Reviewer’s report

Title: Development and validation of a clinical prediction rule to identify suspected breast cancer: a prospective cohort study

Version: 2

Date: 10 June 2014

Reviewer: ELIZABETH SHEPHARD

Reviewer’s report:

Summary of the study: This is a derivation/validation study of 7501 women with breast symptoms referred to a symptomatic breast unit (SBU) in a localised Irish population. The study aimed to develop a clinical prediction rule (CPR) to identify predictors of breast cancer in symptomatic women. The CPR identified increasing age by year, presence of a lump, nipple change and nipple discharge as independent clinical predictors of breast cancer. Incidence of breast cancer rose in the presence of an increasing number of these variables. The CPR model adequately predicted the rate of breast cancer diagnoses in the validation cohort, indicating that the model was a good fit. The incidence of breast cancer was approximately that of the national population.

Summary opinion: There is a dearth of research on the symptoms of cancer in primary care. This paper uses data recorded by primary care GPs of women with breast symptoms which resulted in referral to a SBU in Ireland. Any study investigating the symptoms of cancer is important and welcomed.

The paper is concise, has a clear aim, contains appropriate explanations, good statistics and is very nicely written. The sample size was good. The emphasis for this paper is developing an evidence-based CPR using an Irish population. Generalisability was limited because variables seen to be associated in other studies were not measurable here – as the authors acknowledge. The limitations are detailed below:

1 - This study is similar to that of McCowan et al, who also used a derivation/validation design to develop a clinical prediction rule for diagnosing breast cancer (TF is an author on both papers. The McCowan et al study was cited in this paper). Both studies sought to identify evidence based predictors of breast cancer using their local population: Scottish data for McCowan et al and Irish data in this study. Both studies found similar results but the McCowan study contained some clinical features which were not available for coding in the present study, a disadvantage for this study. The McCowan et al CPR was validated using a different source, whereas validation in this study came from the same data source, albeit from different time periods. This has implications for the
analysis assessing whether the model was a good fit – any idiosyncrasies in the derivation set may also be present if validation uses that same dataset. Validation to another data source would have been more appropriate: indeed cross-validation would have been ideal. It is, therefore, not a surprise that the model was found to be a good fit. This limitation was not adequately addressed in the discussion.

2 - The authors note that some symptoms (and test results) were not available for coding and this is something which could be expanded more, perhaps with reference to other studies. There are very few primary care breast cancer studies published but it would be interesting to have a sentence about the clinical features recorded in other studies(1-3), other than the McCowan et al study which is already discussed. A recent study identifying and quantifying the risk of breast cancer in symptomatic women in primary care has just been accepted for publication by the BJGP. The authors are welcome to contact Prof Willie Hamilton (my colleague) if a copy could be of use.

In conclusion: In light of the fact that some clinically relevant features (breast thickening and nipple retraction) are absent from coding and the data is taken from one geographical area, generalisability is restricted - but this is acknowledged. The results and impact of this study are, however, relevant and of use to identify breast cancer in women with breast symptoms. The predictors are relevant to what is coded in the Irish system. I recommend publication with some amendments.

Suggested amendments:

Introduction – mention clinical features of breast cancer found in other primary care studies (see references below). (Discretionary)

Discussion – context section: Line 240 ‘Irish national incidence of 5.6 %’ – can you clarify this sentence? It reads as if that figure is the annual Irish incidence of female breast cancer. (Essential)

Discussion – strengths/weaknesses section: Add a sentence acknowledging the implication of using the same data source to validate the derivation predictors and the impact that this could have on the model of fit. (Essential)

Discussion – how the clinical predictors relate to those features found in previous studies. (Discretionary).

Some minor spelling/grammar mistakes are also detailed below.

Line 61 – spelling mistake ‘patents’.
Line 62 – ‘In the Irish context…’ Reword this sentence for clarity.
Line 74 – grammar – drop the second ‘are’.
Line 267-8 – font for ‘in our model’ and ‘pathologic discharge’ is different to the rest of the text.
Line 284 – add either ‘from’ or ‘of’ between ‘data over’.
References

1. Eberl MM, Phillips RL, Jr, Lamberts H, Okkes I, Mahoney MC. Characterizing Breast Symptoms in Family Practice


Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I declare that I have no competing interests