ELSEVIER

Artificial Intelligence in Medicine 10 (1997) 129–144

Artificial Intelligence in Medicine

The development and implementation of an expert system for the analysis of umbilical cord blood

Jonathan M. Garibaldi [a,*], Jennifer A. Westgate [b], Emmanuel C. Ifeachor [a], Keith R. Greene [c]

[a] *School of Electronic, Communication and Electrical Engineering, University of Plymouth, Drake Circus, Plymouth, PL4 8AA, UK*

[b] *Department of Obstetrics and Gynaecology, University of Auckland, Middlemore Hospital, Private Bag 93311, Otahuhu, Auckland, New Zealand*

[c] *Postgraduate Medical School, University of Plymouth, Drake Circus, Plymouth, PL4 8AA, UK*

Received 13 November 1996; received in revised form 1 January 1997; accepted 10 February 1997

Abstract

An assessment of neonatal outcome may be obtained from analysis of blood in the umbilical cord of the infant immediately after delivery. This can provide information on the health of the newborn infant, guide requirements for neonatal care, and is recommended practice of the Royal College of Obstetricians and Gynaecologists. However, there are problems with the technique. Samples frequently contain errors in one or more of the important parameters, preventing accurate interpretation and many clinical staff lack the expert knowledge required to interpret error-free results. In this paper the development and implementation of an expert system to overcome these difficulties is described. The expert system validates results, provides a textual interpretation and archives all results to database for audit, research and medico-legal purposes. The system has now been in routine clinical use for over 3 years in Plymouth, and has also been installed in several other hospitals in the UK. Results are presented in which the types and frequency of errors are established and the user acceptance of the system is determined. © 1997 Elsevier Science B.V.

* Corresponding author. Present address: Centre for Intelligent Systems, Charles Cross Centre, Constantine Street, Plymouth PL4 8DE, UK. Tel.: +44 1752 233514; fax: +44 1752 233517; e-mail: jong@cis.plym.ac.uk

PII S0933-3657(97)00390-4

Keywords: Expert systems; Clinical implementation; Umbilical cord acid–base balance; Neonatal outcome

1. Introduction

The umbilical cord vein carries blood from the placenta to the fetus and the two smaller cord arteries return blood from the fetus. The blood from the placenta has been freshly oxygenated, and has a relatively high partial pressure of oxygen (pO_2) and low partial pressure of carbon dioxide (pCO_2). Oxygen in the blood fuels aerobic cell metabolism, with carbon dioxide produced as 'waste'. Thus the blood returning from the fetus has relatively low oxygen and high carbon dioxide content. Some carbon dioxide dissociates to form carbonic acid in the blood, which increases the acidity (lowers the pH). A degree of oxygen starvation for the fetus is a routine occurrence in normal labour. If oxygen supplies are too low, anaerobic (without oxygen) metabolism can supplement aerobic metabolism to maintain essential cell function, but this produces lactic acid as 'waste'. This further acidifies the blood, and can indicate serious problems for the fetus.

Samples of blood may be taken from blood vessels in the umbilical cord of the neonate immediately on delivery, and a blood gas analysis machine measures the pH, partial pressure of carbon dioxide (pCO_2) and partial pressure of oxygen (pO_2). A parameter termed base deficit of extracellular fluid (BDecf) can be derived from the pH and pCO_2 parameters [18]. This can distinguish the cause of a low pH between the distinct physiological conditions of 'respiratory acidosis', due to a short-term accumulation of CO_2, and a 'metabolic acidosis', due to lactic acid from a longer-term oxygen deficiency. Analysis of the acid–base balance of arterial and venous blood from a clamped umbilical cord provides objective information on the severity and duration of any lack of oxygen during labour. Such assessment of the acid–base status of umbilical cord blood has recently been recommended by the Royal College of Obstetricians and Gynaecologists [14].

There are, however, a number of difficulties with the procedure. The cord arteries are very small in comparison to the vein, which can lead to difficulties in obtaining an arterial sample of adequate volume. Due to the narrow diameter of the artery, it is also possible to stick the needle right through the arterial wall and accidentally sample the vein. Two samples, supposedly from each of the artery and vein, can thus actually be from the same vessel, which is usually the vein due to its much larger size. Once the samples are taken it is possible for the pO_2 and pCO_2 values to alter through exposure to air. Blood gas analysis machines require regular internal calibration and external quality control checks to ensure continuing accuracy and precision to the manufacturer's specifications, and failure to perform this routine maintenance can lead to erroneous results.

Cord blood gas analysis has provoked much debate. Many people have proposed a correlation between the acid–base status at delivery and other measures of neonatal condition, including long-term neurological development [5,6,12], and

others have disputed this [4,15]. These differing opinions are possibly caused by the failure to recognise sampling errors, the failure to distinguish arterial samples from venous samples or to distinguish metabolic acidosis from respiratory acidosis [11,8] and a lack of consensus as to what constitutes significant cut-off points for low pH or high BDecf [2,23]. Unfortunately, few of these studies have addressed the problem of the quality of the basic data.

During a recent trial in Plymouth [21], routine cord blood sampling on every delivery was initiated. Careful retrospective analysis of the cord blood gas results highlighted a 25% failure rate to obtain arterial and venous paired samples with all parameters [22]. This sampling error rate is broadly in line with other studies in which the importance of paired samples was recognised and this is despite the fact that the sampling took place within a research study which featured regular staff training sessions. The study also highlighted the facts that clinical staff were not very good at identifying sampling errors (for example asterisks alongside a parameter to indicate it was unreliable were frequently ignored), did not recognise the occurrence of two samples from the same vessel and were poor at interpreting the results.

Many medical expert systems have been developed and described in the last 30 years [17,3], but the number reaching routine clinical use have been few [13,1]. Some of the main reasons for the lack of clinical implementation include: (i) failing to address the need of the users; (ii) appearing to remove the decision making process from the clinician; and (iii) failing to address concerns about adequate validation. The use of a microcomputer for the diagnosis of complex acid–base disorders [16] and an expert system for the interpretation of blood gas analysis in the intensive care unit [24] have previously been described, but both these systems deal only with adult acid–base measurements. The neonate has substantially different blood characteristics from the adult, and therefore umbilical cord acid–base analysis requires specific knowledge of fetal physiology [20].

This paper describes the development and implementation of an on-line expert system for the validation and interpretation of acid–base data for blood taken from the umbilical cord of the neonate immediately after delivery. The project lifecycle was undertaken in compliance with BS5750 quality assurance standards in order to satisfy the requirements of the collaborating blood gas analyser manufacturer. The stages of feasibility study, knowledge elicitation, design and development of the complete system, including the user-interface and actual expert system, and the procedures for software testing, verification and validation are described in detail. The expert system was implemented in the obstetric unit of the local hospital in Plymouth, was externally validated at a nearby hospital in Exeter and has since been installed at over a dozen other hospitals in the UK. A survey of several years results from Plymouth is included, in which the usage of the system and the classifications of the expert system are examined. Finally, some of the factors which have contributed to the successful implementation and user-acceptance of the system are discussed.

2. System requirements and development

2.1. System specification

A feasibility study was initially carried out in the 'C' language on an IBM compatible computer running DOS, in which the communication protocols between the computer and blood gas analyser were investigated and developed. Once the practicality of an on-line connection was established, a high level specification was drawn up to detail the functioning of the program to the informal description of the clinicians. A modular description of the system is shown in Fig. 1. At this stage, the program development was moved to the Microsoft Windows environment, as this was seen as being more suited to the non-sequential aspects of the program identified during the feasibility study and more advanced as a user-interface. As the program was intended to be used by clinical staff, it was designed to be as straight forward and simple to use as possible. The content and appearance of each screen was developed in conjunction with the clinician involved with the project and each was prototyped and refined to the clinician's specifications.

2.2. Expert system design

The expert system module has two main purposes;

- to validate the results and
- to interpret the results.

The development of the expert system was an iterative process that took place in close collaboration with several clinicians experienced in the interpretation of umbilical cord acid–base data. A database of over 2000 cord samples had already been collected, which was used to formulate and verify the rules in conjunction with

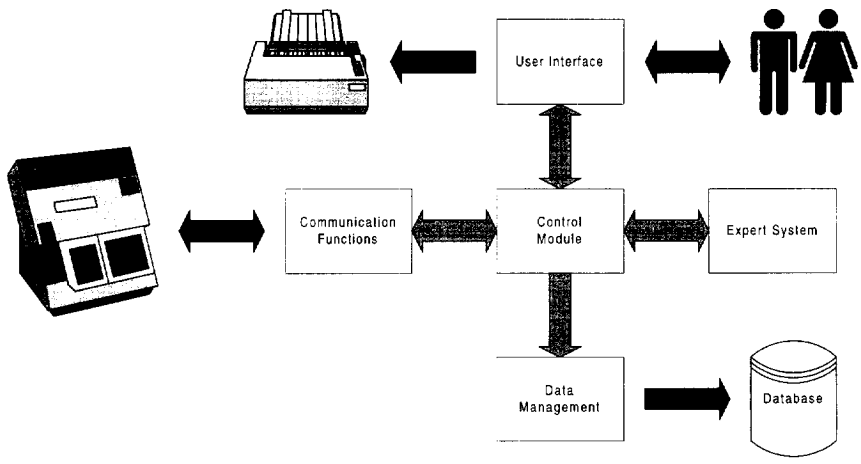

Fig. 1. Conceptual block diagram of expert system.

the experts' knowledge of fetal physiology. Frequency distributions of the pH, pCO_2 and pO_2 values were plotted to establish the median values and lower 2.5th centile ranges of each; means and standard deviations cannot be used on the data as all the distributions are skewed and not Normal. Frequency distributions of the pH, pCO_2 and pO_2 differences between vessels were also plotted which were used to establish the minimum allowable differences for arterial–venous paired samples. The populations were checked against other published data to ensure that they were not specific to our data.

Initially a set of rules was generated by two of the clinicians after a knowledge elicitation session. These rules were then encoded and applied to the database in a variety of ways. Firstly the full results were passed to the expert system and the interpretations recorded. Next each input parameter was marked as containing an error in all combinations and these results were also passed to the expert system. This generated the interpretations of the expert system with successively less information and enabled the internal consistency of the rules to be checked. A number of other techniques, such as passing plausible random numbers to the expert system, were used to examine the behaviour of the expert system rules. The output generated was examined by the clinicians and the rules modified to eliminate inconsistencies and refine interpretations. This process continued iteratively until the rules were deemed acceptable.

Internally, the knowledge representation is conceptually organised as a set of frames, implemented as 'C' structures, with attributes such as pH, pCO_2, BDecf, validation flags and originating vessel (artery or vein) for each sample. The expert system module was tailor written, again in the 'C' language, to simplify the exchange of information with other modules in the system. The expert system features a forward-chaining algorithm, which is suitable for the classification-type rules for both the validation and interpretation.

Each sample's results are passed to the expert system module for error checking. Two classes of errors are detected; analyser errors and physiological errors. Analyser errors are those reported by the analyser at the time of sampling such as caused by the electrodes failing to reach a stable reading, indicated as double asterisks next to parameters. Physiological errors are an additional class of error detected specifically by the expert system by examining whether the results are consistent with the range of possibilities for cord blood. For example, there is a strong relationship between the pH and the pCO_2, as shown in Fig. 2, where the 99.9% confidence of prediction intervals have been calculated by regression analysis. Analysis of the residuals has shown that the variances of the data across the pCO_2 axis are not uniform for all pH, indicating that the linear regression prediction intervals are not strictly valid. Hence, the exclusion limits were constructed beyond these intervals, widening with the increased variance in pCO_2 as pH decreases. Results that fall outside these exclusion limits-caused, for example, by the removal of pCO_2 from a sample through exposure to air-are reported as errors.

Paired results undergo a further, more sophisticated, stage of validation to identify which vessel is which and to ensure that they make 'physiological sense'

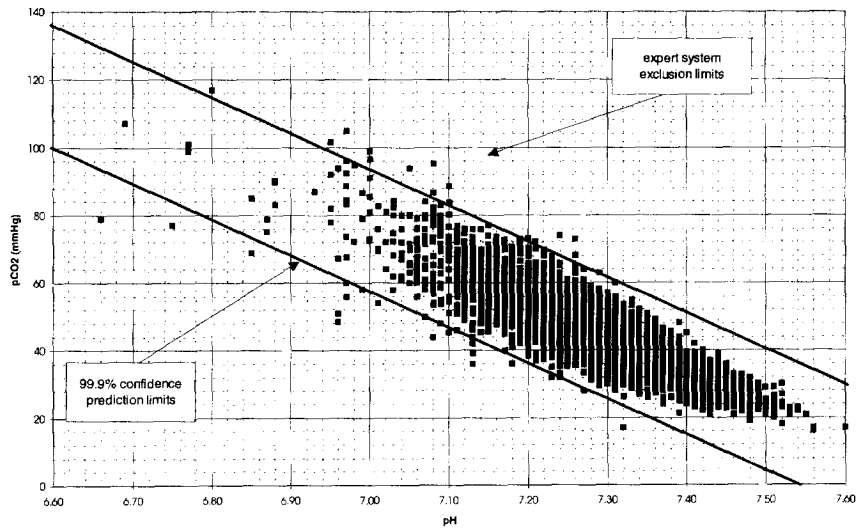

Fig. 2. Scatter diagram of cord blood pH against pCO_2 with 99.9% prediction limits and expert system exclusion limits.

when viewed as arterial and venous blood. For example, the pH of the artery must be lower than the vein whilst the pCO_2 of the artery should be higher; if this is found to not be the case, an error will be marked against the pCO_2 results. An example of such a rule is:

IF (venous − arterial pH) ≥ 0.06 AND (arterial − venous pCO_2) < 4 mmHg)

THEN mark the arterial and venous pCO_2 parameters as errors

If the pH and pCO_2 values for a sample are accepted as valid, the base deficit of the extracellular fluid (BDecf) is calculated by equations from Siggaard-Andersen [19]:

$$\text{BDecf} = -(1 - 0.023\text{Hb})(\text{HCO}_3^- - 24.1 + (2.30\text{Hb} + 7.7)(\text{pH} - 7.40))$$

$$\text{where } \text{HCO}_3^- = 0.23pCO_2 \log^{-1} \text{ and } \text{Hb} = 3.7 \text{ mmol/l}$$

The pH, pCO_2 and BDecf are examined to categorise the results into one of 54 interpretations, ranging from 'normal' to 'severe metabolic acidemia'. Results consistent with respiratory acidosis are distinguished from those indicating metabolic acidosis and the differences between the two vessels, if available, are used to further refine the diagnosis. An interpretation is performed on single samples as well as paired samples, although the information is very much more limited and the user is advised to retry with a paired sample.

2.3. System design

The system was designed to run when the PC is booted and to automatically detect the communications parameters of the blood gas analyser so that the system starts up online without user intervention. The user initially selects the type of sample to be placed through the analyser on the main menu. The program was primarily aimed at cord blood gas analysis, but in order to make the program more general, options were added to the menu system so that samples of any type could be entered. The main sample types identified were:

- paired cord samples
- single cord samples
- fetal blood samples
- neonatal blood samples
- adult blood samples
- other samples
- quality control samples

Every sample is written to a raw data 'master samples' database when it is initially received and to a specific database when processed. It was specified that logging of quality control samples and internal calibrations was essential, but that automatic detection of quality control results falling outside acceptable limits was beyond the scope of the project. The user simply inputs the quality control batch number and, when the results are displayed, selects 'passed' or 'failed' according to whether the results fall within the ranges for the batch. All calibrations are automatically detected and logged to database; the user is informed that a calibration is taking place and temporarily suspended from proceeding with data entry until the calibration sequence is completed. While the analyser is sampling or calibrating, a progress box is displayed indicating the amount of time expected for the sequence and the amount of time elapsed in the form of a moving bar which grows from 0 to 100%. The actual time taken for each sequence is recorded and is used to form a rolling average of elapsed time, which becomes the expected time for the next sequence.

Having selected the type of sample, the user inputs the mother's surname, first name, hospital identification number, and infant number for multiple births, and then is prompted to place the first sample into the analyser. The sample is measured by the analyser and the results transmitted to the computer. These are immediately passed to the expert system module for preliminary error checking. If an error is detected, the user is presented with immediate visual and optional audio feedback. The user is given the choice of retrying the sample if more blood is immediately available, ignoring the error if none is available or abandoning the current measurement if more blood is to be obtained from the cord. The user is prompted again if a second sample is required; once sampling is completed the results are again passed to the expert system for the second stage of validation and interpretation.

The results and interpretation are presented on a screen which identifies which results came from which vessel and displays a brief form of the expert system

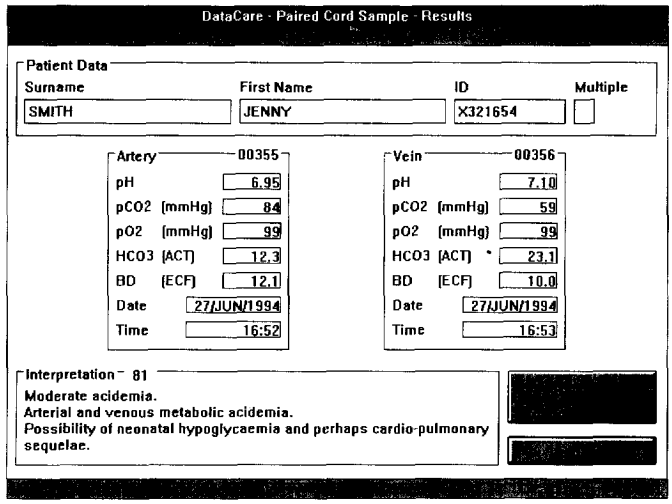

Fig. 3. An example screen showing paired cord results and expert system interpretation.

interpretation (Fig. 3). A more detailed explanation of the interpretation is available on request. The user now has a chance to correct any errors in the patient details and then can print the results to multiple sticky labels (Fig. 4), specifically designed for the mother's and baby's medical notes.

An optional personal audit facility was also included in the design of the system. If this feature is selected, each user must enter a personal identification number (PIN) of up to four characters prior to being allowed to use the system. This allows the individual usage of the system to be tracked, and could for example be used to determine which members of staff are responsible for the most errors. PIN numbers

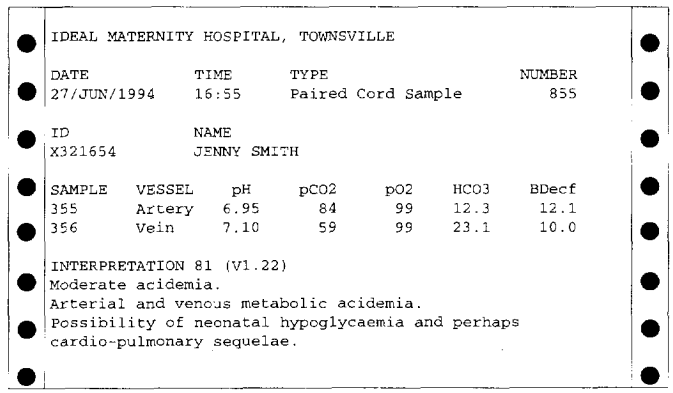

Fig. 4. An example label showing paired cord results.

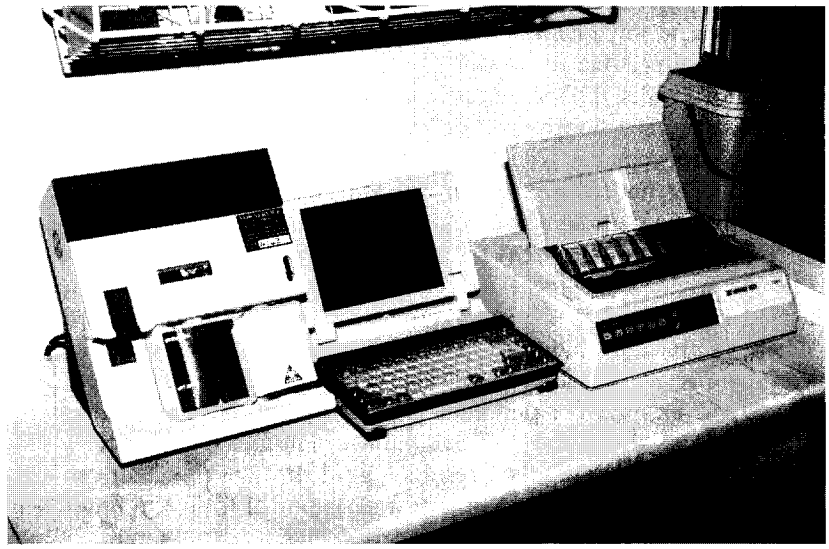

Fig. 5. The expert system implementation at Plymouth.

could be assigned individually, or be assigned to staff groups to determine their status.

3. Testing, validation and clinical trials

The system was put into operation at the local hospital in Plymouth at the beginning of July 1993, to start collecting data and gain feedback from the users (Fig. 5). Cord bloods had been analysed for every delivery in Plymouth since early 1992, so the introduction of the expert system caused no additional clinical work; the only minor change in working practice required was for the auxiliaries, who normally perform the sampling, to note the mother's name and hospital identification number before analysis.

The system was designed to be simple to use, and the users' quick acceptance of the system endorsed this. The multiple sticky labels for patient notes eliminated the time consuming and error prone process of transcription, and this helped significantly with user acceptability. During the trial, deliveries were carried out on two floors, with a blood gas analyser on each. The expert system was linked to only one of the analysers as the maternity department was soon to be moved to a new site at Derriford Hospital. The system proved popular enough that all cord samples were passed through the expert system. Minor changes and enhancements were made to the program in response to users' requests, problems or queries.

A specially designed table was constructed to secure the system with a lockable metal cradle for the PC and space for the blood gas analyser and printer. The whole

table was fixed to the floor to prevent removal. The operating system software was configured to prevent any changes being made to the environment by unauthorised users. A facility is available to download the databases to floppy disk for permanent archiving of results. This requires a software password and the key to unlock the bracket covering the floppy disk drive.

In parallel to the clinical trial, a process of internal software validation and documentation was undertaken to ensure that the software development cycle complied with BS5750 quality standards. This involved extensive 'destruction testing' of the software in which, as far as possible, every aspect of the software was tested. The few minor problems that this process highlighted were corrected.

At the end of September 1993 the system was put into a nearby hospital in Exeter, where routine cord blood analysis was to be initiated. This obviously meant a significant change to clinical practice, combined with the introduction of a novel computer system. Despite the potential pitfalls of such a large scale change, the system proved popular with the staff and gained high user acceptance, with a swift transfer to all deliveries having their results logged through the computer. The trial period, initially planned to last one month, was extended to over 6 months and the system has now been retained on a permanent basis.

4. Results

The introduction of an expert system has allowed, for the first time, the systematic identification of the various errors present in umbilical cord acid–base results, and has provided an accurate distribution of their frequencies. The availability of computerised records for each delivery has also allowed the overall usage of the system to be determined. Some of the issues of user acceptability and user attitudes have been examined through the use of the expert system individual audit facility.

Up to date information on the number of deliveries in Plymouth is currently available up to the end of 1995. In the period between 5th July 1993 and 31st December 1995 there were 11 746 babies delivered, including multiple births and excluding stillbirths. In the same period there were 11 492 patient samples placed through the expert system. Of these, 271 were subsequently repeated because of errors within the samples, resulting in 11 221 distinct samples from actual patients (95.5% of all deliveries). In 120 cases these samples were originally placed through the expert system as two or more 'single' or 'paired' samples, which had to be manually combined into an arterial-venous pair for an infant. As the manual combination process takes place after the data has been downloaded from the system, these paired results do not get an expert system interpretation. One Paired sample produced no results at all, as a result of analyser problems, leaving 11 100 classifiable samples. The validation of these samples is represented in Fig. 6. A terminology for the categorisation of umbilical cord acid–base errors is proposed in Table 1.

From Fig. 6 it can be seen that 9109 samples were categorised as full paired samples from the 11 746 deliveries. Thus 77.5% of all deliveries resulted in a validated arterial and venous paired cord sample. The full paired results receive the most comprehensive interpretation, ranging from category 80 (severe metabolic acidemia in both vessels) to 120 (all results normal), broadly ordered by severity in terms of likelihood of the infant having suffered asphyxial damage during labour. Categories 80 to 84 indicate that an arterial metabolic acidemia was present, of decreasing significance; 90–94 indicate that an arterial acidemia was present, but non-metabolic; and categories 100–113 indicate various states of mixed arterial and venous non-metabolic acidemia. The full breakdown of actual expert system categorisation is shown in Table 2. It can be seen that only 36 infants (0.4% of full paired, 0.3% of all deliveries) were classified in the most serious category 80-severe metabolic acidemia in both vessels.

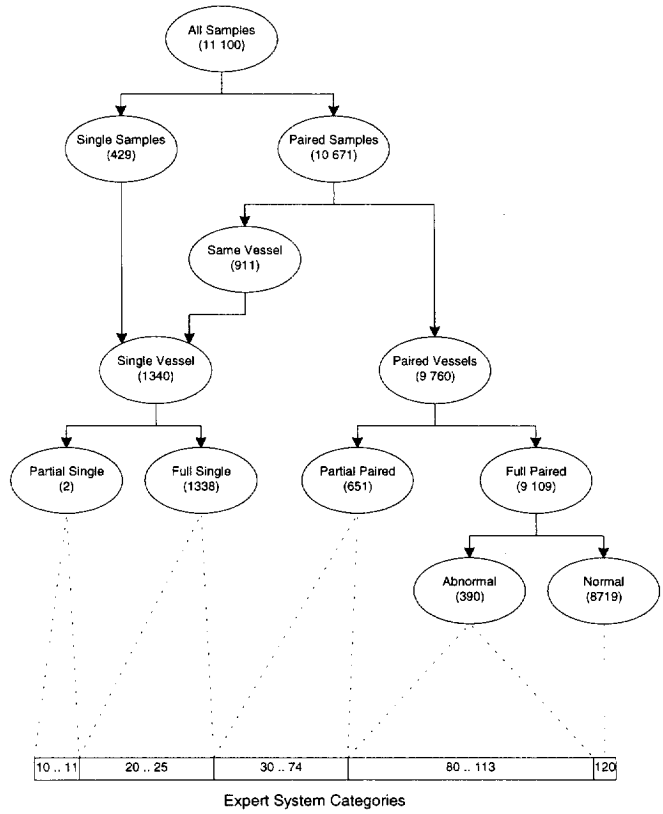

Fig. 6. A representation of the validation and categorisation process of the expert system, figures in brackets show the numbers in each group, and the mapping of the groups onto the expert system categories is shown at the bottom. See Table 2 for a guide to the terminology used.

Table 1
A proposed terminology for the categorisation of umbilical cord acid–base analysis

Terminology	Meaning
Single sample	A case for which only one sample of blood was taken, or more precisely only one sample was ever analysed
Paired sample	A case for which two samples of blood have been taken (presumably intended to be arterial and venous blood, but not necessarily so)
Same vessel	A paired sample in which very similar results from the two samples indicate that they are in fact from the same vessel, a mixed sample or extremely unlikely to have been from different vessels
Single vessel	A single or paired sample which is taken to be a single blood vessel, usually presumed to be the vein
Paired vessel	A paired sample that appears to be both arterial and venous blood
Partial single	A single vessel in which pCO_2 or BDecf errors are present
Full single	A single vessel with fully validated pH, pCO_2 and BDecf
Partial paired	A paired vessel in which one or more pCO_2 or BDecf errors are present
Full paired	A paired vessel with fully validated pH, pCO_2 or BDecf for both arterial and venous samples
Abnormal	A full paired result in which one or more of the parameters indicates some abnormality
Normal	A full paired result in which all parameters indicate a normal (physiological) condition

It had been thought that staff attitudes towards being monitored and audited would prevent the personal auditing feature of the expert system from being used, but when staff opinions were gathered, it was found that the auxiliary staff were keen to receive individual performance feedback. Consequently the audit facility of the expert system was instituted from 5th October 1994, and anonymous league tables of performance were posted at roughly six-monthly intervals. In the year prior to auditing, from 5th October 1993 to 4th October 1994 (inclusive), there were 3633 full paired samples collected from 4774 deliveries (76.1%); in the year after auditing, from 5th October 1994 to 4th October 1995 (inclusive), there were 3698 full paired samples collected from 4619 deliveries (80.1%). Although this improvement is encouraging, it is not statistically significant [χ^2: $P = 0.104$], and could just be an effect of the staff responsible for sampling simply improving over time. Nevertheless, the error rate has dropped steadily throughout the implementation and it is hoped that this progress will continue.

The system is now (October 1996) installed at around 18 hospitals around the UK, and as yet there have been no cases of software problems or maintenance required. In early 1994 the expert system was updated to connect to a new model blood gas analyser-again this modification was incorporated seamlessly and without problem. The lack of bugs and software maintenance requirements is a reflection of the time and effort put into the original software testing and validation process. In terms of project time, at least as much time was occupied by the validation process as for the rest of the project (from specification through knowledge elicitation to coding) combined.

5. Discussion and conclusions

This paper has described the development and implementation of an expert system for the validation and interpretation of umbilical cord blood acid–base. This project, in which engineers and clinicians have successfully collaborated, has resulted in a user-friendly expert system that has been implemented in several hospitals in the UK, and it is hoped that the system will soon become available overseas. The purpose of the system is: (i) to improve the quality of data collected

Table 2
The expert system categorisation of results

ES Category	Class	Description	Number
11	Partial single	Normal pH	2
20	Full single	Severe metabolic acidemia	7
21	Full single	Moderate acidemia	12
22	Full single	Non-significant acidemia	16
23	Full single	Questionable significance	11
24	Full single	Questionable significance	48
25	Full single	Single normal result, arterial acidemia not excluded	1244
30	Partial paired	Significant acidemia, but no base deficits	4
31	Partial paired	Questionable significance	5
32	Partial paired	Acute acidemia with large arterial-venous difference	1
33	Partial paired	Questionable significance	2
34	Partial paired	Results normal, but no base deficits	510
53	Partial paired	Results normal, but no arterial base deficit	6
64	Partial paired	Non-significant acidemia, but no venous base deficit	1
74	Partial paired	Results normal, but no venous base deficit	122
80	Full paired	Severe arterial and venous metabolic acidemia	36
81	Full paired	Moderate arterial and venous acidemia	22
83	Full paired	Acute moderate acidemia	6
84	Full paired	Acute moderate acidemia with large arterial-venous difference	10
90	Full paired	Non-acute moderate acidemia	14
91	Full paired	Moderate acidemia	2
92	Full paired	Significant acidemia	18
93	Full paired	Significant acidemia, but non-metabolic	21
94	Full paired	Non-significant acidemia	63
100	Full paired	Moderate non-acute arterial and venous acidemia	2
101	Full paired	Significant arterial acidemia	16
102	Full paired	Questionable significance	29
104	Full paired	Acute moderate acidemia, with normal venous results	8
105	Full paired	Questionable significance	43
110	Full paired	Significant acidemia, both vessels with venous metabolic	1
111	Full paired	Significant acidemia, with venous metabolic	7
112	Full paired	Non-acute mixed acidemia	40
113	Full paired	Questionable significance	52
120	Full paired	All results normal	8719

in cord blood analysis, in order to clarify the importance of the technique in the appraisal of neonatal outcome; and (ii) to provide an expert interpretation, in order to improve the understanding of cord blood gases during labour. The system also provides an automatic database of all results along with the calibration and quality control results of the machine. The expert system combines knowledge of the errors likely to occur in cord blood gas analysis, physiological knowledge of the reasonable results and the knowledge of the data collected during a large randomised trial in Plymouth to automatically check for errors and then interpret the results of cord blood gas analysis in a consistent and intelligent manner.

Up to now, far too little weight has been placed upon the importance of checking the quality of the data produced from cord blood analysis. In particular, the frequent occurrence of significant machine errors in conjunction with sampling errors leads to erroneous data, which are then interpreted (often poorly) without the full information of the pH's and base deficits (derived from the pH and pCO_2) of both vessels.

The field of obstetrics in general, and acid–base interpretation in particular, is characterised by the large amount of uncertainty in the domain. The input data suffers from the inherent imprecision in the analyser measurement, together with the further imprecision added by poor handling of the samples. During the knowledge elicitation process it became clear that the acid–base knowledge was uncertain, even among respected experts, causing certain combinations of variables to be perplexing and difficult to interpret. Many authors have previously identified the need for uncertainty handling in medical expert systems [9,10,7], and the development of the crisp expert system presented here has confirmed this requirement. Although this system functions at an acceptable level for clinical use, the lack of internal uncertainty handling in the knowledge and data has limited further improvement and future expansion of the system.

The system has been in live clinical use at Plymouth, Exeter and a number of other hospitals within the UK for several years. The system has been well accepted by the staff at Plymouth and Exeter. Indeed, the staff at Exeter found that the expert system considerably eased the introduction of routine cord blood gas analysis into a unit which previously had not performed any, by providing an easily accessible output (a natural language interpretation in clinical terminology) whereas otherwise they would receive a set of numerical data. The extensive software testing and validation process, in conjunction with the overall structured project methodology, has greatly contributed to the software reliability and user-acceptance.

6. Future work

The system described in this paper is the first stage of an ongoing development cycle. As more good quality, validated data is collected and the understanding of the interpretation of acid–base at birth improves, it is anticipated that the system will be regularly updated with new knowledge to become ever more sophisticated. Experimental work has already been completed to quantify the likely errors in each

of the basic input parameters. Current work has concentrated on incorporating fuzzy logic processing to improve the validation and interpretation. The system will then be expanded to include other neonatal data, such as gestational age and birth weight, and other outcome measures, such as Apgar scores and neonatal encephalopathy, as part of a comprehensive reassessment of immediate neonatal outcome.

The centralised collection of large quantities of multi-centre data consisting of cord blood gas results, obstetric information and neonatal progress from around the UK is a long term goal. Only in this way will the interpretation of immediate neonatal outcome and its relationship to long-term neurological development be clarified.

Acknowledgements

Professor Karl Rosén provided additional knowledge and guidance in the expert system development. The authors would like to thank Dr Sarah Beckley and Dr Mark Davies for assistance in the implementation at Plymouth and Exeter, and for providing education to clinical users of the system. We would also like to thank the consultants, midwives, doctors and particularly the auxiliaries of the obstetric unit at Derriford Hospital, Plymouth, for their co-operation throughout this project. This work was supported by EPSRC and Ciba-Corning Diagnostics (now Chiron Diagnostics), Halstead, Essex.

References

[1] K.P. Adlassnig and W. Horak, Development and retrospective evaluation of Hepaxpert-I: a routinely-used expert system for interpretive analysis of hepatitis A and B serologic findings, *Artif. Intell. Med.* 7, (1995) 1–24.

[2] American Academy of Pediatrics, Use and abuse of the Apgar score, *Pediatrics* 78, (1986) 1148–1149.

[3] S.W.K. Chan, K.S. Leung and W.S.F. Wong, An expert-system for the detection of cervical-cancer cells using knowledge-based image analyzer, *Artif. Intell. Med.* 8, (1996) 67–90.

[4] M.J. Dijxhoorn, G.H.A. Visser, H.J. Huisjes, V. Fidler and B.C.L. Touwen, The relationship between umbilical pH values and neonatal neurological morbidity in full term appropriate-for-dates infants, *Early Hum. Dev.* 11, (1985) 33–42.

[5] L.C. Gilstrap, K.J. Leveno, J. Burris, M.L. Williams and B.B. Little, Diagnosis of birth asphyxia on the basis of fetal pH, Apgar score, and newborn cerebral dysfunction, *Am. J. Obstet. Gynecol.* 161, (1989) 825–830.

[6] K.G. Goldaber, L.C. Gilstrap, K.J. Leveno, J.S. Dax and D.D. McIntyre, Pathologic fetal acidemia, *Obstet. Gynecol.* 78, (1991) 1103–1107.

[7] M. Halim, K.M. Ho and A. Liu, Fuzzy logic for medical expert systems, *Ann. Academy Med.* 19, (1990) 672–683.

[8] A. Halligan, M. Connolly, T. Clarke, R.P. Gleeson, M. Holohan, T. Matthews, King M. and M.R.N. Darling, Intrapartum asphyxia in term and post term infants, *Irish Med. J.* 85, (1992) 97–100.

[9] D.L. Hudson and M.E. Cohen, Fuzzy logic in a medical expert system, in: M.M. Gupta and T. Yamakawa, eds., *Fuzzy Computing* (Elsevier Science, North Holland, 1988) 273–284.

[10] C. Hughes, The representation of uncertainty in medical expert systems, *Med. Inform.* 14, (1989) 269–279.
[11] H.J. Huisjes and J.G. Aarnoudse Arterial or venous umbilical pH as a measure of neonatal morbidity, *Early Hum. Dev.* 3, (1979) 155–161.
[12] J.A. Low, R.S. Galbraith, D.W. Muir, H.L. Killen, E.A. Pater and E.J. Karchmar, Motor and cognitive deficits after intrapartum asphyxia in the mature fetus, *Am. J. Obstet. Gynecol.* 158, (1988) 356–361.
[13] R.A. Miller, H.E. Pople and J.D. Myers, Internist-1: an experimental computer-based diagnostic consultant for general internal medicine, *New Engl. J. Med.* 307, (1982) 468–476.
[14] Royal College of Obstetricians and Gynaecologists, Recommendations arising from the 26th RCOG Study Group, in: J.A.D. Spencer and R.H.T. Ward eds., *Intrapartum Fetal Surveillance* (RCOG Press, London 1993), 392.
[15] V.J. Ruth and K.O. Raivio, Perinatal brain damage: predictive value of metabolic acidosis and the Apgar score, *Br. Med. J.* 297, (1988) 24–27.
[16] D.M. Schreck, D. Zacharias and C.F.V. Grunau, Diagnosis of complex acid-base disorders: physician performance versus the microcomputer, *Ann. Emerg. Med.* 15, (1986) 164–170.
[17] E.H. Shortliffe, *MYCIN: A rule-based computer program for advising physicians regarding microbial therapy selection*, (PhD dissertation, Stanford University, Stanford 1974).
[18] O. Siggaard-Andersen, An acid-base chart for arterial blood with normal and pathophysiological reference areas, *Scand. J. Clin. Lab. Invest.* 27, (1971) 239–245.
[19] O. Siggaard-Andersen, *The acid–base status of the blood (4th edn.)* (Munksgaard, Copenhagen 1976) .
[20] M.E. Towell, The rationale for biochemical monitoring of the fetus, J. Perinat. Med. 16, (1988) 55–70.
[21] J.A. Westgate, J.M. Garibaldi and K.R. Greene, Umbilical cord blood gas analysis at delivery: a time for quality data, *Br. J. Obstet. Gynaecol.* 101, (1994) 1054–1063.
[22] J.A. Westgate, M. Harris, Curnow J.S.H. and K.R. Greene, Plymouth randomised trial of cardiotocogram only versus ST waveform plus cardiotocogram for intrapartum monitoring in 2400 cases, *Am. J. Obstet. Gynecol.* 169, (1993) 1151–1160.
[23] C.L. Winkler, J.C. Hauth, J.M. Tucker, Owen J. and C.G. Brumfield, Neonatal complications at term as related to the degree of umbilical artery acidemia, *Am. J. Obstet. Gynecol.* 164, (1991) 637–641.
[24] G. Zarkadakis, E.R. Carson, Cramp D.G. and L. Finkelstein, ANABEL: intelligent blood-gas analysis in the intensive care unit, *Int. J. Clin. Monit. Comput.* 6, (1989) 167–171.