PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Anti-Mullerian hormone (AMH) test information on Australian and New Zealand fertility clinic websites: A content analysis
AUTHORS	Copp, Tessa; Nickel, Brooke; Lensen, Sarah; Hammarberg, Karin; Lieberman, Devora; Doust, Jenny; Mol, Ben; McCaffery, Kirsten

VERSION 1 – REVIEW

REVIEWER	Bungum, Leif
REVIEW RETURNED	24-Feb-2021

GENERAL COMMENTS	In line with the increasing commercialisation of fertility treatments often challenging evidence based practice of treatment, it is important to pinpoint unjustified advertising.Tribute to the authors!

REVIEWER	Wilkinson, Jack University of Manchester, Centre for Biostatistics
REVIEW RETURNED	25-Mar-2021

GENERAL COMMENTS	The authors have undertaken a review and content analysis of AMH test information on Australian and New Zealand fertility clinic websites. For transparency I will state that I have little knowledge of qualitative methodology, including content analysis, and so my ability to identify and comment on particular strengths of this aspect of the work is limited. The work is well motivated in the introduction and the research question is important. I have some minor comments for consideration of the authors, but would consider this to be close to publishable. • The authors have a coding for “Predicts IVF response” and give an example quote stating that the test can predict number of eggs. Would any response occurring during IVF (or ICSI) be counted here? E.g. If claims were made about any of number/ quality of embryos, pregnancy outcomes, or live birth following IVF/ ICSI, would these be included here? Some elaboration of this point would be useful, and it would be similarly useful if the authors could state what outcomes were mentioned (only 9, so shouldn’t be arduous). • In the discussion, the authors suggest that there is supportive evidence of “AMH test being an indicator of ovarian reserve and that it predicts the number of eggs in an IVF cycle”. I was interested to see what this evidence was – the references provided by the authors point to two narrative reviews of AMH rather than to any of the studies supporting the specific claims, which might be preferable? Following the references a bit, I’m not sure I’m quite as compelled by the evidence as the authors and I’m not sure what

	they've stated is quite right. There appears to be some evidence from meta-analysis (although I haven't considered the risk of bias assessments of the included studies, I'm presuming the authors of those reviews and of the current study have done so) that there is a reasonably high Area under ROC curve when considering AMH against poor ovarian response (and perhaps excessive ovarian response?). But high AuROC is a bit different to saying that test is accurate (which requires good calibration (https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-019-1466-7), and in addition, the question is really whether AMH adds substantial predictive value over and above readily available patient characteristics, such as age, duration of infertility etc. Finally, it doesn't appear to be correct to say that the test has been shown to do a good job of predicting the actual 'number of eggs', which isn't the same as saying that the number of eggs is expected to fall above/ below some threshold. On the contrary, at least one study suggested that only a modest portion of variation in number of oocytes following COS is explained even when we consider a large number of variables in combination (Rustamov - https://doi.org/10.1093/hropen/hox018 , with a disclaimer that I co-authored this and am not implying it should be cited). • A lot of readers might not understand what it means for the test to be 'predictive' and might mistakenly think that the fact that AMH is 'significantly associated' with outcomes in some studies demonstrates this. It might be worth inserting a line to correct that misapprehension, which might otherwise cause the premise of the article to be queried. I wouldn't be surprised if some clinicians thought that the test was predictive on the basis of studies showing that it is 'a significant predictor'. • Is it worth mentioning that a key concern is that the test is being used as a way to move healthy individuals into unnecessary fertility care? • The authors could consider adding a simple Figure to show which types of information appeared in combination – clinics on the rows (or columns) and codes on the columns (rows), similar to Figure 2 in Goodman (https://doi.org/10.1111/ajo.13126). But as stated above, I believe this can be published with minor corrections.

REVIEWER	Yuksel, Nese University of Alberta, Canada
REVIEW RETURNED	12-Apr-2021

GENERAL COMMENTS	Thank you for the opportunity to review this manuscript. In this manuscript the authors describe a content analysis of the information on the anti-Mullerian Hormone (AMH) testing found on websites from accredited fertility clinics in Australia and New Zealand. Overall the manuscript is well written, easy to follow and provides interesting findings about the utility of AMH testing that is provided on these websites which is not supported by evidence. I have a few comments to help with future revisions and clarity. Abstract: Design - indicate written information found online (or websites) Introduction:

	p3, line 52 - Did you also look at direct to consumer tests on the websites or is this any information on AMH testing? Methods: p3, lines 59: How were the websites identified/accessed - some details on how the search to identify the fertility clinic websites would be helpful - from the results looks like from the Fertility Society of Australia. Please provide this in the methods. p3, line 60: Please include the inclusion criteria for websites - website had to include mention of the AMH test. If there was no mention of the AMH test on any of the webpages then the website was excluded. This is not clear in the methods (however mentioned in results). p4, line 5 - Were there other exclusions? ie websites that required registration to access information Missing information in methods (based on findings in results): Counting of webpages accessed Assessing promotion or persuasive language on webpages Results: p4, lines 48 - 56 : may be helpful to have flowchart as a figure summarizing the inclusion/exclusion of websites. p4, lines 59 - 60: do you have any data on website/clinic characteristics? For example location of clinics, last update of website etc. p9, lines 44 - 50 - I feel the promotion/persuasion could be teased out a bit more in your findings (could even be a section in itself instead as “additional observations”). This is interesting findings which could be brought out more. Were the websites offering their own AMH testing? Or was this direct to consumer testing? It should also be included as part of your methods. p9, lines 52 - 60 - the conflicting information could be discussed within each of the categories above. In addition, these sentences sound too general for results - do you have the numbers of websites with contradictions for each of these? for example "contradicting statement about whether the AMH test can predict menopause" - only 2 of the websites indicated that it can predict, did the others indicate that it can not predict menopause or was there no mention at all? P9, line 57 - how do you define confusing in this context? Do you mean conflicting information only? Discussion: P11, line 10 - “Our findings of poor quality information” - indicators of “quality” of information/websites were not really looked at in your study. Yes the information provided was misleading and varied but not sure if you can say poor quality information without a validated measure. P11, lines 32 - 36 - other limitations often reported for internet studies include how consumers/women would interpret the data (ie other studies would be needed for this)

	General comments: I am unclear if the AMH testing found on websites is direct to consumer testing that fertility clinics are selling or if AMH testing is being done by the fertility clinics. Also is AMH testing associated with a cost by the consumer in these fertility clinics. There may be different drivers for promotion depending if there is a cost to the consumer and if this is charged by the fertility clinic. Some of these details should be brought out in the introduction as well as in the discussion when discussing promotion or persuasive language.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

In line with the increasing commercialisation of fertility treatments often challenging evidence based practice of treatment, it is important to pinpoint unjustified advertising. Tribute to the authors!

RESPONSE: Thank you.

Reviewer 2

The authors have undertaken a review and content analysis of AMH test information on Australian and New Zealand fertility clinic websites. For transparency I will state that I have little knowledge of qualitative methodology, including content analysis, and so my ability to identify and comment on particular strengths of this aspect of the work is limited. The work is well motivated in the introduction and the research question is important. I have some minor comments for consideration of the authors, but would consider this to be close to publishable.

The authors have a coding for “Predicts IVF response” and give an example quote stating that the test can predict number of eggs. Would any response occurring during IVF (or ICSI) be counted here? E.g. If claims were made about any of number/ quality of embryos, pregnancy outcomes, or live birth following IVF/ ICSI, would these be included here? Some elaboration of this point would be useful, and it would be similarly useful if the authors could state what outcomes were mentioned (only 9, so shouldn’t be arduous).
RESPONSE: Great suggestion thank you. The majority of these websites (n=6) refer to the number of eggs collected in a cycle, however a few (n=3) included vague statements that could easily be inferred to mean live birth (the goal for patients) e.g. “To determine a patient’s response to fertility medication”, “The AMH level in conjunction with a pelvic morphology scan is seen as a good predictor of IVF success”, “It can also help a fertility specialist determine whether a woman is a good candidate to undergo certain fertility treatments and how successful those treatments may be”. We have now added a more

Sydney School of Public Health T +61 2 9351 7789 ABN 15 211 513 464
128C, Edward Ford Building (A27) E tessa.copp@sydney.edu.au CRICOS 00026A

The University of Sydney
NSW 2006
Australia

detailed description of this category and added an additional quote (please see track changes in Table 1 and page 9).

Table 1: “Indicates response to fertility medication (e.g. number of eggs collected, treatment/IVF success)”
Results (p.9): “Other recurring statements included that the test indicates response to fertility treatment (e.g. number of eggs collected (n=6) or vague treatment success statements (n=3) e.g. “…a good predictor of IVF success”; 36%)”

In the discussion, the authors suggest that there is supportive evidence of “AMH test being

an indicator of ovarian reserve and that it predicts the number of eggs in an IVF cycle”. I was interested to see what this evidence was – the references provided by the authors point to two narrative reviews of AMH rather than to any of the studies supporting the specific claims, which might be preferable? Following the references a bit, I’m not sure I’m quite as compelled by the evidence as the authors and I’m not sure what they’ve stated is quite right. There appears to be some evidence from meta-analysis (although I haven’t considered the risk of bias assessments of the included studies, I’m presuming the authors of those reviews and of the current study have done so) that there is a reasonably high Area under ROC curve when considering AMH against poor ovarian response (and perhaps excessive ovarian response?). But high AuROC is a bit different to saying that test is accurate (which requires good calibration (https://protect-au.mimecast.com/s/G5eDCP7LAXfK5MDRvI0cRGu?domain=bmcmedicine.biomedcentral.com), and in addition, the question is really whether AMH adds substantial predictive value over and above readily available patient characteristics, such as age, duration of infertility etc. Finally, it doesn’t appear to be correct to say that the test has been shown to do a good job of predicting the actual ‘number of eggs’, which isn’t the same as saying that the number of eggs is expected to fall above/ below some threshold. On the contrary, at least one study suggested that only a modest portion of variation in number of oocytes following COS is explained even when we consider a large number of variables in combination (Rustamov - https://protect-au.mimecast.com/s/eCjmCQnMBZfkNrDpBsMsGLm?domain=doi.org , with a disclaimer that I co-authored this and am not implying it should be cited).
RESPONSE: Thank you for raising this issue. We have now edited this section to a) better reflect the conflicting, mixed evidence, b) edited our language to reflect an association not necessarily meaning the test is accurate or predictive of exact egg numbers, and c) added a sentence to highlight the key question in whether AMH adds substantial predictive value over and above readily available patient characteristics (please see track changes in discussion and below):
“ Importantly, whilst a number of statements about the utility of the test were made across a number of websites, few are supported by high-level evidence. Statements for which there is some supporting evidence include the AMH test being an indicator of ovarian reserve10 in terms of egg quantity and it being associated with the number of eggs obtained in an IVF cycle,9 24 25 although large variation in ovarian response remains unexplained.25 Statements with mixed evidence include low AMH levels indicating increased risk of miscarriage.26 27 There is preliminary evidence that high levels of AMH indicate PCOS,28 29 however more research is needed to confirm this and current PCOS guidelines recommend against using AMH as a diagnostic tool.30 Statements refuted by existing evidence include the test being able to predict a woman’s future fertility potential or current fertility status,13-15 31 or identifying a woman at risk of early menopause.6 Furthermore, it is important to note that although the AMH may be associated with outcomes at a population level, this does not necessarily mean it has predictive value for individuals. For example, whilst the AMH appears to be associated with age of menopause at a population level, the huge individual variation, imprecision in estimates and limited capacity in predicting the extreme ages of menopause (e.g. it cannot identify those at risk

of early menopause) means its clinical applicability in individual women is limited.32 Questions have also been raised about whether AMH adds substantive predictive value over and above readily available patient characteristics, such as age.9 34”

A lot of readers might not understand what it means for the test to be ‘predictive’ and might mistakenly think that the fact that AMH is ‘significantly associated’ with outcomes in some studies demonstrates this. It might be worth inserting a line to correct that misapprehension, which might otherwise cause the premise of the article to be queried. I wouldn’t be surprised if some clinicians thought that the test was predictive on the basis of studies showing that it is ‘a significant predictor’.

RESPONSE: Great point, we have now inserted a line to correct this misapprehension (please see track changes on page 10 and below):
“Furthermore, it is important to note that although the AMH may be associated with outcomes at a population level, this does not necessarily mean it has predictive value for individuals. For example, whilst the AMH appears to be associated with age of menopause at a population level, the huge individual variation, imprecision in estimates and limited capacity in predicting the extreme ages of menopause (e.g. it cannot identify those at risk of early menopause) means

its clinical applicability in individual women is limited.31”

Is it worth mentioning that a key concern is that the test is being used as a way to move healthy individuals into unnecessary fertility care?
RESPONSE: Great suggestion, we have now expanded on our sentence regarding overdiagnosis/overtreatment in the discussion to explicitly state this point (page 11):
“This could in turn increase women’s perceived need to freeze their eggs,32 try to conceive earlier than they had planned, or pursue fertility treatments when it may not be needed, increasing the risk of healthy individuals receiving unnecessary fertility care.33”

The authors could consider adding a simple Figure to show which types of information appeared in combination – clinics on the rows (or columns) and codes on the columns (rows), similar to Figure 2 in Goodman (https://protect-au.mimecast.com/s/AvvbCRON3D2uvgBK1nSOz4_I?domain=doi.org).
RESPONSE: Thank you for your suggestion. However, as there was 25 websites and 29 categories across the four overarching categories, a potential figure capturing this would be very large/cover over two pages. Our manuscript also does not discuss the types of information that appeared in combination or which websites performed more poorly/better, but instead focuses on the number of websites that contained certain types of information. As such, we would prefer to not include a Figure like this in the manuscript.

Reviewer 3.
Thank you for the opportunity to review this manuscript. In this manuscript the authors describe a content analysis of the information on the anti-Mullerian Hormone (AMH) testing found on websites from accredited fertility clinics in Australia and New Zealand. Overall the manuscript is well written, easy to follow and provides interesting findings about the utility of AMH testing that is provided on these websites which is not supported by evidence. I have a few comments to help with future revisions and clarity.

Abstract: Design - indicate written information found online (or websites)

RESPONSE: Thank you, now specified (page 2).

“Design: Content analysis of online written information about the AMH test on fertility clinic websites.”

Introduction: p3, line 52 - Did you also look at direct to consumer tests on the websites or is this any information on AMH testing?

RESPONSE: We looked at any information about AMH and AMH testing specifically. We have now made a minor edit in the introduction to clarify this (please see page 3 and below).
“Considering the popular narrative that the AMH test can predict fertility, the aim of this study was to systematically record and categorise any written information about the AMH test found on Australian and New Zealand fertility clinic websites.”

Methods: p3, lines 59: How were the websites identified/accessed - some details on how the search to identify the fertility clinic websites would be helpful - from the results looks like from the Fertility Society of Australia. Please provide this in the methods.

RESPONSE: We have now added this information to the methods (please see page 4 and below):

“Setting: Accredited fertility clinics in Australia and New Zealand were identified from the list of accredited practices on the Fertility Society of Australia’s website.23 The websites of those clinics were accessed between April and June, 2020.”

p3, line 60: Please include the inclusion criteria for websites - website had to include mention of the AMH test. If there was no mention of the AMH test on any of the webpages then the website was excluded. This is not clear in the methods (however mentioned in results). RESPONSE: Thank you for highlighting this omission. We have now added this in (please see track changes on page 4):

“All webpages that mentioned the AMH test, including posts or blogs specifically about the AMH test which had been posted since 2015 were scrutinised. Analysis was restricted to written context (i.e. videos and non-text data were excluded). Any webpages described as being specifically for clinicians (e.g. GPs) were also excluded. Websites that did not mention the AMH test were excluded from further analyses.”

p4, line 5 - Were there other exclusions? ie websites that required registration to access information

RESPONSE: There were no other exclusions other than those described (non-written text data (e.g. videos), blogs/posts prior to 2015, webpages explicitly for GPs). No websites required registration to access information.

Missing information in methods (based on findings in results): Counting of webpages accessed, Assessing promotion or persuasive language on webpages

RESPONSE: We have now added this information to the analysis section of the Methods (p.4).

“The analysis involved an iterative process with five members of the study team. After the number of eligible fertility clinic websites were ascertained and the data were extracted by one researcher (TC), content analysis was used to map out the areas of content that emerged and record and categorise the statements made about the AMH test, as well as additional observations.”

Results: p4, lines 48 - 56 : may be helpful to have flowchart as a figure summarizing the inclusion/exclusion of websites.

RESPONSE: Good suggestion thank you, we have now included a figure flowchart to summarise the inclusion/exclusion of websites (please see Figure 1, page 5).

p4, lines 59 - 60: do you have any data on website/clinic characteristics? For example location of clinics, last update of website etc.
RESPONSE: No we did not seek to also capture and categorise this information, aside from when assessing the eligibility of blog posts (excluded if posted prior to 2015). Most accredited fertility clinics were in a number of locations, captured on the one website e.g. City Fertility has clinics in Brisbane City, Brisbane Soutside, Gold Coast, Melbourne, Bundoora VIC, Liverpool NSW, Sydney NSW, Miranda NSW.

p9, lines 44 - 50 - I feel the promotion/persuasion could be teased out a bit more in your findings (could even be a section in itself instead as “additional observations”). This is interesting findings which could be brought out more.

RESPONSE: We labelled these ‘additional observations’ as, unlike the overarching themes, they are not a type of information per se but more an indication of how language is used to promote the test and/or peruade women to use it. We have now added in additional subheadings to give them further emphasis (please see track changes in results page 9).

Were the websites offering their own AMH testing? Or was this direct to consumer testing? It should also be included as part of your methods.
RESPONSE: Unlike online companies selling the test direct to consumers (where you can order and pay for the test online), fertility clinic websites, if specified, suggested readers order the test from their GP or fill out their contact form to make an appointment with one of their specialists to get referred for the test. We have now added a sentence to the introduction (page 3) to clarify this.

“In Australia, AMH testing can occur in several ways, although women are predominantly referred by their GPs or fertility specialists to get the test from pathology laboratories or fertility clinics with in-house pathology. The test is not covered by Australia’s universal health scheme and has out-of-pocket costs.”

p9, lines 52 - 60 - the conflicting information could be discussed within each of the categories above. In addition, these sentences sound too general for results - do you have the numbers of websites with contradictions for each of these? for example "contradicting statement about whether the AMH test can predict menopause" - only 2 of the websites indicated that it can predict, did the others indicate that it can not predict menopause or was there no mention at all?
RESPONSE: Thank you for your suggestions. We would prefer to keep the finding of conflicting and confusing information under “additional observations” as this emerged uring coding, as well as to emphasise the mixed information across and within websites. We have now added the number of websites which explicitly contained the specified information, for example, the websites that either explicitly stated the test could predict menopause or stated it could not predict menopause (please see track changes on pages 9-

10 and below):

“There were also a number of contradictions in the information provided across the websites. These included contradicting statements about whether the AMH test can (n=2) or cannot predict

menopause (n=1), is an indicator (n=1) or is not an indicator of egg quality (n=9), whether the results need to be interpreted by a specialist (n=2) or by a GP (n=3) , and whether the test is reliable (n=6) or can be artificially lower when using oral contraception (n=5). There was even conflicting, ambigious and confusing statements within the same website on three of the websites (12%), with the most common being whether or not the blood sample can be taken whilst using oral contraception and whether the test assesses women's fertility (e.g. "…not a measure of fertility but an important tool in assessing potential fertility" and then in the next paragraph "an AMH test can assess your current fertility")."

P9, line 57 - how do you define confusing in this context? Do you mean conflicting information only? RESPONSE: In addition to conflicting, some of the statements were also very unclear or ambigious e.g. "It's important to understand that ovarian (or egg) reserve is not a measure of fertility but an important tool in assessing potential fertility". We have now added this clarification to the text (page 10):

"There was even conflicting, ambigious and confusing statements within the same website on three of the websites (12%), with the most common being whether or not the blood sample can be taken whilst using oral contraception and whether the test assesses women's fertility (e.g. "…not a measure of fertility but an important tool in assessing potential fertility" and then in the next paragraph "an AMH test can assess your current fertility").

Discussion: P11, line 10 - "Our findings of poor quality information" - indicators of "quality" of information/websites were not really looked at in your study. Yes the information provided was misleading and varied but not sure if you can say poor quality information without a validated measure.

RESPONSE: Although we consider accuracy of information to be an important indicator of quality, we have edited the wording used in the discussion for consistency (please see page 11).

"Our findings of misleading or inaccurate information on fertility clinic websites are similar to recent studies evaluating the quality of website information regarding oocyte cryopreservation and of various interventions used in addition to standard IVF procedures.20 35"

P11, lines 32 - 36 - other limitations often reported for internet studies include how consumers/women would interpret the data (ie other studies would be needed for this) RESPONSE: Great point, thank you. We have now added this to the limitations (please see track changes on page 11).

"A limitation of the study is that it is unclear how consumers would interpret the information. Future studies are needed to assess how women interpret and respond to the information captured."

General comments: I am unclear if the AMH testing found on websites is direct to consumer testing that fertility clinics are selling or if AMH testing is being done by the fertility clinics. Also is AMH testing associated with a cost by the consumer in these fertility clinics. There may be different drivers for promotion depending if there is a cost to the consumer and if this is charged by the fertility clinic. Some of these details should be brought out in the introduction as well as in the discussion when discussing promotion or persuasive language. RESPONSE: Apologies for the confusion. Unlike online websites selling the test direct-to-consumers (e.g. see: https://www.i-screen.com.au/tests/amh-test), none of the accredited fertility clinic websites had the option of paying for and ordering the test online. As discussed above, the majority prompted consumers to make an appointment with one of their

specialists or request the test from their GP. There is a cost associated with the AMH test in Australia, as well as in some locations in New Zealand. We have now added this to the discussion (please see page 11 and below):

“Whilst many clinics do not receive direct financial benefit from ordering the test, clinics would benefit from the outlined potential actions as a result of women getting the test result, such as seeing a fertility specialist, egg freezing or commencing fertility treatment”

We believe that the modifications described above have strengthened the manuscript. We hope that our adjustments make the revised paper suitable for publication in Patient Education and Counselling and we look forward to your decision.

VERSION 2 – REVIEW

REVIEWER	Wilkinson, Jack University of Manchester, Centre for Biostatistics
REVIEW RETURNED	15-Jun-2021

GENERAL COMMENTS	Authors have addressed the minor comments on the initial submission in a satisfactory manner. A figure could have been useful but I don't think the omission is disqualifying. Congratulations to the authors.

REVIEWER	Yuksel, Nese University of Alberta, Canada
REVIEW RETURNED	11-Jun-2021

GENERAL COMMENTS	The authors have adequately addressed all of the reviewer comments and have made all of the revisions. Congratulations to the authors - the manuscript looks really good. I have no further suggestions.