

Zygmunt Szopa, Rajmund Dybczyński

Porównania międzylaboratoryjne

**JAKO BAZA CERTYFIKACJI KANDYDATÓW
NA MATERIAŁY ODNIESIENIA**

Artykuł przedstawia przykłady kryteriów certyfikacji, niektóre możliwości oprogramowania komputerowego, służącego do statystycznej obróbki wyników porównań międzylaboratoryjnych (ILC), oraz ogólny schemat procedury prowadzącej do klasyfikacji pierwiastków do grup mających status wartości „polecanych” i „informacyjnych”.

Oceny większości cech stosowanej metody analitycznej, takich jak precyzja, wykrywalność czy selektywność, może samodzielnie dokonać prawie każde laboratorium analityczne. Wniosek powyższy nie dotyczy oceny dokładności metody i tylko nieliczne laboratoria i instytucje, dysponujące kilkoma niezależnymi metodami analitycznymi i „metodami definitywnymi”, są w stanie ocenić dokładność stosowanej metody we własnym zakresie. Dla pozostałych laboratoriów najpowszechniejszym sposobem oceny dokładności metody jest „równoległe” analizowanie odpowiedniego certyfikowanego materiału odniesienia (CRM) czy materiału odniesienia (RM) i uczestnictwo w porównaniach międzylaboratoryjnych (*interlaboratory comparison* – ILC) lub w ich szczególnym typie, tj. w testach biegłości (*proficiency testing* – PT). Zarówno wykorzystywanie CRM/RM-ów, jak też uczestnictwo w ILC/PT odgrywa ważną rolę nie tylko dla oceny wiarygodności nowo opracowanych metod analitycznych, potwierdzaniu umiejętności nowych laboratoriów, kontroli jakości wykonywanych analiz, kalibracji aparatury i technik analitycznych, ale jest w coraz większym stopniu wymagane do uzyskania, a następnie do utrzy-

mania akredytacji danego laboratorium.

Wśród laboratoriów i instytucji, które organizują ILC w rozmaitych celach, np. oceny biegłości laboratoriów, walidacji metod, technik i aparatury analityczno-pomiarowej, oceny stopnia tzw. nawiązywalności

(spójności) pomiarowej (*traceability*) itd., istnieje grono takich, które wykorzystują wyniki ILC jako bazę certyfikacji kandydatów na materiały odniesienia. Są to jednocześnie producenci CRM/RM-ów, dla których ILC stanowi ważny – lecz nie jedyny – etap w przyjętej strategii postępowania przy certyfikacji kandydata na materiał odniesienia.

Do takich producentów CRM/RM należą między innymi organizacje międzynarodowe, takie jak Międzyna-

rodowa Agencja Energii Atomowej (IAEA, Wiedeń), Biuro Wspólnoty Europejskiej (BCR, Bruksela) i niektóre krajowe laboratoria i organizacje. Organizowane przez nich porównania międzylaboratoryjne (ILC) są – w znakomitej większości – albo otwarte dla wszystkich chętnych, albo takie, w których uczestniczą jedynie laboratoria zaproszone przez organizatora, albo też będące kombinacją obydwu wymienionych typów, tj. otwarte z dodatkowo zaproszonymi „eksperckimi” laboratoriami.

W odróżnieniu od procedury stosowanej przez znanego producenta CRM/RM-ów NIST – Gaithersburg, USA, który wykorzystuje przede wszystkim własne techniki analityczne i pomiary, przez to może osiągnąć względnie szybki w realizacji schemat certyfikacji, stosowane przez innych producentów procedury przygotowania, a następnie certyfikacji kandydata na CRM/RM, są bardziej rozbudowane, a ich realizacja trwa dłużej. Z drugiej strony – po zakończonym porównaniu międzylaboratoryjnym (ILC) dysponują oni bez porównania większymi bazami danych. Właściwa obróbka i interpretacja tych danych prowadzi przede wszystkim do opartej na silnych podstawach statystycznych certyfikacji kandydata na CRM/RM, lecz może też służyć dalszej analizie wyników ILC (jak na przykład udział poszczególnych technik analitycznych i pomiarowych i ich wpływ na wartość atestowaną). Opracowania wyników ILC dokonuje się, zgodnie z wcześniej ustalonymi kryteriami certyfikacji, za pomocą odpowiedniego oprogramowania komputerowego, według określonej procedury.

Na przykładach niektórych testowych porównań międzylaboratoryjnych (ILC) pokażemy, jak wykorzystać można informacje gromadzone przez organizatora ILC do atestacji kandydatów na CRM/RM-y.

KRYTERIA CERTYFIKACJI

Jednym z celów ILC organizowanego dla potrzeb atestacji kandydata na CRM/RM jest zapewnienie jak największej liczby uczestników, którzy dysponują rozmaitymi metodami i technikami analitycznymi, aby późniejsza certyfikacja oparta była na solidnej podstawie statystycznej oraz aby można było atestować materiał dla jak największej ilości pierwiastków (analitów). Organizatorowi ILC zależy także na tym, aby duży w nim udział miały doświadczone laboratoria, o tzw. dobrej



Zestaw polskich CRM-ów
wyprodukowanych
i certyfikowanych przez IChI TS

reputacji, chociaż spełnienie jedynie tego warunku nie zapewnia jeszcze automatycznie powodzenia całego przedsięwzięcia. Dobrowolny uczestnik ILC oprócz możliwości sprawdzenia własnych wyników na tle wyników wszystkich pozostałych uczestników (czyli takiej samej korzyści, jaką daje uczestnictwo w PT), dodatkowo nieodpłatnie otrzymuje przyszły CRM/RM.

Organizator ILC, którego wyniki stanowią podstawę atestacji kandydata na CRM/RM, w zależności od posiadanych narzędzi (w tym odpowiedniego oprogramowania komputerowego), umiejętności (również organizacyjnych) oraz niezbędnego doświadczenia, ustala tzw. zespoły kryteriów certyfikacji. Regułą być powinno, aby organizator ustalał kryteria atestacji, zanim znane są wyniki ILC.

Najogólniej, określają one wymagania, które spełniać powinien zbiór wyników pochodzących z ILC dla danego pierwiastka (analitu), aby na jego podstawie nadać zawartości tego pierwiastka (analitu) status wartości „polecanej” lub „informacyjnej”. Zbiory, które nie spełniają ani kryterium przyjętego dla wartości „polecanej”, ani dla „informacyjnej”, nie mogą być podstawą nie mogą być podstawą do określania wielkości charakteryzujących zawartości analitu w danym materiale. Kryteria stanowią kompromis pomiędzy naturalnym dążeniem organizatora ILC do atestacji jak największej ilości pierwiastków (analitów) a samą jakością atestowanych zawartości. Oczywiście jest, że im zastosowane zostaną ostrzejsze kryteria, tym zawartości mniejszej ilości pierwiastków (analitów) można atestować w danym CRM/RM-ie, lecz zwiększa się za to gwarancja dobrej jakości atestacji. Można zauważyć, że podobne zespoły kryteriów w miarę upływu czasu ulegają pewnym modyfikacjom i uściśleniom w ramach każdej z organizacji – producentów CRM/RM. Nie ma więc jakiegos zespołu kryteriów obowiązujących w skali globalnej.

W tabeli przedstawiono przykład zespołu kryteriów, które stosowano przy atestacji CRM-ów – „Human Hair Reference Materials” – IAEA-085 i IAEA-086. Widać, że sformułowano je w ścisłym powiązaniu z wynikami działania specjalnie w tym celu opracowanego pakietu oprogramowania AQCS-PC oraz że jest to zespół raczej ostrych kryteriów narzuconych przez samego organizatora ILC. I tak na przykład wymóg, że średnia ogólna obliczana jest na bazie co najmniej dziesięciu akceptowanych przez pakiet średnich arytmetycznych przy jednoczesnym spełnieniu warunku, że analizy przeprowadzano więcej niż tylko jedną metodą analityczną, że odrzuconych wyników jest nie więcej niż jedna piąta pierwotnej populacji, że zbiór, który stanowi podstawę atestacji, jest spójny itd. stanowią dobrą gwarancję jakości atestowanej wartości „polecanej”. Zespół kryteriów dla wartości „informacyjnych” jest znacznie łagodniejszy, ale i tutaj (podobnie jak dla wartości „polecanych”) wymaga się, aby atestowana wartość opierała się na względnie silnych podstawach, tj. aby składały się na nią wyniki pochodzące od przy-

najmniej dwóch niezależnych metod analitycznych i co więcej, aby te wyniki nie różniły się istotnie.

Ponieważ rozmaitych (mniej lub bardziej ostrych) kryteriów certyfikacji jest na ogół tyle, ilu jest producentów CRM/RM-ów, ważne jest, aby były one sformułowane jasno i jednoznacznie nie tylko dla potrzeb dokumentacyjnych organizatora ILC, ale aby znane były wszystkim uczestnikom ILC, chociażby z raportu, który otrzymują po zakończeniu kampanii certyfikacyjnej.

STATYSTYCZNA OBRÓBKA WYNIKÓW PORÓWNAŃ MIĘDZYLABORATORYJNYCH

Na podstawie przeprowadzonego ILC, w którym uczestniczą dziesiątki, a czasami nawet setki laboratoriów, organizator buduje odpowiednie bazy danych. Bazy te zawierają nie tylko same wyniki analiz i związane z nimi niepewności, ale również informacje dotyczące laboratoriów uczestniczących (kody laboratoriów), wykorzystywanych metod analitycznych, osiąganych granic oznaczalności, mas analizowanych próbek itp. Wszystkie te informacje gromadzone przez organizatora służą z jednej strony późniejszej wieloaspektowej analizie pochodzących z ILC wyników, z drugiej zaś pełnej dokumentacji pozwalającej prześledzić historię każdego pojedynczego wyniku. Obróbkę takich baz dokonuje się obecnie za pomocą odpowiednich pakietów oprogramowania komputerowego.

Promochem

**WZORCE ANALITYCZNE
INDYWIDUALNE I ZBIORY**

**ROZTWORY PRACOWE I STANDARDY
INDYWIDUALNE I ZBIORY**

**MATERIAŁY DO CODZIENNEJ
KONTROLI JAKOŚCI ANALIZ**

**CERTYFIKOWANE MATERIAŁY
ODNIESIENIA**

Promochem Sp. z o.o.
ul. Annopol 6
03-236 Warszawa
tel. (022) 811-23-99, 814-08-37, fax (022) 679-90-80
<http://www.promochem.com.pl> e-mail: sales@promochem.com.pl

wego, wykorzystujących różne metody statystycznego opracowywania wyników.

PAKIET OPROGRAMOWANIA AQCS-PC

Ten wielofunkcyjny pakiet oprogramowania został zbudowany na bazie podejścia do statystycznej obróbki

Lp.	Kryterium	Wartości „polecane” [recomm. values]	Wartości „informacyjne” [inform values]
1.	Średnia ogólna (arytmetyczna/ważona) obliczana na podst.	$\geq$ akceptowanych laboratoriów	$\geq$ akceptowanych laboratoriów
2	Wielej niż jedna metoda analityczna	Tak	Tak
3	Procent odrzuconych wyników	≤ 20	≤ 30
4	Względna niepewność średniej ogólnej [%]	≤ 20 (dla zawart. ≤ 500 ppm) ≤ 10 (dla zawart. ≥ 500 ppm)	≤ 30 (dla zawart. ≤ 500 ppm) ≤ 20 (dla zawart. ≥ 500 ppm)
5	Sprawdzanie, czy różne metody analityczne różnią się istotnie	Tak	Tak

Przykład zespołu kryteriów
certyfikacji kandydata na CRM/RM
wykorzystywany w IAEA

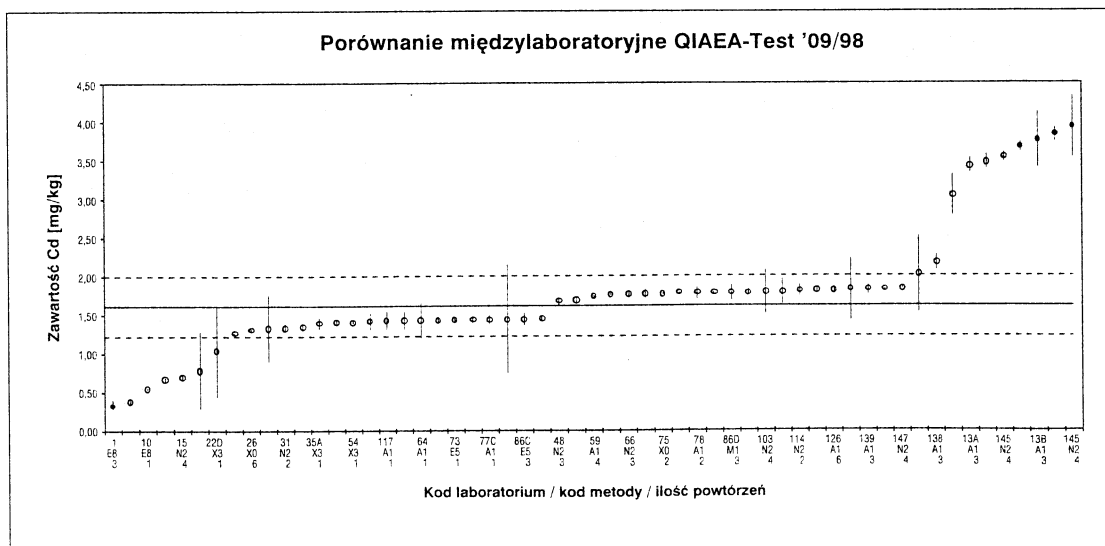
wyników ILC zaproponowanej i wykorzystywanej w IAEA. Jest on obecnie jedynym znanym w literaturze oprogramowaniem opracowanym specjalnie w celu

wych funkcji i opcji z jednej strony uwzględniających nasze doświadczenia zdobywane w trakcie certyfikacji polskich CRM/RM-ów, z drugiej zaś, aby spełnić aktualne wymogi stawiane przez IAEA przed takim oprogramowaniem. Jest on głównie adresowany do grup zajmujących się certyfikacją CRM/RM-ów na bazie organizowanych ILC oraz do organizatorów PT. Jego konstrukcja pozwala na wykorzystywanie go nie tylko na etapie czysto statystycznej obróbki wyników pochodzących z ILC/PT, ale również w procesie dalszej certyfikacji kandydata na CRM/RM. Umożliwia szybką redukcję i interpretację danych oraz współpracuje z powszechnie stosowanymi programami, takimi jak: Excel, Quattro Pro, dBASE, Access, Statgraphics itp., stanowiącymi dla pakietu AQCS-PC oprogramowanie towarzyszące. Poniżej podajemy niektóre przykłady obróbki i graficznej prezentacji wyników ILC otrzymane z pomocą najczęściej wykorzystywanych funkcji i opcji pakietu i oprogramowania towarzyszącego.

Uporządkowane rosnąco średnie laboratoryjne dla danego pierwiastka (analitu) formują tzw. diagram typu Z, odwrotnie zaś, tj. w malejącym porządku – typu S (odpowiednio ang. Z – *shaped plots* lub S – *shaped plots*). Diagramy takie zamieszczane są zwykle w raportach wydawanych po zakończeniu kampanii certyfikacyjnej. Pozwalają uczestnikom ILC między innymi na sprawdzenie własnego wyniku na tle wszystkich pozostałych. Na rysunku 1 przedstawiono taki diagram dla rozkładu wyników analitycznych dla analitu – Cd w testowym ILC. Wzdłuż osi odciętych zaznaczono kody laboratoriów, kody stosowanych przez laboratoria metod analitycznych oraz liczbę niezależnych oznaczeń, które wykonało dane laboratorium. Średnie laboratoryjne przedstawiono w postaci kółek z pionowymi liniami obrazującymi wartości niepewności dla tych średnich. Wyniki wątpliwe, odrzucone przez pakiet z pier-

wotnej populacji, zaznaczono kółkami zaczernionymi. Wartość „polecana” wraz z jej poziomem ufności przedstawiają odpowiednio pozioma linia ciągła i linie przerywane. Widać, że nie jest to typowy „gładki” rozkład typu Z, ale że jest on dwumodalny oraz że obydwa – blisko położone – maksima leżą wewnątrz przedziałów ufności.

Inna, często wykorzystywana opcja pakietu daje możliwość otrzymywania diagramów typu skrzynkowego lub typu skrzynkowego z wcięciami (od-



Rys. 1. Przykład rozkładu wyników
porównania międzylaboratoryjnego
w postaci diagramu typu Z

statystycznej obróbki danych pochodzących z ILC/PT, a wspomagającym podejmowanie decyzji w trakcie certyfikacji. Przez szereg lat sukcesywnie wprowadzano do niego szereg modyfikacji, rozszerzeń, dodatko-

powiednio ang. *box-and-whisker* lub *notched box-and-whisker plots*). Ogólnie, diagram skrzynkowy jest graficzną reprezentacją rozkładu pewnego zbioru wyników. W trakcie analizowania danych pochodzących

z ILC konstruuje się najczęściej diagramy typu skrzynkowego z wcięciami dla podgrup wyników pochodzących z różnych metod analitycznych dla danego pierwiastka (analitu) i umieszcza na jednym rysunku w celu ułatwienia porównań.

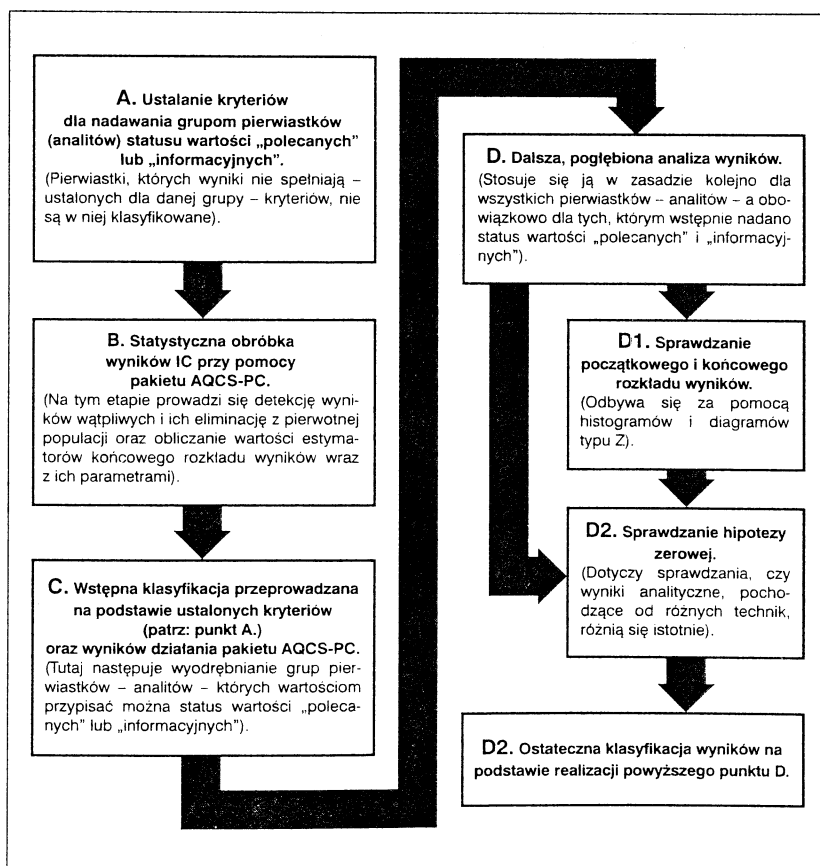
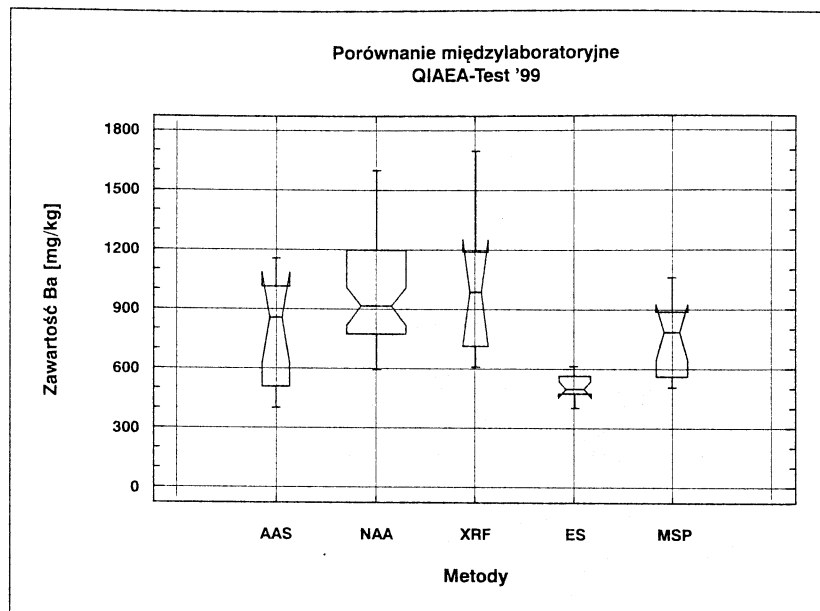
Rysunek 2 przedstawia przykład, gdzie takie diagramy skonstruowano dla zawartości Ba w testowym ILC. Podzbiory wyników dla różnych metod analitycznych (AAS, NAA, XRF... itd. – rozmieszczone wzdłuż osi odciętych) reprezentują odpowiednie diagramy. Ich szerokość jest funkcją liczności populacji, długość zaś przedstawia zakres wyników dla danej podgrupy. Oś rzędnych obrazuje zawartość Ba w analizowanym materiale. Środkowa linia pozioma w diagramie skrzynkowym przedstawia wartość mediany. Skośnie wyprowadzone od mediany linie (tzw. wcięcia) obrazują przedziały ufności. W taki sposób przedstawione diagramy pozwalają na szybkie porównywanie średnich, zakresów, kwartyli itd. poszczególnych podgrup wyników. Natomiast sprawdzanie, czy wyniki analityczne, pochodzące od różnych metod, są istotnie różne czy też nie (patrz: tabela, kryterium 5), odbywa się przez porównanie przedziałów ufności. W naszym przykładzie (rys. 2) widać, że metoda kodowana jako NAA była w tym ILC reprezentowana liczniej niż na przykład metoda XRF. Widać też, że wyniki pochodzące od metod kodowanych jako AAS, NAA i XRF nie różnią się istotnie, co nie jest prawdziwe, jeśli porównuje się na przykład metodę XRF z ES czy NAA z ES.

SCHEMAT OBRÓBK I INTERPRETACJI WYNIKÓW ILC PRZY CERTYFIKACJI KANDYDATA NA CRM/RM

Zespół działań związanych z obróbką i interpretacją wyników ILC przedstawić można w formie 5-etapowego schematu (rys. 3). Opracowano go i sprawdzono w całości w trakcie atestacji dwóch CRM/RM-ów – IAEA-085 i IAEA-086. Niektóre elementy tego schematu stosowano już wcześniej zarówno przy certyfikacji polskich, jak i zagranicznych materiałów odniesienia. Pierwsze dwa etapy obejmują ustalanie kryteriów certyfikacji oraz czysto statystyczną obróbkę wyników ILC za pomocą wielofunkcyjnego pakietu oprogramowania AQCS-PC. W kolejnym etapie (rys. 3, etap C) następuje wstępne wyodrębnienie grup pierwiastków (analitów), którym, według przyjętych kryteriów, przypisać można status wartości „polecanych” i „informacyjnych”. Następny etap zawiera pogłębioną analizę wyników ILC. Na tym etapie – opcjonalnie – sprawdza się charakter początkowego oraz końcowego (tj. po odrzuceniu wyników wątpliwych) rozkładu wyników ILC. Natomiast obligatoryjne na tym etapie jest sprawdzenie hipotezy zerowej – tj. czy wyniki pochodzące od różnych technik analitycznych różnią się istotnie. Odrzucenie tej hipotezy jest warunkiem spełnienia piątego (ostatniego) kryterium certyfikacji w naszym przykładzie (tab.). Wtedy następuje ostateczna klasyfikacja pierwiastków (analitów), które spełniają odpowiednie

kryteria certyfikacji. Jak widać z przedstawionego schematu, procedura obróbki i interpretacji danych pochodzących z ILC w powiązaniu z rozsądnie ustalonymi kryteriami certyfikacji jest niczym innym jak zapla-

Rys. 2. Przykład diagramów typu skrzynkowego z wcięciami



Rys. 3. Schemat obróbki wyników pochodzących z porównania międzylaboratoryjnego

nowanym i logicznym sposobem „filtrowania jakości” danych zgromadzonych w trakcie ILC. Ogólna procedura zaprezentowana w niniejszej pracy mogłaby być wykorzystana w innych, podobnych porównaniach międzylaboratoryjnych lub testach biegłości.