The CTS Proficiency Testing Program

Rise of the blood gas laboratory. Measurements of the respiratory gases in the blood date back to the nineteenth century but the first to be widely used was described by Van Slyke and Neill in 1924. It was technically demanding and time consuming and for several decades was used only for research. During the 1950s, with the success of surgery for congenital heart disease, the Van Slyke technique became clinically useful to identify intracardiac shunting. Every cardiac catheterization laboratory had 'Van Slyke technicians' to run an accurate analysis to confirm the crude oximeters and 'eyeball techniques' that provided a rapid identification of the abnormal step-up of the oxygen saturation in the right heart and pulmonary circulation. This was during the period when a new generation of pulmonologists was emerging, trained not in the tuberculosis sanatorium but in the physiology laboratories of institutions like the Bellevue Hospital and the University of Pennsylvania. Lung function testing was 'hot' and it was in the pulmonary laboratory rather than in clinical pathology where the growing demand for blood gas analysis was satisfied.

During the same period, electrochemical methods to measure blood gases improved with the introduction of the Clark and the Stow-Severinghaus electrodes. Results were obtained within minutes rather than hours on much smaller volumes of blood. Physicians abandoned the cumbersome indwelling Cournand and Riley arterial needles and performed the easy 'one stick method' drawing a couple of millilitres of arterial blood into a glass syringe, the dead space wetted with heparin solution.

The 1960s saw the rise of critical care units, evolving from post-anaesthesia recovery rooms, to coronary care units and then to the pulmonary critical care unit with its intubated patients on mechanical ventilation. In the years before pulse oximetry was available, these patients required frequent blood gas sampling. 'ABGs' were no longer done a few times daily, mostly on outpatients. They became the most frequently ordered test, and incidentally the principal source of revenue, for the pulmonary laboratory. Thus, pulmonologists, trained in internal medicine at the bedside, found themselves directing laboratories that performed clinical chemistry, previously the domain of the pathologist.

The need for quality control. In 1947, Belk and Sunderman reported interlaboratory comparisons of several "common chemical measurements" on unknown samples sent to 59 hospitals and private laboratories. The authors mildly commented that "the accuracy of the measurements is below any reasonable standard". It was in response to findings like this that the College of American Pathologists (CAP) initiated a program of 'multi-laboratory surveys', but resisted the term, 'proficiency testing', considering them a tool for education and quality improvement rather than a regulatory device to distinguish 'good' from 'bad' laboratories.

Nevertheless, in 1967 Congress passed the first Clinical Laboratories Improvement Act (CLIA '67), which included a requirement for proficiency testing. The law applied only to laboratories processing specimens across state borders and it was estimated that only
12,000 of the 200,000 laboratories in business were regulated. For the rest, participation in a proficiency testing program was voluntary, but it required no clairvoyance to see that compulsory participation would arrive before long.

**Blood gas proficiency testing.** Providing specimens for interlaboratory comparison of blood gas analysis presented formidable technical problems because the material needed to be protected from exposure to air. The first blood gas proficiency testing involved 15 hospitals in metropolitan Seattle. Tonometered whole blood was delivered "by taxi, courier, shuttle bus, or ferry" to laboratories up to 25 miles from the reference laboratory 4. Interlaboratory comparisons of blood gas analysers became much more feasible with the introduction of sealed glass ampoules equilibrated with known partial pressures of oxygen and carbon dioxide.

The CTS proficiency testing program for blood gases arose from activities of a CTS committee charged in the mid '70s with writing standards for pulmonary function testing in California. Because of the challenges of transferring the complex techniques for measuring $O_2$ and $CO_2$ to the clinical environment and also because of anecdotal reports that in some clinical settings these measurements were so inaccurate as to be useless, committee members agreed that it would be very useful to offer a program by which the accuracy of these measurements could be objectively assessed. However, because of the complexities of preparing adequate testing samples including the labile relationship between hemoglobin and oxygen, it was concluded that the majority of labs would be unlikely to participate voluntarily because of the cost of such a service.

A key breakthrough was the announcement by Federal officials overseeing clinical laboratories (then HCFA, now CMS) that proficiency testing for blood gas measurements would be required in the future (as was then required for many tests performed in clinical chemistry laboratories). This mandate, which would threaten the continued licensure of labs, including small specialty labs adjacent to critical care units or operating rooms, would assure the participation of blood gas laboratories in new PT programs.

The PFT committee members also recognized that the College of American Pathologists would soon add blood gas measurements to the array of proficiency testing for many tests performed in clinical chemistry labs, and that respiratory specialists would need to move expeditiously in setting up their own proficiency testing programs if they wanted significant control over the quality of such testing, and of the research and education which could evolve from such programs.

The looming Federal mandate assured us of gaining a sufficient number of laboratories to pay for a testing service but did not resolve the challenge of setting up and operating such a testing service. The second key breakthrough was the recognition by some members of the PFT committee that a company already providing samples for the quality control of blood gas measurements would be an ideal partner for a professional society creating such a program. They also believed that the competition would probably be keen for such a partnership and that the QC company that was selected would assume most, if not all, of the costs. A couple of committee members even
opined that the difference between the annual fees paid by each laboratory and the cost per lab for the testing would cover all costs incurred by the professional society for developing and overseeing the program. Perhaps there would even be funds remaining to cover research and education related to blood gas measurements.

**CTS goes it alone.** In 1979, we approached ATS, inviting them to participate in a joint PT program which would serve the interests of patients with lung disease and of the ATS members responsible for providing blood gas measurements. Disappointedly, the ATS declined to do so. Recognizing that there were sufficient numbers of CTS members directing specialty blood gas labs to establish a viable PT program, the CTS moved ahead to set up a program for its membership in California. After a competitive request for proposals, Instrumentation Laboratories (IL) was selected to provide the testing products and statistical services working under the direction of a new committee charged with overseeing CTS blood gas program. We recognized the importance of including in the committee technical staff with hands-on responsibilities for running the blood gas analyzers.

The program was an immediate success. As expected, all of the CTS’s costs for the program were covered from the program’s income, including all costs of the committee meetings and CTS staff support. Soon income from the program even exceeded income from CTS membership fees and continued to do so through the more than three decades of its operation.

CTS and CAP both initiated proficiency testing for blood gases in 1979. Stable material was now available for interlaboratory testing in the form of sealed glass ampoules containing an aqueous buffer solution equilibrated with known partial pressures of oxygen and carbon dioxide. Both programs switched to a fluorocarbon-containing emulsion, CAP after the first year and CTS after the second. The advantage of the fluorocarbon was that it had a much higher oxygen capacity than the buffer solution, with less sensitivity to room air contamination during handling and analysis. The result was a reduction in the variability of the reported PO$_2$ when the same lots were run, both in the 10 reference laboratories, and in the participating laboratories.

We formed a subcommittee from the technical staff of the reference laboratories and they were especially helpful in providing direct support to labs enrolled in the program. The medical directors and technical staff also collaborated in research leading to several publications in peer-reviewed journals 5-13.

**A joint program with ATS.** In part because of pressure from ATS members in other states and in part because of the success of the CTS program, the ATS asked that the CTS join with the ATS is establishing a new nationwide blood gas PT program. To protect the program in the event that ATS was not committed to long-term support of a combined program, the CTS instead invited the ATS to set up a parallel program for states other than California, with their administrative committees meeting jointly. The new ATS-CTS program began in 1984. In 1988 we added proficiency testing for CO-oximetry.
CLIA '88. The original CLIA '67 recommended proficiency testing but for the great majority of laboratories participation was voluntary. Over the next two decades there were further attempts to regulate clinical laboratory testing culminating in a series of articles published in the Wall Street Journal in 1987 decrying the poor performance of many standard tests, especially cervical cancer screening performed in "Pap mills". Congress responded by passing CLIA '88, "short on specifics (this was left for HCFA)" but making proficiency testing mandatory 14. Several of the rules proposed by HCFA posed special problems for blood gas laboratories participating in the ATS-CTS program. "In most cases, the laboratory director would need to be a physician certified in clinical pathology by the American Board of Pathology … Technicians would need to have completed a structured college-level curriculum in medical laboratory techniques or to have a high school diploma plus one year of formal training or to have completed a two-year technician traineeship" 14. It was clear that a blood gas laboratory in which the collection and analysis of specimens was done by respiratory therapists under the direction of an internist-pulmonologist would not satisfy the proposed standards. In the face of strong objections from many segments of the health care industry, but especially from physicians' office laboratories, which were threatened with closure or an intolerable augmentation of their costs, the regulations were greatly modified during the discussion period. The final rules, similar to those we know and love today, were not published until 1992 and came into effect over the next couple of years.

Though the ATS-CTS PT program was fairly similar to what was required under CLIA '88, the regulations were very detailed and very specific. For the better part of two years the meetings of the administrative committee were almost entirely taken up with discussions of how to conform. Members of the committee played an important part in changing the proposed rules so that the technical staff of blood gas laboratories could continue to perform the analyses and qualified pulmonologists could continue to act as medical directors.

Software. IL had provided the grading and reporting software from the beginning of the program. It was run on a mainframe computer and the programmers at IL had difficulty responding to requests for changes from the committee members. The late Bob Crapo was chairman of the ATS component of the committee and his group at LDS Hospital, initially inspired by Steve Berlin, a self-taught amateur, and subsequently managed by their statistician-programmer, Bob Jensen, developed a PC program to do the job. This put us in a strong position to respond to the programming changes needed to satisfy the CLIA regulations. The LDS group did not want the responsibility of processing the data and mailing the reports and so IL agreed run it on their PCs in Massachusetts.

Relationship with CAP. In order to maintain their accreditation by JCAHO, hospitals required inspection of their laboratories during their hospital surveys. The Joint Commission accepted the regular inspections done by the CAP program but did their own inspection of specialty laboratories not enrolled with CAP. Beginning in 1995, laboratories surveyed using JCAHO standards were "deemed to be certifiable" under CLIA '88 rules. If a laboratory was not inspected by CAP, JCAHO would provide the inspection but the cost was no longer included in the basic cost of the hospital survey. CAP declined to inspect a laboratory that was not enrolled in their PT program. This put
directors of blood gas laboratories enrolled in the ATS-CTS program in a bind. They could enroll in both CAP and ATS-CTS PT programs and then would be eligible for inspection by CAP, or they could have their lab separately inspected by JCAHO. Either solution involved added costs and many came under pressure from their cost-conscious administrators to drop our PT program and to switch to CAP. While many directors, out of loyalty to our program or convinced that the cost of enrolling in two programs was justified by its effect on quality, remained with ours, there was a slow but steady decline in the number of enrolled labs, averaging around 7 percent per year through the late 90s.

Since our program was approved by HCFA to perform proficiency testing under the CLIA regulations, we approached CAP to see if they would agree to inspect a blood gas laboratory if it participated in an approved PT program. Our first effort was to ask CAP to recognize alternative PT programs. They declined, and there was no action by HCFA to prohibit what appeared to us to be an illegal restraint of trade by CAP. Subsequently, a lawsuit decision forced CAP to reverse its position, and they began offering applications for alternative PT programs to enter result data in the CAP Laboratory Accreditation Program (LAP) for CAP-regulated labs. A subsequent contact was made informally by an unlikely spokesman, our programmer Bob Jensen. He learned that while CAP would never consider accepting our proficiency program in lieu of theirs as qualifying a laboratory for review by a CAP inspector, they were aware of our program and recognized that our group of pulmonary physiologists and clinicians provided expertise that might be lacking in a program run purely by pathologists. At a further meeting with the ATS leadership, CAP proposed a joint program which would retain the ATS name, and explained that they had undertaken similar joint programs with other non-pathologist groups. A contract was proposed that included protection of revenue CTS had received from the program, at least for the first two years after the merger.

Negotiations proceeded through 1995 but when we believed we were within a few weeks of signing the final document, we learned that CAP decided "not to pursue the possibility of co-sponsorship". Their explanation for the decision was that CAP had misunderstood the numbers provided by ATS-CTS for the number of enrollees in our programs, believing there were 1500 laboratories rather than 1500 instruments.

**More analytes and changing technology.** Through the 1990s the technology of blood gas analyzers was rapidly changing. Laboratories close to the critical care unit provided more rapid reporting of results and instrument manufacturers responded to the demand for multiple tests with single blood specimens run on a single analyzer. New blood gas analyzers appeared that were capable of measuring electrolytes and other analytes such as glucose and lactate, which had never been performed in pulmonary laboratories. A second type of innovation was point-of-care instruments that could provide rapid results in locations remote from the blood gas laboratory, such as operating theatres and emergency rooms. These instruments needed proficiency testing and the test materials were aqueous solutions. The fluorocarbon-based 'abc' that we used was preferable to aqueous solutions for traditional blood gas analyzers but was not supplied in a form that contained electrolytes and the other analytes of interest. An
additional problem was that fluorocarbon test material could damage the sensors used for some of the point-of-care instruments.

**The 'small n problem'**. When an unknown specimen is analyzed for an interlaboratory proficiency testing comparison one needs a target value and a range of values to decide whether the reported result is 'acceptable'. In order to establish the target value and the acceptable range, a 'peer group' is needed, composed of proficient laboratories running samples from the same lot. One could then take, for example, the mean ± 3 standard deviations from the peer group and use this as the acceptable range of values. The peer group might be laboratories using the same instrument model as the instrument being challenged.

In the early days of our program, there were only a few instrument models and we used an 'instrument-specific' target range if there were at least 20 instruments of the same model in the program. Otherwise the target range was derived from the 'all-instrument' mean and standard deviation. When the CLIA '88 regulations came into effect, the number required for a peer group was set at 10 and, for most analytes, fixed limits were established for the target range. For example, for pH the target range was the peer group mean ± 0.03 units. However, for many analytes the between-instrument differences were too variable to allow fixed limits for the target range. For PO2 the target remained the peer group mean ± 3 times the peer group standard deviation.

If there were fewer than 10 instruments of a particular model in the program, then the acceptability of its result would be determined by the all-instrument statistics. So long as the mean value for the instruments of that model was not much different from the all-instrument mean, that would not cause that instrument to 'fail'. In fact, it might improve its chances of an acceptable result since the all-instrument standard deviation was usually greater than the instrument-specific value. In the early days of the program, when all blood gas instruments used similar technology, the 'small n problem' was not really a problem. However, with many new models appearing, some of them using different methodology, our relatively small program had to deal with a new headache. A new model might be represented by, say, half a dozen instruments and most of them would report unacceptable results. However, when we inspected the individual results we would find them in a relatively narrow range, the instrument-specific mean differing significantly from the all-instrument mean against which they were judged.

The committee agonized about how best to deal with the problem. CLIA regulations limited our ability to do something like using the instrument-specific mean even if n was less than 10. We could try to combine the results with those of a model using similar technology but for some models, particularly the point-of-care instruments, that was not possible. Sometimes we had to resort to the unpalatable course of advising the laboratory director that he or she should enroll in a different PT program.

**Switch to aqueous test material**. After a couple of years of discussion, in 1998 we switched to aqueous test material. This allowed us to add to the program PT for electrolytes and certain point-of-care instruments. Most of the instruments were still traditional blood gas analyzers and for them the discontinuation of the fluorocarbon-
based material was a step in the wrong direction. However, to offer a choice of two test materials, in addition to greatly increasing our costs, would exacerbate the small n problem, possibly dividing an adequate peer group into two groups where one or both might be less than the necessary. Though in the short term the fluorocarbon was clearly preferable, in order to future-proof the program as more laboratories switched to the newer technology, we felt we had no choice but to make the change.

ATS drops out. ATS had never shared the enthusiasm that CTS had for the PT program. As a source of revenue it was insignificant compared with the annual meeting and their publications, and we understood that some of their leadership felt that it was not aligned with their primary goals of education and research. ATS had welcomed the negotiations with CAP as offering a face-saving way to end the program and were not worried about any associated loss of income. CTS, on the contrary, were in a way, victims of our own success. The PT program brought in well over half of the total revenue of the society. Few medical specialty groups had separate dues-collecting state organizations and, with membership of CTS slowly declining, it was hard to see how we would be able to maintain our other activities if the PT program were to shut down.

By 1996 the CTS leadership began exploring the possibility of taking over the whole program from ATS. The executive committee was enthusiastic and authorized the PT committee to begin negotiations. No decision was made over the next two years but we learned at the October 1998 meeting of the joint ATS-CTS that "ATS might discontinue its program". In May 1999, faced with an approaching deadline for our HCFA-CDC renewal application, the CTS president sent an urgent letter to the ATS executive director, requesting a prompt decision on whether ATS would continue its program. That finally produced a decision and, according to a memorandum of understanding, originally signed in 1990, CTS had the first right to recruit ATS subscribers who would need to change to another program. A letter was sent to all ATS laboratories under the joint signatures of the chairmen of the ATS and CTS PT committees, inviting them to enroll in the CTS program. The response was gratifying and many ATS laboratories were added to the CTS program, which now became a national operation although the majority of our laboratories, 165 out of a total 280, were still in California.

IL drops out. In the summer of 1999, we faced an even bigger challenge than trying to integrate the ATS program with our own. Our grading software was developed by the LDS Hospital group but was run by the IT staff at IL. With ATS out of the program, Bob Jensen was no longer available to provide program maintenance and, in any case, the software was becoming obsolete. IL was engaged in reprogramming their own QC software and we hoped that they would include updating the PT software as part of the project. The months went by without a definite commitment from IL and our anxiety rose as the deadline for submitting our renewal application approached. It was not until December that we learned that IL, facing a foreign acquisition, had aborted their QC reprogramming project. They recommended that we find an outside contractor to write a new PT program for us. IL agreed to continue their existing service, supplying test material, grading and reporting, but only to the end of 2000. Although it was not until February 2000 that we received final notice that IL would not renew our contract in
2001, the CTS executive committee recognized by the autumn of 1999 that we would need to find new vendors.

The scramble in 2000 to keep the program alive. There were multiple components needed to continue the program. The first, purchasing the vials of test material, was relatively easy because there were several companies that provided suitable QC material. The second was mailing the challenge samples to the participating laboratories. Next was receiving the results from the laboratories, entering the data into a computer and running the grading software. Finally, there was mailing the reports to the laboratories and sending them to the appropriate regulatory bodies. The last was complicated because it required sending the results, not only to the national organizations, but also to several state departments, each with its unique format and delivery schedule.

We first approached Bionostics, a well-regarded company that supplied test material for the CAP program. It seemed at first that they might be able to supply test material, mailing and grading software but, by January, it was clear that they would not be able to handle the software component and that their charges for anything in addition to sending us the test material would strain our budget. We had discussions with a couple of independent programmers but they lost interest after learning how complex a project it would be. We discussed having the CTS office in Tustin do the mailing of samples and processing the reports. That suggestion was vetoed when we realized that it would require contracting or hiring several more permanent staff, since there would be a learning curve that ruled out using temporary help. The need for refrigeration of the CO-oximetry QC material was an additional headache, requiring suitable storage facilities, and imposing a less forgiving delivery schedule.

AccuTest to the rescue. The weeks ticked by and our options seemed few, expensive and unattractive. In February 2000, to our vast relief, Bill Donohue of AccuTest told us he would be prepared to handle the whole package. AccuTest was a small PT provider based in Massachusetts offering a fairly broad menu of PT tests. The program was approved by HCFA and had been submitted to CAP. Combining his database with ours promised a partial solution to the small n problem, which was increasingly troublesome for both programs as the number and heterogeneity of instrument models grew. His company was very responsive to our desire to keep the scoring algorithms and the report formats as similar as possible to the familiar reports that our laboratories had received for years. Through the latter half of 2000 we worked closely with Mr Donohue, hoping for a smooth transition to our new service after two decades of good experience with IL.

The first testing cycle of 2001 was full of glitches. There were delayed delivery of ampoules, wrong addresses, incorrect identification of instrument models and late reports. Many of the problems arose because the CTS database of participant laboratories was incorrectly entered into the AccuTest computers. In May the PT committee met with Bill Donohue with many complaints. Bill assured us that the problems would be fixed with the next cycle and, indeed, things ran more smoothly for the rest of the year, though errors continued. However, the committee felt that we could
not risk loss of HCFA and CAP accreditation. Even if that disaster was avoided, we foresaw that many of our laboratories, unhappy with their experience, might defect to other PT programs. We decided to seek a new provider.

**Wisconsin State Laboratory of Hygiene (WSLH).** WSLH had long experience running a blood gas PT program, dating back to 1980\(^\text{15}\). We worked out a satisfactory contract over the summer of 2001. Their program had a similar number of blood gas instruments as ours and so they welcomed the combination of our databases, which would provide larger peer groups for instrument models affected by the small n problem. The first cycle of 2002, the first with our new provider, was almost free of problems, a huge relief to the CTS staff and to the PT committee.

**The last 15 years.** From then on the program ran remarkably smoothly but it was clear to the committee that the program could not go on for many more years. The older generation of physiologist-pulmonologists was reaching retirement age and their critical care oriented successors did not have the same interest in running a blood gas laboratory. Some laboratories were absorbed by pathologist-directed central labs while others found it administratively convenient to switch all of their PT testing to other services, most commonly CAP. CTS remained highly dependent on revenue from the PT program but we were caught in a vice between slowly rising fixed costs and slowly declining enrollment. In fact, we were able to continue much longer than any of the leadership of CTS would have predicted at the time of our near death experiences of 1999 through 2001.

**The coup de grâce.** For several years we had discussed with WSLH the possibility of allowing web-based entry of results and web-based, instead of mailed, reports. They offered that service to some of their subscribers but the costs involved in reprogramming the reports for CTS seemed excessive. We hoped that WSLH would find it advantageous to offer us that ability at a comparable cost if we waited a year or two, since it would eliminate the expense of printing and mailing reports. In November 2013 they informed us that they were moving to an entirely web-based program that would be provided by an outside vendor. This involved significant added costs although they assured us that we were being charged only a pass-through amount that was billed by their vendor. Unfortunately, the change was not optional and we had already collected the dues from participant laboratories for our PT services though 2014. We were faced with the unpalatable alternatives of sending an additional bill for 2014, which the committee judged would be impossible because the contracts were already signed, or trying to load the 2015 fees enough to make up the difference. The latter course risked laboratories fleeing the program in a state of sticker shock.

As it turned out, we were able to continue through 2015 with an acceptable increase in charges. However, enrollment decreased somewhat more than expected in 2016 to the point where the income from the program looked like it would soon fall below the cost of running it. Reluctantly, the CTS leadership decided the end had come. We sent letters to all laboratories, thanking them for their support over the years and notifying them that they would need to enroll in a different program in 2017.
So, we had a 36-year run. The program has left its mark on the quality of patient care, and in some valuable research published in peer-reviewed journals. We also had some influence on both federal CLIA requirements and specifications of the California Laboratory Field Services.

References