



**Dąbrowskie
Wodociągi**

Dąbrowskie Wodociągi Sp. z o. o.
ul. Powstańców 13
41-300 Dąbrowa Górnicza

www.dabrowskie-wodociagi.pl

PRZEWODNIK PO PROGRAMIE BADANIA BIEGŁOŚCI WASTER

Analizy Fizykochemiczne Ścieków

Wydanie 06 (zastępuje wyd.05 z dnia 09.02.2017)
Data wydania 25.04.2018

**Opracował: Koordynator Programu
mgr inż. Katarzyna Skorek**



1. Wprowadzenie

Ten przewodnik przedstawia zagadnienia dotyczące organizacji i sposobu funkcjonowania Programu badania biegłości WASTER oraz opisuje bilateralny Program badania biegłości WASTER-QR.

Przewodnik jest adresowany do laboratoriów analitycznych uczestniczących w Programie WASTER lub WASTER-QR, klientów tych laboratoriów, organów stanowiących lub udzielających akredytacji oraz innych stron zainteresowanych badaniami biegłości.

2. Tożsamość i kompetencje organizatora Programu

Organizatorem (dostawcą) Programu są Dąbrowskie Wodociągi Sp. z o.o., 41-300 Dąbrowa Górnicza, ul. Powstańców 13.

Organizator posiada w swojej strukturze laboratorium analityczne, akredytowane przez PCA (certyfikat nr AB 709). W zakresie akredytacji mieszczą się oznaczenia wszystkich analizów objętych Programem. Organizator nie korzysta z podwykonawstwa, za wyjątkiem firmy kurierskiej, która dostarcza próbki PT do uczestników.

Metody i wyposażenie stosowane do oceny homogeniczności i stabilności próbek PT są zwalidowane i nadzorowane.

Organizator posiada dostęp do specjalistycznej wiedzy analitycznej i statystycznej poprzez współpracę Koordynatora Programu z Zespołem Doradczym ds. Analitycznych i Statystycznych, który tworzą:

dr inż. Zbigniew Żurek – Biuro Systemów Jakości CERTIQ

dr David Coolens – National Measurement Institute (Holandia)

Koordynatorem Programu jest mgr inż. Katarzyna Skorek, Kierownik Laboratorium Analitycznego Dąbrowskich Wodociągów (tel. 32 63 95 120, e-mail laboratorium@dabrowskie-wodociagi.pl).

Koordynator jest odpowiedzialny za wszystkie działania związane z planowaniem, realizacją i dokumentowaniem Programu oraz za komunikację z uczestnikami.



3. Projekt Programu

Program WASTER jest zaprojektowany tak, aby dostarczał laboratoriom więcej niż tylko ocenę biegłości. W Programie oceniane są również inne aspekty jakościowe wyników raportowanych przez laboratoria, a mianowicie: błąd całkowity, błąd systematyczny, błąd przypadkowy oraz niepewność pomiaru. Taka ocena daje wszechstronny obraz możliwości analitycznych każdego laboratorium. Program WASTER stosuje projekt dwupróbkowy, oparty na koncepcji Youdena [1, 2]. Projekt polega na tym, że:

- 1) Laboratorium bada dwie próbki (A i B), które tworzą tzw. parę Youdena. Próbki A i B są nominalnie identyczne (para zgodna) bądź bardzo zbliżone pod względem stężenia analitów (para rozszczepiona). Laboratorium wykonuje – w tych samych warunkach – pojedynczą analizę każdej próbki i otrzymuje parę wyników (X_A , X_B).
- 2) Przy pomocy techniki graficznej Youdena, z błędu całkowitego, którym obarczona jest para wyników (X_A , X_B) wyodrębniany jest błąd systematyczny i błąd przypadkowy. Na podstawie wielkości tych dwóch błędów, diagnozowany jest błąd, który wnosi dominujący wkład do błędu całkowitego i tym samym jest za ten błąd odpowiedzialny.

4. Cele Programu

Podstawowym celem Programu jest ocena zdolności laboratoriów do kompetentnego wykonywania oznaczeń analitów w ściekach.

Inne cele Programu są następujące:

- 1) sprawdzić dokładność (poprawność i precyzję) wyników analiz wykonanych przez laboratoria
- 2) zdiagnozować przy pomocy diagramu Youdena błędy analityczne, którymi obarczone są wyniki laboratoriów, a tym samym ułatwić laboratoriom identyfikację źródłowych przyczyn tych błędów
- 3) dać laboratoriom możliwość porównania własnych wyników z wynikami innych podobnych laboratoriów
- 4) zweryfikować realistyczność niepewności pomiaru raportowanej przez laboratoria
- 5) umożliwić laboratoriom porównanie biegłości dwóch analityków lub funkcjonowania dwóch metod badawczych
- 6) pomóc laboratoriom w wykorzystywaniu danych – z kilku rund Programu – do ulepszenia funkcjonowania metod badawczych i doskonalenia jakości wyników



5. Zgodność Programu z wymaganiami norm i innych dokumentów

Program jest realizowany według wymagań normy PN-EN ISO/IEC 17043:2011 „Ocena zgodności. Ogólne wymagania dotyczące badania biegłości”.

W obszarze zagadnień statystycznych Program spełnia zalecenia dokumentów:

- IUPAC Technical Report: 2006 “International harmonized protocol for the proficiency testing of analytical chemistry laboratories” [3]
- ISO 13528:2015 „Statistical methods for use in proficiency testing by interlaboratory comparison” [4]
- Eurachem Guide “Selection, use and interpretation of proficiency testing (PT) schemes by laboratories: 2011 [5]

6. Zakres analityczny Programu

Obiektem badania biegłości jest ściek surowy z naturalną matrycą, pobrany z miejsca jego rzeczywistego występowania. Program obejmuje oznaczanie następujących analitów: siarczany, zawiesiny ogólne, azot amonowy, ChZT_{Cr} , BZT_{Cr} , fosfor ogólny, chlorki, azot Kjeldahla, azot ogólny.

Przewidywane zakresy stężeń poszczególnych analitów w próbkach PT są podane na stronie organizatora www.dabrowskie-wodociagi.pl. Sporadycznie, rzeczywiste stężenia analitów mogą nieznacznie różnić się od zakresów przewidywanych.

7. Ramy czasowe Programu

Program jest realizowany w sposób ciągły z częstotliwością dwie rundy w każdym roku (marzec, wrzesień). Harmonogram Programu na dany rok obejmuje dla każdej rundy daty: zgłoszenia do Programu, dystrybucji próbek, rozpoczęcia analiz, przekazania wyników analiz do Koordynatora, dostarczenia do uczestników wstępnego protokołu oceny biegłości oraz raportu z badania biegłości i jest publikowany na stronie organizatora w grudniu roku poprzedniego.

8. Uczestnictwo w Programie

Dostęp do Programu jest otwarty dla każdego laboratorium (bez względu na status akredytacji), które w swojej praktyce wykonuje analizy fizykochemiczne ścieków i zgłosi formalnie udział w Programie (poprzez przekazanie wypełnionego formalnie zgłoszenia do Koordynatora Programu). Koszty udziału w Programie są publikowane na stronie organizatora.



9. Liczba uczestników w rundzie Programu

Minimalna w rundzie Programu liczba par wyników dla każdego badanego analitu, niezbędna do przeprowadzenia oceny biegłości oraz do wykonania analizy graficznej Youdena wynosi $N = 5$. Gdy dla określonego analitu $N < 5$, analit nie bierze udziału w rundzie. Runda Programu jest anulowana w całości, wyłącznie w przypadku, gdy dla każdego badanego analitu liczba par wyników wynosi $N < 5$.

10. Próbki do badania biegłości

10.1. Przygotowanie próbek

Pary próbek (A, B) są przygotowywane przez laboratorium analityczne organizatora. Gdy próbki w parze mają mieć nieznacznie różny poziom stężenia badanych analitów (pary rozszczepione), duża objętość ścieku jest rozdzielana na dwie części do dwóch naczyń (A i B) i ściek w naczyniu B jest nieznacznie rozcieńczany wodą dejonizowaną. Z obydwu naczyń – po starannym ujednorodnieniu ścieku – pobierane są próbki o objętości 2l każda, do opakowań jednostkowych (butelki), oznakowanych odpowiednio A i B, po czym butelki są szczelnie zamykane.

Przygotowywane są pary próbek: przeznaczone dla uczestników, pary rezerwowe oraz pary do badania homogeniczności i stabilności. Szczegóły dotyczące przygotowania próbek są przedstawiane każdorazowo w raporcie z badania biegłości.

10.2. Ocena homogeniczności próbek

Aby upewnić się, że wszyscy uczestnicy otrzymują takie same – pod względem stężenia analitów – pary próbek (A i B), laboratorium analityczne organizatora wykonuje badanie homogeniczności, dla dwóch lub trzech analitów, które w każdej rundzie określa Koordynator Programu.

Badanie homogeniczności jest wykonywane wg procedury IUPAC [3]. Liczba próbek przeznaczonych do badania homogeniczności zależy od wielkości partii próbek przygotowanych i zwykle wynosi $m = 3 \div 5$.

Procedura IUPAC jest wykonywana odrębnie dla próbek A oraz próbek B – w przypadku par rozszczepionych lub tylko dla próbek A – w przypadku par zgodnych i obejmuje:

- wykonanie dublowanych analiz każdej z (m) próbek, w warunkach powtarzalności (analizy są wykonywane na dwóch niezależnych porcjach próbki)
- wyznaczenie różnic (d_i) pomiędzy wynikami dublowanych analiz



- test Cochran dla maksymalnej skwadratowanej różnicy (d_i^2)
- obliczenie wariancji analitycznej wewnątrzpróbkowej, (S_{an}^2) według formuły:

$$S_{an}^2 = \left(\sum_{i=1}^m d_i^2 \right) / 2m$$

- obliczenie wariancji międzypróbkowej „kontaminowanej” ($S_{sam(k)}^2$) według formuły:

$$S_{sam(k)}^2 = \left(\sum_{i=1}^m (\bar{X}_i - \bar{X})^2 \right) / (m-1)$$

gdzie: \bar{X}_i jest średnią dublowanych analiz próbki (i)

\bar{X} jest średnią ze średnich \bar{X}_i

- obliczenie wariancji międzypróbkowej „czystej” (S_{sam}^2)

$$S_{sam}^2 = S_{sam(k)}^2 - \frac{1}{2} S_{an}^2$$

- obliczenie wariancji międzypróbkowej dopuszczalnej (σ_{all}^2)

$$\sigma_{all}^2 = (0,3 \sigma_p)^2$$

gdzie: σ_p jest odchyleniem standardowym do oceny biegłości

Ponieważ zgodnie z protokołem statystycznym stosowanym w Programie (σ_p) jest wyznaczane z wyników uczestników, zatem nie jest znane a priori. Dlatego (σ_p) jest obliczane jako

$$\sigma_p = \bar{X} \cdot CV_{rob}$$

gdzie: $\%CV_{rob}$ jest wyznaczane z zależności regresyjnej $\%CV_{rob}$ ver med, zamodelowanej na podstawie danych z poprzednich rund.

- obliczenie wartości krytycznej (C) do testu Fearnal („wystarczającej” homogeniczności) według formuły:

$$C = F_1 \sigma_{all}^2 + F_2 S_{an}^2$$

gdzie: F_1, F_2 są współczynnikami zależnymi od liczby próbek (m) wyznaczonymi z [3]

- porównanie S_{sam}^2 z wartością krytyczną (C)

jeśli $S_{sam}^2 \leq C$, to próbki są „wystarczająco” homogeniczne do oceny biegłości

jeśli $S_{sam}^2 > C$, to homogeniczność nie została potwierdzona i wówczas wariancja międzypróbkowa „czysta” (S_{sam}^2) jest uwzględniona w ocenie biegłości



10.3. Badanie stabilności próbek

Badanie wykonuje laboratorium analityczne organizatora, aby ocenić czy stężenie analitów w próbkach przekazanych uczestnikom, nie zmienia się podczas ich przechowywania przez czas: od przygotowania próbek (czas t_0) do rozpoczęcia badań przez uczestników (czas t_k). Badanie jest wykonywane według procedury IUPAC [3] tylko na próbkach A (niewielkie rozcieńczenie próbki B nie ma wpływu na stabilność) i dla tych samych analitów, dla których badano homogeniczność. Średnia analiz wykonanych w ramach badania homogeniczności w czasie t_0 (\bar{X}_{t_0}) jest porównana (testem Studenta, na poziomie ufności 95%) ze średnią analiz wykonanych w ten sam sposób (co podczas badania homogeniczności) w czasie t_k (\bar{X}_{t_k}).

Jeżeli różnica między średnimi jest statystycznie nieistotna, przyjmuje się, że próbki są „wystarczająco” stabilne do oceny biegłości. W przeciwnym wypadku odchylenie standardowe różnicy pomiędzy tymi średnimi jest uwzględniane w ocenie biegłości. W niektórych przypadkach stosowane jest zgodnie z [4] kryterium stabilności, oparte na rozrzucie wyników w rundzie. Według tego kryterium stabilność próbek jest potwierdzona, gdy:

$$|\bar{X}_{t_0} - \bar{X}_{t_k}| \leq 0,3 \cdot \min[\text{NIQR} (A), \text{NIQR} (B)]$$

Statystyki NIQR(A), NIQR(B) są wyjaśnione w dalszej części Przewodnika.

10.4. Dystrybucja próbek do uczestników

Dwie próbki, stanowiące parę, oznakowane literą A i B i zapakowane w torbę termoizolacyjną z wkładami chłodzącymi, są dostarczane do uczestników za pośrednictwem firmy kurierskiej. Do czasu dystrybucji, próbki są przechowywane w warunkach nadzorowanych (temperatura 2-4°C, ciemność). Próbki dostarczane są gotowe do analizy. Razem z próbkami dostarczana jest Instrukcja dla Uczestników oraz Karta Wyników.

11. Wykonywanie analiz przez uczestników

Uczestnicy mają obowiązek wykonania analiz metodami, które rutynowo stosują do badania próbek klientów. Wybór metody zależy od laboratorium. Aby ograniczyć rozrzut wyników wynikający ze stosowania przez laboratoria kilku różnych metod do oznaczania tego samego analitu, organizator zaleca stosowanie następujących metod znormalizowanych:



Analit	Zalecana metoda badawcza
Siarczany	PN-ISO 9280:2002
Zawiesiny ogólne	PN-EN 872:2007
Azot amonowy	PN-ISO 5664:2002
ChZT _{Cr}	PN ISO 15705:2005
BZT ₅	PN-EN 1899-1:2002
Fosfor ogólny	PN-EN ISO 6878:2006
Chlorki	PN-ISO 9297:1994
Azot Kjeldahla	PN-EN 25663:2001
Azot ogólny	PN-73/C-04576.14

Uczestnicy są proszeni o wykonanie (dla każdego badanego analitu) pojedynczej analizy próbki A i pojedynczej analizy próbki B. Obydwie analizy muszą być przeprowadzone przez tego samego analityka, w tym samym dniu, przy pomocy tego samego wyposażenia badawczego, z tą samą kalibracją jeśli ma zastosowanie.

12. Protokół statystyczny stosowany w Programie

12.1. Charakterystyka protokołu

Zgromadzone pary wyników (X_A , X_B) tworzą dwa zbiory: zbiór wyników próbki A i zbiór wyników próbki B. Do opisu zbiorów stosowane są: mediana i rozstęp międzykwartyłowy. Statystyki te są odporne na wartości ekstremalne (znacznie odległe od reszty wyników), które zazwyczaj są obecne w zbiorach wyników.

Mediana jest: wynikiem środkowym w uporządkowanym zbiorze wyników (gdy liczba wyników jest nieparzysta) lub średnią arytmetyczną dwóch środkowych wyników (gdy liczba wyników jest parzysta).

Rozstęp międzykwartyłowy (IQR) kwantyfikuje rozrzut 50% środkowych wyników w zbiorze i jest równy różnicy pomiędzy kwartyłem górnym (Q3) i dolnym (Q1): $IQR = Q3 - Q1$.

W statystyce kwartyle są wartościami, które dzielą uporządkowany zbiór wyników na cztery równe części.

Racjonalne stosowanie statystyk odpornych wymaga według [5, 6, 7, 8, 9], aby rozkład wyników był zgrubnie normalny i symetryczny. Wymaganie to jest spełnione, gdy wyniki ekstremalne są rozmieszczone symetrycznie po obydwu stronach średniej. Jednakże, gdy wyniki ekstremalne są rozmieszczone tylko po jednej stronie średniej, wówczas powodują skośność rozkładu (prawo lub



lewostronną). Skośność nie ma wpływu na medianę lecz wpływa na pozycje kwartyła Q1 lub Q3 – w zależności od kierunku skośności. Dlatego rozstęp międzykwartyłowy nie powinien być stosowany – jako miara rozrzutu – w rozkładzie wyników, który wykazuje statystycznie istotną skośność.

Należy dodać, że zgrubnie normalny rozkład wyników próbki A i próbki B jest wymagany także do:

- oszacowania odpornego odchylenia standardowego (σ_p) z rozstępu międzykwartyłowego
- wyznaczenia niepewności standardowej wartości przypisanej
- wyznaczenia elipsy ufności

Protokół statystyczny stosowany w Programie obejmuje:

- diagnostykę symetrii rozkładów
- wyznaczanie wartości przypisanej i odpornego odchylenia standardowego do oceny biegłości
- ocenę wartości przypisanej na podstawie jej niepewności standardowej
- wskaźnikowanie biegłości
- ocenę biegłości

12.2. Diagnostyka symetrii rozkładów wyników próbki A i B

Symetria rozkładu wyników próbki A i rozkładu wyników próbki B jest diagnozowana:

- graficznie przy pomocy dot plotów (analiza rozmieszczenia wyników wokół średniej, identyfikacja asymetrycznych danych ekstremalnych, które mogą przyczyniać się do skośności rozkładu, porównanie położenia mediany względem średniej)
- numerycznie przy pomocy testu D'Agostino [10], który ocenia odchylenie rozkładu wyników od rozkładu normalnego, spowodowane skośnością. Statystyką testową (g) jest stosunek skośności próbkowej (S_k) do błędu standardowego skośności $SE(S_k)$

$$g = \frac{S_k}{SE(S_k)}$$

gdzie:
$$S_k = \frac{N}{(N-1)(N-2)SD^3} \sum_{i=1}^N (X_i - \bar{X})^3$$

$$SE(S_k) = \sqrt{\frac{6N(N-1)}{(N-2)(N+1)(N+3)}}$$

N = liczba wyników w zbiorze ; SD = odchylenie standardowe wyników

X_i = i -ty wynik w zbiorze ; \bar{X} = średnia wyników w zbiorze



Skośność próbkowa (S_k) jest istotna statystycznie (na poziomie $\alpha = 0,05$), gdy prawdopodobieństwo (p) otrzymania statystyki testowej (g) – co najmniej tak dużej jak uzyskana z próby – jest mniejsze niż (α).

Minimalna liczba wyników wymagana dla testu D'Agostino wynosi $N = 8$. Gdy liczba wyników maleje poniżej osiem, to $SE(S_k)$ szybko rośnie, co sprawia, że na podstawie testu, skośność (S_k) jest nieistotna statystycznie, podczas gdy w rzeczywistości rozkład wyników odchyła się istotnie od rozkładu normalnego. Dlatego w przypadku gdy liczba wyników wynosi $5 \leq N < 8$ badana jest normalność rozkładu testem Shapiro – Wilka. Normalność rozkładu jest potwierdzona, gdy wartość obliczeniowa statystyki (W_{obl}), jest większa od wartości krytycznej (W_{kr}), wyznaczonej dla poziomu istotności $\alpha = 0,05$ i liczby wyników (N).

Gdy na podstawie testu D'Agostino, skośność jest istotna statystycznie ($p < 0,05$) lub na podstawie testu Shapiro – Wilka normalność rozkładu nie jest potwierdzona ($W_{obl} < W_{kr}$), wówczas wynik, który przyczynia się do skośności lub „zaburza” normalność rozkładu jest eliminowany, a test D'Agostino lub test Shapiro – Wilka jest powtarzany.

Wylimitowanie określonego wyniku z jednego zbioru, powoduje „automatyczne” wylimitowanie sparowanego z nim wyniku z drugiego zbioru, czyli odrzucenie całej pary wyników. Wyniki odrzucone nie biorą udziału w dalszym postępowaniu statystycznym, jednakże wyniki te są oceniane w Programie i otrzymują wskaźniki biegłości.

W nietypowych sytuacjach decyzję o odrzuceniu lub zachowaniu pary wyników podejmuje Zespół Doradczy.

12.3. Wyznaczenie wartości przypisanej

Wartością przypisaną jest mediana wyznaczona z ważnych wyników uczestników dla próbki A: [med (A)] i dla próbki B: [med (B)].

Liczba ważnych wyników (n_w) wynosi: $n_w = N - n$ gdzie: N = liczba wszystkich wyników;
 n = liczba wyników odrzuconych na etapie diagnozowania symetrii

Spójność pomiarowa wartości przypisanych z wzorcami jednostki miary nie jest zapewniona.

12.4. Wyznaczenie odpornego odchylenia standardowego (σ_p) do oceny biegłości

Odporne odchylenie standardowe (σ_p) dla zbioru wyników próbki A i zbioru wyników próbki B jest wyrażone przez znormalizowany rozstęp międzykwartyłowy (NIQR)

$$\sigma_p = \text{NIQR} \quad \text{NIQR} = 0,7431 \cdot \text{IQR}$$



W kategoriach względnych (σ_p) jest odpornym współczynnikiem zmienności ($\%CV_{rob}$), wyrażonym procentach i obliczonym wg zależności:

$$\%CV_{rob} = \frac{NIQR}{med} 100$$

($\%CV_{rob}$) służy do porównania względnej zmienności w zbiorach wyników próbki A i B.

12.5. Wyznaczenie niepewności standardowej wartości przypisanej dla wyników próbki A i B

Niepewność standardowa wartości przypisanej [$u(\text{med})$] jest obliczana według [11] jako błąd standardowy mediany $SE(\text{med})$, zgodnie z formułą:

$$u(\text{med}) = SE(\text{med}) = \sqrt{\frac{\pi}{2}} \frac{NIQR}{\sqrt{n_w}}$$

12.6. Ocena wiarygodności wartości przypisanej dla próbki A i próbki B

gdym: $u(\text{med}) \leq 0,4 NIQR$, to wartość przypisana jest wiarygodna i $u(\text{med.})$ nie ma wpływu na ocenę biegłości

gdym: $0,4 NIQR < u \leq 0,8 NIQR$, to $u(\text{med.})$ ma wpływ na ocenę biegłości i wówczas odporne odchylenie standardowe do oceny biegłości jest wyznaczane według formuły:

$$\sigma_p^* = \sqrt{NIQR^2 + u^2(\text{med})}$$

gdym: $u(\text{med.}) > 0,8 NIQR$, to wartość przypisana jest niepewna i ocena biegłości nie jest wykonywana

13. Wskaźnikowanie biegłości

Dla każdego wyniku X_A i każdego wyniku X_B , raportowanego przez laboratorium, obliczane są odporne wskaźniki biegłości Z_A i Z_B .

$$Z_A = \frac{X_A - \text{med}(A)}{NIQR(A)} \qquad Z_B = \frac{X_B - \text{med}(B)}{NIQR(B)}$$

gdzie: $\text{med}(A)$, $\text{med}(B)$ – wartości przypisane, wyrażone przez mediany wyników próbki A i B

$NIQR(A)$, $NIQR(B)$ – odporne odchylenia standardowe do oceny biegłości, wyrażone przez znormalizowane rozstępy międzykwartylowe wyników próbki A i B

Gdy niepewność wartości przypisanej ma wpływ na ocenę biegłości, wówczas stosowane są zmodyfikowane wskaźniki biegłości, wyznaczane według formuł:



$$Z'_A = \frac{X_A - \text{med}(A)}{\sqrt{\text{NIQR}^2(A) + u^2(\text{med}(A))}}$$

$$Z'_B = \frac{X_B - \text{med}(B)}{\sqrt{\text{NIQR}^2(B) + u^2(\text{med}(B))}}$$

14. Ocena biegłości

Biegłość laboratorium jest oceniana na podstawie dwóch wskaźników biegłości (Z_A , Z_B) według kryteriów:

- gdy laboratorium otrzymuje obydwa wskaźniki w przedziale $[-2 ; 2]$, wówczas biegłość laboratorium oceniana jest jako zadowalająca:

$$\{|Z_A| \text{ i } |Z_B|\} \leq 2$$

- gdy laboratorium otrzymuje jeden lub obydwa wskaźniki w przedziale $(2 ; 3]$ lub w przedziale $(-2 ; -3]$, wówczas biegłość laboratorium oceniana jest jako wątpliwa:

$$2 < \{|Z_A| \text{ i / lub } |Z_B|\} < 3$$

- gdy laboratorium otrzymuje jeden lub obydwa wskaźniki ≥ 3 bądź ≤ -3 , wówczas biegłość laboratorium oceniana jest jako niezadowalająca:

$$\{|Z_A| \text{ i / lub } |Z_B|\} \geq 3$$

Takie same kryteria oceny biegłości obowiązują w przypadku stosowania wskaźników zmodyfikowanych (Z'_A , Z'_B).

15. Diagram Youdena

15.1. Konstrukcja diagramu

Diagram Youdena służy do graficznej analizy błędów, którymi obarczone są pary wyników raportowane przez laboratoria oraz umożliwia bezpośrednie porównanie wyników pomiędzy laboratoriami.

Diagram składa się z:

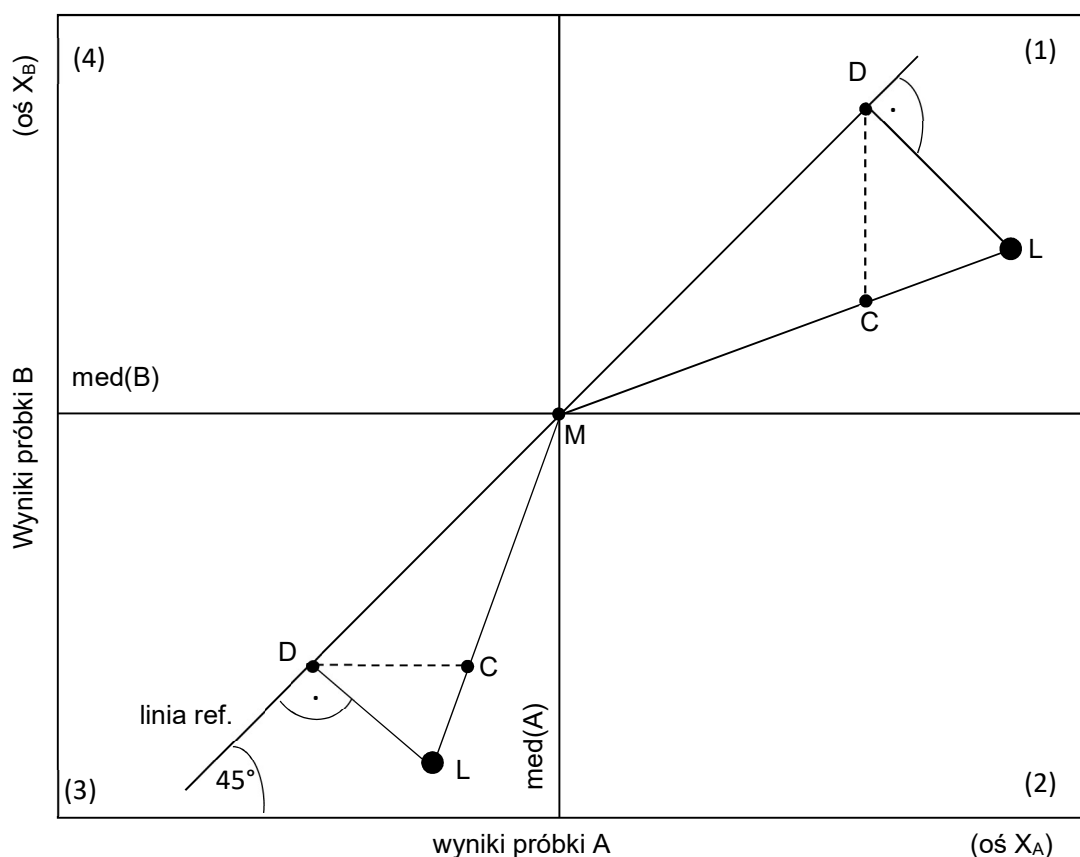
- osi wyników próbki A (oś X_A) oraz osi wyników próbki B (oś X_B), skala obydwu osi jest taka sama
- dwóch wzajemnie prostopadłych linii median (linia med (A) i linia med (B)), które przecinają się w punkcie M zwanym medianą Manhattan i dzielą diagram na cztery tzw. kwadranty, oznakowane (1), (2), (3), (4). Mediana Manhattan jest najbardziej prawdopodobną wartością „prawdziwą” wspólną dla obu próbek
- linii referencyjnej przeprowadzonej przez punkt M pod kątem 45° względem osi X_A . Linia referencyjna pokazuje zgodność pomiędzy wynikami próbki A i próbki B



15.2 Graficzna analiza błędów laboratoryjnych przy pomocy diagramu Youdena.

Na diagram Youdena nanoszone są pary wyników (X_A , X_B) z poszczególnych laboratoriów. Każdy punkt na diagramie ma współrzędne (X_A = wynik próbki A, X_B = wynik próbki B) i przedstawia pojedyncze laboratorium. Obok punktów podawane są kody laboratoriów.

Na podstawie położenia laboratorium na diagramie, możliwe jest oszacowanie wielkości błędu całkowitego laboratorium oraz wielkości dwóch błędów składowych: błędu systematycznego i błędu przypadkowego, wyodrębnionych z błędu całkowitego. Sposób wykonania graficznej analizy błędów, dla laboratorium zlokalizowanego na diagramie Youdena w kwadrancie (1) lub w kwadrancie (3) przedstawia rysunek poniżej:



Objaśnienia dotyczące analizy błędów:

punkt M o współrzędnych [$\text{med}(A)$, $\text{med}(B)$] jest medianą Manhattan

punkt L o współrzędnych (X_A , X_B) reprezentuje lokalizację laboratorium na diagramie

odcinek ML = wielkość błędu całkowitego laboratorium (błąd całkowity jest sumą błędu przypadkowego i systematycznego)



odcinek DM = miara błędu systematycznego (błąd systematyczny jest proporcjonalny do odcinka DM)

odcinek LD = miara błędu przypadkowego (błąd przypadkowy jest proporcjonalny do odcinka LD)

odcinek MC = wielkość błędu systematycznego

odcinek LC = wielkość błędu przypadkowego

Wielkości tych błędów laboratoryjnych można łatwo oszacować poprzez pomiar długości odpowiednich odcinków (ML, MC, LC) i wyrażenie ich w jednostkach (mg/l) przy pomocy skali na osi próbki A lub B.

Z analizy graficznej wynika, że w przypadku obydwu laboratoriów, wkład błędu systematycznego do błędu całkowitego jest znacznie większy od wkładu błędu przypadkowego. Stąd wniosek, że za błąd całkowity odpowiedzialny jest błąd systematyczny.

Ta informacja kieruje uwagę na poszukiwanie i eliminowanie źródeł błędu systematycznego (a nie przypadkowego) i to jest zasadnicza korzyść wynikająca z analizy graficznej Youdena.

15.3 Elipsa ufności na diagramie Youdena

W języku statystyki, elipsa ufności jest dwuwymiarowym obszarem ufności dla wartości „prawdziwej” wspólnej dla obu próbek. Elipsa jest wyznaczana na określonym poziomie ufności, zwykle równym ok. 95%.

W języku praktycznym, 95% elipsa ufności wyznacza granice dopuszczalnego błędu całkowitego, którym mogą być obciążone pary wyników (X_A , X_B).

Elipsa oparta jest na założeniu, że pary wyników podlegają rozkładowi zbliżonemu do dwuwymiarowego rozkładu normalnego. W Programie założenie to jest spełnione, bowiem na etapie diagnozowania symetrii zapewniono, że rozkłady wyników próbki A i próbki B są zgrubnie normalne. Do wyznaczenia parametrów elipsy, to jest: wielkości półosi dużej (a), wielkości półosi małej (b) oraz kąta (α) orientacji elipsy w układzie osi X_A , X_B , stosowane są trzy statystyki:

- wariancja wyników próbki A $\rightarrow \text{var}(A) = \text{NIQR}^2(A)$
- wariancja wyników próbki B $\rightarrow \text{var}(B) = \text{NIQR}^2(B)$
- kowariancja pomiędzy wynikami próbki A i B $\rightarrow \text{cov}(A, B) = \rho_s \cdot \text{NIQR}(A) \cdot \text{NIQR}(B)$

gdzie: ρ_s jest współczynnikiem korelacji rangowej Spearmana, wyznaczonym wg poprzedniego wydania ISO 13528 z roku 2005.



Z tych statystyk zbudowana jest, według [12] macierz kowariancji CVC

$$CVC = \begin{bmatrix} \text{var}(A) & \text{cov}(A, B) \\ \text{cov}(A, B) & \text{var}(B) \end{bmatrix}$$

Pierwiastki kwadratowe z obliczonych wartości własnych tej macierzy (λ_1, λ_2) są równe długościom półosi elipsy, a mianowicie:

$$\text{długość półosi dużej } a = \sqrt{\lambda_1} \quad \text{długość półosi małej } b = \sqrt{\lambda_2}$$

Wyznaczone w ten sposób półosie odpowiadają poziomowi ufności ok. 45%, dlatego aby utworzyć 95% elipsę, długości półosi są zwielokrotnione przy pomocy dwuwymiarowego współczynnika rozszerzenia $k = 2,45$. W przypadku 99% elipsy, współczynnik $k = 3,03$. Orientacja elipsy w układzie osi X_A, X_B jest zdefiniowana przez kąt nachylenia (α) dużej osi elipsy względem:

- osi próbki A odmierzonej w kierunku przeciwnym do ruchu wskazówek zegara, gdy $NIQR(A) > NIQR(B)$
- osi próbki B odmierzonej w kierunku zgodnym z ruchem wskazówek zegara, gdy $NIQR(B) > NIQR(A)$

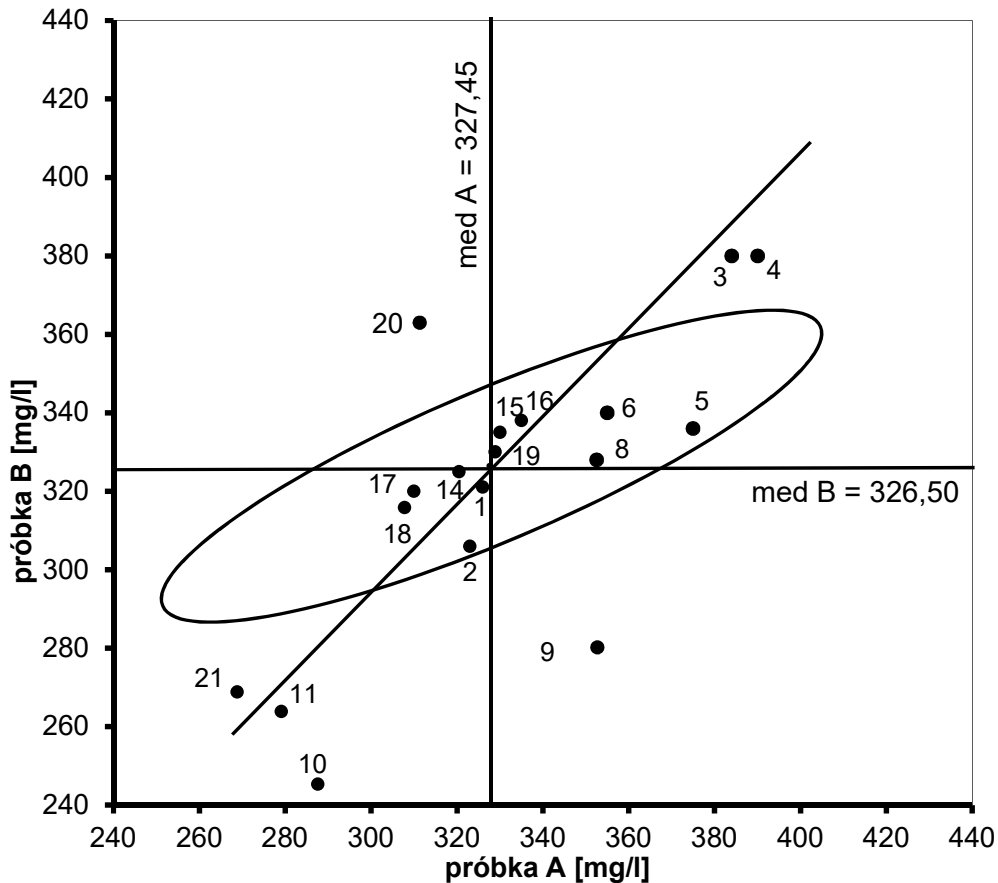
Kąt (α) jest wyznaczany wg [12], z zależności:

$$\alpha = \frac{1}{2} \arctg \left[\frac{2 \cdot \text{cov}(A, B)}{NIQR^2(A) - (NIQR^2(B))} \right]$$

Bardziej szczegółowe rozważania dotyczące wyznaczania elipsy ufności przekraczają ramy tego przewodnika. Jednakże laboratoria nie mają potrzeby samodzielnego wyznaczania elipsy ufności, ponieważ elipsy są wyznaczone i przedstawione na diagramach Youdena w każdym raporcie z badania biegłości.

95% elipsa ufności pozwala łatwo rozgraniczyć laboratoria, z biegłością zadowalającą (zlokalizowane w obszarze elipsy) od pozostałych laboratoriów. Szerokość elipsy pokazuje jak skorelowane są wyniki próbki A i próbki B. Gdy wyniki próbki A i próbki B są bardzo słabo skorelowane (ρ_s bliskie zeru) elipsa przyjmuje formę okręgu.

Przykładowy diagram Youdena z 95% elipsą ufności jest przedstawiony na następnej stronie.



długość półosi dużej = 83,50 mg/l

długość półosi małej = 16,59 mg/l

kąt nachylenia dużej osi elipsy

względem osi próbki A, $\alpha = 23^\circ$

Laboratoria: 3, 4, 10, 11, 21 wykazują błąd całkowity zdominowany przez błąd systematyczny (wyniki obu próbek są zawyżone lub zaniżone wobec odpowiednich wartości przypisanych).

Punkty reprezentujące laboratoria są znacznie odległe od punktu przecięcia linii median oraz zlokalizowane wzdłuż linii referencyjnej. To świadczy o tym, że laboratoria mają problemy z poprawnością analiz.

Laboratoria 9, 20 wykazują błąd całkowity zdominowany przez błąd przypadkowy (wynik jednej próbki jest zawyżony oraz wynik drugiej próbki jest zaniżony). Punkty reprezentujące laboratoria są zlokalizowane w kwadrantach (2) i (4) oraz znacznie oddalone od linii referencyjnej. To świadczy o tym, że laboratoria cechują się słabą precyzją powtarzalności.



16. Wstępny Protokół Oceny Biegłości oraz Raport z Badania Biegłości

Laboratoria, które dla określonego analitu / analitów wykazały w danej rundzie biegłość niezadowalającą lub w danej i poprzedniej rundzie biegłości wątpliwą, otrzymują – w ciągu pięciu dni roboczych od daty zamknięcia rundy – Wstępny Protokół Oceny Biegłości, który zawiera:

- wyniki laboratorium (X_A , X_B)
- wskaźniki (Z_A , Z_B)
- ocenę biegłości
- wybrane statystyki podsumowujące

Protokół jest sporządzany w formacie pdf i przekazywany drogą mailową.

Raport z Badania Biegłości zawiera:

- wprowadzenie
- cele Programu
- opis próbek do badania biegłości (przygotowanie, badanie homogeniczności i stabilności)
- opis postępowania statystycznego
- dot. ploty wyników próbki A i B
- wyniki przekazane przez laboratoria oraz korespondujące z wynikami odporne wskaźniki biegłości
- statystyki opisowe podsumowujące
- diagram Youdena z elipsą ufności
- analizę błędów laboratoryjnych
- błędy obliczone dla laboratoriów, które wykazały biegłość niezadowalającą
- ocenę realistyczności niepewności
- podsumowanie i komentarz edukacyjny

Raport jest sporządzany w formacie pdf i dostarczany do każdego uczestnika drogą mailową.

17. Przeciwdziałanie znowie i fałszowaniu wyników

Fałszowanie wyników i znowa pomiędzy uczestnikami Programu odwraca cele badania biegłości i jest sprzeczna z zasadami profesjonalnej praktyki laboratoryjnej.

Zapis dotyczący zakazu znowy i fałszowania wyników – pod rygorem unieważnienia rundy – jest zawarty w formularzu zgłoszenia do Programu. Podpisanie formularza jest równoznaczne



z zobowiązaniem uczestnika do przestrzegania tego zakazu. Laboratoria, którym udowodniono fałszowanie wyników nie będą mogły uczestniczyć w kolejnych rundach Programu.

Dodatkowym środkiem zapobiegającym jest dostarczanie uczestnikom par próbek, które mogą być zgodne lub rozszczepione. Ponieważ uczestnicy nie mają wiedzy jaki rodzaj par otrzymali, nie mogą porównywać i uzgadniać wyników między sobą.

18. Informacje zwrotne od uczestników

Informacje dotyczące Programu otrzymane od uczestników, mają duże znaczenia dla Koordynatora Programu.

Koordinator jest otwarty na wszelkie spostrzeżenia i sugestie związane z Programem, przekazywane przez uczestników – w dowolnym czasie – drogą mailową lub telefoniczną. Ponadto Koordynator regularnie zbiera informacje dotyczące potrzeb i oczekiwań uczestników, przy pomocy Kwestionariuszy Oceny, które przekazuje do każdego uczestnika po zakończeniu rundy. Informacje zawarte w Kwestionariuszach Oceny są wykorzystywane do doskonalenia Programu.

19. Opis Programu WASTER-QR

19.1. Cel i przeznaczenie Programu WASTER-QR

Program WASTER-QR (zwany dalej w tekście QR) jest bilateralnym badaniem biegłości, organizowanym na życzenie zainteresowanego laboratorium / ów.

Bilateralne badanie biegłości jest zdefiniowane w dokumencie EA-03/04 (pkt 4.3) i zgodnie z tym dokumentem (pkt 6.2a) może być stosowane jako rodzaj badania biegłości, uznawany przez jednostkę akredytującą.

Celem Programu QR jest porównanie wyniku uczestnika z predefiniowaną wartością referencyjną (pochodzącą z zewnątrz) i w granicach zmienności ustalonych przez organizatora przed rozpoczęciem rundy.

Program QR jest Programem suplementarnym do regularnego Programu WASTER i przeznaczonym dla laboratoriów, które potrzebują szybko wyniki oceny biegłości, bez konieczności oczekiwania na planowaną rundę Programu WASTER. Analizy objęte Programem QR wraz z przewidywanym zakresem stężeń są publikowane na stronie organizatora.



Z Programu QR mogą korzystać laboratoria, które chcą, np.

- 1) sprawdzić skuteczność działań korygujących przeprowadzonych w następstwie uzyskania niezadowalającej oceny biegłości w rundzie Programu WASTER lub w innym programie badania biegłości (zgodnie z wymaganiami dokumentu PCA, DA-05 pkt 3.14)
- 2) ocenić umiejętności techniczne nowego personelu
- 3) sprawdzić funkcjonowanie nowej metody, bądź metody do której wprowadzono zmiany
- 4) zademonstrować kompetencje techniczne w związku z rozszerzeniem zakresu akredytacji

19.2. Próbki do badania biegłości

W Programie QR stosowane są próbki syntetyczne w formie: certyfikowanych koncentratów wzorcowych, które wymagają rozcieńczenia przed analizą lub w formie certyfikowanych roztworów wzorcowych gotowych do analizy. Próbki syntetyczne mogą zawierać jeden lub kilka analitów objętych Programem. Zainteresowane laboratorium otrzymuje zwykle parę próbek (A, B) tak jak w Programie WASTER, bądź na życzenie klienta może otrzymać pojedynczą próbkę.

Homogeniczność próbek jest oceniona i potwierdzona przez producenta. Próbki są stabilne podczas przechowywania w warunkach i przez czas określony w certyfikacie roztworu wzorcowego / koncentratu wzorcowego.

19.3. Wartość przypisana i odchylenie standardowe do oceny biegłości

W Programie QR, wartość przypisana (X_{pt}) i odchylenie standardowe do oceny biegłości (σ_{pt}) dla wszystkich analitów objętych Programem są predefiniowane zgodnie z tabelą:



Rodzaj próbki PT	Wartość przypisana X_{pt}^*	Odchylenie standardowe do oceny biegłości, σ_{pt}
Roztwór wzorcowy	Referencyjna wartość stężenia analitu określona w certyfikacie roztworu wzorcowego	Procent wartości przypisanej w zależności od badanego analitu $\sigma_{pt} = \%CV_{rob} \cdot X_{pt}$ $\%CV_{rob}$ jest wyznaczane z równania regresji, które opracowano na podstawie danych z rund Programu WASTER, zgromadzonych na przekroju sześciu lat.
Koncentrat wzorcowy	Certyfikowana wartość stężenia analitu po rozcieńczeniu (według instrukcji producenta), określona w certyfikacie analizy	

(*) wartość przypisana jest spójna do wzorców NIST

Niepewność standardowa wartości przypisanej $u(X_{pt})$ jest otrzymywana z informacji dostarczonej na certyfikacie analizy roztworu wzorcowego / koncentratu wzorcowego (po rozcieńczeniu). Niepewność ta nie bierze udziału w ocenie biegłości uczestnika, natomiast może być wykorzystana – na życzenie uczestnika – do oceny realistyczności niepewności raportowanej przez uczestnika wraz z wynikiem badania.

19.4 System oceny biegłości stosowany w Programie QR

Do oceny biegłości stosowany jest wskaźnik (z), obliczany dla wyniku uczestnika (x), według formuły:

$$z = \frac{X - X_{pt}}{\sigma_{pt}}$$

gdzie: X_{pt} jest wartością przypisaną

σ_{pt} jest odchyleniem standardowym do oceny biegłości

Interpretacja wskaźnika (z) jest następująca:

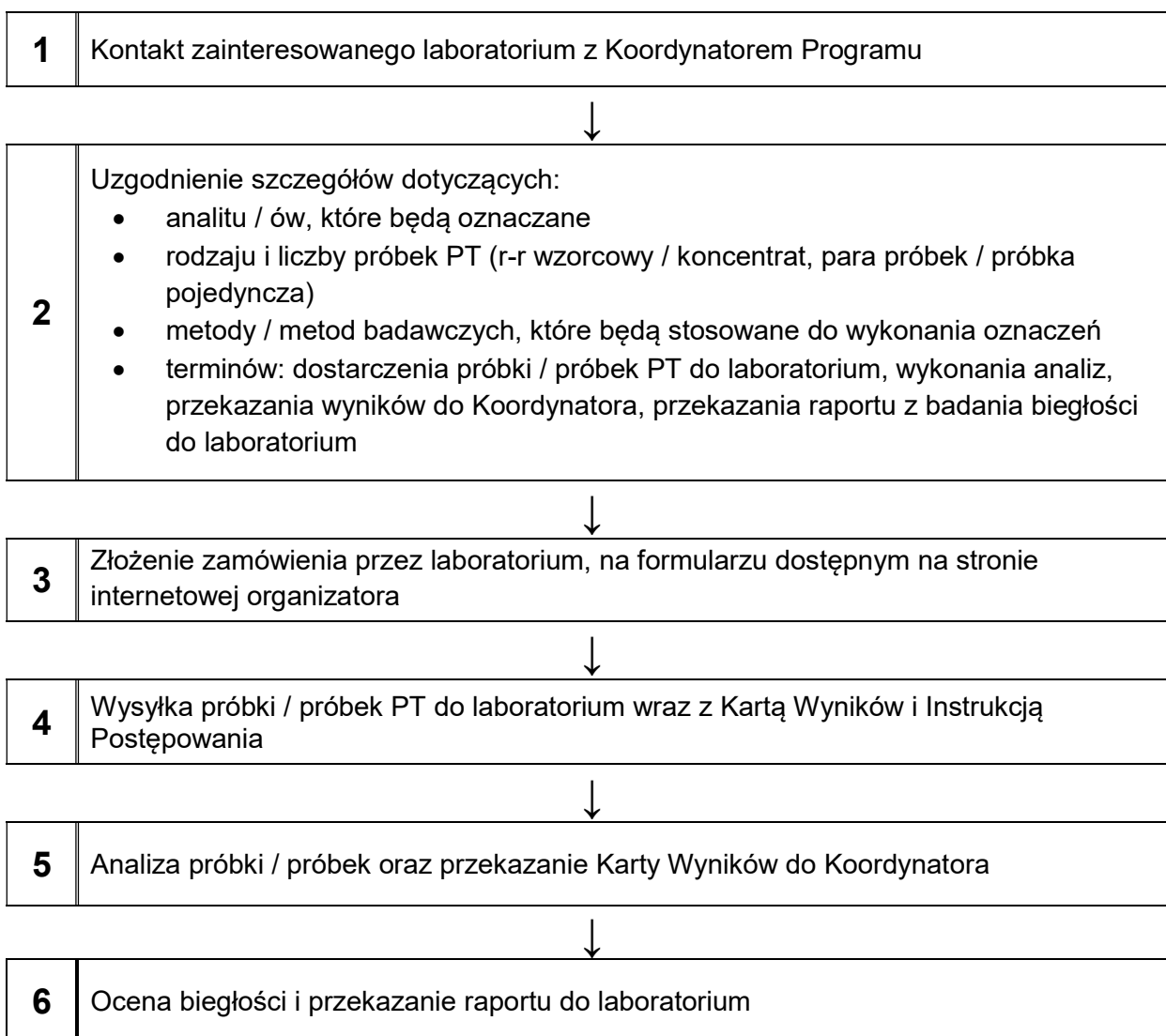
- wynik, który otrzymuje wskaźnik $|z| \leq 2$, jest oceniany jako zadowolający
- wynik, który otrzymuje wskaźnik $2 < |z| < 3$, jest oceniany jako wątpliwy i daje sygnał ostrzegawczy
- wynik, który otrzymuje wskaźnik $|z| \geq 3$, jest oceniany jako niezadowolający i daje sygnał działania

Gdy uczestnik wykonuje (w warunkach powtarzalności) analizę pary próbek (A, B) i tym samym raportuje parę wyników (X_A , X_B), obliczane są dwa wskaźniki biegłości Z_A i Z_B według formuły



przedstawionej powyżej i biegłość uczestnika jest oceniana według takich samych kryteriów jakie są stosowane w Programie WASTER (patrz pkt 14).

19.5. Flowchart funkcjonowania Programu QR



20. Publikacja i aktualizacja przewodnika

Przewodnik jest dokumentem nadzorowanym przez Koordynatora Programu.

Aktualna wersja przewodnika jest publikowana na stronie organizatora.



21. Bibliografia

- [1] Youden W.J., Graphical Diagnosis of Interlaboratory Test Results, Industrial Quality Control, Vol.15, No.11, (1959)
- [2] Youden W. J., The sample, the procedure and the laboratory, Anal. Chem. 32, (1960)
- [3] IUPAC Technical Report: 2006 „International harmonized protocol for the proficiency testing of analytical chemistry laboratories”
- [4] ISO 13528:2015 „Statistical methods for use in proficiency testing by interlaboratory comparisons”
- [5] Eurachem Guide „Selection, use and interpretation of proficiency testing (PT) schemes by laboratories”: 2011
- [6] AMC Technical Brief No.6, (2001)
- [7] Ellison L. R. i inni, Practical Statistics for Analytical Scientist: A Bench Guide, LGC Limited, (2009)
- [8] Huber P. J., Robust Statistics, Wiley, New York, (1981)
- [9] Wilks S. S., Mathematical Statistics, Wiley, New York, (1962)
- [10] D`Agostino R.B., Belanger A. and D`Agostino R.B. Jr.: A suggestion for using powerful and informative tests of normality, The American Statistician, 44(4), (1990)
- [11] Sheskin D. J., Handbook of parametric and non parametric statistical Procedures, (2011)
- [12] Mandel J., Lashof T., Interpretation and Generalization of Youden`s Two – Sample Diagram, Journal of Quality Technology, January, (1974)