



BUDAPESTI MŰSZAKI ÉS GAZDASÁGTUDOMÁNYI EGYETEM
VEGYÉSZMÉRNÖKI ÉS BIOMÉRNÖKI KAR
OLÁH GYÖRGY DOKTORI ISKOLA

Doktori értekezés

Gyógyszerkészítmények stabilitásvizsgálatának statisztikai elemzése

Szerző:

Lakné Komka Kinga
okleveles biomérnök

Témavezető:

Dr. Kemény Sándor
egyetemi tanár

Budapesti Műszaki és Gazdaságtudományi Egyetem
Kémiai és Környezeti Folyamatmérnöki Tanszék

2010

TARTALOMJEGYZÉK

Köszönetnyilvánítás	5
I. Bevezetés	6
I.1. Motiváció	6
I.2. A doktori munka célkitűzései.....	6
I.3. Számítási módszerek	7
II. Irodalmi áttekintés	8
II.1. Gyógyszerkészítmények törzkönyvezése.....	8
II.2. Gyógyszerkészítmények stabilitásvizsgálatának általános leírása.....	9
II.2.1. A stabilitásvizsgálat célja.....	9
II.2.2. A stabilitásvizsgálat típusai	10
II.2.3. Vizsgálandó tulajdonságok	11
II.2.4. Tárolási körülmények	11
II.2.5. Vizsgálati időpontok	12
II.2.6. A vizsgálandó sarzsok kiválasztása	13
II.3. A stabilitásvizsgálat statisztikai értékelésének jelenlegi módszere	13
II.3.1. Alapelvek (egyetlen sarzs vizsgálata).....	13
II.3.1.1. A statisztikai értékelés menete.....	13
II.3.1.2. Az eltarthatósági idő definíciója és meghatározása	14
II.3.1.3. A degradálódási függvény jellegének meghatározása	17
II.3.1.4. A konfidencia-sáv fogalma és tulajdonságai	17
II.3.2. Több sarzs együttes vizsgálata	19
II.3.2.1. A sarzsok adatainak egyesítése.....	19
II.3.2.2. Az egyesítés feltételeinek ellenőrzése	22
II.3.2.3. Az egyesíthetőség vizsgálata során használt statisztikai próbák.....	23
II.3.2.3.1. A meredekségek azonosságának vizsgálata	24
II.3.2.3.2. A tengelymetszetek azonosságának vizsgálata	25
II.3.2.4. A vizsgálatoknál alkalmazott szignifikancia-szint.....	27
II.3.3. Az ICH Guideline statisztikai modellje	27
II.3.4. Többfaktoros kovarianciaanalízis modellek	29
II.3.5. A matrixing és a bracketing módszer.....	29
II.4. Az ICH Guideline kritikája	30
II.4.1. A sarzsok egyesítésének kérdése	30
II.4.2. A szignifikancia-szint kérdése	31
II.4.3. A sarzs: rögzített (fix) vagy véletlen (random) faktor?.....	33
II.4.3.1. A rögzített és a véletlen faktoros megközelítés szemléletmódja a stabilitásvizsgálatnál.....	33
II.4.3.2. A véletlen jellegű sarzs faktor statisztikai kezelése	34
II.4.3.2.1. Statisztikai módszerek a sarzsok közti variancia vizsgálatára	36
II.4.3.2.2. Az eltarthatósági idő becslése véletlen faktoros megközelítés esetén.....	37
II.4.4. A bizonytalanság számszerűsítése: konfidencia-sáv, jóslási sáv vagy tolerancia-sáv?	38
II.4.4.1. Jóslási (predikciós) sáv	39
II.4.4.2. Tolerancia-sáv	39
II.5. A stabilitási adatok ingadozási forrásai	41
III. Kutatási munkám tématerületei	45
IV. A stabilitásvizsgálat statisztikai modelljében szereplő analízis-időpont	47
faktor újraértelmezése	47
IV.1. A probléma felvetése.....	47
IV.2. A Norwood-modell.....	48
IV.3. A Norwod-modell kritikája	49
IV.3.1. A sarzs faktor kérdése	49

IV.3.2. Az analízis-időpont faktor értelmezése	49
IV.4. A Norwood-modell továbbfejlesztése	51
IV.4.1. Az új modell szemléletmódja	51
IV.4.2. Az új modell matematikai alakja	52
IV.4.3. A két időskála	52
IV.4.4. A Norwood- és az új modell mátrix-egyenlete	53
IV.5. A Norwood- és az új modellel történő számítás menete	54
IV.5.1. A véletlen faktorok varianciakomponenseinek becslése	54
IV.5.2. A rögzített faktor hatásának becslése	56
IV.6. Példa	56
V. Az eltarthatósági idő becslése tolerancia-intervallumon alapuló, többfaktoros.....	60
modellel	60
V.1. A tolerancia-intervallum számításának alapjai	60
V.1.1. Egyoldali tolerancia-intervallum a normális eloszlású sokaság elemeinek nagy	60
részének elhelyezkedésére Owen klasszikus cikke alapján	60
V.1.2. Egyoldali tolerancia-tartomány lineáris regresszió esetén	62
V.2. A tolerancia-intervallum stabilitási adatokra való alkalmazásának irodalmi	64
előzményei	64
V.2.1. Chow és Shao kutatásai	64
V.2.2. Kiermeier eredményei	66
V.3. A fejezet célkitűzései	66
V.4. Tolerancia-intervallumos modellek a stabilitásvizsgálatra	67
V.4.1. Az eltarthatósági idő becslése egyetlen ingadozás-forrást tartalmazó modellre	68
(„ A_1 ” modell)	68
V.4.2. Az eltarthatósági idő becslése két ingadozás-forrást tartalmazó modellekre	68
V.4.2.1. Egy sarzs, a sarzson belül kétféle ingadozási forrás (inhomogenitás és analitikai hiba)	70
– a mért hatóanyag-tartalom elhelyezkedésének vizsgálata („ A_2 ” modell)	70
V.4.2.2. Egy sarzs, a sarzson belül kétféle ingadozási forrás (inhomogenitás és analitikai hiba)	71
– a valódi hatóanyag-tartalom elhelyezkedésének vizsgálata („ A_2^* ” modell)	71
V.5. Számolás a tolerancia-intervallumon alapuló stabilitásvizsgálati modellekkel	72
V.6. Szimulációs számítások	73
V.6.1. A szimuláció célja	73
V.6.2. A szimulációs terv leírása	74
V.6.3. A programozandó feladat rövid leírása	75
V.6.4. A szimulációs eredmények és értékelésük	76
V.6.4.1. A szimulációs program verifikálása	78
V.6.4.2. Az új modellek statisztikai tulajdonságainak vizsgálata	79
V.6.4.3. Különböző faktorok hatása a tolerancia-intervallum szélességére	82
V.6.4.3.1. A különböző varianciakomponens-kombinációk vizsgálata	82
V.6.4.3.2. A különböző ismétlésszám-kombinációk vizsgálata	83
V.7. Példa	84
VI. A stabilitási adatok bizonytalanságának előzetes becslése	88
VI.1. A probléma felvetése, célkitűzés	88
VI.2. A stabilitási adatok varianciakomponenseinek előzetes becslése	89
VI.2.1. Egyedi hatóanyag-tartalom meghatározás (Content Uniformity vizsgálat)	89
VI.2.2. Az analitikai módszer validálása	90
VI.2.2.1. Az azonos körülmények közötti ismételtetés vizsgálata (repeatability)	90
VI.2.2.2. A változó körülmények közötti ismételtetés vizsgálata	91
VI.2.2.2.1. A laboratóriumon belüli szórás vizsgálata (intermediate precision)	91
VI.2.2.2.2. A laboratóriumok közötti szórás vizsgálata (reproducibility)	91
VI.2.3. A varianciakomponens-becslés menetének összegzése	92
VI.3. A stabilitási adatok bizonytalanságának hatása a becsült eltarthatósági időre	93
VI.3.1. Egy jövőbeli stabilitási pont varianciájának becslése	93
VI.3.2. A becsült eltarthatósági idő bizonytalanságának számszerűsítése	94

VI.4. Példa	95
VI.4.1. Az analitikai módszer ismételtetésének becslése	95
VI.4.2. A laboratóriumon belüli szórás becslése	96
VI.4.3. A laboratóriumok közötti szórás vizsgálata	98
VI.4.4. Az inhomogenitási varianciakomponens becslése	98
VI.4.5. Egy jövőbeni stabilitási pont várható ingadozása	99
VI.4.6. A stabilitási adatok ingadozásának hatása az eltarthatósági idő becsült hosszára	100
VII. Összefoglalás	103
VIII. Új tudományos eredmények.....	106
IX. Az értekezés témájában megjelent publikációk	108
X. Irodalomjegyzék	109

KÖSZÖNETNYILVÁNÍTÁS

Ezen a helyen szeretnék köszönetet mondani mindazoknak, akik doktori munkám során segítségemre voltak.

Elsősorban köszönöm témavezetőm, Dr. Kemény Sándor szakmai irányítását és mindenre kiterjedő támogatását.

Köszönöm a BME Kémiai és Környezeti Folyamatmérnöki Tanszék korábbi és jelenlegi vezetőjének, Dr. Fonyó Zsoltnak és Dr. Mizsey Péternek, hogy lehetőséget biztosítottak doktori munkám elkészítésére. Külön köszönöm még a Tanszékről Dr. Deák András segítségét.

Köszönöm a Richter Gedeon Nyrt-nek doktori témám gyakorlati háttérének megismerési lehetőségét és a szakmai konzultációkra való nyitottságot. Külön köszönöm Dr. Ganzler Katalin segítségét.

Köszönöm Dr. Boda Krisztinának és Dr. Singer Júliának a számomra nem hozzáférhető SAS és SPSS szoftverrel készített számításait.

Köszönöm Bánfai Baláznak a MATHEMATICA szoftverben való programozásban nyújtott segítségét.

Köszönöm az OTKA (T 033005), a Varga József Alapítvány, a Richter Gedeon Centenárium Alapítvány és az Alapítvány a Klinikai Biostatistika Színvonalának Emeléséért anyagi támogatását.

Végül, de nem utolsósorban köszönöm családom - férjem, gyermekeim és szüleim - önzetlen támogatását.

I. BEVEZETÉS

I.1. Motiváció

A minőség szabályozás világméretű fejlődésében napjainkban egyre inkább teret hódítanak a statisztikai módszerek. Ennek oka abban keresendő, hogy a matematikai statisztika eszköztárának tudatos alkalmazásával idő- és költséghatékony módon javítható a termékek minősége. E tendencia alól a gyógyszeripar sem jelent kivételt, ahol az egyre élesedő piaci verseny ma már elengedhetlenné teszi a statisztikai módszerek alkalmazását.

Bármely gyógyszerkészítmény piacra kerülésének feltétele a törzskönyveztetés. A több száz oldalas törzskönyvi dokumentációban kiemelt helyet foglal el a gyógyszerkészítmény eltarthatósági idejének meghatározására a megfelelő validált vizsgálati módszerekkel végzett stabilitásvizsgálat részletes leírása és statisztikai elemzése. A gyógyszergyárak alapvető piaci érdeke, hogy készítményeik becsült eltarthatósági ideje minél pontosabb és a lehető leghosszabb legyen. E kívánalom eléréséhez a megfelelően választott és helyesen alkalmazott statisztikai módszerek kiváló segítséget nyújtanak.

A stabilitásvizsgálat statisztikai értékelésének jelenlegi, ICH Guideline által előírt módszerével már diplomamunkám keretében megismerkedtem. Az ott szerzett tapasztalatok, a témával foglalkozó irodalom ellentmondásai, valamint a gyógyszergyárak látható igénye adta az indítást a témakörben való elmélyülésre, és új, az eddigi eljárásoknál pontosabb és megbízhatóbb statisztikai módszerek kidolgozására.

Továbbá a doktori munka keretein részben már túlmutató célom, hogy a stabilitásvizsgálat statisztikai elemzésének példáján keresztül felhívjam a figyelmet ipari feladatok megoldásakor a problémafelvetés jelentőségére, a szakmailag releváns kérdés megtalálására, valamint az adekvát statisztikai módszerek alkalmazásának fontosságára és előnyére.

I.2. A doktori munka célkitűzései

Doktori értekezésem fő célkitűzése a stabilitásvizsgálat jelenlegi statisztikai módszerének kritikája és továbbfejlesztése az eltarthatósági idő valóságához és minél megbízhatóbb becslése érdekében.

A fenti általános célkitűzésen belül három konkrét célt fogalmaztam meg, mely egyben a disszertációban bemutatott kutatási munka tagolását is adta.

1. Elsőként a stabilitási adatokat terhelő reprodukálhatósági hiba valóságnak megfelelő figyelembevételét tűztem ki célul, mely a stabilitásvizsgálat statisztikai modelljében szereplő analízis-időpont faktor újradefiniálását jelentette.
2. Másodikként az eltarthatósági idő becslésénél alkalmazandó bizonytalansági tartomány számításánál a szakmai szempontból releváns kérdés megfogalmazására és megoldására törekedtem. Ez az ICH Guideline konfidencia-sávok szemlélete helyett egy tolerancia-intervallumon alapuló megközelítést jelent. A kutatás egyrészt a tolerancia-intervallumos modellbe bevonandó faktorok számának növelésére irányult, másrészt pedig arra, hogy mi legyen a tolerancia-intervallum középpontjában szereplő változó (a mért vagy a valódi hatóanyag-tartalom).
3. A harmadik témakörben - új problémafelvetésként - a stabilitási adatok ingadozásának előzetes becslésére tettem kísérletet. Ezzel az volt a célom, hogy már a stabilitásvizsgálat indítása előtt információnk legyen arról, hogy a jelenlegi gyártási és analitikai módszer bizonytalansága megfelelő-e (azaz nem túl nagy) a kívánt lejáratidő eléréséhez.

I.3. Számítási módszerek

A gyógyszerkészítmények stabilitásvizsgálatának statisztikai elemzése a matematikai statisztika számos részterülete közül a következő főbb témaköröket érinti: regresszióanalízis, varianciaanalízis (ANOVA), ill. e kettő kombinációjaként kapható általános lineáris modell (General linear model), valamint a különböző bizonytalansági intervallumok (konfidencia-, jóslási - és tolerancia-sáv) számítása.

A doktori munkám során végzett statisztikai számítások - ahol ezt külön nem jeleztem - a kutatócsoportban általánosan használt STATISTICA szoftver 6.1 verziójával készültek.

Néhány esetben a felmerülő statisztikai problémák megoldására más szoftverekre is szükségem volt. Az analízis-időpont faktor újradefiniálásakor kapott kevert (mixed) modell elemzéséhez a szoftverbe épített Visual Basic programnyelvet is alkalmaztam, valamint a kapott eredményeket összehasonlítottam más statisztikai szoftverek (SAS, SPSS) eredményeivel. A tolerancia-intervallumos stabilitásvizsgálati modellek statisztikai tulajdonságainak vizsgálatára végzett szimulációs számításokat a MATHEMATICA szoftver 6.0.0 verziójában írt programmal végeztem.

II. IRODALMI ÁTTEKINTÉS

II.1. Gyógyszerkészítmények törzskönyvezése

Bármely gyógyszerkészítmény piacra kerülésének előfeltétele a gyógyszer törzskönyveztetése. A gyógyszertörzskönyvezés a benyújtó gyógyszercég és az illetékes gyógyszerhatóság közötti szerződésnek tekinthető.¹ Egyik oldalról a benyújtó vállalja, hogy a gyógyszerre vonatkozó minden szakmai ismeretét átadja, és felelősséget is vállal, hogy ezek az adatok a legjobb tudása szerint helyesek, ill. az általa majdan szállítandó gyógyszer minden tétele ezeknek megfelel. A másik oldalról a gyógyszerhatóság vállalja, hogy ezen ismeretek birtokában a terméket gyógyszerre nyilvánítja, kihirdetve a forgalmazás és az alkalmazás szakmai feltételeit, és a továbbiakban hatósági jogkörében gondoskodik arról, hogy ezeket betartsák. A forgalomba hozatali engedély tehát arra hatáskörrel és illetékességgel rendelkező hatóság által kiadott, a gyógyszer embergyógyászati célra történő alkalmazhatóságát engedélyező és a törzskönyvi bejegyzést igazoló hatósági határozat.² A hatóság a készítményt akkor jegyzi be a gyógyszertörzskönyvbe, ha annak minősége - ideértve a gyártás körülményeit is - ismert és meghatározott, továbbá hatásossága klinikailag bizonyított és relatív ártalmatlansága igazolt. Egy új gyógyszerkészítmény bevezetéséhez, azaz a törzskönyveztetéséhez több száz vagy akár több ezer oldalas dokumentáció szükséges, melyben a gyógyszerre vonatkozó kémiai, gyógyszerészeti, biológiai, toxikológiai, hatástani (farmakológiai) és klinikai vizsgálatok menetét és eredményeit összegzik.

A törzskönyvi dokumentációban kiemelt helyet foglal el a gyógyszerkészítmény lejárati idejének meghatározására végzett stabilitásvizsgálat. Az ide vonatkozó fejezetnek tartalmaznia kell a stabilitásvizsgálati tervet (protokollt), a lejárati idő végéig érvényes minőségi követelményeket (specifikációkat), a mért stabilitási adatokat, ezek részletes statisztikai elemzését, valamint a becsült lejárati időt és az ajánlott tárolási körülményeket.

A gyógyszerkészítmények törzskönyvezésére, s ezen belül a stabilitásvizsgálatokra vonatkozó követelmények a világon mindenhol szigorodó tendenciát mutatnak. Az elmúlt évtizedben a gyógyszeriparban is megindult a jogharmonizáció, így a korábban ország- ill. régióspecifikus jogszabályok ma már viszonylag egységesek. A harmonizációs folyamatban kiemelkedő szerepet tölt be az ICH (International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human

Use), mely megfelelő fórumot kíván biztosítani az európai, amerikai és japán gyógyszerészeti hatóságok, valamint gyógyszeripari vállalatok közti konstruktív párbeszédhez, valamint a gyógyszeripar tudományos és technikai kérdéseinek megvitatásához. E fő cél érdekében az ICH közös megegyezésen alapuló ajánlásokat készít az eddig régióként különböző szakmai irányelvek és követelmények egységes megfogalmazásának és alkalmazásának, valamint kölcsönös elfogadásának harmonizálására. Az ICH a harmonizált előírásokat (guideline-okat) 4 fő tématerületre osztja:

- A *Quality* (minőség) témakör a kémiai és gyógyszerészeti minőségbiztosítást öleli fel.
- A *Safety* (biztonságosság) témakör az in-vitro és az in-vivo preklinikai vizsgálatokkal kapcsolatos előírásokat tartalmazza.
- Az *Efficacy* (hatásosság) témakör a klinikai (humán) vizsgálatok előírásait gyűjti egybe.
- A *Multidisciplinary* (multidiszciplináris) területre tartoznak azok a témák, amelyek vagy a fenti három témakör egyikébe sem férnek bele, vagy mindhármat érintik.

A gyógyszerkészítmények stabilitásvizsgálatára vonatkozó előírások a minőség témakör első alcsoportját alkotják. A stabilitásvizsgálat kivitelezéséhez és statisztikai értékeléséhez az alábbi két előírás ad támpontot, melyek tartalmát a további fejezetekben részletesen is ismertetem:

- ICH Q1A(R2): Stability Testing of New Drug Substances and Products (Second Revision)³
- ICH Q1E: Evaluation for Stability Data⁴

II.2. Gyógyszerkészítmények stabilitásvizsgálatának általános leírása

II.2.1. A stabilitásvizsgálat célja

Gyógyszerkészítmények stabilitásvizsgálatának célja annak megállapítása, hogy az idő múlásával a különböző környezeti tényezők (hőmérséklet, nedvességtartalom, fény stb.) hatására hogyan változik az adott gyógyszer minősége, és ezen ismeretek alapján az ajánlott tárolási körülmények és az eltarthatósági idő meghatározása³ (annak 1.3. fejezete). Itt jegyzem meg, hogy nemcsak gyógyszerkészítmények stabilitásáról beszélhetünk, hanem standard anyagokról is.^{5,6,7} A disszertációban bemutatott, gyógyszerkészítményekre vonatkozó stabilitásvizsgálati módszerek természetesen standard anyagok esetén is érvényesek.

II.2.2. A stabilitásvizsgálat típusai

Egy teljes stabilitásvizsgálati terv két különböző típusú vizsgálatot foglal magában: hosszúlejárátú (long-term) és gyorsított (accelerated) stabilitásvizsgálatokat. Mivel a különböző típusú vizsgálatok célja más-más, az adatok értékelése is eltérő az egyes vizsgálati típusoknál.

A hosszúlejárátú stabilitásvizsgálat célja azoknak a fizikai és kémiai változásoknak a modellezése, melyek a gyógyszer számára kedvező környezeti paraméterek közötti tárolása során játszódhatnak le. Az így kapott adatok közvetlenül felhasználhatók az adott gyógyszer eltarthatósági idejének becslésére, ezért is nevezik ezeket gyakran elsődleges stabilitási adatoknak. Doktori munkámban kizárólag a hosszúlejárátú stabilitásvizsgálat statisztikai értékelésével foglalkoztam.

A gyorsított stabilitásvizsgálatok célja a gyógyszerben végbemenő fizikai és kémiai változások felgyorsítása, melyet magasabb hőmérsékletű és/vagy nedvességtartalmú tárolás alkalmazásával érnek el. Az így kapott adatoknak két fő felhasználási területe van: egyrészt megfelelő összefüggések segítségével (pl. Arrhenius-egyenlet) már pár hónapos vizsgálati eredmények alapján becsülhetővé válnak a „normál” környezeti körülmények között hosszútávon bekövetkező fizikai és kémiai hatások; másrészt lehetőség nyílik a nem megfelelő szállítás vagy tárolás miatt esetlegesen fellépő, rövidtávú, extrém környezeti hatások modellezésére. A gyorsított stabilitásvizsgálatnak elsősorban a készítményfejlesztés kezdeti szakaszában van jelentősége, hiszen segítségével jóval hamarabb kaphatunk képet a gyógyszerben lejátszódó fizikai és kémiai változásokról. Az pedig, ha a gyorsított stabilitásvizsgálat során sem történik a gyógyszer vizsgált tulajdonságaiban jelentős változás, a hosszúlejárátú vizsgálatok eredményeinek alátámasztására szolgálhat. Fontos azonban megjegyezni, hogy az eltarthatósági idő becslése sohasem alapulhat pusztán gyorsított stabilitásvizsgálatokon, mindig szükség van hosszúlejárátú vizsgálatok indítására is. A gyorsított stabilitásvizsgálatok statisztikai értékelésének szintén kiterjedt irodalma van, melyek közül itt csak néhány könyvet/cikket említek meg.^{8,9,10,11,12}

Létezik ún. közbenső (intermediate) stabilitásvizsgálat is, mely jellege és funkciója szerint a gyorsított stabilitásvizsgálatok körébe tartozik. A közbenső stabilitásvizsgálatra vonatkozó tárolási paraméterek (nevéből is adódóan) a gyorsított és a hosszúlejárátú stabilitásvizsgálatok tárolási paraméterei között találhatóak. Elvégzése akkor kötelező, ha a gyorsított stabilitásvizsgálat során a gyógyszer valamely vizsgálandó paraméterénél

jelentős változást észleltek. Az ICH Q1A(R2) Guideline³ (annak 2.2.7.2. fejezete) az alábbi változásokat minősíti szakmai értelemben jelentősnek:

1. a hatóanyag-tartalom a kiindulási értékhez képest 5%-kal megváltozik,
2. bármely bomlástermék koncentrációja meghaladja az elfogadási kritériumát,
3. a fizikai vizsgálatok eredményei az elfogadási kritérium közelében vannak,
4. a mért pH érték az elfogadási kritérium közelében van,
5. a kioldódási érték az elfogadási kritérium közelében van.

II.2.3. Vizsgálandó tulajdonságok

A stabilitásvizsgálat során a gyógyszerkészítmény mindazon tulajdonságait vizsgálni kell, melyek a tárolás körülményei között megváltozhatnak, és így a gyógyszer minőségét, biztonságát és/vagy hatásosságát befolyásolhatják³ (annak 2.2.5. fejezete). Ennek megfelelően a stabilitásvizsgálat során az alábbi vizsgálatok elvégzése indokolt:

- Fizikai vizsgálatok (pl. küllem, geometriai méretek, átlagtömeg, egyedi tömeg)
- Gyógyszerkémiai vizsgálatok (hatóanyag-tartalom és bomlástermékek vizsgálata)
- Gyógyszerészeti vizsgálatok (kioldódás és szétesés vizsgálata)
- Mikrobiológiai vizsgálatok

Fontos megjegyezni, hogy a gyógyszerkémiai vizsgálatok során csak validált analitikai módszerek alkalmazhatók.

Doktori munkámban csak a gyógyszerkémiai vizsgálatok értékelésével, azon belül pedig főleg a hatóanyag-tartalom vizsgálatából származó stabilitási adatokkal foglalkoztam, így a dolgozatomban található példák többsége is ezekre vonatkozik. Természetesen a többi vizsgálandó tulajdonság tekintetében is hasonló megfontolások és következtetések érvényesek.

II.2.4. Tárolási körülmények

Az ICH Q1A(R2) Guideline³ (annak 2.2.7. fejezete) előírása szerint az egyes stabilitásvizsgálati típusokhoz általános esetben az 1. táblázatban látható tárolási körülmények tartoznak. A hosszúlejáratú stabilitásvizsgálatnál a gyógyszergyár döntheti el, hogy a két alternatíva közül melyik tárolási körülményt választja. Ha a hosszúlejáratú stabilitásvizsgálat a magasabb hőmérsékletű és nedvességtartalmú közegben zajlik, akkor értelemszerűen nincs közbenső stabilitásvizsgálat.

Stabilitásvizsgálat típusa	Tárolási körülmények	A vizsgálat minimális ideje
Hosszúlejáratú	25°C ± 2°C / 60 ± 5% RH vagy 30°C ± 2°C / 65 ± 5% RH	12 hónap
Közbenső	30°C ± 2°C / 65 ± 5% RH	6 hónap
Gyorsított	40°C ± 2°C / 75 ± 5% RH	6 hónap

1. táblázat: A stabilitási minták tárolási körülményei (általános eset)

Fontos hangsúlyozni, hogy a fenti táblázat nem a vizsgálatok tényleges hosszáról, hanem azok minimális időtartamáról rendelkezik. Így abban az esetben, ha a gyógyszer törzskönyvezésekor a hosszúlejáratú stabilitásvizsgálat időtartama még nem fedi le a javasolt eltarthatósági időt, akkor ennek eléréséig a vizsgálatot mindenképpen tovább kell folytatni. Abban az esetben pedig, ha a gyorsított stabilitásvizsgálat során felmerült problémák miatt közbenső stabilitásvizsgálatra is szükség van, akkor ennek előírt minimális időtartamát 6 hónapról 12 hónapra kell növelni.

Az 1. táblázatban ismertetett általános tárolási körülmények természetesen nem minden gyógyszer esetében alkalmazhatók. Indokolt esetben alternatív tárolási körülmények is megengedhetők (pl. ha a gyógyszert magasabb középhőmérsékletű országba szállítják). Néhány speciális esetről pedig, mint az olyan gyógyszerek, melyeket hűtőszekrényben vagy fagyasztószekrényben kell tárolni, ill. melyek csomagolása vízhatlan vagy szemipermeábilis, az ICH Q1A(R2) Guideline külön rendelkezik.

II.2.5. Vizsgálati időpontok

A vizsgálati időpontok gyakoriságára vonatkozó előírások azon alapulnak, hogy az adott gyógyszer bomlási folyamatát megfelelően nyomon lehessen követni.

Az ICH Q1A(R2) Guideline³ (annak 2.2.6. fejezete) előírása szerint mindazon gyógyszerekre, melyek várható eltarthatósági ideje minimum 12 hónap, a hosszúlejáratú stabilitásvizsgálat mintavételi időpontjai a következőképpen alakulnak: az első évben háromhavonta, a második évben félévente, majd ezután évente kell az előírt vizsgálatokat elvégezni. A gyorsított stabilitásvizsgálatoknál minimum három, míg a közbenső stabilitásvizsgálatoknál minimum négy mintavételi időpont (mindkét esetben beleértve a kezdeti és végpontokat is) javasolt.

II.2.6. A vizsgálandó sarzsok kiválasztása

Az ICH Q1A(R2) Guideline³ (annak 2.2.3. fejezete) rendelkezése szerint a stabilitásvizsgálat elvégzéséhez minimálisan három sarzsra (gyártási tételre) van szükség. A stabilitásvizsgálatra eltett sarzsoknak ugyanolyan gyártási technológiával kell készülniük és ugyanolyan csomagolásúaknak kell lenniük, mint amilyen a majdani piacra kerülő termék lesz. A három sarzs közül kettőnek legalább félüzemi léptékűnek kell lennie. A harmadik vizsgált sarzs mérete kisebb is lehet.

Itt említem még meg, hogy az eddig említett, a gyógyszerkészítmény törzskönyvezését megelőző ún. „preapproval” stabilitásvizsgálatok mellett léteznek a gyógyszer törzskönyvezése után indított ún. „postapproval” stabilitásvizsgálatok is, melyek célja a gyógyszerkészítmény monitorozása a gyógyszer dobozán feltüntetett lejáratú időpontig. E stabilitásvizsgálatok során rendszerint lényegesen több (10-20) gyártási sarzsot vizsgálnak, míg a mintavételi időpontok ritkábbak (általában csak évenkéntiek), mint a törzskönyvezését megelőző stabilitásvizsgálatoknál.

II.3. A stabilitásvizsgálat statisztikai értékelésének jelenlegi módszere

II.3.1. Alapelvek (egyetlen sarzs vizsgálata)

II.3.1.1. A statisztikai értékelés menete

Az ICH Q1E Guideline⁴ „A” mellékletében részletes folyamatábra található a stabilitásvizsgálat statisztikai értékelésének menetére, melyen minden vizsgálandó tulajdonság (hatóanyag-tartalom változása, bomlástermékek koncentrációjának változása stb.) esetén végig kell haladni.

Az első lépés a gyorsított stabilitásvizsgálat értékelése, mely során arra a kérdésre keresünk választ, hogy a magasabb hőmérsékletű és/vagy nedvességtartalmú tárolás során a vizsgált tulajdonságban történt-e szakmai értelemben jelentős változás. Ha a gyorsított stabilitásvizsgálat során nem tapasztalunk jelentős mértékű változást, akkor attól függően különböző irányokban léphetünk tovább, hogy a hosszúlejáratú ill. a gyorsított stabilitásvizsgálat során mért értékek az idő múlásával milyen mértékű változást ill. mekkora mértékű ingadozást mutatnak. Ha egyik tárolásnál sem mutatható ki változás ill. ingadozás, vagy a változás ill. ingadozás mértéke kicsi, akkor nincs szükség további

statisztikai elemzésre. Ekkor a gyógyszerkészítmény javasolt eltarthatósági idejének maximuma a hosszúlejáratú stabilitásvizsgálat időtartamának kétszerese (mely azonban nem lépheti túl a vizsgált időtartam + 12 hónapot) adható meg.

Ha a hosszúlejáratú vagy a gyorsított stabilitásvizsgálat során az idő múlásával nagyobb mértékű változás mutatható ki, ill. a stabilitási adatok ingadozása jelentős, akkor további statisztikai elemzésre van szükség. A statisztikai elemzés célja a hosszúlejáratú stabilitásvizsgálatból nyert adatok alapján a gyógyszerkészítmény degradálódási profiljának meghatározása, majd az ICH Guideline definíciója szerinti eltarthatósági idő (ld. később) meghatározása.

A mért stabilitási adatok statisztikai elemzése alapján interpolációval vagy extrapolációval adhatunk becslést a várható eltarthatósági időre. Az extrapoláció során feltételezzük, hogy a gyógyszerkészítményben a tárolás során lezajló változás jellege a vizsgálati időszakon túl is az előzőekhez hasonlóan alakul. Mivel azonban ebben sohasem lehetünk teljesen biztosak, az ICH Q1E Guideline⁴ csak limitált extrapolációt engedélyez. Ez azt jelenti, hogy a javasolt eltarthatósági idő még abban az esetben is csak maximum a hosszúlejáratú stabilitásvizsgálat időtartamának kétszerese (mely azonban nem lépheti túl a vizsgált időtartam + 12 hónapot) lehet, ha a statisztikai elemzés alapján becsült eltarthatósági idő ennél jóval hosszabb. Ha a becsült eltarthatósági idő meghaladja az eddig elvégzett stabilitásvizsgálat vizsgálati időtartamát, akkor a hosszúlejáratú stabilitásvizsgálatot a becsült eltarthatósági időig mindenképpen folytatni kell, hogy az extrapoláció jogosságát igazoljuk.

Ha a gyorsított stabilitásvizsgálat során szakmai értelemben jelentős változást tapasztalunk, akkor a közbenső stabilitásvizsgálat értékelése is szükségessé válik. Ha a közbenső stabilitásvizsgálat során nem tapasztalunk jelentős változást, akkor az előbb említett statisztikai elemzéseket kell elvégeznünk. Ha viszont a közbenső stabilitásvizsgálat során is jelentős változást tapasztalunk, akkor az eltarthatósági idő becsülésénél nem alkalmazhatunk extrapolációt.

II.3.1.2. Az eltarthatósági idő definíciója és meghatározása

Az ICH Q1A(R2) Guideline³ szójegyzékében található definíció szerint egy adott gyógyszerkészítmény eltarthatósági idejének azt az időtartamot tekintjük, mely alatt a

gyógyszer - feltételezve a csomagoláson feltüntetett tárolási körülményeket - a rá vonatkozó elfogadási (specifikációs) kritériumoknak megfelel.

A gyógyszerkészítmények eltarthatósági idejének meghatározására az ICH Q1E Guideline⁴ (annak 2.6. fejezete) az alábbi módszert ajánlja:

- azoknál a kvantitatív tulajdonságoknál, amelyek az idő múlásával vélhetően csökkenni fognak (pl. hatóanyag-tartalom), eltarthatósági időnek azt az időpontot tekintjük, ahol az átlagos degradálódási görbe alsó 95%-os egyoldali konfidencia-sávja metszi az adott tulajdonságra vonatkozó elfogadási kritériumot,
- azoknál a kvantitatív tulajdonságoknál, melyek az idő múlásával vélhetően növekedni fognak (pl. bomlástermékek koncentrációja), eltarthatósági időnek azt az időpontot tekintjük, ahol az átlagos degradálódási görbe felső 95%-os egyoldali konfidencia-sávja metszi az adott tulajdonságra vonatkozó elfogadási kritériumot,
- azoknál a kvantitatív tulajdonságoknál, melyekről előre nem tudjuk, hogy az idő múlásával csökkenni vagy növekedni fognak, eltarthatósági időnek azt az időpontot tekintjük, ahol az átlagos degradálódási görbe 95%-os kétoldali konfidencia-sávja metszi az adott tulajdonságra vonatkozó alsó ill. felső elfogadási kritériumot.

A fenti meghatározások közös jellemzője, hogy az eltarthatósági idő a konfidencia-sáv és a specifikációs határ metszéspontjaként számítódik. A konfidencia-sáv alkalmazásának elméleti háttérét, számítási módját és tulajdonságait a II.3.1.4. fejezetben részletesen ismertetem.

Fontos még tisztázni, hogy a külföldi szakirodalomban három eltarthatósági idő „fogalom” is létezik. Megfelelő magyar nevezéktan híján lefordítottam őket, de az egyértelműség kedvéért zárójelben megadtam a nemzetközileg elfogadott angol elnevezéseket is.

- *elméleti eltarthatósági idő (true shelf-life):*
Az az időpont, ahol a gyógyszer tényleges hatóanyag-tartalma pontosan megegyezik az előírt specifikációs határral, azaz ahol az elméleti degradációs függvény metszi az előírt elfogadási kritériumot.
- *elméleti eltarthatósági idő becslése (estimated shelf-life):*
A degradációs függvény alapján becsült eltarthatósági idő, azaz ahol a mért adatokra illesztett degradációs függvény metszi az előírt specifikációs határt. Ezt az eltarthatósági időt csak abban az értelemben nevezhetjük becsültnek,

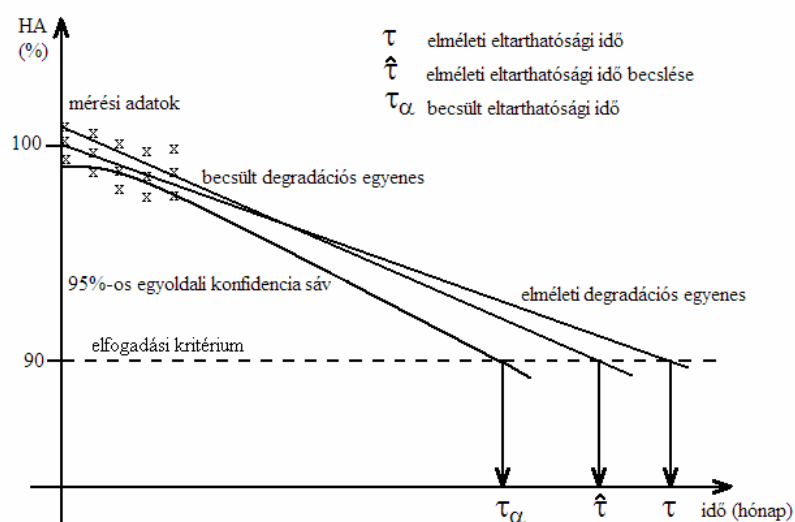
hogy alkalmazásakor a számunkra ismeretlen elméleti degradációs függvényt a rendelkezésünkre álló stabilitási adatok alapján becsüljük. Azonban az így kapott eltarthatósági idő nem megfelelő becslése az elméleti eltarthatósági időnek, mivel nem veszi figyelembe a degradációs függvény becslésének bizonytalanságát. Ez a kategória tehát a gyakorlatban használhatatlan, így a továbbiakban nem foglalkozom vele.

- *becsült eltarthatósági idő (label shelf-life):*

Az ICH Guideline előírásai alapján a degradációs függvény konfidencia-sávjának segítségével becsült eltarthatósági idő, azaz ahol az illesztett degradációs függvény 95%-os konfidencia-sávja metszi az előírt specifikációs határt. Ez az a lejáratási idő, melyet a gyógyszerkészítmény csomagolásán fel kell tüntetni.

Ez az eltarthatósági idő már figyelembe veszi a stabilitási adatok bizonytalanságát, így biztosítva a beteget arról, hogy a gyógyszer a lejáratási időpontig megfelel az előírt specifikációs követelményeknek.

Az alábbi ábra szemléletesen mutatja az eltarthatósági idő becslésének alapelvét - feltüntetve a három megkülönböztetett eltarthatósági idő kategóriát - hatóanyag-tartalom adatok esetén, lineáris degradálódási függvényt feltételezve.



1. ábra: Az eltarthatósági idő becslése az ICH Q1E Guideline⁴ alapján

II.3.1.3. A degradálódási függvény jellegének meghatározása

Az eltarthatósági idő becslésének első lépése a degradálódási függvény meghatározása. A vizsgálandó tulajdonságok egy részénél (pl. hatóanyag-tartalom, bomlástermékek koncentrációja) a kinetika ismert lehet. Például a hatóanyag-tartalom csökkenése általában nullad- vagy elsőrendű reakció szerint zajlik. Kinetikai vizsgálatokkal igazolható, hogy 15%-os bomlásig a két modellel számított eltarthatósági idő között csekély a különbség¹³ (p. 33-35). Mivel a hatóanyag-tartalomra vonatkozó alsó tűréshatár általában 90-95% között van, azaz a maximálisan megengedhető bomlás mindössze 5-10%, így a statisztikai értékeléshez az esetek többségében használható a lényegesen egyszerűbben kezelhető nulladrendű reakció (azaz a lineáris modell)¹⁴ (p. 388). Természetesen a választott degradálódási függvény jóságát megfelelő statisztikai módszerekkel mindig igazolni kell.

Olyan kvantitatív tulajdonságoknál (pl. kioldódási adatok), ahol a kinetika nem ismert, az adatok leírását az egyszerűbb statisztikai kezelés végett célszerű mindig a lineáris modellel kezdeni. Csak akkor lépünk tovább a bonyolultabb négyzetes, exponenciális vagy logaritmusos függvények felé, ha a lineáris függvény jósága statisztikai módszerekkel nem igazolható. Carstensen és társai¹⁵ részletesen is foglalkoznak a pH értékeket és a kioldódási adatokat leíró nullad- és elsőrendű reakciók kinetikájával.

Doktori dolgozatomban kizárólag lineáris degradálódási függvényekkel foglalkozom.

II.3.1.4. A konfidencia-sáv fogalma és tulajdonságai

A mért stabilitási adatok a degradációs függvény (egyenes) körül különböző bizonytalansági források (tabletták inhomogenitása, analitikai mérési hiba stb.) miatt ingadozhatnak. A stabilitásvizsgálat statisztikai értékelésekor a valóság ismeretének ezt a pontatlanságát az illesztett egyenes körüli bizonytalansági tartomány számításával vesszük figyelembe.

Az ICH Q1E Guideline⁴ (annak 2.6. fejezete) előírása alapján ez a bizonytalansági tartomány a konfidencia-sáv. A konfidencia-sáv elképzelés azon a feltételezésen alapul, hogy létezik a gyógyszer igazi összetételére vonatkozó igazi degradálódási függvény, melyet azonban a mért adatok bizonytalansága miatt csak pontatlanul tudunk becsülni. E koncepció szerint az eltarthatósági idő az az időpont, amelyenél az igazi degradálódási egyenes metszi a specifikációs határt. Ismeretünk pontatlansága miatt azonban ezt az időpontot csak bizonytalanul tudjuk megadni, emiatt azt mondjuk, s ez lesz a becsült

eltarthatósági idő, hogy a valódi hatóanyag-tartalom adott időpontbeli konfidencia-tartományának kell elérnie a specifikációs határt. A konfidencia-sáv alkalmazása adja meg tehát azt a biztonságot, hogy a becsült eltarthatósági idő kellő biztonsággal kisebb lesz az elméleti eltarthatósági időnél.

A konfidencia-sáv az a tartomány, amely a valódi degradálódási függvény (jelen esetben egyenes) pontjait $1-\alpha$ valószínűséggel tartalmazza. Lineáris függvényre az egyoldali konfidencia-sáv határpontjai egy adott x helyhez (időponthoz) az alábbi összefüggéssel határozhatók meg:

$$P\left(\hat{y} - t_{\alpha} s_y \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_j (x_j - \bar{x})^2}} < y\right) = 1 - \alpha \quad (1)$$

ahol \hat{y} az adott x helyhez (időponthoz) az illesztett degradációs függvényből számítható érték

t_{α} adott α -hoz és az $n-2$ szabadsági fokhoz tartozó t -érték

s_y a stabilitási adatok szórása

n a mérési pontok száma

x_j mérési időpontok

\bar{x} a mérési időpontok átlaga

A gyógyszergyárak érdeke, hogy a stabilitásvizsgálat során mért adatok alapján becsült eltarthatósági idő a lehető leghosszabb legyen. A fenti képlet szerint - adott meredekségű degradációs egyenes esetén - az eltarthatósági idő annál hosszabb, minél szűkebb a degradálódási egyenes konfidencia-sávja. A konfidencia-sáv szélessége, s így a becsült eltarthatósági idő függ:

- a stabilitási adatok ingadozásától (s_y)
Ennek megfelelően a stabilitási adatok ingadozását - a gyártási folyamat javításával (pl. minél homogénebb termék előállításával) és a lehető legpontosabb analitikai mérési módszer kidolgozásával - amennyire csak lehet, csökkenteni kell.
- a mérési adatok számától (n)
Eszerint célszerű a lehető legtöbb pontban, a lehető legnagyobb ismétlésszámmal méréseket végezni, mert ez csökkenti a gyökjel alatti kifejezés értékét. Szintén n

értékét növeli, ha - amennyiben a később tárgyalandó feltételek megengedik - a statisztikai értékeléshez a vizsgált sarzsok adatait egyesítjük.

- az x_j értékek elhelyezkedésétől

Lineáris degradálódási függvény esetén a mérési időskála két szélén, vagyis a stabilitásvizsgálat kezdetén és végén sok ismételt mérést végezve a $\sum_j (x_j - \bar{x})^2$

nevező értéke növelhető, s így a konfidencia-sáv szűkíthető. Célszerű lenne tehát, ha az általánosan bevett gyakorlat helyett - miszerint minden mérési időpontban ugyanannyi ismételt mérést végeznek - a rendelkezésre álló mintaszámot a vizsgálati időpontok között súlyozottan elosztani.

II.3.2. Több sarzs együttes vizsgálata

Az ICH Q1E Guideline⁴ (annak 2.1. fejezete) előírása szerint a becsült eltarthatósági időnek az összes jövőbeli, hasonló körülmények között gyártandó sarzsra érvényesnek kell lennie. A Guideline elképzelése alapján a vizsgált sarzsok közti természetesen ingadozás biztosítékot nyújt arra, hogy a jövőbeli sarzsok eltarthatósági ideje is a becsült lejárati időn belül maradjon.

Statisztikai szempontból a sarzsok közti egyezőség ill. különbözőség vizsgálatához minél több sarzsot tartalmazó mintára lenne szükség. Nyilvánvaló, hogy egyetlen sarzs vizsgálata alapján a sarzsok közti különbözőség nem értelmezhető, míg két sarzs vizsgálata esetén az egyezőségre vagy különbözőségekre vonatkozó következtetés igen megbízhatatlan lenne. A sarzsszám további növelésével a következtetés pontossága és megbízhatósága fokozatosan javul, ám a végletekig tartó sarzsszám-növelésnek a gazdasági okok és az időbeli korlátok hamar gátat szabnak. E kettősség fényében érthető, hogy az ICH Guideline minimálisan három sarzs vizsgálatára irányuló előírása statisztikai és gazdasági megfontolások kompromisszumaként született¹⁶ (p. 441).

II.3.2.1. A sarzsok adatainak egyesítése

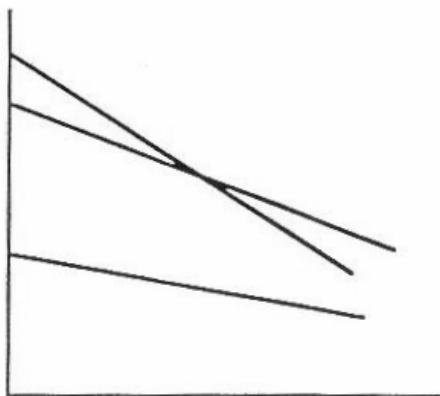
Több sarzs vizsgálata esetén első kérdésünk, hogy a sarzsok adatai kezelhetők-e együtt, azaz illeszthető-e egyetlen közös degradálódási egyenes a több sarzsból származó valamennyi adatra. A stabilitásvizsgálatra eltett sarzsok adatainak egyesítése – a sarzsok

közi ingadozás implicit (szórást okozó) figyelembevételével – azt az előnyt is magában rejti, hogy nő a statisztikai értékeléshez felhasznált adatok száma, ami (azonos meredekségű degradációs egyenes esetén) szűkebb konfidencia-sávot, s így hosszabb becslési eltarthatósági időt eredményez.

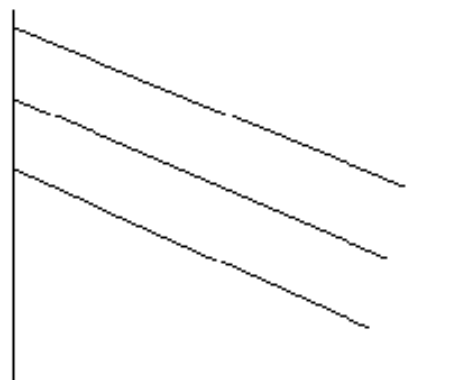
A vizsgált sarzsok egyesítésére akkor van lehetőség, ha a sarzsok közti ingadozás kismértékű. A sarzsok egyesítésének feltétele, hogy a vizsgált sarzsok degradálódási profilja hasonló legyen, azaz az egyes sarzsokra illesztett degradálódási görbe (egyenes) azonos legyen (szignifikánsan ne különbözzön egymástól). A vizsgált sarzsokra illesztett egyenesek akkor tekinthetők azonosaknak, ha paramétereik - meredekségük és tengelymetszetük - azonosak.

A sarzsok degradációs egyeneseinek különbözősége szerint a 2. ábrán szemléltetett négy eset fordulhat elő.

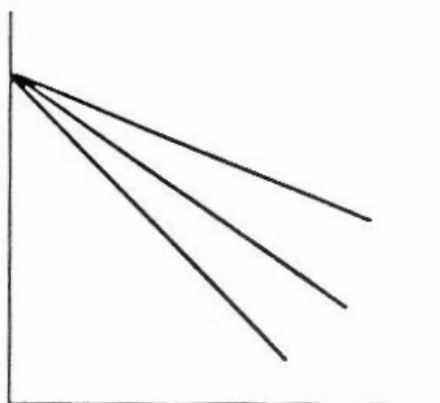
D. Mind a meredekség,
mind a tengelymetszet különböző



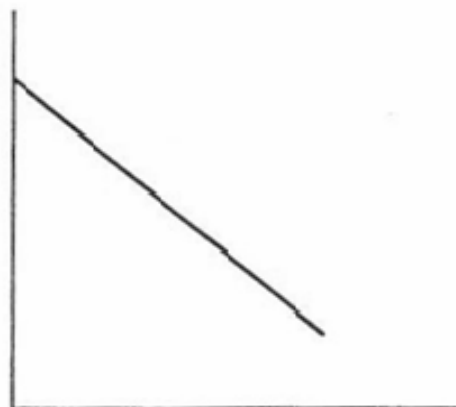
C. Azonos meredekség



B. Azonos tengelymetszet



A. Azonos meredekség és
tengelymetszet



2. ábra: Több sarzs együttes vizsgálatakor kapható regressziós egyenesek¹⁶

A 2. ábrán látható esetekhez az alábbi modellek tartoznak¹⁶.

A kiindulási (*D*) modellnél az adatokra sarzsónként különböző meredekségű és tengelymetszetű egyenest illesztünk:

$$y_{ijk} = \alpha_i + \beta_i x_{ij} + \varepsilon_{ijk} \quad (2)$$

ahol y_{ijk} az *i*-edik sarzs mért hatóanyag-tartalma a *j*-edik vizsgálati időpont *k*-adik mintájának mérésekor

α_i az *i*-edik sarzsra illesztett egyenes tengelymetszete (kiindulási hatóanyag-tartalom)

β_i az *i*-edik sarzsra illesztett egyenes meredeksége (degradálódási ráta)

x_{ij} az *i*-edik sarzs *j*-edik vizsgálati időpontja

ε_{ijk} véletlen hiba

Abban az esetben, ha csak a meredekségek azonosságára vonatkozó nullhipotézis teljesül, a *C* modellel számolunk:

$$y_{ijk} = \alpha_i + \beta x_{ij} + \varepsilon_{ijk} \quad (3)$$

ahol α_i az *i*-edik sarzsra illesztett egyenes tengelymetszete

β az összes adatra illesztett párhuzamos egyenesek közös meredeksége

Abban az esetben, ha csak a tengelymetszetek azonosságára vonatkozó nullhipotézis teljesül, a *B* modellel számolunk:

$$y_{ijk} = \alpha + \beta_i x_{ij} + \varepsilon_{ijk} \quad (4)$$

ahol α az összes adatra illesztett egyenesek közös tengelymetszete

β_i az *i*-edik sarzsra illesztett egyenes meredeksége

Ha pedig mind a meredekségek, mind a tengelymetszetek azonosságát elfogadjuk, akkor a még egyszerűbb (*A*) modellt kapjuk:

$$y_{ijk} = \alpha + \beta x_{ij} + \varepsilon_{ijk} \quad (5)$$

ahol α az egyesített sarzsokra illesztett közös egyenes tengelymetszete

β az egyesített sarzsokra illesztett közös egyenes meredeksége

Megjegyzem, hogy a *B* modellt a gyakorlatban nem szokás használni, helyette a *D* modellt alkalmazzák.

II.3.2.2. Az egyesítés feltételeinek ellenőrzése

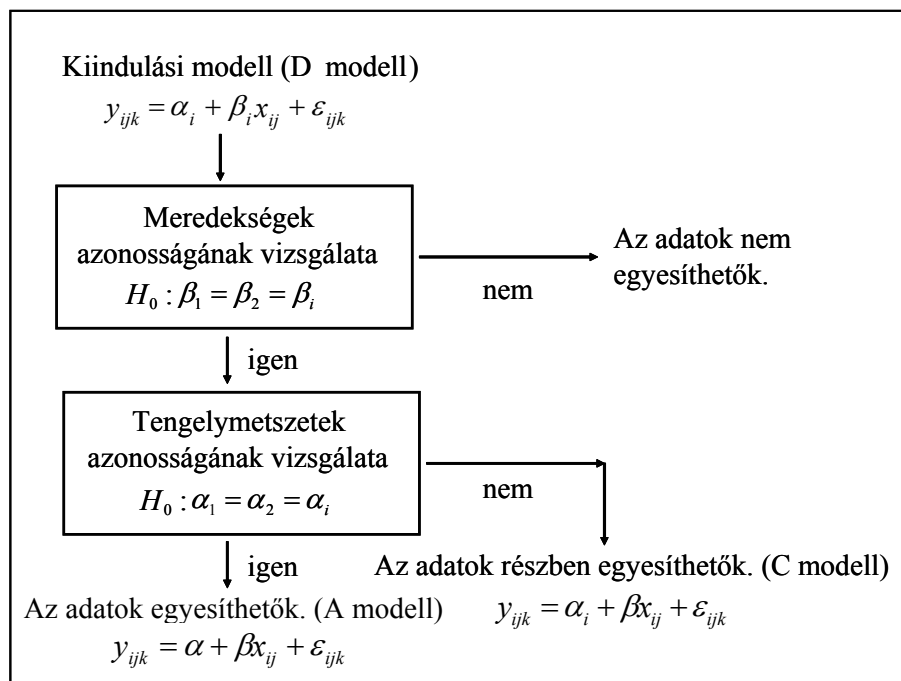
Az ICH Q1E Guideline⁴ (annak B.2.2. fejezete) az egyesítés feltételeinek ellenőrzésére szekvenciális (lépcsőzetes) vizsgálatot javasol, melyben először a meredekségek, majd a tengelymetszetek azonosságát teszteljük.

A vizsgálatok eredményei alapján három esetet különíthetünk el:

1. Ha a meredekségek azonosságára vonatkozó nullhipotézist elutasítjuk, akkor a sarzsok adatait nem egyesíthetjük. Ilyenkor a *D* modell érvényes, azaz minden egyes sarzsra külön-külön ki kell számítani az eltarthatósági időt, sarzsonként egyedi meredekséget és tengelymetszetet alkalmazva. Az így kapott eltarthatósági idők közül a legrövidebb lesz az összes sarzsra elfogadott becsült eltarthatósági idő („minimum approach”).
2. Ha a meredekségek azonosságát elfogadtuk, de a tengelymetszetek azonosságára vonatkozó nullhipotézist elutasítottuk, akkor az adatokat csak a közös meredekség becslésére egyesíthetjük. Ebben az esetben tehát a *C* modell érvényes. Ezután minden egyes sarzsra külön-külön ki kell számítani az eltarthatósági időt, a becsült közös meredekséget és az egyedi tengelymetszetet, valamint a közös reziduális szórásnégyzetet alkalmazva. A kapott eltarthatósági idők közül ismét a legrövidebb lesz az összes sarzsra elfogadott becsült eltarthatósági idő.
Megjegyzendő, hogy a gyógyszergyárak számára ez az eset már kedvezőbb az előzőnél, mivel a degradálódási egyenes meredekségét több adatból pontosabban becsülve szűkebb konfidencia-sávot, s így hosszabb becsült eltarthatósági időt kapunk.
3. Ha mind a meredekségek, mind a tengelymetszetek azonosságát elfogadtuk, a sarzsok adatai egyesíthetők. Ekkor az *A* modell lesz érvényes, azaz az összes adatra egyetlen egyenest illesztve becsülhetjük az eltarthatósági időt.

A gyógyszergyárak számára nyilvánvalóan ez az eset a legkedvezőbb, hiszen mind a meredekség, mind a tengelymetszet becsléséhez felhasználható lényegesen több mérési adat a konfidencia-sáv további szűkülését, s még hosszabb becsült eltarthatósági időt eredményez adott degradálódási ráta esetén.

A szekvenciális vizsgálat egyes lépéseit összefoglalva mutatja a 3. ábra.



3. ábra: A lépcsőzetes vizsgálat menete

II.3.2.3. Az egyesíthetőség vizsgálata során használt statisztikai próbák

A sarzsok egyesíthetőségére végzett statisztikai próbákat az általános lineáris modellel (general linear model) végezhetjük el, mely a varianciaanalízis és a regresszióanalízis kombinációjának tekinthető.¹⁷ Az alábbiakban ennek - az irodalomban általánosan elfogadott - menetét mutatom be Lin és társai könyvfejezete¹⁶ alapján. Megemlítem, hogy Chow és Liu által a gyógyszeripari statisztikai problémákról írt összefoglaló jellegű könyve¹⁴ a 11.3. fejezetében szintén részletesen ismerteti ezt a témakört.

A 2. táblázat összegzi a szekvenciális vizsgálat során számítandó F -próbat statisztikákat.

Nullhipotézis	Alternatív hipotézis	SS	df	MS	F
Közös meredekség és közös tengelymetszet (A)	Közös meredekség és egyedi tengelymetszet (C)	$SS_{AC} = SS_A - SS_C$	$r - 1$	SS_{AC} / df_{AC}	MS_{AC} / MS_D
Közös meredekség és egyedi tengelymetszet (C)	Egyedi meredekség és egyedi tengelymetszet (D)	$SS_{CD} = SS_C - SS_D$	$r - 1$	SS_{CD} / df_{CD}	MS_{CD} / MS_D
Egyedi meredekség és egyedi tengelymetszet (D)		SS_D	$N - 2r$	SS_D / df_D	

Sarzs: $i = 1, \dots, r$ Tárolási idő: $j = 1, \dots, n_i$

2. táblázat: A sarzsok egyesíthetőségére végzett statisztikai számítások

A táblázatban szereplő négyzetösszegek képlete és jelentése az egyes modellekből következnek:

$$SS_D = \sum_{i=1}^r \sum_{j=1}^{n_i} (y_{ij} - \hat{\alpha}_i - \hat{\beta}_i x_{ij})^2,$$

azaz az egyes sarzsokra külön-külön egyeneseket illesztve kapott reziduális négyzetösszeg

$$SS_C = \sum_{i=1}^r \sum_{j=1}^{n_i} (y_{ij} - \hat{\alpha}_i - \hat{\beta} x_{ij})^2$$

azaz a sarzsok adataira azonos meredekségű, de különböző tengelymetszetű egyeneseket illesztve kapott reziduális négyzetösszeg

$$SS_A = \sum_{i=1}^r \sum_{j=1}^{n_i} (y_{ij} - \hat{\alpha} - \hat{\beta} x_{ij})^2$$

azaz az összes adatra egy közös egyenest illesztve kapott reziduális négyzetösszeg

II.3.2.3.1. A meredekségek azonosságának vizsgálata

A meredekségek azonosságának vizsgálatánál arra a kérdésre keressük a választ, hogy leírhatók-e a különböző sarzsból származó adatok azonos meredekségű egyenessel. Ehhez a *közös meredekséggel*, de különböző tengelymetszettel felírt (C) modell és az *egyedi meredekséget* és tengelymetszetet alkalmazó (D) modell reziduális négyzetösszegének különbségéből képzett szórásnégyzetet kell a D modell szórásnégyzetéhez viszonyítani, vagyis a 2. táblázat utolsó két sorát kell összevetni.

A fenti számítás statisztikai programcsomagokkal könnyen kivitelezhető. Az általam használt STATISTICA programban¹⁸ a *General Linear Models (GLM)* modul *Homogeneity-of-slopes* opcióját kell alkalmazni. A vizsgált modell ekkor a következő:

$$y_{ijk} = \alpha_i + \beta x_{ij} + (\alpha\beta)_i x_{ij} + \varepsilon_{ijk}$$

A modell alkalmazásával arra a kérdésünkre kaphatunk választ, hogy van-e kölcsönhatás a folytonos (tárolási idő) és a kategorikus (sarzs) változó között, azaz az egyenesek meredeksége azonosnak tekinthető-e a kategorikus változó (sarzs faktor) mindegyik szintjén. A vizsgált nullhipotézis: $(\alpha\beta)_i = 0$, melyet úgy is megfogalmazhatunk, hogy van-e különbség az adatok leírásában, ha sarzsonként külön egyeneseket veszünk ill. a párhuzamos egyenesek modellje között. Innen ered a párhuzamosságvizsgálat elnevezés is.

II.3.2.3.2. *A tengelymetszetek azonosságának vizsgálata*

Az ICH Q1E Guideline⁴ előírása szerint a tengelymetszetek azonosságának vizsgálata csak a meredekségek azonosságára vonatkozó nullhipotézis elfogadása után következik. Ha a meredekségek azonosságát elutasítjuk, akkor a tengelymetszetek azonosságának vizsgálatára már nincs is szükség, mivel a sarzsok adatai semmiképpen sem egyesíthetők. Mindez azt jelenti, hogy a két statisztikai próba meghatározott (szekvenciális) sorrendben követi egymást.

E szekvenciális szemlélet ellentettje a szimultán vizsgálat lenne, melyben az egyenesek meredekségének és tengelymetszetének azonosságát egyszerre (egy lépésben) ellenőrizzük, azaz tengelymetszetek azonosságának vizsgálatakor nem tételezzük fel előzetesen az egyenesek párhuzamosságát.

Ha az ICH Guideline szekvenciális szemléletét követjük, akkor a tengelymetszetek azonosságának vizsgálatakor arra a kérdésre keressük a választ, hogy a párhuzamos egyeneseknek a tengelymetszete is azonosnak tekinthető-e, azaz a különböző sarzsokból származó adatok leírhatók-e egyetlen közös egyenessel. Ennek a gondolatmenetnek az a számítási módszer felel meg, mely a közös meredekséget és *közös tengelymetszetet* alkalmazó *A* modellel és a közös meredekséggel, de *különböző tengelymetszettel* számoló *C* modellel kapott reziduális négyzetösszegek különbségéből képzett szórásnégyzetet hasonlítja a *C* modell reziduális szórásnégyzetéhez.

Ehhez képest Lin és társai¹⁶ (ahogy a 2. táblázat első sorából láthatjuk) az *F*-próbastatisztikai nevezőjében a *D* modell, vagyis az egyedi meredekséget és tengelymetszetet alkalmazó modell szórásnégyzetét szerepeltetik. Ez pedig azt jelenti, hogy az irodalomban általánosan alkalmazott számítási módszer nem a szekvenciális, hanem a szimultán vizsgálatnak feleltethető meg.

A tengelymetszetek azonosságára végezhető szekvenciális és szimultán vizsgálat közti különbség tehát egyrészt a nullhipotézisek különbözőségében, másrészt az *F*-próba számításában (annak nevezőjében szereplő más-más szórásnégyzet értékben) nyilvánul meg. A szekvenciális vizsgálat esetén a kérdésünk az, hogy párhuzamos (azaz azonos meredekségű) egyeneseket feltételezve van-e különbség a sarzsónként különböző tengelymetszetet tartalmazó modell és a közös tengelymetszettel számoló modell között. Ezzel szemben a szimultán vizsgálatkor azt kérdezzük, hogy a teljes (azaz a különböző meredekséget és tengelymetszetet alkalmazó) modellt alapul véve a sarzsónkénti

különböző tengelymetszet figyelembe vétele jobb-e, mint ha ezt (és csak is ezt) a tagot kihagyjuk a modellből, s így azonos tengelymetszetű (de akár különböző meredekségű) egyenesekkel írjuk le az adatokat.

Megjegyzendő, hogy az ICH Guideline szövegéhez igazodó, de az irodalmi forrásokban nem alkalmazott szekvenciális számítási módszernél a tengelymetszetek azonosságának vizsgálatához alkalmazott próbastatisztika nevezőjében szereplő „error” szórásnégyzet (MS_C) szabadsági foka (df_C) mindig nagyobb, mint a D modellnél használt df_D szabadsági fok, hiszen a különböző (de szignifikánsnak nem minősült) meredekségek okozta ingadozás is a nevezőben szereplő „error” négyzetösszegben jelentkezik ($SS_C = SS_D + SS_{CD}$). A nagyobb szabadsági fokú szórásnégyzet alkalmazása pedig növeli a becslés megbízhatóságát.

Az ICH Guideline-ban említett vizsgálati sorrend („proper order”) érthető úgy is, hogy a tengelymetszetek azonossága csak akkor érdekel minket, ha a meredekségek azonosnak bizonyultak, azaz azt az esetet nem kezeljük külön, amikor a tengelymetszetek azonosak, de a meredekségek nem, akkor végül is elfogadható (vagy legalább is nem helytelen) az a számolási mód, amit az említett irodalmi források alkalmaznak. A kérdés csak az (s erre világít rá a bemutatott másik számolási módszer), hogy a tengelymetszetek azonosságának vizsgálatához miért nem használják fel azt az információt, amit a meredekségek azonosságának elfogadásából nyertek (nagyobb szabadsági fokú becslés).

Statisztikai programcsomagokkal mind a szekvenciális, mind a szimultán számítási mód könnyen kivitelezhető. A tengelymetszetek azonosságának vizsgálata a statisztikai programok nyelvén úgy fogalmazható meg, hogy van-e hatása a kategorikus változónak (vagyis a sarzsak). Ha a vizsgálatot szekvenciális módon kívánjuk elvégezni, akkor az adott statisztikai program által felkínált opciókból olyan modellt kell választanunk, mely nem tételez fel kölcsönhatást a folytonos és a kategorikus változó között, azaz az összes csoportra azonos meredekségű egyeneseket illeszt. Ezt a STATISTICA programban¹⁸ a *General Linear Models (GLM)* modul *Analysis of covariance model* opciójával tehetjük meg.

Ha pedig szimultán kívánjuk elvégezni a vizsgálatot, akkor az alkalmazott statisztikai programnak olyan modellt kell megadni, mely magában foglalja a folytonos és a

kategorikus változó közötti kölcsönhatás tagot is. Ez - ahogy már korábban láthattuk - a STATISTICA program¹⁸ *General Linear Models (GLM)* moduljának *Homogeneity-of-slopes* opciója.

II.3.2.4. A vizsgálatoknál alkalmazott szignifikancia-szint

Több sarzs egyesítésének feltételeit vizsgálva - mint minden más statisztikai próbánál - kétféle hibát követhetünk el:

- elsőfajú hiba: elvetjük a nullhipotézist (a sarzsok degradálódási egyeneseinek azonosságát), pedig igaz
- másodfajú hiba: elfogadjuk a nullhipotézist, holott nem igaz

Az első- ill. másodfajú hiba megengedhető valószínűségének meghatározásakor azt kell figyelembe venni, hogy az egyes hibák elkövetése mekkora kárt okozhat a felhasználónak (a vevőnek) ill. a gyártónak. Az egyesítés feltételeinek ellenőrzésénél nagyon lényeges, hogy a próba képes legyen kimutatni a sarzsok között a valóságban létező különbségeket, tehát a másodfajú hiba elkövetésének valószínűségét kell a lehető legnagyobb mértékben csökkenteni. Stabilitásvizsgálatoknál a másodfajú hiba elkövetésének valószínűsége a relatíve kis mintaelemszám miatt egyébként is igen nagy.

A másodfajú hiba valószínűségének csökkentésére az ICH Q1E Guideline⁴ (annak B.2.2.1. fejezete) által ajánlott, a gyakorlatban általánosan alkalmazott megoldás az, hogy az elsőfajú hiba megengedett valószínűségét a szokásos 0,05 helyett 0,25-ra választják, ugyanis az elsőfajú hiba valószínűségének növelésével a másodfajú hiba elkövetésének valószínűsége csökkenthető, azaz a próba ereje növelhető. Az ICH Guideline ezen ajánlása Bancroft¹⁹ cikkén alapul. A cikk a különböző statisztikai próbák alkalmazásának feltételeit tárgyalja, kitüntetett figyelemmel arra, hogy ezen előzetes vizsgálatokhoz milyen szignifikancia szintet kell választani.

II.3.3. Az ICH Guideline statisztikai modellje

A II.3.2.1. fejezetben a stabilitási adatok statisztikai értékelésére három lehetséges modellt mutattam be. Ezek közül most részletesen is ismertetem a második esetet, vagyis azt, amikor az előzetes, egyesíthetőségre vonatkozó vizsgálatok alapján a meredekségek azonosságát elfogadjuk, de a tengelymetszetek azonosságára vonatkozó nullhipotézist elutasítottuk. Választásom oka az, hogy a stabilitásvizsgálat statisztikai értékelésének

továbbfejlesztésére kidolgozott, a későbbiekben bemutatott új modellem ezen az alapmodellen alapul.

A stabilitásvizsgálat statisztikai modellje lényegében egy varianciaanalízissel (ANOVA) kombinált regresszióanalízis, mely a kategorikus (sarzs faktor) változó mellett folytonos változót (tárolási idő) is tartalmaz. A modell formailag a legegyszerűbb kovarianciaanalízis (ANCOVA) modell: egy faktorról és egy folytonos változóval.

Az ICH Guideline modelljét az alábbi egyenlettel írhatjuk fel:

$$y_{ijk} = \mu_i + \beta x_{ij} + \varepsilon_{ijk}, \quad (6)$$

ahol y_{ijk} az i -edik sarzs mért hatóanyag-tartalma a j -edik tárolási időpont k -edik mintájának mérésekor

μ_i a sarzs faktor i -edik szintje, $i=1, \dots, r$

x_{ij} az i -edik sarzs j -edik tárolási/vizsgálati időpontja, $j=1, \dots, n_i$

és $n = n_1 + \dots + n_r$

β a degradációs ráta

ε_{ijk} a hibatag, mely a véletlen ingadozásokat tartalmazza, $k=1, \dots, n_{ij}$

Ha az ICH Guideline modelljét mátrixformában írjuk fel, akkor az alábbi egyenletet kapjuk:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (7)$$

ahol $\mathbf{y}_{n \times 1}$ vektor tartalmazza a mért hatóanyag-tartalom értékeket,

$\mathbf{X}_{n \times (r+1)}$ mátrix a faktor (sarzs) és a folytonos változó (tárolási idő) tervmátrixa,

$\boldsymbol{\beta}_{(r+1) \times 1} = [\mu_1, \dots, \mu_r, \beta]^T$ a faktor és a folytonos változó paramétervektora,

$\boldsymbol{\varepsilon}_{n \times 1}$ a véletlen hibákat tartalmazó vektor.

A varianciaanalízissel (kovarianciaanalízissel) vizsgált faktorok rögzített (fix) ill. véletlen (random) jellegűek lehetnek. Az ICH Guideline modelljében a sarzs rögzített faktorként szerepel. Ezt az elképzelést számos irodalmi kritika érte, melyekre a II.4.3. fejezetben térek ki részletesebben.

A disszertáció további részében ezt a modellt ((6) egyenlet) tekintem kiindulási alapnak.

II.3.4. Többfaktoros kovarianciaanalízis modellek

Egy adott gyógyszerkészítmény több hatáserősségben ill. többfajta csomagolásban is készülhet. Ilyenkor a stabilitásvizsgálatot minden hatáserősség-csomagolás kombinációra külön-külön el kell végezni.

A kapott adatok statisztikai értékelése kétféleképpen történhet. Az első lehetőség, hogy minden egyes hatáserősség-csomagolás kombinációra külön-külön elvégezzük a lejáratidő becslését. Ez az előző fejezetben ismertetett egyfaktoros kovarianciaanalízis modell többszöri elvégzését jelenti. Ha végeredményként a gyógyszerkészítményre egyetlen lejáratidőt kell adnunk, akkor az egyes hatáserősség-csomagolás kombinációra kapott értékek közül a legrövidebbet választjuk.

A másik lehetőség, hogy a különböző hatáserősség-csomagolás kombinációkból származó adatokat egyesítve kezeljük, s a gyógyszerkészítményre egyetlen eltarthatósági időt adunk meg. Ekkor a statisztikai modellben a folytonos változó mellett már nem csak egyetlen faktor (a sarzs) szerepel, hanem a hatáserősség és a csomagolás hatása (ill. esetlegesen ezek kölcsönhatása) is faktorként jelentkezik. Így kapható ún. többfaktoros kovarianciaanalízis modelleket több cikk is elemzi.^{20,21,22} Természetesen az adatok együttes kezelése előtt mindenképpen meg kell győződni az egyesíthetőség feltételeinek teljesüléséről, vagyis hogy a különböző hatáserősség-csomagolás kombinációból származó adatok leírhatók-e egyetlen közös degradálódási egyenessel. Az egyesíthetőségi vizsgálatok elve megegyezik az egyfaktoros modellnél bemutatott sémával, azzal a különbséggel, hogy most nemcsak a különböző sarzsokra, hanem minden egyes faktorra ill. faktor-kombinációra tesztelni kell a meredekségek és a tengelymetszetek azonosságát. Érdemes kiemelni, hogy az ICH Q1E Guideline⁴ (annak B.3.2.2.1. fejezete) az egyesíthetőségi vizsgálatok szignifikancia-szintjére a faktor típusa szerint két különböző értéket ajánl: a sarzsok közötti különbség tesztelése $\alpha=0,25$ -t ír elő, míg az egyéb faktorok közötti különbség vizsgálatára a szokásos $\alpha=0,05$ szignifikancia-szintet ajánlja.

II.3.5. A matrixing és a bracketing módszer

A stabilitásvizsgálatok rendkívül időigényesek és drágák. A költségek csökkentésének lehetősége a vizsgálandó minták számának csökkentésével valósítható meg, melyre két eljárást, a matrixing és bracketing módszerét dolgozták ki.^{23,24}

A matrixing módszer tulajdonképpen egy rész-faktortervnek tekinthető, mely során egy adott vizsgálati időpontban csupán a vizsgálandó minták egy részét teszteljük. A módszer alkalmazásának fontos követelménye, hogy az egymást követő vizsgálati pontokban a mintáknak mindig különböző részleteit vizsgáljuk meg. A módszer azon a feltételezésen alapul, hogy az adott időpontban vizsgált minták stabilitási tulajdonságai az összes minta stabilitási tulajdonságait reprezentálják. A stabilitási adatoknak a matrixing módszerrel alapuló feldolgozása különösen is hasznos az előző fejezetben említett többfaktoros kovarianciaanalízis modellek esetén, hiszen ilyenkor a sok vizsgált faktor (sarzs, hatáserősség, csomagolás) rengeteg mintavételt jelent. Golden és társai²⁵ erre az esetre ajánlanak egy lehetséges mátrix tervet. Natarajan és társai²⁶ a vizsgálati időpontok megválasztásának, a degradációs rátának és az analízis bizonytalanságának hatását elemzik különböző mátrix tervek esetén. Olivia és társai²⁷ egy konkrét példán bemutatva hasonlítják össze az összes vizsgálati időpontot tartalmazó teljes tervet a különböző lehetséges mátrix tervekkel.

A bracketing módszer egy olyan stabilitásvizsgálati terv, mely során minden vizsgálati időpontban csupán a szélső értékű (pl. hatáserősség vagy csomagolási mód szerinti) mintákat teszteljük. A módszer azon a feltételezésen alapul, hogy a „közbenső” minták stabilitását jellemző tulajdonság-értékek a „szélső” minták stabilitását jellemző tulajdonság-értékek között helyezkednek el.

II.4. Az ICH Guideline kritikája

A stabilitásvizsgálattal foglalkozó ICH Guideline-ok^{3,4} előírásaival kapcsolatban számos kritika fogalmazódhat. Az alábbi alfejezetekben ezeket foglalom össze.

II.4.1. A sarzsok egyesítésének kérdése

A sarzsok - előzetes statisztikai vizsgálatokon alapuló - egyesítése azt az elképzelést tükrözi, hogy az összes sarzs ugyanolyan degradálódási jellegű sarzsok sokaságából származik. Ez a szemléletmód nyilvánul meg az ekkor alkalmazandó, közös meredekséget és tengelymetszetet feltételező statisztikai modellben.

Ha azonban a statisztikai próbák alapján a sarzsok egyesíthetőségére vonatkozó nullhipotézist elutasítjuk, akkor a sarzsok degradációs profilját eltérőnek feltételezve az

egyedülállóan különböző meredekségű és tengelymetszetű modelleket kell illeszteniük. Ebben az esetben az ICH Q1E Guideline⁴ előírása szerint a készítmény összes sarzsára vonatkozó eltarthatósági időt a vizsgálatba bevont leginkább bomló sarzs eltarthatósági ideje alapján kell megadni. Ha jól végiggondoljuk ezt a szakirodalomban a „minimum approach” néven ismeretes szemléletet, akkor ez az elv valójában azt jelenti, hogy a többi sarzs adatait figyelmen kívül hagyva csupán egyetlen egy sarzs (igaz ugyan, hogy a vizsgáltak közül a legrosszabb) alapján adunk becslést a készítmény eltarthatósági idejére. Ez az elképzelés hiányos statisztikai alapokon nyugszik, hiszen semmiféle biztosítékot nem ad arra, hogy az így kapott eltarthatósági idő tényleg kisebb, mint bármelyik, jövőbeni, hasonló technológiával gyártott sarzs valódi lejárat ideje.²⁸ A „minimum approach” elvvel kapcsolatban megfogalmazott kritikák másik fő iránya, hogy mivel csupán a leginkább bomló sarzs adatait veszi figyelembe, az így kapott lejárat idő a vizsgált sarzsok eltarthatósági ideje közül a legrövidebb, vagyis az összes többi vizsgált sarzsra az elvárt 95%-os szintnél nagyobb megbízhatóságot nyújt. Tehát ezen elv alkalmazása az eltarthatósági időre konzervatív becslést ad¹⁴ (p. 372).

A szakirodalomban nem egységes a tekintetben sem, hogy az egyesíthetőségi vizsgálat megfelelő statisztikai próbával történik-e. Ruberg and Hsu²⁹ szerint annak a nullhipotézisnek az elfogadása, hogy a sarzsokra illesztett degradálódási egyenesek meredeksége (ill. tengelymetszete) azonos, nem feltétlenül biztosítja a sarzsok hasonlóságát. Az általuk ajánlott egyesíthetőségi vizsgálat szimultán konfidencia-intervallumok számításán alapul. A módszer szerint a sarzsok akkor egyesíthetőek, ha a minden egyes sarzs és a leginkább degradálódó (azaz a legrövidebb eltarthatósági idővel rendelkező) sarzs meredekségeinek különbségére számolt konfidencia-intervallum felső határa kisebb, mint egy előre meghatározott érték. (Ez az eljárás a „multiple comparison with the worst” néven ismert a szakirodalomban.)

II.4.2. A szignifikancia-szint kérdése

Az ICH Guideline előírása szerint a sarzsok egyesíthetőségére végzett ellenőrző vizsgálatok szignifikancia-szintjét a szokásos 0,05 helyett 0,25-nak kell választani. Ennek háttérben az a megfontolás áll, hogy az elsőfajú hiba megengedett szintjének növelésével a próba ereje is növelhető. Ruberg és Stegeman³⁰ szerint azonban a 0,25-os határ kijelölése teljesen önkényes, sőt paradox helyzetet teremt. Egy igényesen megtervezett és kivitelezett

stabilitásvizsgálat esetén (sok sarzs vizsgálata, hosszú tárolási időtartam, több ismételt mérés, az analitikai mérés jól reprodukálható) ugyanis már a sarzsok közötti kis különbség is könnyen szignifikánsnak minősíthető 0,25-os szinten, hiszen ekkor a „precíz” analitikai mérés miatt az F -próba nevezőjében szereplő viszonyítási szórásnégyzet igen kicsi, és a több mérési pontnak köszönhetően nagyobb a szabadsági foka. Így könnyen előfordulhat, hogy egy szakmai szempontból elhanyagolhatónak tekinthető eltérést tévesen szignifikánsnak minősítünk. Ezzel szemben egy felületesen végzett stabilitásvizsgálat esetén (kevés sarzs vizsgálata, az ismételt mérések hiánya, pontatlan analitikai mérés) a sarzsok egyesítésére vonatkozó nullhipotézis 0,25-os szinten sokkal könnyebben elfogadható. E paradox helyzet következménye, hogy ha a gyógyszergyár a gyógyszer degradációjának minél pontosabb modellezése érdekében a stabilitásvizsgálat igényes elvégzésére törekszik, akkor könnyen kerülhet abba a helyzetbe, hogy a sarzsok egyesíthetőségére vonatkozó nullhipotézis csak ritkán fogadható el, ami számára nyilvánvalóan nem kedvező.

A helyes szignifikancia-szint megválasztására Ruberg és Stegeman³⁰ új megoldást ajánl. Mivel a sarzsok egyesíthetőségi vizsgálatánál elsődleges kérdés, hogy a sarzsok közti szakmai szempontból lényeges különbségeket ki tudjuk mutatni, a vizsgálat kritikus paramétere a másodfajú hiba valószínűsége. Ennek megfelelően az egyesíthetőségi vizsgálatoknál az elsőfajú hiba kockázatát jelentő α szignifikancia-szint helyett a statisztikai próba erejét, vagyis a másodfajú hiba el nem követésének valószínűségét kellene rögzíteni, majd ehhez a fix értékhez kellene kiszámítani a nullhipotézis elutasítására vonatkozó kritikus α szignifikancia-szintet. A módszer alkalmazásának legkényesebb kérdése a másodfajú hiba megadásához szükséges ellenhipotézis értelmes definiálása. A szerzők erre azt a megoldást alkalmazzák, hogy a sarzsok közti különbséget a sarzsokra illeszthető degradálódási egyenesek meredekségeinek különbségeként értelmezik, mégpedig úgy, hogy a meredekségek közti különbséget a sarzsok átlagos hatóanyag-tartalma között maximálisan megengedhető különbség és az elérni kívánt lejárati idő hányadosaként számítják. A konkrét számítás több lépésből áll, melynek ismertetésétől most eltekintek.

II.4.3. A sarzs: rögzített (fix) vagy véletlen (random) faktor?

A rögzített faktorok jellemzője, hogy szintjeiket a kísérletekhez megválaszthatjuk és beállíthatjuk. A véletlen faktorok szintjeit ezzel szemben nem lehet tetszőlegesen beállítani, értékeiket egy elképzelt sokaságból véletlenszerűen választhatjuk ki.³¹

A két faktortípus közti szemléletbeli különbség más-más statisztikai elemzést kíván. Rögzített faktor esetén a statisztikai elemzés célja annak megállapítása, hogy a várható értékben van-e különbség az adott faktor különböző szintjei között, s ha igen, akkor melyik közülük a legjobb, ill. jobb-e valamelyik szint a másiknál, majd becsüljük az egyes szintek hatását.

Véletlen faktor esetén ugyan szintén kíváncsiak vagyunk arra, hogy különböznek-e egymástól a faktor egyes szintjei, de azt nem kérdezzük, hogy valamelyik szintje jobb-e, és nem is becsüljük az egyes szintek hatását. Ehelyett inkább arra vagyunk kíváncsiak, hogy a faktornak a jövőben ugyancsak véletlenszerűen kiválasztandó (vagy előálló) más szintjén milyen mértékű eltérésre kell számítani, azaz mekkora a véletlen faktor szintjeinek ingadozása miatti ingadozás a függő változóban.

II.4.3.1. A rögzített és a véletlen faktoros megközelítés szemléletmódja a stabilitásvizsgálatnál

Egy adott gyógyszerkészítmény sarzsai a természetszerű gyártási ingadozások miatt kisebb-nagyobb mértékben különbözhetnek egymástól.

Ha az ICH Guideline előírását követve először a stabilitásvizsgálatba bevont sarzsok egyesíthetőségét vizsgáljuk, a sarzs lehet rögzített faktor, hiszen előre meghatározott sarzsokat kívánunk összehasonlítani. Kérdésünk ekkor az, hogy a vizsgált sarzsokra illesztett degradációs egyenesek - az ingadozás ellenére - tekinthetők-e azonosnak. Ezután az eltarthatósági idő becsüléséhez az egyesíthetőségi vizsgálat eredményétől függően vagy egyetlen (a legrosszabb) sarzs adatait használjuk fel, vagy tovább már nem foglalkozva az adatok „eredetével” az összes sarzsot együttesen kezeljük. Fontos azonban megjegyezni, hogy az így kapott következtetések valójában csak a vizsgálatba bevont sarzsokra érvényesek.

Ha azonban úgy gondolkozunk, hogy a stabilitásvizsgálatra eltett sarzsok a hasonló technológiával gyártott sarzsok sokaságából véletlenszerűen kiválasztott, azokat

reprezentáló minták, akkor a statisztikai elemzéshez inkább a véletlen faktoros megközelítést kellene alkalmazni. Ennél a szemléletmódnál a sarzsok egyesíthetősége nem merül fel kérdésként, hiszen célunk éppen a sarzsok közti különbség (szórás) számszerűsítése, amihez a sarzsokat - a köztük lévő különbség ellenére - mindenképpen együtt kell kezelni. A másik lényeges, szemléletmódból következő különbség, hogy a véletlen faktoros modellel kapott következtetések már az összes, azonos technológiával gyártott ill. majd a jövőben gyártandó sarzsra érvényesek lesznek.

A véletlen faktoros szemléletmód esetén az eltarthatósági idő meghatározása a sarzsok közti szórás felhasználásával történhet, mégpedig úgy, hogy az ICH Guideline által előírt konfidencia-sávot egy további varianciakomponenssel, a sarzsok közti szórással bővítjük. Az így kapott bizonytalansági sáv a több tagból álló szórás miatt természetesen szélesebb lesz, mint ha azt csak az ICH Guideline előírása szerint számítanánk. Ebből pedig az következik, hogy a véletlen faktoros modell ugyanarra az adatsorra rövidebb (konzervatív) becsült eltarthatósági időt eredményez¹⁶ (p. 437), ami a gyógyszergyárak számára nyilvánvalóan nem kedvező. Ez az ellentmondás csak az eltarthatósági idő újradefiniálásával lenne feloldható.²⁹ A véletlen faktoros megközelítés hátrányaként említik továbbá az irodalomban azt is, hogy a kis sarzsszám miatt a sarzsok közti ingadozás varianciájának becslése igen bizonytalan.²⁹

A sarzsok különbözőségét az ICH Q1E Guideline⁴ sajnálatos módon ellentmondásosan kezeli. A statisztikai modellekben a sarzs rögzített faktorként szerepel, ugyanakkor a Guideline szöveges előírása szerint a becsült eltarthatósági időnek az összes, jövőben bármikor előforduló sarzsra alkalmazhatónak kell lennie. Ez utóbbi (teljesen jogos) követelménnyel viszont a véletlen faktoros megközelítés van összhangban. E kettősség jegyében a szakirodalomban széleskörű vita bontakozott ki a sarzs faktor jellegét illetően. A stabilitásvizsgálat rögzített és a véletlen faktoros szemléletmódjáról részletesen olvashatunk Chow és Liu¹⁴ gyógyszeripari statisztikai alkalmazásokról szóló könyvének 11. ill. 12. fejezetében.

II.4.3.2. A véletlen jellegű sarzs faktor statisztikai kezelése

A disszertáció eddigi fejezeteiben ismertetett statisztikai módszerek, megfontolások a rögzített faktoros szemléleten alapultak. E fejezet további témája a véletlen faktoros megközelítés részletesebb bemutatása.

Ha a sarzs faktort véletlen jellegűnek tekintjük, akkor a stabilitásvizsgálat statisztikai modellje az alábbiak szerint írható fel¹⁴ (p. 395):

$$y_{ijk} = \alpha_i + \beta_i x_{ij} + \varepsilon_{ijk}, \quad (8)$$

ahol a paraméterek értelmezése megegyezik a (2) egyenletnél ismertekkel, azzal a kivétellel, hogy a véletlen faktoros megközelítés miatt α_i és β_i most normális eloszlású paraméterek ($N(\alpha, \sigma_\alpha^2)$ ill. $N(\beta, \sigma_\beta^2)$), továbbá α_i , β_i és ε_{ij} kölcsönösen függetlenek egymástól. A modell az irodalomban „random coefficients model” néven ismeretes.

Ha a sarzsok között nincs különbség, akkor mind σ_α^2 , mind σ_β^2 nullának vehető, azaz modellünk a $y_{ijk} = \alpha + \beta x_{ij} + \varepsilon_{ijk}$ egyenletté egyszerűsödik. Ekkor a degradációs egyenes illesztése a már ismert módon, a legkisebb négyzetek módszerével történik.

Számunkra azonban most fontosabb az az eset, amikor a sarzsok között különbség van, tehát valamelyik (vagy éppen mindkettő) varianciakomponens szignifikánsan eltér nullától. Ekkor az alábbi két aleset állhat fenn:

Ha $\sigma_\alpha^2 > 0$ és $\sigma_\beta^2 = 0$ (azaz ha a sarzsokra illesztett egyenesek tengelymetszete különbözik ugyan egymástól, de a meredekségük nem), akkor a (8) egyenlet az alábbiak szerint alakul:

$$y_{ijk} = \alpha + \beta x_{ij} + u_{ijk},$$

ahol $u_{ijk} = \alpha_i - \alpha + \varepsilon_{ijk}$ alakban felírható normális eloszlású paraméter $N(0, \sigma_\alpha^2 + \sigma_\varepsilon^2)$.

Ebben az esetben a lejáratási idő becslésére még mindig alkalmazható a legkisebb négyzetek módszere.

Ha azonban $\sigma_\beta^2 > 0$ (azaz ha a sarzsokra illesztett egyenesek meredeksége is különbözik egymástól), akkor a degradációs egyenes illesztéséhez a legkisebb négyzetek módszere már nem használható³². Ez a megállapítás rávilágít a $\sigma_\beta^2 = 0$ nullhipotézis tesztelésének különös fontosságára.

A sarzs faktor véletlen jellege tehát jelentheti a tengelymetszet, a meredekség vagy éppen mindkettő paraméter véletlen voltát. A meredekség véletlenszerűsége azon múlik, hogy a bomlás sebessége az egyes sarzsokra véletlenszerűnek tekinthető-e. Ha a bomlás sebessége fizikai-kémiai törvényszerűség, akkor nem. Ha a bomlás sebessége bármilyen értelemben függ attól, hogy az adott sarzs hogyan sikerül, akkor igen. Az utóbbira lehet példa más komponens (pl. víz) koncentrációja, a csomagolás jósága (zár-e) stb.

II.4.3.2.1. Statisztikai módszerek a sarzsok közti variancia vizsgálatára

Chow és Shao³² cikkében három, egyszerűbben számítható, így a gyógyszeripari gyakorlatban könnyebben használható statisztikai módszert ajánl a sarzsok közti variancia vizsgálatára szintén a $\sigma_\beta^2 = 0$ nullhipotézis tesztelésén keresztül. Chow-ék tehát ebben a cikkükben már foglalkoznak a véletlen faktoros megközelítéssel, de ezen belül éppen a sarzsok lehetséges egyesítésére tesznek kísérletet, ami ellenkezik a véletlen faktoros szemlélettel.

A sarzsok közti variancia (σ_α^2 és σ_β^2 varianciakomponensek) konkrét értékének becslésével több cikk^{33,34} is foglalkozik.

Chow és Shao³³ a (8) egyenletben ismertetett modell esetén ad becslést a σ_α^2 és a σ_β^2 varianciakomponensekre. Ehhez a \mathbf{D} kovarianciamátrixot használja fel:

$$\hat{\mathbf{D}}_{n \times n} = \frac{1}{r-1} \sum_{i=1}^r (\mathbf{Y}_i - \bar{\mathbf{Y}})(\mathbf{Y}_i - \bar{\mathbf{Y}})^T, \quad (9)$$

ahol $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{in_i})^T$ vektor az i -edik sarzs mért hatóanyag-tartalom értékeit tartalmazza és

$$\bar{\mathbf{Y}} = \frac{1}{r} \sum_{i=1}^r \mathbf{Y}_i.$$

A \mathbf{D} mátrix elemei (d_{jl}) konzisztens becslései lesznek a σ_{jl} varianciáknak, ahol:

$$\sigma_{jl} = \begin{cases} \sigma_\alpha^2 + \sigma_\beta^2 x_j^2 + \sigma_\epsilon^2 & \text{ha } j = l \\ \sigma_\alpha^2 + \sigma_\beta^2 x_j x_l & \text{ha } j \neq l \end{cases}, \quad (10)$$

ahol x_j ill. x_l a j -edik ill. az l -edik mérési időpontot jelöli és $j = 0, 1, \dots, n-1$.

Látható, hogy a σ_{jl} varianciák σ_α^2 , σ_β^2 és σ_ϵ^2 varianciakomponensek lineáris kombinációi, így a keresett varianciakomponensek a fenti képletekből a legkisebb négyzetek módszerével megkaphatók:

$$\sigma_\alpha^2 = \frac{1}{n-1} \sum_{l=2}^n d_{1l}, \quad (11)$$

$$\sigma_\beta^2 = \frac{\sum_{j=2}^n \sum_{l=j+1}^n d_{jl} - (n-2)/2 \sum_{l=2}^n d_{jl}}{\sum_{j=2}^n \sum_{l=j+1}^n x_j x_l} \quad (12)$$

II.4.3.2.2. Az eltarthatósági idő becslése véletlen faktoros megközelítés esetén

A sarzs faktor véletlen jellegű kezelése esetén az eltarthatósági idő becslésére több módszert is találunk az irodalomban. Az egyes módszerek alkalmazhatósága a stabilitásvizsgálatba bevont sarzsok számától függ.

Chow és Shao³³ cikkében két statisztikai módszert ajánl a nagy sarzsszámú - azaz a gyógyszerkészítmény törzskönyvezése után indított - stabilitásvizsgálatok értékelésére. A stabilitási adatok értékelése mindkét módszernél olyan súlyozott lineáris regresszióval történik, ahol a paraméterek valószínűségi változók. Ez a statisztikai módszer az irodalomban „weighted linear regression model with random coefficients” néven ismert, részletes matematikai leírása Hildreth és Houck³⁵ cikkében található.

Az első módszer az átlagos eltarthatósági idő becslését adja. A degradációs egyenes paramétereinek kiszámításához súlyként az egyes mérési időpontokban a sarzsok közti variancia becslését használja fel, majd az így kapott egyeneshez konfidencia-sávot számít.

A második módszer szintén a súlyozott lineáris regresszió alapul, ám kérdésfeltevése messzebbre vezet. Célja a konfidencia-sáv helyett egy olyan intervallum megadása, mely adott megbízhatósági szinten biztosítja, hogy az egyes sarzsok hatóanyag-tartalma a megadott lejáratú időpontban a specifikációs kritérium felett legyen. Ezt egy ún. tolerancia korrekciós faktor bevezetésével teszi meg, melynek értéke a sarzsok közti szórástól és a megbízhatósági szinttől függ. Kiemelném, hogy az irodalomban először ebben a cikkben találunk konkrét számítást a tolerancia-intervallum stabilitásvizsgálatokra történő alkalmazására.

A fentebb ismertetett súlyozott regresszió csak nagyszámú sarzs esetén alkalmazható¹⁴ (p. 393). Mivel a gyógyszerkészítmény törzskönyvezését megelőző stabilitásvizsgálatoknál leggyakrabban csak három sarzsot vizsgálnak, új statisztikai módszer kidolgozására volt szükség. Ez Shao és Chow³⁶ nevéhez köthető. A módszer a (súlyozatlan) legkisebb négyzetek elvén és ismét a tolerancia-intervallum filozófiáján alapul. Mivel saját munkám egyik része e cikk kritikájára és a módszer helyes alkalmazására épül, a módszert részletesen annak tárgyalásakor, az V.2.1. fejezetben mutatom be. Itt említem még meg Sun és társai cikkét³⁷, melyben a Shao és Chow módszerével becsült eltarthatósági idő eloszlását vizsgálják.

Az ebben a fejezetben említett véletlen sarzs faktort alkalmazó cikkek a sarzsok közti ingadozást véletlen jellegű varianciakomponensként kezelték. Ezekhez képest valamelyest elkülönül Murphy és társai cikke³⁸, mely a sarzsokra illesztett egyenesek meredekségét tekinti véletlen változónak („random slopes model”).

II.4.4. A bizonytalanság számszerűsítése: konfidencia-sáv, jóslási sáv vagy tolerancia-sáv?

Gyógyszerkészítmények eltarthatósági idejének becslése - definíciójának megfelelően (II.3.1.2 fejezet) - az adatokat megfelelően leíró degradálódási függvény megtalálásán, valamint az ismereteink pontatlanságát kifejező bizonytalansági tartomány számításán alapul. Az ICH Q1E Guideline⁴ előírása szerint a stabilitási adatok bizonytalanságának számszerűsítése konfidencia-sávval történhet.

A konfidencia-sáv alkalmazása azon a feltételezésen alapul, hogy létezik egy igazi koncentráció és degradálódási függvény, melytől azonban a mért stabilitási adatok az analízis bizonytalansága miatt eltérnek. Ennek köszönhetően a valódi bomlási függvény paramétereit csak pontatlanul tudjuk becsülni. Az eltarthatósági idő a konfidencia-tartományos koncepció szerint az az időpont, amelynél az igazi hatóanyag-tartalom (az igazi bomlási egyenes) eléri a specifikációs határt. Az ismereteink bizonytalansága miatt azonban ezt az időpontot csak bizonytalanul tudjuk megadni, ezért írjuk elő, hogy a valódi hatóanyag-tartalom adott időpontbeli konfidencia-tartományának kell elérnie a specifikációs határt.

Fontos megjegyezni, hogy a konfidencia-tartomány annál inkább szűkül, minél több mérési adat (több sarzs, több mérési időpont, több minta, több ismételt mérés) áll rendelkezésünkre (ld. (1) egyenlet). Mindez a gyakorlatban azt jelenti, hogy - adott degradációs ráta mellett - a konfidencia-sáv szűkítése, s ezzel a becsült eltarthatósági idő növelése pusztán a mérések számának növelésével megvalósítható, ahelyett, hogy a bizonytalanság tényleges forrásait, vagyis a gyártási folyamat ill. az analitikai mérés ingadozásának okait küszöbölnénk ki. A konfidencia-sáv megközelítés helytelensége végtelenszámú mintánál a legszembetűnőbb, mivel ekkor a konfidencia-sáv zérus szélességig való szűkülése révén az eltarthatósági idő becslése már annyira pontos lesz, hogy a becsült eltarthatósági idő megegyezik a valódi eltarthatósági idővel. Ez viszont azt is jelenti, hogy a valóságban mindig jelenlévő tabletták közti különbözőség miatt ebben a „szélsőséges helyzetben” a tabletták 50%-ának hatóanyag-tartalma a specifikációs határ alatt van.

II.4.4.1. Jóslási (predikciós) sáv

A szakirodalomban széleskörű vita bontakozott ki a konfidencia-sáv alkalmazásának jogosságáról. Egyes szerzők^{28,39,40} a konfidencia-sáv helyett jóslási (predikciós) sávot alkalmaznának.

A jóslási sáv az a tartomány, amely adott időpontra a bomlási egyenes illesztésekor fel nem használt vagy egy jövőbeni mérési adatot adott (pl. 95%-os) valószínűséggel tartalmaz. A fő különbség a két bizonytalansági tartomány szemléletmódjában az, hogy a konfidencia-intervallum az elméleti degradációs egyenes egy pontjának, míg a jóslási tartomány az egyedi mérési adatok elhelyezkedését mutatja.

A két tartomány definíciójából következik, hogy a jóslási sáv mindig szélesebb, mint a konfidencia-sáv, így ugyanarra az adatsorra rövidebb becsült eltarthatósági időt eredményez, azaz konzervatívabb becslést ad.

A másik fontos különbség a két bizonytalansági tartomány között, hogy ugyan mindkettő feltételez egy igazi/valódi bomlási egyenest, de a konfidencia-intervallum a *várható* hatóanyag-tartalom, míg a jóslási intervallum az (egyedi) *mért* hatóanyag-tartalom elhelyezkedésére ad választ.

II.4.4.2. Tolerancia-sáv

Gondolkodjunk el másképpen is a bomlási függvényről és annak bizonytalansági sávjáról. Ha a sarzsok és a tabletták között természetesen meglévő gyártási különbségeket már a bomlási egyenes illesztésekor figyelembe vesszük, akkor az eltarthatósági időt az alapján kellene keresnünk, hogy egy jövőbeli sarzs valamely tablettájának valódi hatóanyag-tartalma mikor éri el a specifikációs határt, s ez az időpont, amit az analitikai (mérési) bizonytalanság miatt csak valamilyen biztonsággal tudunk becsülni. Erre a kérdésre a tolerancia-intervallum adhat választ.

A tolerancia-intervallum definíció szerint az a tartomány, melyben a sokaság legalább p hányada adott/előírt valószínűséggel megtalálható⁴¹. Ezt a definíciót a stabilitásvizsgálatra alkalmazva arra a konkrét kérdésre keressük a választ, hogy a jövőbeli tabletták zömének (pl. legalább 99%-ának) bizonyos időpontbeli hatóanyag-tartalma adott (pl. 95%) valószínűséggel milyen tartományban található. Másként megfogalmazva ez azt jelenti, hogy az az elvárásunk, hogy a lejáratú időpontban a jövőbeli tabletták előírt hányadának (pl. 99%-ának) hatóanyag-tartalma adott (pl. 95%) valószínűséggel a tűréshatár felett legyen.

Az eltarthatósági idő tolerancia-intervallum alapján történő becslésére több szerző^{30,33,36} is utal az irodalomban. Az ezzel a témával foglalkozó legrészletesebb elemzés Kiermeier és társai⁴² cikkében található. Mivel ez a cikk a saját munkám egyik részének kiindulópontja, részletes bemutatását az V.2.2. fejezet tartalmazza.

A konfidencia- és a tolerancia-intervallumos koncepció között lényeges eltérés van a mérési adatok ingadozásának figyelembevételénél. A konfidencia-intervallum egyetlen össz-ingadozással számol, ami ugyan tartalmazza az összes ingadozási forrást (így az inhomogenitást is), ám ezek hatását nem megfelelően összegzi. Ezért számítása csak akkor lesz pontos, ha egyetlen ingadozás-forrást (az analízis bizonytalanságát) feltételezünk, a többit pedig elhanyagolható nagyságúnak tekintjük. A konfidencia-sáv tehát explicite nem veszi figyelembe az inhomogenitást. Ezzel szemben a tolerancia-intervallum számításánál akár több, megkülönböztethető ingadozási forrás (az analitikai hiba mellett pl. a tabletták vagy a sarzsok különbözősége) is figyelembe vehető.

Ez a koncepcióbeli különbség az intervallumok szélességében is megmutatkozik: ugyanis míg a konfidencia-intervallum végtelen számú sarzs ill. tablettá figyelembevétele esetén a végletekig szűkíthető, addig a tolerancia-tartomány szélessége sohasem lehet zérus. A tolerancia sáv minimális szélessége az analízis bizonytalansága mellett figyelembe vett ingadozási források nagyságától (pl. σ_{sarzs}^2 és a $\sigma_{\text{tablettá}}^2$ varianciakomponensek) függ, ehhez járul hozzá a végesszámú analízis okozta bizonytalanság, mely a mérések számától függően itt is tetszőlegesen csökkenthető.

A gyógyszergyárak alapvető (piaci) érdeke, hogy a gyógyszerkészítményeik becsült eltarthatósági ideje a lehető leghosszabb legyen. A tolerancia-intervallum típusú megközelítés - más szemlélete miatt - ugyanarra az adatsorra rövidebb becsült eltarthatósági időt ad, mint ha az ICH Guideline által előírt konfidencia-intervallummal számolnánk. Ez azonban nem jelenti azt, hogy az adott gyógyszer rosszabb minőségű, csak más statisztikai kérdésre kerestük a választ: a konfidencia-intervallumnál a tabletták átlagának elhelyezkedését vizsgáltuk, a tolerancia-intervallum pedig a tabletták nagy hányadának (pl. 99%-ának) elhelyezkedését. Természetes, hogy ugyanannál a stabilitási adatsornál a tabletták nagy hányada szélesebb intervallumban helyezkedik el, mint az átlagértékük, s ezáltal a tolerancia-sáv hamarabb metszi a specifikációs határt rövidebb becsült eltarthatósági időt eredményezve. Az viszont egyáltalán nem lenne jogos, ha a

helyes statisztikai kérdést feltevő tolerancia-sávós szemléletet alkalmazó gyógyszergyárakat rövidebb becsült eltarthatósági idővel büntetnénk, amikor az általuk gyártott gyógyszer minősége ugyanolyan jó. Ezért a tolerancia-intervallumos szemléletre való átállás esetén a specifikációs határok egyidejű megváltoztatására (csökkentésére) lenne szükség, mégpedig úgy, hogy a tolerancia-intervallumos szemlélet szerint megadott lejárati időpontban ugyanakkora valószínűséggel legyünk biztosak abban, hogy az adott gyógyszer tartalmazza az előírt hatóanyag-tartalmat, mint a konfidencia-intervallumos szemléletnél.

II.5. A stabilitási adatok ingadozási forrásai

Gyógyszerkészítmények stabilitási adatainak ingadozását vizsgálva két fő forrást különíthetünk el: az analitikai mérés bizonytalanságát, valamint a gyártási folyamat ingadozását. Mindkét csoport további két-két ingadozási forrásra bontható, így a stabilitási adatok ingadozása összesen négy varianciakomponens eredőjeként adódik. Az egyes ingadozási forrásokat valamint a hozzájuk tartozó elemi varianciakomponenseket az alábbi felsorolás tartalmazza:

- Az analitikai mérés bizonytalansága
 - az analízis hibája ($\sigma_{\text{ismétlés}}^2$)
 - az analízis körülményeinek különbözősége ($\sigma_{\text{körülmények}}^2$: pl. σ_{nap}^2 , σ_{labor}^2)
- A gyártási folyamat ingadozása
 - a sarzsok közötti különbség (σ_{sarzs}^2)
 - a sarzson belüli inhomogenitás ($\sigma_{\text{inhomogenitás}}^2$)

Vizsgáljuk meg részletesebben is az egyes ingadozási forrásokat!

Az analitikai méréseknél mindig jelenlevő és számszerűsíthető ingadozás az analízis hibája. Ez alatt azt a véletlen ingadozást értjük, amely az egyazon mintából azonos analitikai körülmények között (azaz adott napon, ugyanazon laboratóriumban, egy személy által, ugyanazon a készüléken) végzett ismételt analíziseknél tapasztalható. Ezt más néven ismételhetségi (repeatability) hibakomponensnek is nevezik.

Emellett az alap-ingadozás mellett legtöbbször felfedezhető és statisztikailag kimutatható az analízis körülményeinek (melyik napon, mely személy által, mely készüléken, mely

vegyszerek felhasználásával ill. melyik laboratóriumban történt a mérés) ingadozást növelő hatása. Ezt az ingadozási formát összefoglaló néven a mérési módszer reprodukálhatóságának nevezik és két lépcsőben szokás vizsgálni: laboratóriumon belüli ismételhetőség (intermediate precision) ill. laboratóriumok közötti ismételhetőség (reproducibility).

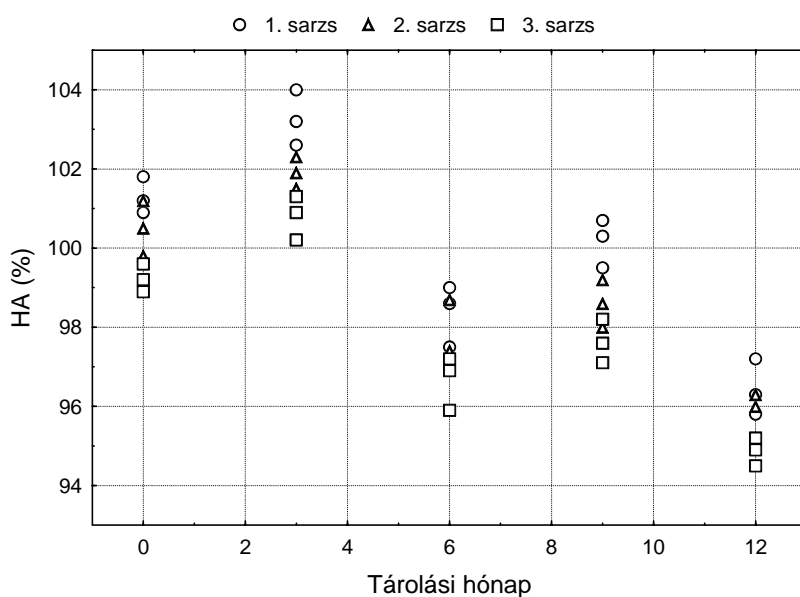
Egy analitikai módszer reprodukálhatósági hibája rendszerint jóval nagyobb az ismételhetőségi hibánál, hiszen az egymástól igen eltérő analitikai körülmények jelentősen növelhetik az adatok szóródását. Renger⁴³ említi, hogy a kromatográfiás elemzéseknél az ismételhetőségi szórásnál általában 50%-kal nagyobb az intermediate precision szórás értéke, és még ennél is 50%-kal nagyobb a reprodukálhatósági szórás.

Az analitikai módszer reprodukálhatóságát befolyásoló tényezők közül a stabilitásvizsgálat szempontjából külön kiemelném az analízis-időpontok különbözőségének hatását, hisz a stabilitásvizsgálatnál az egyes mérések között nemcsak napok, hanem hónapok, akár évek telnek el.⁴⁴ A több hónap különbséggel végzett analíziseknél az analitikai körülmények nagy valószínűséggel sokkal jobban eltérnek egymástól, mint a mérési módszer validálásánál végzett reprodukálhatósági vizsgálatoknál, ahol az ismételt méréseket tipikusan egymást követő napokon végzik.⁴⁵ A stabilitásvizsgálat nagyobb időbeni kihívása miatt érdemes lenne megfontolni (természetesen figyelembevéve ennek költségvonzatát is) az analitikai módszer validálásakor végzett reprodukálhatósági vizsgálat egy új ponttal történő bővítését, miszerint az analízist nemcsak egymást követő napokon, hanem több hét vagy hónap elteltével is meg kellene ismételni.

A gyógyszerkészítmény gyártási folyamatának természetszerű ingadozása miatt az egyes sarzsok, valamint a sarzsokon belül az egyes tabletták hatóanyag-tartalma adott tűréshatáron belül más és más lehet. Mivel a stabilitásvizsgálat során több (általában három) sarzsot vizsgálnak, valamint a különböző időpontokban, mintáknál ill. ismétléseknél mindig más-más tablettákat analizálnak, a stabilitási adatok ingadozásának elemzésekor a sarzsok közti különbség, valamint a sarzson belüli inhomogenitás figyelembevétele feltétlenül indokolt.

A 4. ábra a stabilitásvizsgálat során mért hatóanyag-tartalom értékek tipikus képét mutatja, melyről egyrészt jól látható az adatok szóródásának hármas/négyes szerkezete (ld. következő bekezdés), valamint a különböző analízis-időpontoknak tulajdonítható jelentős ingadozás.

Fontos megemlíteni, hogy a stabilitásvizsgálati mérések során az analitikai mérés ismételhetőségi hibakomponense és az inhomogenitás hatása együttesen jelentkeznek. Ha a minták egy-egy tablettát tartalmaznak, vagy ha ugyan az egyes mintákat több tablettával egyesítéssel készítik, de azokat csak egyszer analizálják, akkor a két ingadozási forrás hatása számszerűleg elválaszthatatlan egymástól. Csak és kizárólag akkor van lehetőség a sarszon belüli inhomogenitás és az analitikai hiba numerikus szétválasztására, ha a mintákat több tablettával elporításával, majd homogenizálásával készítik, és a mintákat többször is analizálják.



4. ábra: Stabilitásvizsgálati adatok tipikus képe

Habár a stabilitási adatok ingadozása a gyakorlatban jelenlevő valóság, hatásával csak igen kevesen foglalkoztak az irodalomban. Bar⁴⁶ a stabilitási adatok időbeni profilját vizsgálva elkülöníti az analitikai módszer bizonytalansága miatti ingadozást a gyógyszerkészítmény degradációja miatt bekövetkező változástól. Cikkében az analitikai módszer ismételhetőségi és reprodukálási varianciájának becslését egy stabil gyógyszerkészítmény stabilitásvizsgálati adatsorán mutatja be. Az így kapott eredmények alapján kiemeli, hogy a stabilitási adatoknál tapasztalható 5%-os hatóanyagtartalom-változást - mely az ICH Q1A(R2) Guideline által szakmai értelemben már jelentősnek minősül (ld. II.2.2 fejezet vége) - akár teljes mértékben az analitikai módszer pontatlansága is okozhatja.

Bar egy későbbi cikkében⁴⁷ egyszerű lehetőséget kínál annak eldöntésére, hogy a stabilitási adatok változását vajon a gyógyszerkészítmény degradációja vagy a stabilitási adatok ingadozása okozza. A gyógyszerkészítményben bekövetkező változás jelzőszámaként a

stabilitási adatokra illesztett egyenes zérus meredekségére vonatkozó nullhipotézist vizsgálja. A stabilitási adatok szóródásának mértékét pedig a folyamatképesség vizsgálatánál használt minőség-képességi indexszel jellemzi.

Van der Veen és társai⁵ részletesen elemzik a standard anyagok stabilitásvizsgálatakor tapasztalható ingadozási forrásokat ill. azok becslési lehetőségét a stabilitási adatokból. Pauwels és társai⁶ szimulációs példákat mutatnak be arra, hogy a degradációs ráta, a stabilitási adatok variációs koefficiense és a konfidencia-sáv megbízhatósági szintje hogyan befolyásolja a referencia anyagok becsült eltarthatósági idejét.

III. KUTATÁSI MUNKÁM TÉMATERÜLETEI

Doktori munkám során a gyógyszerkészítmények stabilitásvizsgálatának statisztikai értékelésének tárgyában részletesen három témakörrel foglalkoztam.

1. Két témakörben a stabilitásvizsgálat jelenlegi statisztikai modelljét kritizálva a modell továbbfejlesztésére, pontosítására törekedtem az eltarthatósági idő valóságnak megfelelőbb becslése érdekében. Az alábbiakban vázaltszerűen bemutatom ezt a két témakört röviden utalva arra, hogy hol támadják a jelenlegi statisztikai elemzési módszert, ill. javaslataim milyen lényegesebb újdonságot tartalmaznak.

1.1. Az ICH Q1E Guideline⁴ által ajánlott statisztikai módszer figyelem kívül hagyja a stabilitási adatok ingadozásának egyik fő (legtöbbször a legjelentősebbnek bizonyuló) forrását: az időben egymástól igen távoli analízis-időpontok okozta különböző analitikai körülmények hatását. Ezt a hiányt pótolta Norwood⁴⁸ cikke, aki a stabilitásvizsgálat statisztikai modelljét egy új faktorról, az analízis időpontjával bővítette. A Norwood-modellben alkalmazott faktor-struktúra azonban nem megfelelő, mivel az egy adott időpontban érvényes reprodukálhatósági hibát sarzsonként különbözőnek tekinti. Az általam ajánlott új modell az analízis-időpont faktor újradefiniálásával a reprodukálhatósági hibát már a valóságnak megfelelően, a sarzstól független faktorként veszi figyelembe. Az itt elért eredményeket a disszertáció IV. fejezete tartalmazza.

1.2. Az ICH Q1E Guideline⁴ előírása szerint az eltarthatósági idő becslése a konfidencia-sáv segítségével történik. Véleményem szerint a stabilitási adatok ingadozását kifejező bizonytalansági tartományra a homogénnek feltételezett tabletta sokaság várható értékének elhelyezkedését mutató konfidencia-sáv helyett megfelelőbb lenne a tabletta sokaság elemeinek adott hányadának hatóanyag-tartalmát adott valószínűséggel tartalmazó tolerancia-sáv. Ezt a felvetést kutatásommal párhuzamosan Kiermeier⁴² is szorgalmazta.

Kiermeier tolerancia-intervallumon alapuló modellje azonban a sarzson belüli ingadozási források közül csak az inhomogenitás, azaz a tabletták különbözőségének hatását veszi figyelembe, az analitikai bizonytalanságot elhanyagolja.

A másik új, eddig csak általam vizsgált kérdés a témakörben, hogy a stabilitásvizsgálat értékelésénél a mért vagy a valódi (tényleges) hatóanyag-tartalomra adjunk-e ingadozási tartományt. Véleményem szerint jogosabb lenne a valódi hatóanyag-tartalom elhelyezkedését vizsgálni, hisz a beteget a tabletta tényleges hatóanyag-tartalma érdekli, nem pedig az analitikai bizonytalansággal terhelt mért hatóanyag-tartalom.

A fenti két megállapítás fényében kutatásom célja egy több ingadozás-forrást figyelembevevő, a valódi hatóanyag-tartalom elhelyezkedését középpontba helyező, tolerancia-sávon alapuló stabilitásvizsgálati módszer kidolgozása és statisztikai tulajdonságainak vizsgálata. Ennek részleteit a disszertáció V. fejezete tartalmazza.

2. A harmadik témakörben a stabilitási adatok ingadozásának előzetes becslésére tettem kísérletet. Céлом ezzel az volt, hogy már a stabilitásvizsgálat indítása előtt látható legyen, hogy az adott gyártási és analitikai módszer bizonytalansága megfelelő-e (azaz nem túl nagy) a kívánt lejárati idő eléréséhez. A becsült eltarthatósági idő ugyanis a jelenlegi ICH Guideline szabályozása szerint a konfidencia-sáv és a tűréshatár metszéspontja. A konfidencia-sáv szélessége, s így a becsült eltarthatósági idő pedig nagymértékben függ a stabilitási adatok ingadozásától. A stabilitási adatok egyes varianciakomponenseinek előzetes becsléséhez a gyógyszerkészítmény fejlesztése során egyébként is elvégzett analitikai mérések, mint például a validálás és a Content Uniformity vizsgálat eredményeit használtam fel. A számítás részleteit a disszertáció VI. fejezetében mutatom be.

IV. A STABILITÁSVIZSGÁLAT STATISZTIKAI MODELLJÉBEN SZEREPLŐ ANALÍZIS-IDŐPONT FAKTOR ÚJRAÉRTELMEZÉSE

IV.1. A probléma felvetése

A stabilitásvizsgálat statisztikai értékelésére az ICH Q1E Guideline⁴ által ajánlott modell (melyet a II.3.3. fejezetben részletesen ismerttettem) két tényezőt vesz figyelembe: a vizsgált sarzsok különbözőségét, valamint a tárolás során jelentkező hatóanyagtartalom-változást. A sarzsok közti ingadozást rögzített jellegű kategorikus változóként, a hatóanyag bomlását a tárolási idő (mint folytonos változó) és a degradációs ráta szorzataként kezeli. Az így kapott kovarianciaanalízis-modell a sarzs faktor és a degradációs ráta mellett még egy tagot tartalmaz, mely a stabilitásvizsgálat során jelentkező véletlen hibákat foglalja magában.

Mint ahogy a II.5. fejezetben említettem, a stabilitási adatok ingadozása négy forrásból ered. Az ICH Q1E Guideline e négy tényezőből viszont csak kettőt/hármat vesz figyelembe: a sarzsok különbözőségét és az analitikai mérés ismétlésekor tapasztalható véletlen hibát, mely együttesen tartalmazza az analitikai mérés ismételhetségi hibakomponensét és a sarzson belüli inhomogenitást. A stabilitásvizsgálat során tapasztalható egymástól igen eltérő analitikai körülmények hatását (ld. 4. ábra) figyelmen kívül hagyja.

Az analízis-időpontok különbözősége okozta jelentős többletingadozás más bizonytalansági forrásoktól való elkülönítésével és a modellben történő figyelembe vételével az adott gyógyszerkészítmény degradációja a valóságnak hűebben írható le, így az eltarthatósági időre pontosabb becslést kaphatunk.

A stabilitási adatok analízis-időpontok szerinti csoportba rendeződésére több szerző^{17,40,48,49} is utal az irodalomban. Többségük^{17,40} a probléma megoldására mélyhűtött minták párhuzamos tárolását ajánlja. A módszer lényege, hogy minden egyes analízis-időpontban az előírt tárolási körülmények között tartott mintákkal párhuzamosan mélyhűtőben tartott mintákat is analizálnak. A módszer azon a feltételezésen alapul, hogy mélyhűtés körülményei között a gyógyszerkészítmény nem degradálódik, így a mélyhűtött

minták adatai csak az analitikai mérés adott időpontbeli rendszeres hibáját tartalmazzák, melyet a „valódi” stabilitási adatok értékelésénél korrekcióba vehetünk.

Az analízis-időpontok különbözősége okozta ingadozás figyelembevételére vonatkozó irodalmi javaslatok közül külön ki kell emelni Norwood⁴⁸ cikkét, mely a stabilitásvizsgálat statisztikai modelljének az analízis időpontjával mint új faktorról való bővítését ajánlja. A módszer részletes elemzését, kritikáját és továbbfejlesztését a következő fejezetekben ismertetem.

IV.2. A Norwood-modell

Norwood⁴⁸ modelljének kiindulási alapja az ICH Q1E Guideline⁴ kovarianciaanalízis-modellje ((6) egyenlet). Norwood ezt a modellt a hibatag két részre bontásával építi tovább, mégpedig úgy, hogy a véletlen hibából kiemeli az analízis-időpontok különbözőségéhez köthető ingadozást (vagyis a reprodukálhatósági hibát), s azt egy új, véletlen jellegű faktorként kezeli. Norwood statisztikai modellje az alábbi egyenlettel írható fel:

$$y_{ijk} = \mu_i + \beta x_{ij} + \delta_{ij} + \varepsilon_{ijk}, \quad (13)$$

ahol y_{ijk} az i -edik sarzs mért hatóanyag-tartalma a j -edik tárolási időpont k -adik ismételt mérésekor

μ_i a sarzs faktor i -edik szintje, $i=1, \dots, r$

β a degradációs ráta

x_{ij} az i -edik sarzs j -edik tárolási/vizsgálati időpontja, $j=1, \dots, n_i$

és $n = n_1 + \dots + n_r$

δ_{ij} a sarzs-idő kombinációkhoz kapcsolódó hibakomponens (azaz a reprodukálhatósági hiba)

ε_{ijk} a sarzs-idő kombinációkon belüli hibakomponens (azaz az ismételhetségi hiba és az inhomogenitás)

A modell feltételezése, hogy a két hibatag normális eloszlású, várható értékük 0, varianciájuk pedig konstans (σ^2 ill. σ_s^2), továbbá hogy egymástól függetlenek.

IV.3. A Norwood-modell kritikája

IV.3.1. A sarzs faktor kérdése

A Norwood-modell az ICH Guideline kontextusában maradván a konfidencia-sávok szemléletét tükrözi. Ebben a megközelítésben a sarzs csakis rögzített faktor lehet. Ennek megfelelően a Norwood-modell a sarzs faktort rögzítettként kezeli.

Az irodalmi részben ismertett okok miatt a sarzs faktor rögzítettként való kezelése statisztikai szempontból ugyan nem legmegfelelőbb (ld. II.4.3. fejezet), de az (időben) első kutatási témában célt csupán a másik faktor, az analízis-időpont faktor vizsgálata volt.

A sarzs faktor véletlenként való kezelése már távolabb, a tolerancia-intervallumos szemlélethez vezet, melyre való áttérés kutatási munkám kezdetén még nem volt célt.

A véletlen sarzs faktor bevezetése továbbá felveti azt a problémát is, hogy az ipari gyakorlatban jelenleg alkalmazott kevés sarzs szám mellett a sarzs hatását (variánciakomponensét) csak igen pontatlanul tudjuk becsülni. Ez egyben azt is jelenti, hogy a véletlen sarzs faktoros modellek alkalmazása a jelenleginél lényegesen nagyobb sarzs-számot tenné szükségessé.

Végül a rögzített sarzs faktor megtartása mellett szólt az is, hogy a három modell (ICH Guideline modellje, Norwood-modell és annak továbbfejlesztése) összehasonlítása is így volt a legegyszerűbb.

A fentieket összegezve elmondható, hogy a Norwood modell továbbfejlesztésekor célt a modell egy ponton (az analízis-időpont faktor értelmezésén) történő javítása volt, nem pedig az ICH Guideline által ajánlott és a szakmai körökben általánosan elfogadott konfidencia-sávok szemlélet teljes megváltoztatása.

IV.3.2. Az analízis-időpont faktor értelmezése

A stabilitásvizsgálat statisztikai értékelésére szolgáló modell Norwood általi módosításának lényege az analízis-időpontok különbözősége okozta rendszeres analitikai hiba elkülönítése az egy adott időpontban végzett ismételt mérésekből adódó hibától. Elemezzük a továbbiakban részletesen az így kapott, reprodukálhatósági hibát jellemző új faktort (δ_{ij})!

Norwood elképzelése szerint a modellbe bevont új faktor a sarzs-idő kombinációkhoz kapcsolódik. Erre utal az új faktor indexe (ij) is, miszerint minden egyes sarzs-idő ($i-j$)

kombináció a faktor újabb és újabb szintjét jelöli. Mindez azt jelenti, hogy az új faktor (δ_{ij}) az egyes sarzs-idő kombinációk (pontosabban a hozzájuk tartozó analitikai körülmények) közti különbség, míg a fennmaradó hibatag (ε_{ijk}) az egyes sarzs-idő kombinációkon belüli ismételt mérések eltéréseinek leírására szolgál.

Norwood modelljében tehát minden egyes sarzs-idő kombináció egy újabb esetet (más analitikai körülményt) jelöl, s így az új, reprodukálhatósági hibát jellemző faktor más és más, azaz egymással össze nem vethető szintjéhez tartozik. Mindezt szakmai nyelven megfogalmazva azt mondhatjuk, hogy a két faktor egymással hierarchikus viszonyban van, mégpedig úgy, hogy a tárolási/mintavételi idő szintjei a sarzs faktor szintjeibe vannak ágyazva. Az angol nyelvű szakirodalom a véletlen hibakomponensek ilyen felépítésű szerkezetét „nested-error structure”-nek nevezi⁵⁰.

A Norwood-modell szemlélete hétköznapi nyelven pedig azt jelenti, mintha az egyes időpontokban aktuálisan érvényes analitikai körülmények a sarzs faktor különböző szintjein mások és mások lennének. Tehát hiába indulna minden sarzs stabilitásvizsgálata ugyanabban az időpontban és zajlanék azonos tárolási terv szerint, azaz hiába történnének sarzsonként azonos időpontban a mintavételek és azt követő analízisek, a modell az egy adott időpontban érvényes reprodukálhatósági hibát sarzsonként különbözőnek tekinti, ami egyáltalán nem indokolható feltételezés.

Az analitikai mérések során tapasztalható reprodukálhatósági hiba jellemzője éppen az, mint azt a 4. ábra is sugallja, hogy az egy adott időpontban végzett méréseknél - az analízis körülményeinek azonossága miatt - a valódi értéktől azonos mértékű eltérést eredményez. Tehát, ha egy adott időpontban több, akár különböző sarzsokból származó mintát elemzünk, akkor ezeket a méréseket ugyanaz a reprodukálhatósági hibakomponens terheli. Ezt figyelembe véve a reprodukálhatósági hibának a modellben sarzstól független (nem pedig abba ágyazott) faktorként kell szerepelnie. Ez pedig reményt ad arra, hogy a reprodukálhatósági hibát jellemző faktor (δ) értéke kisebb lesz, szűkítve ezzel a konfidencia-sávot.

Összefoglalva megállapítható, hogy Norwood ugyan helyesen látta meg a reprodukálhatósági hibát jellemző új faktor bevezetésének szükségességét, de a faktor szintjeit nem megfelelően határozta meg.

IV.4. A Norwood-modell továbbfejlesztése

IV.4.1. Az új modell szemléletmódja

A Norwood-modell módosításának célja a reprodukálhatósági hiba (vagyis az analízis-időpont hatásának) valóságnak megfelelő figyelembevételére, mely a reprodukálhatósági hibát jellemző analízis-időpont faktor újradefiniálását jelenti. A modellt úgy célszerű módosítani, hogy az egy adott analízis-időponthoz tartozó, s ezért ugyanazzal a reprodukálhatósági hibával terhelt mérési eredmények az analízis-időpont faktor (δ) azonos szintjéhez tartozzanak. Természetesen ennek feltétele, hogy az adott (tárolási) időponthoz tartozó mérések ugyanabban a laboratóriumban, ugyanazon a készüléken, ugyanazon a napon, ugyanazon személy által történjenek.

Az alábbi példa (3. táblázat) szemléletesen mutatja a két koncepció közti különbséget. Norwood modelljében (bal oldali ábra) az analízis-időpont faktor egy szintjének egy sarzs- δ idő (ij) kombináció, azaz egy cella felel meg. Ennek megfelelően Norwood elgondolása szerint ebben a példában a faktor szintjeinek száma $3 \times 3 = 9$ lesz. Az új modell (jobb oldali ábra) szemléletmódja szerint az analízis-időpont faktor szintjeit csak a mintavételi időpontok száma határozza meg, hiszen az azonos időponthoz tartozó analíziseket - sarzstól függetlenül - ugyanaz a reprodukálhatósági hiba terheli, így ugyanahhoz az analízis-időpont faktor szinthez tartoznak. Az új modellben a faktor-szintek száma tehát megegyezik a tárolási/mintavételi időpontok számával, ami ebben a példában 3. Látható, hogy a faktor-szintek száma a Norwood-modellben a cellák számával, míg az új modellben a sorok számával egyenlő.

Analízis időpontja (hónap) (j)	Sarzs (i)		
	A	B	C
0			
3			
6			

A Norwood-modell koncepciója

Analízis időpontja (hónap) (j)	Sarzs (i)		
	A	B	C
0			
3			
6			

Az új modell szemléletmódja

3. táblázat: Az analízis-időpont faktor egy-egy szintje a két modellben (szürkével kiemelve)

IV.4.2. Az új modell matematikai alakja

Az előző fejezetek gondolatmenete alapján látható, hogy a Norwood-modell hibája a helytelen faktor-struktúra (hierarchikus terv keresztosztályozás helyett). A stabilitásvizsgálat leírására szolgáló új modell tehát az analízis-időpont faktor leírásában tér el Norwood modelljétől, mégpedig úgy, hogy az analízis-időpont faktor szintjeire az eddigi kettő (ij) helyett egy, eddig nem használt indexet (t) vezet be. Így az alábbi modellt kapjuk:

$$y_{ijk} = \mu_i + \beta x_{ij} + \delta_t + \varepsilon_{ijk}, \quad (14)$$

ahol a változók jelentése - az analízis-időpont hatását kifejező δ_t faktor kivételével - megegyezik a Norwood-modellnél ((13) egyenlet) felsoroltakkal. Az analízis-időpont faktor új indexe (t) az analitikai mérés konkrét időpontjára (dátumára) utal, így a δ_t faktor az adott időpontban tapasztalható - sarzstól független - reprodukálhatósági hibát jelenti.

IV.4.3. A két időskála

A stabilitási adatok leírására szolgáló módosított modell újszerű szemlélete abban nyilvánul meg, hogy - az irodalomban található eddigi modellekkel ellentétben - két különböző időskálát alkalmaz: az abszolút időskálát és a tárolási időskálát. Az abszolút időskála (t) a naptári időnek (dátumnak) felel meg, és a modellben az analízis-időpontok különbözőségét (a reprodukálhatósági hibát) jellemző véletlen faktorhoz kapcsolódik. Megjegyzem, hogy az abszolút időskála egy adott időpontja az akkor érvényes analitikai körülmények együttesét jelenti. A tárolási időskála (j) a gyógyszer gyártása óta eltelt időt, vagyis a sarzs korát mutatja, és a modellben a degradációs rátához kötődik. A két időskála csak akkor lesz azonos egymással, ha az összes vizsgált sarzs tárolása ugyanabban az időpontban indul. Ha viszont valamelyik sarzs stabilitásvizsgálata a többi sarzsétól eltérő időpontban indul, akkor a később indult sarzsnál a két időskála már eltér egymástól. A stabilitásvizsgálatra eltett sarzsok különböző időpontokban történő indítása ma még ugyan nem általános gyakorlat a gyógyszeriparban, de nem is kivételes, egyedi eset. Alkalmazását a későbbiekben vázolt előnyök miatt érdemes megfontolni.

Lássunk egy példát a különböző időpontokban indított sarzsok statisztikai kezelésére!

Abszolút időskála (t)	Tárolási időskála (j)		
	„A” sarzs	„B” sarzs	„C” sarzs
0	0		
3	3	0	
6	6	3	0
9	9	6	3
12	12	9	6
18	18	15	12
24	24	21	18

4. táblázat: A két időskála (hónap)

A 4. táblázatban bemutatott példánál a „B” sarzs stabilitásvizsgálata 3, míg a „C” sarzs stabilitásvizsgálata 6 hónappal az „A” sarzs stabilitásvizsgálatának kezdete után indult. A két időskála eltérése nyilvánvaló: az abszolút időskála egy adott pontjához (ahogy a szürke színnel jelölt sor is mutatja) a tárolási időskála három különböző pontja tartozik a három különböző időpontban indított sarzs esetén.

Az abszolút időskála kezdőpontját mindig ahhoz az időponthoz rögzítjük, amikor a stabilitásvizsgálatra legelőször eltett sarzs (itt az „A” sarzs) tárolását indítottuk. Ennek következtében az elsőnek indított sarzsnál a két időskála megegyezik egymással. Ezzel szemben a később induló sarzsok esetén a két időskála már eltér egymástól, mégpedig úgy, hogy az abszolút időskála egy adott dátumhoz (vagyis az analízis-időponthoz) képest a tárolási idő(tartam) épp annyival marad el, mint amennyivel az adott sarzs stabilitásvizsgálata a legelsőként indított sarzs stabilitásvizsgálata után indult.

IV.4.4. A Norwood- és az új modell mátrix-egyenlete

A Norwood-modell és a módosításaként kapott új modell mátrixjelölésekkel az alábbi egyenlet formájában írható fel:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon}, \quad (15)$$

ahol $\mathbf{y}_{n \times 1}$ vektor tartalmazza a mért hatóanyag-tartalom értékeket

(ahol n a mérési adatok száma),

$\mathbf{X}_{n \times (r+1)}$ mátrix a rögzített faktorok (sarzs, tárolási idő) tervmátrixa

(ahol r a sarzsok száma),

$\boldsymbol{\beta}_{(r+1) \times 1} = [\mu_1, \dots, \mu_r, \beta]^T$ a rögzített faktorok paramétervektora,

$\mathbf{Z}_{n \times p}$ a véletlen faktor (analízis-időpont) tervmátrixa

(ahol p a véletlen faktor szintjeinek száma)

$\boldsymbol{\delta}_{p \times 1}$ a véletlen faktor paramétervektora,

$\boldsymbol{\varepsilon}_{n \times 1}$ a véletlen hibákat tartalmazó vektor.

A két modell közti különbség a véletlen jellegű analízis-időpont faktor szintjeinek értelmezésében és számában van, mely mátrix-felírásban a p változó értékében, s így a $\mathbf{Z}_{n \times p}$ tervmátrix és a $\boldsymbol{\delta}_{p \times 1}$ paramétervektor méretében jelentkezik.

Megjegyzem, hogy az ICH Guideline modelljének mátrix-egyenlete ((7) egyenlet) a fenti egyenlet egyszerűsítéseként adódik, mivel az az analízis-időpont hatásának figyelmen kívül hagyása miatt csak rögzített faktorokat, s az ennek megfelelő $\mathbf{X}_{n \times (r+1)}$ tervmátrixot és $\boldsymbol{\beta}_{(r+1) \times 1}$ paramétervektort tartalmazza. Az $\mathbf{X}_{n \times (r+1)}$ tervmátrix tehát mindhárom modellben azonos.

IV.5. A Norwood- és az új modellel történő számítás menete

Az ICH Guideline rögzített faktorokat tartalmazó modelljét a véletlen jellegű analízis-időpont faktorról bővítve, ún. kevert (mixed) modellt kapunk. A rögzített és véletlen faktorokat egyaránt tartalmazó kevert modellek statisztikai kezelése lényegesen nehezebb feladat a csupán azonos jellegű faktorokat tartalmazó modellek elemzésénél. Különösen igaz ez a megállapítás a nem kiegyensúlyozott ill. a hiányzó cellákat tartalmazó kísérleti tervek esetén. A stabilitásvizsgálatra eltett sarzsok különböző időpontokban történő indítása - melynek statisztikai kezelésére az új modell megoldást jelent - viszont mindenképpen hiányzó cellákat tartalmazó tervet eredményez.

A kevert modellek statisztikai elemzését - az irodalmi ajánlás⁵¹ (p. 276) szerint - a modell véletlen részével kell kezdeni: az első lépésünk tehát a véletlen jellegű analízis-időpont faktor varianciakomponensének becslése lesz.

IV.5.1. A véletlen faktorok varianciakomponenseinek becslése

A különböző modellek varianciakomponenseinek becslésére számos módszer ismert a szakirodalomban. Bonyolultabb modellek esetére Henderson⁵² három módszert is javasol, melyek közül a kevert modellek varianciakomponenseinek becslésére leginkább Henderson

III. módszerét (az ún. „fitting constants method-ot”) ajánlja a szakirodalom^{53*}. A módszer alapelve, hogy a teljes modellt részmodellekre bontva, a különböző modellek illesztésével kapott maradék négyzetösszegek különbségeit egyenlővé teszi azok várható értékével, s az egyenletből az ismeretlen varianciakomponenst számítja. A számítás menetét Searle⁵³ cikke alapján az alábbiakban röviden ismertetem.

Általános esetben a teljes modell a következőképpen írható fel: $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, ahol az \mathbf{y} vektor tartalmazza a függő változó értékeit, \mathbf{X} a tervmátrix, $\boldsymbol{\beta}$ a paramétervektor, $\boldsymbol{\varepsilon}$ pedig a hibamátrix. Bontsuk a faktorokat két csoportra ($\mathbf{X}_1, \mathbf{X}_2$), így a következő modellt kapjuk: $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$. A faktorok kettéosztása teljesen önkényes lehet, független attól, hogy mely hatások rögzítettek és melyek véletlenek. Majd számítsuk ki azt a négyzetösszeget $[R(\boldsymbol{\beta}_2|\boldsymbol{\beta}_1)]$, melyet úgy kapunk, hogy kivonjuk egymásból a teljes modellel ($R(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$), valamint a $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}$ részmodellel kapott maradék-négyzetösszegeket ($R(\boldsymbol{\beta}_1)$). Az így számított négyzetösszeg várható értéke a következő mátrixalgebrai kifejezéssel egyenlő:

$$E[R(\boldsymbol{\beta}_2|\boldsymbol{\beta}_1)] = E[R(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) - R(\boldsymbol{\beta}_1)] = \text{tr} \left\{ \mathbf{X}_2' [\mathbf{I} - \mathbf{X}_1 (\mathbf{X}_1' \mathbf{X}_1)^{-1} \mathbf{X}_1'] \mathbf{X}_2 E(\boldsymbol{\beta}_2 \boldsymbol{\beta}_2') \right\} + \sigma_e^2 [r(\mathbf{X}) - r(\mathbf{X}_1)], \quad (16)$$

ahol \mathbf{I} az egységmátrix, r pedig az adott mátrix rangját jelöli.

A fenti egyenletből látható, hogy a $R(\boldsymbol{\beta}_2|\boldsymbol{\beta}_1)$ négyzetösszeg várható értéke csak két paraméter, $E(\boldsymbol{\beta}_2 \boldsymbol{\beta}_2')$ és σ_e^2 függvénye, $\boldsymbol{\beta}_1$ pedig nem szerepel benne. Ennek a megállapításnak akkor van kiemelkedő jelentősége, ha a kiindulási modellben a faktorokat úgy csoportosítottuk, hogy a $\boldsymbol{\beta}_1$ paramétervektor csak a rögzített faktorokat, a $\boldsymbol{\beta}_2$ pedig csak a véletlen faktorokat tartalmazza. Ekkor az $R(\boldsymbol{\beta}_2|\boldsymbol{\beta}_1)$ négyzetösszeg várható értéke csak a σ_e^2 hibavarianciát és a véletlen faktorokhoz tartozó varianciakomponenseket tartalmazza, a rögzített faktorokat pedig nem. E kedvező tulajdonsága teszi a módszert közkedvelté a kevert modellek használatakor. Az általános eset eredményét felhasználva a 15. egyenletben ismertetett új modellnél az $\mathbf{X}_1 = \mathbf{X}$, az $\mathbf{X}_2 = \mathbf{Z}$, a $\boldsymbol{\beta}_1 = \boldsymbol{\beta}$ és a $\boldsymbol{\beta}_2 = \boldsymbol{\delta}$ megfeleltetéseket kell alkalmazni.

* Tisztában vagyok vele, hogy a fentebb említett hivatkozások régiek. A kevert modellek statisztikai elemzésére ma már több statisztikai szoftver (pl. az ingyenes és mára igen elterjedt R-program) képes. Doktori munkámat azonban 10 évvel ezelőtt kezdtem, s a három fő kutatási témakör közül ezzel foglalkoztam a legkorábban (2003-ig). Akkoriban még csak a dolgozatban említett irodalmi hivatkozások és statisztikai szoftverek álltak rendelkezésemre, az R-program például épp akkor jelent meg.

IV.5.2. A rögzített faktor hatásának becslése

A számítás a modell rögzített részének elemzésével folytatódik, melyhez már felhasználjuk a véletlen hatások varianciakomponenseire kapott becsült értékeket is. A rögzített hatásokat tartalmazó β paramétervektort az általánosított legkisebb négyzetek módszerével kaphatjuk meg:

$$\hat{\beta} = (\mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{y}, \quad (17)$$

$$\text{ahol } \mathbf{V}_{n \times n} = \text{Var}(\mathbf{Y}) / \sigma^2 = \mathbf{Z}\mathbf{Z}^T \sigma_{\delta}^2 / \sigma^2 + \mathbf{I}_n. \quad (18)$$

IV.6. Példa

Abszolút időskála (az analízis időpontja)	Tárolási idő (kor)		Hatóanyag-tartalom (%)			
	„A” és „C” sarzs	„B” sarzs	„A” sarzs	„B” sarzs	„C” sarzs	
0	0		100,4 100,0 100,4 100,1		105,6 101,6 103,8 104,8 103,0 102,6 103,2 103,8 98,8 98,0 103,4 104,0 100,4 104,2 101,2 103,8	
0,5	0,5	0	97,1 96,4 96,8 99,0	98,1 97,4 98,1 97,4	96,0 101,8	
2,0	2,0	1,5	96,5 94,4 95,1 91,7 94,5 93,2	96,4 96,8 96,6 96,8	98,6 98,0 99,8 98,0	
2,5	2,5	2,0	100,0 98,1 98,0 95,9 98,3 101,7	93,2 95,2 93,3 94,9		
3,0	3,0	2,5	97,2 97,5 99,9 101,1	100,0 100,8 100,1 100,0		
3,5	3,5	3,0		96,4 95,8 95,8 94,9		

5. táblázat: Hatóanyag-tartalom adatok (Norwood⁴⁸ cikkéből) valamint a két időskála értékei

Az analízis-időpont faktor nélkül ill. annak kétféleképpen való figyelembevételével kapott statisztikai modellek (az ICH Guideline modellje, a Norwood-modell valamint az új modell) közti különbséget a Norwood-cikkben⁴⁸ szereplő példán szemléltetem. Az adatok - melyeket a

5. táblázat mutat - egy gyógyszerkészítmény stabilitásvizsgálatának hatóanyag-tartalmi méréséből származnak.

Az 5. táblázat a mért hatóanyag-tartalom értékeken kívül a két időskála önkényes egységekben kifejezett értékeit is tartalmazza. A két időskála összevetéséből látható, hogy a „B” sarzs stabilitásvizsgálata a másik két sarzséhoz képest 0,5 egységgel később indult, tehát ennél a sarzsnál a tárolási időskála egy adott értékéhez az abszolút időskálán mindig 0,5 egységgel nagyobb érték tartozik. A másik két sarzsnál a két időskála megegyezik egymással. Az 5. táblázatból jól látható a kísérleti terv hiányzó cellákat tartalmazó szerkezete. Az adatok kiértékelését a terv nem kiegyensúlyozott volta (azaz a cellánkénti nem azonos ismétlésszám) is nehezíti.

Az ICH Guideline egyszerű ANCOVA modelljének megoldása a legkisebb négyzetek módszerével, míg a Norwood- és az új modell megoldása az előző fejezetben ismertetett módon történt. Norwood modelljében az analízis-időpont faktor szintjeinek száma 13 (egyenlő a cellák számával), míg az új modellben csak 6 (megegyezik a táblázat sorainak - azaz az analízis-időpontoknak a - számával).

	ICH Guideline modellje	Norwood- modell	Új modell
μ_A	98,8	98,7	98,7
μ_B	98,1	98,1	98,2
μ_C	101,8	100,6	101,2
β	-0,66	-0,51	-0,57
σ_δ^2	-	5,27	4,03
σ^2	6,01	2,63	3,63
Konf. sáv szélessége 3 időegységnél	0,89	1,06	1,06

6. táblázat: A három modellel kapott eredmények (STATISTICA Visual Basic programmal)

A konkrét számításokat a rendelkezésemre álló STATISTICA szoftver¹⁸ 6.1. verziójával végeztem. Mivel kevert modellek esetén a szoftver e verziója csak a varianciakomponensek (σ_δ^2 és σ^2) becslésére alkalmas, így a meredekség (β paraméter) számítására a szoftverbe épített Visual Basic programnyelven programokat kellett írnom. A programozáshoz a véletlen

hatások becslésénél a momentumok módszerét, a rögzített faktorok becslésénél pedig az általánosított legkisebb négyzetek módszerét használtam.

Az egyes modellek paramétereinek becsült értékét (a sarzs faktor egyes szintjeinél a kiindulási koncentráció várható értékét, a degradációs egyenes meredekségét, az analízis-időpont hatás varianciakomponensét, valamint a hibavarianciát) a 6. táblázat mutatja.

A számítási eredményekből látható, hogy a modell-választás ténylegesen befolyásolja a paraméterek becsült értékeit.

A degradációs egyenes becsült meredeksége modellenként különböző: a legnagyobb értéket (abszolút értékben) az ICH Guideline modelljével kaptam, azaz a hatóanyag-tartalom becsült bomlása itt a leggyorsabb. Mindez azt is jelenti, hogy ha a jelenlegi szabályozás szerint az eltérő analízis-időpontok okozta reprodukálhatósági hibát figyelmen kívül hagyjuk, a helytelen modellválasztás ennél a példánál a legrövidebb becsült eltarthatósági időt eredményezi, ha a három modellre azonos szórást feltételezünk.

Példaként kiszámítottam a konfidencia-sáv szélességét 3 időegységénél, ennek eredményét a 6. táblázat utolsó sora mutatja.

Az analízis időpont hatását kifejező faktor varianciakomponensére (σ_{δ}^2) kapott érték az új modellben kisebb, mint a Norwood-modellben, ami nem véletlen, hiszen az új modellben az analízis időpont faktornak csak annyi szintje van, ahány analízis időpont (és ennek megfelelő különböző analitikai körülmény) ténylegesen is volt.

Saját számításaim ellenőrzéséhez segítséget kértem két biostatistikus kollégától*, akik a számomra nem hozzáférhető SAS⁵⁴ és SPSS szoftverben is elvégezték a szükséges számításokat. A különböző statisztikai szoftverekkel kapott eredmények összefoglalását a 7. táblázat mutatja. (A SAS és az SPSS szoftverrel kapott eredmények megegyeztek egymással, így a táblázatban csak az előbbieket közlöm.)

A 7. táblázatból látható, hogy a különböző statisztikai szoftverekkel kapott eredmények némileg eltérnek egymástól. Az ICH Guideline kovarianciaanalízis modelljére még mindhárom szoftverrel egyező eredményt kapunk, ám a másik két modell szoftveres megoldásánál már számos különbséget láthatunk. Egyrészt - mint már említettem - a STATISTICA szoftver a bonyolultabb számítást igénylő Norwood- és új modellnél

* Köszönöm Dr. Boda Krisztina és Dr. Singer Júlia segítségét.

önmagában nem alkalmas minden paraméter becslésére. (Erre utalnak a kérdőjelek a táblázat adott celláiban.)

Másrészt a SAS szoftver kevert modellekre alkalmazható két modulja (*GLM* és *Mixed*) egymástól elérő eredményt ad mind a varianciakomponensekre, mind pedig az illesztett egyenes meredekségére. Ennek oka a két modul eltérő becslési mechanizmusában keresendő. A *GLM* (*General Linear Model*) modul a véletlen hatások becslését a momentumok módszerével végzi, míg a *Mixed* modul a maximum likelihood módszer alapján. A rögzített faktorok becslése esetén a *GLM* modul a legkisebb négyzetek módszerével, a *Mixed* modul pedig az általánosított legkisebb négyzetek módszerével számol. Ha mindezt összevetjük a kevert modellek számítására ajánlott irodalmi módszerekkel (ld. IV.5. fejezet), akkor már érthető, hogy a helyes eredményt miért a SAS szoftver említett két moduljának kombinációjával kaphatjuk meg.

	ICH Guideline modellje				Norwood-modell				Új modell			
	SVB	STAT	SAS		SVB	STAT	SAS		SVB	STAT	SAS	
			GLM	Mixed			GLM	Mixed			GLM	Mixed
β	-0,66	-0,66	-0,66	-0,66	-0,51	?	-0,20	-0,51	-0,57	?	-1,03	-0,57
σ_{δ}^2	6,01	6,01	6,01	6,01	5,27	5,27	5,27	4,86	4,03	4,03	4,03	3,52
σ^2					2,63	2,63	2,63	2,63	3,63	3,63	3,63	3,62

7. táblázat: A különböző statisztikai szoftverekkel kapott eredmények összefoglalása

(SVB: STATISTICA Visual Basic makró,

STAT: STATISTICA szoftver, SAS: SAS szoftver)

Összességében tehát megállapítható, hogy „egyszerűbb” modellek esetén (mint például az itt szereplő egy faktort és egy folytonos változót tartalmazó kovarianciaanalízis modell) a statisztikai szoftverek egymással összhangban levő, az átlagos felhasználó számára is egyértelmű eredményt adnak. Összetettebb, több különböző típusú faktort tartalmazó modellek esetén a programcsomagok használata már mélyebb statisztikai ismereteket igényel.

V. AZ ELTARTHATÓSÁGI IDŐ BECSLÉSE TOLERANCIA-INTERVALLUMON ALAPULÓ, TÖBBFAKTOROS MODELLEL

V.1. A tolerancia-intervallum számításának alapjai

A magyar szakirodalomban tudomásom szerint a tolerancia-intervallum nem kellően ismert és használatos, ezért ebben a fejezetekben röviden áttekintem a nemzetközi szakirodalomban található, a témához kapcsolódó tolerancia-intervallum számításokat.

V.1.1. Egyoldali tolerancia-intervallum a normális eloszlású sokaság elemeinek nagy részének elhelyezkedésére Owen klasszikus cikke alapján⁵⁵ (p. 446-447)

Ha egy normális eloszlású sokaság μ várható értéke és σ^2 varianciája ismert, a standard normális eloszlás (jelölése ezentúl Z -eloszlás) eloszlásfüggvényének (és $Z_{1-\gamma}$ kvantilisének) ismeretében pontosan megadhatjuk azt az $x_U = \mu + Z_{1-\gamma}\sigma$ felső határt, amely alatt az X valószínűségi változó értékei $1-\gamma$ valószínűséggel megtalálhatók (azaz az X valószínűségi változó $X_{1-\gamma}$ kvantilisét).

Ha azonban a sokaság μ várható értéke és a σ^2 varianciája előzetesen nem ismertek, így a mintából becsüljük őket, akkor a fenti összefüggés már nem érvényes. Ilyenkor az egyoldali felső tolerancia-határt $x_U = \bar{x} + ks$ alakban keressük, ahol a \bar{x} minta átlaga, s pedig a szórása. Mivel \bar{x} és s valószínűségi változók, a határ is valószínűségi változó lesz, így állításunkat csak úgy fogalmazhatjuk meg, hogy a kérdéses intervallumban a sokaság adott hányada $(1-\gamma)$ előírt valószínűséggel $(1-\alpha)$ megtalálható (azaz az x valószínűségi változó $x_{1-\gamma}$ kvantilisének konfidencia-intervallumát adjuk meg).

A feladat matematikailag úgy fogalmazható meg, hogy keressük azt az $x_U = \bar{x} + ks$ felső határértéket, amelynél a sokaság elemeinek legalább $1-\gamma$ hányada kisebb $1-\alpha$ valószínűséggel, ill. keressük az e valószínűségekhez tartozó k értéket:

$$P[P(X \leq x_U) \geq 1 - \gamma] = 1 - \alpha .^* \quad (19)$$

* Megjegyzem, hogy Owen cikkében a sokaság adott hányadát P -vel, a megbízhatósági szintet pedig γ -val jelöli.

Áttérve a Z standard normális változóra, s behelyettesítve az $z_U = \frac{x_U - \mu}{\sigma} = \frac{\bar{x} + ks - \mu}{\sigma}$

kifejezéseket:

$$P[z_U \geq Z_{1-\gamma}] = P\left[\frac{x_U - \mu}{\sigma} \geq Z_{1-\gamma}\right] = P\left[\frac{\bar{x} + ks - \mu}{\sigma} \geq Z_{1-\gamma}\right] = 1 - \alpha.$$

Továbbalakítva a fenti kifejezést:

$$P\left[\frac{\bar{x} - \mu}{\sigma} - Z_{1-\gamma} \geq -\frac{ks}{\sigma}\right] = 1 - \alpha,$$

$$P\left[\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} - Z_{1-\gamma}\sqrt{n} \geq -\frac{ks}{\sigma/\sqrt{n}}\right] = P\left[\frac{\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} - Z_{1-\gamma}\sqrt{n}}{s/\sigma} \geq -k\sqrt{n}\right] = 1 - \alpha, \quad (20)$$

ahol n a minta elemeinek száma.

A (20) egyenlet második tagjának bal oldala nem-centrális t -eloszlású valószínűségi változó.

Ennek magyarázata:

A t -eloszlás képlete ismeretes: $t = \frac{Z}{\sqrt{\chi^2/v}} = \frac{\bar{x} - \mu}{s/\sqrt{n}}$.

A nem-centrális t -eloszlás alapképlete a következő:

$$t_{nc}(\delta) = \frac{Z + \delta}{\sqrt{\chi^2/v}} = \frac{Z + \delta}{s/\sigma} = \frac{\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} + \delta}{s/\sigma}, \quad (21)$$

ahol δ a nem-centralitási paraméter.

Ennek megfelelően a (20) egyenlet bal oldala nem-centrális t -eloszlású valószínűségi változó

$n-1$ szabadsági fokkal és $\delta = -Z_{1-\gamma}\sqrt{n}$ nem-centralitási paraméterrel, azaz

$$P[t_{nc}(n-1, -Z_{1-\gamma}\sqrt{n}) \geq -k\sqrt{n}] = 1 - \alpha.$$

A végeredményt ekvivalens átalakítások útján kapjuk⁵⁵ (p. 465):

$$\begin{aligned} P[t_{nc}(n-1, -Z_{1-\gamma}\sqrt{n}) \geq -k\sqrt{n}] &= 1 - P[t_{nc}(n-1, Z_{1-\gamma}\sqrt{n}) \geq k\sqrt{n}] = \\ &= P[t_{nc}(n-1, Z_{1-\gamma}\sqrt{n}) \leq k\sqrt{n}] = 1 - \alpha \end{aligned} \quad (22)$$

azaz az ismeretlen k paraméter a mintaelemszám (n), az $1-\gamma$ sokaságbeli hányad és az $1-\alpha$ valószínűség ismeretében a nem-centrális t -eloszlás segítségével számítható.

Természetesen a fenti levezetéshez hasonlóan járunk el, ha egy normális eloszlású sokaság alsó tolerancia-intervallumára vagyunk kíváncsiak. Ekkor az alsó tolerancia-határt $x_L = \bar{x} - ks$ alakban keressük, kiindulási képletünk pedig így alakul:

$$P[P(X \geq x_L) \geq 1 - \gamma] = 1 - \alpha,$$

azaz $1 - \alpha$ annak valószínűsége, hogy a sokaság legalább $1 - \gamma$ hányada az alsó határérték (x_L) felett található.

Az előző pontban ismertetett levezetéssel analóg módon kapott végeredmény k értékére természetesen megegyezik a felső határ esetén kapottal:

$$P[t_{nc}(n-1, Z_{1-\gamma} \sqrt{n}) \leq k \sqrt{n}] = 1 - \alpha.$$

V.1.2. Egyoldali tolerancia-tartomány lineáris regresszió esetén

Az egyenes körüli tolerancia-tartomány számításának részletes leírása Owen⁵⁵ (p. 462-463) cikkében található. Gryze és munkatársai⁵⁶ egy összefoglaló jellegű cikkben elemzik a lineáris regresszió esetén használható különböző intervallum-bebecsléseket, kitérve a tolerancia-intervallumok használatára is.

Az egyenes körüli tolerancia-tartomány ilyen módon való számításának feltétele, hogy az adatok azonos varianciájú, normális eloszlású sokaságból származzanak, valamint, hogy a mérések egymástól függetlenek legyenek.

A regressziós eseteknél az adott x -hez tartozó egyoldali alsó tolerancia-határt

$$y_L = \hat{y}_x - ks_y \tag{23}$$

alakban keressük, ahol \hat{y}_x (a továbbiakban rövidítve \hat{y}) a regressziós függvény alapján becsült érték az x időpontban, s_y pedig az y bizonytalanságának szórása.

Az egyenes körüli alsó tolerancia-tartomány azt mutatja, hogy az adott x -ekhez tartozó y függő változó alkotta sokaság legalább $1 - \gamma$ hányada adott ($100(1 - \alpha)\%$) valószínűséggel milyen határérték felett található:

$$P[P(Y \geq y_L) \geq 1 - \gamma] = 1 - \alpha. \tag{24}$$

A levezetés a V.1.1. fejezetben alkalmazott módon, a Z standard normális eloszlásra való áttéréssel történik. Mivel $E(y) = Y$ és $Var(y) = \sigma_y^2$, így $z_L = \frac{y_L - Y}{\sigma_y} = \frac{\hat{y}_x - ks_y - Y}{\sigma_y}$.

Ezt behelyettesítve a (24) képletbe:

$$P[z_L \leq Z_{1-\gamma}] = P\left[\frac{\hat{y} - ks_y - Y}{\sigma_y} \leq Z_{1-\gamma}\right] = P\left[\frac{\hat{y} - Y}{\sigma_y} - Z_{1-\gamma} \leq \frac{ks_y}{\sigma_y}\right] = 1 - \alpha, \quad (25)$$

ahol $Z_{1-\gamma}$ a standard normális eloszlás γ kvantilise.

A fenti képletben szereplő $\frac{\hat{y} - Y}{\sigma_y}$ tag nem normális eloszlású, mivel \hat{y} (a regressziós

egyenes egy pontjának) szórása nem σ_y , hanem $\sqrt{A}\sigma_y$, ahol

$$A = \frac{1}{mp} \left[\frac{1}{q} + \frac{(x - \bar{x})^2}{\sum_j (x_j - \bar{x})^2} \right], \quad (26)$$

ahol m a minták száma az egyes mérési időpontokban, p a mintánként ismételt mérések száma, q pedig a mérési időpontok száma.

A végeredményt a (25) képlet átalakításával kapjuk:

$$P\left[\frac{\hat{y} - Y}{\sqrt{A}\sigma_y} - \frac{Z_{1-\gamma}}{\sqrt{A}} \leq \frac{ks_y}{\sqrt{A}\sigma_y}\right] = P\left[\frac{\frac{\hat{y} - Y}{\sqrt{A}\sigma_y} - \frac{Z_{1-\gamma}}{\sqrt{A}}}{s_y/\sigma_y} \leq \frac{k}{\sqrt{A}}\right] = P\left[t_{nc}\left(\nu, -\frac{Z_{1-\gamma}}{\sqrt{A}}\right) \leq \frac{k}{\sqrt{A}}\right] = 1 - \alpha, \quad (27)$$

ahol t_{nc} a nem-centrális t -eloszlás ν szabadsági fokkal és a $-\frac{Z_{1-\gamma}}{\sqrt{A}}$ nem-centralitási

paraméterrel. Az egyenletből a k paraméter értéke a mérési időpontok (q), a minták (m) és az ismétlések (p) száma, valamint az $1-\gamma$ sokaságbeli hányad és a $100(1-\alpha)\%$ valószínűség ismeretében a nem-centrális t -eloszlás segítségével meghatározható.

V.2. A tolerancia-intervallum stabilitási adatokra való alkalmazásának irodalmi előzményei

A gyógyszerkészítmények eltarthatósági idejének becslése - ahogy a II.3.1.2. fejezetben már említettem - az adatokat megfelelően leíró degradálódási függvény megtalálásán, valamint az ismereteink pontatlanságát kifejező bizonytalansági tartomány számításán alapul. Az ICH Q1E Guideline⁴ előírása szerint ez a számítandó bizonytalansági tartomány a konfidencia-sáv. Az ICH Guideline előírásait érő kritikák egyik fő csoportja épp a konfidencia-sáv jogosságát vitatja (ld. II.4.4. fejezet). A szerzők egy része a konfidencia-sáv helyett a tolerancia sávot ajánlja, mely arra a - véleményem szerint is helyes - kérdésre adja meg a választ, hogy a jövőbeli tabletták zömének (pl. 99%-ának) hatóanyag-tartalma adott valószínűséggel (pl. 95%) milyen tartományban található.

A következőkben röviden ismertetem a stabilitásvizsgálatokhoz kapcsolódó, irodalomban fellelhető tolerancia-intervallum számításokat.

V.2.1. Chow és Shao kutatásai

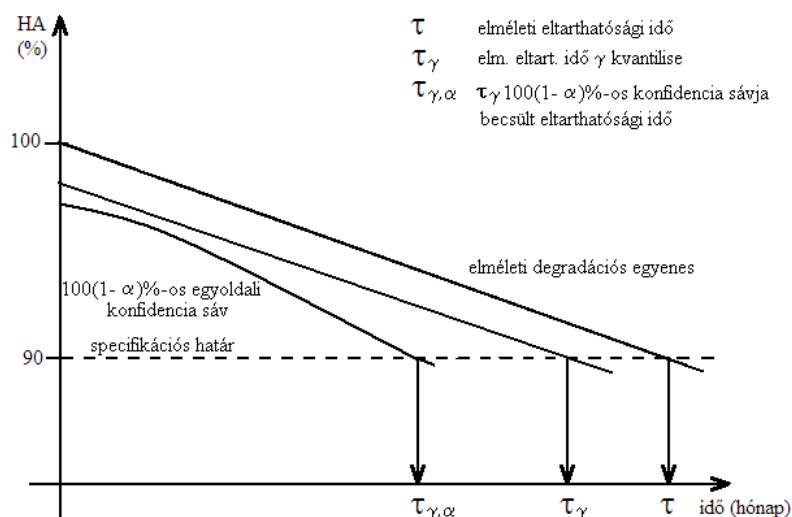
Chow és Shao^{14(12.fejezet),33,36} több írásában is foglalkozik a tolerancia-sáv stabilitásvizsgálatokbeli alkalmazásával.

Elképzelésük szerint a becsült eltarthatósági idő ($\tau_{\gamma,\alpha}$) az elméleti eltarthatósági idő (τ) γ kvantilisének $100(1-\alpha)\%$ -os alsó konfidencia-intervalluma, ahol γ egy kis pozitív konstans¹⁴ (p. 405):

$$P\{\tau_{\gamma,\alpha} \leq \tau\} \geq 1 - \alpha, \quad (28)$$

$$P\{\tau \leq \tau_{\gamma}\} = \gamma \quad (29)$$

Az 5. ábra az eltarthatósági idő tolerancia-sáv alkalmazásán alapuló becslését szemlélteti. (Vessük össze az 1. ábrával, mely a lejáratási idő konfidencia-sávon alapuló becslését mutatja!)



5. ábra: Az eltarthatósági idő becslése a tolerancia-intervallumos szemlélet szerint

Chow és Shao tolerancia-intervallum számítása a gyógyszerkészítményre vonatkozó ún. igazított alsó specifikációs határ számítására épül:

$$\tilde{y}_{spec}(t) = y_{spec} + c_K(\gamma, \alpha) z_\gamma \sqrt{v(t)}, \quad (30)$$

ahol y_{spec} a gyógyszerkészítményre vonatkozó specifikációs határ,

$c_K(\gamma, \alpha)$ a tolerancia faktor,

z_γ a standard normális eloszlás γ kvantilise,

$v(t)$ a stabilitási adatok adott (t) időpontbeli szórása.

A tolerancia faktort pedig az alábbi képlet alapján számítják:

$$c_K(\gamma, \alpha) = \frac{1}{\sqrt{K} z_\gamma} t(\alpha, K-1, \sqrt{K} z_\gamma), \quad (31)$$

ahol K a sarzsok száma, $t(\alpha, K-1, \sqrt{K} z_\gamma)$ pedig a nem-centrális t -eloszlás felső α kvantilise, $K-1$ szabadsági fokkal és $\sqrt{K} z_\gamma$ nem-centralitási paraméterrel.

Összevetve a (31) egyenletet a (22) egyenlettel látható, hogy a Chow-ék által használt tolerancia faktor analóg a normális eloszlású sokaság elemeinek nagy részének elhelyezkedésére vonatkozó tolerancia-intervallum képletével, ami nem helytálló, mivel a stabilitásvizsgálatnál nem egy adott sokaság, hanem az adatokra illeszthető egyenes tolerancia-intervallumáról beszélünk.

* Megjegyzem, hogy Chow és Shao a specifikációs határt η -val, a sokaság adott hányadát ε -nal jelölik.

Összefoglalva tehát megállapítható, hogy Chow és Shao helyesen ismerték fel a tolerancia-intervallum alkalmazásának szükségességét, számítási összefüggéseik azonban rosszak, mivel a stabilitási adatokra mechanikusan alkalmazzák a legegyszerűbb, a normális eloszlású sokaság elemeinek nagy részének elhelyezkedésére vonatkozó eljárás képletét (ld. (22) egyenlet) az egyenes esetén használandó regressziós képlet (ld. (27) egyenlet) helyett.

V.2.2. Kiermeier eredményei

A tolerancia-intervallum stabilitásvizsgálatoknál való alkalmazásának lehetőségét (Chow és Shao kezdetleges eredményei után jó húsz évvel) 2004-ben Kiermeier és kutatócsoportja vetette fel újra. Az ő kutatásaik egybeestek az én kezdeti eredményeimmel, melyet azonban akkor már nem sikerült folyóiratban publikálnom, mivel személyes okokból (gyermekeket szültem) szüneteltetnem kellett a kutatómunkát.

A Kiermeier és társai cikkében⁴² szereplő tolerancia-intervallumos modell egy sarzsot vesz figyelembe, s azon belül pedig egyetlen ingadozási forrással, a tabletták különbözőségével számol. A többi varianciakomponens közül az analízis hibáját az inhomogenitás mellett elhanyagolhatónak tekinti (e feltétel teljesülésére egy adott mérési pontban több ismételt mérés végzését ajánlja, hisz úgyszólván az átlagot használjuk fel az elemzéshez), míg az analízis időpontok különbözősége okozta ingadozással egyáltalán nem törődik. A cikkben ismertetett számítás - azaz, hogy egy sarzs jövőbeli mérési adatainak adott hányada (pl. 99%-a) adott (pl. 95%) biztonsággal milyen intervallumban várható - korrekt.

Kiermeier a disszertációjában⁵⁷ több sarzsra is kiterjeszti számításait, de az analízis hibáját továbbra sem tekinti külön ingadozási forrásnak.

V.3. A fejezet célkitűzései

Az előző fejezetekből látható, hogy a tolerancia-intervallum stabilitási adatokra való alkalmazása, bár a stabilitásvizsgálat valóságos feladatát tekintve a leginkább adekvát megközelítés, még egyáltalán nem kiforrott az irodalomban. Ezért éreztem szükségesnek a stabilitásvizsgálatokhoz alkalmazható tolerancia-intervallumos modellek további kutatását. A következőkben ismertetett eredmények időben hosszú és gyermekeim születése után többször újrakezdett kutatási munka gyümölcsei.

A stabilitási adatok értékelésére alkalmazható tolerancia-intervallumos modellt két ponton fejlesztettem tovább:

1. Egyrészt a figyelembe vett ingadozási források számát bővítve egy több-ingadozás-forrásos tolerancia-intervallumon alapuló modellt javasoltam.

Kiermeier modelljeiben az egy időpontban vett stabilitási minták különbözőségét okozó két varianciakomponens közül csak a tabletták különbözőségét vette figyelembe. Tény, hogy a sarzson belüli inhomogenitás mértéke a valóságban általában tényleg nagyobb, mint az analitikai mérés ismételési hibája, ám ez utóbbi mértéke sem tekinthető elhanyagolhatónak. Több varianciakomponens figyelembevétele pedig az eltarthatósági idő pontosabb becslését teszi lehetővé.

2. A másik fontos és az irodalomban eddig még nem vizsgált kérdés, hogy a stabilitásvizsgálat értékelése során a mért vagy a valódi (tényleges) hatóanyag-tartalom ingadozását vizsgáljuk.

Mért hatóanyag-tartalom alatt az adott vizsgálati időpontban, az adott mintából, az adott ismételt analízisnél kapott mérési eredményt értem.

Valódi hatóanyag-tartalom alatt a tabletták tényleges hatóanyag-tartalmát értem, melyet azonban az analízis hibája miatt nem ismerünk. Ugyanakkor ez az az érték, ami a páciens/beteg szervezetébe kerül, amikor beveszi a gyógyszert.

Véleményem szerint a valódi hatóanyag-tartalomra kellene ingadozási tartományt adni, hiszen a beteget (orvost) a bevett tabletták tényleges hatóanyag-tartalma érdekli. Megjegyzem, hogy Kiermeier-nél ez a probléma nem merülhetett fel, mivel ő hibamentes analízist feltételezett, így nála a valódi és a mért hatóanyag-tartalom megegyezik.

V.4. Tolerancia-intervallumos modellek a stabilitásvizsgálatra

Gyógyszerkészítmények stabilitásvizsgálatának tolerancia-intervallumon alapuló értékelésénél abban szeretnénk adott valószínűséggel biztosak lenni, hogy a gyógyszerkészítmény hatóanyag-tartalma a tabletták adott (nagy) részénél a lejáratási időpontban az alsó specifikációs határ felett van. Ennek az elvárásnak megfelelően az eltarthatósági idő becslése a degradálódási egyenes körüli alsó egyoldali tolerancia-tartomány számításával történik.

V.4.1. Az eltarthatósági idő becslése egyetlen ingadozás-forrást tartalmazó modellre („A₁” modell)

Először vizsgáljuk meg a legegyszerűbb stabilitásvizsgálati modellt, mely egy sarzs adatait vizsgálja, és azon belül egy ingadozási forrást (a tabletták különbözőségét, ha emellett az analitikai mérés bizonytalanságát elhanyagolhatónak tekintjük vagy az analitikai mérési hibát, ha az inhomogenitástól tekintünk el) tartalmaz. Ez az eset matematikailag az egyváltozós lineáris regresszióknak felel meg, melynek elemzését az V.1.2. fejezetben már részletesen bemutattam. Így az ekkor érvényes modell a (27) egyenlet, azzal a kitételrel, hogy az ott szereplő A paramétert itt A_1 -gyel jelölöm.

Fontos külön is hangsúlyozni, hogy ez egy, a végletekig leegyszerűsített stabilitásvizsgálati modell, így az eltarthatósági idő valóságghú becslésére nem alkalmas. A továbbiakban csak kiindulási ill. összehasonlítási alapként tekintem.

Megemlítem továbbá, hogy ebből a modellből kapható meg Kiermeier és társai modellje⁴² az alábbi két megszorítással:

- a feltételezett egyetlen ingadozási forrás a tabletták különbözősége (az analitikai mérés hibáját emellett elhanyagolhatónak tekintik)
- a stabilitásvizsgálat során ugyan minden mérési időpontban több mintát vesznek, de mindegyiket csak egyszer analizálják.

V.4.2. Az eltarthatósági idő becslése két ingadozás-forrást tartalmazó modellekre

Az ebben a fejezetben ismertetett modelleknél a stabilitási adatok bizonytalanságát befolyásoló tényezők közül (ld. II.5. fejezet) az alábbi két ingadozási forrást vettem figyelembe:

- a tabletták különbözősége, azaz a sarzson belüli inhomogenitás (σ_{inhom}^2)
- az analitikai mérés ismételhetségi hibakomponense (röviden: az analízis hibája, σ_{ism}^2)

A két (több) ingadozás-forrást tartalmazó stabilitásvizsgálati modellek azonos varianciájú, normális eloszlású adatokat feltételeznek, továbbá azt hogy, hogy a vizsgált ingadozási források egymástól függetlenek. Az ingadozási források függetlensége jelen esetben azt jelenti, hogy az analízis hibájának nagysága nem függ össze a tabletták hatóanyag-

tartalmával, azaz pl. a kisebb hatóanyag-tartalom nem jelent kisebb analitikai hibát. A stabilitásvizsgálatoknál ez a feltétel teljesül.

Fontos továbbá, hogy a két (több) ingadozás-forrást tartalmazó stabilitásvizsgálati modellek elemzéséhez vizsgálati időpontonként többszöri mintavétel és az egyes minták többszöri analízise szükséges, mivel az ingadozási források numerikus szétválasztása csak ebben az esetben lehetséges. Ezt a kikötést már a stabilitásvizsgálat tervezésekor figyelembe kell venni.

A két ingadozási forrást (a tabletták különbözőségét valamint az analitikai hibát) tartalmazó stabilitásvizsgálati modellek tolerancia-intervallumának számítása analóg az egy ingadozási forrásos modellével. Az alapképlet továbbra is a (27) egyenlet marad, csak a nevezőben szereplő A paraméter értéke módosul az adott modellnek megfelelően. Az egyes modellekhez tartozó A paraméter értékének meghatározásához minden modell esetén az alábbi három szórás ill. variancia értékét kell definiálni :

1. a tolerancia-intervallum képletében ((23) egyenlet) szereplő szórás (s_y)

E szórás - önkényes - definíciója attól függ, hogy az adott modellben a mért vagy a valódi hatóanyag-tartalomra kívánunk-e ingadozási tartományt adni. Ha az intervallum a mért hatóanyag-tartalom ingadozását mutatja, akkor a szórásnégyzet a két ingadozási forrás összegeként adódik ($s_y^2 = s_{ism}^2 + s_{inhom}^2$), ha azonban a valódi hatóanyag-tartalom elhelyezkedésére vagyunk kíváncsiak, akkor ez a szórásnégyzet csak a tabletták különbözőségét számszerűsítő inhomogenitási tagból áll ($s_y^2 = s_{inhom}^2$).

2. az egyedi stabilitási adat varianciája ($\sigma_{y_{jkl}}^2$),

ahol $j=1\dots q$ a vizsgálati időpontok számát, $k=1\dots m$ az egyes vizsgálati időpontokban a minták számát és $l=1\dots p$ az egyes minták ismételt analíziseinek számát jelenti

3. a stabilitási adatokra illesztett egyenes varianciája egy adott x időpontban ($\sigma_{\hat{y}}^2$)

Ez utóbbi két variancia ($\sigma_{y_{jkl}}^2$ és $\sigma_{\hat{y}}^2$) között épp az A paraméter teremti meg a kapcsolatot (ld. (26) egyenlet), azaz megállapítható, hogy a két ingadozás-forrást tartalmazó modellek kulcsfontja az A paraméter meghatározása. A fejezet további részében részletesen is ismertetem ennek menetét.

V.4.2.1. Egy sarzs, a sarzson belül kétféle ingadozási forrás (inhomogenitás és analitikai hiba) – a mért hatóanyag-tartalom elhelyezkedésének vizsgálata („A₂” modell)

Az egyedi (mért) stabilitási adat összvarianciája ($\sigma_{y_{jk}}^2$, ezután rövidítve σ_y^2 -ként jelölöm) egyszerűen a két vizsgált ingadozási forrás összegeként adódik:

$$\sigma_y^2 = \sigma_{\text{ism}}^2 + \sigma_{\text{inhom}}^2 \quad (32)$$

A fenti képlet abban az esetben érvényes, ha a minta egyetlen tablettát tartalmaz. Ha a minta tabletták homogenizált elegyét tartalmazza, akkor a (32) egyenlet az alábbiak szerint módosul:

$$\sigma_y^2 = \sigma_{\text{ism}}^2 + \sigma_{\text{inhom}}^2 / d, \quad (33)$$

ahol d a mintánkénti tabletták száma.

A stabilitási adatokra illesztett egyenes varianciájának számítása az alábbi módon történhet. Illesszünk egyenest a minták és az analízisek szerint átlagolva kapott $y_{j..}$ adatokra! Ekkor:

$$\sigma_{y_{j..}}^2 = \sigma_{y_{j..}}^2 \left[\frac{1}{q} + \frac{(x - \bar{x})^2}{\sum_j (x_j - \bar{x})^2} \right] = \frac{\sigma_{\text{ism}}^2 + p\sigma_{\text{inhom}}^2}{mp} \left[\frac{1}{q} + \frac{(x - \bar{x})^2}{\sum_j (x_j - \bar{x})^2} \right] = A_1 (\sigma_{\text{ism}}^2 + p\sigma_{\text{inhom}}^2) \quad (34)$$

ahol $\sigma_{y_{j..}}^2$ az $y_{j..}$ átlag varianciája: $\sigma_{y_{j..}}^2 = \frac{\sigma_{\text{ism}}^2}{mp} + \frac{\sigma_{\text{inhom}}^2}{m}$. Ha ezt a variancia értéket megszorozzuk mp -vel, a stabilitásvizsgálatba bevont összes adat varianciáját kapjuk, ami természetesen nem azonos az egyedi stabilitási adat varianciájára kapott σ_y^2 értékkel.

Az A_2 paraméter értéke a (32) és (34) egyenletekkel meghatározott varianciák hányadosa lesz:

$$A_2 = \frac{\sigma_{y_{j..}}^2}{\sigma_y^2} = A_1 \frac{\sigma_{\text{anal}}^2 + p\sigma_{\text{inhom}}^2}{\sigma_{\text{anal}}^2 + \sigma_{\text{inhom}}^2} \quad (35)$$

A (27) egyenletben A paraméter helyére A_2 értékét helyettesítve, a mért hatóanyag-tartalomra vonatkozó tolerancia-intervallum számítható.

Az itt bemutatott modellt a továbbiakban röviden „A₂” modellnek nevezem.

V.4.2.2. Egy sarzs, a sarzson belül kétféle ingadozási forrás (inhomogenitás és analitikai hiba) – a valódi hatóanyag-tartalom elhelyezkedésének vizsgálata („ A_2^* ” modell)

Mivel a modell felépítése megegyezik az előző pontban bemutatott esettel, hiszen továbbra is két különböző ingadozási forrást - az analízis hibáját és a tabletták különbözőségét - vesszük figyelembe, a stabilitási adatokra illesztett egyenes varianciája (σ_y^2) most is ugyanúgy számítandó, mint az előző esetben (ld. (34) egyenlet).

A különbség az egy adott stabilitási adat varianciájának ($\sigma_{y_{jkl}}^2$) becslésében lesz, hiszen most a gyógyszerkészítmény tényleges (nem pedig a mért) hatóanyag-tartalmának ingadozását vizsgáljuk. A valódi (tényleges) stabilitási adatok ingadozását a két vizsgált varianciakomponens közül csak a tabletták különbözősége okozza. (Mivel nem mért értékekről van szó, az analitikai mérés hibája okozta bizonytalanságot nem kell figyelembe venni.) A fentieknek megfelelően az A paraméter nevezőjében levő variancia valamint az A_2^* paraméter értéke a következőképpen alakul:

$$\sigma_{y^*}^2 = \sigma_{\text{inhom}}^2, \quad (36)$$

valamint

$$A_2^* = \frac{\sigma_y^2}{\sigma_{y^*}^2} = \frac{1}{mp} \left[\frac{1}{q} + \frac{(x - \bar{x})^2}{\sum_j (x_j - \bar{x})^2} \right] \frac{\sigma_{\text{ism}}^2 + p\sigma_{\text{inhom}}^2}{\sigma_{\text{inhom}}^2} = A_1 * \frac{\sigma_{\text{ism}}^2 + p\sigma_{\text{inhom}}^2}{\sigma_{\text{inhom}}^2}. \quad (37)$$

A (27) egyenletben A paraméter helyére A_2^* értékét helyettesítve, a valódi hatóanyag-tartalomra vonatkozó tolerancia-intervallumot kapjuk.

Erre a modellre a későbbiekben „ A_2^* ” néven hivatkozom.

Összefoglalva a két - két varianciakomponenst tartalmazó - modell közti különbséget elmondható, hogy az, hogy a mért vagy a valódi (tényleges) hatóanyag-tartalom ingadozását vizsgáljuk, csak az A paraméter nevezőjében szereplő variancia értékére van hatással. Ugyanis ez az a variancia, ami az egyedi stabilitási adat ingadozását mutatja, ennek értéke függ attól, hogy az a változó, aminek elhelyezkedését vizsgáljuk, hány varianciakomponenst tartalmaz: a valódi érték csak az inhomogenitást, a mért érték pedig az inhomogenitás mellett az analitikai hibát is.

Az A paraméter számlálójában szereplő variancia a mért stabilitási adatokra illesztett egyenes bizonytalanságát mutatja. Ennek értéke a modell felépítésétől függ (azaz, hogy hány

ingadozási forrást vesszünk figyelembe), attól viszont független, hogy a valódi vagy a mért hatóanyag-tartalom elhelyezkedésére kívánunk-e tolerancia-intervallumot adni.

V.5. Számolás a tolerancia-intervallumon alapuló stabilitásvizsgálati modellekkel

A tolerancia-intervallumos szemléletben a gyógyszerkészítmény becsült eltarthatósági ideje az az időpont, ahol tabletták előre meghatározott hányadának $(1-\gamma)$ hatóanyag-tartalmát adott valószínűséggel $(100(1-\alpha)\%)$ tartalmazó tolerancia-intervallum alsó határa (y_L) metszi a készítményre vonatkozó specifikációs határt (y_{spec}) .

Az egyoldali, alsó tolerancia-határt $y_L = \hat{y} - ks_y$ alakban képzeljük el (ld. (23) egyenlet). A feladat a fenti képletben található k érték meghatározása, mely az előző fejezetekben ismertetett módon, a nem-centrális t -eloszlás segítségével történik:

$$P\left[t_{nc}\left(\nu, -\frac{Z_{1-\gamma}}{\sqrt{A}}\right) \leq \frac{k}{\sqrt{A}}\right] = 1 - \alpha, \quad (38)$$

ahol az egyes modellekhez tartozó A paraméter értékeit a (26), (35) és (37) egyenletek tartalmazzák.

A fenti képletből látható, hogy a nem-centrális t -eloszlás k/\sqrt{A} kvantilisének kiszámításához az A paraméter értéke valamint a megfelelő szabadsági fok (ν) meghatározására van szükség.

Az egy ingadozás-forrást tartalmazó modellben A_1 értéke a stabilitásvizsgálati terv alapján a vizsgálati időpontok számának (q) , valamint az egyes vizsgálati időpontokhoz tartozó minta- (m) és ismétlésszámok (p) ismeretében könnyen meghatározható (ld. (26) egyenlet).

A két ingadozás-forrást tartalmazó modellekben A_2 és A_2^* paraméterek kiszámításához azonban már σ_{ism}^2 és σ_{inhom}^2 varianciák értékére is szükség van. A két varianciakomponens becsült értéke a stabilitási adatokból az alábbi varianciaanalízis-moddal nyerhető:

$$y_{jkl} = \mu + \beta_j x_j + \delta_{k(j)} + \varepsilon_{jkl}, \quad (39)$$

ahol μ a tengelymetszet, $\beta_j x_j$ a j -edik hónap hatása, $\delta_{k(j)}$ a j -edik hónap k -adik mintájának hatása (ez tartalmazza a tabletták különbözőségét) és ε_{jkl} az analitikai hiba. Két faktorunk van: a rögzítettnek tekintett vizsgálati hónap és a véletlen jellegű minta. Mivel minden

időpontban más-más mintákat analizálnak, hierarchikus osztályozásról van szó: a minta faktort kell a hónap faktorba ágyazni.

A nem-centrális t -eloszlás szabadsági foka (ν értéke a (38) egyenletben) az adott modellhez tartozó s_y^2 szórásnégyzet szabadsági fokával lesz egyenlő.

Egy (össz)ingadozás feltételezése esetén („A₁” modell) egyszerű lineáris regresszióval van dolgunk, ekkor az s_y^2 reziduális szórásnégyzet szabadsági foka $N - 2$, ahol N a stabilitási adatok száma ($N = qmp$).

Két ingadozási forrás esetén azonban már bonyolultabb a helyzet, ugyanis s_y^2 értéke szórásnégyzetek lineáris kombinációjaként adódik ($\sum a_i s_i^2$). Ilyenkor a megfelelő szabadsági fok a Satterthwaite-közelítéssel⁵⁸ számolható:

$$\nu = \frac{\left(\sum_i a_i s_i^2 \right)^2}{\sum_i \frac{(a_i s_i^2)^2}{\nu_i}}, \quad (40)$$

A számításhoz szükséges együtthatókat (a_i) az alábbi formulákból kapjuk:

$$\text{„A}_2\text{” modell esetén: } s_y^2 = s_{\text{ism}}^2 + s_{\text{inhom}}^2 = s_{\text{ism}}^2 + \left(\frac{s_{\text{minta}}^2 - s_{\text{ism}}^2}{p} \right) = s_{\text{ism}}^2 \left(1 - \frac{1}{p} \right) + \frac{s_{\text{minta}}^2}{p} \quad (41)$$

$$\text{„A}_2^*\text{” modell esetén: } s_y^2 = s_{\text{inhom}}^2 = \frac{s_{\text{minta}}^2 - s_{\text{ism}}^2}{p} \quad (42)$$

A Satterthwaite-közelítés sajátossága miatt a két ingadozás-forrást tartalmazó modelleknél a nem-centrális t -eloszlás szabadsági foka általában nem egész szám lesz.

V.6. Szimulációs számítások

V.6.1. A szimuláció célja

Az előző fejezetekben ismertetett három, tolerancia-intervallumon alapuló stabilitásvizsgálati modell összehasonlítására a MATHEMATICA programcsomaggal⁵⁹ szimulációs számításokat végeztem. A különböző modellekkel számított tolerancia-intervallumok szélessége (és így a becsült eltarthatósági idő) elvileg analitikus összefüggésekkel is összehasonlítható, ám a módszerekben alkalmazott közelítések miatt (pl. Satterthwaite-

közelítés) mindez reménytelenül bonyolult lenne. Ezért döntöttem a szimulációs vizsgálat mellett. A szimuláció célja a bemutatott modellek statisztikai tulajdonságainak vizsgálata: különös tekintettel a modellek lefedési valószínűségére (hogy pl. az a tartomány, amelyben a tabletták 99%-ának hatóanyag-tartalma pl. 95% valószínűséggel megtalálható, tényleg ilyen valószínűséggel tartalmazza az eredményeket), valamint a tolerancia-intervallumok szélességének alakulására az egyes modellek ill. paraméter-kombinációk esetén.

Jogosan merülhet fel a kérdés, hogy a szimulációs eredményekből miért nem a becsült eltarthatósági időt határoztam meg. Az eltarthatósági idő becslése valóban logikusabb lenne, ám jóval több számítást is igényel: a függő változó (itt a hatóanyag-tartalom) egy adott időpontbeli tolerancia-intervallumának meghatározása helyett annak sokszori megismétlését a teljes degradációs függvény mentén. A számomra fontos kérdés, azaz a különböző stabilitásvizsgálati modellek jellemzői, megfelelően tanulmányozhatók akkor is, ha a becsült eltarthatósági idő helyett a tolerancia-intervallumok egy adott időpontbeli szélességét vizsgáljuk.

V.6.2. A szimulációs terv leírása

A (38) egyenletből látható, hogy a tolerancia-intervallum szélességét befolyásoló k paraméter értéke a tabletták γ hányadának, az α valószínűségnek és az A paraméter értékének ismeretében számítható. Figyelembe véve a (26), (35) és (37) egyenleteket, a különböző A paraméterek értékét az alábbi tényezők befolyásolják:

- a mérési időpontok száma (q) és elhelyezkedése (x_j),
- a minták száma (m) és az egyes mintákból végzett ismételt analízisek száma (p) egy adott vizsgálati időpontban
- a tabletták különbözőségére valamint az analitikai hiba mértékére vonatkozó varianciakomponensek értéke (σ_{inhom}^2 és σ_{ism}^2).

E tényezők közül a $\gamma = 0,01$ hányadot és az $\alpha = 0,05$ valószínűséget rögzítettem, azaz az összes szimulációs eredmény olyan alsó, egyoldali tolerancia-intervallumot mutat, mely a tabletták 99%-át 95%-os valószínűséggel tartalmazza.

A többi befolyásoló tényező értékét az alábbiak szerint adtam meg ill. változtattam a szimuláció során:

1. Minden esetben lineáris degradációt feltételeztem. A kiindulási hatóanyag-tartalom (β_0) mindig 100% volt, a degradációs ráta (β_1) pedig 0,1 vagy -0,2 %/hónap volt.
2. Két vagy három évig tartó stabilitásvizsgálatokat elemeztem. A vizsgálati időpontok gyakorisága az ICH Guideline előírásainak megfelelően alakult: $x_j = 0, 3, 6, 9, 12, 18, 24$ vagy 36 hónap.
3. A stabilitásvizsgálatok gyakorlati kivitelezésénél indokolt gazdasági megfontolások miatt az egyes vizsgálati időpontokban végzett mérések számát ($r = mp$) 10-ben maximalizáltam. Ezt a korlátot figyelembe véve az alábbi három minta- (m) és ismételt analízis számmal (p) dolgoztam: $m = p = 3$; $m = 2$ és $p = 5$; $m = 5$ és $p = 2$.
4. Gyógyszeripari tapasztalatok alapján a két vizsgált varianciakomponens értékét a következők szerint változtattam: $\sigma_{\text{inhom}}^2 = 1, 2, 4$ és $\sigma_{\text{ism}}^2 = 1, 2$. A varianciák mértékegysége hatóanyag-százalék a négyzetben.

A szimuláció során minden egyes paraméter-kombinációnál 1000-szeri futtatást végeztem.

V.6.3. A programozandó feladat rövid leírása

A MATHEMATICA szoftverben⁵⁹ írt programom* egy futási ciklusa az alábbi lépésekből állt:

1. A stabilitásvizsgálati adatsor előállítás.

Az adatok generálása során először az adott paraméterű lineáris degradálódási függvénnyel előállítottam a tárolási hónapokhoz tartozó hatóanyag-tartalom értékeket tartalmazó vektort. Ez lett a hatóanyag-tartalom adott időpontbeli várható értéke. Ezután a vektor egyes elemeihez mintánként, majd a mintákon belül ismételt analízisenként normális eloszlású véletlen számokat („hibákat”) adtam az adott futási ciklushoz tartozó varianciákkal.
2. A generált adatok elemzése: az adott futtatás során kapott stabilitási adatok becült varianciakomponenseinek meghatározása.

Ez egyrészt a lineáris degradálódási függvény, másrészt a (39) egyenletben bemutatott varianciaanalízis modell illesztését jelentette.

* Köszönöm Bánfai Balázsnak a programozásban nyújtott segítségét.

3. A tolerancia-sáv adott időpontbeli szélességének (ks_y) számítása.

A tolerancia-sáv k paraméterének értéke a (38) egyenlet szerint a nem-centrális t -eloszlás segítségével határozható meg, így ez a programrészlet a nem-centrális t -eloszlás függvény alkalmazására épült.

(Itt jegyzem meg, hogy a szimulációs számítások elvégzésére épp azért választottam a MATHEMATICA szoftvert, mert ennek programba épített nem-centrális t -eloszlás függvénye képes kezelni a nem egészszámú szabadsági fokokat is.)

V.6.4. A szimulációs eredmények és értékelésük

A továbbiakban minden egyes szimulációs esetet (paraméter-kombinációt) egy-egy kóddal jelölök, mely a négy legfontosabb paraméter értékét a következő sorrendben tartalmazza:

- a minták száma az egyes vizsgálati időpontokban (m)
- a mintánkénti ismételt analízisek száma (p)
- a tabletták különbözőségére vonatkozó variancia feltételezett értéke (σ_{inhom}^2)
- az analízis hibájának feltételezett értéke (σ_{ism}^2)

Így például az 5-2-2-1 kód azt a szimulációs esetet jelenti, ahol mérési időpontonként 5-5 mintát vizsgálnak, minden egyes mintát kétszer elemeznek, valamint a feltételezett varianciakomponensek értéke a következő: $\sigma_{\text{inhom}}^2 = 2$ és $\sigma_{\text{ism}}^2 = 1$.

(A könnyebb áttekinthetőség kedvéért a későbbi eredménytáblázatok is tartalmazzák a kódok jelentését.)

Az egyes szimulációs esetek 1000-szeri futtatásából mindhárom stabilitásvizsgálati modell („A₁”, „A₂” és „A₂^{*}” modellek) esetén az alábbi két értéket határoztam meg:

- az átlagos tolerancia-intervallum-szélességet (ks_y) a 24. vagy a 36. vizsgálati hónapnál
- a fenti értékre vonatkozó lefedési valószínűséget (*coverage probability*)

A lefedési valószínűség (*coverage probability*) azon esetek aránya, ahol a tolerancia-intervallum számított értéke (y_L) kisebb, mint az elméleti tolerancia-intervallum ($y_{L,z}$ az „A₁” és az „A₂” modellek esetén valamint $y_{L,z}^*$ az „A₂^{*}” modell esetén). Az elméleti tolerancia-intervallum az a határérték, mely felett a tabletták 99%-ának hatóanyag-tartalma

található. Ha a mért hatóanyag-tartalom értékeket vizsgáljuk, akkor értéke az alábbi képlettel számítható:

$$y_{L,z} = \hat{Y}_t - z_{0,99} \sqrt{\sigma_{\text{inhom}}^2 + \sigma_{\text{ism}}^2} \quad (43)$$

ahol $\hat{Y}_t = \beta_0 + \beta_1 * t$ a várható hatóanyag-tartalom a t időpontban.

A valódi (tényleges) hatóanyag-tartalom vizsgálata esetén az elméleti tolerancia-intervallum a következőképpen számítható:

$$y_{L,z}^* = \hat{Y}_t - z_{0,99} \sqrt{\sigma_{\text{inhom}}^2} \quad (44)$$

A (43) és (44) egyenletekből látható, hogy $y_{L,z}^*$ értéke mindig nagyobb $y_{L,z}$ értékénél, mivel az utóbbinál a gyökjel alatt csak egy varianciakomponens szerepel. Például két éves tárolás, 0,1%/hónap degradációs ráta, valamint $\sigma_{\text{inhom}}^2 = 4$ és $\sigma_{\text{ism}}^2 = 1$ varianciakomponensek esetén az elméleti tolerancia-intervallum értékei a következők: $y_{L,z} = 92,40$ és $y_{L,z}^* = 92,95$.

A lefedési valószínűség a modell jóságát jellemzi. Értékének a lehető legközelebb kell lennie az előírt 0,95-os értékhez, mert így biztosított 95%-os megbízhatósággal, hogy a tabletták adott hányadának (jelen esetben 99%-ának) hatóanyag-tartalma a számított tolerancia-határ felett van.

Mivel csak korlátozott számú szimulációs futtatásra volt lehetőségem, melyekre a lefedési valószínűség (a mintavételi ingadozás miatt) csak közelítőleg adhatja a névleges értéket, meghatároztam a lefedések számának 95%-os konfidencia-intervallumát annak eldöntésére, hogy a névleges értéktől eltérő talált lefedési valószínűség valóságos különbséget tükröz-e. A lefedések száma binomiális eloszlású valószínűségi változó. A lefedések számának 95%-os konfidencia-intervallumát könnyen kiszámíthatjuk, ha a binomiális eloszlást normális eloszlással közelítjük: $\hat{p} \pm z_{0,95} * \sqrt{\frac{\hat{p}(1-\hat{p})}{1000}}$, ahol \hat{p} az 1000 futtatásból számított lefedési valószínűség. A lefedési valószínűség kívánt tartománya tehát $0,95 \pm 0,014$.⁴²

Megjegyzem, hogy az egzakt konfidencia-intervallum számítását azért nem tartom indokoltnak, mert a szimulációs futtatások száma ahhoz viszont elég nagy volt, hogy a normális eloszláson alapuló szokásos közelítést megnyugtatóan lehessen alkalmazni.

V.6.4.1. A szimulációs program verifikálása

Elsőként saját szimulációs programom alkalmazásának igazolását végeztem el a Kiermeier és társai⁴² által publikált szimulációs vizsgálat egy esetének reprodukálásával. A kiválasztott eset jellemzői a következők:

- $\beta_0 = 100$, $\beta_1 = -0,2$, $x_j = 0, 3, 6, 9, 12, 18, 24, 36$ hónap,
- a stabilitási adatok ingadozását jelző variancia, mely jelen esetben csak a tabletták különbözőségét tartalmazza: $\sigma_y^2 = \sigma_{inhom}^2 = 4$,
- az egy vizsgálati időpontban végzett ismétlések száma: $r = 10$.

Mivel Kiermeiernél csak egy feltételezett ingadozási forrás van, mégpedig a tabletták/minták különbözőségét jellemző inhomogenitási variancia, ezért Kiermeier modelljének ellenőrzéséhez a szimulációs programomat egy kicsit módosítani kellett, mégpedig úgy, hogy csak egy (össz)ingadozási szórásnégyzettel számoljon. A módosított program kódja tehát csak háromtagú lesz: $m - p - \sigma_y^2$. Az ismétlésszámoknak is Kiermeier modelljéhez kell igazodnia, azaz az egy időpontban végzett 10 ismétlés, 10 különböző mintát jelent ($m = 10$). Ez viszont ugyanakkor azt is jelenti, hogy ennél az esetnél minden egyes mintát csak egyszer analizálnak ($p = 1$). Összefoglalva a fenti megállapításokat elmondható, hogy a kiválasztott eset a 10-1-4 kódszámú esetnek felel meg.

Természetesen Kiermeier modellje esetén a tolerancia-intervallumos modellek közül csak az „A₁” modell kiszámítására van lehetőség, mivel az „A₂” és „A₂*” modellek két megkülönböztethető ingadozási forrást és 1-nél nagyobb minta- és ismétlésszámot feltételeznek.

A 8. táblázat első két sorából látható, hogy a Kiermeier modelljének megfelelő (10-1-4 kódú) eset szimulációjakor kapott lefedési valószínűség érték pontosan megegyezik a Kiermeier és társai cikkében szereplő értékkel, azaz sikerült Kiermeier szimulációs számításait reprodukálni.

	az eset kódja				lefedési valószínűség az egyes modellek esetén		
	m	p	σ_{inhom}^2	σ_{ism}^2	„A ₁ ”	„A ₂ ”	„A ₂ [*] ”
Kiermeier és társai cikkében ⁴²	10	1	4	-	0,95	-	-
saját programommal	10	1	4	-	0,95	-	-
	2	5	2	2	0,84	0,95	0,99
	5	2	2	2	0,91	0,95	0,99

8. táblázat: Egy kiválasztott eset elemzése Kiermeier és társai⁴² szimulációs vizsgálatából

$$(\beta_0 = 100, \beta_1 = -0,2, t = 36 \text{ hónap}, r = 10, \sigma_y^2 = 4)$$

V.6.4.2. Az új modellek statisztikai tulajdonságainak vizsgálata

Az új modellek bevezetésének értelmére és „jobb voltuk” (a valóságot hűbben leíró voltuk) szemléltetésére Kiermeier kiválasztott szimulációs esetét két megkülönböztethető ingadozási forrás (az inhomogenitás és az analitikai hiba) figyelembevételével is elemeztem. Ennek érdekében a stabilitási adatok (össz)ingadozását jelző varianciát ($\sigma_y^2 = 4$) két egyenlő részre osztottam: $\sigma_{ism}^2 = 2$ és $\sigma_{inhom}^2 = 2$.

Az új modellek alkalmazásának természetesen feltétele, hogy minden vizsgálati időpontban több mintát vegyenek és azokat többször is analizálják. Itt az egy adott vizsgálati időpontban végzett ismétlések ($r = 10$) két lehetséges felbontását vizsgáltam: $m = 5$ és $p = 2$ ill. $m = 2$ és $p = 5$ eseteket.

Ekkor tehát a szimuláció során az adatok generálását két megkülönböztethető ingadozási forrás figyelembevételével végeztem, majd az így kapott adatokat mindhárom, tehát az egyetlen varianciakomponenssel számoló modellel is elemeztem. A kapott lefedési valószínűség értékeket a 8. táblázat utolsó két sora mutatja. Az adatokból megállapítható, hogy:

- az ajánlott két ingadozás-forrásos modellek („A₂” ill. „A₂^{*}” modellek) lefedési valószínűsége az elvárt 0,95 érték körül ill. fölötte van;
- ha viszont a számítás során az ingadozási források struktúráját figyelmen kívül hagyva a stabilitási adatokat csupán egy összingadozással jellemezzük („A₁” modell), akkor a lefedési valószínűség értéke jóval kisebb az elvárt 0,95-től.

(Természetesen az A₁ modelltől nem is vártam el, hogy megfelelően leírja a két

varianciakomponens figyelembevételével generált adatokat, tehát itt előzetesen is 0,95-nél jóval kisebb lefedési valószínűsége számítottam, de kíváncsi voltam, hogy a varianciakomponensek nem megfelelő figyelembevétele mekkora hibát okoz.)

Mindez azt is jelenti, hogy két megkülönböztethető ingadozás-forrás jelenléte esetén az „A₂” és „A₂^{*}” modellek jól leírják a szimulációs esetet, s így reális eltarthatósági idő becslést tesznek lehetővé, míg a csak egy összingadozással számoló „A₁” modell nem alkalmas a két ingadozós szimulációs esetek megfelelő leírására. Így kijelenthető, hogy mivel a valóságban a stabilitásvizsgálatok során mindig több ingadozás-forrás hatása jelentkezik, a gyógyszeripari gyakorlatban az „A₁” modell nem megfelelő a stabilitási adatok leírására.

Kiermeier és társai szimulációs vizsgálatából kiválasztott esetet tovább elemezve a tolerancia-intervallumok átlagos szélességét (ks_y értéke az 1000 futtatás szerint átlagolva) a 9. táblázat tartalmazza.

Az adatokból látható, hogy a tolerancia-intervallum szélessége (egy értéket leszámítva) az „A₁” modellnél a legszűkebb. Ez ugyan egyben hosszabb becsült eltarthatósági időt is jelent, ami nyilván kedvező a gyógyszergyárak számára, de a fentiekben vázolt indokok miatt már tudjuk, hogy az „A₁” modell nem alkalmas ezen adatok valóság-hű (szimuláció-hű) leírására.

Összehasonlítva az új, azaz az „A₂” és „A₂^{*}” modellekkel kapott értékeket megállapítható, hogy a tolerancia-intervallumok szűkebbek az „A₂^{*}” modellnél. Ez várható is volt, mivel az „A₂^{*}” modell az analízis hibájával nem terhelt tényleges hatóanyag-tartalomra ad ingadozási tartományt. Ismételten megjegyzem, hogy a beteg kérdésére épp az „A₂^{*}” modell adja meg a helyes választ, hiszen a beteg a tablettá tényleges hatóanyag-tartalmára kíváncsi, nem pedig az analitikai mérés hibájával terhelt értékre.

az eset kódja				Az ingadozási tartomány átlagos szélessége			
				tolerancia-intervallum			alsó 95%-os konfidencia-intervallum
m	p	σ_{inhom}^2	σ_{ism}^2	„A ₁ ”	„A ₂ ”	„A ₂ [*] ”	
10	1	4	-	5,75	-	-	
2	5	2	2	5,74	6,53	5,95	0,81
5	2	2	2	5,74	5,97	5,21	0,82

9. táblázat: A tolerancia-és konfidencia-intervallumok átlagos szélessége

$$(\beta_0 = 100, \beta_1 = -0,2, t = 36 \text{ hónap}, r = 10, \sigma_y^2 = 4)$$

Annak érdekében, hogy a tolerancia-intervallumos modellekkel kapott eredmények összehasonlíthatók legyenek az ICH jelenlegi szabályozásával⁴, a szimuláció során meghatároztam az alsó, 95%-os konfidencia-intervallumokat is. A 9. táblázatból látható, hogy a tolerancia-intervallumok jóval szélesebbek, mint a jelenlegi gyakorlatban alkalmazott konfidencia-intervallum. Ez egyáltalán nem meglepő eredmény, hiszen két különböző típusú intervallumról van szó, melyek más-más valószínűségi változó elhelyezkedését vizsgálják, így teljesen más kérdésekre adnak választ.

Azt viszont feltétlenül meg kell jegyezni, hogy ha az eltarthatósági időt a tolerancia-intervallum alapján határozzuk meg, akkor ugyanarra az adatsorra lényegesen rövidebb becsült eltarthatósági időt kapunk a jóval szélesebb intervallumok miatt. Határozzuk meg például az 5-2-2-2 kódú esetre 3 éves tárolásnál az „A₂^{*}” modellel kapott intervallumok szélességét, a hatóanyag-tartalomra szokásos 90%-os specifikációs határt figyelembe véve!

- A jelenlegi (konfidencia-intervallumon alapuló) szabályozás szerint:

$$\hat{Y}_{36} - \text{konf} = 100 - 0,2 * 36 - 0,82 = 91,98 > 90,$$

azaz a három éves becsült eltarthatósági idő alátámasztható.

- A tolerancia-intervallumos szemlélet szerint:

$$\hat{Y}_{36} - \text{tol} = 100 - 0,2 * 36 - 5,21 = 87,59 < 90,$$

azaz három évnél a tolerancia-intervallum alsó határa már bőven a specifikációs határ alatt van.

Egy adott gyógyszerkészítmény stabilitásvizsgálatának értékelésekor a tolerancia-intervallumos modellel kapható rövidebb becsült eltarthatósági idő a gyógyszergyáraknak nyilvánvalóan előnytelen. Holott az így kapott rövidebb becsült eltarthatósági idő nem jelenti azt, hogy a beteg rosszabb gyógyszert kap. Ezért elsődleges feladat lenne annak biztosítása, hogy ugyanarra a készítményre ugyanolyan hosszú becsült eltarthatósági időt kapjunk a tolerancia- ill. a konfidencia-intervallumos módszerrel. Ennek pedig az a feltétele, hogy a tolerancia-intervallumos modellnél a specifikációs határokat is a tolerancia-intervallumos szemlélethez igazítsák. A fentieket összefoglalva kijelenthető, hogy a tolerancia-intervallumos módszerre való áttérés csak a specifikációs határok egyidejű módosításával lehetséges.

V.6.4.3. *Különböző faktorok hatása a tolerancia-intervallum szélességére*

V.6.4.3.1. *A különböző varianciakomponens-kombinációk vizsgálata*

Az új modellekben szereplő két ingadozási forrás (a tabletták különbözősége valamint az analízis hibája) egymáshoz viszonyított mértékének hatását a következő hat varianciakomponens-kombinációval vizsgáltam: $\sigma_{inhom}^2 = 1, 2, 4$ és $\sigma_{ism}^2 = 1, 2$.

A szimuláció során az alábbi paramétereket rögzítettem: $\beta_0 = 100$, $\beta_1 = -0,1$, $x_j = 0, 3, 6, 9, 12, 18, 24$ hónap, $m = 3$ és $p = 3$.

A három vizsgált stabilitásvizsgálati modellel kapott lefedési valószínűség értékeket a 10. táblázat tartalmazza. Ahogy az várható volt, az „A₁” modellel kapott lefedési valószínűség értékek jóval 0,95 alatt vannak. Az „A₂” modellel kapott lefedési valószínűség értékek az előírt 0,95 érték 95%-os konfidencia-tartományába esnek (0,936-0,964).

Az „A₂^{*}” modell esetén azonban néhány kiugróan magas lefedési valószínűség értéket is megfigyelhetünk. A 10. táblázat adataiból megállapítható, hogy akkor kapunk extrém magas (0,99 körüli) lefedési valószínűség értéket, ha a szimuláció során feltételezett összvariancia kicsi és a két ingadozási forrás feltételezett varianciakomponensei közti kapcsolatra az jellemző, hogy $\sigma_{inhom}^2 \leq \sigma_{ism}^2$.

A szimulációs program részletesebb elemzésével megállapítottam, hogy ezeknél a szimulációs eseteknél az 1000 futtatás során többször is extrém széles tolerancia-intervallumot kapunk. (Egy extrém széles tolerancia-intervallum alsó határa pedig biztosan alatta lesz elméleti tolerancia-intervallum határértékének, tehát „jónak” minősülve növeli a lefedési valószínűség értékét.) Extrém széles tolerancia-intervallumot akkor kapunk, ha a nem-centrális *t*-eloszlás szabadsági foka (Satterthwaite-közelítéssel számolva) igen kicsi (<1,5). Ez pedig akkor fordulhat elő, ha az egy adott futtatás során a számított inhomogenitási szórásnégyzet (s_{inhom}^2) értéke kisebb 0,5-nél. A gyógyszeripari gyakorlatban a tabletták különbözőségét jellemző inhomogenitási variancia általában lényegesen nagyobb az analitikai mérés bizonytalanságánál, így a valóságban igen kicsi az esélye annak, hogy extrém széles tolerancia-intervallumokat kapjunk.

Ha a problémát mégis kezelni akarnánk, több (rokon gondolatmeneten alapuló) közelítő módszer áll rendelkezésre^{51(p.255),60,61}.

az eset kódja				lefedési valószínűség az egyes modelleknél			tolerancia-intervallum átlagos szélessége az egyes modelleknél		
m	p	σ_{inhom}^2	σ_{ism}^2	„A ₁ ”	„A ₂ ”	„A ₂ [*] ”	„A ₁ ”	„A ₂ ”	„A ₂ [*] ”
3	3	1	1	0,89	0,94	0,99	4,15	4,49	4,14
3	3	1	2	0,92	0,95	1,000	5,05	5,31	5,16
3	3	2	1	0,87	0,94	0,97	5,07	5,67	5,31
3	3	2	2	0,87	0,94	0,99	5,82	6,29	5,76
3	3	4	1	0,83	0,94	0,96	6,51	7,48	7,21
3	3	4	2	0,86	0,94	0,97	7,11	7,93	7,42

10. táblázat: A lefedési valószínűség és a tolerancia-intervallum átlagos szélessége különböző varianciakomponens-kombinációk esetén

$$(\beta_0 = 100, \beta_1 = -0,1, t = 24 \text{ hónap}, m = 3, p = 3)$$

A 10. táblázat utolsó három oszlopa a számított tolerancia-intervallumok átlagos szélességét mutatja a három vizsgált modell esetén. Az adatok alapján megállapítható, hogy:

- Minden varianciakomponens-kombinációnál az „A₂^{*}” modell esetén szűkebb a tolerancia-intervallum az „A₂” modellhez képest.
- Az összvariancia növekedésével nő a tolerancia-intervallumok szélessége.
- Két megkülönböztethető ingadozási forrás feltételezése és adott összvariancia esetén a tolerancia-intervallum abban az esetben lesz szűkebb, ha az inhomogenitási szórásnégyzet kisebb az analitikai hibához képest. Ez azonban a gyakorlatban nem reális eset.

V.6.4.3.2. A különböző ismétlésszám-kombinációk vizsgálata

Az egy adott időpontban vett minták (m) valamint a mintákból végzett ismételt analízisek számának (p) a tolerancia-intervallum szélességére gyakorolt hatását az alábbi ismétlésszám-kombinációknál vizsgáltam:

- $r = 9 \rightarrow m = 3$ és $p = 3$
- $r = 10 \rightarrow m = 2$ és $p = 5$
- $r = 10 \rightarrow m = 5$ és $p = 2$
- $r = 10 \rightarrow m = 10$ és $p = 1$

A szimuláció során a következő paramétereket rögzítettem: $\beta_0 = 100$, $\beta_1 = -0,2$, $x_j = 0, 3, 6, 9, 12, 18, 24$ hónap. A feltételezett varianciakomponensek: $\sigma_{inhom}^2 = 4$ és $\sigma_{ism}^2 = 1$ voltak.

Az eredményeket a 11. táblázatban foglaltam össze.

Ahogy az várható volt, az „A₂^{*}” modellel számolt tolerancia-intervallum minden minta-és ismétlésszám kombináció esetén szűkebb az „A₂” modellel számoltnál.

A mintaszám növekedésével a tolerancia-intervallum mindkét modellnél szűkül. Ez azt jelenti, hogy a stabilitásvizsgálat tervezésekor célszerűbb lenne az egy adott időpontban rendelkezésre álló ismétlésszámot ($r = m * p$) úgy elosztani, hogy a lehető legtöbb mintát vegyük, s az azokból végzett ismételt analízisek számát csökkentjük. Ekkor ugyanis szűkebb tolerancia-intervallumot kapunk, mintha a gyógyszergyári gyakorlatban megszokott módon a minták és az ismételt analízisek számát egyenlően osztjuk el.

az eset kódja				lefedési valószínűség az egyes modelleknél			tolerancia-intervallum átlagos szélessége az egyes modelleknél		
m	p	σ_{inhom}^2	σ_{ism}^2	„A ₁ ”	„A ₂ ”	„A ₂ [*] ”	„A ₁ ”	„A ₂ ”	„A ₂ [*] ”
3	3	4	1	0,83	0,94	0,96	6,51	7,48	7,21
2	5	4	1	0,76	0,92	0,94	6,35	7,99	7,87
5	2	4	1	0,89	0,94	0,95	6,44	6,92	6,55
10	1	5	-	0,96	-	-	6,46	-	-

11. táblázat: A lefedési valószínűség és a tolerancia-intervallum átlagos szélessége különböző ismétlésszám-kombinációk esetén

$$(\beta_0 = 100, \beta_1 = -0,2, t = 24 \text{ hónap}, \sigma_{inhom}^2 = 4 \text{ és } \sigma_{ism}^2 = 1)$$

V.7. Példa

Az ismertett tolerancia-intervallumos modelleket valamint a jelenleg használt konfidencia-intervallumos módszert az alábbi modell példákon keresztül hasonlítottam össze. Mivel a valódi gyógyszeripari adatok titkosak, így három mesterségesen előállított, de a saját tapasztalataim alapján valóság-hű adatsort elemeztem. Mindhárom esetben a tablettákat szobahőmérsékleten tárolták. A stabilitásvizsgálati terv hét mintavételi időpontot ($q = 7$), időpontonként három mintát ($m = 3$) és mintánként három ismételt analízist ($p = 3$) tartalmazott.

Az első példa mért hatóanyag-tartalom értékeket a 12. táblázat mutatja.

Minta	Tárolási hónap						
	0	3	6	9	12	18	24
1	97,65	98,34	100,89	98,95	101,31	95,70	97,45
	98,55	98,12	101,00	98,65	99,84	97,55	95,60
	99,31	97,70	100,87	98,60	102,05	94,93	94,81
2	98,34	99,60	101,12	99,23	96,03	97,98	97,78
	100,49	100,30	98,62	97,48	96,35	97,70	96,89
	98,30	100,10	98,99	98,84	95,60	96,29	98,39
3	100,41	99,35	97,53	98,80	96,41	98,75	98,66
	101,01	99,10	100,91	101,43	96,96	98,23	97,79
	100,44	99,63	99,77	100,14	97,15	98,99	97,35

12. táblázat: Stabilitási adatok: hatóanyag-tartalom értékek (%)

A számítást az V.5. fejezetben írtak szerint végeztem. A hatóanyag-tartalomra vonatkozó alsó specifikációs határt $\kappa=90\%$ -nak vettem. Mind a konfidencia-, mind a tolerancia-intervallumon alapuló módszernél $\alpha = 0,05$ valószínűséggel számoltam. A tolerancia-intervallumokat kétféle γ értékre (0,01 ill. 0,05) is meghatároztam.

A stabilitási adatsor becsült degradációs egyenesére a következő egyenletet kaptam: $\hat{y} = 99,75 - 0,113x$. A becsült varianciakomponensek varianciaanalízissel $s_{ism}^2 = 0,82$ -nak és $s_{inhom}^2 = 1,43$ -nak adódtak. A különböző módszerekkel becsült eltarthatósági időket a 13. táblázat tartalmazza.

γ	Konfidencia-intervallum módszer	Tolerancia-intervallum modellek		
		„A ₁ ”	„A ₂ ”	„A ₂ [*] ”
0,01	66	42	37	39
0,05		49	44	45

13. táblázat: A 12. táblázatban ismertetett stabilitási adatok becsült eltarthatósági ideje (hónapban) a különböző modellekre ($\alpha = 0,05$)

A 13. táblázatból látható, hogy $\gamma = 0,01$ esetén a becsült eltarthatósági idő az „A₁” modellel 42 hónap, az „A₂” modellel 37 hónap, az „A₂^{*}” modellel pedig 39 hónap. Ha a számított tolerancia-intervallum a tabletták 99%-a helyett csak 95%-át tartalmazza, akkor a becsült eltarthatósági idő 6-7 hónappal nő.

Azt is megfigyelhetjük, hogy az ICH Guideline által előírt konfidencia-intervallum alapján a becsült eltarthatósági idő jóval hosszabbnak (66 hónapnak) adódik. A különbség oka a két intervallum eltérő szemléletében keresendő: a konfidencia-intervallum a tabletták várható értékének, míg a tolerancia-intervallum a tabletták nagy hányadának (pl. 99%-ának) elhelyezkedését mutatja. Mivel azonban természetesen mindkét intervallum esetén ugyanarról az adatsorról (ugyanarról a gyógyszerkészítményről) van szó, nem lenne jogos, ha a tolerancia-sávós szemléletet alkalmazó gyógyszergyárakat rövidebb becsült eltarthatósági idővel büntetnénk, amikor az általuk gyártott gyógyszer minősége ugyanolyan jó. Ezért lenne szükség a tolerancia-intervallumos szemléletre való átállás esetén a specifikációs határok egyidejű megváltoztatására (csökkentésére), mégpedig úgy, hogy a tolerancia-intervallumos szemlélet szerint megadott lejáratú időpontban is ugyanakkora valószínűséggel legyünk biztosak abban, hogy az adott gyógyszer tartalmazza az előírt hatóanyag-tartalmat, mint a konfidencia-intervallumos szemléletnél. Ezt ki is számítottam ennél a példánál a $\gamma = 0,01$ esetre. Eredményül azt kaptam, hogy ahhoz, hogy a konfidencia- ill. a tolerancia-sávós elképzeléssel ugyanolyan hosszú becsült eltarthatósági időt kapjunk, a tolerancia-sávós modelleknél a specifikációs határt az alábbiak szerint kellene módosítani: az „ A_1 ” modellnél 86,4%-ra, az „ A_2 ” modellnél 85,1%-ra, az „ A_2^* ” modellnél pedig 85,4%-ra (a konfidencia-sáv esetén megszokott 90% helyett).

A tolerancia-intervallum szemléletéből adódóan mindig szélesebb lesz a konfidencia-intervallumnál, s így mindig rövidebb becsült eltarthatósági időt eredményez. A tendencia tehát általános, de a két szemléletmód alapján becsült eltarthatósági idők közti különbség értéke minden egyes konkrét stabilitási adatsor esetén más és más lesz. A különbség nagysága függ a stabilitási adatok ingadozását okozó varianciakomponensek egymáshoz viszonyított arányától és az illesztett degradációs egyenes paramétereitől. Ennek szemléltetésére mutatom be a további két modell példát. A példák konkrét hatóanyag-tartalom adatait helyhiány miatt nem ismertetem, csak az adatsorok főbb becsült paramétereit és az eltarthatósági időket. Az összehasonlításhoz olyan adatsorokat választottam, ahol a tolerancia-intervallumos modellekben használt két varianciakomponens aránya példánként jelentősen eltérő. (A már ismertetett 1. példában az inhomogenitási szórásnégyzet valamivel nagyobb az ismételtetési szórásnégyzetnél, a 2. példában a két szórásnégyzet értéke szinte megegyezik egymással, a 3. példánál pedig az inhomogenitási szórásnégyzet jelentősen (majdnem 3-szor) nagyobb az ismételtetési szórásnégyzetnél.

Példa	A becült degradációs egyenes	A becült variancia-komponensek	Konfidencia-intervallum módszer	Tolerancia-intervallum modellek		
				„A ₁ ”	„A ₂ ”	„A ₂ [*] ”
1.	$\hat{y} = 99,75 - 0,113x$	$s_{inhom}^2 = 1,43$ $s_{ism}^2 = 0,82$	66	42	37	39
2.	$\hat{y} = 99,68 - 0,079x$	$s_{inhom}^2 = 1,10$ $s_{ism}^2 = 1,08$	85	54	48	53
3.	$\hat{y} = 99,04 - 0,055x$	$s_{inhom}^2 = 2,56$ $s_{ism}^2 = 0,89$	92	47	37	38

14. táblázat: A három példára a különböző modellekkel kapott becült eltarthatósági idők (hónapban) ($\alpha = 0,05$, $\gamma = 0,01$)

VI. A STABILITÁSI ADATOK BIZONYTALANSÁGÁNAK ELŐZETES BECSLÉSE

VI.1. A probléma felvetése, célkitűzés

A gyógyszergyárak alapvető érdeke, hogy gyógyszerkészítményeik becsült eltarthatósági ideje a lehető leghosszabb legyen. Az ICH Q1E Guideline⁴ előírása szerint a becsült lejárat idő a stabilitási adatokra illesztett degradálódási függvény konfidencia-sávjának és a vizsgált tulajdonságra vonatkozó elfogadási kritériumnak a metszéspontja. (ld. II.3.1.2. fejezet). A hatóság által kötelezően előírt specifikációs határokat adottnak tekintve, a becsült eltarthatósági idő a gyógyszer degradációjának jellegétől és mértékétől, a gyártás inhomogenitásától és az analitikai mérés bizonytalanságától, valamint az ismereteink pontatlanságát kifejező bizonytalansági sáv szélességétől függ.

Adott meredekségű degradációs egyenest feltételezve a becsült eltarthatósági idő annál hosszabb, minél szűkebb a stabilitási adatokra illesztett egyenes konfidencia-sávja. A konfidencia-sáv szélessége a stabilitási adatok ingadozásától (beleértve az inhomogenitást és az analitikai bizonytalanságot is), valamint a mérési adatok számától és elhelyezkedésétől függ (ld. (1) egyenlet). A mintavételi időpontok száma és sűrűsége az ICH Q1A(R2) Guideline³ előírásai, valamint gazdasági megfontolások miatt csak szűk határok között változtatható, így a konfidencia-sáv szűkítésére, s ezáltal a becsült eltarthatósági idő növelésére leginkább a stabilitási adatok ingadozásának csökkentése kínál lehetőséget. Ez a megállapítás ugyanakkor azt is jelenti, hogy a stabilitási adatok bizonytalansága - adott degradációs ráta esetén - behatárolja a becsült eltarthatósági idő hosszát.

A stabilitásvizsgálat során mért adatok kisebb-nagyobb véletlen ingadozását mutatnak. Mint azt a II.5. fejezetben már ismertettem, a stabilitási adatok szóródása négy, a gyógyszerkészítmény gyártásához, valamint a minták analitikai elemzéséhez kapcsolódó ingadozási forrásnak tulajdonítható.

Az irodalmi részben már említett cikkek^{5,6,46,47} a stabilitásvizsgálatok során jelentkező ingadozási források mértékére magából a stabilitásvizsgálatból következtettek. Ha az ingadozási források varianciakomponenseinek nagyságrendjét előzetesen ismernénk, akkor már a stabilitásvizsgálat kezdete előtt képet kaphatunk a majdani stabilitási adatok ingadozásának mértékéről, és így eldönthetnénk, hogy a jelenlegi gyártási folyamat és az analitikai mérési módszer bizonytalansága megfelelő-e (azaz nem túl nagy) a kívánt lejárat

idő eléréséhez. Ennek igen nagy jelentősége lenne, hiszen így még idejében látni lehetne, ha a gyógyszerkészítmény várhatóan megbukik a stabilitásvizsgálaton ill. ennek fényében lehetne fejleszteni a gyártási technológiát ill. az analitikai mérési módszert.

A fenti célt szem előtt tartva dolgoztam ki az ebben fejezetben bemutatott számítási módszert. A módszer igen előnyös tulajdonsága, hogy nem igényel újabb idő- és költségigényes analitikai vizsgálatokat, mivel a gyógyszerkészítmény fejlesztése során egyébként is elvégzett analitikai vizsgálatok (validálás és egyedi hatóanyag-tartalom meghatározás) eredményeiből teszi lehetővé a stabilitási adatok ingadozását okozó elemi varianciakomponensek becslését.

VI.2. A stabilitási adatok varianciakomponenseinek előzetes becslése

A stabilitási adatok ingadozását okozó varianciakomponensek egymástól való elkülönítéséhez és számszerű becsléséhez az alábbi vizsgálatok eredményeit használtam fel:

- egyedi hatóanyag-tartalom meghatározás (Content Uniformity vizsgálat)
- az analitikai módszer validálásából
 - az ismételhetőség azonos körülmények között (repeatability)
 - az ismételhetőség változó körülmények között
 - laboratóriumon belül (intermediate precision)
 - laboratóriumok között (reproducibility)

VI.2.1. Egyedi hatóanyag-tartalom meghatározás (Content Uniformity vizsgálat)

Az egyedi hatóanyag-tartalom meghatározása⁶² során 3 sarzsból, sarzsonként 10-10 tablettá egyenkénti analizisét végzik el. Az adatok statisztikai feldolgozásával (egy faktoros varianciaanalízissel) két varianciakomponens különíthető el: σ_{sarzs}^2 és σ_{CU}^2 .

A σ_{sarzs}^2 variancia a sarzs-faktor hatását, azaz a sarzsok különbözőségét mutatja. A σ_{CU}^2 maradék variancia a tabletták különbözősége mellett az analitikai mérés ismételhetőségi hibakomponensét is tartalmazza. A két varianciakomponens hatása ugyanis numerikusan nem választható el egymástól, mivel az egyes tabletták analizise roncsolásos úton történik, így a mérés nem ismételhető meg ugyanazon tablettával.

VI.2.2. Az analitikai módszer validálása

Az analitikai módszer validálása során számos teljesítményjellemzőt vizsgálnak. Ezek közül számunkra most a véletlen hibákkal összefüggő analitikai teljesítményjellemző, a precizitás (precision) érdekes. A precizitás meghatározása több lépésben történik⁶³.

VI.2.2.1. Az azonos körülmények közötti ismételhetőség vizsgálata (repeatability)

Az azonos körülmények közötti ismételhetőség vizsgálatokor egy adott személy, egy adott laboratóriumban és műszeren, egy nap alatt, egy adott sarzsból több (általában 4-6) ismételt mérést végez. Az egyes mérések több tabletta elporításából készült homogén mintákból történnek.

A kapott adatok szóródását jellemző ismételhetőségi variancia (σ_{repeat}^2) szintén együttesen tartalmazza az inhomogenitást (azaz az egyes minták, áttételesen pedig a tabletták különbözőségének) és az analízis hibájának hatását, mivel ennél a vizsgálatnál az adott tablettákból álló mintákat csak egyszer analizálják, s így lehetetlen a két ingadozási forrást egymástól numerikusan elkülöníteni.

Szerencsére az egyedi hatóanyag-tartalom meghatározás és a validálás ismételhetőségi vizsgálatának kivitelezése közti kicsiny eltérés lehetőséget biztosít az inhomogenitás okozta bizonytalanság és az analitikai mérés ismételhetőségi hibakomponensének szétválasztására.

Az egyedi hatóanyag-tartalom meghatározása ugyanis - a vizsgálat nevének megfelelően - a tabletták egyenkénti analízisével történik, így a σ_{CU}^2 variancia az inhomogenitás hatásának és az analízis ismételhetőségi hibájának egyszerű összegeként adódik:

$$\sigma_{\text{CU}}^2 = \sigma_{\text{inhom}}^2 + \sigma_{\text{ism}}^2 \quad (45)$$

Ezzel ellentétben az azonos körülmények közötti ismételhetőség vizsgálatokor egy-egy mérésnél több tablettából álló mintákat analizálnak, így ezekben az esetekben az inhomogenitás hatása csak csökkentve jelentkezik:

$$\sigma_{\text{repeat}}^2 = \sigma_{\text{inhom}}^2 / d + \sigma_{\text{ism}}^2 \quad (46)$$

ahol d a tabletták száma az egyes mintákban.

Megjegyzem, hogy a (46) egyenlet abban az esetben érvényes, ha minden mérés új mintából történik. Ha a mérés ismétléséhez ugyanazt a mintát használják, akkor természetesen

$$\sigma_{\text{repeat}}^2 = \sigma_{\text{ism}}^2 \cdot$$

Összevetve a (45) és (46) egyenleteket az inhomogenitás és az analitikai mérés ismételhetőségi hibájának elemi varianciakomponense már elkülönítve kapható meg:

$$\sigma_{\text{inhom}}^2 = (\sigma_{\text{CU}}^2 - \sigma_{\text{repeat}}^2) / (1 - 1/d), \quad (47)$$

$$\sigma_{\text{ism}}^2 = \sigma_{\text{CU}}^2 - \sigma_{\text{inhom}}^2. \quad (48)$$

VI.2.2.2. A változó körülmények közötti ismételhetőség vizsgálata

Az analitikai módszer változó körülmények közötti ismételhetőségét két fokozatban szokás vizsgálni.

VI.2.2.2.1. A laboratóriumon belüli szórás vizsgálata (intermediate precision)

A laboratóriumon belüli szórás vizsgálatok a méréseket ugyan egy adott laboratóriumon belül, de vagy különböző személyek, vagy különböző műszereken, vagy különböző napokon végzik el. A vizsgálat eredményeit egy faktoros varianciaanalízissel értékelhetjük. A kapott varianciák: az ismétlés különböző körülményének hatását jelző $\sigma_{\text{körülmények}}^2$ variancia és a maradék-variancia, mely tartalmilag megegyezik az (az azonos körülmények közti) ismételhetőségre vonatkozó, az előző alpontban részletesen ismertetett σ_{repeat}^2 varianciával. Az analízis változó körülményeire vonatkozó variancia ($\sigma_{\text{körülmények}}^2$) - attól függően, hogy a konkrét helyzetben a felsorolt vizsgálati lehetőségek közül melyiket vagy melyek kombinációit választottuk - a személyek ($\sigma_{\text{személy}}^2$), a műszerek (σ_{muszer}^2), ill. a napok (σ_{nap}^2) különbözősége okozta mérési bizonytalanságot mutatja.

VI.2.2.2.2. A laboratóriumok közötti szórás vizsgálata (reproducibility)

A reprodukálhatóság vizsgálatok már csak a vizsgálati módszer azonos, de az ismételt méréseket különböző laboratóriumokban, természetesen különböző műszereken, különböző személyek végzik el. Az adatok feldolgozása itt is varianciaanalízissel történik, a kapott varianciák: a laboratóriumok közti különbséget mutató σ_{labor}^2 variancia és az (az azonos körülmények közti) ismételhetőségre vonatkozó σ_{repeat}^2 variancia.

Végignézve az analitikai módszer pontosságára irányuló mérésorozatot megállapítható, hogy az azonos körülmények közti ismételhetőség mértékére (σ_{repeat}^2) mind a három vizsgálatnál kapunk egy-egy becslést. Ez felveti annak lehetőségét is, hogy az azonos körülmények közti ismételhetőséget külön nem vizsgálva (azaz a *repeatability* vizsgálat elhagyásával) információveszteség nélkül csökkentsük az elvégzendő mérések számát.

VI.2.3. A varianciakomponens-becslés menetének összegzése

A stabilitási adatok ingadozását okozó varianciakomponensek becslésének menetét a 15. táblázatban összegeztem. Az általam kidolgozott számítási módszer lényege a második és harmadik oszlop közti különbségen alapul, hisz az egyes analitikai vizsgálatokból közvetlenül meghatározható varianciakomponensek nem minden esetben egyeznek meg a stabilitási adatok ingadozását okozó elemi varianciakomponensekkel.

Analitikai vizsgálat	A mérésből közvetlenül meghatározható varianciakomponensek	Elemi varianciakomponensek
Egyedi hatóanyag-tartalom meghatározása	σ_{sarzs}^2	σ_{sarzs}^2
	σ_{CU}^2	σ_{inhom}^2 és σ_{ism}^2
Validálás: A módszer pontosság vizsgálata	σ_{labor}^2	σ_{labor}^2
	σ_{nap}^2 , $\sigma_{\text{személy}}^2$	σ_{nap}^2 , $\sigma_{\text{személy}}^2$
	σ_{repeat}^2	σ_{inhom}^2 és σ_{ism}^2

15. táblázat: A varianciakomponensek becslésének menete

Megjegyzem, hogy a számítás során a táblázatban látható varianciáknak természetesen csak a mérési adatokból számítható becsléseit ($\hat{\sigma}$) kapjuk meg. Fontos még azt is megemlíteni, hogy a számított varianciák az elméleti érték pontbecslései, melyek bizonytalansága a szintek kis száma (pl. kevés nap) miatt igen nagy is lehet. A bizonytalanság hatását az eltarthatósági idő becslésénél próbálom meg figyelembe venni.

VI.3. A stabilitási adatok bizonytalanságának hatása a becsült eltarthatósági időre

A stabilitási adatok ingadozását okozó varianciakomponensek azonosítása és számszerűsítése után azt vizsgálom, hogy a stabilitási adatok bizonytalansága milyen hatással van az eltarthatósági idő becsült időtartamára. Ehhez először egy jövőbeli stabilitási pont várható varianciáját ($\hat{\sigma}$) kell kiszámítani.

VI.3.1. Egy jövőbeli stabilitási pont varianciájának becslése

A gyógyszerkészítmények stabilitásvizsgálata során az előírt vizsgálati időpontokban sarzsonként több mintát vesznek, mintánként több tablettát analizálva. Egy jövőbeli stabilitási adat várható mért értéke tehát függ attól, hogy a mintát melyik sarzsból vették, a vett minta milyen hatóanyag-tartalmú tablettákat tartalmaz, valamint attól, hogy az analitikai mérés ismételhősége ill. körülményei (mikor mértek, melyik laboratóriumban) milyenek voltak. Ezeket az ingadozási forrásokat a 15. táblázat harmadik oszlopában látható elemi varianciakomponensekkel jellemezhetjük. Az elemi varianciakomponensek megfelelő összegzésével pedig egy jövőbeli stabilitási adat várható varianciáját kaphatjuk meg.

Ha a stabilitásvizsgálat értékelésekor minden egyes adatot külön kezelünk, akkor egy jövőbeli stabilitási pont becsült varianciája az alábbi képlettel számítható:

$$\hat{\sigma}^2 = \frac{\hat{\sigma}_{\text{inhom}}^2}{d} + \hat{\sigma}_{\text{ism}}^2 + \hat{\sigma}_{\text{sarzs}}^2 + \hat{\sigma}_{\text{nap}}^2 + \hat{\sigma}_{\text{labor}}^2 \quad (49)$$

Az adatok értékelésekor azonban általában az egy adott időpontban vett minták átlagértékét veszik figyelembe. Ekkor egy jövőbeli stabilitási pont bizonytalanságának becslése az alábbi képlet alapján történhet:

$$\hat{\sigma}^2 = \frac{\hat{\sigma}_{\text{ism}}^2 + \frac{\hat{\sigma}_{\text{inhom}}^2}{d}}{m} + \hat{\sigma}_{\text{sarzs}}^2 + \hat{\sigma}_{\text{nap}}^2 + \hat{\sigma}_{\text{labor}}^2 \quad (50)$$

ahol m a minták száma.

Ha a stabilitásvizsgálat értékelése során a sarzsokat külön kezelve csak a legrosszabb sarzs adatait használjuk fel, akkor a fenti képletekben a σ_{sarzs}^2 varianciakomponenst természetesen nem vesszük figyelembe.

VI.3.2. A becsült eltarthatósági idő bizonytalanságának számszerűsítése

A stabilitási adatok bizonytalanságának a becsült eltarthatósági időre való hatását a degradálódási függvény (egyenes) 95%-os alsó konfidencia-sávjának várható szélességével jellemzem. A konfidencia-sáv szélessége (ld. (1) egyenlet) többek között a stabilitási adatok regressziós feldolgozásakor kapott s_y reziduális szórás értékétől függ.

Jelen esetben azonban nem áll rendelkezésemre egy konkrét tárolási adatsor, hiszen céлом éppen az, hogy még a stabilitásvizsgálat kezdete előtt (tehát az adatok hiányában) becsüljem a stabilitási adatok várható bizonytalanságát és ennek hatását az eltarthatósági idő hosszára. Ennek érdekében a konfidencia-sáv szélességének számításánál a reziduális szórást olyan szórás értékkel kell helyettesítenem, mely a stabilitási adatok várható ingadozását megfelelően jellemzi. Ezt kétféleképpen tettem meg.

1. Az első megközelítésnél a reziduális szórást egyszerűen a majdani stabilitási terv egy pontjának becsült varianciájával (ld. (49) vagy (50) egyenlet) helyettesítem. A konfidencia-sáv képletében szereplő egyéb paraméterek (pl. szabadsági fok, mérési időpontok száma) értékét pedig a majdani stabilitási tervnek megfelelően tételeztem fel.
2. A második megközelítésnél az egyedi hatóanyag-tartalom meghatározásából és az analitikai módszer validálásából becsült elemi varianciakomponensekkel (ld. 15. táblázat) stabilitási adatsorokat szimuláltam. Az így kapott hatóanyag-tartalom adatokra egyenest illesztve a konfidencia-sávhoz szükséges reziduális szórás számítható.

Jogosan merülhet fel a kérdés, hogy a stabilitási adatok bizonytalanságának a becsült eltarthatósági időre való hatása hogyan alakul a tolerancia-intervallumos szemléletben. Az ismertett számítási módszert természetesen a tolerancia-intervallumos modellekre is lehetne alkalmazni, hisz a validálásból és a Content Uniformity vizsgálatból kapott varianciakomponenseket éppúgy be lehet helyettesíteni a tolerancia-sáv képletébe, mint a konfidencia-sávba. Részletesen erre most nem térek ki, mert céлом ezzel az újszerű problémafelvetéssel annak bemutatása volt, hogy az ICH Guideline jelenlegi gyakorlatát (azaz a konfidencia-sávós szemléletet) követve hogyan tudnák a gyógyszergyárak előzetes megjósolni a stabilitási adatok várható ingadozását, s így már a stabilitásvizsgálat indítása előtt látni, hogy az adott készítmény várhatóan átmegy vagy megbukik a stabilitásvizsgálaton.

VI.4. Példa

Az alábbiakban egy példán mutatom be a stabilitási adatok ingadozását okozó varianciakomponensek előzetes becslésének menetét valamint az általuk okozott bizonytalanságnak a becsült eltarthatósági időre gyakorolt hatását.

Az első lépés tehát az ingadozási források elemi varianciakomponenseinek meghatározása, melyhez a VI.2. fejezetben említett módon a validálás és a Content Uniformity vizsgálat eredményeit használom fel.

Az elemzendő példa konkrét adatait titoktartási kötelezettség miatt nem adhatom meg. Az adatok egy gyógyszerkészítmény HPLC-s hatóanyag-tartalom meghatározásából származnak. Az adott gyógyszert két hatáserősségben gyártják (I. és II.) és két hatóanyagot (A és B) tartalmaz. A hatóanyag pontos meghatározását nehezítette annak kis koncentrációja és a mátrix-komponensek zavaró hatása.

Az analitikai mérési módszer validálása során a precizitás vizsgálata az alábbi lépésekből állt. Az azonos körülmények közti ismételhetőség (*repeatability*) vizsgálatára egy analitikus, egy adott napon, azonos sarzsból származó tablettákból 6 ismételt mérést végzett, mintánként 5-5 tablettát homogenizálva. A laboratóriumon belüli szórás (*intermediate precision*) vizsgálatakor 3 napon át, naponta 3-3 mérést végeztek. A laboratóriumok közötti szórás (*reproducibility*) vizsgálatát pedig 2 laboratóriumban, laboratóriumonként 6-6 mérésből végezték.

Az adott analitikai módszer validálása során a laboratóriumon belüli szórás vizsgálatát az egyedi hatóanyag-tartalomra is elvégezték: 3 napon át, naponta 10-10 tablettát egyenként is analizáltak. Ez a fajta vizsgálat csupán annyiban tér el a VI.2.1. fejezetben említett Content Uniformity vizsgálatától, hogy a tabletták egyenkénti analízisét nem 3 sarzsból, hanem 3 különböző napon végezték el, így ennek a vizsgálatnak az eredményei ugyanúgy felhasználhatók az inhomogenitási variancia meghatározására. Természetesen a vizsgálat eredményeit a laboratóriumon belüli szórás (jelen esetben a napok közti különbség hatásának) számítására is felhasználhatjuk.

VI.4.1. Az analitikai módszer ismételhetőségének becslése

Az analitikai módszer azonos körülmények közötti ismételhetőségének varianciájára (σ_{repeat}^2) a validálás precizitás-vizsgálatának (ld. VI.2.2. fejezet) mindhárom lépésében kapunk

becslést. Ugyanis mindhárom vizsgálatnál - azonos vagy változó körülmények között, de - ismételt mérések vannak és ez lehetőséget biztosít az ismételhetőségi variancia közvetlen becslésére. Mivel három becslésünk is van az ismételhetőségi varianciára, így a megbízhatóbb (nagyobb szabadsági fokú) becslés érdekében célszerű ezeket egyesíteni. Az egyesítést a Fisher-Cochran tétel alapján végeztem el.

Az összefoglalt eredményeket a 16. táblázat mutatja. A táblázat utolsó sora tartalmazza a három vizsgálatból származó ismételhetőségi szórásnégyzetek egyesített értékét.

Az adatokból látható, hogy az ismételhetőségi és a laboratóriumok közötti szórás vizsgálatból az ismételhetőségre kapott becslt varianciák egymással jó egyezést mutatnak, míg a laboratóriumon belüli szórás vizsgálata esetén ezek az értékek valamivel alacsonyabbak.

Az adatokat elemezve azt is megállapíthatjuk, hogy az egyesített szórásnégyzet értékek nagyjából megegyeznek a négy vizsgált esetben (oszlopban), azaz a gyógyszerkészítmény mindkét hatáserősségének mindkét hatóanyagánál az ismételhetőségi varianciakomponens megközelítőleg azonos.

Mely vizsgálatból becsüljük?	I. hatáserősség		II. hatáserősség		Szabadsági fok
	A	B	A	B	
Ismételhetőségi vizsgálat	1,67	1,13	1,13	0,96	5
Laboratóriumon belüli szórás vizsgálata	0,71	0,41	0,76	0,64	6
Laboratóriumok közötti szórás vizsgálata	1,38	0,90	1,27	1,06	10
Egyesített szórásnégyzet	1,26	0,82	1,09	0,92	21

16. táblázat: Az ismételhetőségi varianciakomponens becslt értékei (hatóanyag-tartalom %)

VI.4.2. A laboratóriumon belüli szórás becslése

Mint már említettem, ennél a példánál az analitikai módszer laboratóriumon belüli szórásának vizsgálatakor az analízis lehetséges változó körülményei közül a naponkénti ismételhetőség vizsgálatát választották.

Mivel a laboratóriumon belüli szórás vizsgálatát mind a több tablettából álló mintákra (a validálás során), mind pedig az egyedi hatóanyag-tartalomra elvégezték, a napok

különbözösége okozta mérési bizonytalanságra minden hatásereőség-hatóanyag kombinációnál két becslést kapunk. A napok különbözöségének hatását mutató becslült varianciakomponenseket (σ_{nap}^2) a 17. táblázatban összegzem.

Mely vizsgálatból becsüljük?	I. hatásereőség		II. hatásereőség	
	A	B	A	B
Laboratóriumon belüli szórás vizsgálata	4,48	2,26	1,81	0,99
Az egyedi hatóanyag- tartalom naponkénti vizsgálata	3,76	1,61	0*	0*

17. táblázat: A naponkénti ismételtetés becslült varianciakomponensei (hatóanyag-tartalom %)

*a nap mint faktor hatása nem szignifikáns 0,05 szinten

A 17. táblázatból látható, hogy az analitikai mérés egyik fontos körülményének, az analízis időpontjának jelentős hatása van a mért hatóanyag-tartalomra. A napok különbözösége okozta többletingadozás az I. hatásereőség „A” hatóanyagánál különösen is magas. Még inkább szembetűnő a kapott eredmény, ha meggondoljuk, hogy jelen esetben az ismételt analízisek egymást követő napokon történtek, amikor a mérési körülmények valószínűleg sokkal inkább hasonlóak voltak, mint a majdani stabilitásvizsgálat méréseinek több hónapos időközéinél. Ez a példa is alátámasztja a különbözö analízis-időpontok okozta mérési bizonytalanság stabilitásvizsgálatoknál történő figyelembevételének fontosságát, melyet a dolgozatomban már többször is hangsúlyoztam.

A 17. táblázat sorait összehasonlítva megállapítható, hogy az egyedi hatóanyag-tartalom naponkénti vizsgálatok kapott varianciakomponensek értékei alacsonyabbak (ha egyáltalán a napok különbözöségének hatása kimutatható). Ennek oka a két vizsgálati terv különbözöségében keresendő. Az egyfaktoros varianciaanalízis esetén ugyanis a vizsgált faktor hatásának kimutatása lényegesen nehezebb, ha az azonos körülmények közti ismételtetés varianciája nagyobb. A tabletták egyenkénti vizsgálatánál az inhomogenitás hatása a vizsgálat jellegéből adódóan fokozottabban jelentkezik, ezáltal megnő a csoportokon belüli hibavariancia értéke, s könnyebben elfedi a vizsgált faktor, azaz a napok különbözöségének hatását.

VI.4.3. A laboratóriumok közötti szórás vizsgálata

A laboratóriumok közötti szórás vizsgálatával kapott adatok varianciaanalízissel történő feldolgozása alapján megállapítottam, hogy a laboratóriumnak mint faktornak 5 %-os szignifikancia szinten nincs hatása a mérési eredményekre (a vizsgált példa mind a négy esetben). Tehát $\sigma_{\text{labor}}^2 = 0$ tekinthető, azaz a laboratóriumok közötti különbség a már eddigi bizonytalansági forrásokhoz képest nem okoz többletingadozást.

VI.4.4. Az inhomogenitási varianciakomponens becslése

Az inhomogenitási varianciakomponens (azaz a tabletták különbözőségének) becsléséhez egyedi tabletták analizésére is szükség van. Ez ebben a példában egyedi tabletták naponkénti vizsgálatával valósult meg.

Az adatok varianciaanalízissel történő feldolgozásával (a nap mint faktor hatása mellett) közvetlenül csak az egyedi hatóanyag-tartalom ingadozásának becslése (σ_{CU}^2) kapható meg, mely természetesen elválaszthatatlanul tartalmazza a tabletták különbözőségének és az analízis hibájának hatását (ld. VI.2.1. fejezet).

Az inhomogenitási és az ismételhetőségi varianciakomponens numerikus elkülönítésére a (47) és (48) egyenletek adnak lehetőséget. A számításhoz felhasznált becült variancia-értékeket ($\hat{\sigma}_{\text{CU}}^2$ és $\hat{\sigma}_{\text{repeat, egyesített}}^2$) valamint az eredményül kapott elemi varianciakomponensek értékeit ($\hat{\sigma}_{\text{inhom}}^2$ és $\hat{\sigma}_{\text{ism}}^2$) a 18. táblázatban összesítettem.

Az adatokból látható, hogy a σ_{CU}^2 variancia becült értékei a vizsgált gyógyszer mindkét hatáserősségének mindkét hatóanyagánál igen magasak. Ennek köszönhető, hogy jelen gyógyszer és analitikai vizsgálati módszer esetében a tabletták különbözőségének hatása sokkal-sokkal jelentősebb, mint az analitikai módszer véletlen hibája.

	I. hatáserősség		II. hatáserősség	
	A	B	A	B
$\hat{\sigma}_{\text{CU}}^2$	5,44	4,99	4,73	4,23
$\hat{\sigma}_{\text{repeat}}^2$ (egyesített)	1,26	0,82	1,09	0,92
$\hat{\sigma}_{\text{inhom}}^2$	5,23	5,21	4,55	4,14
$\hat{\sigma}_{\text{ism}}^2$	0,21	-0,22	0,18	0,09

18. táblázat: Az inhomogenitási és az ismételhetőségi varianciakomponensek becült értékei (hatóanyag-tartalom %)

A VI.4. fejezet eddigi alfejezeteiben az adott gyógyszerkészítmény majdani stabilitási adatainak ingadozását okozó elemi varianciakomponensek ($\hat{\sigma}_{inhom}^2$, $\hat{\sigma}_{ism}^2$, $\hat{\sigma}_{nap}^2$, $\hat{\sigma}_{labor}^2$) előzetes becslésének menetét mutattam be. A számításokhoz a validálás precizitás-vizsgálatának és a Content Uniformity vizsgálatnak eredményeit használtam fel.

Mivel ebben példában az egyedi hatóanyag-tartalom meghatározása nem sarzsokként, hanem naponként történt, a sarzsok közti ingadozás varianciakomponensére ($\hat{\sigma}_{sarzs}^2$) nincs becslésünk, így annak értékét a további számításoknál 1,5-nek vettem.

VI.4.5. Egy jövőbeni stabilitási pont várható ingadozása

A példa folytatásaként azt vizsgálom, hogy ezek az elemi varianciakomponensek a majdani stabilitásvizsgálat során mért értékeknél mekkora ingadozást okozhatnak, s ez a bizonytalanság miként befolyásolja a becsült eltarthatósági időt.

Ehhez képzeljük el az alábbi - a gyakorlatban általános alkalmazott - stabilitásvizsgálati tervet: Az adott gyógyszer 3 sarzsának egy éves (0, 3, 6, 9, 12 hónapos) stabilitási adatai (hatóanyag-tartalom értékek) állnak majd rendelkezésünkre. Minden vizsgálati időpontban sarzsokként három-három mintát vesznek, az egyes minták pedig 5 tableta homogenizálásával készülnek.

A stabilitási adatok várható bizonytalansága egy jövőbeni stabilitási pont becsült varianciájával jellemezhető (ld. (49) vagy (50) képlet attól függően, hogy az egy adott időpontban mért stabilitási adatokat egyenként vagy csak az átlagukat vesszük figyelembe). A számításokat az alábbi két esetre végeztem el:

- A sarzsok adatai nem egyesíthetőek, így az eltarthatósági idő becslésénél csak a leggyorsabban bomló sarzsot vesszük figyelembe. Ekkor a sarzsok közti ingadozás varianciakomponensét ($\hat{\sigma}_{sarzs}^2$) a számításnál nem kel figyelembe venni.
- A sarzsok adatai egyesíthetőek, azaz az eltarthatósági időt három sarzs adatai alapján becsüljük.

A számítások eredményeit a 19. táblázat tartalmazza.

A sarzsok egyesíthetőek-e?	Az egy adott időpontban vett minták kezelése	I. hatásereőség		II. hatásereőség	
		A	B	A	B
nem	egyenként	5,74	3,26	2,90	1,91
	átlaguk	4,90	2,59	2,17	1,30
igen	egyenként	7,24	4,76	4,40	3,41
	átlaguk	6,40	4,09	3,67	2,80

19. táblázat: Egy jövőbeli stabilitási pont becsült varianciája (hatóanyag-tartalom %)

A 19. táblázatból látható, hogy

- ha az egy adott időpontban vett stabilitási mintákat egyenként kezeljük, akkor egy jövőbeli stabilitási pont becsült varianciája mindig nagyobb lesz, mintha az átlagértéket vizsgálnánk, hiszen ekkor a képletben nem szerepel osztóként a mintaszám
- ha a sarzsokat egyesítve kezeljük, akkor egy jövőbeli stabilitási pont becsült varianciája természetesen nagyobb lesz, mintha csak egy sarzs adatait elemeznénk, hiszen ekkor a sarzsok közti különbség varianciakomponensét is hozzá kell adnunk

VI.4.6. A stabilitási adatok ingadozásának hatása az eltarthatósági idő becsült hosszára

A stabilitási adatok ingadozásának az eltarthatósági idő becsült hosszára gyakorolt hatását a 2 éves tárolásnál számított egyoldali, 95%-os konfidencia-sáv szélességével jellemzem (ld. VI.3.2. fejezet).

A 20. táblázatban az első megközelítés szerinti eredményeket ismertetem, amikor is a konfidencia-sáv képletében (ld. (1) egyenlet) szereplő reziduális szórását a 19. táblázatban szereplő becsült varianciákkal helyettesítettem. A konfidencia-sáv többi paramétere (mérési időpontok helye és száma, szabadsági fok) az elképzelt stabilitásvizsgálati tervből adódik.

A sarzsok egyesíthetőek-e?	Az egy adott időpontban vett minták kezelése	I. hatásereőség		II. hatásereőség	
		A	B	A	B
nem	egyenként	4,77	3,60	3,39	2,75
	átlaguk	10,14	7,38	6,75	5,22
igen	egyenként	2,94	2,38	2,29	2,02
	átlaguk	5,04	4,03	3,82	3,33

20. táblázat: Az egyoldali, 95%-os konfidencia-sáv szélessége 2 évnél (hatóanyag-tartalom %)

A 20. táblázatból látható, hogy

- A konfidencia-sáv szűkülésének kedvez, ha a sarzsok adatait egyesíteni lehet ill. ha az egy adott időpontban vett mintákat egyenként kezeljük. Ennek oka abban keresendő, hogy mindkét előbbi feltételnél lényegesen nő a mérési pontok száma, s ez jelentősen csökkenti a konfidencia-sáv képletben ((1) képlet) szereplő gyökjel alatti kifejezés értékét.
- A legszűkebb konfidencia-intervallumot akkor kapjuk, ha a sarzsokat egyesíthetjük és az egy adott időpontban vett mintákat egyenként kezeljük (20. táblázat 3. sora). Ekkor a konfidencia-sáv szélessége 2 évnél 2-3 HA% körül lesz, tehát 5-7%-os hatóanyag-tartalom bomlás esetén a tervezett 2 éves eltarthatósági idő még tartható.
- Ha a sarzsokat egyesíthetjük, de az egy adott időpontban vett minták átlagértékével számolunk (ami pedig csak az adatfeldolgozáson, azaz rajtunk múlik, ellentétben a sarzsok egyesíthetőségével, ami a jelenlegi szabályozás szerint feltételekhez kötött (ld. II.3.2. fejezet), akkor a konfidencia-intervallum szélesség már 3,5-5% körül lesz.
- Ha a sarzsok nem egyesíthetőek, akkor mindenféleképpen célszerű a stabilitási adatokat együttesen kezelni, mert akkor a konfidencia-sáv szélessége 3-5% körül lesz, míg ha csak az átlagértékeket vizsgáljuk, akkor a szélesség már 5-10% között mozog.
Az 5-10% körüli konfidencia-intervallum szélesség pedig azt jelenti, hogy - figyelembe véve, hogy a hatóanyag-tartalomra vonatkozó alsó specifikációs határ általában 90-95% - még abban az esetben is, ha a gyógyszerkészítmény kiindulási hatóanyag-tartalma 100% és a gyógyszerkészítmény tárolása során nem következik be degradáció, azaz az adatokra vízszintes egyenes illeszthető, a konfidencia-sáv alsó határa már a tervezett 2 év környékén metszi a hatóanyag-tartalomra vonatkozó alsó specifikációs határt.

Fontosnak tartom megjegyezni, hogy ugyan a 19. táblázatból úgy tűnt, hogy az egy adott időpontban vett minták egyenkénti kezelése ill. a sarzsok egyesítése nem kedvező, mivel növeli egy jövőbeli stabilitási pont becsült varianciáját, ám ez a hatás a konfidencia-sáv szélességének számításánál bőven kompenzálódik a mérési pontok számának növekedése adta előnnyel.

A 21. táblázat a második megközelítés (szimulációs vizsgálat) szerinti eredményeket mutatja a gyógyszerkészítmény I. hatáserősségének „A” hatóanyagánál. Itt az egyedi hatóanyag-tartalom meghatározásából és az analitikai módszer validálásából becsült elemi varianciakomponensekkel (ld. 15. táblázat, konkrét értékeit erre a példára pedig a 17. és 18.

táblázatban találjuk) stabilitási adatsorokat szimuláltam (1000 futtatást végeztem). Ezután minden egyes futtatásnál a kapott hatóanyag-tartalom adatokra (az egy adott időpontban vett minták átlagára) egyenest illesztettem, majd meghatároztam a reziduális szórás értékét. Végül a reziduális szórás és az elképzelt stabilitásvizsgálati terv paraméterei (mérési időpontok ill. minták száma) alapján kiszámítottam az egyoldali, 95%-os konfidencia-sávot 2 évnél.

I. hatásereőség „A” hatóanyag	Az egyoldali, 95%-os konfidencia-sáv szélessége 2 évnél	
	Átlag*	Szórás*
A sarszok nem egyesíthetők	8,98	3,84
A sarszok egyesíthetők	4,06	1,24

21. táblázat: A szimulációs vizsgálat eredménye

(* 1000 futtatás adataiból számítva)

Összehasonlítva a 20. és 21. táblázatot megállapíthatjuk, hogy a kétféle megközelítéssel számított konfidencia-intervallumok szélessége közelítőleg megegyezik.

VII. ÖSSZEFOGLALÁS

Doktori munkám során a gyógyszerkészítmények piacra kerülhetőségének, azaz a törzskönyvezésnek egyik fontos részterületével, a stabilitásvizsgálat statisztikai elemzésével foglalkoztam. A stabilitásvizsgálat kivitelezésére és statisztikai elemzésére vonatkozó jelenleg érvényes és (Európában, az USA-ban és Japánban) egységes előírásokat az ICH Q1A és Q1E Guideline-k tartalmazzák. A stabilitási adatok statisztikai értékelésére vonatkozó szabályozás számos pontjáról azonban a szakirodalomban ma is élénk vita folyik.

A disszertációban bemutatott kutatási munka három részre tagolható aszerint, hogy a gyógyszerkészítmények stabilitásvizsgálatának statisztikai elemzését mely ponton érinti. Két témakörben a stabilitásvizsgálat jelenlegi statisztikai modelljét kritizálva a modell továbbfejlesztésére, pontosítására törekedtem az eltarthatósági idő valóságnak megfelelőbb becslése érdekében. A harmadik témakörben - új problémafelvetésként - a stabilitási adatok ingadozásának előzetes becslésére tettem kísérletet.

Az alábbiakban röviden összefoglalom az egyes témakörökben elért eredményeimet.

Az ICH Guideline által ajánlott statisztikai modell csak a vizsgált sarzsok különbözőségét, valamint a tárolás során jelentkező hatóanyagtartalom-változást veszi figyelembe. A stabilitási adatok ingadozásának egyik fő (legtöbbször a legjelentősebbnek bizonyuló) forrásának, az analízis-időpontok különbözősége okozta reprodukálhatósági hibának modellbe vételére csak néhány utalást találunk az irodalomban. Ezek közül ki kell emelni Norwood⁴⁸ cikkét, aki a reprodukálhatósági hibát egy új, véletlen jellegű faktorként kezelte, ám a faktorok struktúráját rosszul határozta meg: az egy adott időpontban érvényes reprodukálhatósági hibát sarzsonként különbözőnek tekintette, azaz hierarchikus osztályozást alkalmazott keresztosztályozás helyett.

Az általam ajánlott új modell az analízis-időpont faktor újradefiniálásával a reprodukálhatósági hibát már a valóságnak megfelelően, a sarzstól független faktorként veszi figyelembe.

Az analízis-időpont faktor újradefiniálásával a modellben két külön időskála jelenik meg, melyre az irodalomban eddig nem találunk példát. Az abszolút időskála a naptári évnek felel meg, és az analízis időpontok különbözőségét (azaz az eltérő analitikai körülmények okozta reprodukálhatósági hibát) jellemző véletlen faktorhoz kapcsolódik. A tárolási időskála a

gyógyszer gyártása óta eltelt időt mutatja, és a degradációs rátához kötődik. Könnyen belátható, hogy a két időskála nem azonos egymással, ha a vizsgált szarzsok tárolása különböző időpontokban indul.

A gyógyszerkészítmények eltarthatósági idejének becslésénél az adatok ingadozását kifejező bizonytalansági tartományra az ICH Guideline a konfidencia-sávot javasolja. Az ICH Guideline elképzelése azonban nem valóság, mert nem a szakmailag releváns kérdésre adja meg a választ. A valóságban ugyanis nem a homogénnek feltételezett tabletták sokaság várható értékének időbeni változására vagyunk kíváncsiak, hanem arra, hogy a tabletták sokaság elemei nagy hányadának hatóanyag-tartalma mely időpontig lesz egy elfogadható tartományon belül. Erre a kérdésre a tolerancia-intervallumon alapuló eltarthatósági idő becslés adhat választ.

Kezdeti kutatásommal párhuzamosan Kiermeier és társai⁴² is szorgalmazták a tolerancia-intervallum stabilitásvizsgálatokra való alkalmazását. Kiermeier tolerancia-intervallumon alapuló modellje azonban csak egy ingadozási forrást, a szarzsokon belüli inhomogenitást veszi figyelembe, az analitikai bizonytalanságot elhanyagolja.

Az általam kidolgozott tolerancia-intervallumon alapuló modell egyetlen szarzsokon belül is két ingadozási forrással - a tabletták különbözőségével és az analitikai mérés hibájával - számol. A modell szerkezete lehetővé teszi további ingadozási források (pl. a szarzsok különbözősége) egyszerű bevonását.

A másik, eddig csak általam vizsgált kérdés e témakörben, hogy a stabilitásvizsgálat értékelésénél a mért vagy a valódi (tényleges) hatóanyag-tartalomra adjunk-e ingadozási tartományt. Véleményem szerint jogosabb lenne a valódi hatóanyag-tartalom elhelyezkedését vizsgálni, hisz a beteg nem a tabletták laboratóriumban mért hatóanyag-tartalma, hanem a bevett tabletták tényleges hatóanyag-tartalma érdekli.

A gyógyszergyárak alapvető érdeke, hogy gyógyszerkészítményeik becsült eltarthatósági ideje a lehető leghosszabb legyen. A stabilitási adatok ingadozása azonban jelentősen behatárolhatja a becsült eltarthatósági idő hosszát.

Ha a stabilitási adatok ingadozási forrásainak nagyságrendjét előzetesen ismernénk, akkor már a stabilitásvizsgálat indítása előtt képet kaphatnánk a majdani stabilitási adatok ingadozásának mértékéről, és így eldönthetnénk, hogy a jelenlegi gyártási folyamat és az analitikai mérési módszer bizonytalansága megfelelő-e (azaz nem túl nagy) a kívánt lejáratási idő eléréséhez.

A fenti célt szem előtt tartva kidolgoztam egy - a varianciaanalízisre és a varianciakomponens-elemzésre épülő - számítási módszert, mely a stabilitásvizsgálat kezdetekor már rendszerint rendelkezésre álló gyártási tapasztalatokból (vagy az addigra már elvégzett Content Uniformity vizsgálatból) származó sarzsok közötti és sazson belüli inhomogenitási információ és az analitikai módszer kidolgozásakor elvégzett validálás eredményeinek felhasználásával lehetővé teszi a stabilitási adatok ingadozását okozó elemi varianciakomponensek becslését.

VIII. ÚJ TUDOMÁNYOS EREDMÉNYEK

1. tézis

A gyógyszerkészítmények stabilitásvizsgálatának elemzésére szolgáló statisztikai modellben továbbfejlesztettem Norwood módszerét az analízis-időpont okozta hiba figyelembe vételére. Az ICH Guideline kontextusában maradva, azaz a konfidencia-sávok szemléletét megtartva, az általa javasolt hierarchikus modell helyett keresztosztályozást alkalmaztam. Ezzel együtt járt az analízis-időpont faktor szintjeinek számának megváltozása, melyet az új modellben a mintavételi/mérési időpontokhoz rögzítettem.

Az így kapott modell az irodalomban fellelhető modellekkel ellentétben két időskálát tartalmaz, mely lehetővé teszi a sarzsok korának és az analízis-időpontok különbözőségének elkülönült figyelembevételét [2].

2. tézis

A gyógyszerkészítmények eltarthatósági idejének becslésénél az adatok ingadozását kifejező bizonytalansági tartományra az ICH Guideline által ajánlott konfidencia-sáv helyett a tolerancia-sávot javasoltam.

Az irodalomban található egyetlen ingadozás-forrást tartalmazó modellek helyett egy olyan többfaktoros tolerancia-intervallumon alapuló modellt dolgoztam ki, ami a sarzson belüli két ingadozási forrást - a tabletták különbözőségét és az analitikai mérés hibáját - egyaránt figyelembe veszi [1].

3. tézis

Az irodalomban elsőként ajánlottam, hogy a gyógyszerkészítmények stabilitásvizsgálatának elemzésére szolgáló tolerancia-intervallum a mért helyett a valódi hatóanyag-tartalomra vonatkozzék.

A két ingadozási forrást figyelembevevő tolerancia-intervallumok szélessége függ attól, hogy a mért vagy a valódi hatóanyag-tartalom elhelyezkedését vizsgáljuk. A valódi hatóanyag-tartalom tolerancia-intervalluma szűkebb a mérhető adatra vonatkozóénál [1].

4. tézis

A tolerancia-intervallumon alapuló stabilitásvizsgálati modellek statisztikai tulajdonságait szimuláció segítségével vizsgálva megállapítottam [1], hogy:

- A. Két megkülönböztethető ingadozási forrás jelenléte esetén a két ingadozási forrást tartalmazó modellek jól leírták a szimulációs eseteket (a lefedési valószínűség értéke az elvárt 0,95 körül volt), míg a csak egy összingadozással számoló modell nem volt alkalmas a két ingadozást feltételező szimulációs esetek megfelelő leírására (a lefedési valószínűség értéke jóval kisebb volt az elvárt 0,95-től). Így kijelenthető, hogy mivel a stabilitásvizsgálatok során a valóságban mindig több ingadozási forrás hatása jelentkezik, az irodalomban eddig fellelhető, egy ingadozási forrást tartalmazó modell nem megfelelő a stabilitási adatok leírására.
- B. A különböző varianciakomponens-kombinációkat vizsgálva megállapítottam, hogy egyrészt az összvariancia növekedésével nő a tolerancia-intervallumok szélessége, másrészt az említett két ingadozási forrás feltételezése és adott összvariancia esetén a tolerancia-intervallum abban az esetben lesz szűkebb, ha az inhomogenitási szórásnégyzet kisebb az analitikai hibához képest.
- C. A különböző ismétlésszám-kombinációk vizsgálva megállapítottam, hogy az egy sarzból egy időpontban vizsgált minták számának növekedésével a tolerancia-intervallum lényegesen szűkül, az analízisek számának növelésével nem. Ez azt jelenti, hogy a stabilitásvizsgálat tervezésekor célszerűbb az egy adott időpontban rendelkezésre álló analízis-számot (a gyógyszergyári gyakorlattól eltérő módon) úgy elosztani, hogy a lehető legtöbb mintát vegyük, s az azokból végzett ismételt analízisek számát csökkentjük.

5. tézis

A stabilitási adatok bizonytalanságának előzetes becslésére alkalmas számítási módszert dolgoztam ki, melyhez az analitikai módszer kidolgozásakor elvégzett validálás eredményeit és a szokásos gyártási tapasztalatokból (vagy a Content Uniformity vizsgálatból) származó inhomogenitási információt használtam fel.

Módszerem alkalmazásával még a stabilitásvizsgálat kezdete előtt lehetővé válik annak vizsgálata, hogy az adott gyógyszerkészítmény jelenlegi gyártási folyamatának és analitikai mérési módszerének bizonytalansága megfelelő-e (azaz nem túl nagy) a kívánt lejáratú idő eléréséhez [3].

IX. AZ ÉRTEKEZÉS TÉMÁJÁBAN MEGJELENT PUBLIKÁCIÓK

1. **K. Komka**, S. Kemény, B. Bánfai: Novel tolerance interval model for the estimation of the shelf-life of pharmaceutical products, *Journal of Chemometrics*, 24, 131-139 (2010). **(IF: 1.415)**
2. **K. Komka**, S. Kemény: A modified error model for the assessment of stability of pharmaceutical products, *Chemometrics and Intelligent Laboratory Systems*, 72, 161-165 (2004). **(IF: 1.940)**
3. **K. Komka**, S. Kemény: Effect of the uncertainty of the stability data on the shelf life estimation of pharmaceutical products, *Periodica Polytechnica Ser. Chemical Engineering*, 48(1), 41-52 (2004).

X. IRODALOMJEGYZÉK

- ¹ Paál Tamás Dr.: Az OGYI Gyógyszertörzskönyvező tevékenysége, *Gyógyszereink*, 45, 1-5 (1995)
- ² 2005. évi XCV. törvény az emberi alkalmazásra kerülő gyógyszerekről és egyéb, a gyógyszerpiacot szabályozó törvények módosításáról (5. § (1) pont)
- ³ ICH Harmonised Tripartite Guideline Q1A(R2): Stability Testing of New Drug Substances and Products (Second Revision), 2003
- ⁴ ICH Harmonised Tripartite Guideline Q1E: Evaluation for Stability Data, 2003
- ⁵ Van der Veen, A.M.H.; Linsinger, T.P.J.; Lamberty, A.; Pauwels, J.: Uncertainty calculations in the certification of reference materials 3. Stability study, *Accreditation and Quality Assurance*, 6(6), 257-263 (2001)
- ⁶ Pauwels, J.; Lamberty, A.; Schimmel, H.: Quantification of the expected shelf-life of certified reference materials, *Fresenius Journal of Analytical Chemistry*, 361, 395-399 (1998)
- ⁷ Lamberty, A.; Schimmel, H.; Pauwels, J.: The study of the stability of reference materials by isochronous measurements, *Fresenius Journal of Analytical Chemistry*, 360, 359-361 (1998)
- ⁸ Chow, S.C.; Liu, J.P.: Accelerated Testing, Chapter 9 in Statistical Design and Analysis in Pharmaceutical Science, *Marcel Dekker, Inc.*, New York, 1995, p. 295-327
- ⁹ Davies, O.L.; Hudson, H.E.: Stability of drugs: accelerated storage tests, Chapter 15 in Statistics in the Pharmaceutical Industry, ed. Buncher, C.R. and Tsay, J.Y., *Marcel Dekker, Inc.*, New York, 1981, p. 355-395
- ¹⁰ Davies, O.L.; Budgett, D.A.: Accelerated storage tests on pharmaceutical products: Effect of error structure of assay and errors in recorded temperature, *Journal of Pharmacy and Pharmacology*, 32, 155-159 (1980)
- ¹¹ Yoshioka, S.; Aso, Y.; Uchiyama, M.: Statistical evaluation of nonisothermal prediction of drug stability, *Journal of Pharmaceutical Sciences*, 76(10), 794-798 (1987)
- ¹² Ertel, K.D.; Carstensen, J.T.: Examination of a modified Arrhenius relationship for pharmaceutical stability prediction, *International Journal of Pharmaceutics*, 61, 9-14 (1990)
- ¹³ Carstensen, J.T.: Drug Stability, *Marcel Dekker, Inc.*, New York, 1995
- ¹⁴ Chow, S.C.; Liu, J.P.: Statistical Design and Analysis in Pharmaceutical Science, *Marcel Dekker, Inc.*, New York, 1995

- ¹⁵ Carstensen, J.T.; Franchini, M.; Ertel, K.: Statistical approaches to stability protocol design, *Journal of Pharmaceutical Sciences*, 81(3), 303-308 (1992)
- ¹⁶ Lin, K.K.; Lin, T.Y.D.; Kelly, R.E.: Stability of Drugs, Chapter 21A in *Statistics in the Pharmaceutical Industry*, ed. Buncher, C.B.; Tsay, J.Y., *Marcel Dekker, Inc.*, New York, 1994, p. 419-444
- ¹⁷ Langenbucher, F.: Statistical analysis of stability data by means of general linear models (GLM), *Drug Development and Industrial Pharmacy*, 17(2), 165-181 (1991)
- ¹⁸ StatSoft, Inc. (2003). STATISTICA (data analysis software system), version 6.1. www.statsoft.com.
- ¹⁹ Bancroft, T.A.: Analysis and inference for incompletely specified models involving the use of preliminary test(s) of significance, *Biometrics*, 20, 427-442 (1964)
- ²⁰ Fairweather, W.R.; Lin, T.Y.D.; Kelly, R.: Regulatory, design, and analysis aspects of complex stability studies, *Journal of Pharmaceutical Sciences*, 84(11), 1322-1326 (1995)
- ²¹ Ahn, H.; James, J.C.; Lin, T.Y.D.: A two-way analysis of covariance model for classification of stability data, *Biometrical Journal*, 39(5), 559-576 (1997)
- ²² Chen, J.J.; Ahn H.; Tsong Y.: Shelf-life estimation for multifactor stability studies, *Drug Information Journal*, 31, 573-587 (1997)
- ²³ Helboe, P.: New designs for stability testing programs: Matrix and factorial design. Authorities viewpoint on the predictive value of such studies, *Drug Information Journal*, 26, 629-634 (1992)
- ²⁴ ICH Harmonised Tripartite Guideline Q1D: Bracketing and Matrixing Designs for Stability Testing of New Drug Substances and Products, 2002
- ²⁵ Golden, M.H.; Cooper, D.C.; Riebe, M.T.; Carswell, K.E.: A matrix approach to long-term stability testing of pharmaceutical products, *Journal of Pharmaceutical Sciences*, 85(2), 240-245 (1996)
- ²⁶ Natarajan, J.; Altan, S.; Raghavarao, D.: Expiration dating of pharmaceutical compounds in relation to analytical variation, degradation rate and matrix designs, *Drug Information Journal*, 31, 589-595 (1997)
- ²⁷ Oliva, A.; Farina, J.; Llabres M.: Comparison of shelf-life estimates for a human insulin pharmaceutical preparation using the matrix and full-testing approaches, *Drug Development and Industrial Pharmacy*, 29(5), 513-521 (2003)
- ²⁸ Shao, J.; Chen, L.: Prediction bounds for random shelf-lives, *Statistics in Medicine*, 16, 11167-11173 (1997)

- ²⁹ Ruberg, S.J.; Hsu, J.C.: Multiple comparison procedures for pooling batches in stability studies, *Technometrics*, 34, 465-472 (1992)
- ³⁰ Ruberg, S.J.; Stegeman, J.W.: Pooling data for stability studies: testing the equality of batch degradation slopes, *Biometrics*, 47, 1059-1069 (1991)
- ³¹ Kemény S., Deák A.: Kísérletek tervezése és értékelése, *Műszaki Könyvkiadó*, Budapest, 2000
- ³² Chow, S.C.; Shao, J.: Test for batch-to-batch variation in stability analysis, *Statistics in Medicine*, 8, 883-890 (1989)
- ³³ Chow, S.C.; Shao, J.: Estimating drug shelf-life with random batches, *Biometrics*, 47, 1071-1079 (1991)
- ³⁴ Chow, S.C.; Wang, S.G.: On the estimation of variance components in stability analysis, *Communication in Statistics - Theory and Methods*, 23(1), 289-303 (1994)
- ³⁵ Hildreth, C.; Houck, J.P.: Some estimators for a linear model with random coefficients, *American Statistical Association Journal*, 63, 584-595 (1968)
- ³⁶ Shao, J.; Chow, S.C.: Statistical inference in stability analysis, *Biometrics*, 50, 753-763 (1994)
- ³⁷ Sun, Y.; Chow, S.C.; Li, G.; Chen, K.W.: Assessing distributions of estimated drug shelf lives in stability analysis, *Biometrics*, 55, 896-899 (1999)
- ³⁸ Murphy, J.R.; Hofer, J.D.: Establishing shelf life, expiry limits and release limits, *Drug Information Journal*, 36, 769-781 (2002)
- ³⁹ Carstensen, J.T.; Nelson, E.: Terminology regarding labeled and contained amounts in dosage forms, *Journal of Pharmaceutical Sciences*, 65(2), 311-312 (1976)
- ⁴⁰ Wessels, P.; Holz, M.; Erni, F.; Krummen, K.; Ogorka, J.: Statistical evaluation of stability data of pharmaceutical products for specification setting, *Drug Development and Industrial Pharmacy*, 23(5), 427-439 (1997)
- ⁴¹ Hahn, G. J.: Statistical intervals for normal population, Part I. Tables, Examples and Applications, *Journal of Quality Technology*, 2(3), 115-125 (1970)
- ⁴² Kiermeier, A.; Jarett, R.G., Verbyla, A.P.: A new approach to estimating shelf-life, *Pharmaceutical Statistics*, 3, 3-11, (2004)
- ⁴³ Renger, B.: System performance and variability of chromatographic techniques used in pharmaceutical quality control, *Journal of Chromatography B*, 745, 167-176 (2000)
- ⁴⁴ Ermer, J.; Arth, C.; De Raeve, P.; Dill, D.; Friedel, H-D.; Höwer-Fritzen, H.; Kleinschmidt, G.; Köller, G.; Köppel, H.; Kramer, M.; Maegerlein, M.; Schepers, U.; Wätzig, H.: Precision

from drug stability studies Investigation of reliable repeatability and intermediate precision of HPLC assay procedures, *Journal of Pharmaceutical and Biomedical Analysis*, 38, 653-663 (2005)

⁴⁵ Renger, B.: Chromatography in pharmaceutical analysis: Uncertainty versus specifications, *Pharmaceutical Technology Europe*, 9, 36-44 (1997)

⁴⁶ Bar, R.: Statistical evaluation of stability data for assessing variability of an analytical method, *Accreditation and Quality Assurance*, 4(6), 235-239 (1999)

⁴⁷ Bar, R.: Statistical evaluation of stability data: Criteria for change over-time and data variability, *PDA Journal of Pharmaceutical Science and Technology*, 57(5), 369-377 (2003)

⁴⁸ Norwood, T.E.: Statistical analysis of pharmaceutical stability data, *Drug Development and Industrial Pharmacy*, 12(4), 553-560 (1986)

⁴⁹ Holz, M., Dill, D., Frank, R., Wember, T.: Auswertung und Optimierung von Haltbarkeitsuntersuchungen (Teil 1 Grundlagen), *Pharmazeutische Industrie*, 63(1), 85-93 (2001)

⁵⁰ Fuller, W.A.; Battese, G.E.: Transformations for estimation of linear models with nested-error structure, *Journal of the American Statistical Association*, 63, 626-632 (1973)

⁵¹ Milliken, G.A.; Johnson, D.E.: Analysis of Messy Data, Vol.I.: Designed Experiments, *Chapman&Hall*, London, 1992

⁵² Henderson, C.R.: Estimation of variance and covariance components, *Biometrics*, 9(2), 226-252 (1953)

⁵³ Searle, S.R.: Topics in Variance Component Estimation, *Biometrics*, 27, 1-76 (1971)

⁵⁴ SAS/STAT software version 8.2 (2003)

⁵⁵ Owen, D.B.: A Survey of Properties and Applications of the Noncentral *t*-Distribution, *Technometrics*, 10(3), 445-478 (1967)

⁵⁶ De Gryze S, Langhans I, Vandebroek M.: Using the correct intervals for prediction: A tutorial on tolerance intervals for ordinary least-squares regression, *Chemometrics and Intelligent Laboratory Systems*, 87, 147-154 (2007)

⁵⁷ Kiermeier, A.: Shelf-life: Designing and Analysis Stability Trials, PhD Thesis (2003)

⁵⁸ Satterthwaite F.E.: An approximate distribution of estimates of variance components, *Biometrics Bulletin*, 2, 110-114 (1946)

⁵⁹ Wolfram Mathematica. version 6.0.0 (2007)

⁶⁰ Wang, C.M.; Iyer, H.K.: Tolerance intervals of the distribution of true values in the presence of measurement error, *Technometrics*, 36, 162-170 (1994)

⁶¹ Williams, J.S.: A confidence interval for variance components, *Biometrika*, 49, 278-81 (1962)

⁶² European Pharmacopoeia (Ph. Eur.): Supplement 6.1 (official April 2008) Uniformity of Dosage Units (reference 01/2008: 20940)

⁶³ ICH Harmonised Tripartite Guideline Q2(R1): Validation of Analytical Procedures: Text and Methodology, 2005