Development of a risk prediction model for endometrial carcinoma among postmenopausal women in the Western province of Sri Lanka

Withanage Iresha Udayangani Jayawickrama[1], Chrishantha Abeysena[2]

(Index words: endometrial cancer, high-risk population, risk prediction model, postmenopausal women)

Abstract

Introduction: Globally, endometrial carcinoma is the most common reproductive tract cancer among women. Risk prediction model is a simple, low-cost tool to identify women with increased risk of developing endometrial carcinoma.

Objectives: The aim of the study was to develop a model to predict the risk of endometrial carcinoma among postmenopausal women in Sri Lanka.

Methods: A case control study was conducted. The cases and the controls were defined as postmenopausal women who had and had not been diagnosed as endometrial carcinoma based on histological confirmation respectively. Variable selection was done considering the objectivity and feasibility of the measurements in addition to the statistical criteria. A scoring system [0-9] was designed based on weighted score of each risk predictor. Predictive validity of the model was tested by calibration and discrimination. Receiver Operator Characteristic (ROC) curve was used to determine the cut-off value.

Results: The developed model consisted of six predictors; Age >55 years, never conceived, age at menarche ≤11 years, ever experienced postmenopausal bleeding, having family history of any type of cancer among first degree relative, generalized obesity. Discrimination of the model was measured by the area under the ROC curve (0.92, 95% Confidence Interval: 0.88-0.95). Calibration with goodness of fit by Hosmer and Lemeshow test (p=0.72) was satisfactory. The tool demonstrated a good predictive ability with sensitivity of 79.5% (CI: 68.9%-87.3%) and specificity of 90.7% (CI: 86.8%-93.5%) at the cut-off point of 4.5.

Conclusions: Model demonstrated good discrimination and well calibration. It can be used in screening of high-risk women for developing endometrial carcinoma.

Introduction

Endometrial carcinoma (EC) is generally diagnosed at an early stage (nearly 80%) in developed countries due to their early presentation to a health care facility [1]. Early-stage endometrial carcinoma shows a good prognosis; therefore, the importance of early detection of the disease is felt vital. The five-year survival of stage I, II, III and IV of the disease are nearly 85%, 70-75%, 45% and <30% respectively [2]. Though, it is hardly found any evidence to show the proportion of women diagnosed at an early stage and their prognosis in developing countries.

The most common presentation of endometrial carcinoma patients is postmenopausal bleeding and it is defined as an episode of bleeding, 12 months after the last menstrual period. Evidence shows that 10% of the women with post-menopausal bleeding have EC as the diagnosis [3,4]. Transvaginal Ultrasound Scanning (TVS) provides the best screening test for early detection of EC [5]. With its introduction, the survival of women with endometrial carcinoma has improved dramatically over the years [6]. However, the use of TVS is limited only to symptomatic women. Thus, establishing an organized EC screening programme in a country is less feasible due to financial, human resource and infrastructure shortages [7], but screening of only high-risk women may be an alternative for the resource limited countries.

Cancer risk prediction provides an estimate of the risk for developing cancer and can be used in reducing the disease burden. Prediction of cancer risk using a risk prediction model can help to identify individuals at high risk for any cancer, and follow them up closely with periodic screening and counsel them on behavioural changes to reduce risk. It is a less invasive and low-cost mechanism to reduce morbidity and mortality related to EC with early detection of the cancer [8].

Risk prediction models are developed based on several risk factors associated with individual characteristics that are proven to be linked with the health event of interest [9]. A comprehensive analysis of risk factors should be ideally carried out by conducting a prospective study i.e. cohort study, however, long follow up and less feasibility discourage the prospective studies and suggest retrospective studies i.e. case control study designs in developing risk prediction models [10-14]. Predictors can vary from subject characteristics such as age and sex, history, physical examination, to imaging, electrophysiology, blood, urine or even genetic markers like advanced techniques.

There are several risk predictions models available to predict the risk of a woman developing endometrial carcinoma, thus the risk predictors included in the models and weight assigned for each predictor by the model are specific for the setting in which the models were developed limit its common use in other settings [4,13,15,16]. Most of the models have been developed to predict the risk of endometrial carcinoma among symptomatic women with a combination of risk factor assessment, ultrasound imaging and presence of biomarkers. A limitation inherent to these tools is that it has been developed based on their country-specific risk factors where it cannot be applied in another population without doing a proper external validation. None of these available risk prediction models were developed or validated in developing countries. Therefore, the objective of this study was to develop a country specific model to predict the risk of endometrial carcinoma among postmenopausal women in Sri Lanka.

Methods

A hospital-based unmatched case-control study was carried out from September 2016 to March 2017 in the Western Province. The required sample size was calculated for several common risk factors of endometrial carcinoma and the largest calculated value was selected as the required sample size for the study. The values for odds ratios (OR) for different risk factors, their community prevalence was based on the available literature. Case to control ratio was taken as 1:4. For the calculation of the sample size, power was taken as 80%. The largest sample size calculated was 79 and adding 5% for non-response, the study included histologically confirmed 83 newly diagnosed EC cases and 332 controls, confirmed as not having EC by performing hysteroscopy, endometrial biopsy or curettage, from fourteen secondary and tertiary hospitals in the government sector where a Consultant Obstetrician and Gynaecologist and a Consultant Pathologist were available. Details of the methods have been published elsewhere [17]. Outcome of interest was defined as whether the patient was diagnosed of EC with histological confirmation. Assessment of Outcome variable was blinded to information about predictors as the disease was already assessed before commencing the study.

An interviewer-administered questionnaire was used to assess potential risk predictors. Firstly, potential risk factors consistently identified in the literature [17-24] and secondly, few additional risk predictors with the consensus of a panel of experts were included. Information on socio-demographic factors, reproductive factors, lifestyle related factors, biological factors, genetic factors and co-morbid factors were obtained with verification through medical records when necessary. Trained five pre-intern medical officers collected data. Blinding of data collectors to the outcome while collecting predictor data was not done.

Quality of diet was assessed into optimal or suboptimal consumption of energy dense food, food containing dietary fibre and anti-oxidants using validated Food Frequency Questionnaire. Overconsumption of energy dense food and inadequate consumption of food containing dietary fibre and anti-oxidants were considered as sub-optimal consumption. Physical activity was defined as life time total physical activity in terms of occupational, household, sports and exercise activities and assessed by calculating the average Metabolic Equivalent of Task (MET) value for total physical activities for a week during a year by validated Life Time Total Physical Activity Questionnaire. The Metabolic Equivalent values less than 25th percentile was considered as low, more than 75th percentile was high and in between 25th to 75th percentile was taken as average physical activity in life. Exposure to electromagnetic field was defined as a distance less than 100m from home to high tension wires and telecommunication towers. Exposure to outdoor air pollution was defined as having the house within 100m from main road or having a large industry within one km. Long term illness was defined as diagnosis of diabetes mellitus, hypertension and hyperlipidaemia. Generalized obesity was categorized based on Body Mass Index (BMI) according to the World Health Organization (WHO) definitions for adult Asians, Body Mass Index value of 25 kg/m^2 or more was categorized as obese. Central obese was defined as waist circumference of 80 cm or more based on American Diabetes Association Criteria.

Data analysis was done using the Statistical Package for the Social Sciences (SPSS)-16th version. Multiple logistic regression was performed to identify the independent risk predictors for endometrial carcinoma. The independent variables used in the LR analysis were the variables that showed a statistically significant association with endometrial carcinoma at a significant level of < 0.2 in the bivariate analysis. Logistic regression analysis was carried out by purposeful selection of variables method. Goodness of fit of the LR model was assessed by Hosmer and Lameshow test. The values of Cox and Snell R Square test and Negelkerke R Square test were given an indication of the amount of variation in the dependent variable explained by the model. The model with the best goodness-of-fit in the LR analysis was selected as the final LR model. Selection of the risk

predictors in the model was done based on the feasibility and reliability of the independent variables retained in the multivariable analysis. The variables entailed recall bias (Eg: x-ray exposure), and could change over time (Eg: income) were not included in the model. Two models were developed including six variables as risk predictors. For the first model, BMI and age were included as categorical variables and for the second model, as continuous variables. However, the later stage of the development the continuous predictors were categorised into several categories. Two scoring methods were used to calculate the individual score to each predictor variable.

Four models were developed and assessed for model performance by discrimination and calibration. The discriminative performance of the model was assessed by Receiver Operator Characteristic (ROC) curve. Calibration curve and Hosmer-Lemeshow goodness-of-fit test were used to assess the calibration of the model [25].

In the final risk prediction model, each risk predictor carried a weighted score. This weight was based on the β-coefficient obtained for that variable in the LR model. Out of the two scoring methods used in giving the weighted score based on the ROC values the best scoring method was selected. One method was converting the value of regression coefficient to the closest integer and second method was calculating point values based on the smallest β coefficient [25]. At the end, a single total risk score which predicted the overall risk of an individual was calculated by adding up all those weighted scores given for each risk predictor that was relevant to the individual whom the risk prediction model applied.

Based on the scores assigned to each predictor in the model, every individual was assigned a total risk score. ROC curve was plotted and the point on the curve with the minimum distance (d^2) to the upper left corner (0,1) of the ROC plane was used to determine the optimal cut off point to discriminate the cases from non-cases. At-risk or not at-risk categories that were defined by the risk prediction model were tested against the true presence (cases) or absence (controls) of endometrial carcinoma. Sensitivity, specificity, likelihood ratios and predictive values were also presented with 95% confidence intervals. Reliability of the developed risk prediction model was assessed by re-administrating the risk prediction model, employing test-retest method among the cases and the controls.

Results

The total study sample for the development study consisted of 415 postmenopausal women, while 83 were cases diagnosed with endometrial carcinoma and 332 were controls who were free of endometrial carcinoma at the time of recruitment. The response rate of the cases and controls were 100%. A majority of the sample was above 55 years (n=223,53.7%), Sinhalese (n=364, 87.7%), Buddhist (n=333, 80.2%), and had a family income of more than twenty thousand rupees (n=229, 55.2%).

The risk predictors included in the model were age 55 years or more, never conceived, age at menarche ≤11 years, family history of any type of cancer, generalized obesity and ever experienced postmenopausal bleeding. The unadjusted odds ratios of the risk predictors included in the model were shown in Table 1. The size of the sample included in the multivariable analysis was 415 postmenopausal women.

Selection of best risk prediction model

Two models with BMI and age as categorical variables and as continuous variables were shown in Table 2 and Table 3 respectively. However, the later stage of the development the continuous predictors were categorised into several categories and it is shown in Table 4. Further, Table 2 shows the final scores allocated to each variable using two scoring methods in model one. The total score of the method I was 11 and the method II was 9. The final scores of the model two by two scoring method were shown in table 4, and the maximum total scores of the method I and II were 13 and 20 respectively. The area under the curve values of ROC curves drawn to model I and model II with scoring systems separately are given in Table 5. The highest value of AUC was 0.92 (95% CI: 0.88, 0.95) (Figure 1) which was corresponding to the model I (categorical variables) with scoring system II. Therefore, it was considered as the best model to be selected as the final risk prediction model. The results of the Hosmer-Lemeshow goodness-of-fit test indicated a significance of 0.72. The calibration curve to provide insight into the calibrating potential of the model is shown in Figure 2.

Development of a cut-off point to predict the risk of developing endometrial carcinoma

Cut off values of total risk score ranged from -1 to 9. The shortest distance (d^2) in the ROC curve for the final model was 0.051 (Table 6). It corresponded with the total risk score of 4.5, indicating that 4.5 to be the optimal cut off value to categorize each participant into "at-risk" of developing endometrial carcinoma or "not at-risk" of developing endometrial carcinoma.

Table 7 shows the validity indicators of the final risk prediction model estimated based on the calculated cut off of 4.5. The sensitivity of the model was 79.5% (95% CI:68.9,87.3) and the specificity was 90.7% (95% CI:86.8,93.5). Likelihood ratio positive, which explains how much to increase the probability of having a disease, given a positive test result was 8.55 (95% CI:5.99, 12.11) and likelihood ratio negative, which interprets the decreased probability of having a disease, given the negative test result was 0.23 (95% CI:0.15,0.35). These estimates were indicative of a good predictive ability of the risk prediction model based on the development sample. The results of the test-retest reliability demonstrated good test-retest reliability with the correlation coefficient of 0.85, at 0.05 significance level.

Table 1. List of variables with a significant unadjusted OR in bivariate analysis

Variable	Categories	Un-adjusted ORs	CI
Age	More than 55	3.66	2.10-6.38
History of conception	Never conceived	3.31	1.68-6.50
Age at menarche	≤11 years	4.92	2.01-12.02
Family history of any cancer	Yes	9.82	5.17-18.66
BMI (kg/m^2)	≥25	7.92	4.14-15.15
Post-menopausal bleeding	Ever experienced	13.10	6.34-27.10

Table 2. Model I-Parameter estimates, their significance of the logistic regression model as age and BMI as categorical variables and scoring values

Predictor variable	β	SE (β)	AOR*	95% CI for AOR* Lower	95% CI for AOR* Upper	p value	Score Method I	Score Method II
Age more than 55	1.74	0.37	5.71	2.75	11.87	0.000	2	1
Never conceived	1.34	0.52	3.82	1.39	10.48	0.009	1	1
Age at menarche ≤11 years	1.82	0.68	6.14	1.62	23.27	0.008	2	1
Ever experienced postmenopausal bleeding	2.44	0.43	11.50	4.94	26.79	0.000	2	2
Having family history of any type of cancer	2.35	0.45	10.47	4.30	25.49	0.000	2	2
Generalized obesity BMI ≥25 Kgm^{-2}	2.25	0.43	9.51	4.08	22.18	0.000	2	2
Constant	-6.36	0.66	0.002			0.000		
							Total 11	Total 9

Table 3. Model 2-Parameter estimates and their significance of the logistic regression model, age and BMI as continuous variables

Predictor variable	β	SE (β)	AOR*	95% CI for AOR* Lower	95% CI for AOR* Upper	p value
Age	0.08	0.02	1.08	1.04	1.13	0.000
Never conceived	1.47	0.50	4.36	1.62	11.70	0.003
Age at menarche ≤11 years	1.91	0.63	6.72	1.94	23.27	0.003
Ever experienced postmenopausal bleeding	2.40	0.44	11.03	4.70	25.90	0.000
Having family history of any type of cancer	2.31	0.46	10.10	4.11	24.78	0.000
Body Mass Index (BMI)	0.27	0.04	1.31	1.20	1.43	0.000
Constant	-15.72	0.66	0.002			0.000

AOR* = Adjusted Odds Ratio

Table 4. Modified model II with final score by scoring method I and II

Predictor variable	β	Range	Score method I	Score method II
Age		0.08	43-49	0
		50-59	1	1
		60-69	1	2
		70-79	2	3
		80-84	3	4
Never conceived	1.47		1	2
Age at menarche ≤11 years	1.91		2	3
Ever experienced postmenopausal bleeding	2.40		2	3
Having family history of any type of cancer	2.31		2	3
Body Mass Index (BMI)	0.27	16-24.9	0	
		25-38.9	3	5
Total (maximum)			**13**	**20**

Table 5. AUC values of model I and model II

Model discrimination	Model I (Categorical variables)	Model II (Continuous variables)
Scoring system I	0.91 (95% CI: 0.87, 0.94)	0.90 (95% CI: 0.86, 0.93)
Scoring system II	**0.92 (95% CI: 0.88, 0.95)**	0.90 (95% CI: 0.87, 0.94)

Table 6. Sensitivity, specificity and corresponding d^2 each cut-off value

Cut off value	Sensitivity	1 - Specificity	Distance from curve($d2$)
-1.00	1.000	1.000	1.0
0.50	1.000	0.877	0.769
1.50	1.000	0.660	0.436
2.50	0.964	0.428	0.184
3.50	0.892	0.238	0.068
4.50	0.795	0.093	0.051
5.50	0.410	0.021	0.348
6.50	0.217	0.006	0.613
7.50	0.072	0.000	0.861
9.00	0.000	.000	1.000

Table 7. Validity indicators of the final risk prediction model at the optimal cut off level

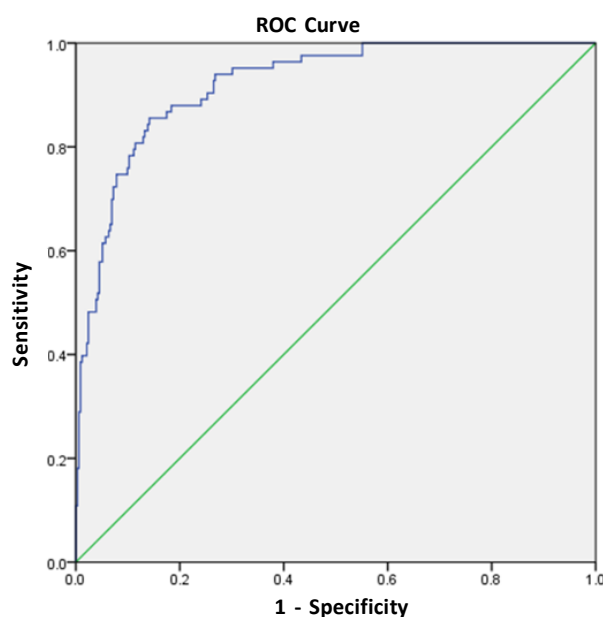| Validity indicators	Value (95% CI)
Sensitivity	79.5 (68.9,87.3)
Specificity	90.7 (86.8,93.5)
Positive Predictive Value	68.0 (57.7,76.9)
Negative Predictive Value	94.7 (91.4, 96.7)
Likelihood Ratio Positive	8.55 (5.99, 12.11)
Likelihood Ratio Negative	0.23 (0.15,0.35)

Figure 1. ROC curve for summary risk scores against the presence of endometrial carcinoma among the study population.
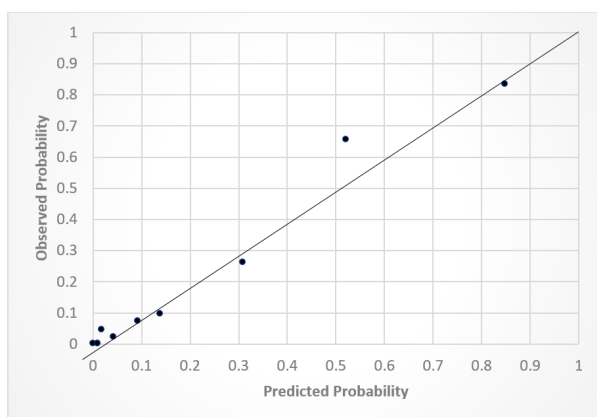

Figure 2. Calibration plot for predicted probability by risk prediction model vs observed frequencies.

Discussion

The aim of the present study was to develop a simple, low cost and user-friendly risk prediction model for endometrial carcinoma. Predictors included in the final risk prediction model of the present study were age more than 55 years, history of never conceived, age at menarche ≤11 years, having a family history of any type of cancer, postmenopausal bleeding and generalized obesity with BMI ≥25 kgm^{-2}.

In the present study, a prognostic model was developed to predict the risk of a postmenopausal woman in developing endometrial carcinoma in future. The most preferable study design to be used in developing a

prediction model is a prospective cohort study, as it facilitates the optimal measurement of predictors and the outcomes [26]. It should be ideally developed using individuals in good health [27]. However, a case control study was carried out to develop the prediction model in the current study due to the limited time and other logistic constraints which is a limitation of the present study. However, case control study design being an alternative to longitudinal designs in developing risk prediction models [14], it had been used to develop risk prediction models for other carcinomas in literature [10-12].

The present study was conducted based on scientifically estimated sample size to minimize the chance during sampling. Case to control ratio of 1:4 was taken in the sample size calculation to strengthen the statistical power of the study. Several pieces of research had suggested estimating the sample size based on rule 'at least ten events per variable' [26,28,29]. In this study, 83 cases of endometrial carcinoma had been used to develop the model resulting in only six variables in the final risk prediction model indicating the adequacy of the sample size.

The primary objective of developing risk prediction model was to predict 'at risk' population using a simple, less time-consuming tool with low cost. Therefore, it was ensured that developed risk prediction model was a user friendly, easy to be applied in a community setting or in a clinic setting as a screening tool to identify 'at risk' women for developing endometrial carcinoma among postmenopausal women. Ideally, the development of the risk prediction model to screen the high-risk women before symptoms occur should be carried out by including risk factors of endometrial carcinoma to prevent the effect of inflating the predictive value of the tool by including presenting symptoms. In this study, one of the common presenting symptoms of endometrial carcinoma, postmenopausal bleeding had been included as a risk predictor, thus this model was developed as a detection tool to improve the detection rate of endometrial carcinoma with or without symptoms but had not sought medical care. Further, the risk predictors in the current study had also been found in the other risk prediction models for endometrial carcinoma [4,13,15]. Endometrial thickness was a common predictor in many of these risk prediction models [4,13,15] and models incorporating endometrial thickness were more accurate in predicting endometrial carcinoma, Though, in a low resource country like Sri Lanka, to use a risk prediction model with ultrasound imaging at the community or even at a clinic is not feasible. Hence, a model based on clinical characteristics would be a better option at early triage of high-risk women. Potential bias introduced by predictor assessment contaminated by knowledge on outcome is also a limitation of the study.

Consistency was observed in methods to assess the model performance of the present study compared to similar studies to develop risk prediction models for EC. An AUC of 0.92 indicated a good discriminative power compared

to other models to differentiate between individuals who are at-risk from not at-risk of developing EC. The calibration power was also assessed by calculating the agreement between the predicted and observed probabilities of developing EC. Ideally, the model should be highly sensitive as well as highly specific, though this becomes impossible due to the trade-off between sensitivity and specificity. The model developed in the present study indicated a high specificity at the expense of sensitivity. It is well accepted that the specificity of the model should be increased at the expense of sensitivity when the cost and the risk associated with further investigations are considerable [30]. The women with increased risk of developing EC identified by the model need to be closely followed up and undergone invasive and costly investigations for definitive diagnosis. Hence, a model with high specificity compared to low sensitivity is justifiable.

This present study was carried out in the Western province. The socio-economic status of residents in the Western province is different from the residents of the other provinces. District of Colombo being highly commercial and exposed to mixture of modern life influence the lifestyle of residents in the Western province. Exposure to several risk factors related to lifestyle and environment among women of Western province may differ from the women of other provinces. Hence, the magnitude of the risk factors identified in the risk profile for endometrial carcinoma cannot be generalized to the women in the rest of the country. This was identified as a limitation of the study.

The developed risk prediction model can be incorporated into the existing healthcare system at the community setting as well as at the clinical setting. The Well Women Clinics (WWC), Healthy Lifestyle Centres (HLC) and Medical Officer of Health (MOH) office are common community settings where postmenopausal women are attending to get the health services in Sri Lanka. In addition, gynaecology clinics, medical clinics, eye clinics, outpatient departments and primary healthcare units are the most common clinical settings where that the model can be applied to identify postmenopausal women at increased risk of developing EC in Sri Lankan healthcare setting.

The risk prediction model developed to predict the risk of EC among postmenopausal women comprised of six risk predictors. The model demonstrated good discrimination and well calibration. It entails many features of a model that can be applied in the community as well as in the clinical setting with low cost and minimum time consumption.

The performance of a predictive model is overestimated when simply determined on the sample of subjects that is used to develop the model [31]. After developing a prediction model, it is strongly recommended to evaluate the performance of the model in another set of participants' data than is used for the model development [32]. Therefore, temporal validation was carried out as a technique of external validation of the developed risk prediction model before application in reality [33].

References

1. Arora V, Quinn, MA. Endometrial cancer. *Best Pract Res Clin Obstet Gynaecol*. 2012; **26**(3): 311-24. doi: 10.1016/j.bpobgyn.2011.12.007

2. Sorosky JI. Endometrial Cancer. *Obstet Gynecol.* 2008; **111**(2, Part 1): 436-47. doi:10.1097/aog.0b013e318162f690

3. Munot S, Lane G. Modern management of postmenopausal bleeding. *Trends Urol Gynaecol Sex Health* 2008; **13**: 20-24. https://doi.org/10.1002/tre.84

4. Burbos N, Musonda P, Giarenis I, Shiner AM, Giamougiannis P, Morris EP, Nieto JJ. Predicting the risk of endometrial cancer in postmenopausal women presenting with vaginal bleeding: The Norwich DEFAB risk assessment tool. *Br J Cancer*. 2010; **102**(8): 1201. doi: 10.1038/sj.bjc.6605620

5. Jacobs I, Gentry-Maharaj A, Burnell M, Manchanda R, Singh N, Sharma A, Brunell C. Sensitivity of transvaginal ultrasound screening for endometrial cancer in postmenopausal women: a case-control study within the UKCTOCS cohort. *Lancet Oncol*. 2011; **12**(1): 38-48. doi:10.1016/S1470-2045(10)70268-0

6. American Cancer Society. Endometrial Cancer [internet]. Atlanta: American Cancer Society; 2014; 1-25. Available from http://www.cancer.org/cancer/endometrialcancer/detailedguide/endometrial-uterine-cancer-what-is-endometrial-cancer

7. Anothaisintawee T, Teerawattananon Y, Wiratkapun C, Srinakarin J, Woodtichartpreecha P, Hirunpat S, Thakkinstian A. Development and validation of a breast cancer risk prediction model for Thai women: a cross-sectional study. *Asian Pac J Cancer Prev*. 2014; **15**(16): 6811-7.

8. National Cancer Institute. Cancer risk prediction and assessment [internet]. United States of America: 2018. Available from https://epi.grants.cancer.gov/cancer_risk_prediction/

9. Pavlou M, Ambler G, Seaman SR, Guttmann O, Elliott P, King M, Omar RZ. How to develop a more accurate risk prediction model when there are few events. *BMJ*. 2015; **351**: h3868. http://doi.org/10.1136/bmj.h3868

10. Kumari PBVR. Risk factors and risk assessment of breast cancer among women in the district of Colombo: Post Graduate Institute of Medicine, University of Colombo, Colombo; 2013.

11. Samarakoon YM, Gunawardena NS, Pathirana A, Perera MN, Hewage SA. Prediction of colorectal cancer risk among adults in a lower middle-income country. *J Gastrointest Oncol* 2019; **10**(3): 445-52. doi: 10.21037/ jgo.2019.01.27

12. Thalagala IA. Risk factors, risk prediction and quality of life related to oesophageal carcinoma among adults in the Western province: Post Graduate Institute of Medicine, University of Colombo, Colombo; 2016.

13. Wong ASW, Cheung CW, Fung LWY, Lao TTH, Mol BWJ, Sahota DS. Development and validation of prediction models for endometrial cancer in postmenopausal bleeding. *Eur J Obstet Gynecol Reprod Biol*. 2016; **203**: 220-4. http://doi.org/10.1016/j.ejogrb.2016.05.004

14. Moons KG, Altman DG, Reitsma JB, *et al*. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015; **162**(1): W1-W73. doi:10.7326/M14-0698

15. Giannella L, Mfuta K, Setti T, Cerami LB, Bergamini E, Boselli F. A risk-scoring model for the prediction of endometrial cancer among symptomatic postmenopausal women with endometrial thickness> 4 mm. *Biomed Res Int*. 2014: 1-7. doi:10.1155/2014/130569

16. Opolskiene G, Sladkevicius P, Valentin L. Prediction of endometrial malignancy in women with postmenopausal bleeding and sonographic endometrial thickness ≥4.5 mm. *Ultrasound Obstet Gynecol*. 2011; **37**(2): 232-40. doi:10.1002/uog.8871.

17. Jayawickcrama WIU, Abeysena C. Risk factors for endometrial carcinoma among postmenopausal women in Sri Lanka: a case control study. *BMC Public Health* 2019; **19**: 1387 doi:10.1186/s12889-019-7757-2

18. Xu WH, Xiang YB, Ruan ZX, Zheng W, Cheng JR, Dai Q, *et al*. Menstrual and reproductive factors and endometrial cancer risk: Results from a population-based case-control study in urban Shanghai. *Int J Cancer*. 2004; **108**(4): 613-9. doi:10.1002/ijc.11598

19. Van Doorn LC, Opmeer B, Duk M, Kruitwagen R, DIJKHUIZEN FPHLJ, Mol BWM. The relation between age, time since menopause, and endometrial cancer in women with postmenopausal bleeding. *Int J Gynecol Cancer*. 2007; **17**: 1118-23.

20. Xu WH, Xiang YB, Zheng W, Zhang X, Ruan ZX, Cheng JR, *et al*. Weight history and risk of endometrial cancer among Chinese women. *Int J Epidemiol*. 2006; **35**(1): 159-66.

21. Friberg E, Mantzoros CS, Wolk A. Physical activity and risk of endometrial cancer: A population-based prospective cohort study. *Cancer Epidemiol Biomarkers Prev*. 2006; **15**(11): 2136-40.

22. Inoue M, Okayama A, Fujith M, Enomata T, Tanizawa O, Ueshima H. A case control study on risk factors for uterine cancer in Japan, *Jpn J Cancer Res*. 1994; **85**(4), 346-50.

23. Terry P, Baron JA, Weiderpass E, *et al* (1999). Lifestyle and endometrial cancer risk: a cohort study from the Swedish Twin Registry. *Int J Cancer*. 1999; **82**: 38-42.

24. Friberg, E, Mantzoros CS, Wolk A. Diabetes and risk of endometrial cancer: A population-based prospective cohort study. *Cancer Epidemiol Biomarkers Prev*. 2007a; **16**(2): 276-80. https://doi.org/10.1158/1055-9965.EPI-06-0751

25. Han K, Song K, Choi BW. How to Develop, Validate, and Compare Clinical Prediction Models Involving Radiological Parameters: Study Design and Statistical Methods. *Korean J Radiol*. 2016; **17**(3): 339-50. http://doi.org/10.3348/kjr.2016.17.3.339

26. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. (2009). Prognosis and prognostic research: what, why, and how? *BMJ*. 2009: 338-75. http://doi.org/10.1136/bmj.b375

27. Collins GS, Reitsma JB, Altman DG, Moons KGMM. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD Statement. *Ann Intern Med*. 2015; **162**(1): 55-63. https://doi.org/10.7326/M14-0697

28. Hendriksen JMT, Geersing GJ, Moons KGM, Groot JA. (2013). Diagnostic and prognostic prediction models. *J Thromb Haemost*. 2013; **11**: 129-41. doi:10.1111/jth.12262

29. Lee Y, Bang H, Kim DJ. How to Establish Clinical Prediction Models. *J Clin Endocrinol Metab*. 2016; **31**(1): 38. https://doi.org/10.3803/EnM.2016.31.1.38

30. Hennekens CH, Buring JE, Mayrent SL. Epidemiology in Medicine. 1987. 1st ed. Boston/Toronto: Little Brown and Company.

31. Steyerberg EW, Bleeker S E, Moll HA, Grobbee DE, Moons KG (2003). Internal and external validation of predictive models: a simulation study of bias and precision in small samples. *Journal of Clinical Epidemiology* 2003; **56**(5): 441-7. doi:10.1016/S0895-4356(03)00047-7

32. Collins GS, Reitsma JB, Altman DG, Moons KGMM. (2015). Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD Statement. *Annals of Internal Medicine* 2015; **162**(1): 55-63. https://doi.org/10.7326/M14-0697

33. Jayawickrama WIU, Abeysena C (2022). Validation of a risk prediction model for endometrial carcinoma among postmenopausal women in Sri Lanka. *Journal of the College of Community Physicians of Sri Lanka* 2022; **27**(5): 71-8. DOI: http://doi.org/10.4038/jccpsl.v27i5.8443