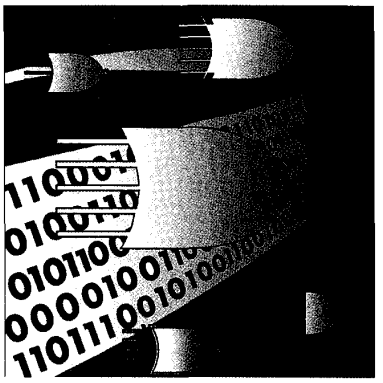


# Fuzzy Support for Serodiagnosis: The ONSET Program

Friedrich Steimann*
Michael Hayde**
Birgit Panzenböck**
Klaus-Peter Adlassnig*
and Arnold Pollak**
*Department of Medical Computer Sciences,
University of Vienna,
**Department of Neonatology,
University Childrens Hospital, Vienna, Austria

Infection with *Toxoplasma gondii*, a parasite widespread all over the world, is usually of little danger to the immunocompetent person. However, deliberately immunosuppressed organ transplant recipients or patients with acquired immune deficiency syndrome (AIDS) are exposed to its devastating effects. Furthermore, and the subject of this article, pathogens of a postconceptionally infected mother (acute gestational primoinfection) may cross the placental barrier and afflict the unborn [3, 7]. Clinical symptoms of fetal infection range from fetal death and stillbirth to clinically healthy newborns with an 80% chance of developing ocular toxoplasmosis and blindness in adulthood. General toxoplasmosis screening programs to detect acute infection of pregnant women have therefore been devised.

## Problem

While acute toxoplasma infection often remains asymptomatic and methods for direct detection of antigen are not routinely available, tests reactive to specific antibodies in human serum provide indirect serological evidence of infection [7]. However, because specific antibodies usually persist for the whole lifetime of the infected person, seropositivity alone is not indicative of acute infection. For a more specific diagnosis, i.e., to assess the time of onset of infection relative to the date of conception, knowledge of the development of antibodies over time must be employed. Unfortunately, as pointed out in [1], serological findings cannot unmistakably be linked with a time of onset, so that correct assessment is intrinsically difficult to achieve.

In this article, we present a computer-based method that encircles the time of onset of infection as closely as possible. It is based upon combination of evidence from serological findings.

## Methods

Both serological and computational methods are involved in our study. The former are described here only to an extent necessary to understand the latter.

### Performed Serological Tests

For the detection of toxoplasma-specific antibodies, a variety of tests are available. Tests differ in the type of antibody they respond to (e.g., IgG, IgM, IgA, or IgE), and in the quality of their response [7].

In our study, three tests were employed:

1. Sabin-Feldman dye test (DT), the reference IgG test suggested by the World Health Organization (WHO),

2. immunosorbent agglutination assay (ISAGA) for the detection of IgM antibodies, and

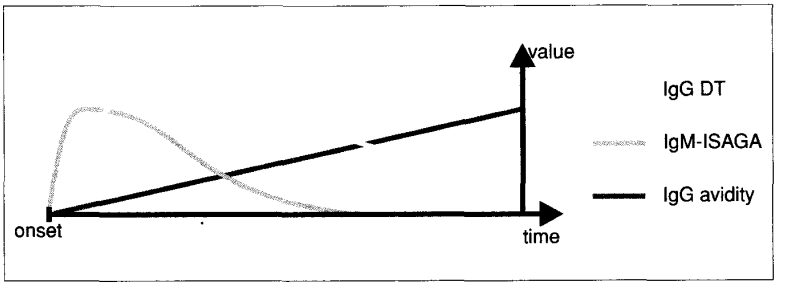
3. solid-phase enzyme immunoassay for the determination of IgG antibody avidity.

All test results were obtained and further processed in their quantitative form, i.e., IgG DT as a titer, IgM-ISAGA as an index, and IgG avidity as a percentage.

DT is highly sensitive and specific, i.e., positive titers prove toxoplasma infection while negative exclude it. Because this test requires living parasites, it can only be performed at few reference laboratories. Unfortunately, its capability of discriminating acute from chronic infections based on a single serum is limited, as individual immune response to toxoplasma antigens varies considerably in both strength and speed so that reliable assessment of acuteness can only be based on relative changes. The IgM-ISAGA test is commonly used to rule out or confirm suspected acute infection. In the initial phase of infection, this test yields high positive results that soon become negative with ongoing infection. However, because occasional high titer persistency lasting for several years is documented, its predictive value is also limited. Finally, IgG avidity test is a new technique for the measurement of the antigen-binding avidity (functional affinity) of IgG, distinguishing low-affinity antibodies at an early stage of infection from those with a higher binding

affinity reflecting preexisting immunity [6]. Because this test is not yet commonly accepted, it has only been included in this study for cases involved in a field trial also conducted at our laboratory.

Figure 1 depicts the ideal course of infection as observable through the mentioned tests.

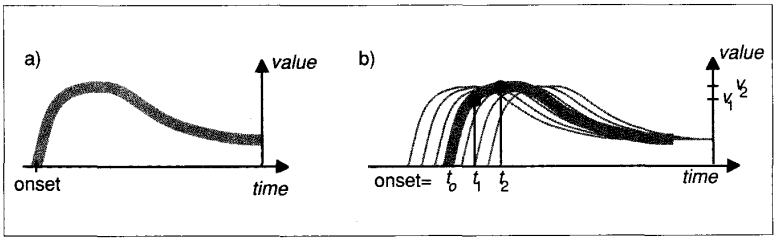



**1. Idealized antibody response to acute toxoplasma infection (adapted from [3] and [6])**

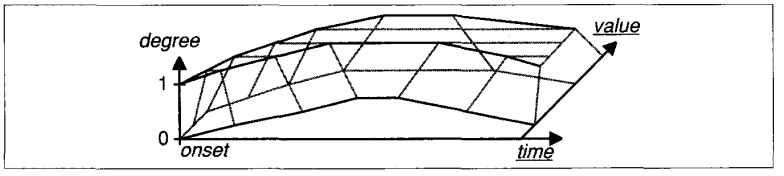
## From Serological Findings to the Onset of Infection

If an individual's exact course of antibody production in response to primary infection were known, times of onset could easily be determined by matching serological findings with the course: Fig. 2 illustrates the matching process for two findings, $v_1$ and $v_2$ observed at times $t_1$ and $t_2$, respectively.

Mathematically, the matching problem can be stated as follows. The course of immune response as observable through an antibody test is defined by a function $c$: $T \rightarrow V$ that maps time, $T$, (modelled by the real line) into a respective value space, $V$. For convenience, we assume that the course is left-aligned to time zero so that the onset corresponds to the origin. If the test results obtained from a patient are interpreted as points $(t_1, v_1),..., (t_n, v_n)$ within the time-value space, onsets of infection are determined by finding offsets $t_o$ such that all points lie on $c$ translated by $t_o$, i.e., $\forall$ $(t_i, v_i)$ , $1 \leq i \leq n : c(t_i - t_o) = v_i$ (see Fig. 2 for illustration).




**2. Deriving the time of onset by matching the course with findings, a) patient's course of infection as observable through a test (DT), b) matching by "sliding" the course over the findings**




**3. A fuzzy prototypical course assigning a degree to every time-value pair**

However, due to a wide physiological variability, there is no unique course of infection: some individuals respond with a rather vivid antibody production, while others exhibit comparatively moderate reactions. Because there are no analytical or statistical models available based upon which the course could be predicted, the real course of an individual's immune response is never known in advance. Nevertheless, if only sufficient information is available, an experienced clinician can restrict the range of possible onsets by applying general knowledge about typical and possible courses. The following provides means to explicate the clinician's knowledge and reasoning.

A *fuzzy prototypical course* is a formal representation of the expert's understanding of possible courses of infection as observable through a particular test. It provides the basis for deriving a *degree* or *grade of compatibility* of a time of onset with a series of findings. The degree reflects how close these findings are to the expected course starting at that time. The fuzzy prototypical course may be visualized as a band, the borders of which are blurred rather than sharp, so that transition from full compatibility with the course (*degree* = 1) to no compatibility (*degree* = 0) is gradual. Figure 3 shows a three-dimensional depiction of such a course.

Formally, a fuzzy prototypical course, $\tilde{C}$, is defined by a fuzzy relation, $\tilde{f}$, that assigns a degree of membership, $\mu_{\tilde{f}}$, to every pair $(t, v)$ of the time-value space. Again, the course is left-aligned so that $t = 0$ corresponds to the onset of infection. The degree of compatibility of a time of onset $t_o$, with an infection characterized by $n$ points $p_i = (t_i, v_i)$, $1 \leq i \leq n$ in $\tilde{C}$ is then defined as: $\bigwedge_{1 \leq i \leq n} \mu_{\tilde{f}}(t_i - t_o, v_i)$, where $\wedge$ denotes a suitable fuzzy conjunctive connective, a so-called *t-norm [4]*.

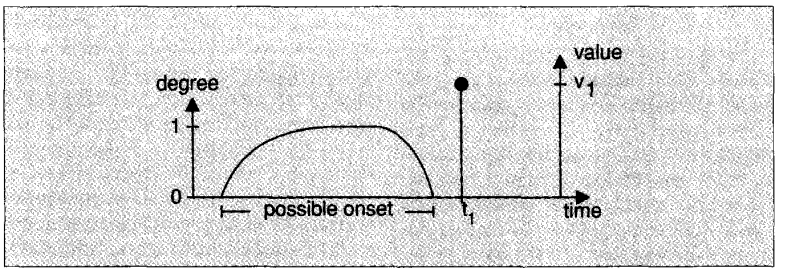
Deriving the times of onset from fuzzy prototypical courses and findings is analogous to the non-fuzzy case depicted in Fig. 2: the course is "slid" over the findings, and the respective degrees of compatibility are recorded along the time-scale, resulting in a distribution of degrees of compatibility over time, as shown in Fig. 4.

If more than one test is performed on a patient, the method is applied to all tests and the resulting distributions are combined to arrive at a single assessment. As rule of combination, we adopted the expert's reasoning, which combines evidence in a conjunctive fashion: if one test excludes infection at a certain time, then it overrules other tests—even if they suggest a possible infection at that time. Figure 5 depicts the result of combination using the minimum operator on a case involving two of the tests employed in our study.

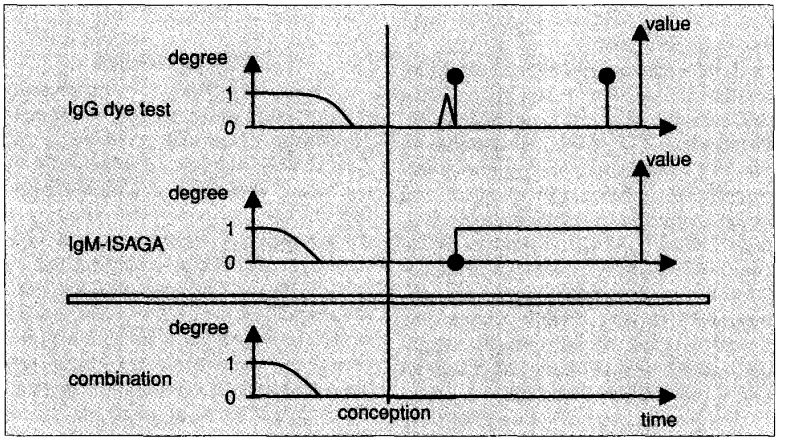
So far, definition of fuzzy prototypical courses merely reflects a clinician's intuitive understanding; next we will demonstrate how it can be based upon or derived from real cases.

## Obtaining Fuzzy Prototypical Courses

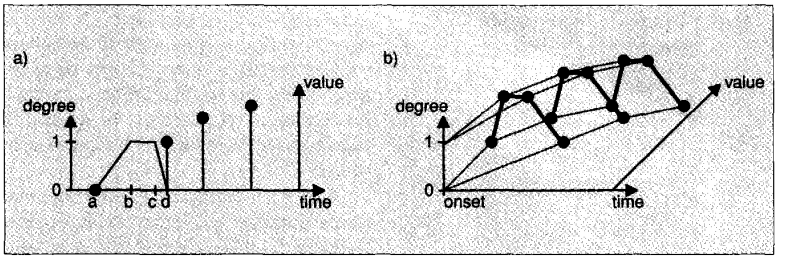
Even for an expert in the field of toxoplasmosis, it is a nontrivial task to




**4. Degree of compatibility of given findings with a fuzzy prototypical course relative to possible onsets**




**5. Combination of evidence as obtained from different tests (in this case excluding postconceptional infection)**




**6. "Learning" a fuzzy prototypical course from an acute infection a) possible onsets assessed for a seroconversion b) fragmentary course constructed from a)**

specify appropriately the required prototypical courses. In particular, a reasonable trade-off between sensitivity (reflected in a wide spread leading to unspecific results) and specificity (with narrow ranges excluding rare yet possible cases) is hard to find. We therefore propose a simple method to support the specification process, based on data obtained from patients where acute infection is evident.

When presented with an obvious case of acute infection, a clinician feels confident of correctly assessing possible times of infection onset. To account for gradation in the degree of compatibility, we asked the clinician to designate two nested temporal intervals $[a,d]$ and $[b,c]$, the outer one to denote those times starting at which a course could at all be called compatible (with degrees above zero), and the inner one specifying those times of onset starting at which the course would be fully compatible (with a degree of one). These intervals, interpreted as trapezoidal fuzzy sets as depicted in Fig. 6 a), specify the result expected of ONSET when presented with the given findings.

ONSET can now learn the fuzzy prototypical courses by reversing the assessment provided by the clinician: rather than the onset being regarded as the variable of the course, it is aligned to time zero and a fragmentary fuzzy prototypical course is constructed by translating the findings by the assessed times *a, b, c,* and *d,* as depicted in Fig. 6b).

Note that when applied to the whole series of findings or to any single one of its points, the matching procedure described above arrives at the same degrees of compatibility as initially specified by the clinician. Additional courses are also covered, as neither the relative temporal distance between any two nor the number of findings itself needs to be maintained to achieve a match.

A complete fuzzy prototypical course can now be obtained by constructing a fuzzy envelope, including all fragmentary courses derived as above from evidently acute cases. The more cases are included, the more general the obtained course will be. Note that constructing the envelope does not take relative frequency into account— due to relatively small numbers of documented acute infections, this does not seem worth consideration.

## Evaluation

To evaluate the performance of ONSET, a retrospective study was conducted involving 1000 patients randomly chosen from our internal database who had follow-up serology. From these patients, a total of 2373 sera had been drawn and just as many IgG Sabin-Feldman dye tests, 425 IgM-ISAGA, and 172 IgG avidity tests had been performed.

Congruence of the results produced by ONSET with the clinician's diagnosis is difficult to measure: while a clinician is used to clear statements such as "acute" or "latent infection", the outcome of ONSET covers a continuous spectrum of finely graduated diagnoses (cf. Fig. 5). Although this may be an advantage in clinical practice—a high resolution compatibility distribution is more differentiated than a binary *yes/no*-answer and can provide the basis for discussion and estimation of risk—it presents a problem to the evaluation process. We therefore mapped ONSET's output onto diagnostic classes representing clinically relevant diagnoses.

Depending on the compatibility of findings with onsets before and after the date of conception, five mutually exclusive cases can be identified, each of which establishes one diagnostic class:

1. If test results indicate infection and constrain the onset to times after concep-

tion, then acute infection is certain. This case is called *acute*.

2. If times of onset both before and after conception are compatible with the findings, acute infection cannot be excluded. To be on the safe side, this case is called *suspected acute*.

3. Compatibility of onset with times before conception only excludes acute infection. This case is called *latent*.

4. The status where all sera remain negative throughout pregnancy is described as *seronegative*.

5. If none of the above applies, i.e., at least one serum is positive but no onset is compatible with the findings, then the conclusion must be *inconsistent data*.

## Results

ONSET diagnosed the selected cases with a total accuracy of 91.4%. Table 1 summarizes the diagnostic performance of ONSET contrasted with that of a clinician.

Most notably, none of the acute cases was misclassified. In a binary classification scheme in which *acute* and *suspected acute* are comprised as one and the remainder as its complement, ONSET achieved a sensitivity of 97.5% and a specificity of 91.9%. Table 2 lists sensitivity, specificity, and accuracy as obtained for every diagnostic class alone.

## Discussion

Comparatively low accuracy for the classes *latent* and *suspected acute* goes back to 74 latent cases falsely classified by ONSET as *suspected acute*, though their initially low titer did not rise significantly in follow-up serology, a sign commonly agreed to exclude acute infection. This result was put into relation by the fact that in most of these cases, ONSET could derive only little evidence for acute infection, reflected in low degrees of compatibility for onsets after the date of conception, as shown in Fig. 7. However, due to the rigid assignment to diagnostic classes these subtleties were lost.

The fundamental problem of determining the time of onset of infection relative to the date of conception is illustratively discussed in [1]. There, all attempts to arrive at a sufficiently precise assessment of onset solely based on serological tests are questioned from a mathematical standpoint. Even if idealized models of the course of infection applied, simple decision rules would inevitably lead to erroneous results. ONSET gracefully accounts for this circumstance by delivering a compatibility distribution of onsets rather than an error-prone definite diagnosis.

The problem addressed by ONSET is a special case of a general trend detection task, which can be reduced to a multiple gradual class assignment problem based on fuzzy sets [8, 9]. In this particular application, only one trend—the fuzzy prototypical course—is specified per performed test. The goal of trend detection is not the classification of the observed course, but rather the determination of onsets, making the findings compatible with that trend. On-line detection of trends in a real-time environment, i.e., the multiple gradual assignment of an observed course to a collection of trends, where the time of onset of a trend is predetermined by the actual time and the duration of each trend, is the contrary problem.

**Table 1**
**Performance of ONSET vs. a clinician**

| Clinician → onset ↓ | acute | Suspected Acute | Latent | Sero-negative | Inconsistent Data | Σ |
|---|---|---|---|---|---|---|
| Acute | 5 | 0 | 0 | 0 | 0 | 5 |
| Suspected Acute | 0 | 37 | 77 | 0 | 1 | 115 |
| Latent | 0 | 1 | 259 | 0 | 2 | 262 |
| Seronegative | 0 | 0 | 0 | 606 | 0 | 606 |
| Inconsistent Data | 0 | 0 | 5 | 0 | 7 | 12 |
| Σ | 5 | 38 | 341 | 606 | 10 | 1000 |

**Table 2**
**Sensitivity, specificity and accuracy of ONSET**

| Diagnostic Class | Sensitivity (%) | Specificity (%) | Accuracy (%) |
|---|---|---|---|
| Acute | 100.0 | 100.0 | 100.0 |
| Suspected Acute | 97.4 | 91.9 | 92.1 |
| Latent | 76.0 | 99.5 | 91.5 |
| Seronegative | 100.0 | 100.0 | 100.0 |
| Inconsistent Data | 70.0 | 99.5 | 98.2 |

## Other Approaches

A comprehensive survey of trend-detection methodologies in biomedical monitoring systems is provided in [2]. This review addresses a wide range of mathematical methods and their applicability to trend analysis and forecasting of densely observed variables in a noisy environment. Although serological test results are inevitably subject to "noise" (due to variability in the quality of test material as well as imprecise evaluation methods such as cell counts), they are undoubtedly typical representatives of sparse sampling, which makes them unsuited for the reviewed methods.

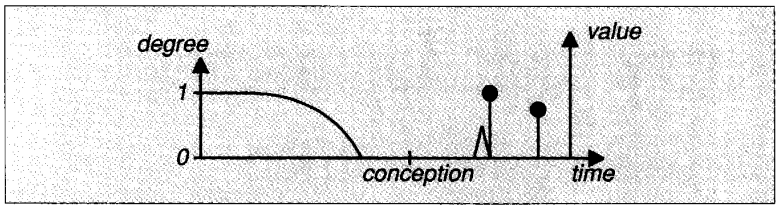
$TrenD_X$ [5] is a well-devised approach to trend detection from sparse samples, explicitly accounting for different phases of a trend. It decomposes each trend into a collection of temporal intervals, each of which constrains a number of parameter values. Temporal intervals may be of indeterminate length, their bounds are then related to other intervals or landmark points through additional temporal constraints.

Trends are detected by assigning time-stamped data to suitable intervals. For this purpose, $TrenD_X$ recursively generates and prunes hypotheses, instantiations of trend templates whose intervals are adapted to fit the data times and values.

Although $TrenD_X$ aims at the classification of courses rather than the determination of the onset of a trend, it can also be used to generate hypotheses, leading to different onsets. However, due to its interval-based explicit temporal logic, generation of all feasible hypotheses, particularly with a resolution comparable to ONSET's, is computationally expensive.

## Conclusion

ONSET derives its diagnosis from the combined evidence of contemporaneous and successive findings, which is in accordance with the behaviour of a clinician.




**7. Frequent false diagnoses of** ONSET

Although it does not increase the diagnostic capabilities in principle, i.e., it cannot perform better than an expert reasoning on the basis of the same knowledge, computational power can be exploited to perform the matches and combine the results with greater precision, resolution, reliability, and transparency. Fundamental improvements of diagnosis, however, can only be expected from better serological tests or increased testing frequency.

ONSETrelies on domain knowledge explicated in a non-verbalized, analytical form. Knowledge representation—although mathematical—is intuitively clear and easy to conceive, making it equally accessible to both man and machine, and hence subject to objective evaluation and criticism.

## References

1. **Ades AE:** Evaluating the sensitivity and predictive value of tests of recent infection: toxoplasmosis in pregnancy. *Epidemiol Infect* 107:527-535, 1991.

2. **Avent RK:** A Critical review of trend-detection methodologies for biomedical monitoring systems. *Critical Reviews in Biomedical Engineering* 17(6):621-659, 1990.

3. **Desmonts G, Couvreur J:** Toxoplasmosis. In: Conn RB (ed): *Current Diagnosis* 7, W.B. Saunders Company, Philadelphia 1985.

4. **Dubois D, Prade H:** *Fuzzy Sets and Systems: Theory and Applications,* Academic Press, New York 1980.

5. **Haimowitz IJ, Kohane IS:** Automated Trend Detection with Alternate Temporal Hypotheses. *Proc. of the 13th IJCAI* 146-151, 1993.

6. **Lappalainen M, Koskela P, Koskiniemi M, Ämmälä P, Hiilesmaa V, et al.:** Toxoplasmosis acquired during pregnancy: improved serodiagnosis based on avidity of IgG, *J Infect Dis* 167:691-697, 1993.

7. **McCabe RE, Remington JS:** The diagnosis and treatment of toxoplasmosis. *Eur J Clin Microbiol* 2(2):95-104, 1983.

8. **Pedrycz W:** Fuzzy sets in pattern recognition: methodology and methods. *Pattern Recognition* 23:121-146, 1990.

9. **Steimann F, Adlassnig KP:** *Diagnostic monitoring based on fuzzy state transitions.* Technical Report 1-94, Department of Medical Computer Sciences, University of Vienna, Austria 1994.

*Freidrich Steimann* was born in 1965 in Heidelberg, Germany. After graduating in Informatics from the University of Karlsruhe, Germany, in 1991, he joined the Alcatel Austria Elin Research Center, Vienna, as a software engineer to work on object-oriented analysis and design methods. In 1992, he changed to the Department of Medical Computer Sciences of the University of Vienna. Since then, he has been working on fuzzy medical patient records and trend detection methods in both intensive care unit monitoring and screening applications. He can be reached at the Department of Medical Computer Sciences, University of Vienna, Waehringer Guertel 18-20 A-1090 Vienna, Austria e-mail: steimann@vm.akh-wien.ac.at

*Michael Hayde* was born in 1954 in Vienna, Austria. He received his MD from the University of Vienna Medical School in 1981 and joined the University Childrens Hospital, where he is still working as a neonatologist and general pediatrician. His scientific interests span research in diet therapy of patients with inborn errors of metabolism, surfactant therapy in premature infants, as well as detection and treatment of prenatal infections and fetal disorders. As a collaborator of Austria's Toxoplasmosis Reference Laboratory, he is evaluating new serological test for potential use within screening programs.

*Birgit Panzenböch* was born in 1967 in Neunkirchen, Austria. She received her MTA degree in 1988. Since then, she has been working in Austria's Toxoplasmosis Reference Laboratory at the University Children's Hospital in Vienna. Together with Dr. Hayde, she is evaluating new toxoplasma testing methods.

*Klaus-Peter Adlassnig* was born in Germany in 1950. He received the M.Sc. and Ph.D. degrees in computer science in 1974, and 1983, respectively. He joined the Department of Medical Computer Sciences as of the University of Vienna Medical School in 1976, where he became professor of medical informatics in 1992. Since 1988, he has been head of the Medical Expert and Knowledge-Based Systems Unit. He was a research associate at the University of California, Berkeley, from 1984-86 and a visiting professor at Stanford University in 1993. He has published more than 100 articles in the areas of medical informatics, medical expert systems, and fuzzy set theory.

*Arnold Pollak* was born in 1945 in Cochabamba, Bolivia. He received his MD from the University of Vienna Medical School in 1971, where he became professor of pediatrics in 1987. He has been head of Neonatology and NICU since 1992. He published more than 300 papers in international journals, gave more than 100 presentations at national and international conferences, was a visiting professor in the United States three times, and received six medical awards.