No health outcomes; cost per patient estimated only
Methods used to generate results	Investigative pathway cost-comparison
Software packages used	Microsoft Excel 2013

AFC = antral follicle count; E2 = estradiol; FSH = follicle-stimulating hormone

LITERATURE REVIEW

No studies were identified that conducted economic analysis comparing AMH in addition to other ovarian reserve tests with other ovarian reserve tests alone. Only one economic analysis was found reporting cost-effectiveness of customised FSH dosing based on AMH test results compared to standard FSH dosing before IVF/ICSI in the general population (van Tilborg et al. 2017). The study population, context and structure of this study are not relevant for the present report.

D.4. INPUTS TO THE ECONOMIC EVALUATION

TEST COSTS AND ASSOCIATED COSTS

The resource use and associated costs considered in the economic analysis are presented in Table 63.

For the AMH test, the proposed MBS item schedule fee of \$100 is used. For the proposed and comparator scenario, the other ovarian reserve test costs are based on the average provider fee, which takes into account bulk billing and patient contributions above the schedule fee. These are MBS items 55065 and 55067 for ultrasound and AFC, and MBS item 66701 for blood test measuring FSH and E2. MBS item 55067 is used for pelvic ultrasound performed with older equipment. MBS statistics for the average number of services performed for these items indicate that only 0.1% of ultrasounds were performed using older equipment (item 55067). Therefore only MBS item 55065 is included in the cost comparisons. In addition, pathology tests are routinely associated with the cost

of consultation by the referring doctor (MBS item 104) and a consult to review the test results (MBS item 105). The schedule fees for these items are \$85.55 and \$43 respectively.

Table 63 Various costs associated with ovarian reserve testing

Test	Base case	Source
AMH	\$100.00	Proposed fee (Section A.3)
AFC	\$117.02	MBS item 55065 (Average provider fee July 2011–June 2016) ¹
FSH and E2	\$49.72	MBS item 66701 (Average provider fee July 2011–June 2016) ¹
Specialist consult for referral	\$85.55	MBS item 104
Specialist consult to review results	\$43.00	MBS item 105

AMH = anti-Müllerian hormone; AFC = antral follicle count; E2 = estradiol; FSH = follicle-stimulating hormone; MBS = Medicare Benefits Schedule

¹Data provided by the Medical Benefits Division, Australian Government Department of Health.

While these resource items have been identified as associated with AMH testing, the proposal is that AMH be used **in addition** to ovarian reserve tests already in current use. Therefore, it is assumed that it will not require additional consultations for referral or review beyond those that already occurring in current practice, and there would be no net change in these associated costs with the proposed use of AMH testing.

D.5. RESULTS OF THE ECONOMIC EVALUATION

The overall costs and incremental costs as calculated for the proposed use of AMH and the comparator are shown in Table 64. The use of the proposed AMH test would be expected to increase the cost of current practice by \$100 per patient, i.e. only the proposed fee for the AMH test.

Table 64 Costs associated with testing ovarian reserve, and incremental cost per patient

	AMH + current practice ^a	Current practice
Specialist consultations for referral and review	\$128.55	\$128.55
Test costs	\$266.74	\$166.74
Total cost per patient	\$395.29	\$295.29
Incremental cost per patient		\$100.00

AMH = anti-Müllerian hormone; AFC = antral follicle count; E2 = estradiol; FSH = follicle-stimulating hormone

^a AFC and FSH+E2 measurements are considered as current practice for ovarian reserve testing and the intervention includes all these tests and AMH.

D.6. SENSITIVITY ANALYSES

The base case analysis assumes AMH is used as an additional test to the current practice, i.e. AFC and FSH+E2. However, alternative scenarios where AMH replaces one of the existing tests used to estimate ovarian reserve are costed. These are presented in Table 65.

Table 65 Scenario analyses: AMH is used as replacement test to either AFC, or FSH+E2

Scenario:	Current practice^a	AMH replacing AFC	AMH replacing FSH+E2
Test combinations	AFC and FSH+E2	AMH and FSH+E2	AMH and AFC
Specialist consultations	\$128.55	\$128.55	\$128.55
Test costs	\$166.74	\$149.72	\$217.02
Total cost per patient	\$295.29	\$278.27	\$345.57
Incremental cost per patient compared with current practice	-	-\$17.02	\$50.28

AMH = anti-Müllerian hormone; AFC = antral follicle count; E2 = estradiol; FSH = follicle-stimulating hormone

^a AFC and FSH+E2 measurements are considered as current practice for ovarian reserve testing.

If AMH were to replace AFC in current practice, it would result in a cost saving of \$17 per patient. However, if AMH were used to replace FSH+E2, i.e. proposed intervention being the use of AMH and AFC, there would still be a net increase in costs of approximately \$50 per patient.

SECTION E

FINANCIAL IMPLICATIONS

E.1. JUSTIFICATION OF THE SELECTION OF SOURCES OF DATA

An epidemiological approach has been used to estimate the financial implications of the proposed MBS funding of AMH testing to inform fertility management in female patients preceding or following gonadotoxic treatment. AMH testing is currently performed in this population using private funding.

The sources for data used in the financial analysis are presented in Table 66.

Table 66 Data sources used in the financial analysis

Data	Source
Cost of AMH to the MBS	85% of the proposed schedule fee, assuming that tests are performed in an outpatient setting, consistent with the setting for other blood tests (MBS data for items 66695, 66701 and 66707, 2011–12 to 2015–16)
Patient co-payment for AMH service	15% of the proposed schedule fee
Number of incident cases of cancer	AIHW cancer books(Australian Institute of Health and Welfare 2017)
Proportion of cancer patients undergoing gonadotoxic treatment	50% (Abdallah et al. 2017; Oktay & Sonmezer 2008)
Proportion of patients with non-malignant conditions considering fertility preservation or ovarian function monitoring	22% (Pacheco & Oktay 2017)
Cancer patients undergoing fertility consultation (referral rate)	59% (Logan et al. 2017)
AMH uptake rate, current	53% (in female oncology patients, aged 0–45 years, who attend fertility clinic consultation), based on local data provided by clinical expert ⁶
AMH uptake rate, projected	90% (Tobler et al. 2015)
Average number of tests per patient	3, data provided by clinical expert ⁷

AIHW = Australian Institute of Health and Welfare; AMH = anti-Müllerian hormone; MBS = Medicare Benefits Schedule

E.2. USE AND COSTS OF AMH TESTING

ESTIMATED USE OF AMH TESTING

The MBS funded AMH test is proposed as an additional test to AFC, FSH and E2 prior to or following gonadotoxic treatment. While other tests of ovarian reserve (AFC, FSH and E2) are currently funded by the MBS, they are not restricted to the proposed population. Therefore, an epidemiological approach, combined with uptake estimates, has been used to project the number of services of AMH that would be MBS funded under the proposed restriction.

⁶ Personal communication with A/Prof Kate Stern, email received on 27th August 2017.

⁷ Personal communication with A/Prof Kate Stern, email received on 13th October 2017.

Cancer patients

Eligible population

Data on the number of incident cases of cancer in females aged under 45 years is sourced from Australian cancer data (Australian Institute of Health and Welfare 2017). This data shows an average growth rate of 0.36% in incident cases between 2004 and 2017. This growth rate is applied to project the number of incident cases of cancer in females aged 0 to 44 years over the next five years. The majority of cases of cancer in females under 44 years occur post-puberty. More than 95% of these cases are in females over the age of 15 years old.

It is anticipated that approximately 50% of women of reproductive age who are diagnosed with a malignant condition will undergo gonadotoxic treatment compromising their fertility (Abdallah et al. 2017; Oktay & Sonmezer 2008). Therefore, this number is applied to the incident cases of cancer (females 0–44 years) to estimate the number of patients potentially eligible for AMH testing. If the rate of gonadotoxic treatment in pre-pubescent cancer cases is different (an estimate was not identified) this will make little difference on the overall estimate of the eligible population.

Table 67 presents the number of incident cases of cancer (females 0–44 years) and number of patients eligible for AMH testing based on the incident estimates for 2017 (Australian Institute of Health and Welfare 2017).

Table 67 Incident cases of cancer (females 0–44 years) estimated to undergo gonadotoxic treatment (2017)

Age group (years)	Number of incident cases of cancer	% of total
0–14	322	4.9
15–24	432	6.6
25–39	3,320	50.9
40–44	2,446	37.5
Total (0–44)	6,520	100
% of patients undergoing gonadotoxic treatment	50%	
Number of patients eligible for AMH testing	3,260^a	

AMH = anti-Müllerian hormone

^a 6,520 * 50%

Source: Chapter 3; Table A3.1: Estimated incidence rates of all cancers combined, by age at diagnosis and sex, 2017 (Australian Institute of Health and Welfare 2017).

Proportion of female cancer patients referred to fertility specialists

There is little evidence on the proportion of oncology patients being specifically referred for oncofertility counselling. A proportion of patients who are too young, or do not desire fertility preservation due to advanced age or having had completed their families, may not want AMH tests. A recently published systematic review on oncofertility support needs for cancer patients of a reproductive age (14–45 years) reported that the proportion of women referred to a fertility specialist or service ranged from 14% to 67% between 2012 and 2016. In the case where 67% of the newly diagnosed breast cancer patients were referred for fertility preservation, 59% actually attended a fertility consultation (Logan et al. 2017).

In the base case analysis, it is assumed that around 59% of the oncology patients would attend counselling for fertility preservation. Referral rates of 14% and 67% are assessed in sensitivity analysis.

Non-cancer patients

In addition to malignant conditions, there are number of non-malignant indications that may require gonadotoxic treatment. Data reporting this proportion in Australian settings could not be identified. Demeestere et al reported that non-oncological conditions represented nearly 20% of the indications in the studied population of patients requiring fertility preservation (Demeestere et al. 2009). A recent meta-analysis conducted to determine cohort epidemiological characteristics and success rates of autologous ovarian tissue transplantation reported that approximately 78% of the women undergoing fertility preservation due to gonadotoxic treatment had malignant conditions and 22% had non-malignant indications (Pacheco & Oktay 2017). Based on the proportion reported in this meta-analysis, it is assumed that the malignant indications will represent 78% of the target population and 22% of the patients will have non-malignant indications. Using this approach approximately 544 patients with non-malignant indications undergoing gonadotoxic treatment would have been counselled for fertility preservation in addition to 1,930 oncology patients in 2017.

In non-academic literature available to the public, the AMH test is known as the 'Egg-Timer Test' and is promoted as providing useful information to women who are delaying pregnancy (Monash IVF 2017). Also, Tobler et al in 2015 (Tobler et al. 2015) identified that 91% of surveyed Australian infertility clinics considered an AMH test appropriate for first-line management of infertility. Therefore there is potential for considerable leakage beyond those patients who have received gonadotoxic treatment to a broader population. This is examined in the sensitivity analyses in Section E.6.

Uptake of AMH testing

Data provided by a clinical expert indicated that nearly 53% of female oncology patients aged 0 to 44 years who sought fertility specialist advice, underwent AMH testing at least once over the past six years (2011–2017). While some patients are considered too young, other reasons why the test was not done included a lack of availability (in the first few years of data collection) and the expense.⁸ Therefore, it is assumed that if the AMH test is MBS funded, the uptake will increase. The data by Tobler et al, indicating that over 90% of Australian IVF clinics considered AMH a first-line and relevant or extremely relevant test, and that it would be used routinely if available at no cost (Tobler et al. 2015). Based on this data, it is estimated that in patients who receive fertility advice, the uptake of AMH testing would increase to 90%. Uptake rates of 53% and 100% are assessed in the sensitivity analysis (see Section E.6).

⁸ Personal communication with A/Prof Kate Stern, email received on 27th August 2017.

Number of AMH tests per patient

AMH testing is also proposed for the assessment of ovarian reserve, at a maximum of one AMH test per year, following gonadotoxic treatment to assess the gonadotoxic effects of treatment, pubertal delay, ovarian failure, and/or the need for fertility preservation following treatment, and to assess the need for assisted reproductive treatment for family planning. Not all women in the target population will undergo AMH testing annually and many will consider it only when they start planning to have children. Data on the current usage of AMH tests for both pre-gonadotoxic treatment and post-gonadotoxic treatment in Australia are lacking.

Clinical expert advice was that three AMH tests per patient, on average, may be performed.⁹ To estimate the overall number of AMH tests performed annually, both prior to and post-gonadotoxic treatment, the number of patients undergoing gonadotoxic treatment for malignant and non-malignant indications expected to uptake AMH tests, was multiplied by three, the average number of tests per patient. This approach is uncertain, but is conservative as AMH tests performed clinically may be less than estimated.

Table 68 shows the steps taken to project the total number of AMH tests that will be performed annually over the first five years of the listing.

Table 68 Estimate of number of AMH tests performed, 2018–19 to 2022–23

Row ^a	Description	2017–18	2018–19	2019–20	2020–21	2021–22	2022–23
A	Projected incident cases of cancer in females 0–44 years ^b	6,520	6,543	6,567	6,591	6,614	6,638
B	Proportion undergoing gonadotoxic treatment (B = A * 50%)	3,260	3,272	3,284	3,295	3,307	3,319
C	Proportion of female cancer patients referred/counselled for fertility preservation (C = B * 59%)	1,923	1,930	1,937	1,944	1,951	1,958
D	Number of patients with non-malignant indications (D = C*22% / 78%)	542	544	546	548	550	552
E	Total number of patients counselled for fertility preservation prior to gonadotoxic treatment (E = C + D)	2,466	2,475	2,484	2,493	2,502	2,511
F	Uptake of AMH testing in counselled patients prior to gonadotoxic treatment (F = E * 90%)	1,307 ^c	2,227	2,235	2,243	2,251	2,260
G	Estimated number of AMH tests performed in a year	3,921	6,682	6,706	6,730	6,754	6,779

⁹ Personal communication with A/Prof Kate Stern, email received on 13th October 2017.

Row ^a	Description	2017–18	2018–19	2019–20	2020–21	2021–22	2022–23
	(G = F *3)						

AMH = anti-Müllerian hormone

^a Row numbers are used to represent the calculations.

^b Incident cases of cancer are projected by applying 0.36% growth rate on 2017 incident estimates.

^c Uptake rate in 2017 is assumed to be 53% and for the next five years of listing it is assumed to be 90%.

ESTIMATED COSTS OF AMH TESTING

The applicant has proposed the scheduled fee for an AMH test is \$100. This is higher than the fee currently charged by some providers offering the service to privately funded patients (Table 69).

Table 69 Proposed or advertised fees^a for AMH testing

Provider	Cost	Source
Applicant	\$100	Proposed fee in the protocol ^b
IVF Australia	\$80	https://www.ivf.com.au/ovarian-reserve-amh-test
Fertility North	\$55	http://www.fertilitynorth.com.au/amh-test/
Repromed	\$98	http://repromed.com.au/what-to-expect/preliminary-investigations/amh-blood-test/
Clinpath Pathology	\$60	http://www.clinpathpathology.com.au/media/96085/anti-mullerian%20hormone%20(amh).pdf
i-screen	\$85	https://www.i-screen.com.au/app/register/amh-test

AMH = anti-Müllerian hormone

^a Advertised fees, as per websites accessed on 2 November 2017. The applicant has proposed the scheduled fee for an AMH test is \$100.

^b MSAC application 1434 (Protocol): Anti-Müllerian Hormone (AMH) MBS listing for female patients preceding or following gonadotoxic treatment, Australian Government Department of Health, Canberra.

The MBS usage data¹⁰ on the other additional blood tests likely taken in conjunction with AMH testing (MBS item 66701 for three blood tests described in item 66695) for the years 2011–12 to 2015–16 indicates that the services are performed in outpatient settings 99% of the time, and 95% of these services are bulk-billed. Therefore, it is assumed that the AMH test would also generally be performed in the outpatient setting and have a similar bulk billing rate. When not bulk-billed, the average patient contribution per service for MBS item 66701 was \$19, which is higher than the 15% patient contribution gap, which would be \$9. In the base case financial analysis, the patient contribution is assumed to be 15% of the scheduled fee, but higher patient contributions and different bulk billing rates are assessed in the sensitivity analysis.

Table 70 summarises the estimated costs of AMH testing to the MBS, i.e. government expenditure, and patients, i.e. out-of-pocket expenditure. The base case analysis estimates that AMH testing will cost approximately \$570,000 to the MBS each year.

Table 70 Estimated costs of AMH testing, 2018–19 to 2022–23

	2018–2019	2019–2020	2020–2021	2021–2022	2022–2023

¹⁰ Data provided by Australian Government Department of Health.

	2018–2019	2019–2020	2020–2021	2021–2022	2022–2023
Projected number of AMH tests	6,682	6,706	6,730	6,754	6,779
Cost of AMH to the MBS	\$567,961	\$570,005	\$572,057	\$574,117	\$576,184
Cost of AMH to the patients	\$100,228	\$100,589	\$100,951	\$101,315	\$101,679
Total cost of AMH test	\$668,189	\$670,595	\$673,009	\$675,432	\$677,863

AMH = anti-Müllerian hormone; MBS = Medicare Benefits Schedule

E.3. CHANGES IN USE AND COST OF OTHER MEDICAL SERVICES

The AMH test is proposed as an additional test to AFC and FSH+E2. Therefore there are no cost-offsets, nor additional costs expected to be associated with AMH expenditure.

E.4. FINANCIAL IMPLICATIONS FOR THE MBS

The total financial implications to the MBS resulting from the proposed listing of AMH testing are as estimated in Table 70.

E.5. FINANCIAL IMPLICATIONS FOR GOVERNMENT HEALTH BUDGETS

No other financial implications of AMH testing for other health budgets have been identified.

E.6. IDENTIFICATION, ESTIMATION AND REDUCTION OF UNCERTAINTY

In the base case analysis there are a number of uncertainties in the estimates and alternative scenarios to the base case analysis that may occur in clinical practice, should the AMH test be listed on the MBS.

ALTERNATIVE REFERRAL AND UPTAKE RATES

Table 71 presents sensitivity analyses around referral and uptake rates in the financial model.

Table 71 Financial implications of listing the AMH test on the MBS, sensitivity analysis varying referral and uptake rates

	Year 1	Year 2	Year 3	Year 4	Year 5
<i>Base case</i>					
Cost of AMH testing to the MBS	\$567,961	\$570,005	\$572,057	\$574,117	\$576,184
<i>Referral rate for fertility preservation: 14% (base case 59%)</i>					
Cost of AMH testing to the MBS	\$134,770	\$135,256	\$135,742	\$136,231	\$136,722
<i>Referral rate for fertility preservation: 67% (base case 59%)</i>					
Cost of AMH testing to the MBS	\$644,972	\$647,294	\$649,625	\$651,963	\$654,310
<i>Uptake of AMH testing: 53% (base case 90%)</i>					
Cost of AMH testing to the MBS	\$334,466	\$335,670	\$336,878	\$338,091	\$339,308
<i>Uptake of AMH testing: 100% (base case 90%)</i>					
Cost of AMH testing to the MBS	\$631,068	\$633,339	\$635,619	\$637,908	\$640,204

REDUCED USE IN YOUNGER AND OLDER WOMEN

If AMH testing is assumed to only occur in women aged between 25 and 39 years, rather than the base case including eligible females aged 0 to 44 years, then the estimated cost to the MBS decreases, as shown in Table 72.

LEAKAGE

Although not the intention of the listing, given the large number of women seeking fertility information and treatment, and the broadly available public information about the test, there is some risk that AMH tests could be provided under the MBS to females who do not meet the restriction criteria. In 2013, 16,357 women were identified as having their first ever fresh autologous IVF cycle, many other women per year would also seek fertility advice (Fitzgerald et al. 2017).

The financial impact of leakage of 15,000 additional AMH tests per year (acknowledged to be in the extreme upper range of potential leakage), is examined in the scenario analysis presented in Table 72, below.

Table 72 Financial implications of MBS listing the AMH test: scenario analysis of leakage to other indications

	Year 1	Year 2	Year 3	Year 4	Year 5
<i>Base case</i>					
Cost of AMH testing to the MBS	\$567,961	\$570,005	\$572,057	\$574,117	\$576,184
<i>If AMH is only used in women aged 25–39 years (base case 0–44 years)</i>					
Cost of AMH testing to the MBS	\$289,207	\$290,248	\$291,293	\$292,342	\$293,394
<i>If AMH is used beyond the proposed restriction in other women seeking fertility advice or treatment (15,000 per year)</i>					
Cost of AMH testing to the MBS	\$1,874,619	\$1,881,368	\$1,888,141	\$1,894,938	\$1,901,760

AMH = anti-Müllerian hormone; MBS = Medicare Benefits Schedule

AMH AS A REPLACEMENT TEST

While the application proposes that AMH test is used as an additional test to AFC and FSH+E2, if the AMH test were to replace any of the tests being performed currently, the following cost-offsets per test replaced would apply to MBS expenditure (Table 73). This would result in the net **expenditure to the MBS** being reduced by approximately \$68,000 per year, in the case of replacing AFC (Table 74), or an increase of \$310,00 per year, if FSH+E2 tests are not referred (Table 75).

Table 73 Scenario analyses, potential cost-offsets per test to MBS expenditure, if AMH is a replacement test

Tests	Cost-offsets	Source
AFC	\$95	Average benefits paid for MBS item 55065, July 2011–June 2016 ¹
FSH+E2	\$49	Average benefits paid for MBS item 66701, July 2011–June 2016 ¹

AMH = anti-Müllerian hormone; AFC = antral follicle count; E2 = estradiol; FSH = follicle-stimulating hormone; MBS = Medicare Benefits Schedule

¹ Data provided by the Medical Benefits Division, Australian Government Department of Health.

Table 74 Scenario analyses, potential cost-offsets to MBS expenditure, if AMH is a replacement test for AFC

	Year 1	Year 2	Year 3	Year 4	Year 5
Projected number of AMH tests	6,682	6,706	6,730	6,754	6,779
Cost of AMH testing to MBS	\$567,961	\$570,005	\$572,057	\$574,117	\$576,184
Cost of AFC test to MBS	\$634,780	\$637,065	\$639,358	\$641,660	\$643,970
Net costs to MBS if AMH test replaces AFC	-\$66,819	-\$67,059	-\$67,301	-\$67,543	-\$67,786

AMH = anti-Müllerian hormone; AFC = antral follicle count; MBS = Medicare Benefits Schedule

Table 75 Scenario analyses, potential cost-offsets to MBS, if AMH is a replacement test for FSH+E2

	Year 1	Year 2	Year 3	Year 4	Year 5
Projected number of AMH tests	6,682	6,706	6,730	6,754	6,779
Cost of AMH testing to MBS	\$567,961	\$570,005	\$572,057	\$574,117	\$576,184
Cost of FSH+E2 tests to MBS	\$327,413	\$328,591	\$329,774	\$330,961	\$332,153
Net costs to MBS if AMH test replaces FSH+E2	\$307,367	\$308,474	\$309,584	\$310,699	\$311,817

AMH = anti-Müllerian hormone; E2 = estradiol; FSH = follicle-stimulating hormone; MBS = Medicare Benefits Schedule

Patients should be counselled before receiving the AMH test, as it is difficult to interpret the results of the test. Girls and women who receive a test result indicating a low AMH level could become unnecessarily anxious about their fertility without the appropriate professional advice.

One abstract was identified investigating the psychological impact of ovarian reserve testing (O'Brien, Wingfield & Kelleher 2016). This study was conducted in Ireland. Interview analysis showed that in women who attended a fertility clinic and underwent AMH testing, the overall awareness of the clinical relevance of the test was low. Their main source of information was the internet, rather than their medical practitioner. Their feelings about the test were mostly determined by the test result. It was reported that women with a low AMH level had feelings of isolation, loss of femininity and purpose, and devastation. Women with a normal test result felt reassured and surprised that their result was normal. In general, regardless of their test results, women indicated that knowledge about their ovarian reserve was important and that it impacted their decision-making regarding childbearing intentions. No evidence was found on the psychological impact of AMH testing in women undergoing gonadotoxic treatment. It is not known whether the psychological impact in the target population will be similar, as the current priorities of women visiting a fertility specialist are likely to be quite different compared to those women with a potentially serious disease requiring gonadotoxic treatment for whom biological children are not wanted, at least in the near future.

APPENDIX A CLINICAL EXPERTS AND ASSESSMENT GROUP

CLINICAL EXPERTS

<u>Name</u>	<u>Expertise or affiliation</u>
Dr. Antoinette Anazodo	Paediatric and adolescent oncologist; Director, Sydney Youth Cancer Service; NSW Lead, Youth Cancer in NSW and ACT
Assoc. Prof. Cathryn Stern	Head, Fertility Preservation Service, Royal Women's Hospital and Melbourne IVF
Prof. William Ledger	Senior fertility specialist and gynaecologist, IVF Australia; Head and Professor of Obstetrics and Gynaecology, Royal Hospital for Women; Head, Discipline of Obstetrics and Gynaecology, University of New South Wales
Dr. David Molloy	Gynaecologist and fertility specialist; Clinical Director, Queensland Fertility Group

ASSESSMENT GROUP

<u>Name</u>	<u>Position</u>
Sharon Kessels	Senior Research Officer
Ruchi Mittal	Health Economist
Judy Morona	Senior Research Officer
Skye Newton	Team Leader (Medical HTA)
Camille Schubert	Team Leader (Economics)
Jo Milverton	Research Officer
Tracy Merlin	Managing Director

Noted conflicts of interest

There were no conflicts of interest.

APPENDIX B

SEARCH STRATEGIES

BIBLIOGRAPHIC DATABASES

Electronic bibliographic databases were searched studies meeting the inclusion criteria to address each of the research questions developed for this MSAC assessment. These databases are described in Table 76.

Table 76 Bibliographic databases

Electronic database	Time period
Cochrane Library, including Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, the Cochrane Central Register of Controlled Trials (CENTRAL), the Health Technology Assessment Database, the NHS Economic Evaluation Database	Inception–6/2017
CINAHL	Inception–6/2017
Current Contents	Inception–6/2017
Embase (including Embase and Medline)	Inception–6/2017
PubMed	Inception–6/2017
Web of Science - Science Citation Index Expanded	Inception–6/2017
PsycINFO (for ethical issues only)	Inception–6/2017

Additional literature, including peer-reviewed or grey literature, was sought from the sources outlined in Table 77, and from the health technology assessment agency websites provided in Table 78. Websites of specialty organisations were also searched for any potentially relevant information (Table 79).

Table 77 Additional sources of literature

Source	Location
--------	----------

Source	Location
Internet	
NHMRC- National Health and Medical Research Council (Australia)	http://www.nhmrc.gov.au/
US Department of Health and Human Services (reports and publications)	http://www.hhs.gov/
New York Academy of Medicine Grey Literature Report	http://www.greylit.org/
Trip Database	http://www.tripdatabase.com
Current Controlled Trials metaRegister	http://controlled-trials.com/
National Library of Medicine Health Services/Technology Assessment Text	http://text.nlm.nih.gov/
U.K. National Research Register	http://www.nihr.ac.uk/Pages/NRRArchive.aspx
Google Scholar	http://scholar.google.com/
Australian and New Zealand Clinical Trials Registry	www.anzctr.org.au
World Health Organization	http://www.who.int/en/
Pearling	
All included articles will have their reference lists searched for additional relevant source material	

Table 78 Specialty websites

Clinical Oncology Society of Australia	https://www.cosa.org.au/
American Society of Clinical Oncology	https://www.asco.org/
European Society for Medical Oncology	http://www.esmo.org/

Table 79 HTA websites

INTERNATIONAL	
International Network of Agencies for Health Technology Assessment	http://www.inahta.org/
AUSTRALIA	
Australian Safety and Efficacy Register of New Interventional Procedures-Surgical (ASERNIP-S)	http://www.surgeons.org/for-health-professionals/audits-and-surgical-research/asernip-s/
Centre for Clinical Effectiveness, Monash University	http://www.monashhealth.org/page/Health_Professionals/CCE/
Centre for Health Economics, Monash University	http://www.buseco.monash.edu.au/centres/che/
AUSTRIA	
Institute of Technology Assessment / HTA unit	http://www.oew.ac.at/ita
CANADA	
Institut National d'Excellence en Santé et en Services Sociaux (INESSS)	http://www.inesss.qc.ca/en/publications/publications/
Alberta Heritage Foundation for Medical Research (AHFMR)	http://www.ahfmr.ab.ca/publications.html
Alberta Institute of Health Economics	http://www.ihe.ca/
The Canadian Agency for Drugs And Technologies in Health (CADTH)	http://www.cadth.ca/index.php/en/
The Canadian Association for Health Services and Policy Research (CAHSPR)	http://www.cahspr.ca/
Centre for Health Economics and Policy Analysis,	http://www.cheqa.org

McMaster University	
Health Utilities Index, McMaster University Centre for Health Services and Policy Research, University of British Columbia	http://www.ices.mcmaster.ca/hug/index.htm
Institute for Clinical and Evaluative Studies	http://www.ices.on.ca
Saskatchewan Health Quality Council (Canada)	http://www.hqc.sk.ca
DENMARK	
Danish National Institute Of Public Health	http://www.si-folkesundhed.dk/?lang=en
FINLAND	
Finnish National Institute for Health and Welfare	http://www.thl.fi/en/web/thlfi-en/
FRANCE	
L'Agence Nationale d'Accréditation et d'Evaluation en Santé (ANAES)	http://www.anaes.fr/
GERMANY	
German Institute for Medical Documentation and Information (DIMDI) / HTA	http://www.dimdi.de/static/en/index.html
Institute for Quality and Efficiency in Health Care (IQWiG)	http://www.iqwig.de
THE NETHERLANDS	
Health Council of the Netherlands Gezondheidsraad	http://www.gezondheidsraad.nl/en/
Institute for Medical Technology Assessment (Netherlands)	http://www.imta.nl/
NEW ZEALAND	
New Zealand Health Technology Assessment (NZHTA)	http://www.otago.ac.nz/christchurch/research/nzhta/
NORWAY	
Norwegian Knowledge Centre for the Health Services	http://www.kunnskapsenteret.no
SPAIN	
Agencia de Evaluación de Tecnologías Sanitarias, Instituto de Salud "Carlos III"/Health Technology Assessment Agency (AETS)	http://www.isciii.es/
Andalusian Agency for Health Technology Assessment (Spain)	http://www.juntadeandalucia.es/
Catalan Agency for Health Technology Assessment (CAHTA)	http://www.gencat.cat
SWEDEN	
Center for Medical Technology Assessment, Linköping University	http://www.cmt.liu.se/?l=en&sc=true
Swedish Council on Technology Assessment in Health Care (SBU)	http://www.sbu.se/en/
SWITZERLAND	
Swiss Network on Health Technology Assessment (SNHTA)	http://www.snhta.ch/
UNITED KINGDOM	
National institute for Health Research, Health Technology Assessment Programme	http://www.hta.ac.uk/
NHS Quality Improvement Scotland	http://www.nhshealthquality.org/
National Institute for Clinical Excellence (NICE)	http://www.nice.org.uk/

The European International Network on New and Changing Health Technologies	http://www.euroscan.bham.ac.uk/
University of York NHS Centre for Reviews and Dissemination (NHS CRD)	http://www.york.ac.uk/inst/crd/
UNITED STATES	
Agency for Healthcare Research and Quality	http://www.ahrq.gov/clinic/techix.htm
Harvard School of Public Health	http://www.hsph.harvard.edu/
Institute for Clinical and Economic Review (ICER)	http://www.icer-review.org/
Institute for Clinical Systems Improvement	http://www.icsi.org
Minnesota Department of Health (US)	http://www.health.state.mn.us/
National Information Centre of Health Services Research and Health Care Technology (US)	http://www.nlm.nih.gov/nichsr/nichsr.html
Oregon Health Resources Commission (US)	http://www.oregon.gov/oha/OHPR/HRC/Pages/index.aspx
Office of Health Technology Assessment Archive (US)	http://ota.fas.org/
U.S. Blue Cross/ Blue Shield Association Technology Evaluation Center (Tec)	http://www.bcbs.com/blueresources/tec/
Veteran's Affairs Research and Development Technology Assessment Program (US)	http://www.research.va.gov/default.cfm

SEARCH TERMS

Search terms used in the literature searches are shown below (Table 80 and Table 81).

Table 80 Suggested search terms for AMH testing to measure ovarian reserve to determine the requirement for cryopreservation of ovarian tissue or oocytes prior to or following gonadotoxic treatment, PubMed

Element of clinical question	PubMed/Medline search terms
Population	<p>"Neoplasms" [MeSH] OR cancer OR tumour OR tumor OR neoplasm OR neoplastic OR "Inflammatory bowel disease" OR "ulcerative colitis" OR "Crohn's disease" OR "rheumatoid arthritis" OR "asplenic anemia" OR vasculitis OR lupus OR "metabolic disease" OR "endometriosis" OR "Antineoplastic Agents" [MeSH] OR "Radiotherapy"[MeSH] OR "anti-mitotic drugs" OR irradiation OR gonadotoxic*</p> <p>AND</p> <p>"Ovarian Follicle" [MeSH] OR oocyte* OR oogonium OR ovarian OR follicle* OR gonadal OR egg* OR adolescent* OR paediatric OR young</p>
Intervention	<p>"Ovarian Reserve" [MeSH] OR "anti-Müllerian hormone" OR "antimullerian hormone" OR AMH</p> <p>OR</p> <p>"Cryopreservation"[MeSH] OR "Fertility Preservation" [MeSH] OR "Oocyte Retrieval" [MeSH] OR cryopreserv* OR "fertility preservation" OR "fertility restoration" OR "vitrification" OR oncofertility</p>
Comparator (if applicable)	-
Outcomes (if applicable)	-
Limits	Humans

MeSH = Medical Subject Heading, based on a Medline/PubMed platform

Table 81 Suggested search terms for AMH testing to measure ovarian reserve to determine the requirement for cryopreservation of ovarian tissue or oocytes prior to or following gonadotoxic treatment, Embase

Element of clinical question	Embase search terms
Population	neoplasm'/exp OR cancer OR tumour OR tumor OR neoplastic OR 'Inflammatory bowel disease' OR 'ulcerative colitis' OR 'Crohns disease' OR 'rheumatoid arthritis' OR 'asplenic anemia' OR vasculitis OR lupus OR 'metabolic disease' OR 'endometriosis' OR 'antineoplastic agent'/exp OR 'antimitotic agent'/exp OR 'radiotherapy'/exp OR irradiation OR 'gonadotoxicity'/exp OR gonadotoxic AND 'ovary follicle cell'/exp OR oocyte OR oogonium OR ovarian OR follicle* OR gonadal OR adolescent* OR egg OR eggs OR paediatric OR pediatric OR young
Intervention	'ovarian reserve'/exp OR 'anti-Müllerian hormone' OR 'antimullerian hormone' OR AMH OR 'fertility preservation'/exp OR 'oocyte retrieval'/exp OR 'cryopreservation'/exp OR 'fertility restoration' OR 'vitrification' OR 'follicular aspiration'/exp OR oncofertility
Comparator (if applicable)	-
Outcomes (if applicable)	-
Limits	Humans, exp terms also as free text

APPENDIX C

STUDIES INCLUDED IN THE SYSTEMATIC REVIEW

Table 82 Profiles of studies on analytical validity studies providing 2x2 or ROC data included in the systematic literature review

Authors Publication year Location	Study design Level of evidence ^a Risk of bias ^b	Study population characteristics	Intervention	Comparator	Reference standard	Relevant outcomes assessed	Measurement of outcomes and methods of analysis
Decanter et al. (2014) France	Prospective case series Level III-2 diagnostic accuracy evidence Quality: Low risk of bias	N=58 selected serum samples stored at -80°C collected from 3 to 24 months after the end of chemotherapy in 30 women regardless of their menstrual status. n=17 lymphoma n=13 early breast cancer The distribution of the samples was as follows: 3 months: 9; 6 months: 13; 9 months: 11; 12 months: 18; 18 months: 2; and 24 months after the end of chemotherapy: 5	<u>EIA AMH/MIS assay</u> The dynamic range of the standard curve was 3–150 pmol/L. Between-run reproducibility determined by measuring three quality control samples (mean concentrations, 9.8, 18.6, and 37.1 pmol/L) in duplicate. The CoVs were 14%, 13%, and 12.6% for the three levels assessed. LoQ corresponding to lowest concentration measurable with acceptable performance was determined from the precision profile curve = 2.5 pmol/L. Results under the first calibrator value (i.e., 3.0 pmol/L) expressed as undetectable.	<u>Pico-AMH ELISA</u> Hypersensitive ELISA assay standardised against recombinant human AMH. Dynamic range of the standard curve was 0.07–6.5 pmol/L. The between-run reproducibility was determined by duplicate measuring two quality control samples (mean concentrations, 0.69 and 2.0 pmol/L). The CoVs were 1.38% and 3.84% for the two levels assessed. The LoQ that corresponds to the lowest concentration that can be measured with an acceptable performance was determined from the precision profile curve = 0.034 pmol/L. Results under the first calibrator value (i.e. <0.07) expressed as undetectable.	Normal menstruation: Samples were selected to constitute two equally sized groups, according to the menstrual status of the patients at the time of sampling: n=30 amenorrhea n=28 spontaneous resumption of menses	Sensitivity specificity LR+ LR-	AMH levels 2 × 2 data for AMH detectable/not detectable vs amenorrhea/normal cycle
Miyoshi et al. (2013) Japan	Case series Level III-2 diagnostic	N=53 female patients who had survived >2 years after childhood cancer treatment	<u>EIA AMH/MIS ELISA</u> The lower and upper limits of detection for AMH were 1 and	<u>Access FSH</u> , a chemiluminescent enzyme immunoassay Sensitivity of 0.2 mIU/mL. The	NA	Sensitivity specificity	AMH levels FSH levels

Authors Publication year Location	Study design Level of evidence ^a Risk of bias ^b	Study population characteristics	Intervention	Comparator	Reference standard	Relevant outcomes assessed	Measurement of outcomes and methods of analysis
	accuracy evidence Quality: Low risk of bias	reviewed retrospectively n=21 solid tumours n=19 haematological n=13 brain tumour Median age at evaluation = 17.4 years (range 4.0–29.6) Median age at diagnosis of underlying disease = 6.3 years (range 0–12.9) Median follow-up duration from completion of therapy to hormonal evaluation = 8.8 years (range 2.3–26.1).	150 pmol/l, respectively. Concentrations were compared to normal values in healthy paediatric females, with a cut- off value below the 2.5th percentile being defined as low AMH. Undetectably low AMH levels (<1 pmol/l) were recorded as 1 pmol/l.	reference values of FSH for were as follows: 4.5–11.0 mIU/mL during the follicular phase, 3.6–20.6 mIU/mL during the ovulatory phase, 1.5–10.8 mIU/mL during the luteal phase, 2.0– 21.9 mIU/mL during perimenopause, and 21.5–159.0 mIU/mL during menopause. High FSH was defined as >10 mIU/mL prior to menarche and >20 mIU/mL for females under estrogen replacement therapy in this study.		LR+ LR-	2 × 2 data for low AMH vs low FSH
Su et al. (2011) USA	Case series Level III-2 diagnostic accuracy evidence Quality: Low risk of bias	N=56 female post- chemotherapy breast cancer survivors All subjects underwent cyclophosphamide-based chemotherapy regimens Median age at chemotherapy = 43.6 years (range 30–56). Median age at study assessment = 48 years (range 35–62)	<u>DSL ACTIVE® AMH/MIS ELISA</u> The lower limit of detection for AMH was 25 pg/mL, and the intra-assay CoV was 2%. For each measure of ovarian reserve, a cut-point was selected to optimize the positive predictive value for CRA.	<u>DSL inhibin B ELISA</u> The intra- and inter-assay CoV were 7.9% and 8.4%, respectively. The lower limit of detection was 5 pg/mL. E2 and FSH were measured by radioimmunoassay using Coat-A- Count commercial kits. The intra- and inter-assay CoV were less than 5%. Values below detection thresholds were given half of the threshold value in analyses.	CRA Self-reported menstrual pattern and pelvic ultrasound	Sensitivity specificity PPV NPV	AMH levels FSH levels Inhibin B levels E2 levels AUC data

AFC = antral follicle count; AMH = anti-Müllerian hormone; AUC = area under the curve; CRA = chemotherapy-related ovarian failure; CoV = coefficient(s) of variation; E2 = estradiol; EIA = enzyme immunoassay; ELISA = enzyme-linked immunosorbent assay; FSH = follicle-stimulating hormone; LR = likelihood ratio; LoQ = limit of quantification; MII oocytes = mature, metaphase II oocytes; MIS = Müllerian-inhibiting substance; NPV = negative predictive value; PPV = positive predictive value; ROC = receiver operator characteristic

^a Source: See NHMRC hierarchy of evidence

^b Risk of bias as it relates to primary outcomes of the systematic review

Table 83 Profiles of studies on diagnostic concordance between different AMH tests included in the systematic literature review

Authors Publication Year Location	Study design Level of evidence ^a Risk of bias ^b	Study population characteristics	AMH tests	Measurement of relevant outcomes
de Souza et al. (2015) Brazil	Case series Quality: Low risk of bias	N=8 8 female polymyositis patients who are or had treatment with immunosuppressive agents and corticosteroids Mean age = 31.4 ± 6.5 years N=16 healthy volunteer age-matched women Mean age = 30.7 ± 6.2 years	Ovarian function was assessed by determining serum hormone levels during the early follicular phase of the menstrual cycle and was blinded to the other parameters of ovarian function. AMH levels were evaluated in duplicate samples by two different AMH tests AMH Gen II ELISA Intra- and inter-assay CoVs were limited to 5.4% and 5.6 %, respectively. Ultrasensitive AMH ELISA Intra- and inter-assay CoVs were limited to 3.1 and 2.7 %, respectively. AMH levels less than 1.0 ng/mL were regarded as low as suggested by the manufacturers of both ELISA tests.	Correlation: Spearman's rank correlation test
Su et al. (2014) USA	Case series Quality: Low risk of bias	N=90 newly diagnosed breast cancer patients prior to cancer treatment Mean age = 38.3 years No periods in past year: n=2 had 1–3 n=4 had 4–9 n=79 had ≥10	Due to urgency in starting chemotherapy, enrolment blood specimens were drawn across the menstrual cycle and not timed to the early follicular phase. Serum was frozen in aliquots and stored at -80°C until assayed for AMH. AMH Gen II ELISA The standard curve range was 0.16–22.5 ng/mL and the levels of detection 0.08 ng/mL. The inter-assay CoVs were 5.6% and 4.5% at 4.42 and 14.0 ng/mL, respectively. Ultrasensitive AMH ELISA The standard curve range was 0.1–14 ng/mL and the levels of detection 0.07 ng/mL. The inter-assay CoVs were 4.6%, 4.8%, 2.0% at 0.346, 0.715 and 1.85 ng/mL, respectively. pico-AMH ELISA The standard curve range was 6–746 pg/mL and the levels of detection 0.01 ng/mL. The inter-assay CoVs were 4.5%, 2.2%, 3.8% at 22.6, 86.5 and 373 pg/mL, respectively.	Correlation: Pearson's bivariate correlation test

AMH = anti-Müllerian hormone; CoV = coefficient(s) of variation; EIA = enzyme Immunoassay; ELISA = enzyme-linked immunosorbent assay; MIS = Müllerian-inhibiting substance; PCOS = polycystic ovary syndrome

Table 84 Profiles of studies on diagnostic concordance between AMH and other tests included in the systematic literature review

Authors Publication year Location	Study design Level of evidence ^a Risk of bias ^b	Study population characteristics	Intervention	Comparator	Measurement of relevant outcomes
--	--	----------------------------------	--------------	------------	--

Authors Publication year Location	Study design Level of evidence ^a Risk of bias ^b	Study population characteristics	Intervention	Comparator	Measurement of relevant outcomes
Beneventi et al. (2014) Italy	Case series Quality: Low risk of bias	N=135 female survivors treated for childhood malignant and non-malignant diseases and patients who had received BMT for thalassaemia major or sickle cell anaemia. 92 (68.1%) diagnosed before menarche 43 (31.85%) diagnosed after menarche Median age at enrolment = 19 years (IQR 16–21) Median age at start of treatment = 10 years (IQR 6–16) Median time since treatment = 9 years (IQR 6–12).	Blood samples were collected in the early follicular phase in patients with spontaneous menstrual cycles, and in withdrawal bleeding in patients treated with oral contraceptives or HRT. <u>AMH Gen II ELISA</u> According to manufacturer's instructions.	<u>Inhibin B Gen II ELISA</u> According to manufacturer's instructions	Correlation: Spearman's rank correlation test
Biacchiardi et al. (2011) Italy	Case series Quality: Low risk of bias	N=43 normo-ovulatory women aged 18–42 years who were affected by one or more ovarian endometriomas Mean age = 34.2 ± 5.4 years n=7 ASRM stage 2 n=26 ASRM stage 3 n=10 ASRM stage 4	Blood samples were drawn during the early follicular phase (day 3) of the month in which surgery was scheduled. Sera were frozen at -20°C for subsequent centralised testing <u>EIA AMH/MIS assay</u> Intra-assay and inter-assay CoV of 12.3% and 14.2%, respectively.	<u>AFC TVUS</u> Transvaginal ultrasound examination was performed in the early follicular phase in order to estimate the AFC. AFC was determined by counting the follicles >3 mm diameter that were visible through a complete scanning of both ovaries. In order to avoid any operator-linked bias, all ultrasound examinations were done by the same investigators using the same equipment.	Correlation: Pearson's bivariate correlation test

Authors Publication year Location	Study design Level of evidence ^a Risk of bias ^b	Study population characteristics	Intervention	Comparator	Measurement of relevant outcomes
Decanter et al. (2014) France	Case series Quality: Low risk of bias	N=58 samples drawn at least 3 months after the end of chemotherapy in 30 women with either breast cancer (n=13) or haematological malignancies (n=17).	<p><u>pico-AMH ELISA</u></p> <p>The dynamic range of the standard curve was 0.07–6.5 pmol/L.</p> <p>The between-run CoVs were 1.38% and 3.84% at 0.69 and 2.0 pmol/L</p> <p>The lowest detectable concentration was 0.034 pmol/L.</p> <p>For this study, results, <0.07 were expressed as undetectable.</p> <p><u>EIA AMH/MIS assay</u></p> <p>The dynamic range of the standard curve was 3–150 pmol/L.</p> <p>The between-run CoVs were 14%, 13%, and 12.6% at 9.8, 18.6, and 37.1 pmol/L.</p> <p>The lowest detectable concentration was 2.5 pmol/L.</p> <p>For this study, results <3.0 pmol/L) were expressed as undetectable.</p>	<p><u>FSH test</u></p> <p>Not reported</p>	Correlation: Pearson's bivariate correlation test
Fabbri et al. (2014) Italy	Case series Quality: Low risk of bias	N=86 women with various non-gynaecological malignancies who underwent ovarian tissue cryopreservation.	<p><u>AMH Gen II ELISA</u></p> <p>The lowest detection limit of the assay was 0.08 ng/mL; intra- and inter-assay CoV were 5.4 and 5.6%, respectively, at a concentration of 4.42 ng/mL.</p>	<p><u>Density of primordial follicles from biopsy</u></p> <p>An ovarian biopsy (from both the right and left ovary) was collected by laparoscopy for ovarian tissue cryopreservation.</p> <p>For each sample, one 0.5-µm thick section out of every 30 was collected and stained with toluidine blue for light microscopic examination to identify and count the follicles.</p> <p><u>FSH Elecsys assay</u></p> <p><u>LH Elecsys assay</u></p> <p><u>E2 Elecsys assay</u></p>	Correlation: Spearman's rank correlation test
Kim et al.	Case series	N=32 pre-menopausal women with clinical stage III hormone receptor-positive	<u>USCN AMH ELISA</u>	<u>BlueGene inhibin B ELISA</u>	Correlation: Spearman's

Authors Publication year Location	Study design Level of evidence ^a Risk of bias ^b	Study population characteristics	Intervention	Comparator	Measurement of relevant outcomes
(2016) Korea	Quality: Low risk of bias	invasive ductal breast cancer treated by neoadjuvant chemotherapy FSH level ≤ 30 mIU/mL at the time of diagnosis, was considered pre-menopausal Median age = 41.5 years (range 27–50).	Standards or samples were added to the wells along with an AMH-specific, biotin-conjugated polyclonal antibody and avidin conjugated to horseradish peroxidase. After incubation, a substrate solution was added to the wells, resulting in a colorimetric reaction. The optical intensity, which is proportional to the amount of AMH bound in the initial step, was measured at 405 nm in a microplate reader; and AMH concentrations were extrapolated from standard curves.	Standards or samples were added to microtiter wells pre-coated with an inhibin B-specific monoclonal antibody. An HRP-conjugated polyclonal antibody specific for inhibin B was then added to the wells to sandwich the inhibin B bound to the immobilized monoclonal antibody. After incubation, the wells were thoroughly washed and a substrate solution was added. The antibody- substrate reaction was terminated by adding a sulfuric acid solution, and optical intensity was measured spectrophotometrically at 450 nm. The inhibin B concentration in each sample was extrapolated from a standard curve. <u>FSH Elecsys assay</u> The limit of detection was 0.100 mIU/mL. <u>E2 Elecsys assay</u> The limit of detection was 5.00 pg/mL.	rank correlation test
Lee et al. (2011) USA	Case series Quality: Low risk of bias	N=41 women with breast cancer before adjuvant treatment n=7 had oocyte freezing n=29 had embryo cryopreservation n=4 underwent both Mean age = 34.8 ± 4.7 years (range 24– 44).	<u>DSL ACTIVE® AMH ELISA</u>	<u>AFC TVUS</u> AFC was performed on menstrual cycle day 2 <u>Inhibin B assay</u> Not reported <u>FSH assay</u> Not reported	Correlation: Spearman's rank correlation test
Lutchman Singh et al. (2007) UK	Case series Quality: Low risk of bias	N=22 pre-menopausal women with breast cancer pre- and post-treatment In group 1, patients were offered ovarian reserve testing before receiving chemotherapy and further testing immediately following chemotherapy In group 2, patients were tested for ovarian	<u>EIA AMH/MIS assay</u> The sensitivity of the assay was 0.098 ng/mL. The intra- and inter-assay CoV were <15% using an in- house quality control pool.	<u>DSL inhibin B ELISA</u> The sensitivity of the assay was 10 pg/mL. The intra- and inter-assay CoV were <10%. <u>DPC Immuno-radiometric FSH assay</u> The sensitivity of the assay was 0.06 mIU/mL and the intra- and inter-assay CoV was <7%. <u>IBL E2 ELISA</u>	Correlation: Spearman's rank correlation test

Authors Publication year Location	Study design Level of evidence ^a Risk of bias ^b	Study population characteristics	Intervention	Comparator	Measurement of relevant outcomes
		reserve before chemotherapy and after chemotherapy (once regular menstrual cycles had resumed).		The sensitivity of the assay was 4.6 pg/mL and the intra- and inter-assay CoV was <6%.	
Nielsen et al. (2013) Denmark	Case series Quality: Low risk of bias	N=71 female childhood cancer survivors who were less than 15 years old at the time of diagnosis, were treated with radiotherapy and/or chemotherapy, are in complete remission and were at least 18 years old at study inclusion.	<u>EIA AMH/MIS assay</u> According to the manufacturer, the intra- and inter-assay CoVs were ≤12.3% and ≤14.2%, respectively, and the analytical sensitivity, defined as the lowest AMH concentration from the zero calibrator, was 0.7 pmol/l. The functional sensitivity, defined as the lowest concentration that gives a day-to-day CoV ≤25%, was estimated to be 3 pmol/l.	<u>AFC by TVUS</u> The number of small antral follicles with a size of 2–10 mm was counted for each ovary and the AFC was recorded as the number of follicles in both ovaries.	Correlation: Spearman's rank correlation test
Paradisi et al. (2016) Italy	Case series Quality: Low risk of bias	N=191 patients with non-gynaecological malignancies before chemo-radiotherapy and OTC who had regular menstrual cycles (26–32 days); no evidence of endocrine/metabolic diseases; no ovarian abnormalities; no ovarian surgery; no previous chemo-radiotherapy and no hormonal therapy in the three months preceding OTC Mean age = 26.4 ± 6.9 years (range 12–38).	A blood sample was used to analyse the serum levels of FSH, E2, inhibin B and AMH. All serum measurements were performed in duplicate at the Central Laboratory. <u>AMH Gen II ELISA</u> The sensitivity of the methods and the respective intra- and inter-assay CoV for the hormonal assays were in agreement with those reported by the manufacturer.	<u>AFC by TVUS</u> A transvaginal ultrasound was performed. Only one experienced sonographer did the ultrasound to avoid the inter-observer variability. The limit of sensitivity was 2 mm and the intra-observer CoV was 7%. AFC was defined as the total number of visible intra-ovarian sonolucent structures with diameter 2–9mm in both ovaries. <u>Inhibin B Gen II ELISA</u> <u>FSH Elecsys assay</u> <u>E2 Elecsys assay</u>	Correlation: Spearman's rank correlation test
Partridge et al. (2010) USA	Case series Quality: Low risk of bias	N=20 breast cancer survivors with continued menses after chemotherapy Mean age = 36.8 years (range 31–42).	<u>DSL ACTIVE® AMH ELISA</u> The minimum reportable concentration of this test is 0.03 ng/mL. Testing is monitored using quality control sera (two levels); the intra-assay CoV is <6% and the inter-assay CoV is <12%.	<u>AFC by TVUS</u> AFC was performed by two reproductive endocrinologists. Standardisation of caliper placement for follicle measurement was confirmed before the start of the study, and was consistent with clinical AFC measurements used for standard follicle monitoring for fertility patients. Ovaries were scanned cephalad to caudad in the coronal plane, to ensure complete visualization of	Correlation: Spearman's rank correlation test

Authors Publication year Location	Study design Level of evidence ^a Risk of bias ^b	Study population characteristics	Intervention	Comparator	Measurement of relevant outcomes
				the ovary, and avoid missing follicles.	
van Beek et al. (2007) the Netherlands	Case series Quality: Low risk of bias	32 women treated with chemotherapy for Hodgkin's lymphoma during childhood Median age at diagnosis = 14 years (range 5.0–17.2) Median age at follow-up = 25 years (range 19.2–40.4).	<u>DSL ACTIVE® AMH ELISA</u> Intra- and inter-assay CoVs <5 and <8%.	<u>Serotec inhibin B ELISA</u> Intra- and inter-assay CoVs were <9 and <15% <u>Immulite 2000 FSH immunoassay</u> Intra- and inter-assay CoVs were <3 and <8%	Correlation: Spearman's rank correlation test

AFC = antral follicle count; AMH = anti-Müllerian hormone; ASRM = American Society for Reproductive Medicine classification of endometriosis; BMT = bone marrow transplant; CoV = coefficient(s) of variation; E2 = estradiol; EIA = enzyme immunoassay; ELISA = enzyme-linked immunosorbent assay; FSH = follicle-stimulating hormone; HRT = hormone replacement therapy; LH = luteinising hormone; MIS = Müllerian-inhibiting substance; OTC = ovarian tissue cryopreservation; TVUS = transvaginal ultrasound examination

^a Source: See NHMRC hierarchy of evidence

^b Risk of bias as it relates to primary outcomes of the systematic review

Table 85 Prognostic evidence on ovarian function / amenorrhea / menses

Author, year, country Study design Level of evidence Quality appraisal	Population characteristics	Objectives Eligibility criteria	AMH test and comparator	Prognostic outcomes assessed
Anders et al. (2008) USA <u>Level:</u> IV (prospective study) <u>Quality:</u> Moderate risk of bias	<u>Population:</u> 44 pre-menopausal women aged 18–55, with histologically confirmed, stage I–II diagnosis of operable breast adenocarcinoma. Age: 40 (range 21–51) <u>Treatment:</u> a planned course of chemotherapy with a ≥25% risk of permanent amenorrhea.	<u>Objective:</u> to determine prospectively-validated, predictive markers of CRA. <u>Inclusion criteria:</u> Karnofsky score ≥70, ability to provide informed consent, planned follow-up at Duke University <u>Exclusion criteria:</u> history of ovarian tumour, current pregnancy, oral contraceptives within 30 days of study enrolment.	<u>Test characteristics:</u> AMH was measured via ELISA (Diagnostic Systems Laboratories, Inc). Sensitivity was 0.017 ng/mL for AMH. <u>Comparator:</u> inhibin B, FSH, estradiol, Inhibin A <u>Analysis:</u> Medians of each serum marker at each visit on all 6 end-points were calculated according to menstrual status and according to age group. Differences between CRA and age subgroups on all end-points were tested with the non-parametric Wilcoxon two-sample test using a two-sided alpha of 0.10. The Wilcoxon test was also used to test for subgroup differences on change in hormone level	<u>CRA defined as:</u> absence of menses one year post-chemotherapy. <u>Outcomes:</u> Baseline median serum hormone values (FSH, estradiol, inhibin A, inhibin B, AMH). Subgroups of baseline median serum hormone values: CRA and no CRA one year post-

Author, year, country Study design Level of evidence Quality appraisal	Population characteristics	Objectives Eligibility criteria	AMH test and comparator	Prognostic outcomes assessed
			<p>from pre-chemotherapy to post-chemotherapy.</p> <p><u>Follow-up period:</u> Patients were evaluated at four study time points: pre-chemotherapy (≤ 3 weeks of chemotherapy), post-chemotherapy (3 to 7 weeks post-chemotherapy), 6 months post-chemotherapy (18–30 weeks post-chemotherapy), and one year post-chemotherapy (48–60 weeks post-chemotherapy).</p>	<p>chemotherapy.</p> <p>Relative risk of CRA based on median hormone values.</p>
<p>Anderson & Cameron. (2011) UK <u>Level:</u> II (prospective study) <u>Quality:</u> Low risk of bias</p>	<p><u>Population:</u> 56 pre-menopausal women with early operable breast cancer.</p> <p><u>Treatment:</u> Three received neo-adjuvant adriamycin and cyclophosphamide for six cycles, and 38 had post-operative chemotherapy of which three received six cycles of cyclophosphamide, methotrexate, and 5-fluorouracil (CMF), 25 received sequential anthracycline-CMF for eight to 12 cycles, and 11 were given eight cycles of sequential anthracyclines and taxanes. Thereafter, four women received no hormonal treatment, with the remainder generally treated with tamoxifen with additional goserelin in eight. One woman was treated with the aromatase inhibitor (AI) letrozole throughout the study after completion of chemotherapy. An</p>	<p><u>Objective:</u> Determine whether pre-treatment AMH concentration predicts long-term ovarian function in women treated with chemotherapy for early breast cancer more accurately than age.</p> <p><u>Inclusion criteria:</u> ability to provide informed consent, operable breast cancer, regular menses, no hormonal contraception.</p>	<p><u>Test characteristics:</u> AMH was measured by ELISA (Active MIS/AMH; Beckman Coulter, Chaska, MN), with a sensitivity of 0.05 ng/mL, and inter-assay and intra-assay CoV were 4.0 and 3.6%, respectively.</p> <p><u>Analysis:</u> Proportions of women with amenorrhea were analysed by Fisher's exact test. Initial analysis of hormonal potential predictors of amenorrhea was analysed by Student's t-test after log transformation to correct for heterogeneity of variance. Where this indicated a significant relationship between the pre-treatment hormone variable and late amenorrhea (i.e. AMH and FSH but not inhibin B or estradiol), multivariate logistic regression analysis was performed to determine which factors independently predicted amenorrhea.</p> <p><u>Follow-up period:</u> Patients re-attended at 2, 3, 4, and 5 years after starting chemotherapy for their breast cancer. Return visits were scheduled to be in the early follicular phase (day 2–5) if patients continued to have menstrual cycles.</p>	<p><u>CRA defined as:</u> Patients kept a menstrual diary. Amenorrhea was defined as no further bleeding subsequently over 6 months.</p> <p><u>Outcomes:</u> Mean pre-treatment hormone concentrations compared to ovarian function at 4–5 years. Relationships between pre-chemotherapy age, AMH, and FSH with later ongoing menses (multivariate regression analysis). Sensitivity and specificity of AMH for predicting amenorrhea (ROC analysis).</p>

Author, year, country Study design Level of evidence Quality appraisal	Population characteristics	Objectives Eligibility criteria	AMH test and comparator	Prognostic outcomes assessed
	additional 11 women were treated with an AI during the second half of the study, generally changing from tamoxifen at between 36 and 56 months of the study except in one case who changed at month 18.			
Anderson et al. (2013) UK <u>Level:</u> II (prospective study) <u>Quality:</u> Low risk of bias	<u>Population:</u> 59 pre-menopausal women with early breast cancer (data available from n=55 at 1 year and n=46 at 2 years) Mean age: 42.6 years (range 23.3–52.5) <u>Treatment:</u> 44 women received tamoxifen treatment following chemotherapy, and seven received goserelin (only one woman received goserelin but not tamoxifen) and one woman was treated with anastrozole in addition to tamoxifen.	<u>Objective:</u> test whether AMH measured at the time of diagnosis would be a clinically useful predictor of amenorrhea after chemotherapy for early breast cancer, in comparison to age at diagnosis or other biochemical markers of ovarian reserve. <u>Inclusion criteria:</u> primary operable breast cancer without evidence of metastases, being pre-menopausal, absence of sex steroid contraception, or pre-menopausal gonadotrophin and estradiol concentrations. <u>Exclusion criteria:</u> previous surgery to either ovary or previously received chemotherapy.	<u>Test characteristics:</u> AMH was measured by the Gen II ELISA kit (Beckman Coulter, Chaska, MN). This has a sensitivity of 0.16 ng/mL (1.1 pmol/L) and in-house intra- and inter-assay CoV of <6%. <u>Analysis:</u> Initial analysis of predictors of amenorrhoea (i.e. the primary objective of the study) was performed by Student's t-test, with log transformation of hormonal data to correct for heterogeneity of variance. Because of relationships between the variables, a multivariate logistic regression analysis was performed to determine which factors independently predicted amenorrhoea. Data of this study was combined with data from Anderson & Cameron (2011), giving a cohort of n=75, and analysis of the AUC of ROC curve plots for age and AMH as separate predictors were done. Furthermore, the relative importance of age and AMH was calculated by the use of Random Forests to derive 2000 classification trees each of which uses AMH and age to predict amenorrhea. <u>Follow-up period:</u> 2 years.	<u>CRA defined as:</u> patients kept a menstrual diary. Amenorrhea was defined as no bleeding for the previous 6 months. <u>Outcomes:</u> Mean pre-treatment hormone concentrations compared to ovarian function at 1 and 2 years. Relationships between pre-chemotherapy age, AMH, inhibin B and FSH with amenorrhea (logistic regression). AUC of AMH and age predicting amenorrhea at 2 year follow-up.
Chai et al. (2014) UK <u>Level:</u> IV (prospective study) <u>Quality:</u> Low risk of bias	<u>Population:</u> Pre-menopausal women with early operable breast cancer, two cohorts (n=53 and n=42).	<u>Objective:</u> to investigate AMH assay in the assessment of ovarian function after chemotherapy in women with early breast cancer, in combination with other current markers of ovarian reserve.	<u>Test characteristics:</u> AMH was assayed using the Ansh labs pico-AMH ELISA kit (Ansh Catalog no. AL-124, Webster, TX). For comparison with previous assays, serum samples taken at 2 years in cohort 1 also analysed using the Active MIS/AMH ELSIA (Beckman Coulter, Chaska,	<u>CRA defined as:</u> no ongoing menses (amenorrhea) <u>Outcomes:</u> AMH levels in patients with

Author, year, country Study design Level of evidence Quality appraisal	Population characteristics	Objectives Eligibility criteria	AMH test and comparator	Prognostic outcomes assessed
	Age (mean \pm SEM): Cohort 1 (n=53): 42.5 \pm 0.9 years Cohort 2 (n=42): 43.0 \pm 0.9 years <u>Treatment:</u> chemotherapy regimens: sequential anthracycline-CMF or anthracyclines and taxanes in most, with 14 women in cohort 1 not given any chemotherapy; therapeutic decisions were not influenced by this study.	<u>Inclusion criteria:</u> regular menses, absence of hormonal contraception. <u>Exclusion criteria:</u> -	MN), sensitivity 0.05 ng/mL <u>Analysis:</u> The primary analysis was comparison of hormone concentrations versus menstrual function at 2 years post diagnosis, with groups compared by ANOVA; in longitudinal analyses time points were compared to data at 2 years using Dunn's post-hoc test. <u>Follow-up period:</u> patients were evaluated at 2, 3, 4, and 5 years after diagnosis in cohort 1, and after 1 and 3 years in cohort 2.	CRA compared to AMH levels in patients with ongoing menses.
D'avila et al. (2015) Brazil <u>Level:</u> IV (prospective study) <u>Quality:</u> Moderate risk of bias	<u>Population:</u> 52 women with breast cancer, undergoing chemotherapy with cyclophosphamide Mean age: 35.3 \pm 3.8 years (range 27–40 years) <u>Treatment:</u> chemotherapy with cyclophosphamide. 40% of patients underwent breast-conserving surgery prior to or followed by chemotherapy treatments and 75% underwent adjuvant radiotherapy.	<u>Objective:</u> Determine which ovarian reserve measurement can be used as a predictor for anovulation 6 months after chemotherapy with cyclophosphamide in women with breast cancer. <u>Inclusion criteria:</u> younger than 40 years of age, requiring chemotherapy containing cyclophosphamide, no previous chemotherapy treatment. <u>Exclusion criteria:</u> -	<u>Test characteristics:</u> AMH was measured through ELISA (Beckman Coulter, Genese Immunotech®, France) <u>Comparator:</u> AFC and FSH <u>Analysis:</u> Results are presented as median and interquartile range (25–75 %) because the data of this study do not show Gaussian normal distribution. The data were tested with the Mann-Whitney test and the multiple comparisons were corrected through Bonferroni. The categorical variables were analysed through Pearson Chi-square test. A logistic regression for independent samples and determination of the ROC curve were performed. <u>Follow-up period:</u> Median 14 \pm 3 months (results were based on 6 month follow-up)	<u>Outcomes:</u> Baseline AMH and AFC values and menstrual outcomes 6 months after chemotherapy OR, sensitivity and specificity of AMH, AFC and FSH when certain cut-off values are used for predicting amenorrhea or oligomenorrhea. ROC AUC of AMH, AFC and FSH for predicting amenorrhea or oligomenorrhea.
Dezellus et al. (2017) France <u>Level:</u> IV (prospective)	<u>Population:</u> 249 women with breast cancer undergoing chemotherapy. Mean age: 34.8 \pm 3.9	<u>Objective:</u> to evaluate the serum AMH level at diagnosis, its evolution throughout chemotherapy and its long-	<u>Test characteristics:</u> All assays were performed in the same centralised laboratory with 1 st generation AMH/Müllerian-Inhibiting Substance Enzyme-Immuno	<u>CRA defined as:</u> chemotherapy-related amenorrhea (not further defined)

Author, year, country Study design Level of evidence Quality appraisal	Population characteristics	Objectives Eligibility criteria	AMH test and comparator	Prognostic outcomes assessed
cohort) <u>Quality:</u> Low risk of bias	<u>Treatment:</u> 75 (30.1%) had neoadjuvant chemotherapy, 174 (69.9%) had adjuvant chemotherapy.	term evolution during a 24 month follow-up in women of reproductive age treated with chemotherapy in adjuvant or neoadjuvant settings for breast cancer. <u>Inclusion criteria:</u> patients aged 18–39 years, diagnosed with breast cancer (T0-T4, N1-N3, M0), prior to treatment with chemotherapy in an adjuvant/neoadjuvant setting. <u>Exclusion criteria:</u> menopause, breast cancer not treated with chemotherapy, history of previous malignancy treated with chemotherapy and comorbidities related to fertility.	Assay (MIS EIA) Immunotech Beckman Coulter™ method (Marseille, France) according to the manufacturer instructions. The lower limit of detection was 0.14 ng/mL. The lower limit of quantification was 0.42 ng/mL. The analytical measurement range was 0.42–21 ng/mL. Repeatability (intra-assay precision) and reproducibility (inter-assay precision) coefficients were below or equal to 12.3% and 14.2%, respectively. <u>Analysis:</u> Mann-Whitney and Kruskal Wallis tests were used for group comparison and Spearman coefficient for correlation studies <u>Follow-up period:</u> 6, 12 and 24 months.	<u>Outcomes:</u> Mean AMH scores in the CRA and no CRA groups (mean and SD, p-value).
Henry et al. (2014) USA <u>Level:</u> II (prospective study) <u>Quality:</u> Low risk of bias	<u>Population:</u> N=29 pre- and peri-menopausal women with newly diagnosed breast cancer 25–34 years: n=6 35–39 years: n=8 40–44 years: n=7 45–50 years: n=8 <u>Treatment:</u> adjuvant or neoadjuvant chemotherapy	<u>Objective:</u> to test the hypothesis that low pre-chemotherapy serum concentrations of AMH and inhibin B would predict lack of recovery of ovarian function following chemotherapy. <u>Inclusion criteria:</u> women aged 25-50 years with a menstrual cycle within 3 months prior to study entry, diagnosed with stage I–III breast cancer, scheduled for neoadjuvant or adjuvant chemotherapy <u>Exclusion criteria:</u> prior cytotoxic chemotherapy, bilateral oophorectomy, hysterectomy, pelvic radiation.	<u>Test characteristics:</u> Serum AMH concentrations were measured using the AMH Gen II ELISA (A73818; Beckman Coulter, Fullerton, CA). the minimum detectable concentration was 0.16 ng/mL, the upper LoQ was 22.5 ng/mL, and the inter- and intra-assay CoVs were 7.1% and 2.9%, respectively <u>Comparator:</u> inhibin B was measured with Gen II ELISA (A81301; Beckman Coulter), the minimal detectable concentration was 10 pg/mL, the upper LoQ was 1,000 pg/mL, and the inter- and intra-assay CoVs were 4.3% and 2.8%, respectively. Serum FSH concentrations were measured using a two-site chemiluminescence sandwich assay, which has a minimum detectable concentration of 0.3 mIU/mL and an upper LoQ of 200 mIU/mL. Inter- and intra-assay CoVs	<u>Ovarian function defined as:</u> recurrence of menses or serum estradiol concentration >10 pg/mL <u>Outcomes:</u> Univariate analysis of the effect of potential covariates on recovery of ovarian function after chemotherapy.

Author, year, country Study design Level of evidence Quality appraisal	Population characteristics	Objectives Eligibility criteria	AMH test and comparator	Prognostic outcomes assessed
			<p>were 8.1% and 3.5%, respectively. Serum estradiol concentrations were measured using a gas chromatography tandem mass spectroscopy assay (InVentiv Health Clinical, Princeton, NJ, which has a minimal detectable concentration of 0.625 pg/mL.</p> <p><u>Analysis:</u> In univariate analysis, recovery of ovarian function was the response variable and was tested against each potential covariate. The multivariate analysis included only those factors that were significant in the univariate analysis.</p> <p><u>Follow-up period:</u> 18 months (average 13.6 months)</p>	
<p>Ruddy et al. (2014) USA <u>Level:</u> II (prospective study) <u>Quality:</u> Moderate risk of bias</p>	<p><u>Population:</u> 124 pre-menopausal women with breast cancer and 12-month menses data (n=100 women provided 18 month menses data) Median age: 46 (range 25–55) years. Only 36 women were ≤40 years.</p> <p><u>Treatment:</u> doxorubicin-cyclophosphamide followed by paclitaxel with either placebo or one of two durations of bevacizumab therapy.</p>	<p><u>Objective:</u> to investigate whether pre-chemotherapy AMH is a biomarker for CRA in breast cancer patients.</p> <p><u>Inclusion criteria:</u> pre-menopausal, breast cancer</p> <p><u>Exclusion criteria:</u> patients who underwent ovarian suppression or bilateral salpingoophorectomy prior to the 12-month follow-up, and patients who did not give consent and/or who had no stored serum available.</p>	<p><u>Test characteristics:</u> AMH was measured by two-site ELISA (Diagnostic Systems Laboratory, BeckmanCoulter, Webster, TX) from baseline serum samples drawn and banked before chemotherapy. Testing was monitored using quality control sera (two levels); the intra-assay CoV was <6 % and the inter-assay CoV was <12 %.</p> <p><u>Comparator:</u> NA</p> <p><u>Analysis:</u> Wilcoxon rank sum and Fisher's exact tests were used to assess for univariate associations between 12- and 18-month CRA. Multivariate logistic regression was used to assess for predictors of CRA at both time points including baseline AMH, age, race, whether or not the patient received bevacizumab, and use of tamoxifen</p> <p><u>Follow-up period:</u> 12 (n=124) and 18 (n=100) months</p>	<p><u>CRA defined as:</u> having not experienced a period within the six months prior to the follow-up.</p> <p><u>Outcomes:</u> Univariate predictors of 12 month and 18 month CRA (AMH and age) Multivariate predictors of 12 and 18 month CRA.</p>
<p>Su et al. (2014)</p>	<p><u>Population:</u> N=109 women with</p>	<p><u>Objective:</u> To examine the association</p>	<p><u>Test characteristics:</u> AMH was measured using the AMH</p>	<p><u>Return of ovarian function was</u></p>

Author, year, country Study design Level of evidence Quality appraisal	Population characteristics	Objectives Eligibility criteria	AMH test and comparator	Prognostic outcomes assessed
USA <u>Level:</u> II (prospective study) <u>Quality:</u> Low risk of bias	breast cancer and at least 3 months of amenorrhea with chemotherapy <u>Treatment:</u> chemotherapy.	of pre-chemotherapy AMH, FSH and inhibin B levels with the timing of post-chemotherapy ovarian function in young women with breast cancer, and to generate a prognostic score for ovarian recovery. <u>Inclusion criteria:</u> aged 18–45 years, diagnosed with early stage breast cancer (stage I–III), had a uterus and at least one ovary, and reported at least one menses over the previous 12 months. Patients had to have received chemotherapy and experienced secondary amenorrhea with chemotherapy. <u>Exclusion criteria:</u> Pregnancy, breast feeding, use of psychotropic drugs known to impact ovulation, and history of prior cancer, chemotherapy or pelvic radiation.	enzyme-linked immunosorbent assay kit, which has a limit of detectability of 0.17 ng/mL (AMH Gen II assay; Beckman Coulter Inc., Brea, Calif) <u>Comparator:</u> FSH was measured by direct immunochemiluminometric assay using the automated Immulite system (Siemens Medical Solutions, Los Angeles, Calif) E2 was measured by radioimmunoassay after an organic solvent extraction step and had a limit of detectability of 3 pg/mL. Inhibin B assays used a monoclonal, 2-site ELISA with a limit of detectability of 9.4 pg/mL (Diagnostic Systems Laboratories, Webster, Tex). Inter-assay CoV were <10% for all assays. <u>Analysis:</u> Time-to-event methods were used. Kaplan-Meier survival curves were generated, and the time to return of menses was compared according to baseline characteristics using log rank tests. A Cox proportional-hazards regression model was developed to examine predictors of return of ovarian function and to control for confounding. All variables with p<0.05 based on the Wald test from univariate analysis and known potential confounders (body mass index, race, chemotherapy regimen, and tamoxifen) were included in the multivariable model. <u>Follow-up period:</u> median follow-up was 163 days (range 4–1009 days) after chemotherapy	<u>defined as:</u> from the end-of-chemotherapy treatment to the first episode of vaginal bleeding. <u>Outcomes:</u> Pre-chemotherapy levels of AMH, FSH, inhibin B and estradiol and unadjusted HRs of return of ovarian function Multivariable model of time to return of ovarian function (adjusted HRs).
Su et al. (2010)	<u>Population:</u> 127 post-chemotherapy	<u>Objective:</u> to determine the impact of	<u>Test characteristics:</u> AMH was assayed using AMH ELISA	<u>CRA defined as:</u> menstrual

Author, year, country Study design Level of evidence Quality appraisal	Population characteristics	Objectives Eligibility criteria	AMH test and comparator	Prognostic outcomes assessed
USA <u>Level:</u> IV <u>Quality:</u> Low risk of bias	breast cancer survivors Median age at initiation of chemotherapy (range): 43.2 (26.7–57.8) years <u>Treatment:</u> cyclophosphamide-based adjuvant chemotherapy.	breast cancer treatment on hormones, to determine the association between hormones and CRA in cancer survivors, and to examine whether hormones can predict subsequent menstrual patterns in cancer survivors. <u>Inclusion criteria:</u> Cancer stage I–III breast cancer, pre-menopausal status at cancer diagnosis, cyclophosphamide-based adjuvant chemotherapy, the presence of a uterus and at least 1 ovary, initiation of adjuvant chemotherapy 1–4 years before enrolment. <u>Exclusion criteria:</u> -	kits (Diagnostic Systems, Webster, Tex). The lower limit of detection for AMH was 25 pg/mL (SI conversion: 1 ng/mL = 7.14 pmol/L), and the intra-assay CoV was 2%. <u>Comparator:</u> Inhibin B was assayed using inhibin B ELISA kits (Diagnostic Systems). The intra-assay and inter-assay CoV were 7.9% and 8.4%, respectively. The lower limit of detection was 5 pg/mL. Estradiol and FSH were measured by radioimmunoassay using Coat-A-Count commercial kits (Diagnostic Products, Los Angeles, Calif). The intra- and inter-assay CoV were <5%. <u>Analysis:</u> the association between assessment 1 hormones and change in CRA status between assessments 1 and 2 in breast cancer subjects was examined using the Student t-test. Assessment 2 CRA status was categorized as “no change” from assessment 1 CRA status,” “CRA reversal,” or “CRA progression.” <u>Follow-up period:</u> median years of follow-up from initiation of chemotherapy to the first assessment (AMH testing) was 2.1 (range 1.0–4.9) Second follow-up was at 2–7 years after chemotherapy.	history and defined as ≥ 12 months of amenorrhea occurring after start of chemotherapy. <u>Outcomes:</u> Median hormone values in patients with CRA vs hormone values in patients with no CRA on second follow-up.
Yu et al. (2010) USA <u>Level:</u> IV <u>Quality:</u> Moderate risk of bias	<u>Population:</u> 26 pre-menopausal women undergoing adjuvant chemotherapy for early stage breast cancer <u>Treatment:</u> adjuvant chemotherapy.	<u>Objective:</u> to evaluate the effect of adjuvant chemotherapy on markers of ovarian reserve and endocrine function in women with breast cancer, and to evaluate the predictive potential of changes in these biomarkers for determining future menstrual status.	<u>Test characteristics:</u> AMH was measured using an enzyme-linked immunoadsorbent assay (Diagnostic Systems Laboratories, Inc, Webster, Tex) with sensitivity of 0.05 ng/mL. <u>Comparator:</u>	<u>CRA defined as:</u> the absence of menses at 52 weeks <u>Outcomes:</u> AMH levels at baseline by

Author, year, country Study design Level of evidence Quality appraisal	Population characteristics	Objectives Eligibility criteria	AMH test and comparator	Prognostic outcomes assessed
		<p><u>Inclusion criteria:</u> enrolled after surgery and before initiation of chemotherapy, aged <40 years</p> <p><u>Exclusion criteria:</u> -</p>	<p>E2 was measured using a radioimmunoassay (Siemens Medical Solutions Diagnostics, Los Angeles, Calif) with a sensitivity of 5 pg/mL. FSH was measured using chemiluminescent enzyme immunoassays (Immulite; Siemens Medical Solutions Diagnostics, Los Angeles, Calif) with a sensitivity of 20 pg/mL or 0.1 mIU/mL.</p> <p><u>Analysis:</u> Differences between the amenorrhea and age subgroups on all time points were tested with the non-parametric Wilcoxon 2-sample test using a 2-sided α of 0.05. The Wilcoxon test was also used to test for differences on change in hormone levels at different time points before, during, and after chemotherapy.</p> <p><u>Follow-up period:</u> 52 weeks after initiation of chemotherapy</p>	<p>subgroup (CRA or no CRA)</p> <p>Resumption of menstrual function</p>

AFC = antral follicle count; AMH = anti-Müllerian hormone; ANOVA = analysis of variance; AUC = area under the curve; BMI = body mass index; CRA = chemotherapy-related amenorrhea; CoV = coefficient(s) of variation; E2 = estradiol; EIA = enzyme Immunoassay; ELISA = enzyme-linked immunosorbent assay; FSH = follicle-stimulating hormone; HR = hazard ratio; HRT = hormone replacement therapy; LH = luteinising hormone; LoQ = level of quantification; NA = not available; OR = odds ratio; ROC = receiving operating characteristic

Table 86 Predictive evidence: AMH levels predicting ovarian response

Author, year, country Study design Level of evidence Quality appraisal	Population characteristics	Objectives Eligibility criteria	AMH test	Prognostic outcomes assessed
<p>Lee et al. (2011) USA <u>Level:</u> IV (retrospective study) <u>Quality:</u> Low risk of bias</p>	<p><u>Population:</u> 41 women with stage ≤3 breast cancer undergoing cryopreservation. Seven women has oocyte freezing, 29 had embryo crypreservation and four underwent both embryo and oocyte cryopreservation. Mean age: 34.8 ± 4.7 years <u>Treatment:</u> NR</p>	<p><u>Objective:</u> to predict embryo/oocyte cryopreservation cycle outcomes in breast cancer patients stimulated with letrozole and FSH for fertility preservation based on observed AMH levels and AFC. <u>Inclusion criteria:</u> age <45 years, breast cancer stage ≤3, no prior chemotherapy, no prior history of ovarian surgery or infertility and availability of serum AMH level <u>Exclusion criteria:</u> -</p>	<p><u>Test characteristics:</u> AMH was performed using highly specific mono/mono two-site ELISA method (DSL AMH assay). <u>Analysis:</u> Mann-Whitney U-test or Student's t-test was performed to analyse differences in mean values (presented as mean ± SD). Non-parametric correlations were analysed with Spearman's test and parametric correlations with Pearson's test. Multiple linear regression models were developed using a stepwise procedure introducing variables with a cut-off p=0.10 but requiring p=0.05 in the final model.</p>	<p>Comparison of number of MII oocytes retrieved, total number of oocytes retrieved, maturation rate (number of mature oocytes/number of total oocytes), fertilisation rate (number of fertilised oocytes/number of oocytes undergoing ICSI) and total number of embryos cryopreserved following ICSI.</p>
<p>Manno et al. (2016) Italy <u>Level:</u> IV (retrospective) <u>Quality:</u> Moderate risk of bias</p>	<p><u>Population:</u> 38 patients with estrogen receptor-positive (n=17), estrogen receptor negative (n=6) breast cancer, Hodgkins lymphoma (n=8) or non-Hodgkins lymphoma (n=2) or other types of cancer (n=5). Mean age: 31.2 years (range 16-39 years) <u>Treatment:</u> chemotherapy</p>	<p><u>Objective:</u> optimising fertility preservation counselling by presenting correlations between fertility preservation cycle outcomes, humoral and biophysical markers, and methods of ovarian stimulation. <u>Inclusion criteria:</u> NR <u>Exclusion criteria:</u> NR</p>	<p><u>Test characteristics:</u> AMH values were determined using random sampling and each sample was stored at -20°C before assay with Beckman Coulter Generation II ELISA (BC Gen II) system and, after its recent implementation, Elecsys Roche automated method. <u>Analysis:</u> Pearson correlation <u>Follow-up period:</u> NR</p>	<p>Correlation between oocytes retrieved/vitrified and AMH values, AFC values and peak E2 values.</p>
<p>Sonigo et al. (2016) France <u>Level:</u> IV (prospective)</p>	<p><u>Population:</u> 340 cancer patients. 100 women (29%) already had at least one child. Indications were breast cancer (n=300), haematological malignancies (n=14), and other diseases requiring emergency</p>	<p><u>Objective:</u> to determine the threshold values of AFC and AMH levels, for ensuring the cryopreservation of sufficient numbers of in vitro matured oocytes for cancer patients as candidates for fertility</p>	<p><u>Test characteristics:</u> Serum AMH levels were determined using an ultrasensitive ELISA (Beckman Coulter, Villepinte, France). Intra- and inter-assay variation coefficients were 6 and 10%, respectively, the lower detection limit was 0.13</p>	<p>Association between number of in vitro matured oocytes cryopreserved and AFC and serum AMH levels</p>

Author, year, country Study design Level of evidence Quality appraisal	Population characteristics	Objectives Eligibility criteria	AMH test	Prognostic outcomes assessed
study) <u>Quality:</u> Low risk of bias	chemotherapy (n=26). 301 women opted for oocyte cryopreservation, 39 patients chose embryo freezing and 47 had ovarian tissue cryopreservation in combination with in vitro maturation. Mean age: 31.8 ± 4.5 years. <u>Treatment:</u> chemotherapy	preservation. <u>Inclusion criteria:</u> two ovaries present, visible and easily accessible ultrasound-guided puncture; AFC >8 follicles; no previous history of chemotherapy <u>Exclusion criteria:</u> -	ng/mL and linearity was up to 21 ng/mL. <u>Analysis:</u> Relationships between continuous variables were assessed using the Spearman test to determine coefficients of correlation. For each oocyte threshold, a univariate logistic regression was performed and effect measures were estimated using exposure ORs and their corresponding 95% CIs. ROC curves were plotted and AUC was determined. ROC curves were calculated to examine the diagnostic test performance (i.e. the ability to discriminate either patients for whom at least 8, 10 or 15 mature oocytes are cryopreserved, or those with ≤2). AUC was calculated and sensitivity as well as 1-specificity were plotted at each AFC and AMH threshold value. Sensitivities and specificities were estimated with the choice of maximising sensitivity (≈80%) with acceptable specificity (≈60%). All tests were two-sided at a 0.05 significance level.	Threshold values and their sensitivity and specificity of AFC and serum AMH levels for obtaining ≤2 or ≥8, 10 or 15 mature oocytes frozen.
Takae et al. (2015) Japan <u>Level:</u> IV (retrospective study) <u>Quality:</u> Moderate risk of bias	<u>Population:</u> 27 breast cancer patients with regular menstrual cycles. (7 patients had already been pregnant at least once) Mean age: 33.7 ± 3.8 years. <u>Treatment:</u> NR	<u>Objective:</u> to elucidate the factors that correlate to the efficacy of various combined fertility preservation methods on young breast cancer patients who received ovarian tissue cryopreservation for fertility preservation. <u>Inclusion criteria:</u> Breast cancer patients referred to the oncofertility unit at the Center for Reproductive Medicine from February 2010 to March 2014.	<u>Test characteristics:</u> Serum AMH levels were measured using a commercial assay kit (AMH Gen II ELISA, Beckman Coulter, Brea, CA, USA) according to the manufacturer's protocol. The detection limit of this kit was 0.16 ng/mL. <u>Analysis:</u> to investigate the correlations between AMH levels, age, BMI, ovarian tissue volume, and the number of extracted oocytes, a non-parametric Spearman's correlation coefficient by rank test was used.	Correlation between the patient's AMH level and the number of extracted oocytes.

Author, year, country Study design Level of evidence Quality appraisal	Population characteristics	Objectives Eligibility criteria	AMH test	Prognostic outcomes assessed
		<u>Exclusion criteria:</u> -		

AFC = antral follicle count; AMH = anti-Müllerian hormone; AUC = area under the curve; BMI = body mass index; CI = confidence interval; CRA = chemotherapy-related amenorrhea; E2 = estradiol; EIA = enzyme Immunoassay; ELISA = enzyme-linked immunosorbent assay; FSH = follicle-stimulating hormone; ICSI = intracytoplasmic sperm injection; LH = luteinising hormone; NA = not available; NR = not reported OR = odds ratio; ROC = receiving operating characteristic

Table 87 Prognostic evidence: AMH levels predicting pregnancy

Author, year, country Study design Level of evidence Quality appraisal	Population characteristics	Objectives Eligibility criteria	AMH test	Prognostic outcomes assessed
Hamy et al. (2016) France <u>Level:</u> IV (retrospective study) <u>Quality:</u> Moderate risk of bias	<u>Population:</u> 134/146 consecutive women aged 26–43 years who received adjuvant or neoadjuvant chemotherapy for breast cancer were included. Median age (range): 35.5 (26–43) <u>Treatment:</u> adjuvant or neoadjuvant chemotherapy.	<u>Objective:</u> to investigate the relationship between the biological assessment of AMH and the occurrence of a pregnancy after treatment with chemotherapy in a cohort of women with breast cancer of reproductive age. <u>Inclusion criteria:</u> - <u>Exclusion criteria:</u> metastases at diagnosis.	<u>Test characteristics:</u> The analyses were carried out in duplicate using Immunotech A11893 kits (Beckman Coulter, Marseille, France). The lower limit of detection for AMH was 0.14 ng/mL; according to the manufacturer's instructions. <u>Analysis:</u> Data are presented as a count (per cent) or median (range). The occurrence of pregnancy during follow-up was analysed within a competing risk framework, with local tumour recurrence as a competing event. Indeed, it was assumed that tumour recurrence would prevent or modify the probability of pregnancy, particularly because of new chemotherapy received. Analyses of both cumulative incidence functions and cause-specific hazards were conducted, Cumulative incidence functions were estimated using standard methodology and compared between groups using Gray tests. Cox proportional cause-specific hazards models were then used. The results are expressed as (cause-specific) HRs versus a reference category, each being tested using a Wald test.	<u>Outcomes:</u> Pregnancies achieved Spontaneous pregnancy rate. Time from chemotherapy to pregnancy. Pregnancy outcomes. Association of baseline AMH and end-of-chemotherapy AMH with occurrence of pregnancy.

Author, year, country Study design Level of evidence Quality appraisal	Population characteristics	Objectives Eligibility criteria	AMH test	Prognostic outcomes assessed
			<p>Follow-up period: Blood samples for AMH assessment were retrieved the day of the first chemotherapy (baseline n=135: one patient had two samples retrieved), during treatment (two to four samples; n=393) and during follow-up (4 months to 5.5 years after the end of the chemotherapy; n=343).</p> <p>Median follow-up: 59 months (range 11–104).</p>	
<p>Iwase et al. (2016) Japan Level: IV Quality: Moderate risk of bias</p>	<p><u>Population:</u> N=58 patients with endometrioma</p> <p><u>Treatment:</u> laparoscopic cystectomy.</p>	<p><u>Objective:</u> To evaluate the serum AMH levels after cystectomy for endometriomas with respect to the postsurgical pregnancy rate and recurrence of endometriomas.</p> <p><u>Inclusion criteria:</u> uni/bilateral endometrioma(s) diagnosed by two or more ultrasound examinations and by magnetic resonance imaging, women aged 20–42 years with regular menstrual cycles, no evidence of any other endocrine disorders or apparent infertility causes.</p> <p><u>Exclusion criteria:</u> previous history of adnexal surgery and any suspicious findings of malignant ovarian diseases.</p>	<p><u>Test characteristics:</u> The serum AMH concentrations were measured with an enzyme immunoassay kit (EIA AMH/MIS; Immuntotech, Marseille, France). AMH Gen II (Beckman Coulter, Inc., Brea, CA), a similar enzyme immunoassay kit for AMH, was used in the latter half of the study due to the EIA AMH/MIS no longer being available. The intra-assay and inter-assay CoV were below 12.3% and 14.2% for the EIA AMH/MIS kit and 5.4% and 5.6% for the AMH Gen II, respectively.</p> <p><u>Analysis:</u> Student's t-test, the Mann-Whitney U-test and the Fisher's exact test were used to compare the patient characteristics and variables between the pregnancy and non-pregnancy groups.</p> <p><u>Follow-up period:</u> >1 year.</p>	<p>Mean AMH values at baseline, at 1 month post-operative and at 1 year post-operative, grouped by pregnancy or non-pregnancy (and p-values).</p>
<p>Lind et al. (2016) Sweden Level: IV Quality: Moderate risk of bias</p>	<p><u>Population:</u> 45 women of reproductive age (18–44 years) with a desire to have children with benign ovarian cysts.</p> <p>Mean age at surgery: 30.4 ± 5.9 years</p>	<p><u>Objective:</u> to investigate the impact on ovarian cyst surgery on ovarian reserve and on reproductive outcome of women that attempted to achieve pregnancy within a two-year post-operative follow-up.</p>	<p><u>Test characteristics:</u> AMH tests were done using an ELISA kit (ACTIVE AMH gen II ELISA, Beckman Coulter Inc. Webster, NY). The intra-assay and inter-assay CoV were 5.4% and 5.6%, respectively.</p> <p><u>Analysis:</u> Independent t-test, Chi-square test/Fischer's</p>	<p>Reproductive behaviour, pregnancy rate and live births per AMH category group.</p>

Author, year, country Study design Level of evidence Quality appraisal	Population characteristics	Objectives Eligibility criteria	AMH test	Prognostic outcomes assessed
	<p>Complete data for 34 women.</p> <p><u>Treatment:</u> ovarian cyst surgery (laparoscopy or laparotomy).</p>	<p><u>Inclusion criteria:</u> women with pain, an ovarian cyst and fear of cancer, scheduled for ovarian cyst surgery.</p> <p><u>Exclusion criteria:</u> pregnancy at inclusion, unilateral oophorectomy, requirement of repeated surgery for new ovarian cysts during follow-up, no attendance at follow-up, malignancy.</p>	<p>exact test or Pearson Chi-square test were used for comparisons. One-way ANOVA test, the Kruskal-Wallis test and Pearson's Chi-squared test were used to compare subgroups based on desire for children at baseline, achievement of pregnancy and absolute AMH concentration at the six-month follow-up visit</p> <p><u>Follow-up period:</u> follow-up visits were scheduled at 6 and 24 months post-surgery.</p>	
<p>Ozaki et al. (2016) Japan <u>Level:</u> IV (prospective) <u>Quality:</u> Moderate risk of bias</p>	<p><u>Population:</u> Women with endometriomas and no pre-surgical diminished ovarian reserve. 35/112 women tried to have a spontaneous pregnancy following surgery.</p> <p>Mean age: 33.5 ± 4.9</p> <p><u>Treatment:</u> laparoscopic cystectomy</p>	<p><u>Objective:</u> Investigate the factors associated with diminished ovarian reserve and the potential risk of becoming a poor ovarian responder before and after laparoscopic cystectomy.</p> <p><u>Inclusion criteria:</u> symptomatic ovarian endometrioma with a cyst >4 cm, non-pregnant</p> <p><u>Exclusion criteria:</u> patients with pre-surgical adverse diminished ovarian reserve.</p>	<p><u>Test characteristics:</u> The serum AMH concentrations were measured using an enzyme immunoassay kit according to the manufacturer's instructions (EIA AMH/MIS; Immunotech, Marseille, France). The intra- and inter-assay CoV for the AMH assay were below 12.3 and 14.2 %, respectively.</p> <p><u>Analysis:</u> Unpaired Student's t-tests or the Mann-Whitney U-test was performed to compare consecutive variables, and chi-square tests or Fisher's exact test was performed to compare categorical variables.</p> <p><u>Follow-up period:</u> pregnancy rates were measured 2 years post-surgery.</p>	<p>Cumulative spontaneous pregnancy rate (low AMH compared to normal/high AMH).</p>
<p>Pup et al. (2014) Italy <u>Level:</u> IV (retrospective) <u>Quality:</u> Moderate risk of bias</p>	<p><u>Population:</u> 17 women of childbearing age affected by lymphoma. 12 women had AMH testing before treatment.</p> <p><u>Treatment:</u> Haematopoietic cell</p>	<p><u>Objective:</u> to report on the experience of fertility preservation after HCT and to validate specific tools that may predict long-term ovarian function (i.e. AMH).</p> <p><u>Inclusion criteria:</u> women of childbearing</p>	<p><u>Test characteristics:</u> NR</p> <p><u>Analysis:</u> Standard descriptive methods were used (median, ranges, Chi-squared test, and Fischer's exact test)</p>	<p>Pregnancy rate (low AMH compared to normal/high AMH).</p>

Author, year, country Study design Level of evidence Quality appraisal	Population characteristics	Objectives Eligibility criteria	AMH test	Prognostic outcomes assessed
	transplantation. Patients received chemotherapy before HCT. Twelve patients received radiotherapy.	age affected by lymphoma and submitted to haematopoietic cell transplantation <u>Exclusion criteria:</u> -	<u>Follow-up period:</u> NR	
Stochino-Loi et al. (2017) France <u>Level:</u> IV (retrospective study) <u>Quality:</u> Low/moderate/high risk of bias	<u>Population:</u> 180 women with stage 3 and 4 endometriosis and pregnancy intention Mean age: 30.5 ± 4.1 years <u>Treatment:</u> ovarian endometrioma ablation.	<u>Objective:</u> to investigate whether surgery for severe endometriosis may be proposed in women with low ovarian reserve with good fertility outcomes. <u>Inclusion criteria:</u> deep infiltrating endometriosis or endometriomas measuring >3 cm, pregnancy intention before surgery, benefited from pre- and post-operative assessment of AMH, minimum 12 month follow-up. <u>Exclusion criteria:</u> women with superficial endometriosis and hydro salpinx.	<u>Test characteristics:</u> NR <u>Analysis:</u> Fisher's exact test was used to compare qualitative variables, Student's t-test and Mann-Whitney U-test were used to compare continuous variables. Kaplan-Meier curves were done to estimate the probability of non-pregnancy according to post-operative time and were compared using the log rank test. <u>Follow-up period:</u> >1 year. Post-operative follow-up was based on data from questionnaires completed at 1, 3, and 5 years after surgery.	Patients were enrolled in groups (low AMH vs normal/high AMH) Time from surgery to first pregnancy Pregnancy outcomes (delivery, miscarriage, ectopic pregnancy, ongoing pregnancies) Conception mode (spontaneous or ART) Pregnancy rate

AMH = anti-Müllerian hormone; ANOVA = analysis of variance; ART = assisted reproductive technology; CoV = coefficient(s) of variation; EIA = enzyme Immunoassay; ELISA = enzyme-linked immunosorbent assay; HCT = haematopoietic cell transplantation; NR = not reported

APPENDIX D

EVIDENCE PROFILE TABLES

Table 88 Evidence profile table for the analytical validity of AMH testing for girls or women undergoing gonadotoxic treatment

Outcome	No. of participants, no. of studies and study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Results for AMH testing	Results for comparator test	Analytical validity QoE	Importance
Analytical validity of AMH tests compared with clinical reference standard	n=30 k=1 Case series	0 Low risk of bias	NA	-1 No evidence of incremental value of AMH was provided	-1 Only one small study	No	There was a 40-fold difference in level of detectable AMH between the two AMH tests. Predictably, more samples had undetectable levels with the less sensitive EIA AMH/MIS assay (53 samples) compared with the pico-AMH ELISA (33 samples).	NA	⊕○○○ Very low quality	Important (to know if different AMH assays are comparable)
AUC of AMH to discriminate between women with and without menses	n=56 k=1 Case series	0 Low risk of bias	NA	-1 No evidence of incremental value of AMH was provided	-1 Only one small study	No	AUC of AMH ≤ 25 pg/ml was 0.71	Comparator AUCs: AFC (total 2–10 mm) <1: 0.82 FSH ≥40IU/L: 0.72 Inhibin B ≤5 pg/ml: 0.63 AFC <1 + AMH ≤25 pg/ml: 0.87 AFC <1 + FSH ≥40IU/L: 0.87 AMH ≤25 pg/ml + FSH ≥40IU/L: 0.74	⊕○○○ Very low quality	Not important (unknown how to use this evidence in clinical practice)
Sensitivity of AMH compared to other ovarian reserve tests in distinguishing	n=56 k=1 Case series	0 Low risk of bias	NA	-1 No evidence of incremental value of AMH was provided	-1 Only one small study,	No	Sensitivity AMH ≤ 25 pg/mL 76%.	Sensitivities comparators: AFC (total 2–10 mm) < 1: 79%	⊕○○○ Very low quality	Not important (unknown how to use this evidence in clinical

Outcome	No. of participants, no. of studies and study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Results for AMH testing	Results for comparator test	Analytical validity QoE	Importance
CRA					no CI reported			FSH \geq 40IU/L: 78% Inhibin B \leq 5pg/ml: 54%		practice)
Specificity of AMH compared to other ovarian reserve tests in distinguishing CRA	n=56 k=1 Case series	0 Low risk of bias	NA	-1 No evidence of incremental value of AMH was provided	-1 Only one small study, no CI reported	No	Specificity AMH \leq 25 pg/mL 60%.	Specificities comparators: AFC (total 2–10 mm) <1: 89% FSH \geq 40IU/L: 64% Inhibin B \leq 5pg/ml: 64%	⊕○○○ Very low quality	Not important (unknown how to use this evidence in clinical practice)
Correlation between different AMH tests	n=98 k=2 Case series	0 Low risk of bias	0	-1 No evidence of incremental value of AMH was provided	-1 Low number of patients included	No	AMH Gen II ELISA versus Ansh Labs AMH ELISA: rho=0.964, p<0.0001 AMH Gen II ELISA versus Ultrasensitive AMH ELISA: rho=0.92, p<0.001 AMH Gen II ELISA vs pico-AMH ELISA: rho=0.92, p<0.001 Ultrasensitive AMH ELISA vs pico-AMH ELISA: rho = 0.99, p<0.001	NA	⊕○○○ Very low quality	Important (to know if different AMH assays are comparable)
Concordance between AMH and FSH	n=53 k=1 Case series	0 Low risk of bias	NA	-1 No evidence of incremental value of AMH was provided	-1 Low number of patients included , only one	No	The overall, positive and negative per cent agreements between low AMH and high FSH levels were 77.4%, 57.1% and 67.6%, respectively.	See results for AMH testing' column	⊕○○○ Very low quality	Not important (unknown how to use this evidence in clinical practice)

Outcome	No. of participants, no. of studies and study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Results for AMH testing	Results for comparator test	Analytical validity QoE	Importance
					study					
Correlation between AMH and AFC	n=366 k=5 Case series	0 Low risk of bias	0 All positive moderate to strong correlations	-1 No evidence of incremental value of AMH was provided	0	No	Correlation (range): Positive: 0.62 (0.44–0.83), k=4 Pre-cancer: 0.48 (0.44–0.52), k=2 Post-cancer: 0.72, k=1 Post-childhood cancer: 0.83, k=1 Pre-endo surgery: 0.842 (k=1)	See results for AMH testing' column	⊕○○○ Very low quality	Not important (unknown how to use this evidence in clinical practice)
Correlation between AMH and FSH	n=462 k=7 Case series	0 Low risk of bias	-1 Inconsistent results between studies	-1 No evidence of incremental value of AMH was provided	-1 Non-significant and inconsistent results	No	Negative: -0.50 (-0.52, -0.47; k=2), NS (k=3) Pre-cancer: -0.47, k=1, NS, k=2 Post-cancer: NS, k=1 Post-childhood cancer: -0.52, k=1 Serum AMH versus serum FSH levels: NS, k=3 Detectable AMH versus serum FHS: -0.55, k=1	See results for AMH testing' column	⊕○○○ Very low quality	Not important (unknown how to use this evidence in clinical practice)
Correlation between AMH and E2	n=331 k=4 Case series	0 Low risk of bias	-1 Inconsistent results between studies	-1 No evidence of incremental value of AMH was provided	-1 non-significant and inconsistent results	No	Positive: 0.30 (0.15–0.44), k=2 Negative: -0.64, k=1 Pre-cancer -0.64 and 0.15, k=2 Positive post-cancer: 0.44, k=1 Pre- and post-cancer: NS (k=1)	See results for AMH testing' column	⊕○○○ Very low quality	Not important (unknown how to use this evidence in clinical practice)

Outcome	No. of participants, no. of studies and study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Results for AMH testing	Results for comparator test	Analytical validity QoE	Importance
Correlation between AMH and inhibin B	n=353 k=6 Case series	0 Low risk of bias	0 All positive moderate to strong correlations	-1 No evidence of incremental value of AMH was provided	0	No	Positive: 0.72 (0.44–0.87), k=5 Pre-cancer: 0.32 (0.22–0.42), k=2 Post-cancer(4.8 years): 0.52, k=1 Post-childhood cancer: 0.44 (0.43–0.44), k=2 Positive pre- and post-cancer: 0.84, k=1	See results for AMH testing' column	⊕○○○ Very low quality	Not important (unknown how to use this evidence in clinical practice)

AFC = antral follicle count; AMH = anti-Müllerian hormone; AUC = area under the curve; CI = confidence interval; CRA = chemotherapy-related amenorrhea; E2 = estradiol; EIA = enzyme immunoassay; ELISA = enzyme-linked immunosorbent assay; FSH = follicle-stimulating hormone; LH = luteinising hormone; NA = not available; NR = not reported; QoE = quality of evidence

Table 89 Evidence profile table for the prognostic and predictive evidence of AMH testing for women undergoing gonadotoxic treatment

Outcome	No. of participants, No. of studies and study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Results for AMH testing	Results for comparator test	Predictive validity QoE	Importance
Baseline AMH levels in breast cancer patients with CRA at follow-up compared with baseline AMH levels in patients without CRA	n=555 k=7 Case control studies	-1 Moderate risk of bias	0 Most (6/7) studies showed lower baseline AMH in women with CRA at follow-up	-1 No evidence of incremental value of AMH was provided	0	No	All but one of these studies (6/7, 85%) reported a statistically significant lower baseline AMH in women who had CRA at follow-up, compared to women who continued menstruation.	Of k=4 measuring baseline FSH in addition to AMH, k=2 (total n=54) identified higher baseline FSH levels in women with CRA at follow-up. Of k=2 measuring E2 levels at baseline, k=1 (n = 21) found a higher baseline E2 level in women with a menstrual cycle at follow-up; k=3 included inhibin B in the analyses; k=1 showed baseline inhibin B was lower in women with CRA	⊕○○○ Very low quality	Not important (unknown how to use this evidence in clinical practice)

Outcome	No. of participants, No. of studies and study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Results for AMH testing	Results for comparator test	Predictive validity QoE	Importance
								at follow-up (n=21); k=2 (total n=88) found no difference.		
Prediction of CRA (multivariate analysis)	n=341 k=5 Cohort studies	0 Low risk of bias	0 In most (4/5) studies, AMH remained a significant predictor for continued or return of ovarian function	0	0	No	AMH remained a significant predictor for continued or return of ovarian function in four out of five studies. Only the study with the smallest study population did not find significant factors for predicting ovarian function.	Age and FSH were also identified as significant predictors in multivariate analysis in k=2 and k=1, respectively.	⊕⊕⊖⊖ Low quality	Important
Accuracy of AMH for predicting ovarian function (ROC)	n=127 k=2 Cohort studies	0 Low risk of bias	0	-1 No evidence of incremental value of AMH was provided	0	No	The AUC for predicting ongoing menses at 2 years follow-up was 0.90 (95%CI 0.82, 0.97) for AMH. AUC for AMH predicting amenorrhea or oligomenorrhea after 6 months was 0.86.	Follow-up at 2 years (AUC): Age: 0.88 (95%CI 0.78, 0.97) Follow-up at 6 months (AUC): AFC: 0.81	⊕⊖⊖⊖ Very low quality	Not important (no evidence of incremental value, unknown how to use this evidence in clinical practice)
Average AMH levels in patients with CRA after treatment compared with baseline AMH levels in patients without CRA	n=111 k=1 Cohort study	0 Low risk of bias	NA	-1 No evidence of incremental value of AMH was provided	0	No	Compared to women without CRA, the average AMH value in 4 women with CRA was significantly lower: 25.2 pg/mL (range <25–233.5 pg/mL) vs 179.4 pg/mL (range 96.2–334.1 pg/mL; p=0.03).	FSH was higher in women with CRA: 48.1 IU/L (range 13.3–173.7 IU/L) vs 17.4 IU/L (range, 12.2–24.7 IU/L); p=0.04.	⊕⊖⊖⊖ Very low quality	Not important (no evidence of incremental value, unknown how to use this evidence in clinical practice)
Accuracy of AMH for predicting ovarian function after gonadotoxic	n=39 k=1 Cohort study	0 Low risk of bias	NA	-1 No evidence of incremental value of AMH	-1 Due to few patients	No	AMH AUC: 0.99 (95%CI 0.97, 1.01), p <0.0001.	FSH AUC: 0.86 (95%CI 0.73, 0.98), p=0.001 E2 AUC: 0.93 (95 5CI 0.83, 1.03); p <0.001	⊕⊖⊖⊖ Very low quality	Not important (no evidence of incremental value, unknown

Outcome	No. of participants, No. of studies and study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Results for AMH testing	Results for comparator test	Predictive validity QoE	Importance
treatment (ROC)				was provided	included and only one study			Inhibin B AUC: 0.74 (95%CI 0.54, 0.94); p=0.03.		how to use this evidence in clinical practice)
Predictive value of AMH predicting pregnancy outcomes	n=395 k=5 Case control studies	-1 Moderate risk of bias	0 In most (4/5) studies, AMH did not show a relationship between AMH values and pregnancy.	-1 No evidence of incremental value of AMH was provided	-1	No	4/5 studies (n=360) did not report a statistically significant pregnancy rate difference in women with higher AMH values, compared to women with lower AMH values. One small study (n=35) reported a higher pregnancy rate in women with high AMH.	NA	⊕○○○ Very low quality	Important (no evidence of incremental value, however predicting pregnancy is a primary outcome)
Association between matured oocytes for cryopreservation and AMH levels	n=340 k=1 Cohort study	0 Low risk of bias	NA	-1 No evidence of incremental value of AMH was provided	0	No	Odds ratios and the AUCs for number of oocytes frozen (Table 50). AMH would be moderate (0.7–0.8) to good (0.8–0.9) at predicting oocyte yields above or below a threshold	Odds ratios for AFC and the AUCs for number of oocytes frozen (Table 50). AFC would have a good (0.8–0.9) accuracy for predicting an oocyte yield.	⊕○○○ Very low quality	Important (no evidence of incremental value, however possibility for change in management)
Correlation between AMH and number of oocytes retrieved	n=106 k=3	-1 Moderate risk of bias	0	-1 No evidence of incremental value of AMH was provided	-1	No	The correlations between AMH value and oocytes retrieved are between 0.45 and 0.60 in all three studies, which indicates a moderate linear relationship.	Comparison was made between AFC and oocyte retrieval in two studies, and FSH, inhibin B and E2 and oocyte retrieval in one study.	⊕○○○ Very low quality	Not important (no evidence of incremental value, unknown how to use this evidence in clinical practice)

AFC = antral follicle count; AMH = anti-Müllerian hormone; AUC = area under the curve; CI = confidence interval; CRA = chemotherapy-related amenorrhea; E2 = estradiol; EIA = enzyme immunoassay; ELISA = enzyme-linked immunosorbent assay; FSH = follicle-stimulating hormone; LH = luteinising hormone; NA = not available; NR = not reported OR = odds ratio; QoE = quality of evidence; ROC = receiving operating characteristic

Table 90 Evidence profile table for the clinical validity of AMH testing for girls or women undergoing gonadotoxic treatment

Outcome	No. of participants, No. of studies and study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Results for AMH testing	Clinical validity QoE	Importance
Clinical validity of AMH testing measured prior to gonadotoxic treatment (resumption of menses as RS)	n=96 k=3	0 Low risk of bias	0	-1 No evidence of incremental value of AMH was provided	-1 Due to few patients included and unclear CIs	No	First generation AMH tests showed lack of utility AMH measured both prior to and after treatment compared to resumption of menses/CRA. Second generation tests perform better, however the positive test result was only clinically useful in women at high risk of ovarian failure, whereas the negative test result was only useful in the group at low risk of ovarian failure.	⊕⊖⊖⊖ Very low quality	Unclear
Clinical validity of AMH testing measured after gonadotoxic treatment (resumption of menses as RS)	n=214 k=4	0 Low risk of bias	0	-1 No evidence of incremental value of AMH was provided	0	No	First generation AMH tests showed lack of utility AMH measured both prior to and after treatment compared to resumption of menses/CRA. Second generation tests perform better, however the positive test result was only clinically useful in women at high risk of ovarian failure, whereas the negative test result was only useful in the group at low risk of ovarian failure.	⊕⊖⊖⊖ Very low quality	Unclear

AMH = anti-Müllerian hormone; CI = confidence interval; CRA = chemotherapy-related amenorrhea; QoE = quality of evidence; RS = reference standard

There were no evidence profile tables for change in management or impact of change in management are provided, due to lack of evidence in the proposed study population.

FULL TEXT COULD NOT BE RETRIEVED

Sandya, M & Kumar, P 2016, 'Size of endometrioma and number does influence the ovarian reserve: A prospective observational study', *International Journal of Infertility and Fetal Medicine*, vol. 7, no. 1, 2016-1-1, pp. 14-18.

NO EXTRACTABLE DATA FOR ANALYTICAL VALIDITY SECTION OR ONLY PREDICTIVE/PROGNOSTIC DATA

Anders, C, Marcom, PK, Peterson, B, Gu, L, Unruhe, S, Welch, R, Lyons, P, Kimmick, G, Shaw, H, Snyder, S, Antenos, M, Woodruff, T & Blackwell, K 2008, 'A pilot study of predictive markers of chemotherapy-related amenorrhea among pre-menopausal women with early stage breast cancer', *Cancer Investigation*, vol. 26, no. 3, 2008-1-1, pp. 286-295.

Chai, J, Howie, AF, Cameron, DA & Anderson, RA 2014, 'A highly-sensitive anti-Müllerian hormone assay improves analysis of ovarian function following chemotherapy for early breast cancer', *European Journal of Cancer*, vol. 50, no. 14, 2014-1-1, pp. 2367-2374.

D'Avila Â, M, Biolchi, V, Capp, E & Corleta, HVE 2015, 'Age, anti-müllerian hormone, antral follicles count to predict amenorrhea or oligomenorrhea after chemotherapy with cyclophosphamide', *Journal of Ovarian Research*, vol. 8, no. 1, 2015-1-1.

Henry, NL, Xia, R, Schott, AF, McConnell, D, Banerjee, M & Hayes, DF 2014, 'Prediction of postchemotherapy ovarian function using markers of ovarian reserve', *Oncologist*, vol. 19, no. 1, 2014-1-1, pp. 68-74.

INCORRECT POPULATION

Christiansen, SC, Eilertsen, TB, Vanky, E & Carlsen, SM 2016, 'Does AMH reflect follicle number similarly in women with and without PCOS?', *PLoS ONE*, vol. 11, no. 1, 2016-1-1.

Leonhardt, H, Hellstrom, M, Gull, B, Lind, AK, Nilsson, L, Janson, PO & Stener-Victorin, E 2014, 'Ovarian morphology assessed by magnetic resonance imaging in women with and without polycystic ovary syndrome and associations with antimüllerian hormone, free testosterone, and glucose disposal rate', *Fertil Steril*, vol. 101, no. 6, 2014-1-1, pp. 1747-1756 e1741-1743.

Pigny, P, Gorisse, E, Ghulam, A, Robin, G, Catteau-Jonard, S, Duhamel, A & Dewailly, D 2016, 'Comparative assessment of five serum antimüllerian hormone assays for the diagnosis of polycystic ovary syndrome', *Fertil Steril*, vol. 105, no. 4, 2016-1-1, pp. 1063-1069 e1063.

Sahmay, S, Aydin, Y, Atakul, N, Aydogan, B & Kaleli, S 2014, 'Relation of antimüllerian hormone with the clinical signs of hyperandrogenism and polycystic ovary morphology', *Gynecol Endocrinol*, vol. 30, no. 2, 2014-1-1, pp. 130-134.

INCORRECT INTERVENTION

Yding Andersen, C, Rosendahl, M & Byskov, AG 2008, 'Concentration of anti-Müllerian hormone and inhibin-B in relation to steroids and age in follicular fluid from small antral human follicles', *J Clin Endocrinol Metab*, vol. 93, no. 6, 2008-1-1, pp. 2344-2349.

NO COMPARATOR

Jantke, A, Rendtorff, R, Dittrich, R, Müller, A, Pfitzer, C, Hohmann, C, Keil, T & Borgmann-Staudt, A 2012, 'Association between self-reported questionnaire data on fertility and results of hormone analyses in women after childhood cancer: A cross-sectional study', *Journal of Obstetrics and Gynaecology Research*, vol. 38, no. 10, 2012-1-1, pp. 1254-1259.

INCORRECT OUTCOMES

Alborzi, S, Keramati, P, Younesi, M, Samsami, A & Dadras, N 2014, 'The impact of laparoscopic cystectomy on ovarian reserve in patients with unilateral and bilateral endometriomas', *Fertil Steril*, vol. 101, no. 2, 2014-1-1, pp. 427-434.

Anderson, RA, Themmen, APN, Al-Qahtani, A, Groome, NP & Cameron, DA 2006, 'The effects of chemotherapy and long-term gonadotrophin suppression on the ovarian reserve in pre-menopausal women with breast cancer', *Human Reproduction*, vol. 21, no. 10, 2006-1-1, pp. 2583-2592.

Rosendahl, M, Andersen, CY, Ernst, E, Westergaard, LG, Rasmussen, PE, Loft, A & Andersen, AN 2008, 'Ovarian function after removal of an entire ovary for cryopreservation of pieces of cortex prior to gonadotoxic treatment: a follow-up study', *Hum Reprod*, vol. 23, no. 11, 2008-1-1, pp. 2475-2483.

Sabek, EAS, Saleh, OI & Ahmed, HA 2015, 'Ultrasound in evaluating ovarian reserve, is it reliable?', *Egyptian Journal of Radiology and Nuclear Medicine*, vol. 46, no. 4, 2015-1-1, pp. 1343-1348.

Salihoğlu, KN, Dilbaz, B, Cirik, DA, Ozelci, R, Ozkaya, E & Mollamahmutoğlu, L 2016, 'Short-Term Impact of Laparoscopic Cystectomy on Ovarian Reserve Tests in Bilateral and Unilateral Endometriotic and Nonendometriotic Cysts', *Journal of Minimally Invasive Gynecology*, vol. 23, no. 5, 2016-1-1, pp. 719-725.

Tokmak, A, Timur, H, Aksoy, RT, Çınar, M & Yılmaz, N 2015, 'Is anti-mullerian hormone a good diagnostic marker for adolescent and young adult patients with polycystic ovary syndrome?', *Türk Jinekoloji ve Obstetrik Derneği Dergisi*, vol. 12, no. 4, 2015-1-1, pp. 199-204.

EXCLUDED FROM PREDICTIVE/PROGNOSTIC SECTION

NOT PREDICTIVE/PROGNOSTIC

Behringer, K, Thielen, I, Mueller, H, Goergen, H, Eibl, AD, Rosenbrock, J, Halbsguth, T, Eichenauer, DA, Fuchs, M, Reiners, KS, Renno, JH, van der Ven, K, Kuehr, M, von Wolff, M, Diehl, V, Engert, A & Borchmann, P 2012, 'Fertility and gonadal function in female survivors after treatment of early unfavorable Hodgkin lymphoma (HL) within the German Hodgkin Study Group HD14 trial', *Annals of Oncology*, vol. 23, no. 7, 2012-1-1, pp. 1818-1825.

Dillon, KE, Sammel, MD, Ginsberg, JP, Lechtenberg, L, Prewitt, M & Gracia, CR 2013, 'Pregnancy after cancer: Results from a prospective cohort study of cancer survivors', *Pediatric Blood and Cancer*, vol. 60, no. 12, 2013-1-1, pp. 2001-2006.

El-Shalakany, AH, Ali, MS, Abdelmaksoud, AA, El-Ghany, SA & Hasan, EA 2013, 'Ovarian function in female survivors of childhood malignancies', *Pediatric Hematology and Oncology*, vol. 30, no. 4, 2013-1-1, pp. 328-335.

Giuseppe, L, Attilio, G, Edoardo, DN, Loredana, G, Cristina, L & Vincenzo, L 2007, 'Ovarian function after cancer treatment in young women affected by Hodgkin disease (HD)', *Hematology*, vol. 12, no. 2, Apr, pp. 141-147.

Greve, T, Schmidt, KT, Kristensen, SG, Ernst, E & Andersen, CY 2012, 'Evaluation of the ovarian reserve in women transplanted with frozen and thawed ovarian cortical tissue', *Fertility and Sterility*, vol. 97, no. 6, 2012-1-1, pp. 1394-+.

Loverro, G, Guarini, A, Di Naro, E, Giacomantonio, L, Lavopa, C & Liso, V 2007, 'Ovarian function after cancer treatment in young women affected by Hodgkin disease (HD)', *Hematology*, vol. 12, no. 2, 2007-1-1, pp. 141-147.

Lunsford, AJ, Whelan, K, McCormick, K & McLaren, JF 2014, 'Antimüllerian hormone as a measure of reproductive function in female childhood cancer survivors', *Fertil Steril*, vol. 101, no. 1, 2014-1-1, pp. 227-231.

Miyoshi, Y, Ohta, H, Namba, N, Tachibana, M, Miyamura, T, Miyashita, E, Hashii, Y, Oue, T, Isobe, A, Tsutsui, T, Kimura, T & Ozono, K 2013, 'Low serum concentrations of anti-müllerian hormone are common in 53 female childhood cancer survivors', *Hormone Research in Paediatrics*, vol. 79, no. 1, 2013-1-1, pp. 17-21.

Nielsen, SN, Andersen, AN, Schmidt, KT, Rechnitzer, C, Schmiegelow, K, Bentzen, JG & Larsen, EC 2013, 'A 10-year follow-up of reproductive function in women treated for childhood cancer', *Reproductive BioMedicine Online*, vol. 27, no. 2, 2013-1-1, pp. 192-200.

Rosendahl, M, Andersen, CY, Ernst, E, Westergaard, LG, Rasmussen, PE, Loft, A & Andersen, AN 2008, 'Ovarian function after removal of an entire ovary for cryopreservation of pieces of cortex prior to gonadotoxic treatment: a follow-up study', *Hum Reprod*, vol. 23, no. 11, 2008-1-1, pp. 2475-2483.

DATA NOT EXTRACTABLE

Bastings, L, Liebenthron, J, Westphal, JR, Beerendonk, CC, van der Ven, H, Meinecke, B, Montag, M, Braat, DD & Peek, R 2014, 'Efficacy of ovarian tissue cryopreservation in a major European center', *J Assist Reprod Genet*, vol. 31, no. 8, Aug, pp. 1003-1012.

Greve, T, Schmidt, KT, Kristensen, SG, Ernst, E & Andersen, CY 2012, 'Evaluation of the ovarian reserve in women transplanted with frozen and thawed ovarian cortical tissue', *Fertility and Sterility*, vol. 97, no. 6, 2012-1-1, pp. 1394-+.

Gupta, AA, Lee Chong, A, Deveault, C, Traubici, J, Maloney, AM, Knight, S, Lorenzo, A & Allen, L 2016, 'Anti-Müllerian Hormone in Female Adolescent Cancer Patients Before, During, and After Completion of Therapy: A Pilot Feasibility Study', *J Pediatr Adolesc Gynecol*, vol. 29, no. 6, 2016-1-1, pp. 599-603.

Hamre, H, Kiserud, CE, Ruud, E, Thorsby, PM & Fosså, SD 2012, 'Gonadal function and parenthood 20 years after treatment for childhood lymphoma: A cross-sectional study', *Pediatric Blood and Cancer*, vol. 59, no. 2, 2012-1-1, pp. 271-277.

Jadoul, P, Anckaert, E, D, eleer, A, Steffens, M, Dolmans, MM, Vermeylen, C, Smits, J, Donnez, J & Maiter, D 2011, 'Clinical and biologic evaluation of ovarian function in women treated by bone marrow transplantation for various indications during childhood or adolescence', *Fertil Steril*, vol. 96, no. 1, 2011-1-1, pp. 126-133 e123.

Janse, F, Donnez, J, Anckaert, E, de Jong, FH, Fauser, BC & Dolmans, MM 2011, 'Limited value of ovarian function markers following orthotopic transplantation of ovarian tissue after gonadotoxic treatment', *J Clin Endocrinol Metab*, vol. 96, no. 4, 2011-1-1, pp. 1136-1144.

Lunsford, AJ, Whelan, K, McCormick, K & McLaren, JF 2014, 'Antimüllerian hormone as a measure of reproductive function in female childhood cancer survivors', *Fertil Steril*, vol. 101, no. 1, 2014-1-1, pp. 227-231.

Morel, N, Bachelot, A, Chakhtoura, Z, Ghillani-Dalbin, P, Amoura, Z, Galicier, L, Aumaitre, O, Piette, JC, Pourrat, J, Boutin, D, Sacre, K, Kahn, JE, Duhaut, P, Farge, D, Frances, C, Guettrot-Imbert, G, Harle, JR, Lambotte, O, Le Guern, V, Sene, D, Trad, S, Vidal, E, Sarrot-Reynauld, F, Gompel, A, Tanguy, ML, Touraine, P, Lacorte, JM, Costedoat-Chalumeau, N & Grp, P 2013, 'Study of Anti-Müllerian Hormone and Its Relation to the Subsequent Probability of Pregnancy in 112 Patients With Systemic Lupus Erythematosus, Exposed or Not to Cyclophosphamide', *Journal of Clinical Endocrinology & Metabolism*, vol. 98, no. 9, 2013-1-1, pp. 3785-3792.

Ozaki, R, Kumakiri, J, Tinelli, A, Grimbizis, GF, Kitade, M & Takeda, S 2016, 'Evaluation of factors predicting diminished ovarian reserve before and after laparoscopic cystectomy for ovarian endometriomas: A prospective cohort study', *Journal of Ovarian Research*, vol. 9, no. 1, 2016-1-1.

Taniguchi, F, Sakamoto, Y, Yabuta, Y, Azuma, Y, Hirakawa, E, Nagira, K, Uegaki, T, Deura, I, Hata, K & Harada, T 2016, 'Analysis of pregnancy outcome and decline of anti-Müllerian hormone after laparoscopic cystectomy for ovarian endometriomas', *Journal of Obstetrics and Gynaecology Research*, vol. 42, no. 11, 2016-1-1, pp. 1534-1540.

INCORRECT POPULATION OR STUDY POPULATION TOO SMALL (N<10)

Ghorani, E, Ramaswami, R, Smith, RJ, Savage, PM & Seckl, MJ 2016, 'Anti-Müllerian Hormone in Patients Treated with Chemotherapy for Gestational Trophoblastic Neoplasia Does Not Predict Short-Term Fertility', *Journal of Reproductive Medicine*, vol. 61, no. 5, 2016-1-1, pp. 205-209.

Kunt, C, Ozaksit, G, Keskin Kurt, R, Cakir Gungor, AN, Kanat-Pektas, M, Kilic, S & Dede, A 2011, 'Anti-Müllerian hormone is a better marker than inhibin B, follicle-stimulating hormone, estradiol or antral follicle count in predicting the outcome of in vitro fertilization', *Archives of Gynecology and Obstetrics*, vol. 283, no. 6, 2011-1-1, pp. 1415-1421.

Lee, JR, Kim, SH, Kim, SM, Jee, BC, Ku, SY, Suh, CS, Choi, YM, Kim, JG & Moon, SY 2008, 'Follicular fluid anti-Müllerian hormone and inhibin B concentrations: comparison between gonadotropin-releasing hormone (GnRH) agonist and GnRH antagonist cycles', *Fertility and Sterility*, vol. 89, no. 4, 2008-1-1, pp. 860-867.

APPENDIX F

QUALITY APPRAISAL OF PROGNOSTIC EVIDENCE

Table 91 Quality appraisal for studies included in predictive and prognostic evidence section of B4, measured using QUIPS tool

Study	Biases (risk of bias)						Overall risk of bias
	Study participation	Study attrition	Prognostic factor measurement	Outcome measurement	Study confounding	Statistical analysis and reporting	
Anders et al. (2008)	Moderate	High	Low	Low	Moderate	Moderate	Moderate
Anderson & Cameron (2011)	Moderate	Moderate	Low	Low	Low	Low	Low
Anderson et al. (2013)	Moderate	Moderate	Low	Low	Low	Low	Low
Chai et al. (2014)	Moderate	Low	Low	Low	Low	Low	Low
D'avila et al. (2015)	Moderate	Moderate	Low	Moderate	Low	Low	Moderate
Dezellus et al. (2017)	Moderate	Low	Low	Moderate	Low	Low	Low
Hamy et al. (2016)	Low	Moderate	Moderate	Low	Moderate	Low	Moderate
Henry et al. (2014)	Low	Moderate	Low	Low	Low	Low	Low
Iwase et al. (2016)	Moderate	Low	Moderate	Low	Moderate	Low	Moderate
Lee et al. (2011)	Moderate	Low	Low	Low	Moderate	Low	Low
Lind et al. (2016)	Moderate	High	Low	Low	Moderate	Moderate	Moderate
Manno et al. (2016)	Moderate	Low	Low	Low	Moderate	Moderate	Moderate
Ozaki et al. (2016)	Moderate	Moderate	Moderate	Low	Moderate	Moderate	Moderate
Pup et al. (2014)	Moderate	Low	High ^a	Moderate	Moderate	Moderate	Moderate
Ruddy et al. (2014)	Moderate	Moderate	Low	Moderate	Low	Low	Moderate
Sonigo et al. (2016)	Moderate	Moderate	Low	Low	Low	Low	Low
Stochino-Loi et al. (2017)	Moderate	Low	High ^a	Low	Low	Low	Moderate
Su et al. (2010)	Moderate	Moderate	Low	Low	Low	Low	Low
Su et al. (2014)	Moderate	Low	Low	Low	Low	Low	Low
Takae et al. (2015)	Moderate	Moderate	Low	Low	Moderate	Moderate	Moderate
Yu et al. (2010)	Moderate	High	Low	Moderate	Low	Moderate	Moderate

^a It was not reported how AMH levels were measured, i.e. which assay/kit was used

APPENDIX G

CORRELATION DATA

Table 92 Correlation of AMH and AFC

Study	Comparison	Population	Correlation	Coefficient
<i>Pre-treatment</i>				Spearman's rank correlation
Lee et al. (2011)	DSL ACTIVE® AMH ELISA versus AFC by TVUS	41 women with breast cancer before adjuvant treatment	The was a correlation between early follicular phase serum levels of AMH and total AFC	rho=0.444, p=0.005
Paradisi et al. (2016)	AMH Gen II ELISA versus AFC by TVUS	191 cancer patients before chemo-radiotherapy and OTC and 43 controls	There was a correlation between serum levels of AMH and total AFC (2–9 mm) throughout the whole menstrual cycle in the entire study population	rho=0.52, p<0.001
Post-treatment				-
Partridge et al. (2010)	DSL ACTIVE® AMH ELISA versus AFC by TVUS	20 breast cancer survivors with continued menses after chemotherapy	The AFC and AMH levels were highly correlated	rho=0.72, p<0.0001
Post-childhood cancer				-
Nielsen et al. (2013)	EIA AMH/MIS assay versus AFC by TVUS	71 female childhood cancer survivors	There was a correlation between AMH and total AFC (2–10 mm)	rho=0.83, p<0.001
<i>Pre- surgery</i>				Pearson's bivariate correlation
Biacchiardi et al. (2011)	EIA AMH/MIS assay versus AFC TVUS	43 normo-ovulatory women aged 18–42 years who were affected by one or more ovarian endometriomas	There were positive correlations between early follicular phase serum levels of AMH and total AFC (≥3mm)	r= -0.312, p=0.038
<i>Post-surgery</i>				
Biacchiardi et al. (2011)	EIA AMH/MIS assay versus AFC TVUS	43 normo-ovulatory women aged 18–42 years who were affected by one or more ovarian endometriomas	There were positive correlations between early follicular phase serum levels of AMH and total AFC (≥3mm) 3 months post-surgery	r= -0.819, p<0.001

AFC = antral follicle count; AMH = anti-Müllerian hormone; EIA = enzyme Immunoassay; ELISA = enzyme-linked immunosorbent assay; k = number of studies; MIS = Müllerian-inhibiting substance; OTC = ovarian tissue cryopreservation; TVUS = transvaginal ultrasound

Table 93 Correlation between AMH and ovarian biopsy (primordial follicle count)

Study	Comparison	Population	Correlation	Coefficient
-------	------------	------------	-------------	-------------

Study	Comparison	Population	Correlation	Coefficient
<i>Pre-treatment</i>				Pearson's bivariate correlation
Fabbri et al. (2014)	AMH Gen II ELISA versus Primordial follicle count from biopsy	86 women with various non-gynaecological malignancies who underwent OTC	Density of primordial follicles positively correlated with AMH levels	r=0.23, p=0.03

AMH = anti-Müllerian hormone; ELISA = enzyme-linked immunosorbent assay; OTC = ovarian tissue cryopreservation

Table 94 Correlation between AMH and FSH levels

Study	Comparison	Population	Correlation	Coefficient
<i>Pre-treatment</i>				Spearman's rank correlation
Fabbri et al. (2014)	AMH Gen II ELISA versus FSH Elecsys assay	86 women with various non-gynaecological malignancies who underwent OTC	Spearman's test found no significant correlation was found between early follicular phase AMH and FSH levels	NS
Lee et al. (2011)	DSL ACTIVE® AMH ELISA versus FSH assay not reported	41 women with breast cancer before adjuvant treatment	There was a negative correlation between early follicular phase serum levels of AMH and FSH	rho= -0.47, p<0.0001
Paradisi et al. (2016)	AMH Gen II ELISA versus FSH Elecsys assay	191 cancer patients before chemo-radiotherapy and OTC and 43 controls	Spearman's test found no correlation between serum levels of AMH and FSH throughout the whole menstrual cycle in the entire study population	NS: p>0.05
<i>Post-treatment</i>				
Kim et al. (2016)	USCN AMH ELISA versus FSH Elecsys assay	32 pre-menopausal women with clinical stage III hormone receptor-positive invasive ductal breast cancer treated by neoadjuvant chemotherapy	Spearman's test found no correlation between serum levels of AMH and FSH	NS: p>0.05
<i>Post-childhood cancer</i>				
van Beek et al. (2007)	DSL ACTIVE® AMH ELISA versus Immulite 2000 FSH immunoassay	32 women treated with chemotherapy for Hodgkin's lymphoma during childhood	There was a correlation between early follicular phase serum levels of AMH and FSH	rho= -0.52, p<0.01
<i>Post-treatment</i>				Pearson's bivariate correlation
Decanter et al. (2014)	Pico-AMH ELISA versus FSH test not reported	58 samples drawn at least 3 months after the end of chemotherapy in 30 women with either breast cancer (n=13) or haematological malignancies (n=17)	No (Pearson's) correlation was found between the absolute values of AMH (when detectable) and FSH serum levels. There was a significant and independent association between detectable/not detectable AMH and serum FSH level	AMH serum levels NS: p>0.05 Detectable AMH r= -0.546, p<0.001

Study		Population	Correlation	Coefficient
Decanter et al. (2014)	EIA AMH/MIS assay versus FSH test not reported	58 samples drawn at least 3 months after the end of chemotherapy in 30 women with either breast cancer (n=13) or haematological malignancies (n=17)	Pearson's correlation test found FSH was not significantly associated to a detectable AMH level	NS: p>0.05
<i>Pre- and post-treatment</i>				
Lutchman Singh et al. (2007)	EIA AMH/MIS assay versus DPC Immuno-radiometric FSH assay	22 young women with breast cancer pre- and post-treatment and 24 controls	Pearson's correlation test showed no correlation between early follicular phase serum levels of AMH and FSH	NS: p>0.05

AFC = antral follicle count; EIA = enzyme Immunoassay; ELISA = enzyme-linked immunosorbent assay; FSH = follicle-stimulating hormone; MIS = Müllerian-inhibiting substance; NS = not significant; OTC = ovarian tissue cryopreservation;

Table 95 Correlation between AMH and E2 levels

Study		Population	Correlation	Coefficient
<i>Pre-treatment</i>				Spearman's rank correlation
Fabbri et al. (2014)	AMH Gen II ELISA versus E2 Elecsys assay	86 women with various non-gynaecological malignancies who underwent OTC	AMH levels were inversely correlated to E2 levels in the early follicular phase.	rho= -0.64, p=0.007
Paradisi et al. (2016)	AMH Gen II ELISA kit versus E2 Elecsys assay	191 cancer patients before chemo-radiotherapy and OTC and 43 controls	There was a weak correlation between serum levels of AMH and E2 throughout the whole menstrual cycle in the entire study population.	rho=0.15, p=0.044
<i>Post-treatment</i>				
Kim et al. (2016)	USCN AMH ELISA versus E2 Elecsys assay	32 pre-menopausal women with clinical stage III hormone receptor-positive invasive ductal breast cancer treated by neoadjuvant chemotherapy	There was a correlation between serum levels of AMH and E2.	rho=0.441, p<0.05
<i>Pre- and post-treatment</i>				Pearson's bivariate correlation
Lutchman Singh et al. (2007)	EIA AMH/MIS assay versus IBL E2 ELISA	22 young women with breast cancer pre- and post-treatment and 24 controls	Pearson's correlation test showed no correlation between early follicular phase serum levels of AMH and E2.	NS: p>0.05

AMH = anti-Müllerian hormone; E2 = estradiol; EIA = enzyme Immunoassay; ELISA = enzyme-linked immunosorbent assay; MIS = Müllerian-inhibiting substance; NS = not significant; OTC = ovarian tissue cryopreservation

Table 96 Correlation between AMH and inhibin B levels

Study, tests		Population	Correlation	Coefficient
<i>Pre-treatment</i>				Spearman's rank correlation
Lee et al. (2011)	DSL ACTIVE® AMH ELISA versus Inhibin B assay not reported	41 women with breast cancer before adjuvant treatment	The was a correlation between early follicular phase serum levels of AMH and inhibin B	rho=0.417, p<0.005
Paradisi et al. (2016)	AMH Gen II ELISA versus Inhibin B Gen II ELISA	191 cancer patients before chemo-radiotherapy and OTC and 43 controls	There was a correlation between serum levels of AMH and inhibin B throughout the whole menstrual cycle in the entire study population	rho=0.22, p=0.003
<i>Post-treatment</i>				
Kim et al. (2016)	USCN AMH ELISA versus BlueGene inhibin B ELISA	32 pre-menopausal women with clinical stage III hormone receptor-positive invasive ductal breast cancer treated by neoadjuvant chemotherapy	The was a correlation between serum levels of AMH and inhibin B	rho=0.515, p<0.05
<i>Post-childhood cancer</i>				
Beneventi et al. (2014)	AMH Gen II ELISA versus Inhibin B Gen II ELISA	135 female survivors treated for childhood malignant and non-malignant diseases. 92 (68.1%) had before menarche and 43 (31.85%) after menarche	Concentrations of AMH and inhibin B were strongly correlated	rho=0.44, p<0.001
van Beek et al. (2007)	DSL ACTIVE® AMH ELISA versus Serotec inhibin B ELISA	32 women treated with chemotherapy for Hodgkin's lymphoma during childhood	The was a correlation between early follicular phase serum levels of AMH and inhibin B	rho=0.43, p<0.05
<i>Pre- and post-treatment</i>				Pearson's bivariate correlation
Lutchman Singh et al. (2007)	EIA AMH/MIS assay DSL inhibin B ELISA	22 young women with breast cancer pre- and post-treatment and 24 controls	The was a correlation between early follicular phase serum levels of AMH and inhibin B	r=0.842, p<0.001

AMH = anti-Müllerian hormone; EIA = enzyme Immunoassay; ELISA = enzyme-linked immunosorbent assay; MIS = Müllerian-inhibiting substance; OTC = ovarian tissue cryopreservation

REFERENCES

Abdallah, Y, Briggs, J, Jones, J, Horne, G & Fitzgerald, C 2017, 'A nationwide UK survey of female fertility preservation prior to cancer treatment', *Hum Fertil (Camb)*, May 26, pp. 1-8.

Allegra, A, Marino, A, Volpes, A, Coffaro, F, Scaglione, P, Gullo, S & La Marca, A 2017, 'A randomized controlled trial investigating the use of a predictive nomogram for the selection of the FSH starting dose in IVF/ICSI cycles', *Reprod Biomed Online*, vol. 34, no. 4, Apr, pp. 429-438.

American Society for Reproductive Medicine Practice Committee 2015, 'Testing and interpreting measures of ovarian reserve: a committee opinion', *Fertil Steril*, vol. 103, no. 3, Mar, pp. e9-e17.

Amir, E, Freedman, O, Allen, L, Colgan, T & Clemons, M 2010, 'Defining ovarian failure in amenorrheic young breast cancer patients', *Breast*, vol. 19, no. 6, Dec, pp. 545-548.

Anders, C, Marcom, PK, Peterson, B, Gu, L, Unruhe, S, Welch, R, Lyons, P, Kimmick, G, Shaw, H, Snyder, S, Antenos, M, Woodruff, T & Blackwell, K 2008, 'A pilot study of predictive markers of chemotherapy-related amenorrhea among premenopausal women with early stage breast cancer', *Cancer Investigation*, vol. 26, no. 3, 2008-1-1, pp. 286-295.

Anderson, RA & Cameron, DA 2011, 'Pretreatment serum anti-mullerian hormone predicts long-term ovarian function and bone mass after chemotherapy for early breast cancer', *J Clin Endocrinol Metab*, vol. 96, no. 5, May, pp. 1336-1343.

Anderson, RA, Rosendahl, M, Kelsey, TW & Cameron, DA 2013, 'Pretreatment anti-Mullerian hormone predicts for loss of ovarian function after chemotherapy for early breast cancer', *Eur J Cancer*, vol. 49, no. 16, Nov, pp. 3404-3411.

Anderson, RA & Wallace, WH 2013, 'Antimullerian hormone, the assessment of the ovarian reserve, and the reproductive outcome of the young patient with cancer', *Fertil Steril*, vol. 99, no. 6, May, pp. 1469-1475.

Australian Institute of Health and Welfare 2017, 'Cancer in Australia 2017', *Cancer series no. 101. Cat. no. CAN 100*, AIHW, Canberra, p. 205 <<https://www.aihw.gov.au/getmedia/3da1f3c2-30f0-4475-8aed-1f19f8e16d48/20066-cancer-2017.pdf.aspx?inline=true>>.

Beneventi, F, Locatelli, E, Giorgiani, G, Zecca, M, Locatelli, F, Cavagnoli, C, Simonetta, M, Bariselli, S, Negri, B & Spinillo, A 2014, 'Gonadal and uterine function in female survivors treated by chemotherapy, radiotherapy, and/or bone marrow transplantation for childhood malignant and non-malignant diseases', *BJOG: An International Journal of Obstetrics and Gynaecology*, vol. 121, no. 7, 2014-1-1, pp. 856-865.

Biacchiardi, CP, Piane, LD, Camanni, M, Deltetto, F, Delpiano, EM, Marchino, GL, Gennarelli, G & Revelli, A 2011, 'Laparoscopic stripping of endometriomas negatively affects ovarian follicular reserve even if performed by experienced surgeons', *Reproductive BioMedicine Online*, vol. 23, no. 6, 2011-1-1, pp. 740-746.

Bozza, C, Puglisi, F, Lambertini, M, Osa, EO, Manno, M & Del Mastro, L 2014, 'Anti-Müllerian hormone: Determination of ovarian reserve in early breast cancer patients', *Endocrine-Related Cancer*, vol. 21, no. 1, 2014-1-1, pp. R51-R65.

Chada, M, Prusa, R, Bronsky, J, Pechova, M, Kotaska, K & Lisa, L 2003, 'Inhibin B, follicle stimulating hormone, luteinizing hormone, and estradiol and their relationship to the regulation of follicle development in girls during childhood and puberty', *Physiol Res*, vol. 52, no. 3, pp. 341-346.

Chai, J, Howie, AF, Cameron, DA & Anderson, RA 2014, 'A highly-sensitive anti-Müllerian hormone assay improves analysis of ovarian function following chemotherapy for early breast cancer', *European Journal of Cancer*, vol. 50, no. 14, 2014-1-1, pp. 2367-2374.

D'Avila, ÂM, Biolchi, V, Capp, E & Corleta, HVE 2015, 'Age, anti-müllerian hormone, antral follicles count to predict amenorrhea or oligomenorrhea after chemotherapy with cyclophosphamide', *Journal of Ovarian Research*, vol. 8, no. 1, 2015-1-1.

de Souza, FHC, da Silva, CA, Yamakami, LYS, dos Santos Trindade Viana, V, Bonfá, E & Shinjo, SK 2015, 'Reduced ovarian reserve in patients with adult polymyositis', *Clinical Rheumatology*, vol. 34, no. 10, 2015-1-1, pp. 1795-1799.

Decanter, C, Morschhauser, F, Pigny, P, Lefebvre, C, Gallo, C & Dewailly, D 2010, 'Anti-Mullerian hormone follow-up in young women treated by chemotherapy for lymphoma: preliminary results', *Reprod Biomed Online*, vol. 20, no. 2, 2010-1-1, pp. 280-285.

Decanter, C, Peigne, M, Mailliez, A, Morschhauser, F, Dassonneville, A, Dewailly, D & Pigny, P 2014, 'Toward a better follow-up of ovarian recovery in young women after chemotherapy with a hypersensitive antimullerian hormone assay', *Fertil Steril*, vol. 102, no. 2, 2014-1-1, pp. 483-487.

Demeestere, I, Simon, P, Emiliani, S, Delbaere, A & Englert, Y 2009, 'Orthotopic and heterotopic ovarian tissue transplantation', *Hum Reprod Update*, vol. 15, no. 6, Nov-Dec, pp. 649-665.

Dezellus, A, Barriere, P, Campone, M, Lemanski, C, Vanlemmens, L, Mignot, L, Delozier, T, Levy, C, Bendavid, C, Debled, M, Bachelot, T, Jouannaud, C, Loustalot, C, Mouret-Reynier, MA, Gallais-Umbert, A, Masson, D & Freour, T 2017, 'Prospective evaluation of serum anti-Mullerian hormone dynamics in 250 women of reproductive age treated with chemotherapy for breast cancer', *Eur J Cancer*, vol. 79, 2017-1-1, pp. 72-80.

Dillon, KE, Sammel, MD, Prewitt, M, Ginsberg, JP, Walker, D, Mersereau, JE, Gosiengfiao, Y & Gracia, CR 2013, 'Pretreatment antimullerian hormone levels determine rate of posttherapy ovarian reserve recovery: acute changes in ovarian reserve during and after chemotherapy', *Fertil Steril*, vol. 99, no. 2, 2013-1-1, pp. 477-483.

Everhov, AH, Bergmark, K, Smedby, KE, Hirschberg, AL & Radestad, AF 2014, 'Anti-Mullerian hormone in premenopausal women following treatment of uterine cervical cancer', *Acta Obstetrica Et Gynecologica Scandinavica*, vol. 93, no. 9, 2014-1-1, pp. 949-953.

Fabrizi, R, Macciocca, M, Melotti, C, Pasquinelli, G, Vicenti, R, Reggiani, MLB, Terribile, D & Venturoli, S 2014, 'Anti-Müllerian hormone as an ovarian reserve marker in young cancer women who undergo ovarian tissue cryopreservation', *Future Oncology*, vol. 10, no. 8, 2014-1-1, pp. 1343-1351.

Fitzgerald, O, Harris, K, Paul, R & Chambers, G 2017, *Assisted reproductive technology in Australia and New Zealand 2015*, Sydney: National Perinatal Epidemiology and Statistics Unit, the University of New South Wales Sydney.

Hamre, H, Kiserud, CE, Ruud, E, Thorsby, PM & Fosså, SD 2012, 'Gonadal function and parenthood 20 years after treatment for childhood lymphoma: A cross-sectional study', *Pediatric Blood and Cancer*, vol. 59, no. 2, 2012-1-1, pp. 271-277.

Hamy, AS, Porcher, R, Cuvier, C, Giacchetti, S, Schlageter, MH, Coussieu, C, Gronier, H, Feugeas, JP, Adoui, N, Lacorte, JM, Poirot, C, Habdous, M & Espie, M 2014, 'Ovarian reserve in breast cancer: assessment with anti-Mullerian hormone', *Reproductive BioMedicine Online*, vol. 29, no. 5, 2014-1-1, pp. 573-580.

Hamy, AS, Porcher, R, Eskenazi, S, Cuvier, C, Giacchetti, S, Coussy, F, Hocini, H, Tournant, B, Perret, F, Bonfils, S, Charveriat, P, Lacorte, JM & Espie, M 2016, 'Anti-Mullerian hormone in breast cancer patients treated with chemotherapy: a retrospective evaluation of subsequent pregnancies', *Reprod Biomed Online*, vol. 32, no. 3, 2016-1-1, pp. 299-307.

Hayden, JA, van der Windt, DA, Cartwright, JL, Cote, P & Bombardier, C 2013, 'Assessing bias in studies of prognostic factors', *Ann Intern Med*, vol. 158, no. 4, Feb 19, pp. 280-286.

Henry, NL, Xia, R, Schott, AF, McConnell, D, Banerjee, M & Hayes, DF 2014, 'Prediction of postchemotherapy ovarian function using markers of ovarian reserve', *Oncologist*, vol. 19, no. 1, 2014-1-1, pp. 68-74.

Hsu, A, Arny, M, Knee, AB, Bell, C, Cook, E, Novak, AL & Grow, DR 2011, 'Antral follicle count in clinical practice: analyzing clinical relevance', *Fertil Steril*, vol. 95, no. 2, Feb, pp. 474-479.

Iwase, A, Nakamura, T, Kato, N, Goto, M, Takikawa, S, Kondo, M, Osuka, S, Mori, M & Kikkawa, F 2016, 'Anti-Müllerian hormone levels after laparoscopic cystectomy for endometriomas as a possible

predictor for pregnancy in infertility treatments', *Gynecological Endocrinology*, vol. 32, no. 4, 2016-1-1, pp. 293-297.

Iwase, A, Nakamura, T, Nakahara, T, Goto, M & Kikkawa, F 2014, 'Assessment of ovarian reserve using anti-Müllerian hormone levels in benign gynecologic conditions and surgical interventions: A systematic narrative review', *Reproductive Biology and Endocrinology*, vol. 12, no. 1, 2014-1-1.

Kim, HA, Seong, MK, Kim, JH, Kim, YG, Choi, HS, Kim, JS, Park, IC, Jin, HO, Lee, JK & Noh, WC 2016, 'Prognostic value of anti-müllerian hormone and inhibin b in patients with premenopausal hormone receptor-positive breast cancer', *Anticancer Research*, vol. 36, no. 3, 2016-1-1, pp. 1051-1058.

Kim, YJ, Cha, SW & Kim, HO 2017, 'Serum anti-Müllerian hormone levels decrease after endometriosis surgery', *Journal of Obstetrics and Gynaecology*, vol. 37, no. 3, 2017-1-1, pp. 342-346.

Kondapalli, LA, Dillon, KE, Sammel, MD, Ray, A, Prewitt, M, Ginsberg, JP & Gracia, CR 2014, 'Quality of life in female cancer survivors: is it related to ovarian reserve?', *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation*, vol. 23, no. 2, 2014-1-1, pp. 585-592.

Lee, S, Ozkavukcu, S, Heytens, E, Moy, F, Alappat, RM & Oktay, K 2011, 'Anti-Mullerian hormone and antral follicle count as predictors for embryo/ovocyte cryopreservation cycle outcomes in breast cancer patients stimulated with letrozole and follicle stimulating hormone', *J Assist Reprod Genet*, vol. 28, no. 7, 2011-1-1, pp. 651-656.

Lind, T, Lampic, C, Olofsson, JI & Rodriguez-Wallberg, KA 2016, 'Postoperative AMH reduction is not associated with reduced fecundity two years following ovarian cyst surgery', *Gynecological Endocrinology*, vol. 32, no. 9, 2016-1-1, pp. 745-748.

Logan, S, Perz, J, Ussher, JM, Peate, M & Anazodo, A 2017, 'A systematic review of patient oncofertility support needs in reproductive cancer patients aged 14 to 45 years of age', *Psychooncology*, Jul 22.

Loren, AW, Mangu, PB, Beck, LN, Brennan, L, Magdalinski, AJ, Partridge, AH, Quinn, G, Wallace, WH, Oktay, K & American Society of Clinical, O 2013, 'Fertility preservation for patients with cancer: American Society of Clinical Oncology clinical practice guideline update', *J Clin Oncol*, vol. 31, no. 19, Jul 01, pp. 2500-2510.

Lunsford, AJ, Whelan, K, McCormick, K & McLaren, JF 2014, 'Antimullerian hormone as a measure of reproductive function in female childhood cancer survivors', *Fertil Steril*, vol. 101, no. 1, 2014-1-1, pp. 227-231.

Lutchman Singh, K, Muttukrishna, S, Stein, RC, McGarrigle, HH, Patel, A, Parikh, B, Groome, NP, Davies, MC & Chatterjee, R 2007, 'Predictors of ovarian reserve in young women with breast cancer', *Br J Cancer*, vol. 96, no. 12, 2007-1-1, pp. 1808-1816.

Manno, M, Tomei, F, Puglisi, F, Zaja, F, Metus, P, Tozzoli, R, Mascarin, M & Manno, M 2016, 'Ovarian reserve biomarkers usefulness for optimization of counselling in a public network for fertility preservation in oncological patients', *Italian Journal of Gynaecology and Obstetrics*, vol. 28, no. 5, 2016-1-1, pp. 15-22.

Miyoshi, Y, Ohta, H, Namba, N, Tachibana, M, Miyamura, T, Miyashita, E, Hashii, Y, Oue, T, Isobe, A, Tsutsui, T, Kimura, T & Ozono, K 2013, 'Low serum concentrations of anti-müllerian hormone are common in 53 female childhood cancer survivors', *Hormone Research in Paediatrics*, vol. 79, no. 1, 2013-1-1, pp. 17-21.

Monash IVF 2017, *Understanding the Egg Timer Test – what is Ovarian Reserve & AMH?*, viewed 2 November 2017, <<https://monashivf.com/understanding-egg-timer-test-ovarian-reserve-amh/>>.

MSAC 2005, *Guidelines for the assessment of diagnostic technologies*, Commonwealth of Australia, Canberra, ACT.

Munoz, M, Santaballa, A, Segui, MA, Beato, C, de la Cruz, S, Espinosa, J, Fonseca, PJ, Perez, J, Quintanar, T & Blasco, A 2016, 'SEOM Clinical Guideline of fertility preservation and reproduction in cancer patients (2016)', *Clin Transl Oncol*, vol. 18, no. 12, Dec, pp. 1229-1236.

Nelson, SM & La Marca, A 2011, 'The journey from the old to the new AMH assay: how to avoid getting lost in the values', *Reprod Biomed Online*, vol. 23, no. 4, Oct, pp. 411-420.

Nielsen, SN, Andersen, AN, Schmidt, KT, Rechner, C, Schmiegelow, K, Bentzen, JG & Larsen, EC 2013, 'A 10-year follow up of reproductive function in women treated for childhood cancer', *Reproductive BioMedicine Online*, vol. 27, no. 2, 2013-1-1, pp. 192-200.

Nyboe Andersen, A, Nelson, SM, Fauser, BC, Garcia-Velasco, JA, Klein, BM, Arce, JC & group, E-s 2017, 'Individualized versus conventional ovarian stimulation for in vitro fertilization: a multicenter, randomized, controlled, assessor-blinded, phase 3 noninferiority trial', *Fertil Steril*, vol. 107, no. 2, Feb, pp. 387-396 e384.

O'Brien, Y, Wingfield, M & Kelleher, C 2016, 'The psychological impact of ovarian reserve testing: a qualitative analysis', *Fertility and Sterility*, vol. 106, no. 3, Supplement, 2016/09/01/, p. e47.

Oktay, K & Sonmezer, M 2008, 'Chemotherapy and amenorrhea: risks and treatment options', *Curr Opin Obstet Gynecol*, vol. 20, no. 4, Aug, pp. 408-415.

Ozaki, R, Kumakiri, J, Tinelli, A, Grimbizis, GF, Kitade, M & Takeda, S 2016, 'Evaluation of factors predicting diminished ovarian reserve before and after laparoscopic cystectomy for ovarian endometriomas: A prospective cohort study', *Journal of Ovarian Research*, vol. 9, no. 1, 2016-1-1.

Pacheco, F & Oktay, K 2017, 'Current Success and Efficiency of Autologous Ovarian Transplantation: A Meta-Analysis', *Reprod Sci*, vol. 24, no. 8, Aug, pp. 1111-1120.

Papaleo, E, Zaffagnini, S, Munaretto, M, Vanni, VS, Rebonato, G, Grisendi, V, Di Paola, R & La Marca, A 2016, 'Clinical application of a nomogram based on age, serum FSH and AMH to select the FSH starting dose in IVF/ICSI cycles: a retrospective two-centres study', *Eur J Obstet Gynecol Reprod Biol*, vol. 207, Dec, pp. 94-99.

Paradisi, R, Macciocca, M, Vicenti, R, Rossi, S, Morselli-Labate, AM, Mastroroberto, M, Seracchioli, R & Fabbri, R 2016, 'New insights in the selection and management of cancer patients applicants for ovarian tissue cryopreservation', *Gynecol Endocrinol*, vol. 32, no. 11, 2016-1-1, pp. 881-885.

Partridge, AH, Ruddy, KJ, Gelber, S, Schapira, L, Abusief, M, Meyer, M & Ginsburg, E 2010, 'Ovarian reserve in women who remain premenopausal after chemotherapy for early stage breast cancer', *Fertility and Sterility*, vol. 94, no. 2, 2010-1-1, pp. 638-644.

Pup, LD, Zanet, E, Rupolo, M, Talamini, R, Tirelli, U, Mazzucato, M, Steffan, A, Zanussi, S, Doretto, P & Michieli, M 2014, 'Which tools may help physicians in female fertility prediction after autologous bone marrow transplantation for lymphoma? A pilot study', *J Chemother*, vol. 26, no. 5, 2014-1-1, pp. 293-299.

Roberts, J, Ronn, R, Tallon, N & Holzer, H 2015, 'Fertility preservation in reproductive-age women facing gonadotoxic treatments', *Curr Oncol*, vol. 22, no. 4, Aug, pp. e294-304.

Ruddy, K, O'Neill, A, Miller, K, Schneider, B, Baker, E, Sparano, J, Dang, C, Northfelt, D, Sledge, G & Partridge, A 2014, 'Biomarker prediction of chemotherapy-related amenorrhea in premenopausal women with breast cancer participating in E5103', vol. 144, no. 3, pp. 591-597, viewed 2014-1-1, <<http://onlinelibrary.wiley.com/doi/10.1002/ajco.22585>>.

Somigliana, E, B, a, N, Benaglia, L, Viganò, P, Vercellini, P & Fedele, L 2012, 'Surgical excision of endometriomas and ovarian reserve: A systematic review on serum antimüllerian hormone level modifications', *Fertility and Sterility*, vol. 98, no. 6, 2012-1-1, pp. 1531-1538.

Sonigo, C, Simon, C, Boubaya, M, Benoit, A, Sifer, C, Sermondade, N & Grynberg, M 2016, 'What threshold values of antral follicle count and serum AMH levels should be considered for oocyte cryopreservation after in vitro maturation?', *Hum Reprod*, vol. 31, no. 7, 2016-1-1, pp. 1493-1500.

Stochino-Loi, E, Darwish, B, Mircea, O, Touleimat, S, Millochau, JC, Abo, C, Angioni, S & Roman, H 2017, 'Does preoperative antimullerian hormone level influence postoperative pregnancy rate in women undergoing surgery for severe endometriosis?', *Fertil Steril*, vol. 107, no. 3, 2017-1-1, pp. 707-713 e703.

Su, HC, Haunschild, C, Chung, K, Komrokian, S, Boles, S, Sammel, MD & DeMichele, A 2014, 'Prechemotherapy antimullerian hormone, age, and body size predict timing of return of ovarian function in young breast cancer patients', *Cancer*, vol. 120, no. 23, 2014-1-1, pp. 3691-3698.

Su, HI 2010, 'Measuring ovarian function in young cancer survivors', *Minerva Endocrinologica*, vol. 35, no. 4, 2010-1-1, pp. 259-270.

Su, HI, Chung, K, Sammel, MD, Gracia, CR & DeMichele, A 2011, 'Antral follicle count provides additive information to hormone measures for determining ovarian function in breast cancer survivors', *Fertil Steril*, vol. 95, no. 5, 2011-1-1, pp. 1857-1859.

Su, HI, Sammel, MD, Green, J, Velders, L, Stankiewicz, C, Matro, J, Freeman, EW, Gracia, CR & DeMichele, A 2010, 'Antimullerian hormone and inhibin B are hormone measures of ovarian function in late reproductive-aged breast cancer survivors', *Cancer*, vol. 116, no. 3, 2010-1-1, pp. 592-599.

Su, HI, Sammel, MD, Homer, MV, Bui, K, Haunschild, C & Stanczyk, FZ 2014, 'Comparability of antimullerian hormone levels among commercially available immunoassays', *Fertility and Sterility*, vol. 101, no. 6, 2014-1-1, pp. 1766-U1352.

Takae, S, Sugishita, Y, Yoshioka, N, Hoshina, M, Horage, Y, Sato, Y, Nishijima, C, Kawamura, K & Suzuki, N 2015, 'The role of menstrual cycle phase and AMH levels in breast cancer patients whose ovarian tissue was cryopreserved for oncofertility treatment', *J Assist Reprod Genet*, vol. 32, no. 2, 2015-1-1, pp. 305-312.

Taniguchi, F, Sakamoto, Y, Yabuta, Y, Azuma, Y, Hirakawa, E, Nagira, K, Uegaki, T, Deura, I, Hata, K & Harada, T 2016, 'Analysis of pregnancy outcome and decline of anti-Müllerian hormone after laparoscopic cystectomy for ovarian endometriomas', *Journal of Obstetrics and Gynaecology Research*, vol. 42, no. 11, 2016-1-1, pp. 1534-1540.

Tobler, KJ, Shoham, G, Christianson, MS, Zhao, Y, Leong, M & Shoham, Z 2015, 'Use of anti-mullerian hormone for testing ovarian reserve: a survey of 796 infertility clinics worldwide', *J Assist Reprod Genet*, vol. 32, no. 10, Oct, pp. 1441-1448.

van Beek, RD, van den Heuvel-Eibrink, MM, Laven, JSE, de Jong, FH, Themmen, APN, Hakvoort-Cammel, FG, van den Bos, C, van den Berg, H, Pieters, R & Keizer-Schrama, S 2007, 'Anti-mullerian hormone is a sensitive serum marker for gonadal function in women treated for Hodgkin's

lymphoma during childhood', *Journal of Clinical Endocrinology & Metabolism*, vol. 92, no. 10, 2007-1-1, pp. 3869-3874.

van der Kooi, ALF, van den Heuvel-Eibrink, MM, van Noordwijk, A, Neggers, S, Pluijm, SMF, van Dulmen-den Broeder, E, van Dorp, W & Laven, JSE 2017, 'Longitudinal follow-up in female Childhood Cancer Survivors: no signs of accelerated ovarian function loss', *Human Reproduction*, vol. 32, no. 1, 2017-1-1, pp. 193-200.

van Tilborg, TC, Oudshoorn, SC, Eijkemans, MJC, Mochtar, MH, van Golde, RJT, Hoek, A, Kuchenbecker, WKH, Fleischer, K, de Bruin, JP, Groen, H, van Wely, M, Lambalk, CB, Laven, JSE, Mol, BWJ, Broekmans, FJM, Torrance, HL & group, Os 2017, 'Individualized FSH dosing based on ovarian reserve testing in women starting IVF/ICSI: a multicentre trial and cost-effectiveness analysis', *Hum Reprod*, vol. 32, no. 12, Dec 1, pp. 2485-2495.

Visser, JA, Schipper, I, Laven, JS & Themmen, AP 2012, 'Anti-Mullerian hormone: an ovarian reserve marker in primary ovarian insufficiency', *Nat Rev Endocrinol*, vol. 8, no. 6, Jan 10, pp. 331-341.

Whiting, PF, Rutjes, AW, Westwood, ME, Mallett, S, Deeks, JJ, Reitsma, JB, Leeflang, MM, Sterne, JA & Bossuyt, PM 2011, 'QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies', *Ann Intern Med*, vol. 155, no. 8, Oct 18, pp. 529-536.

Yates, AP, Roberts, SA & Nardo, LG 2012, 'Anti-Mullerian hormone-tailored stimulation protocols improve outcomes whilst reducing adverse effects and costs of IVF', *Hum Reprod*, vol. 27, no. 2, Feb, p. 629.

Yates, AP, Rustomov, O, Roberts, SA, Lim, HY, Pemberton, PW, Smith, A & Nardo, LG 2011, 'Anti-Mullerian hormone-tailored stimulation protocols improve outcomes whilst reducing adverse effects and costs of IVF', *Hum Reprod*, vol. 26, no. 9, Sep, pp. 2353-2362.

Yu, B, Douglas, N, Ferin, MJ, Nakhuda, GS, Crew, K, Lobo, RA & Hershman, DL 2010, 'Changes in markers of ovarian reserve and endocrine function in young women with breast cancer undergoing adjuvant chemotherapy', *Cancer*, vol. 116, no. 9, 2010-1-1, pp. 2099-2105.