INTELLIGENT FETAL HEART RATE ANALYSIS

J F Skinner[1], J M Garibaldi[2], J Curnow[3], E C Ifeachor[1]

[1]University of Plymouth, UK; [2]De Montfort University, UK; [3]Plymouth Hospitals NHS Trust, UK

ABSTRACT

The cardiotocogram (CTG) consists of a continuous recording of fetal heart rate and maternal contractions during labour. Changes in the fetal heart rate pattern relative to contractions provide an indication of fetal condition. There are two rôles in which the CTG can be used. The first is to identify cases of fetal compromise *during* labour, used by clinicians to determine the need for clinical intervention. The second is as a retrospective record of how labour was managed, used for clinical audit and potentially in litigation. CTG interpretation is a difficult task, requiring clinical experience and significant expertise. A crisp expert system for assessment of the CTG during labour had previously been developed. This paper describes the findings of a research project with two main aims: to investigate whether fuzzy logic could offer an improvement in CTG analysis over the crisp expert system; and to investigate whether retrospective analysis of complete CTG traces could be automated. Two studies are presented and the findings of each are summarised and discussed. It was found that fuzzification of the crisp expert system improved system performance and that, although fuzzy logic methodology is promising, retrospective analysis of the CTG requires considerably more work before reaching the status of a useful clinical tool.

INTRODUCTION

Childbirth is a critical time for infant and mother. Clinical outcome is usually good for both, but problems may occur that result in permanent fetal organ damage or even death. A number of fetal monitoring technologies have been developed which seek to assist clinical decision making during this period. The most common of these is the cardiotocogram (CTG), which consists of a continuous recording of fetal heart rate and maternal contractions. Changes in the fetal heart rate pattern relative to contractions provide an indication of fetal condition which can be used to identify those cases needing intervention. CTG interpretation is a difficult task, requiring clinical experience and significant expertise. Studies have shown that this expertise is often lacking in delivery units, with CTG misinterpretation implicated in a large number of preventable fetal deaths (CESDI, 1) and unnecessary operative interventions (Neilsen and Grant, 2).

As a result, many computerised systems have been developed to encapsulate expert interpretation of the CTG (3-10). These range from simple feature extraction and classification systems to intelligent expert systems that assess the CTG along with clinical information to provide management advice. Preliminary evaluation of these systems has shown some promising results, but despite over three decades of development none has been adopted into widespread use or demonstrated to have clinical benefit. One of the main problems that has impeded progress is the inherent uncertainty in clinical knowledge relating to CTG interpretation. This uncertainty has not been effectively represented in any automated CTG system developed to date.

It is also becoming increasingly important to have objective information relating to all areas of clinical practice. Trends such as audit and evidence based medicine require objective information and interpretation of clinical data, as does the growing area of medical litigation. A summary analysis of the whole of labour based on the CTG has great potential to provide this objective data. Furthermore, an interpretation of fetal response to labour would provide useful information for planning neonatal care and allow instant feedback to labour-ward staff on their clinical management. Retrospective assessment also avoids many of the pitfalls of real-time decision support and may lead to better understanding of the CTG.

The aims of this work are to investigate uncertainty handling in a rule based system for CTG analysis and to examine how retrospective analysis may be used to summarise whole CTG data.

UNCERTAINTY HANDLING IN CTG ANALYSIS

One of the most widely used techniques for handling uncertainty is fuzzy logic (Zadeh, 11). This has shown benefit for decision support systems in many application areas but has yet to be applied to CTG analysis. In this paper we present the development of a fuzzy expert system for CTG assessment, including preliminary system evaluation and demonstrate how the system can be used to produce an objective report on fetal condition during labour. The objective of this work was to determine whether fuzzification of the existing crisp system could create a better model of current clinical knowledge, with the potential to improve system performance.

Methods

The developed fuzzy system is an evolution of the front-end feature classification module of an existing crisp CTG expert system (Keith et al., 7). This crisp system has been previously validated and is undergoing continuous development and refinement. An example of 15-minutes of CTG trace, with the important clinical features labelled, is shown in Figure 1. The crisp system extracts the basic features in 5 minute segments; *Baseline*, *Variability*, *Accelerations*, *Decelerations* and *Contractions,* which are classified according to the following clinical terms:

- *Baseline {Bradycardia, Slight Bradycardia, Normal, Slight tachycardia, Tachycardia}*
- *Variability {Absent, Reduced, Normal, Increased}*
- *Accelerations {Absent, Present}*
- *Decelerations {Absent, Present, Severe}*

There are thus 120 possible feature combinations, although clinically speaking some are extremely unlikely to occur. Each possible feature combination is classified overall using the terms:- *Classification {Normal, Intermediate, Abnormal, Severely Abnormal}*. The crisp expert system considers the overall segment and feature classifications in the context of clinical events to provide management recommendations. The feature classification is based on well-defined recommendations, but feature identification is linguistic and inexact. For example, a normal heart rate baseline is defined as 110-160 bpm, however there are no precise definitions of how to identify the baseline. In consequence, the feature classification and any interpretation based on this classification will be uncertain.

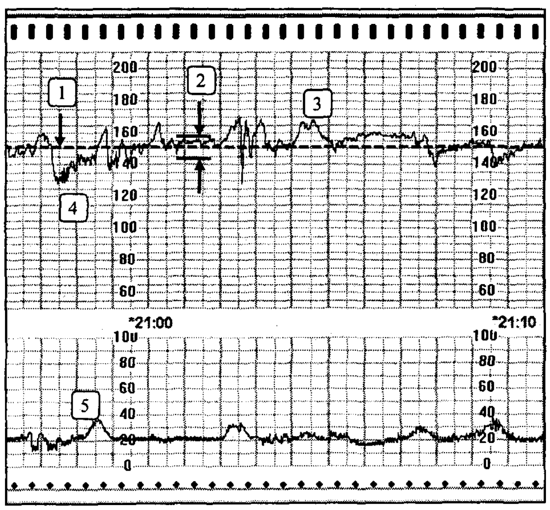

Figure 1: An example 15 minute segment of cardiotocogram with important features. The top trace is fetal heart rate and the bottom trace maternal contractions. The features of interest are: (1) *Baseline*, the basic heart rate value about which the heart rate pattern fluctuates; (2) *Variability*, peak to peak amplitude of high frequency perturbations about the baseline; (3) *Accelerations*, relatively long term transient increases in heart rate from the baseline; (4) *Decelerations*, relatively long term transient decreases in heart rate from the baseline; (5) *Contractions*

Feature characterisation. The simplest parameters to represent in the fuzzy system are the baseline and variability. There is no standard method to calculate either but there is good consensus on the clinical classification terms and the ranges of values associated with each term (Rooth et al., 12) (see Tables 1 and 2).

Since the existing baseline and variability algorithms are considered to be good, membership functions for the corresponding fuzzy terms were created directly from the crisp terms. The position and width of the terms were determined by the cut-offs in the crisp definitions. For example the transition from *Baseline {Slight Bradycardia}* to *Baseline {Normal}* is at 110 bpm. The term-sets for fuzzy baseline are shown in Figure 2. Sigmoid curves were used as they were felt to represent the clinical model more realistically than triangular functions.

TABLE 1 - Classification terms for heart rate baseline

Variability (beats per minute)	Crisp Classification
<2	Absent
2-5	Reduced
6-25	Normal
>25	Increased

TABLE 2 - Classification terms for heart rate variability

Baseline (beats per minute)	Crisp Classification
<90	Bradycardia
90-109	Slight Bradycardia
110-159	Normal
160-179	Slight Tachycardia
>180	Tachycardia

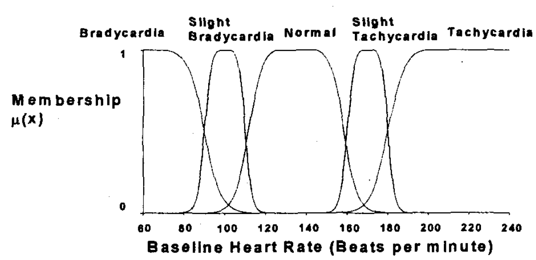

Figure 2: Membership sets for baseline heart rate

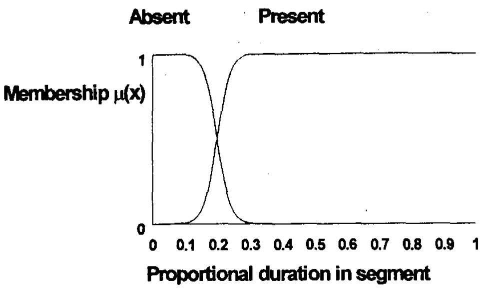

Figure 3: Membership sets for accelerations and each of the types of decelerations

In the crisp system accelerations are identified by their duration and height above the baseline and classified overall for the segment as *Accelerations {Absent, Present}*. This assessment does not allow variation in acceleration pattern to be represented. In the fuzzy system the terms *Accelerations {Absent, Present}* are defined according to the total proportional duration of identified accelerations in a segment. The terms for the *Accelerations* membership function are shown in Figure 3. These are also sigmoid functions where the intersection of the terms corresponds to the cut-off duration of a single acceleration in the crisp system.

Individual decelerations are classified by duration, area and depth below the baseline and their timing relative to a maternal contraction using the terms; *Deceleration {Absent, Early, Late, Severe Early, Severe Late}*. All identified decelerations are assessed using a small rule set to provide the overall deceleration classification:- *Deceleration Class {Absent, Present, Severe}*. In the fuzzy system, each of the four types of deceleration are described by their duration in the segment in the same manner as accelerations (see Figure 3). The existing deceleration rule set is then applied to these terms to give the overall deceleration classification:- *Deceleration Class {Absent, Present, Severe}*. The membership functions for the deceleration classification output term-sets are sigmoid functions with arbitrarily determined width and position.

Rule determination. The rules used in the fuzzy system are identical to those in the two crisp sub-systems for deceleration and overall segment classification. The in-house fuzzy system development shell allows either crisp or fuzzy input variables, i.e. scalar values or fuzzy membership functions. This allows the output set of one fuzzy system to be used as input to another without intermediate de-fuzzification. The deceleration classification output set can thus be forward chained directly into the overall segment classification. This has the advantage of retaining the initial uncertainty in deceleration classification when the second rule set is applied.

Rules for deceleration classification were used directly, but constructing 120 explicit rules for the segment classification was felt to be extremely inefficient given that there are only four possible output categories. To minimise the number of rules required, the ID3 rule induction algorithm (Quinlan, 13) was applied to the existing rule set. This reduced the number of rules required to classify all possible inputs from 120 to 33. The segment classification rule set thus uses the fuzzy variables *Baseline, Variability, Accelerations* and the fuzzy set *Deceleration Class*. The terms of the overall classification are:- *Segment Classification {Normal, Intermediate, Abnormal, Severely Abnormal}*. The membership functions for this output term-set are sigmoid with arbitrarily determined width and position. Max-Prod inference is used in both rule based sub-systems. The segment classification output set is defuzzified using the centroid method to give a scalar value (0-100), to provide an overall index of the segment abnormality.

Evaluation of segment classification

Methods. Preliminary evaluation was performed in a study comparing the assessment of the fuzzy and crisp systems to human experts. Ninety five 15-minute segments of CTG trace were chosen from a database of approximately 6,500 hours of digitised CTG data. The segments selected had not been used in system development and were chosen to represent as many of the 120 potential combinations of crisp features as possible. Fifteen minute segments were used as this was considered to be the minimum clinically useful length of trace.

The selected segments were independently evaluated by three clinicians. Two were senior grade obstetricians, the third a Professor of clinical physiology. All three are actively involved in fetal monitoring research and have significant experience of CTG interpretation. The reviewers were required to assess each 15-minute segment and allocate a score from 0-39, indicating how normal or abnormal they considered the heart rate pattern to be. The relationship between reviewer scores and overall linguistic categorisation of the trace is shown in Table 3. These linguistic categories mirror the output of the crisp system and were used to guide the reviewers.

Since the two expert systems work on 5-minute segments it was necessary to combine the assessments of

TABLE 3 - Scoring protocol for reviewers showing the relationship between score and linguistic description of the trace

Reviewer Score	Linguistic Classification
0-9	Normal
10-19	Intermediate
20-29	Abnormal
30-39	Severely abnormal

TABLE 4 - Ranking of three 5-minute segment combinations of crisp system output classification

Classifications			Rank	Classifications			Rank
N	N	N	1	I	I	S	11
N	N	I	2	I	A	A	12
N	N	A	3	N	A	S	13
N	N	S	4	I	A	S	14
N	I	I	5	A	A	A	15
I	I	I	6	A	A	S	16
N	I	A	7	N	S	S	17
N	I	S	8	I	S	S	18
I	I	A	9	S	S	A	19
N	A	A	10	S	S	S	20

TABLE 5 - Mean and SD of scores for each reviewer

Reviewer	A	B	C
Score Mean	24.23	12.44	15.42
Score Std Dev	10.17	9.39	11.04

three consecutive segments to provide a 15-minute assessment. In this way the systems could be compared directly with the reviewers. For the fuzzy system the score for a 15-minute segment was calculated as the mean of the scores for three 5-minute segments. For the crisp system a more complex combination procedure was required due to the fact that the crisp system produces an ordinal classification of *Normal (N)*, *Intermediate (I)*, *Abnormal (A)* or *Severely Abnormal (S)* for each five minute segment. Each possible combination of three 5-minute classifications were ranked from most normal to most abnormal by rank-ordering each combination pair (see Table 4). For example, consider the combination *NNS* and *NII*. For this pair *N* would rank 2 (three *N*'s tie for ranks 1, 2 and 3), *I* would rank 4.5 (two *I*'s tie for ranks 4 and 5) and *S* would rank 6. Summing these ranks gave 10 for *NNS* and 11 for *NII*, hence *NNS* was ranked lower than *NII*.

Results. The mean scores for each reviewer are shown in Table 5. The distribution of each set of scores was analysed using ANOVA and showed that A's mean is higher than B's ($p < 10^{-13}$) and C's ($p < 10^{-7}$) at the 5% significance level. B and C's means are not significantly different ($p > 0.05$).

The traces were then ranked by score for each reviewer and by the output scores derived for the two systems. The agreement in case ranking between each reviewer and the systems, calculated by the Spearman rank correlation statistic with correction for tied ranks, is shown in Table 6.

The use of absolute scoring systems to assess CTG traces has been known to be problematic for many years, often with wide inter-observer variation. This is highlighted in the difference in scores between reviewer

TABLE 6 - Spearman rank correlation (R) of segment rankings for each reviewer and the two systems

Reviewer	A	B	C	Crisp System	Fuzzy System
A	-	0.80	0.37	0.51	0.73
B	0.80	-	0.59	0.49	0.67
C	0.37	0.59	-	0.46	0.50
Crisp System	0.51	0.49	0.46	-	0.83
Fuzzy System	0.73	0.67	0.50	0.83	-

A and reviewers B and C. Overall, reviewer A has assigned significantly higher scores to the traces than reviewers B and C. If such a scoring system were used as the basis for clinical decision making, the scores assigned by A would result in approximately twice as many interventions as those of B and C, for a given threshold.

However, when trace ranking is used to examine relative assessment, reviewer A achieves good correlation with the others, particularly reviewer B. This implies that common criteria are being used for relative assessment of traces, but the assignment of an overall linguistic description to a trace is inconsistent. The difference in meaning that domain experts attach to linguistic classifications is a significant problem for knowledge engineers and system designers.

Reviewer C has much lower agreement in ranking than the other reviewers. It is interesting to note that this reviewer, while experienced in CTG assessment, is not involved in CTG interpretation and decision making in the clinical setting. It is likely that different criteria for assessment are being used, which may not be directly applicable for clinical use. The two practising clinicians, A and B, have very high agreement in ranking.

The agreement of the fuzzy system with each of the reviewers, particularly the two practising clinicians, is higher than that of the crisp system. This result is encouraging and demonstrates the potential of fuzzy techniques to improve on the performance of crisp rule-based systems. The fuzzy system assessment has a very high correlation with that of the crisp system. This is likely to be because the current fuzzy system is an evolution of the crisp system, using the same rules and with membership functions that are largely based on existing crisp classifications.

It is impossible to predict clinical outcome or to produce full management recommendations from a single 15-minute segment of trace, unless it is catastrophically pre-terminal. However, since the fuzzy score agrees well with experts, a continuous CTG assessment based on this measure has the potential to simplify the information presented to clinicians without loss in meaning. Clinicians are used to assessing the CTG, but it is an arbitrary method of presenting information about the fetal heart rate. Simpler data in the form of an abnormality score is potentially just as effective as the CTG at identifying first line abnormality. Appropriate protocols for assessing the continuous CTG index could then form the basis of management decision support without requiring the system to provide a definitive interpretation of the trace.

RETROSPECTIVE ASSESSMENT OF WHOLE CTG RECORDS

Having determined that the fuzzy CTG system provided a better classification of individual segments than the crisp system, the feasibility of using the fuzzy score to assess a whole CTG trace was investigated. The aims of this study were to establish:

- whether any correlation exists between the fuzzy CTG abnormality index and newborn infant condition, and
- if a continuous fuzzy abnormality index for the CTG could potentially reduce the amount of data that the clinician is required to assess.

Comparison with Clinical Outcome

Methods. The fuzzy CTG abnormality score was generated for each 5-minute segment in 1,123 digitised traces producing 79,030 segment scores. The abnormality scores for a typical trace are shown in Figure 4.

Objective measurement of newborn infant condition is a complex problem and for this study umbilical cord blood acid-base analysis at delivery is used. While this provides an objective measurement, the four-dimensional data (pH artery, base deficit artery, pH vein, base deficit vein) still causes significant interpretation difficulties. A separate fuzzy expert system to interpret blood gas results has been developed in

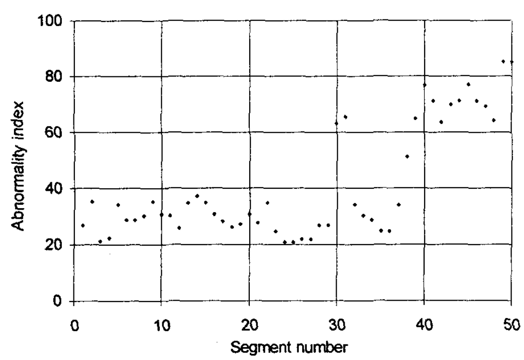

Figure 4: Example of time sequence of the fuzzy CTG abnormality index for a single case

Plymouth (Garibaldi, 14). This system has been previously validated against clinical expert opinion and provides a continuous metric called the Possibility of Intrapartum Asphyxial Damage (PIAD). This ranges from 0 to 1 and indicates how compatible the blood gas results are with damage from oxygen deprivation during birth.

To allow easier visual comparison between the fuzzy CTG abnormality index and PIAD scores, the CTG score was normalised to the range 0 to 1. The resulting index is called the Possibility of an Abnormal Trace (PAT) and represents the possibility that the CTG is compatible with the clinical description 'abnormal trace'. As the PAT is ordinal data the median PAT index for a trace, or for the last hour of a trace was used.

In order to allow a comparison with the crisp expert system, again a technique needed to be found to combine the crisp system 'raw' output (*N*, *I*, *A*, or *S* for each 5-minute segment) into an overall assessment for a complete CTG trace. This was achieved using the same process of repeated pair-wise rank ordering described earlier, except that mean ranks were used to account for the varying number of segments in different CTG's.

Results. The relationship between the fuzzy system's median PAT score for each trace and the PIAD is shown in Figure 5. As can be seen, the relationship is far from straight forward. Consequently, the median PAT score in the last hour of trace was also examined (Figure 6), as an assessment of the last hour of CTG may be more closely associated with clinical outcome.

The Spearman rank correlation of the outputs of the crisp and fuzzy systems, for both the whole trace and the last hour of trace, against the PIAD was calculated, as shown in Table 7. Three things can be seen from these results. Firstly, there is weak correlation between the expert systems' assessment of CTG abnormality and clinical outcome expressed as PIAD. This indicates that neither of the systems is particularly useful as a predictor of outcome. Secondly, the correlation of the last hour's assessment by each system is slightly higher than that for the whole trace. Since the CTG trace immediately prior to delivery should better reflect fetal condition at birth, this indicates that, although the absolute correlations are low, the systems are indeed assessing CTG abnormality to a degree. Thirdly, the fuzzy expert system achieves better correlation than the crisp system in each case.

TABLE 7 - Spearman rank correlations of crisp & fuzzy systems taken over the whole CTG trace & last hour of trace against clinical outcome (expressed by PIAD)

System	Whole CTG trace	Last hour of CTG trace
Fuzzy	0.27	0.29
Crisp	0.20	0.26

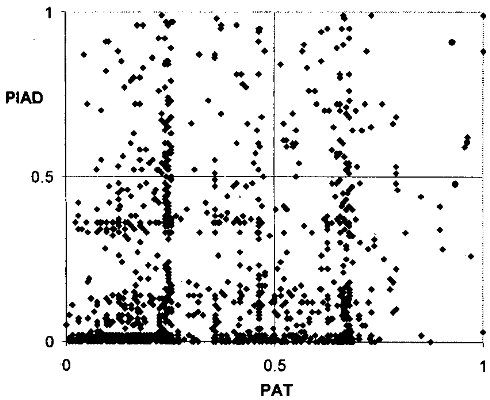

Figure 5: Possibility of Intrapartum Damage vs median Possibility of Abnormal Trace

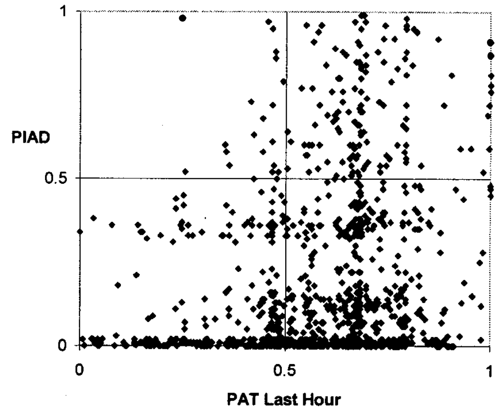

Figure 6: Possibility of Intrapartum Damage vs median Possibility of Abnormal Trace in last hour

In medicine it is important to be able to evaluate the performance of a diagnostic test. There is no perfect test, i.e. one that correctly identifies all cases as being healthy or diseased against some clinical 'gold standard'. To assess test performance the measures *sensitivity*, *specificity*, *positive predictive value* and *negative predictive value* are used. Sensitivity is the is the proportion of cases correctly identified as being diseased out of all cases that are actually diseased. Specificity is the is the proportion of cases correctly identified as being healthy out of all cases that are actually healthy. Positive and negative predictive value indicate how well a positive or negative test result predicts outcome correctly. The CTG is known to have a high sensitivity and low specificity, i.e. poor clinical outcome almost always has an abnormal trace, but an abnormal trace has low association with poor outcome. A normal trace almost always results in good clinical outcome. To evaluate the use of the fuzzy PAT measure for prediction of outcome, sensitivity and specificity analysis were performed using the PIAD index and median PAT index for the whole trace and the last hour of trace. The classification cut-offs were chosen at PIAD=0.5 and PAT=0.5. Cases with less than 2 hours of CTG trace were excluded from analysis of the last hour score. The results are shown in Table 8.

TABLE 8 - Sensitivity, specificity, positive and negative predictive value of the fuzzy PAT score for clinical outcome (PIAD)

Measure	Median PAT score	Median PAT score in last hour
Sensitivity	0.39	0.86
Specificity	0.73	0.31
NPV	0.88	0.93
PPV	0.19	0.17

The median PAT score in the last hour of trace agrees well with existing knowledge of the CTG; a normal trace is highly predictive of good outcome (high NPV), and a poor outcome is almost always associated with an abnormal trace (high sensitivity). However the specificity and positive predictive value are low. It is therefore not possible to use the last hour PAT score to identify abnormal cases with any degree of accuracy.

The median PAT score for the whole trace produced quite different results. A low PAT score is still reasonably predictive of good outcome, but less so than for the last hour of trace (0.88 predictive cf. 0.93). More strikingly, few cases with a poor outcome have a poor overall average HR trace (0.39 of cases cf. 0.86 of last hour traces). The implication of this result is that the CTG trace at the end of labour correlates with outcome better than the whole trace.

Comparison with Clinical Experts.

Methods. To establish whether the fuzzy CTG system models current clinical knowledge, a preliminary comparison was made between the fuzzy CTG system, the crisp system and clinical expert assessment. In a previous evaluation of the crisp system, 487 complete CTG traces were reviewed by three expert obstetricians; two consultants and one senior registrar from three centres. Each 15-minute segment of trace was scored from 1-5 depending on the reviewer's concern for fetal condition. Although the score protocol was based around clinical management, the scores from 1-5 are ordinal and reflect increasing concern for the fetus. The reviewers had access to a full clinical history and the important labour events of each case, as well as fetal scalp blood sample results where available. Any correlation between the system and reviewer scores would therefore be extremely encouraging since the current system bases its assessment on CTG data only.

The median fuzzy CTG score was calculated for each case and used to rank all 487 traces. The reviewer rank order was found from the mean rank of all the scores in each case. The rank order of the crisp system assessment was determined by repeated pair-wise comparison, as before.

Results. The Spearman rank correlation (*R*) between the systems and the reviewers' scores was calculated, as shown in Table 9. It can be seen that the three clinicians agree relatively well on these cases, and that the fuzzy system had a higher correlation with each of the three experts than the crisp system. The fuzzy system also had a high absolute correlation with the crisp system. This is to be expected since the rules used for segment classification were identical in both systems.

TABLE 9 - Spearman rank correlation (R) of cases for three clinical experts (D,E,F) and crisp & fuzzy expert systems

Reviewer	D	E	F	Crisp System	Fuzzy System
D	-	0.71	0.73	0.60	0.67
E	0.71	-	0.72	0.56	0.62
F	0.73	0.72	-	0.56	0.61
Crisp System	0.60	0.56	0.56	-	0.84
Fuzzy System	0.67	0.62	0.61	0.84	-

DISCUSSION AND CONCLUSIONS

When considering CTG analysis it is important to remember that the CTG is just a recording of a physiological variable. The performance of CTG assessment is totally dependant upon the clinical knowledge used to interpret it. Knowledge elicitation and representation are crucial in the future development of automated assessment. This work has demonstrated that existing clinical knowledge is better represented in a fuzzy expert system than in a crisp expert system. Comparison of a crisp expert system for umbilical acid-base analysis (Garibaldi, 15) with its derivative fuzzy expert system (Garibaldi, 14) has previously shown a similar improvement through the use of fuzzy logic.

The low correlation between the generated CTG index and clinical outcome suggests that the fuzzy system cannot be used to predict outcome reliably. However, it would be surprising if the heart rate trace assessment correlated strongly with clinical outcome since this would assume that clinical management has no effect on outcome. It must be remembered that these are retrospective cases where outcome has already been determined to some extent by management of labour, which in turn will have been affected by clinical interpretation of the CTG trace. The fact that there is a positive correlation between the fuzzy CTG index (PAT) and fuzzy umbilical acid-base index (PIAD) is encouraging.

The fuzzy score produced by the system has good correlation with the assessment of clinical experts and also the existing crisp system. A generated measure of the whole trace, the Possibility of an Abnormal Trace (PAT), models the current clinical reality of CTG interpretation in that it has a high sensitivity and low specificity at predicting clinical outcome. Furthermore a low PAT score for a trace is highly predictive of good clinical outcome. To this extent, retrospective CTG assessment shows some promise and on-going research is exploring new ways of representing the complex information in the CTG in a simpler format, including automated linguistic reporting.

Most of the existing knowledge about the CTG is based on empirical observations made in the 1960's that certain heart rate features correlate with poor outcome. Applying these guidelines does not guarantee successful assessment and it is apparent that there is still much to learn about fetal heart rate. Recent advances in machine learning and data mining may allow new relationships between fetal heart rate features and clinical outcome to be detected using large databases of cases. Any new knowledge will have inherent uncertainty which will need to be represented using a technique such as fuzzy logic.

REFERENCES

1. CESDI, 1995, "Confidential Enquiry into Stillbirths and Deaths in Infancy. Annual report for 1 January - 31 December 1993. Part 1: Summary of methods and main results", Department of Health, UK

2. Neilsen, J.P., and Grant, A.M., 1993, In "Intrapartum fetal surveillance", Spencer, J.A.D. and Ward, R.H.T. (eds), RCOG Press, London, 77-93

3. Chung, T.K.H., Mohajer, M.P., Yang, Z.J., et al., 1995, Br J Obstet Gynaecol, 102, 454-460

4. Dawes, G.S., Moulden, M., Redman, C.W.G., 1991, J Perinat Med, 19, 47-51

5. Ulbricht, C., Doffner, G., Lee, A., 1998, Artificial Intelligence in Medicine, 12, 271-84

6. Krause, W., 1990, In: "Computers in perinatal medicine", Maeda, K. (ed), Elsevier Science Publishers BV Amsterdam, 103-111

7. Keith R.D.F., Beckley, S., Garibaldi, J.M., et al., 1995, Br J Obstet Gynaecol, 102, 688-700

8. Bernardes, J., Ayres-de-Campos, D., Costa-Pereira, A., et al., 1998, Int J Gynecol Obstet, 62, 141-147

9. Alonso-Betanzos, A., Devoe, L., Castillo, R., et al., 1989, Artificial Intelligence in Medicine, 1, 93-99

10. Hamilton, E., Kiminani, E.K., 1994, In "Analysis of Complex Data and Artificial Intelligence in Obstetrics", Chang and Rogers (eds), Bailliere's Clinical Obstetrics and Gynaecology (International Practice and Research), Bailliere Tindall, London, 567-581

11. Zadeh, L.A., 1983, Fuzzy Sets and Systems, 11, 199-227

12. Rooth, G.A., Huch, A., Huch, R., 1987, Int J Gynecol Obstet, 25, 159-167

13. Quinlan, J.R., 1986, Machine Learning, 1, 81-106

14. Garibaldi, J.M., Ifeachor, E.C., 1999, IEEE Trans Fuzzy Systems, 7, 72-84

15. Garibaldi, J.M., Westgate, J.A., Ifeachor, E.C., 1999, Artificial Intelligence In Medicine, 17, 109-130