

Valutazione Esterna di Qualità: esperienza con il programma del College of American Pathologists

D. Giavarina, G. Soffiati

Laboratorio di Chimica clinica ed Ematologia, Ospedale San Bortolo, Vicenza

Le prove di valutazione esterna di qualità (VEQ), che molto spesso coincidono con prove di abilità o proficiency testing (PT), forniscono una misura dell'efficacia dei programmi di controllo di qualità dei laboratori (1). Superare una prova di valutazione esterna di qualità significa dimostrare che i programmi di controllo di qualità interni, attuati dal laboratorio, ottengono il loro obiettivo di garantire il risultato "esterno". I risultati delle PT sono anche utili per determinare, nel tempo, se un metodo è eseguito secondo i parametri previsti dal produttore e con buona pratica di laboratorio.

Oltre a questo, i risultati di questi esercizi possono definire l'imprecisione interlaboratorio di un metodo ed anche documentare le preferenze attuali relativamente ai metodi stessi, dimostrando quali siano le strumentazioni, i metodi, le procedure più diffuse nell'area geografica interessata dal programma.

Minori sono invece le possibili ricadute all'interno del laboratorio: i risultati di una PT non possono essere usati per valutare l'accuratezza di un metodo o l'utilità clinica, né per confrontare l'accuratezza di metodi diversi (2). Le PT sono, in realtà degli indicatori delle performance di laboratorio, ma non possono costituire l'unico sistema di misura. Alcune delle limitazioni delle PT possono essere ricondotte a quattro principali problemi (3):

1. Gli esercizi sono in genere incompleti. Non tutti i processi di analisi né tutti i metodi sono sempre previsti. Molti test non hanno valutazione.
2. I materiali per le prove sono trattati in maniera differente rispetto ai campioni.
3. I criteri di accettabilità di una prova sono spesso legati alla media e alla varianza dei risultati di tutti

i partecipanti. Questo rende il risultato dipendente dal numero dei partecipanti, dalla diffusione del metodo, dalla variabilità analitica dello stesso. Per superare questo punto fondamentale di valutazione di una PT, è stato proposto che il livello target sia determinato con un metodo di riferimento ampiamente accettato, mentre l'intervallo di accettabilità dovrebbe essere in relazione con le esigenze cliniche piuttosto che con variabilità statistiche (4).

4. L'effetto matrice affligge la qualità di molte prove. Ad esempio, è stato dimostrato un bias dovuto all'effetto matrice nel 70% di 37 sistemi strumento/reagente specifici per l'analisi dei livelli di colesterolo (5).

I programmi di Valutazione Esterna di Qualità (VEQ) prevedono l'esecuzione di esercizi comuni da parte di gruppi di laboratori più o meno omogenei, ad intervalli variabili.

Gli obiettivi che si può porre il gestore di un programma di VEQ sono molteplici e a volte non sempre ben definiti. In alcuni paesi, come gli Stati Uniti, la partecipazione alle PT è obbligatoria (CLIA '67 e '88) e, ai fini di ottenere il riconoscimento al rimborso da parte del sistema sanitario o assicurativo, è necessario non solo la partecipazione ma anche il superamento delle "prove di abilità" eseguite (6).

La valutazione esterna della qualità del servizio di laboratorio ha in Italia, ed in particolare nel Veneto, una definizione incerta (DPCM 18 febbraio 1984) e più diretta alle realtà private che a quelle pubbliche. Pur con caratteristiche "educative", ossia non coercitive e limitative dell'attività stessa, alcune regioni hanno un qualche sistema di VEQ più o meno obbligatorio. La partecipazione a questi programmi è talvolta non soddisfacente, riproponendo alcuni dei problemi sopra elencati:

1. La tipologia delle prove di abilità richieste da questi esercizi spesso si limita a pochi, anche se importanti, settori del laboratorio (chimica clini-

Corrispondenza: dott. Davide Giavarina
Laboratorio di Chimica clinica
ed Ematologia, Ospedale San Bortolo, Vicenza
Tel. 0444-993372; e-mail: giavarina@libero.it

ca, ematologia, coagulazione, immunometria).

2. Il numero delle prove per esercizio è talora esiguo, su pochi livelli diversi, a volte unico.
3. La cadenza temporale degli esercizi è molto variabile e spesso non regolare.
4. La qualità dei materiali di controllo non sempre è soddisfacente.
5. La partecipazione ad un esercizio, divisa per metodo, presenta un numero di partecipanti che può non essere sufficiente per una corretta e possibile valutazione statistica.
6. Le prove sono limitate alla fase analitica dell'attività di laboratorio.
7. La gestione delle diverse prove è disomogenea sia nella ricezione dei materiali, sia nell'esecuzione, sia nella interpretazione e rappresentazione dei risultati, poiché eseguito da ditte diverse, secondo direttive di centri diversi.

Le necessità statunitensi di "abilitare e verificare nel tempo" i laboratori per ogni test eseguito, ma anche per ogni risposta diagnostica oltre il test, in maniera omogenea su tutto il territorio e tra i diversi laboratori, ha probabilmente permesso lo sviluppo di programmi di VEQ molto robusti, in particolare se confrontati con analoghi europei.

Pur avendo una valenza amministrativa ed anche legale, il più grosso programma di Proficiency Testing è gestito non da un ente governativo ma da un grande organismo professionale, il College of American Pathologists (CAP).

Tabella I: elenco riassuntivo del numero di esercizi diversi disponibili dal CAP per l'anno 2001

| Disciplina | Numero di esercizi diversi |
|-------------------------------|----------------------------|
| Emogasanalisi | 2 |
| Chimica | 5 |
| Markers cardiaci | 5 |
| Coagulazione | 2 |
| Microscopia clinica | 3 |
| Tipizzazione HLA DR/DQ | 2 |
| Elettroforesi | 2 |
| Citometria a flusso | 3 |
| Istocompatibilità | 4 |
| Epatite C carica virale | 2 |
| HIV carica virale | 2 |
| Analisi d'immagine | 2 |
| Medicina trasfusionale | 2 |
| Immunometria | 3 |
| Immunometria SI | 2 |
| Oncologia molecolare | 3 |
| Diagnostica immunologica | 11 |
| Analisi del seme | 4 |
| Analisi del sudore | 4 |
| Markers tumorali | 2 |
| Markers virali | 5 |
| Coagulazione su sangue intero | 3 |
| Immunometria speciale | 4 |
| TDM | 5 |

Il Programma

Il "CAP Survey & Educational Anatomic Pathology Programs" presenta una grande tipologia di test e prove per i laboratori. Come dichiarato dallo stesso programma, "costituisce un impegno verso la qualità della cura del paziente, integrandolo nel lavoro quotidiano dei laboratori. ...rappresenta il più alto standard per il trattamento dei pazienti, includendo le più recenti ed innovative tecniche e tecnologie."

Il programma fornisce materiali di controllo per moltissimi test di laboratorio, dalla chimica alle proteine, ai liquidi biologici, markers oncologici e di danno miocardico, tossicologia, farmacologia, elettroforesi emoglobinica, sangue occulto, urine, screening materno/fetali, citofluorimetria, immunologia e autoimmunità, ossimetria, immunodosaggi, enzimologia, allergologia, anticorpi antifosfolipidi, e altri. Il programma offre anche controlli per la microbiologia e l'anatomia patologica. La tabella I riassume le possibilità per l'anno in corso.

Ogni esercizio prevede tre invii l'anno, con calendario definito. Gli esercizi hanno da tre a cinque livelli per costituente. Ciascun invio è accompagnato da un documento di alcune pagine dove sono elencati gli obiettivi dell'esercizio, il modo di esecuzione e compilazione del modulo di risposta, la codifica dei metodi e delle strumentazioni, la codifica degli eventuali problemi riscontrati. I moduli per le risposte sono strutturati secondo uno stesso standard per tutti gli esercizi, in modo da facilitarne l'uso.

Oltre alle prove di abilità analitica, ampie e articolate, ciò che rende interessanti questi esercizi sono le collegate prove di abilità diagnostica e clinica. Per molte prove è richiesto di dire se il risultato ottenuto è classificato dal laboratorio come "normal" o "abnormal"; per alcune prove è richiesto il risultato diagnostico finale più che il livello misurato; gli screening materno/fetali sono strutturati come un test reale, con i dati della paziente da integrare ai dati analitici per ottenere la valutazione del rischio; alcuni esercizi richiedono, oltre alla valutazione clinica del risultato, anche il consiglio da dare al paziente o al medico curante, (come ad es. per le emoglobine patologiche), considerando la consulenza un'attività non derogabile da parte di un laboratorio che si occupi di questa diagnostica.

Sono valutate anche alcune abilità degli operatori. Ad esempio, preparati microscopici di ematologia o preparati per immunofluorescenza microscopica valutano le capacità di operatori specialisti. Altri esercizi puramente teorici sono proposti come quiz per la verifica della capacità diagnostica.

Esperienza di Partecipazione dall'Italia

Il laboratorio di Chimica clinica ed Ematologia dell'Ospedale San Bortolo di Vicenza ha aderito al programma di VEQ del CAP a partire dal 1999. Poiché

uno degli obiettivi principali era di avere prove per tutti i test eseguiti, la fase iniziale è stata rivolta all'analisi dei test eseguiti e alla ottimizzazione dei diversi "Survey" cui abbonarsi. Alla fine di questo processo è stato possibile individuare una prova relativamente al 98% circa dei test eseguiti dal laboratorio. Considerando due ripetizioni per chimica clinica ed ematologia di routine e di urgenza, sono risultati necessari 32 esercizi diversi(a).

La possibilità di avere un controllo per ogni test non è stato l'unico obiettivo.

Nel complesso gli obiettivi prefissati sono stati i seguenti:

1. istituire una VEQ per ogni test eseguito nel laboratorio;
2. estendere le "prove di abilità" (PT) oltre la sola fase analitica;
3. rientrare in gruppi omogenei di numerosità sufficiente per un valido giudizio statistico dei risultati;
4. uniformare le procedure di VEQ del laboratorio;
5. razionalizzare la spesa.

Valutazione del raggiungimento degli obiettivi

Alla fine del primo anno di esperienza abbiamo tentato una valutazione complessiva.

- 1) Come si era calcolato nella fase di acquisizione, l'abbonamento a questo tipo di servizio ha permesso di sottoporre a prova circa il 98% dei test eseguiti presso in nostro laboratorio. La tipologia dei test per i quali non è stato possibile applicare un esercizio appropriato è risultata essenzialmente composta da test su materiali particolari, come l'esame di calcoli, o da test rari, come nel caso della zincoproporfirina. Un'eccezione è stata relativa alla VES, che eseguita con una strumentazione di relativa nuova introduzione (TEST1, SIRE Analytical System, Udine, I), non ha permesso l'uso dei materiali proposti.
- 2) La fase preanalitica rimane essenzialmente al di fuori della possibilità di valutazione con questi controlli. Molto spesso sono richieste alcune valutazioni sullo stato dei campioni all'arrivo, compresa in qualche caso la misura della temperatura della confezione, nonché una certa tempestività di esecuzione. Le specifiche sono però lontane dalle problema-

tiche preanalitiche che si hanno con pazienti e campioni reali.

In alcuni esercizi è invece applicata un'impegnativa prova di abilità post test, come nel caso dell'interpretazione e consiglio al medico curante per le emoglobine patologiche o nello screening materno-fetale.

- 3) Poiché negli Stati Uniti questo tipo di controllo è obbligatorio, mediamente la numerosità è oltremodo sufficiente per la valutazione statistica, arrivando anche a oltre 1000 partecipanti per esercizio. In qualche caso, metodi anche piuttosto diffusi in Europa non hanno mercato negli Stati Uniti, e in quest'evenienza la numerosità è risultata insufficiente.
- 4) L'uniformità delle procedure di VEQ è stata ottenuta grazie ad un iter standard di processo dei campioni e di compilazione dei report, che sono simili per i diversi controlli ed accompagnati da estese spiegazioni. Inoltre è stata introdotta una procedura di indagine sui risultati non accettabili secondo lo standard NCCLS GP27-A, che ha contribuito all'ulteriore standardizzazione delle procedure. L'utilizzo di un modulo standard per la valutazione della risposta ai programmi permette alcuni miglioramenti: in primo luogo ogni controllo sarà adeguatamente esaminato nella risposta e valutato nella sostanza; l'analisi delle risposte non conformi spinge ad un'indagine in ogni settore ed in ogni campo interessato alla produzione di quel risultato, motivando e coinvolgendo tutto il personale sia nelle problematiche teoriche della qualità relative all'esercizio in esame che per gli esercizi successivi. Inoltre, a livello documentativo si dimostra un'azione compiuta a fronte di un risultato non atteso al fine del miglioramento della qualità. Archiviare risultati di prove di abilità falliti senza alcuna documentazione di quanto si è fatto per capire le cause e correggere gli errori potrebbe avere l'effetto di documentare un'inefficienza del laboratorio che aderisce a programmi di valutazione di qualità e abilità. Noi utilizziamo questa procedura anche per i risultati accettabili, nella convinzione che "... gli utilizzatori possono scoprire problemi dalla interpretazione dei loro propri risultati, anche se nessuno dei cinque risultati dei livelli della PT ha superato l'errore accettabile (8)". La figura 1 riporta il modulo di istruttoria utilizzato per l'analisi dei risultati.

^{a)} gli esercizi erano:

Anti-Cardiolipin
Critical Care/Aqueous Blood gas
Chemistry General
Coagulation-Limited Cerebrospinal Fluid
Chemistry Occult Blood Reticulocyte
Immunology Special
Diagnostic Allergy
TDM-General
TDM Special

Serum Alcohol/Volatiles
Cardiac Iso Immunochemical Hematology-
Comprehensive Autodiff
Flow Cytometry Lymph. Phenotype
Body Fluid
Maternal Screening Glycohemoglobin
Hemoglobinopathy
Immunology-General
Antinuclear Antibody
Blood Lead

Ligand-General
Ligand-Special
Ligand Special (Rap)
Blood Oximetry
Serum Protein Electrophoresis
Tumor Markers
Troponins I And T
Urine Chemistry General
Urine Chemistry-Special
Urine Drug Testing (Screening)

Figura 1: Modulo di istruttoria per risultato inaccettabile di valutazione esterna di qualità. Il modulo è utilizzato anche come approvazione di risultati accettabili da parte del direttore.

Servizio Sanitario Nazionale – Regione Veneto **ULSS N. 6 “VICENZA”**
PRESIDIO OSPEDALIERO DI VICENZA – Laboratorio di Chimica clinica ed Ematologia
 ISTRUTTORIA PER RISULTATO INACCETTABILE DI VALUTAZIONE ESTERNA DI QUALITÀ
STANDARD NCCLS GP27-A

| | |
|---|--|
| Data dell'indagine | |
| Identificativo del campione di VEQ | |
| Data dell'analisi | |
| Risultato inaccettabile | |
| Risultato atteso | |
| Intervallo di accettabilità | |
| Risultati inaccettabili precedenti o trend per questo esame | |
| Revisione delle trascrizioni: FASE POST-ANALITICA | |
| Indagine sulla FASE ANALITICA | |
| CONCLUSIONI | |
| Effetto sui risultati dei pazienti | |
| Classificazione del problema | <input type="checkbox"/> errore di trascrizione <input type="checkbox"/> metodologico <input type="checkbox"/> tecnico <input type="checkbox"/> del materiale VEQ <input type="checkbox"/> della valutazione nella VEQ <input type="checkbox"/> inspiegabile |
| Azioni correttive | |
| Approvazioni | Supervisore data Direttore data |

5) La spesa complessiva non è particolarmente onerosa alla fonte. Il costo complessivo per il programma 2000 è di \$ 12.988. La valenza economica può sembrare non conveniente, se confrontata con altri programmi italiani o europei. Il problema del costo è anche in parte contingente, perché legato all'attuale debolezza di cambio dell'euro rispetto al dollaro. La "copertura" di quasi tutti i test del nostro servizio richiede l'adesione a 32 esercizi (96 prove l'anno). Considerando il gran numero di prove e di test verificati, il costo per test diviene competitivo. Esistono problemi relativi al costo di sdoganamento ed invio dei campioni dagli aeroporti italiani al Laboratorio. La possibilità di un'agenzia italiana che potesse gestire per l'Italia gli abbonamenti ai diversi Surveys, senza dover direttamente importare dagli Stati Uniti, sarebbe di grande aiuto, soprattutto per il superamento delle difficoltà amministrative di importazione.

Conclusioni

È stato scritto su questo giornale che la partecipazione ad un programma di VEQ è vissuta spesso alterando l'atteggiamento di chi adempie ad un noioso

obbligo burocratico a quello di chi teme impietose sanzioni (9). L'atteggiamento sanzionatorio non è ancora entrato nella procedura sanitaria italiana e, forse giustamente, non entrerà. La possibilità di aderire a programmi di VEQ che realmente facciano crescere e migliorare la qualità per il paziente e non soddisfino solamente la verifica di metodi, test e preferenze commerciali forse potrà contribuire a ridurre la sensazione "burocratica" dell'esecuzione di PT.

L'esperienza della verifica delle prove eseguite secondo lo standard NCCLS GP27-A è stata parte integrante di un processo complessivo di rivalutazione di questi esercizi. Eseguire un'istruttoria accurata sulle eventuali cause di un risultato non accettabile è stato motivo di valutazione della qualità, ma anche di discussione tra tutto il personale coinvolto nella produzione dei risultati di laboratorio sul ruolo che questi controlli hanno. Questa procedura è divenuta credibile perché integrata in un programma globale, che non escludeva alcun ambito e non rendeva singole prove prive di un qualche significato qualitativo.

Le aziende e gli organismi che si occupano di VEQ potrebbero considerare la necessità di programmi globali ed organizzati per la valutazione non solo dei risultati di laboratorio, ma anche dell'informazione diagnostica prodotta.

Bibliografia

1. Jenny RW, Jackson-Tarentino KY. Causes of unsatisfactory performance in proficiency testing. *Clin Chem* 2000; 46:89-99.
2. Miller WG, Ross JW. Pushing proficiency testing results to their limits. *CAP Today* 1999; 13:93-5.
3. Shahangian S. Proficiency testing in laboratory medicine: used and limitations. *Arch Pathol Lab Med* 1998; 112:15-30.
4. Tiets NW, Rodgerson DO, Laessig RH. Are clinical Laboratory proficiency tests as good as they can be? *Clin Chem* 1992; 38:473-5
5. Ross Jw, Myers GL, Gilmore BF, Cooper GR, Naito HR, Eckfeldt J. Matrix effects and the accuracy of cholesterol analysis. *Arch Pathol Lab Med* 1993; 117:393-400.
6. US Dept of Health and Human Services. Medicare, Medicaid and CLIA Programs: Regulations implementing the Clinical Laboratory Improvement Amendments of 1988 (CLIA). Final Rule. *Federal Register* 1992; 57:7151-62.
7. Using proficiency testing (PT) to improve the clinical laboratory. Approved guideline. P27-A. Villanova, PA: NCCLS, 1998.
8. Cembrowski GS, Carey RN. Adding value to proficiency testing programs. *Clin Chem* 2000; 46:7-8.
9. Pradella M. Come si partecipa ad un programma di Valutazione Esterna di qualità (VEQ) o ad una Prova di Abilità (proficiency test, PT): standard NCCLS GP27-A. *Med Lab* 1999; 7:203-6.