Quality of interpretative commenting on common clinical chemistry results in the Asia-Pacific region and Africa

Samuel D. Vasikaran1,2,*, Leslie C. Lai3, Sunil Sethi4, Joseph B. Lopez5 and Kenneth A. Sikaris6

1 Department of Core Clinical Pathology and Biochemistry, PathWest Laboratory Medicine, Royal Perth Hospital, Perth, WA, Australia
2 School of Pathology and Laboratory Medicine, Faculty of Medicine and Dentistry, University of Western Australia, Nedlands, WA, Australia
3 Gleneagles Intan Medical Center, Kuala Lumpur, Malaysia
4 Department of Laboratory Medicine, National University Health System, Singapore
5 Department of Biomedical Sciences, MAHSA University College, Damansara Heights, Kuala Lumpur, Malaysia
6 Melbourne Pathology, Collingwood, Australia

Abstract

Background: Interpretative commenting is an important activity of the clinical diagnostic laboratory. We describe a study of interpretative commenting abilities among senior laboratory professionals in the Asia-Pacific region and Africa.

Methods: Five sets of laboratory results reflecting common and important problems encountered in clinical chemistry were distributed at 4-weekly intervals to 31 registered participants from countries in the Asia-Pacific region and Africa. Participants were asked to attach an interpretative comment to the results assuming that the requesting doctor had asked for an interpretation of the result.

Results: Twelve pathologists and 19 scientists from seven countries registered to participate and the overall reply rate was ~50% for the five cases. The quality of the comments returned by participants was diverse and some reflected incorrect or misleading interpretation and advice.

Conclusions: While interpretative commenting is an important laboratory activity, the results of this study suggest that there is room for improvement in the quality of interpretative comments offered by senior laboratory professionals, even for commonly reported results relating to most prevalent and important public health conditions. Interpretative commenting should be formally taught during training of pathologists and scientists, and continuing professional development in this area is required for the provision of a quality interpretative service.

Keywords: continuing professional development; interpretative commenting; narrative interpretations; post-analytic; quality assessment.

Introduction

The role of the clinical diagnostic laboratory includes offering advice on the selection and interpretation of laboratory tests. While most anatomical pathology laboratory reports carry an interpretative comment, clinical pathology reports mostly do not, nor is this possible given the volume of reports produced. However, clinicians appreciate receiving interpretative comments on the more complex tests and feel that they help or influence patient management, save them time, and improve the diagnostic process (1, 2). While there are no randomized control studies to confirm the effectiveness of interpretative comments on improving outcomes, there are some observational data to suggest that this may be the case (3). This is not surprising given that doctors at graduation often lack adequate training in test ordering and result interpretation (4). However, inappropriate commenting by inadequately trained laboratory personnel can be dangerous (5, 6). Training in pathology consultation has to be undertaken in a formal manner for this activity to be performed effectively (7). There is now some acceptance in the clinical chemistry profession for the need and ability to assess the quality of interpretative services offered by the clinical chemistry laboratory (6, 8, 9). We describe a study of interpretative commenting ability among laboratory personnel in the Asia-Pacific region and Africa.

Materials and methods

Expressions of interest were invited from individuals and laboratories from the Asia-Pacific region via the Asian and Pacific Federation of Clinical Biochemistry (APFCB) council representatives and the APFCB Committee on Laboratory Management (C-LM) representatives of each country to participate in the Commenting Educational Program. The representative from each APFCB member country had discretion to decide on how to publicize this program and invite participants in their country. Participation was free, but prior registration with the APFCB C-LM was required.

One set of laboratory results with some clinical background was distributed by e-mail to registered participants 5 times during 2008 at approximately 4-weekly intervals. The
format of the report was an electronic version of a set of laboratory results together with any relevant clinical information (Table 1). Participants were invited to assume that the requesting doctor had asked for an interpretation of the results and were asked to attach an interpretative comment (maximum 500 characters including spaces) to the report and return to the organizers by e-mail within a fortnight. This may represent a written, printed or verbal comment from the reporting laboratory (Pathologist/Scientist) to the requesting doctor. A total of five cases were circulated over 5 months (Table 2).

Once the replies from participants were received, they were reviewed by the Commenting Educational Program Committee (consisting of SDV, KAS, LCL and SS). An example of a good interpretative comment and an overview and discussion of the case, together with reference to any relevant publications, were returned to participants by e-mail within a few weeks of the closing date for replies. Anonymity of participants was maintained. Thus, participants were able to review their interpretation of results and communication to the requesting doctor and compare it with current best practice in the opinion of the Program Committee. For the purpose of this publication, the individual comments were scored on a scale of 0–4 (0 being a bad comment and 4 an ideal comment) independently by two members of the Program Committee and the scores averaged. The scoring of comments by the two committee members were compared by statistical analysis. The average score for all the comments by participants who were pathologists was calculated and compared with the average score for all comments by participants who were scientists (Student’s t-test).

Results

A total of 31 people registered for the program; 27 were from the APFCB region and an additional four people from Africa who had heard about the program by word of mouth also registered after the program had started (Table 3). Most of the participants were senior laboratory staff, with 12 of the registrants being pathologists and 19 scientists. A sample of replies received from participants together with the mean score awarded to each comment by the committee are listed in Tables 4–8 (comments longer than 500 characters have been truncated at 500).
Table 3 The number of participants listed by country of origin.

<table>
<thead>
<tr>
<th>Country</th>
<th>Number of registrants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hong Kong</td>
<td>2</td>
</tr>
<tr>
<td>India</td>
<td>11</td>
</tr>
<tr>
<td>Malaysia</td>
<td>7</td>
</tr>
<tr>
<td>Pakistan</td>
<td>1</td>
</tr>
<tr>
<td>Sri Lanka</td>
<td>6</td>
</tr>
<tr>
<td>Kenya</td>
<td>1</td>
</tr>
<tr>
<td>South Africa</td>
<td>3</td>
</tr>
</tbody>
</table>

with the comment composed by the Program Committee for each case. Participation fluctuated at around 50% which, while disappointing, is not dissimilar to other programs where participants pay to register (6).

**Case 1**

The first case had 14 responses out of which 10 had a diagnosis of impaired fasting glycemia (IFG) (Table 4). One comment had a diagnosis of “borderline diabetes”, one “mild diabetes”, one “slightly high fasting glucose” and two had impaired glucose tolerance (IGT). Suggestions for further investigations included measurement of insulin in one comment, “urine sugar” in one, repeat fasting glucose without a time stipulation, or at 1 or 6 months, and repeat oral glucose tolerance testing (OGTT) at 1 or 12 months. Five of the comments had no suggestion for further investigation or follow-up. Only five comments had suggestions for lifestyle modification.

**Case 2**

The second case elicited 16 responses which included a diagnosis of subclinical hypothyroidism in 10 comments, hypothyroidism in one comment, “normal results” in one comment and early or mild hypothyroidism in three comments (Table 5). Four comments included a suggestion to consider replacement therapy, seven suggested repeat testing, six suggested thyroid peroxidase (TPO) antibody testing and two each suggested free triiodothyronine (FT3) testing or imaging studies. One comment had a diagnosis only with no suggestion for further investigation or follow-up.

**Case 3**

This case elicited 17 responses, 12 of which stated that the result would be above the age related upper reference limit (Table 6). Two comments stated the results were normal and two suggested the results indicated benign hyperplasia. Three comments listed other causes for a transient increase in prostate specific antigen (PSA). Eight comments suggested performing a free PSA (PSA) measurement and 11 suggested repeat PSA testing ranging from 1 week, 1 month, and 6 months to 1 year. One comment included an overt suggestion for urology referral while three suggested prostatic biopsy, four advised digital rectal examination (DRE) and two an ultrasound examination.

**Case 4**

We received 15 responses of which seven referred to decreased renal function or estimated glomerular filtration rate (eGFR) and five mentioned stage 3 chronic kidney disease (CKD) while one referred to grade 2 CKD and five stated the renal function was normal for the patient’s age (Table 7). Only one comment mentioned race may be a confounder, as the estimate may

---

### Table 4 Examples of participant comments with designation of participant (D – Medical, S – Scientist) and mean score given by the Program Committee, together with the comment drafted by the Program Committee for case 1.

<table>
<thead>
<tr>
<th>D/S</th>
<th>Mean score</th>
<th>Participants’ comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>0</td>
<td>This is a case of “impaired glucose tolerance (IGT) and mild diabetes.” At this stage the patient does not need treatment.</td>
</tr>
<tr>
<td>S</td>
<td>1</td>
<td>Patient’s fasting glucose is high but 2 h post 75 g glucose result is lower than the diabetic diagnosis criteria (&lt;7.7). This condition shows that patient is having impaired glucose tolerance. Suggest to do serum insulin.</td>
</tr>
<tr>
<td>D</td>
<td>2</td>
<td>Laboratory diagnosis impaired glucose tolerance. The accepted fasting, 1st and 2nd hour levels are 5.5, 9.4 and 7.7 mmol/L, respectively. The fasting is higher than the normal reference range. The 2nd hour sample is well within the acceptable range of 7.7.</td>
</tr>
<tr>
<td>S</td>
<td>3</td>
<td>The report revealed that patient had impaired fasting glucose (IFG) levels and 60 min levels suggest lag storage response. It is recommended to undergo repeat oral glucose tolerance testing (OGTT) after 6 months for follow-up.</td>
</tr>
<tr>
<td>D</td>
<td>4</td>
<td>This patient has an IFG on her OGTT. It is recommended that the OGTT be repeated annually and she be advised on lifestyle modification and cardiovascular risk factors assessment.</td>
</tr>
</tbody>
</table>

Comment drafted by the Program Committee
This patient has an IFG on her OGTT. It is recommended that the OGTT be repeated annually and she be advised on lifestyle modification and cardiovascular risk factors assessment.
not be appropriate for blacks. None of the comments mentioned the appropriateness or otherwise of the estimate in an Asian population. Six comments mentioned that creatinine was stable in the last 12 months and eight comments mentioned diabetes and/or hypertension as possible causes of renal damage. Three comments suggested there was likely to be no renal damage, and all the comments had suggestions for further evaluation, investigation or follow-up. Three comments suggested avoiding nephrotoxic drugs, two suggested ultrasound investigation and two advised specialist referral. One comment advised GFR measurement.

Case 5

The fifth and final case elicited 13 responses of which nine stated the low-density lipoprotein (LDL) was above target value (Table 8). Four suggested reviewing/optimizing drug therapy, three suggested reviewing lifestyle measures and two suggested looking at secondary causes. Two advised checking liver function tests and/or muscle enzymes. Four comments stated that target value had been achieved and one of these suggested a trial cessation of lipid lowering therapy. Two comments advised measures to increase high-density lipoprotein (HDL).
The two members of the Program Committee who independently scored comments gave identical scores for 35 out of the total of 75 comments and gave scores that were different by one for the rest of the comments, with each adjudicator giving a higher score for an equal number ($n=20$) of comments. The Wilcoxon sign rank test showed no difference in the scores: $z$-value $0.2688$ ($0.7872 < p < 0.7949$) and Spearman’s rank order correlation was $0.6633$ ($p<0.001$). The mean scores obtained by medically trained participants ($2.0$) was not significantly different from the mean score obtained by scientifically trained participants ($1.8$).

**Discussion**

Interpretative commenting is an integral part of the activities and responsibilities of the clinical diagnostic laboratory. Our program was developed as an educational activity for interested participants from the Asia-Pacific region, but the results also provide a
snapshot of commenting ability among pathologists and laboratory scientists in the region. The comments of the participants were not individually assessed as the main educational component was the general feedback offered in the form of a commentary by the Program Committee as well as a case discussion (Table 9). Scoring the comments or some feedback to individual participants on the appropriateness may be done in the future (6, 8). The five reports circulated reflected common and important problems encountered in clinical chemistry investigations with which practising clinical chemists should be familiar. The somewhat artificial environment of the study has to be acknowledged, since in many real-life settings the interpreter would have the opportunity to review the entire medical record of the patient if necessary, and may also be able to discuss issues of concern with the ordering clinician. However, the quality of the comments attached to the cases by participants in this study was diverse; some responses received from the participants reflected suboptimal or misleading interpretation and advice which cannot be accounted for by lack of information.

The effectiveness of a diagnostic clinical laboratory depends on the selection of appropriate tests, accurate and timely measurement and reporting of results for the correct patients, and appropriate interpretation of the results and action. The laboratory has a role in each of these steps. It is an interesting paradox that while major progress is being made in all aspects of measurement in order to produce accurate and precise results, with ever shorter turn-around of results due to improvements in information technology and auto validation rules, and appropriate test selection due in part to electronic order entry, the interpretation of results has not seen similar improvements. Interpretation has remained largely a manual activity despite some attempts at the development of expert systems (21). In any case, even the quality of the interpretation offered by an expert system will be depend-

Table 9 A brief discussion of the five cases circulated in the program.

<table>
<thead>
<tr>
<th>Case number</th>
<th>Discussion of case</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The WHO defined values for IFG are a fasting plasma glucose concentration in the range of 6.1–6.9 mmol/L, and a 2-h post-glucose load (if measured) &lt;7.8 mmol/L (11). The alternative definition of IFG by the American Diabetes Association of a fasting plasma glucose concentration between 5.6 mmol/L and 6.9 mmol/L and a 2-h post-glucose load &lt;7.8 mmol/L makes no difference to the interpretation of this case (12). About 3%–10% of people per year with IFG develop diabetes (13). There are limited data on the preventability of progression of IFG to diabetes by lifestyle measures. The cardiovascular risk profile of people with IFG should be checked (blood pressure measurement, serum lipids etc.). IFG is also associated with the metabolic syndrome; impairment of glucose metabolism is an optional criterion for the definition of the metabolic syndrome. Retesting may be performed in 12 months’ time to assess change in glucose tolerance.</td>
</tr>
<tr>
<td>2</td>
<td>Hypothyroidism whether overt or subclinical can have an adverse impact on the course of pregnancy or fetal development and should be corrected before initiation of pregnancy (14). The Endocrine Society Clinical Practice Guideline recommends adjustment of thyroid (T4) therapy to attain a thyroid stimulating hormone (TSH) &lt;2.6 mU/L before conception based on the finding that babies born to women with overt or subclinical hypothyroidism have an average intelligence quotient (IQ) score 7 points below the mean IQ score of babies born to healthy women and women on T4 replacement (15). Fetal thyroid develops in the second trimester of pregnancy followed by fetal thyroid hormone production in mid-gestation.</td>
</tr>
<tr>
<td>3</td>
<td>Use of age-specific PSA reference intervals appropriate for the assay used increases the sensitivity of the test in younger men while maintaining 95% specificity. Those men found to have a prostate cancer with a PSA 2.6–4.0 µg/L have an 81% chance of their cancer being confined to the prostate gland (16). For a raised PSA that is below 10 µg/L, free PSA (fPSA) may help in risk stratification. Likelihood of prostate cancer is high if fPSA &lt;10% and low if &gt;25%. This patient should at least have a repeat PSA ± fPSA in 1 month and ideally should be referred to a urologist, especially if there is a strong family history. The urologist can repeat one or more tests and recommend a transrectal ultrasound and biopsy of the prostate if necessary.</td>
</tr>
<tr>
<td>4</td>
<td>According to the National Kidney Foundation Kidney Disease Outcomes Quality Initiative CKD classification these eGFR results would place this patient in stage 3 CKD (17). While age-related reference intervals for serum creatinine or eGFR are not recommended (17), eGFR does decline with age generally, although it is possible that the decline is seen mainly in those with hypertension (18). Deciding if the decline is due to pathology or not is not straightforward, and the results should be interpreted with caution. The Australian Creatinine Consensus Working Party has concluded that “it was appropriate to advise medical practitioners that, in people aged ≥70 years, eGFR values in the range 45–59 mL/min/1.73 m², if stable over time and unaccompanied by other evidence of kidney damage (e.g., proteinuria, hematuria etc.), may be interpreted as typical for this age group and are unlikely to be associated with CKD complications (19).”</td>
</tr>
<tr>
<td>5</td>
<td>The main therapeutic target for lipid lowering therapy in coronary heart disease (CHD) is LDL cholesterol. LDL-lowering therapy reduces total mortality, coronary mortality, major coronary events, coronary artery procedures, and stroke in persons with established CHD (20). The National Cholesterol Education Program Adult Treatment Panel (NCEP ATP) III guidelines suggest an LDL cholesterol level of &lt;2.6 mmol/L as optimal for secondary prevention in established CHD and for persons with CHD risk equivalents (20). National guidelines may vary e.g., the National Heart Foundation of Australia guidelines recommend an LDL cholesterol target &lt;2.0 mmol/L in therapy for secondary prevention (10).</td>
</tr>
</tbody>
</table>

IFG, impaired fasting glucose; PSA, prostate specific antigen; CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate; LDL, low-density lipoprotein; T4, thyroxin.
ent on the knowledge and commenting skills of the person building the knowledge base. Proficiency testing of this activity has only recently been attempted (22). Narrative interpretative comments are still a voluntary and unregulated aspect of a laboratory result. Lack of clinical information provided with test requests and medico-legal implications of inappropriate interpretative advice are some limiting factors. In the future, interpretative commenting may become part of a contractual requirement – as tests become more complex, population-specific and qualitative in their nature. Education in these activities is an important part of training as well as continuing professional development of laboratory personnel. The APFCB and its members represent postgraduate specialist societies comprised of practising laboratory professionals, and the individual participants in this program were pathologists and mainly senior scientists. The comments received in this program are reflective of laboratory practice in the Asia-Pacific region.

The voluntary nature of the study might be expected to be biased toward laboratories that perform better than average with their interpretive comments. Therefore, the results of this program as well as other published studies suggest that there is considerable room for improvement in the quality of interpretative comments offered by laboratory professionals (6). Professional organizations have a role in the education of laboratory personnel, both pathologists and scientists, in interpretative commenting on laboratory results. The performance of pathologists was not significantly different to that of the scientists in terms of the mean score obtained. This may be surprising, but underscores the need for training pathologist in interpretative commenting as part of their education. Whilst the case studies point out some regional differences in guidelines in interpretation, the discussion shows that such a program can operate across borders, especially as professionals should be aware of any existing regional debates regarding interpretation.

The educational value of programs such as this would be enhanced by individualized feedback about participants’ own comments. Scoring comments is an obvious strategy, although it could be somewhat subjective and may lack consistency. While giving a score for comments may be helpful to enable participants to assess their standard of commenting, individual feedback on each aspect of a comment would be more useful educationally. However, such individualized feedback is very labor intensive. Another approach used is the analysis of comments by classifying common key-phrases in the comments as preferred, inappropriate, etc.; which may retain the educational usefulness of the feedback to participants and be applicable where there are large numbers of participants (6). Further work is needed to explore and develop appropriate and useful methods for evaluation of comments and educational feedback to laboratory professionals who participate in this activity. Whatever the strategy employed, the results of our study indicate that training and continuing professional development in the area of interpretative commenting is important and needed in the regions studied.

Statement

There are no conflicts of interest for any of the authors, and none of the authors have been employed or financed by anyone with economic interests. We have no other conflicts of interest.

Acknowledgements

We thank Dr. Amanda Hooper for proof reading the manuscript.

References