

Praktyczna statystyka dla farmacji

Ocena Stabilności Produktów Leczniczych

Michał Iwaniec, Mirosław Popieluch
StatSoft Polska



Wraz z rosnącymi wymaganiami rynku oraz coraz bardziej rygorystycznymi normami narzucanymi przez organy regulacyjne, firmy wytwarzające środki farmaceutyczne stają przed coraz to nowymi wyzwaniami technologiczno-organizacyjnymi. W szczególności rosnąca ilość analiz, porównań, zestawień i raportów, które należy wykonywać okresowo dla celów kontrolnych i dokumentacyjnych, generują coraz większą ilość rutynowo wykonywanej pracy. Zadania te dotyczą zarówno obszarów organizacyjnych,

jak i technologicznych związanych z jakością wytwarzanych produktów.

O ile kwestie związane z czysto formalną dokumentacją dotyczącą organizacji procesu produkcyjnego nie przysparzają problemów, to w dziedzinie analizy danych produkcyjnych są obszary wymagające znajomości zaawansowanych technik statystycznej analizy danych, umiejętności wykonywania analiz statystycznych i wyciągania na ich podstawie wniosków. Analiza danych to nie tylko zestawienia dotyczące danych historycznych. Dzięki niej możemy pójść o krok dalej i na podstawie danych historycznych prognozować wartość określonych parametrów. Właśnie takie podejście jest stosowane przy ocenie stabilności produktów leczniczych.

Zbieranie i obróbka danych

Analiza danych nie może oczywiście istnieć bez samych danych. O ile procedury zbierania danych do celów badania stabilności są dobrze ustalone, metody analityczne służące do wykonania pomiarów są poznane, to zdarza się, że wyciągnięcie wniosków na podstawie wyników pomiarów często stwarza wiele problemów. Wpływa na to wiele czynników, takich jak: nieznanostwo metod statystycznych, brak doświadczenia w interpretacji wyników oraz w stosowaniu dodatkowych analiz przygotowujących dane do właściwej analizy (np. wykrywanie wartości odstających) i wreszcie brak przekonania analityka, co do poprawności wykonywanych analiz.

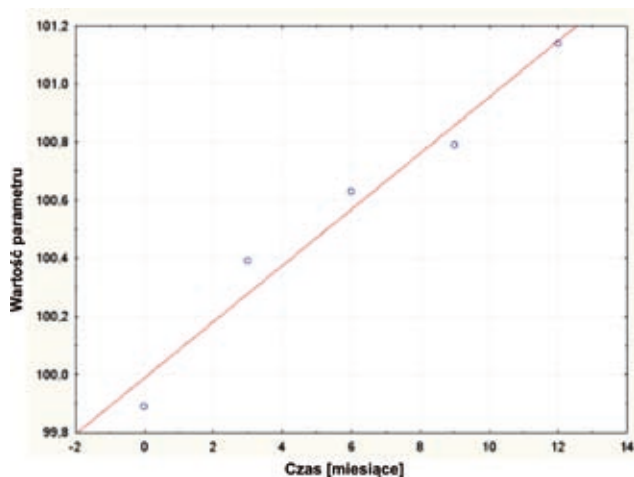
Ze statystycznego punktu widzenia większość analiz proponowanych do oszacowania stabilności produktu leczniczego w założonym okresie czasu nie jest skomplikowana, ale biorąc pod uwagę pewne czynniki związane ze specyfiką tej analizy i kilka możliwych ścieżek postępowania, poprawne przeprowadzenie obliczeń statystycznych może być czasochłonne lub nawet niemożliwe do wykonania bez wykorzystania odpowiedniego oprogramowania.

Ogólne podejście do analizy danych w ocenie stabilności

Analiza stabilności polega na określeniu, po jakim czasie analizowany parametr przekroczy określone dla niego limity (granice) tolerancji (specyfikacji). Przykładowo jeżeli określono, że zawartość zanieczyszczeń nie może być większa niż 2%, to będziemy chcieli określić czas, po jakim ta wartość zanieczyszczeń zostanie przekroczona.

Chcemy zatem powiązać nie tyle aktualną wartość parametru z określonym limitem co zmianę parametru w czasie z limitem. Dodatkowo w niektórych

przypadkach będziemy chcieli sprawdzić, czy parametry leku lub substancji po np. 24 miesiącach będą mieścić się w limitach, jeżeli mamy dane zebrane tylko po 0, 3, 6, 9 i 12 miesiącach przechowywania. Żeby to określić, będziemy potrzebowali modelu zależności pomiędzy badanym parametrem a czasem przechowywania. Najprostszym podejściem, proponowanym np. przez ICH (1), jest założenie, że parametr zmienia się z czasem w sposób liniowy. Mówiąc inaczej, zakładamy, że wykres zależności pomiędzy czasem i badanym parametrem można przybliżyć linią prostą:



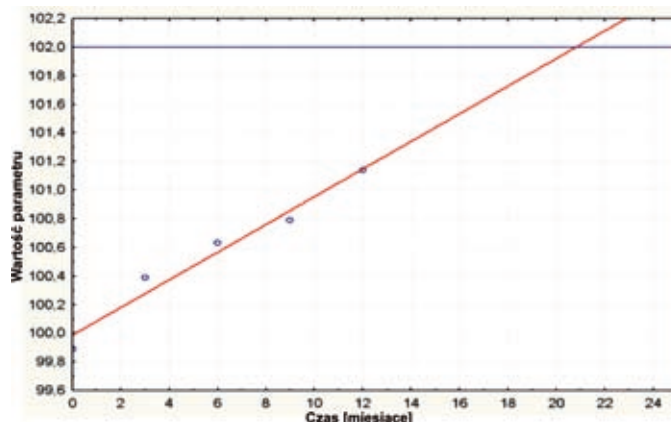
Dopasowanie prostej regresji
– wykres z programu STATISTICA

Tak jak to widać na rysunku powyżej, dane wydają się podlegać zależności liniowej, ale nie jest to zależność funkcyjna, tylko statystyczna. Oznacza to, że punkty nie leżą na linii prostej, ale są wokół niej rozrzucone. Zmienność pomiarów w czasie może zależeć od bardzo wielu czynników, takich jak błąd podczas pomiaru lub niestabilność warunków przechowywania. Wszystko to prowadzi do tego, że będziemy musieli poradzić sobie ze zjawiskiem, które obarczone jest pewną niepewnością (zmiennością). Najlepiej z takimi zagadnieniami radzi sobie statystyka.

Jeżeli zakładamy, że rozważana zależność jest liniowa (w pierwszej kolejności musimy to założenie zweryfikować, używając odpowiedniego testu statystycznego) to możemy skorzystać z dobrze opisanego modelu regresji liniowej. Zauważmy, że jeżeli zmiana parametru w czasie nie jest liniowa, wtedy do każdej grupy zmienności w czasie musimy zastosować oddzielny aparat matematyczny, który będzie opisywał zmienność lepiej niż linia prosta. Zazwyczaj będzie to jednak dużo trudniejsze obliczeniowo.

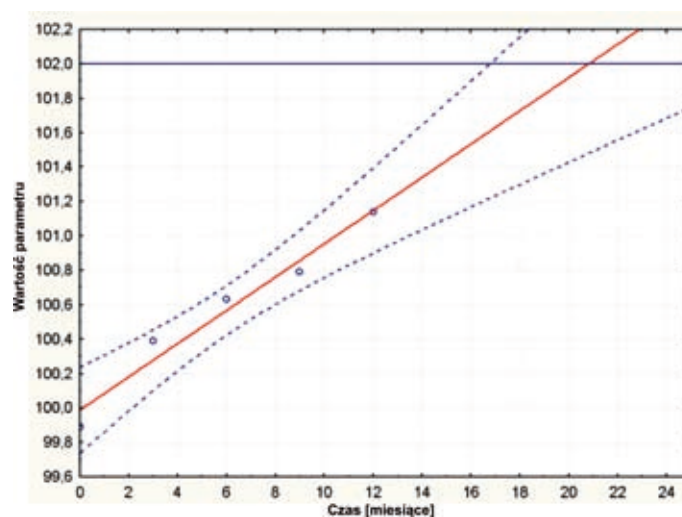
Jeżeli rozpatrywana zależność jest liniowa, wtedy do danych możemy dopasować linię prostą, stosując regresję liniową. Oczywiście dopasowana w ten sposób linia nie będzie pokrywać się ze wszystkimi punktami – będą one rozrzucone wokół prostej, co jest cechą zależności statystycznej. Żeby oszacować, kiedy model

zmian naszego parametru (linia prosta $y=at+b$) przecnie limit (np. górny), możemy w najprostszym przypadku wyznaczyć ten punkt przecięcia metodą graficzną:



Prosta regresji przecinająca górny limit
– wykres z programu STATISTICA

Z powyższego rysunku wynika, że przecięcie limitu następuje po ok. 21 miesiącach. Musimy jednak wziąć pod uwagę to, że nasz model (dopasowana do danych prosta) jest obarczony pewnym błędem, lub mówiąc inaczej, wyznaczona prosta nie odwzorowała prawdziwej zależności, a tylko ją przybliżyła. Powinniśmy zatem uwzględnić niepewność, jaką ta analiza niesie ze sobą, wyznaczając obszar ufności dla prostej regresji. Wytyczne sugerują, aby wyznaczony obszar był obszarem 95%, co oznacza, że prosta wyznaczająca rzeczywistość (niestety nieznaną) zależność na 95% powinna znaleźć się w wyznaczonym obszarze. Po wyznaczeniu obszaru ufności powinniśmy uwzględnić miejsca przecięcia linii wyznaczających granicę obszaru z liniami określonymi przez limity (u nas, tak jak na rys. powyżej, górnym limitem dla parametru jest wartość 102).



Przecięcie granicy obszaru ufności z limitem
– wykres z programu STATISTICA

Na powyższym rysunku dla rozważanych danych został wyznaczony obszar ufności. Widać, że linia

wyznaczająca górne ograniczenie obszaru ufności spotka się z linią limitu po ok. 17 miesiącach. Jak można było się domyślać, czas uległ skróceniu, ale jednocześnie wynik jest bardziej prawdopodobny, co w farmacji ma zasadnicze znaczenie.

Obliczenia te nie są skomplikowane, ale trudności mogą się pojawić podczas wyznaczania punktu przecięcia górnej lub dolnej linii ograniczającej obszar ufności z limitami (górnym lub dolnym). Wyniki obliczeń oraz dokładne punkty przecięcia są standardowo podawane w programie *STATISTICA Ocena Stabilności*.

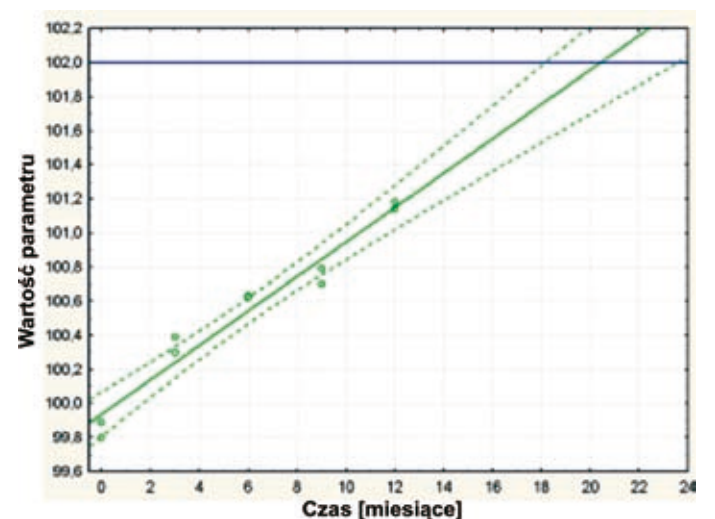
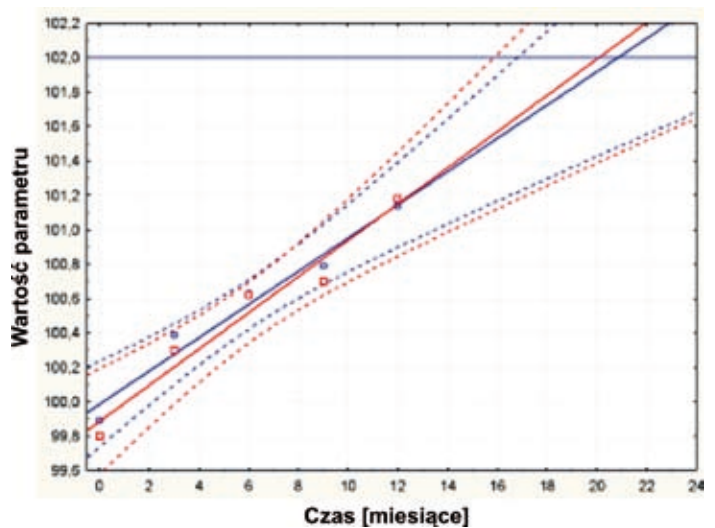
Analizy statystyczne w przypadku większej liczby danych

Według wytycznych w celu wykazania stabilności parametru w produkcie lub substancji należy przeprowadzić analizy, opierając się na danych pochodzących z co najmniej trzech serii. W takim przypadku możemy postępować analogicznie jak wyżej w celu oszacowania czasu przekroczenia limitów przez dany parametr, traktując każdy zbiór danych oddzielnie (analizując osobno dane z każdej z serii), a następnie przyjmując najkrótszy oszacowany czas.

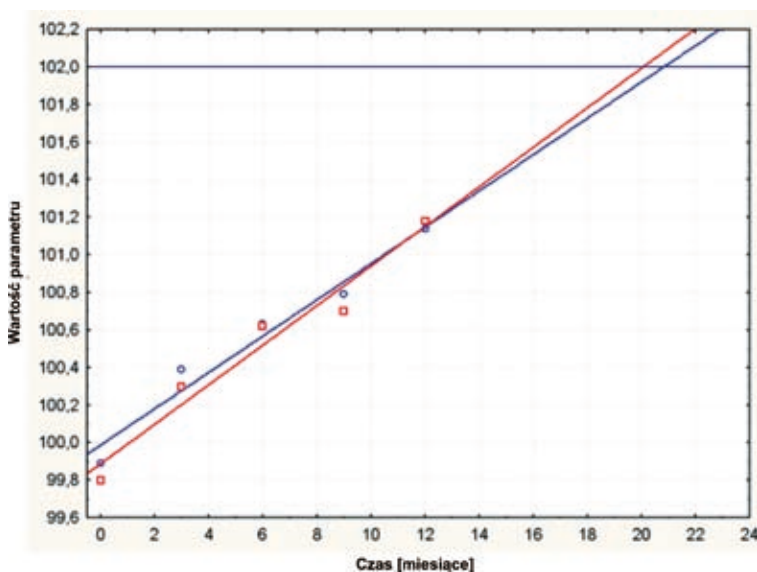
Według wytycznych ICH dane, pochodzące z różnych serii, możemy potraktować jako jednolity zbiór danych (można analizować je wspólnie). Jednak aby można było tak zrobić, dane z poszczególnych serii muszą być porównywalne, a dokładniej mówiąc, porównywalne muszą być zależności pomiędzy badanym parametrem a czasem dla poszczególnych serii.

Wcześniej założyliśmy, że zależności są liniowe i że stosować będziemy model regresyjny, tzn. do danych dopasowywać będziemy prostą metodą regresji liniowej. Żeby można było powiedzieć, że zależności liniowe w dwóch zbiorach danych są równoważne, musimy przeprowadzić testy statystyczne badające, czy (1) współczynniki kierunkowe wyznaczonych prostych oraz (2) czy wyrazy wolne wyznaczonych prostych nie różnią się od siebie w sensie statystycznym. Rozważmy sytuację jak na rysunku poniżej:

Do danych rozważanych wyżej dodane zostały dane pochodzące z innej serii tego samego produktu w tym samym opakowaniu. Do nowych danych została dopasowana druga prosta metodą regresji liniowej. Widać gołym okiem, że dwie proste nie pokrywają się, ale czy nachylenia i położenie prostych są od siebie rzeczywiście różne, biorąc pod uwagę pewien losowy rozrzut punktów w każdej grupie danych wokół linii prostych? Okazuje się, że w sensie statystycznym zależności pomiędzy czasem a badanym parametrem w obu seriach nie są od siebie istotnie różne. Jeżeli nie jesteśmy w stanie wykazać istotnych różnic, to możemy potraktować te dane jako jednolity zbiór i wspólnie analizować. Jaki z tego jest zysk, możemy zobaczyć, porównując rysunki poniżej:



Porównania analizy dla danych traktowanych rozłącznie i wspólnie - wykresy z programu *STATISTICA*



Wykres dopasowania prostej regresji dla dwóch serii – wykres z programu *STATISTICA*

ANALIZA DANYCH W PRZEMYŚLE FARMACEUTYCZNYM

W przypadku potraktowania danych oddzielnie musimy do danych dopasować dwie proste, wyznaczyć dla nich dwa obszary ufności i jako szacowany czas przekroczenia limitów powinniśmy wziąć ten mniejszy. W wyniku tak przeprowadzonej analizy czas stabilności parametru wyniósłby ok. 16 miesięcy. Jeżeli dane potraktujemy wspólnie, czyli do wszystkich dostępnych danych dopasujemy prostą regresji i wyznaczymy dla niej jeden obszar ufności, wtedy oszacowany w ten sposób okres stabilności wyniesie ok. 18 miesięcy. Skąd ta różnica? Otóż w przypadku analizy stabilności borykamy się zazwyczaj z małą ilością danych, co wpływa na dokładność analizy statystycznej, czego skutkiem jest poszerzenie obszaru ufności. Im danych jest więcej, tym analiza jest „mocniejsza” i obszar ufności zawęża się, co zazwyczaj powoduje zwiększenie szacowanego czasu stabilności. Ogólnie mówiąc im więcej danych i im bardziej są one skupione, tym obszar ufności będzie węższy i wnioski wyciągnięte z analizy bardziej prawdopodobne.

Podsumowanie

Cały cykl czynności związanych z analizą danych i przygotowaniem raportu jest niezwykle pracochłonny. Na szczęście na rynku dostępny jest dedykowany pakiet oprogramowania *STATISTICA*, który wspiera użytkowników poprzez automatyzację procesu analizy danych produkcyjnych, skracając czas przygotowania poszczególnych raportów, nawet z 5 dni do 5 minut! *STATISTICA Zestaw Farmaceutyczny (www.StatSoft.pl)* to pakiet programów przeznaczonych do realizacji zadań analitycznych związanych z oceną stabilności produktów, walidacją procesu technologicznego oraz porównywaniem profili uwalniania. Zawiera zestaw odpowiednio dobranych metod i analiz statystycznych, które automatycznie wykonują rutynowe działania związane z analizowaniem danych i tworzą gotowy do wydruku raport, który może być bezpośrednio umieszczony w dokumentacji medycznej i walidacyjnej przygotowywanej na potrzeby audytów.

Organizacje zajmujące się opracowywaniem i produkcją środków farmaceutycznych na całym świecie wykorzystują oprogramowanie *STATISTICA* jako zintegrowaną platformę analizy danych, wspierającą powstawanie nowych leków oraz sterowanie procesem i jakością.



Badanie stabilności produktów
Porównywanie profili uwalniania
Walidacja procesów technologicznych



Walidacja metod pomiarowych
Stabilność metod pomiarowych
Kalibracja liniowa



Automatyzacja analiz (np. APR lub PPR)
Systemy zbierania, gromadzenia i analizy pomiarów
Współpraca z systemami LIMS

STATISTICA®



StatSoft®

StatSoft Polska, ul. Kraszewskiego 36, 30-110 Kraków, faks 012 4284301
tel. 012 4284300, 0601 414151, e-mail: info@statsoft.pl, www.StatSoft.pl