An Expert System to Detect Uterine Cancer under Uncertainty

Muhammed Jamshed Alam Patwary [1], Subrina Akter [2] Tanjim Mahmud [3]

[1](Computer Science & Engineering, International Islamic University Chittagong, Bangladesh)
[2](Computer Science & Engineering, International Islamic University Chittagong, Bangladesh)
[3](Computer Science and Engineering, University of Information Technology and Sciences, Bangladesh)

Abstract: *Uterine cancer is one of the conspicuous cancers for women in both developed and developing countries including Bangladesh. Now, in the world it is the sixth most common cancer among women and fourteenth most common cancer overall. The high occurrence of uterine cancer in women has increased significantly in the last years. This involves many factors to be measured and evaluated, which are related to the signs and symptoms of this disease. These factors usually expressed in quantitative and qualitative ways. In addition, a hierarchical relationship exists among these factors. Since qualitative factors cannot be measured in a quantitative way, resulting various types of uncertainties such as incompleteness, vagueness, imprecision. Therefore, it is necessary to address the issue of uncertainty by using appropriate methodology; otherwise, the conclusion to detect uterine cancer will become inaccurate. There exist many systems to address the issue presented in this paper. However, none of them is able to address the issue of uncertainty. Therefore, this paper demonstrates the application of a novel method, named belief rule-based inference methodology-RIMER, which is capable of addressing the uncertainties in both clinical domain knowledge and clinical data. This paper reports the development of a Belief Rule Based Expert System (BRBES) using RIMER approach, which is capable of detecting the presence of uterine cancer by taking account of signs and symptoms. The system has been validated by using real patient data and it has been observed that the results generated by the BRBES are more reliable than the manual system usually carried out by a physician.*
Keywords: *Expert System, Evidential Reasoning, RIMER, Signs and Symptoms, Uncertainty, Uterine Cancer*

I. Introduction

Uterine cancer is the most common cancer, recognized in the reproductive system of women. It is also known as endometrial cancer in clinical term. Uterine cancer begins when normal cells of the uterus get changing and grows irrepressibly and thus, forming a mass called tumor. A tumor can be benign or malignant. Benign tumor is noncancerous but malignant tumor grows rapidly and damages the surrounding tissues by attacking them. It is the sixth most common cancer among women, while fourteenth most in overall [1]. From a survey, it has been observed that in 2012 the cases of uterine cancer were found 5% among all types of cancers [1].The cases of this cancer is the highest in Northern America and Europe, while it is lowest in Africa and Asia [2]. In most developed countries the percentage of uterine cancer occurrence is about 53 among all types of cancers [2]. Most of the women who are suffering from this cancer in Bangladesh died. The dead rate is increasing day by day [3].

The identification of uterine cancer is very complex. Usually, by looking at the signs and symptoms of this disease, a physician determines the suspicion of this disease. These signs and symptoms are measured by using linguistic terms such as high, medium low and hence, they are qualitative in nature. Consequently, the determination of suspicion of uterine cancer contains various types of uncertainties such as vagueness, imprecision, and incompleteness, resulting inaccurate determination of this disease. Signs are measured by physician and symptoms are expressed by the patient. An example of sign is abdominal enlargement, which is measured by the physician and expressed in linguistic terms such as high, medium and low. Pain is an example of symptom expressed by the patient again in linguistic terms as mentioned. Details of the signs and symptoms and their associated uncertainties are discussed in the next section. In Bangladesh, the physician usually decides whether a patient will go for the expensive lab investigation by looking at his signs and symptoms. In most of the cases, it has been observed that patients usually suffer from benign tumor which is noncancerous and hence, resulting enormous financial loss to the patient.

From the above it can be inferred that the determination of suspicion of uterine cancer in a woman is carried out based on human judgment or physician judgment. However, this judgment is inaccurate because of the inability of the physician in processing uncertainty that exists with the signs and symptoms as discussed. This inability can be overcome by developing an expert system with the ability to process uncertain knowledge associated with the signs and symptoms of uterine cancer. IF-Then rules are widely used to capture human knowledge. An extended version of the IF-Then rule, known as Belief Rule Base (BRB) can be employed to

build the knowledge base of this expert system, which has the appropriate schema to capture uncertainty. Evidential Reasoning (ER) is considered as the inference engine in this BRB, which is capable of handle, various types of uncertainty.

In this paper, Section II describes an overview of uterine cancer with signs, symptoms. The review of the existing work relevant to this proposed system is presented in section III. Section IV introduces the RIMER methodology. A Belief Rule Based Expert System (BRBES) is presented in Section V. Section VI presents the result and discussion while VII concludes the paper.

II. Uterine cancer (UC)

Uterine cancer (UC) forms in the endometrium, known as the wall of the uterus. The UCs those begin in the endometrium, called endometrial carcinomas. However, some UCs can initiate in the uterus muscle layer [4][5]called sarcomas[6].In most cases UCs develop in women with aged between 50s and 75s[7]. UC may develop in women under the age of 50[7], which is recognized as the special case. UC treatment usually carried out in the emergency department of a hospital. It can be recoverable if timely suspicion of UC can be assessed. However, the assessment of suspicion of UC depends on the appropriate evaluation of the signs, symptoms as well as on the overall health condition of a patient. Therefore, this section presents a brief discussion on the common signs and symptoms of UC and their associated uncertainties. The procedures of determination of UC suspicion are also presented.

When a woman noticed the signs and symptoms of UCs after or before menopause, she needs to visit the physician   immediately. During the assessment of UC suspicion, a physician makes physical and pelvic investigation, as well as talk to the patient to know the signs and symptoms, risk factors, and family medical history [6]. While the suspicion is found positive the patient is advised to go for ultrasonography. If tumor detected during ultrasonography, biopsy carried out to test the tissue of the tumor. After testing the tissue, if the tumor identified as malignancy, then it can be inferred that the patient has UC. In addition to ultrasonography, the physician may recommend other types of test such as transvaginal ultrasound, hysteroscopy, and/or blood tests to learn more. All these tests are very expensive. However, hysteroscopy is not available in Bangladesh. For this reason, all uterine tumor patients whether it is benign or malignant need to go for expensive operation [8][9]. It is interesting to observe that although benign tumor is not responsible for UC, a patient with this tumor needs to go for expensive and unnecessary operation.

2.1. Uncertainty in UC Suspicion from Signs and Symptoms

According to expert doctors (gynecologists) the symptoms of uterine cancer are pelvic pain, menopausal types, post-menopausal vaginal bleeding, Para, abnormal bleeding (in between periods), abnormal vaginal discharge (watery or blood-tinged) and family medical history etc[8][9][10]. Menopause is typically a regular change; it usually arises in women approximately in their age of late 40s or early 50s, signing the infertile phase of a woman's life[11].Para means a woman who has borne one or more infants[12]. Two types of Para are considered here. One is Nullipara (Para 0)[13], a woman who has never given birth to a viable infant and Multipara, a woman who given birth to one or more viable infants[12].

All symptoms have been investigated but two symptoms including menopausal types and Para and one sign and symptom post-menopausal vaginal bleeding are given most importance, because the remaining symptoms except these three are not common in every patient. Post-menopausal vaginal bleeding is expressed by the patient by using linguistic terms such as High, Medium and Low and hence its expression is ambiguous, vague and imprecise. For example, the patients may consider her bleeding as high, on the other hand the experts may consider the patients bleeding as medium. Therefore, it can't be measured with 100% certainty. Menopausal types are expressed by the patient in terms of 'yes' (pre) or 'no' (post) while asked by the physician. Similarly, Para can also be expressed by 'yes' (nullipara) or 'no' (multipara). Therefore, in evaluating the value of the three symptoms, it can be observed that menopausal types and Para can be evaluated with 100% certainty while post-menopausal vaginal bleeding can't be evaluated with 100% certainty. But it is one of the major symptoms of UC [4][8][9].

III. Related Work

The introduction of expert systems (ES) can help simplifying the health care processes and substantially facilitate clinical practice and reduce medical errors. Till- now, there is no computer aided system developed for detecting uterine cancer diagnosis But there are numerous examples of expert systems[14] in health care which have successfully improved the quality of health care. Several research groups working world wide in developing Computer Aided System for detecting cervical cancer and ovarian cancer. A hybrid decision support system in soft computing paradigm for detecting the different stages of cervical cancer has been developed [15]. Another method using neuro-fuzzy model for preoperative prediction of malignancy in ovarian tumor has been proposed [16].This model predicts ovarian cancer malignancy by using demographic, serum

tumor marker and ultrasound-derived criteria. A new method named, CLFNN, to aid existing diagnosis methods for detecting ovarian cancer also developed [17]. A method has been developed to detect cervical cancer which based on Fuzzy Logics and Image Interpretation [18]. Artificial Neural Network also used for early detection of ovarian cancer [19]. In this paper data mining technique used to diagnosis of ovarian cancer and identified the most informative points of the mass-spectrum curve, then used student t-test and neural networks to determine the differences between the curves of cancer patients and healthy people. Though ESs are promising in helping facilitate evidence based medicine and reducing patient safety incidents, there are challenges in this research area that have made few ESs widely applied in practice. The reason for this is that uncertainties in clinical signs, clinical symptoms and clinical domain knowledge [20], the complexity involved in inference mechanism, and difficulties with domain selection and knowledge base construction and maintenance impede development and implementation of ESs. In the literature, different methods have been proposed to model uncertainties such as certian factors were used in MYCIN[21][22] which was an rule-based expert system developed[23] in the early 1970s at Stanford University to identify causes of severe infections and to recommend antibiotics. An example of another method [24] is in abdominal pain diagnosis. Bayesian theory was employed in Iliad [25], which was an ES developed at the University of Utah to teach medical decision-making [26] [27] and to provide assistance in differential diagnosis across the domain of internal medicine. All of the mentioned methods can handle some types of uncertainty, but these methods cannot handle all types of uncertainties. Fuzzy information may exist with ignorance, leading to the stimulation of knowledge with uncertainty but only with degrees of belief [14][28]. In Fuzzy set theory, reasonable memberships functions are not always clear [14]. Various methods have been developed [29][30], but none of the general methods are able to address the various types of uncertainty[40]. Exponential number of prior probabilities is required, which is determined by statistical analysis in Bayesian Network and this is the major problem of Bayesian approach [30]. The nature of input data are both qualitative and quantitative in nature to assess the suspicion of Uterine Cancer and therefore, RIMER methodology has been employed to develop an Expert System, allowing the modeling of clinical domain knowledge under uncertainties[32]. Such a rule base is capable of capturing vagueness, incompleteness, and imprecision, while traditional 'IF-THEN' rules can be represented as a special case. The next section introduces the RIMER methodology.

IV. Overview Of Rimer Methodology

In RIMER, Belief Rule Base (BRB) can capture complicated nonlinear causal relationships between antecedent attributes and consequents, which are not possible in traditional IF-THEN rules. BRB is used to model domain specific knowledge under uncertainty, and the ER approach is employed to facilitate inference. This section introduces BRB as a knowledge representation schema under uncertainty as well as inference procedures of RIMER.

4.1. Modeling domain knowledge using BRB

Belief Rules are the key constituents of a BRB, which include belief degree. This is the extended form of traditional IF-THEN rules. In a belief rule, each antecedent attribute takes referential values and each possible consequent is associated with belief degrees [33]. The knowledge representation parameters are rule weights; attribute weights and belief degrees embedded with consequent attribute of a rule, which are not available in traditional IF-THEN rules. A belief rule can be defined in the following way.

$$R_k : \begin{cases} IF(P_1 \ is \ A_1^k) \cap (P_2 \ is \ A_2^k) ... ... \cap (P_{T_k} \ is \ A_{T_k}^k) \\ THEN\{(C_1, \beta_{1k}), (C_2, \beta_{2k}), ...(C_N, \beta_{Nk})\} \end{cases} \qquad (1)$$

$$(\beta_{jk} \geq 0, \sum_{j=1}^{N} \beta_{jk} \leq 1), with \ a \ rule \ weight \ \theta_k, attribute \ weights \ \delta_{k1}, \delta_{k2}, \delta_{k3}, ...\delta_{kT_k} \ k \in \{1, ..., L\}$$

Where $p_1, p_2, p_3, ..., p_{T_k}$ represent the antecedent attributes in the kth rule. $A_i^k (i = 1, ..., T_k, k = 1, ..., L)$ represents one of the referential values of the ith antecedent attribute P_i in the kth rule. C_j, is one of the consequent reference values of the belief rule.

$\beta_{jk} (j = 1, ..., N, \ k = 1, ... L)$ is one of the belief degrees to which the consequent reference value C_j is believed to be true. If $\sum_{j=1}^{N} \beta_{jk} = 1$ the kth rule is said to be complete; otherwise, it is incomplete. T_k, is the total number of antecedent attributes used in kth rule. L is the number of belief rules in the rule base. N is the

number of all possible consequent reference values of the consequent attribute.

For example a belief rule to assess pain of a patient with uterine cancer can be written in the following way.

$$R_k : \begin{cases} IF \quad Back \ pain \ is \quad High \ AND \quad Pelvic \ pain \ is \quad High \\ THEN \quad Pain \quad is \\ \{(High,(1.00)),(Medium,(0.00)),(Low,(0.00))\} \end{cases} \tag{2}$$

Where {(High, 1.00), (Medium, 0.00), (Low, 0.00)} is the belief distribution for the pain consequent, stating that the degree of belief associated with High is 100%, 0% with Medium and 0% with Low. In this belief rule, the total degree of belief is (1+0+0) =1, hence, the assessment is complete.

4.2. BRB Inference Procedures

The ER approach [32][35][36][37] developed to handle multiple attribute decision analysis (MADA) problem having both qualitative and quantitative attributes. Different from traditional MADA approaches, ER presents MADA problem by using a decision matrix, or a belief expression matrix, in which each attribute of an alternative described by a distribution assessment using a belief structure. The inference procedures in BRB inference system consists of the procedures including input transformation, rule activation weight calculation, rule update mechanism, followed by the aggregation of the rules of a BRB by using ER [33][34][38].

The input transformation of a value of an antecedent attribute Pi consists of distributing the value into belief degrees of the referential values of that antecedent [39]. The input value of P_i, which is the ith antecedent attribute of a rule, along with its belief degree ε_i is shown below by equation (3). The belief degree ε_i to the input value is assigned by the expert in this research.

$$H(P_i, \varepsilon_i) = \{(A_{ij}, \alpha_{ij}), j =, ..., j_i\}, i = 1, ..., T_k \tag{3}$$

Here H is used to show the assessment of the belief degree assigned to the input value of the antecedent attribute. In the above equation A_{ij} (ith value) is the jth referential value of the input P_i. α_{ij}, is the belief degree to the referential value A_{ij} with $\alpha_{ij} \geq 0$. $\sum_{j=1}^{j_i} \alpha_{ij} \leq 1(i = 1, ..., T_k)$, and j_i is the number of the referential values.

In this research the input value of an antecedent attribute is collected from the patient or from the physician in terms of linguistic terms such as 'High', 'Medium' and 'Low'. This linguistic value is then assigned degree of belief ε_i by taking account of expert judgment. This assigned degree of belief is then distributed in terms of belief degree α_{ij} of the referential values A_{ij} [High (H), Medium (M), Low, (L)] of the antecedent attribute. The above procedure of input transformation is elaborated below through equations (4) and (5)

$$if \ H \ value \geq input \ value(\varepsilon_i) \geq M \ value \ Then, Medium = \frac{H \ value - input \ value}{H \ value - M \ value},$$

$$High = (1 - Medium), Low = 0.00 \tag{4}$$

$$if \ M \ value > input \ value(\varepsilon_i) \geq L \ value \ Than, Low = \frac{M \ value - input \ value}{M \ value - L \ value},$$

$$Medium = (1 - Low), High = 0.00 \tag{5}$$

For example, if the degree of belief ε_i is assigned with reference to a linguistic term by the expert as 0.80 then this can be considered as the input value of the antecedent attribute. This input value can be transformed or distributed over the referential values of this antecedent attribute by using either equation (4) or (5). The following is the distribution of the degree of belief associated with the referential values of this antecedent attribute. Here, "High" referential value is given utility score as "1" and "Medium" is given utility score as "0.5". While "Low" referential value is given utility score as "0".

Medium= (1.00-0.80) / (1-0.5) =0.40, High=1-Meduim=1-0.4=0.60

Hence, {(High, 0.60), (Medium, 0.40), (Low, 0.00)}

It is important to note that the rule (**k**th) of an initial rule base of a BRB is constructed by taking account of only one of the referential values (A_i^k, which is the element of A_{ij}) of an antecedent attribute (**Pi**).

Therefore, it is necessary to find out the degree of belief α_i^k (which is the element of α_{ij}) of this referential value (A_i^k), which can be defined as the matching degree at which the belief is matched. α_{ij} can be calculated by using (4) and (5). When matching degree is assigned to the referential values of the antecedent attributes of a rule (**k**th) then the rule is considered as activated. This phenomenon is called Packet Antecedent of a rule. When the kth rule is activated, the weight of activation of the kth rule, w_k is calculated by using the flowing formula[33].

$$w_k = \frac{\theta_k \alpha_k}{\sum_{j=1}^{L} \theta_j \alpha_j} = \frac{\theta_k \prod_{i=1}^{T_k} (\alpha_i^k)^{\overline{\delta}_{ki}}}{\sum_{j=1}^{L} \theta_j [\prod_{i=1}^{T_k} (\alpha_i^k)^{\overline{\delta}_{ji}}]}$$

And $\overline{\delta}_{ki} = \frac{\delta_{ki}}{\max\{\delta_{ki}\}}, i = 1,...,T_k$ (6)

Where $\overline{\delta}_{ki}$ is the relative weight of P_i used in the kth rule, which is calculated by dividing weight of P_i with maximum weight of all the antecedent attributes of the kth rule. By doing so, the value of $\overline{\delta}_{ki}$ becomes normalize, meaning that the range of its value should be between 0 and 1. $\alpha_k = \prod_{i=1}^{T_k} (\alpha_i^k)^{\overline{\delta}_{ki}}$ is the combined matching degree, which is calculated by using multiplicative aggregation function.

When the kth rule is activated, the incompleteness of the consequent of this rule can also be happen due to the absence of input data of its antecedent attributes. An incomplete input to an attribute will lead to an incomplete output in each of the rules in which the attribute is used. Therefore, it is necessary to update the original belief degree. Hence, original belief degree $\overline{\beta}_{ik}$ in the ith consequent C_i of the kth rule is updated based on the actual input information as[33].

Table 1: Rule Base for Pain with Update Belief Degree

W_k	Antecedent	Initial Belief Degree	Updated Belief Degree
W_1=0.88	(D_1 is H^ D_2 is H)	Pain is {(H, 0.90)}	Pain is {(H, 1)}
W_2=0.73	(D_1 is H^ D_2 is M)	Pain is{(H, 0.66),(M, 0.10)}	Pain is{(H, 0.73),(M, 0.15)}
W_3=0	(D_1 is H^ D_2 is L)	Pain is{(M, 0.45),(L, 0.35)}	Pain is{(M, 0.53),(L, 0.25)}
W_4=0	(D_1 is M^ D_2 is H)	Pain is{(H, 0.70),(M, 0.12)}	Pain is{(H, 0.73),(M, 0.15)}
W_5=0.44	(D_1 is M^ D_2 is M)	Pain is{(H, 0.12),(M, 0.72),(L,0.1)}	Pain is{(H, 0.10),(M, 0.73),(L,0.17)}
W_6=0	(D_1 is M^ D_2 is L)	Pain is{(H, 0),(M, 0.450)(L,0.56)}	Pain is{(H, 0),(M, 0.50)(L,0.5)}
W_7=0	(D_1 is L^ D_2 is H)	Pain is{(H, 0.3),(M, 0.5)}	Pain is{(H, 0.3),(M, 0.15)}
W_8=0.22	(D_1 is L^ D_2 is M)	Pain is{(M, 0.2),(L, 0.10)}	Pain is{(M, 0.3),(L, 0.15)}
W_9=0	(D_1 is L^ D_2 is L)	Pain is{(H, 0.3),(M, 0.15)}	Pain is{(H, 0.3),(M, 0.15)}

$$\beta_{ik} = \overline{\beta}_{ik} \frac{\sum_{t=1}^{T_k} (\tau(t,k) \sum_{j=1}^{J_t} \alpha_{tj})}{\sum_{t=1}^{T_k} \tau(t,k)}$$ (7)

Where, $\tau(t,k) = \begin{cases} 1, \text{if } U_t \text{is used in defining } R_k \ (t = 1,...,T_k) \\ 0, \text{otherwise.} \end{cases}$

Here β_{ik} is the updated belief degree.

Due to the incomplete input for 'Back Pain', the belief degree of the associated rules needs to be updated to demonstrate the incompleteness by using (7)

$$\beta_{ik} \equiv \overline{\beta}_{ik} \frac{1.6}{2} = \overline{\beta}_{ik} * 0.8, \qquad i = 1,2,3; k = 1,...9.$$

Therefore $0 < \sum_{i=1}^{3} \beta_{ik} < 1$ for all rules that are associated with 'Back pain'. The sub rule base for 'pain' with the updated belief degree is illustrated in table 1.

ER approach is used to aggregate all the packet antecedents of the **L** rules to obtain the degree of belief of each referential values of the consequent attribute by taking account of given input values P_i of antecedent attributes. This aggregation can be carried out either using recursive or analytical approach. Analytical approach [40] has been considered since it is computationally efficient than recursive approach [35][36][37][41]. The conclusion O(Y), consisting of referential values of the consequent attribute, is generated by using equation (8).

$$O(Y) = S(P_i) = \{(C_j, \beta_j), j = 1, ..., N\} \tag{8}$$

Where β_j denotes the belief degree associated with one of the consequent reference values such as C_j and β_j is calculating by analytical approach [42] as illustrated in equation (9).

$$\beta_j = \frac{\mu \times [\prod_{k=1}^{L}((w_k \beta_{jk} + 1 - w_k \sum_{j=1}^{N} \beta_{jk})) - \prod_{k=1}^{L}(1 - w_k \sum_{j=1}^{N} \beta_{jk})]}{1 - \mu \times [\prod_{k=1}^{L}(1 - w_k)]} \tag{9}$$

With $\mu = \left[\sum_{j=1}^{N} \prod_{k=1}^{L}(w_k \beta_{jk} + 1 - w_k \sum_{j=1}^{N} \beta_{jk}) - (N-1) \times \prod_{k=1}^{L}(1 - w_k \sum_{j=1}^{N} \beta_{jk}) \right]^{-1}$

The output of the BRB system is not crisp/numerical value. Hence, this output can be converted into crisp/numerical value by assigning utility score to each referential value of the consequent attribute[33] -

$$H(A^*) = \sum_{j=1}^{N} u(C_j) B_j \tag{10}$$

Where $H(A^*)$, is the expected score expressed as numerical value and $u(C_j)$ is the utility score of each referential value

V. Brbes To Assess Uterine Cancer
The design and implementation of the Belief Rule Based Expert System (BRBES) to assess uterine cancer is presented in this section.
5.1. BRBES Architecture
Architectural design represents the structure of data and program components that are required to build a computer-based system. It also considers the pattern of the system organization, known as architectural style. BRBES adopts the three-layer architecture[43], which consist of presentation layer, application layer and data processing layer as shown in figure 1.
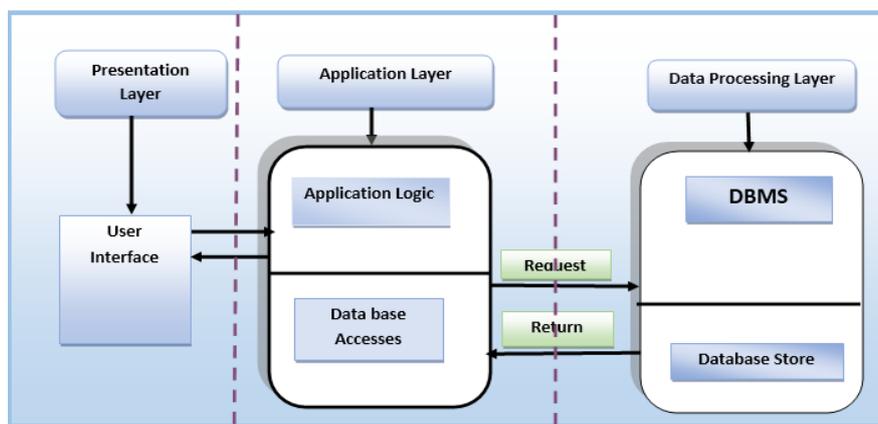
Fig 1: BRBES Architecture

Here, presentation layer takes the input and display the results. Application layer consists of application logic and database access method. It requests data processing layer for data. Data processing layer consists of DBMS and database store.

5.2. Input Value Transformation

The input information that used to perform the uterine cancer suspicion are classified into two categories namely symptoms and signs. The input antecedents consisting of signs and symptoms are D1(Back Pain), D2 (Pelvic Pain), S1(Watery Discharge), S2(Fever), P1(Pre-menopausal), P2(Menopausal/Post-menopausal), N1(Nullipara), N2(Multipara), C1(1st generation), C2(Before 1st Generation), R1(Lower abdomen), and R2(Upper Abdomen)(Binaghi and Madella, 1999). The input values of these antecedent attributes are transformed into referential values by using equations (5) and (6) on the basis of belief degree provided by the physician as mentioned earlier. The input clarifications of this BRB system transformed to referential values is shown in table 2.

Table 2: The Input are Transformed into Referential Value

Sl. No.	Input Antecedent	Expert Belief	Referential Value		
			High	Medium	Low
0	D1	1.0	1	0	0
1	D2	0.5	0.1	0.7	0.2
2	S1	0.8	0.5	0.5	0
3	S2	0.5	0.1	0.8	0.1
4	P1	1	0.8	0.2	0
5	P2	0.5	0.1	0.4	0.5
6	N1	1	0.8	0.2	0
7	N2	1	0.8	0.2	0
8	C1	0.4	0.1	0.5	0.4
9	C2	0.5	0.5	0.4	0.1
10	R1	0.3	0.6	0.3	0.1
11	R2	0.5	0.1	0.8	0.1

5.3. Knowledge Base Construction

In order to construct BRB knowledge base to assess suspicion of Uterine Cancer (UC), a BRB framework has been developed as shown in figure 2) by taking account of opinion of the clinical domain experts. From the framework, it can be observed that input factors that determine this cancer include antecedent attributes such as D1(Back Pain), D2 (Pelvic Pain), S1(Watery Discharge), S2(Fever), P1(Pre-menopausal),P2(Menopausal/Post-menopausal),N1(Nullipara),N2(Multipara),C1(1st generation), C2 (Before 1st Generation), R1(Lower Abdomen), and R2(Upper Abdomen).
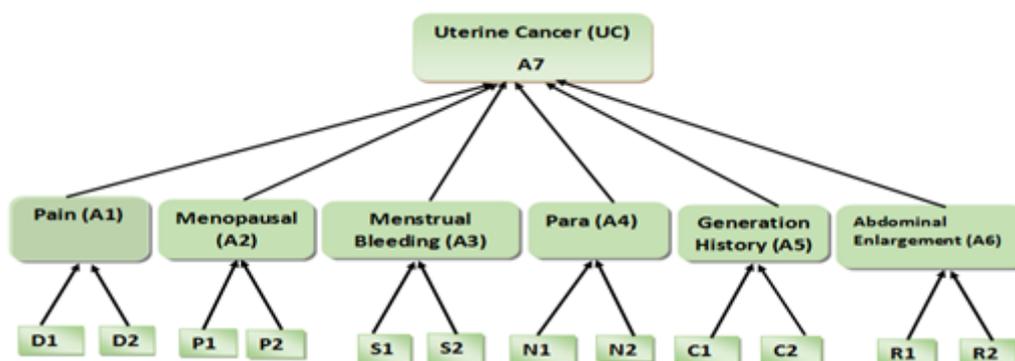
Fig 2: BRB Framework to Assess Uterine Cancer

A BRB can be established in the following four ways[33]- (1) Extracting belief rules from expert knowledge (2) Extracting belief rules by examining historical data; (3) Using the previous rule bases if available, and (4) Random rules without any pre-knowledge. BRB has been developed by using domain expert knowledge. This BRB consists of seven sub-rule-bases namely Pain, Excessive Menstrual Bleeding, Menopausal, Para, Generation History, Abdominal pain.

Table 3: Initial Belief Rules of Sub-Rule-Base (Excessive Menstrual Bleeding)

Rule No.	Rule Weight	IF		THEN		
		S1	S2	Excessive menstrual bleeding		
				High	Medium	Low
0	1	High	High	0.8	0.2	0
1	1	High	Medium	0.4667	0.5333	0
2	1	High	Low	0.0667	0.9333	0
3	1	Medium	High	0	0.9333	0.1
4	1	Medium	Medium	0	0.8	0.2
5	1	Medium	Low	0	0.6667	0.3
6	1	Low	High	0.3333	0.6667	0
7	1	Low	Medium	0	0.9333	0.1
8	1	Low	Low	0	0.8	0.2

Table 4: Initial Belief Rules of Sub-Rule-Base (Menopausal)

Rule No.	Rule Weight	IF		THEN		
		P1	P2	Menopausal		
				High	Medium	Low
0	1	Yes	No	0.333	0.333	0.333
1	1	No	Yes	0.80	0.20	0

Table 5: Initial Belief Rules of Sub-Rule-Base (Abdominal Enlargement)

Rule No.	Rule Weight	IF		THEN		
		R1	R2	Abdominal Enlargement		
				High	Medium	Low
0	1	High	High	0.3	0	0.6667
1	1	High	Medium	0.5	0	0.5333
2	1	High	Low	0.6	0	0.4
3	1	Medium	High	0.7	0	0.2667
4	1	Medium	Medium	0.1	0	0.9333
5	1	Medium	Low	0	0.0667	0.9333
6	1	Low	High	0.1	0	0.9333
7	1	Low	Medium	0.2	0	0.8
8	1	Low	Low	0	0.0667	0.9333

Enlargement and Uterine Cancer. Pain sub-rule-base has two antecedent attributes. Each antecedent attribute consists of three referential values. Hence, this sub-rule-base consists of 9 rules. The entire BRB (which consists of seven sub-rule bases) consists of (9+9+2+2+4+9+216) = 251 belief rules. It is assumed that all belief rules have equal rule weight; all antecedents have equal weight, and the initial belief degree assigned to each possible consequent by two experts (doctor) from accumulated patient data. To better handle uncertainties, each belief rule considered the three referential values including High (H), Medium (M) and Low (L).

An example of a belief rule taken from Table 3 is illustrated below.

R1: **IF** S1 is 'H'AND S2 is 'H'

    THEN Excessive menstrual bleeding is {H (0.8), M (0.2), L (0.00)}

In the above belief rule, the belief degrees are attached to more than two referential values of the consequent attribute. For example, if we have 100 patients whose clinical information such as Pelvic pain is Menstrual Bleeding is H, menopausal is No, Para is Yes, Generation history is Yes and Abdominal enlargement is H; meaning that clinical information of the antecedent parts of the rule is evaluated with high grade of the referential values of the corresponding antecedent attribute of the rule then all patients to be judged as =High'to

Uterine Cancer (UC) symptoms. Consequently, we can assign the initial belief degrees to the consequents in the rule as 1 to H, 0 to M and 0 to L.

5.4. Inference Engine using ER Approach

Evidential Reasoning algorithm [31] [33] has been used as the inference engine for this BRB expert system as elaborated in section 4.2. This inference engine works in the following ways: i) it first reads input data, acquired from the patient or the physician, ii) the input data are then transformed by using (4) (5), iii) it calculates the activation weight of all rules using (6), iv) it updates the belief degree of consequence in rules using (7) and v) it finally aggregates all rules by using (8) (9).

5.5. System Interface

System interface is a media that ensures the interaction between the user and the system. Figure 3 represents the BRBES's interface.

This interface facilitates the acquiring of the leaf nodes' data of the BRB framework (figure 2). Figure 3 illustrates the output for the leaf nodes (D1 =0.8, D2= 0.9) associated with '**pain**' sub-rule-base. It can be observed from figure 3 that the belief degree obtained for the referential values of the consequent attribute "**pain**" of this sub-rule-base is {(High,0.70), (Medium,0.20),(Low, 0.10)}. Similarly, the belief degree of the referential values of other consequent attributes can be understood from figure 4. The degree of belief of the intermediate level's consequent nodes (pain, menopausal, bleeding, para, generation history and abdominal enlargement) can be obtained from their respective child nodes. Intermediate level nodes values cannot be assigned directly like leaf nodes. The belief degrees of root node can be obtained by aggregating lower level to higher level attributes. Figure 3 also illustrates the combined or aggregated assessment, which is {(High, 0.80), (Medium, 0.10), (Low, 0.10)}. This is transformed into crisp value by using equation (10), which is 85% as shown in figure 3.
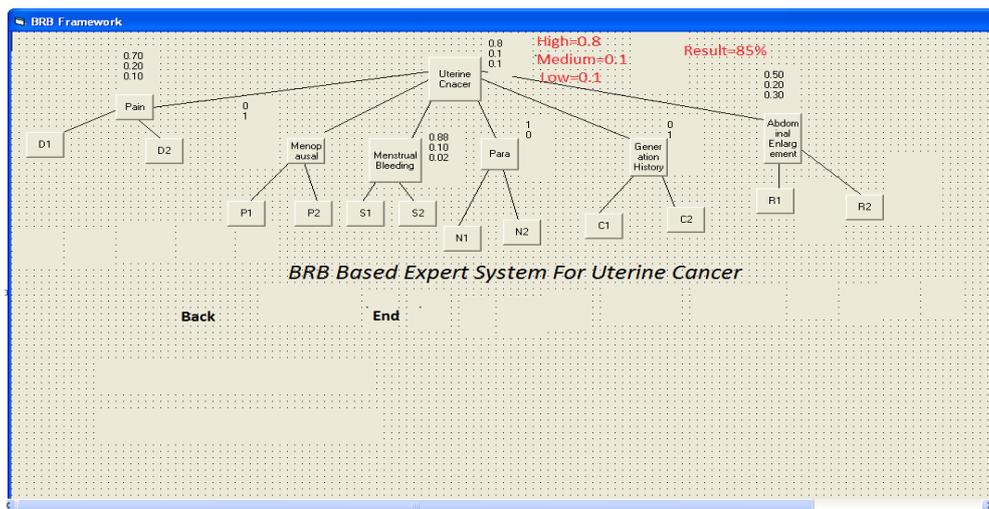
Fig 3: GUI of the Expert System

VI. Result And Discussion

The clinical data set variables used in this paper to determine the suspicion of uterine cancer are patient's clinical signs and symptoms, which include pain, Excessive Menstrual Bleeding, Menopausal, Para, Generation History and Abdominal Enlargement. The real data set of 36 patients having uterine cancer was collected. The data set of five patients with diagnosis outcome is presented as example in table II. This table represents overall uterine cancer diagnosis outcome from patient's information. The result of this system is measured in percentage for recommendation. The output of this system was generated based on output utility equation (10). In this paper, the utility score of 100% assigned to 'High', 50% assigned to 'Medium', and 0% assigned to 'Low'. For example, we can estimate overall system output uterine cancer as 99% , if the Fuzzy result of the system is {(High, 0.90), (Medium, 0.10), (Low, 0)}.

Table 6: Example Data Set from Signs and Symptoms, and Comparison of Results among Manual System, BRBES and Benchmark data

Patient ID	Age	D1	D2	S1	S2	A3	P1	P2	N1	N2	C1	C2	R1	R2	Manual System Result	Benchmark Result	BRB System Result	Risk Stage
1	50	H	H	H	H	H	Y	N	Y	N	Y	Y	H	H	80.44%	92.55%	91.55%	High>70%
2	67	M	L	M	H	M	N	Y	N	Y	Y	N	L	M	67.33%	69.55%	66.99%	Medium<70%
3	47	L	H	M	L	M	Y	N	N	Y	N	N	M	L	57.55%	50%	52.66%	Low<50%
4	52	M	M	H	L	H	N	Y	Y	Y	Y	Y	H	H	85%	90%	87%	High>70%
5	40	L	L	L	M	L	Y	N	Y	N	Y	N	L	M	52.11%	55.23%	54%	Low<50%
6	53	H	H	H	H	H	N	Y	Y	N	Y	Y	H	H	85%	90%	87%	High>70%
7	36	M	L	M	H	M	N	Y	N	N	Y	Y	L	M	52.11%	43.23%	45%	Low<50%
8	63	L	H	M	L	M	N	Y	N	Y	N	N	M	L	61.66%	67.43%	69.88%	Medium<70%

In the case study, the assessment of suspicion of uterine cancer of 36 patients carried out by using BRBES and doctors' opinion (manual system) as shown in table 6. The clinical historical results were considered as benchmark. From table 6 it can be observed that BRBES's generated result has less deviation from benchmark result, while doctor's opinion has much deviation. Hence, it can be argued that BRBES's output is more reliable than manual system. Therefore, it can be concluded that if the assessment of uterine cancer is carried out by using the BRBES, eventually this will play an important role in taking decision to avoid unnecessary costly lab investigation.
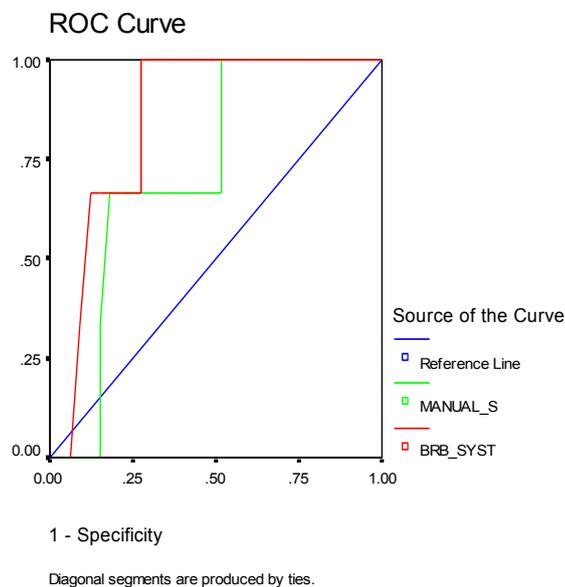
Fig 4: ROC Curves Illustrate the Performance of BRB Expert System and Manual System

Figure 4 illustrates two ROC (Receiver Operating Characteristics Curve) curves – one represents the suspicion performances of the BRB system and the other manual system. The ROC curve in red illustrates the BRB system ACS suspicion result while the yellow curve illustrates the manual system Uterine Cancer suspicion result. The AUC for the BRB system is 0.980 (95% confidence intervals 0.950 – 1.009) and the AUC of the manual system is 0.960 (95% confidence intervals 0.914 – 1.006). From the AUC of the two systems, it can be observed that the AUC of the BRB expert system is greater than the AUC of the manual system. This implies that results generated by the BRB expert system is better than the results generated by the manual system, the latter of which uses traditional rule without taking account of uncertainty. SPSS 11.5 has been used to construct the ROC curve and to calculate the AUC of these curves.

VII.   Conclusion

The development and application of a Belief Rule Based Expert System (BRBES) to assess suspicion of uterine cancer by using signs, and symptoms of patients have been presented. The prototype BRBES is embedded with a novel methodology, known as RIMER, allows the handling of various types of uncertainty and hence, be considered as a robust tool can be utilized in detecting uterine cancer. Consequently, the prototype ES can handle various types of uncertainties found in clinical domain knowledge as well as in signs and symptoms of a patient. Most importantly, the system will play an important role in reducing the cost of lab investigations. The system will facilitate patients in taking precautionary measures well in advance. It can also provide a

percentage of risk recommendation, which is more reliable and informative than from the traditional expert's opinion. The prototype BRBES can only is used to detect the suspicion of uterine cancer by using signs and symptoms of a patient.

In further, research will consider the uterine cancer accurate detection system using signs and symptoms. The validation of the system is necessary by using real patient historical data and the knowledge representation parameters should be trained by real clinical data using the BRB optimal learning training model. At last, the employing of RIMER to develop BRBES for uterine cancer suspicion is a valid new approach and real data can be used to training in future research.

Acknowledgement

References

[1]   World Cancer Statistics April,2014 http://www.wcrf.org/cancer_statistics/world_cancer_statistics.php#Women.
[2]   Cancer Statistics According to Specific Cancer April, 2014 http://www.wcrf.org/cancer_statistics/data_specific_cancers/endometrial_cancer_statistics.php.
[3]   Cancer Statistics by Country April, 2014 http://www.rightdiagnosis.com/e/endometrial_cancer/statscountry.html.
[4]   Uterine or Endometrial Cancer Treatment April, 2014,:http://www.uabmedicine.org/conditions-and.../cancer-gyn-uterine-endometrial.
[5]   Information and Resources for Various Types of Cancer April, 2014:http://www.cancer.org.
[6]   Information about Uterine Cancer and its Diagnosis April, 2014 http://www.cancer.org/cancer/endometrialcancer/detailedguide/endometrial-uterine-cancer-diagnosis.
[7]   Uterine Cancer Mortality Statistics by Age April, 2014 http://www.cancerresearchuk.org/cancerinfo/cancerstats/types/uterus/mortality/uk-uterine-cancer-mortality-statistics#By2.
[8]   Dutta D.C.(2003) "Text book of gynaecology"4th Edition,Published by new central book agency ,calcutta (Book).
[9]   Jeffcoat.(2008) "Jeffcoat principle of Gynaecology"7th edition ,printed and bounded by ajanta offset ltd.,India(Book).
[10]   Sugary Drinks Increases Risk of Endometrial Cancer, April 2014 blog.naturaltherapyforall.com/.../sugary-drinks-increases-risk-of-endometrial-cancer/.
[11]   Menopause related Information April 2014 http://en.wikipedia.org/wiki/Menopause
[12]   Medical Dictionary April, 2014 http://medical-dictionary.thefreedictionary.com/para
[13]   Medical Dictionary about Types of Para April, 2014 http://medical-dictionary.thefreedictionary.com/nullipara.
[14]   Yuan, Y., Feldhamer, S., Gafni, A., Fyfe, F., Ludwin, D., 2002. The development and evaluation of a fuzzy logic expert system for renal transplantation assignment: is this a useful tool? Euro. J. Oper. Res. 142, 152–173.
[15]   Pabitra Mitra, Sushmita Mitra and Sankar K. Pal(2000) "Staging of Cervical Cancer with Soft Computing"IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 47, NO. 7,
[16]   Preoperative Uterine Cancer Diagnosis using Neuro Fuzzy Model April, 2014, ftp://ftp.esat.kuleuven.be/pub/sista/ida/reports/05-196.pdf .
[17]   Tuan Zea Tan, Chai Quek, Geok See Ng, Khalil Razvi ,Ovarian cancer diagnosis with complementary learning fuzzy neural network, Centre for Computational Intelligence, School of Computer Engineering, Nanyang Technological University, Block N4, #2A-32, Nanyang Avenue, Singapore 639798, Singapore,15 April 2008.
[18]   Karem R. Domínguez Hernández,Alberto A. Aguilar Lasserre,Rubén Posada Gómez, José A. Palet Guzmán,and Blanca E. González Sánchez ,Development of an Expert System as a Diagnostic Support of Cervical Cancer in Atypical Glandular Cells, Based on Fuzzy Logics and Image Interpretation, ,17th February,2013.
[19]   Ankita Thakur,1,2 Vijay Mishra,2 and Sunil K. Jain ,Feed Forward Artificial Neural Network: Tool for Early Detection of Ovarian Cancer, Jul 5, 2011.
[20]   Musen, M., Shahar, Y. Shortliffe. E. 2006. Clinical Decision-Support Systems. In: Shortliffe, E.H. Cimino, J.J. (Eds.) Biomedical Informatics.
[21]   Shortliffe, E. H. Computer-Based Medical Consultations: MYCIN, New York, Elsevier , 1976
[22]   Buchanan and Shortliffe, Stanford University , MYCIN expert system, 1984.
[23]   Kumar, K. A., Singh, Y. & Sanyal, S. (2009) Hybrid approach using case-based reasoning and rule-based reasoning for domain independent clinical decision support in ICU. Expert Systems with Applications, 36, 65-71.
[24]   De Dombal, F. T., Leaper, D. J., Staniland, J.R., Mccann, A. P. & Horrocks, J. C. Computer-aided diagnosis of acute abdominal pain. British Medical Journal,2(5804) 5 -9, 1972.
[25]   Warner, H.R.J., 1989. Iliad: moving medical decision-making into new frontiers. Methods Inform. Med. 28, 370–372.
[26]   Graham, T., Bullard, M., Kushniruk, A., Holroyd, B. & Rowe, B. Assessing the sensibility of two clinical decision support systems. Journal of Medical Systems,32, 361-368, 2008.
[27]   Mack, E. H., Wheeler, D. S. & Embi, P. J. Clinical decision support systems in the pediatric intensive care unit. Pediatric Critical Care Medicine, 10, 23-28, 2009.
[28]   E. Binaghi and P. Madella, "Fuzzy Dempster–Shafer reasoning for rule-based classifiers," Int. J. Intell. Syst., vol. 14, no. 6, pp. 559–583, 1999.
[29]   J. B. Yang, Y. M. Wang, D. L. Xu and K. S. Chin, "The evidential reasoning approach for MCDA under both probabilistic and fuzzy uncertainties,"Eur. J. Oper. Res., Vol. 171, No.1, pp.309-343, 2006.
[30]   STEVEN J. HENKIND AND MALCOLM C. HARRISON, "An analysis of four uncertainty calculie" IEEE transtactions on systems…. Vol.18,no.5, September/October 1988.
[31]   Yu-Wang Chen, Jian-Bo Yang, Dong-Ling Xu, Zhi-Jie Zhou, Da-Wei Tang, Inference analysis and adaptive training for belief rule based systems, Expert Systems with Applications 38; 12845–12860, 2011.
[32]   Yang, J. B. ; Sen, P. (1994) A general multi-level evaluation process for hybrid MADM with uncertainty. Systems, Man and Cybernetics, IEEE Transactions on, 24, 1458-1473.
[33]   Yang, J. B., Liu, J., Wang, J., Sii, H. S. & Wang, H. W. (2006) Belief rule-base inference methodology using the evidential reasoning approach -RIMER. IEEE Transactions on Systems Man and Cybernetics Part A-Systems and Humans, 36, 266-285.

[34]   D.L. Xu, J. Liu, J.B. Yang, G.P. Liu, J. Wang, I. Jenkinson, J. Ren,   Inference and learning methodology of belief-rule-based expert system for pipeline leak detection, Expert Systems with Applications 32 (2007) 103–113.

[35]   Mahmud,T.;Hossain,M.S.(2012).”An Evidential Reasoning-based Decision Support System to Support House Hunting.” International Journal of Computer Applications 57(21):51-58,. Published by Foundation of Computer Science, New York, USA.

[36]   Mahmud,T.;Rahman,K.N.;Hossain,M.S.(2013).”Evaluation of Job Offers Using The Evidential Reasoning Approach”. Global Journal of Computer Science and Technology. Volume 13 Issue 2 Version 1.0 . Published by Global Journals Inc. (USA).

[37]   Mahmud,T & Sikder,J (2013) “Intelligent Decision System for Evaluation of Job Offers” 1st National Conference on Intelligent Computing and Information Technology(NCICIT) , November 21, CUET, Chittagong-4349, Bangladesh.

[38]   J. B. Yang, ―Rule and utility based evidential reasoning approach for multi-attribute decision analysis under uncertainties,‖Eur. J. Oper. Res., vol. 131, no. 1, pp. 31–61, 200.

[39]   Grams, R. R. Clinical laboratory test reference (CLTR). Journal of Medical Systems,17,59-67, 1993.

[40]   A. M. Nonvich and T. B. Turksen, “A model for the measurement of membership and the consequences **of** its empirical implementa-tion.”Fuzz,. Sers Syst., vol. 12, pp. 1-25. 1984.

[41]   Wang, Y.M., Yang, J.B., Xu, D.L., 2006. Environmental impact assessment using the evidential reasoning approach. Eur. J. Oper. Res. 174, 1885–1913.

[42]   B. Yang, J. Liu, D.L. Xu, J. Wang, H.W. Wang, Optimal learning method for training belief rule based systems, IEEE Transactions on Systems, Man,and Cybernetics (Part A) 37 (2007) 569–585.

[43]   Roger S. Pressman.(2005) Software engineering: a practitioner‘s approach.— 6th ed. p. cm.— (McGraw-Hill series in computer science) Includes index. ISBN 0 -07-365578-3, p373- 374.