

Benchmarking as part of competency testing

October 2003



Stephen Frost

Department of Clinical Biochemistry and Immunology
Brighton and Sussex University Hospitals NHS Trust
The Princess Royal Hospital
Haywards Heath
West Sussex
RH16 4EX
UK

Like many hospitals throughout the world, the Princess Royal Hospital in Sussex, England has set up a quality assurance (QA) scheme to verify the competency of glucose-meter operators. Their scoring system, however is anything but typical.

Instead of relying solely on individual results, the system also uses ward-based scores and cumulative reports. The system has led to a higher involvement among senior nurse staff groups, as it has turned them into active stakeholders in the process of evaluating and improving operator competency.

Recently, acute**care**testing.org interviewed Dr Stephen Frost from the Princess Royal Hospital's Department of Clinical Biochemistry and Immunology about setting up the system.

[acute**care**testing.org](http://acutecaretesting.org): Describe the benchmarking scheme you have set up.

Stephen Frost: The Princess Royal has about 35 wards

and departments that use glucose meters, with some local general practices and a private hospital also supported. The QA schemes at the Brighton Hospitals still operate slightly differently in some respects, as they evolved independently.

As is good practice, we long ago standardized on one type of glucose meter. I would say that the basic QA scheme is similar to that used in many hospitals these days.

Samples, which are animal-sera-based, are provided monthly on an external quality assurance scheme provided by the manufacturer. Samples are sent to each ward from the Biochemistry Department, with instructions that all trained staff on duty should measure the material once each and return the results to the laboratory for analysis.

Reports are processed by a small team of lab staff, usually a week or so later, and issued to the wards.

An aspect that is less typical is that we have introduced our own ward scoring system as an add-on to the reports for individual nurses. Individual nurse reports rather than ward or departmental reports have been the standard, in my experience, in most commercial packages for glucose QA as supplied by the glucose-meter firms.

Our ward-based reports are in effect a form of benchmarking for the wards rather than for individual nurses.

acutecaretesting.org: How are the scores calculated?

Stephen Frost: The scoring scheme is designed to be easy to calculate and assigns a score for each meter, irrespective of how many nurses participate. If a ward has two meters it gets two reports, one for each meter. The calculation is done in a simple Microsoft Excel program written in-house.

Although we still are using a paper-based system for these QA reports, it should easily be incorporated into electronic 'connectivity' approaches. The numbering used for the scoring system is somewhat arbitrary and has evolved pragmatically from experience, with minor changes over the years.

For the monthly results returned from each ward, we assess and score the performance of each nurse. The ward is given a score of 0 to 5 for that month.

Until recently this was calculated as follows: all results acceptable (within 10 % of the mean) = 5; at least 80 % acceptable = 3; at least 50 % acceptable = 2; less than 50 % acceptable = 1; no return = 0.

Because glucose meters have become easier to operate, we have recently introduced a modification to the scoring. A score of 5 is now given to wards with all participants within 5 % of the target mean, while those just reaching the 'good' 10 % target for all participants are now awarded a score of 4.

This reflects the fact that most wards can now achieve the old 10 % target provided they remember to send back their results. We prefer a range of scores rather than everyone achieving 100 %, which would give a false sense of security. I also like to award a nominal prize such as a box of chocolates every so often to the highest scoring ward, so it helps not to have everyone achieving 100 %.

In addition to this monthly score we provide a cumulative score based on the last six monthly returns. This is calculated as the sum of the last six scores for the ward, expressed as a percentage of the maximum possible (30). Most wards maintain a score of over 70 % and usually a few high-fliers have 100 % scores.

A ward's cumulative score is considered unacceptable if it falls to below 50 % and a score below 40 % requires the laboratory and senior nursing staff to agree on an action plan.

The scoring system rewards accurate results and penalizes off-target results. We have adopted this simple scoring system, rather than a more rigorous statistical approach, since poor results are often in the form of 'blunders' (that is, gross errors) rather than imprecision in the normal laboratory sense.

A 'blunder' is a result well off target, which could be due to faulty technique, transcription error, incorrect or out-of-date sample, and so on. Often these 'blunders' are excluded from statistical analysis because they are more than 3 SD from the target, distorting statistical data.

Statistical analysis, on the other hand, is more useful in comparing instruments, work practices, etc.[1,2]. A scoring system also has the advantage of being more easily understandable by those without statistical knowledge.

The scoring system penalizes poor participation rates. This is intentional and is a significant drive to ensuring that all wards participate often enough to obtain acceptable scores.

acutecaretesting.org: Who receives the reports?

Stephen Frost: Our QA reports, including a sheet giving the current ward score for this distribution (0-5) together with the cumulative score for the last six returns, are sent to link nurses for each ward or department.

The scores for all the other wards and departments are included, but only identified by a semi-confidential code letter. The link nurse is usually a senior nurse or sister.

acutecaretesting.org: Who takes corrective actions and how?

Stephen Frost: If their cumulative score is a cause for concern, normally a senior colleague or I discuss the issue with the link nurse and we agree on an action plan. Often this is due to poor participation and almost always the contact is sufficient to restore it.

Occasionally there have been genuine issues, such as difficulties transporting samples, which have needed addressing. A copy of the scores for all wards is sent to senior nursing and medical staff but their intervention is seldom required.

I also chair a multidisciplinary point-of-care testing group with access to senior nurse management, clinical risk managers, and other committees if referral to these is ever needed. However, I would see that as a last resort.

acutecaretesting.org: Why introduce a system based on ward results and cumulative reports?

Stephen Frost: There were two main reasons.

The first was to involve senior nursing staff. Usually the senior nursing staff groups in each ward are those with the most influence and it is important to make them stakeholders of the scheme.

The second reason was to give a reflection of performance over a period of time and not just a snapshot that is easily forgotten by the time of the next distribution.

The cumulative score addresses persistent poor performance and gives a simple numerical value that is hard for managers, either on the ward or in the laboratory, to ignore.

acutecaretesting.org: What were the benefits of setting a scoring systems based on ward results and cumulative reports?

Stephen Frost: Firstly, from a laboratory or scheme operator's point of view, there is a big advantage in having a numerical score reflecting cumulative performance. Otherwise it is very easy to forget about previous poor performance, or lack of participation, of a certain ward.

Before you know it a ward might have failed for many months without action being taken, particularly if the lab staff change.

For lab managers who send the report to nurse managers and clinicians in the hospital, it also focuses their mind. If a ward has a low score you probably will make sure you have taken action and can justify the report you produce.

At ward level, I feel that individual nurse participants will be motivated to perform well if they know their senior nursing staff groups are stakeholders in the QA scheme. The motivation of the senior nurse is probably the more complex and no doubt varies between individuals.

In a real world some will be motivated more by altruism and some by self-interest. Those who have commented to me have seen their report as adding value compared with a nurse-based report. There will often on their part be a genuine desire to perform well and to improve their ward's efficiency.

A related motivation is the desire to be seen to perform well, and to avoid the disapproval of their superiors.

The cumulative score introduces an element of competition with fellow wards and departments. Because it is cumulative, they know that a poor score

won't easily be forgotten. They also know the report will go to senior managers but not exactly who!

Finally, if they are seen as helping supervise QA this may improve their status with their more junior colleagues, however slightly.

acutecaretesting.org: How have the different staff groups reacted to this scheme?

Stephen Frost: We are fortunate that the ward-benchmarking element to our QA scheme was introduced in the early stages so nursing staff have always accepted it.

During a brief period we stopped this element in favor of only issuing reports to individual nurses and in fact participation fell. After six months we reintroduced it and participation more than recovered [3].

At that stage we asked for feedback and those link nurses that replied preferred a report with the ward-based benchmark element.

acutecaretesting.org: Do you have any plans to optimize the benchmarking system? If yes, how?

Stephen Frost: The next step forward almost certainly will come with connectivity and developments in electronic reporting of QA data.

This will remove the labor-intensive process of producing paper reports and enable more sophisticated benchmarking at an individual nurse level, perhaps related to re-validation, as well as at ward level.

This in turn will depend on progress with IT both within and outside the laboratory and with harmonization across the new Trust.

acutecaretesting.org: This scheme is only used for glucose testing. Do you have plans to include other types of point-of-care test in the benchmarking scheme? Why/why not?

Stephen Frost: The scheme's use for glucose is largely for historical reasons.

Glucose was the first analyte to be measured in large numbers of instruments outside the laboratory. Blood gas analyzers also have been subject to QA, but in the past we only had at most a handful of machines so the data from each machine could be more easily scrutinized without this kind of manipulation.

Recently, blood gas analyzers seem to be getting smaller, require less maintenance and are easier to support on multiple sites. The days of a single machine per hospital, perhaps in the lab, are fading.

At present in the Princess Royal we have eight blood gas machines, while on the Brighton sites they number in double figures. Procurement proposals are appearing frequently. Benchmarking can only help rationalize this situation and may tease out analytical weaknesses of any of the current machines.

I would expect a benchmarking approach to become more widespread as the numbers of blood gas instruments increase.

The benchmarking scheme could also be applied to any of the other types of POCT analysis that requires QA in large numbers, such as urinalysis, pregnancy testing and so on. For this to happen the facility for connectivity will have to be incorporated into these types of devices.

Princess Royal Hospital

Until 2002, the Princess Royal Hospital was a medium-sized district general hospital with about 400 beds. However, in April 2002 the hospital merged with the Brighton Hospitals to form The Brighton and Sussex University Hospitals Trust. The whole trust now includes six hospitals with 1,140 patient beds and a local population of 460,000. Since the merger, Clinical Biochemistry services are becoming organized as a single combined department across these sites.

Interviewer

Ana Cristina Magalhães

Radiometer Medical ApS

Åkandevej 21

2700 Brønshøj

Denmark

References

1. Frost SJ, Firth GB. A year's experience of the Roche Advantage II glucose test strip. *Clin Chim Acta* 2001; 307: 69-73.
2. Frost SJ, Wheatley T, Harris C, Firth GB. The performance of diabetic patients in a glucose meter quality assurance scheme. *Proc ACB Natl Meeting* 1993: 88.
3. Frost SJ, Firth GB. Nurses' participation in a glucose meter quality assurance scheme appears to be improved by a report to senior ward staff. *Point of Care* 2003; 2: 195-200.