



Közzététel: 2023. november 22.

A tanulmány címe:

**Profil-likelihood konfidenciaintervallum alkalmazása betegség prevalenciájára, ha a diagnosztikai teszt hibával terhelt**

Szerzők:

**HÁRSFALVI PÉTER**

a BiTrial Kft. kutatási igazgatója, az Állatorvostudományi Egyetem Biostatistika Tanszékének doktorandusza  
E-mail: peter.harsfalvi@bitrial.hu

**REICZIGEL JENŐ**

az Állatorvostudományi Egyetem Biostatistika Tanszékének egyetemi tanára  
E-mail: reiczigel.jeno@univet.hu

DOI: <https://doi.org/10.20311/stat2023.11.hu1033>

**Az alábbi feltételek érvényesek minden, a Központi Statisztikai Hivatal (a továbbiakban: KSH) Statisztikai Szemle c. folyóiratában (a továbbiakban: Folyóirat) megjelenő tanulmányra. Felhasználó a tanulmány vagy annak részei felhasználásával egyidejűleg tudomásul veszi a jelen dokumentumban foglalt felhasználási feltételeket, és azokat magára nézve kötelezőnek fogadja el. Tudomásul veszi, hogy a jelen feltételek megszegéséből eredő valamennyi kárért felelősséggel tartozik.**

1. A jogszabályi tartalom kivételével a tanulmányok a szerzői jogról szóló 1999. évi LXXVI. törvény (Szjt.) szerint szerzői műnek minősülnek. A szerzői jog jogosultja a KSH.
2. A KSH földrajzi és időbeli korlátozás nélküli, nem kizárólagos, nem átadható, térítésmentes felhasználási jogot biztosít a Felhasználó részére a tanulmány vonatkozásában.
3. A felhasználási jog keretében a Felhasználó jogosult a tanulmány:
  - a) oktatási és kutatási célú felhasználására (nyilvánosságra hozatalára és továbbítására a 4. pontban foglalt kivétellel) a Folyóirat és a szerző(k) feltüntetésével;
  - b) tartalmáról összefoglaló készítésére az írott és az elektronikus médiában a Folyóirat és a szerző(k) feltüntetésével;
  - c) részletének idézésére – az átvevő mű jellege és célja által indokolt terjedelemben és az eredetihez híven – a forrás, valamint az ott megjelölt szerző(k) megnevezésével.
4. A Felhasználó nem jogosult a tanulmány továbbértékesítésére, haszonszerzési célú felhasználására. Ez a korlátozás nem érinti a tanulmány felhasználásával előállított, de az Szjt. szerint önálló szerzői műnek minősülő mű ilyen célú felhasználását.
5. A tanulmány átdolgozása, újra publikálása tilos.
6. A 3. a)–c) pontban foglaltak alapján a Folyóiratot és a szerző(ke)t az alábbiak szerint kell feltüntetni:  
„*Forrás: Statisztikai Szemle c. folyóirat 101. évfolyam 11. számában megjelent, Hársfalvi Péter–Reiczigel Jenő által írt, Profil-likelihood konfidenciaintervallum alkalmazása betegség prevalenciájára, ha a diagnosztikai teszt hibával terhelt* című tanulmány (link csatolása)”
7. A Folyóiratban megjelenő tanulmányok kutatói véleményeket tükröznek, amelyek nem feltétlenül esnek egybe a KSH vagy a szerzők által képviselt intézmények hivatalos álláspontjával.

Hársfalvi Péter – Reiczigel Jenő

## Profil-likelihood konfidenciaintervallum alkalmazása betegség prevalenciájára, ha a diagnosztikai teszt hibával terhelt

### Profile likelihood confidence interval for the prevalence assessed by an imperfect diagnostic test

Hársfalvi Péter, a BiTrial Kft. kutatási igazgatója, az Állatorvostudományi Egyetem Biostatisztika Tanszékének doktorandusza

E-mail: peter.harsfalvi@bitrial.hu

Reiczigel Jenő, az Állatorvostudományi Egyetem Biostatisztika Tanszékének egyetemi tanára

E-mail: reiczigel.jeno@univet.hu

A tanulmányban ismertetünk egy új konfidenciaintervallumot a betegségek prevalenciabecslésének azon eseteire, amikor a felhasznált diagnosztikai teszt szenzitivitását és specificitását a vizsgálati mintától független mintákból becsüljük. Az új eljárás alapja a profil-likelihood módszer volt, az intervallum lefedési valószínűségének javítása érdekében pedig további korrekciót is alkalmaztunk. Az intervallum lefedési valószínűségét és várható hosszát szimulációval értékeltük, és összehasonlítottuk két másik, ugyanezre a problémára javasolt módszerrel, nevezetesen *Lang és Reiczigel (2014)*, illetve *Flor és munkatársai (2020)* módszereivel. Az új intervallum várható hossza rövidebb, mint a Lang–Reiczigel-féle intervallumé, miközben lefedési valószínűségük nagyjából megegyezik. A Flor-féle intervallummal összehasonlítva a várható hossz hasonló, de az új intervallum lefedési valószínűsége nagyobb. Összességében tehát az új intervallum mindkét versenytársánál jobbnak bizonyult.

Tárgyszavak: profil-likelihood konfidenciaintervallum, téves besorolások, becsült szenzitivitás és specificitás

We present a new confidence interval for the prevalence of a disease for a situation when sensitivity and specificity of the diagnostic test are estimated from validation samples independent of the study sample. The new interval is based on profile likelihood and incorporates an adjustment improving the coverage probability. Its coverage probability and expected length were assessed by simulation and compared to two other methods for this problem, namely those by *Lang and Reiczigel (2014)* and *Flor et al. (2020)*. Expected length of the new interval is less than that of the Lang and Reiczigel interval while its coverage is about the same. Comparison to the Flor interval resulted in similar expected length but higher coverage probabilities for the new interval. All in all, the new interval proved to be better than both its competitors.

Keywords: profile likelihood confidence interval, misclassification, estimated sensitivity and specificity

Egy betegség prevalenciájának becslése alapvető feladat az epidemiológiai vizsgálatokban. Prevalenciának nevezzük egy meghatározott időpontban az illető betegségben szenvedők részarányát egy vizsgált populációban. Mivel a legtöbb diagnosztikai teszt nem tévedhetetlen, alkalmazásuk során előfordulhatnak téves besorolások (más néven félreosztályozások). A statisztikai tesztekhez hasonlóan a téves diagnózis is kétféle lehet, téves pozitív vagy téves negatív, ezért a diagnosztikai eljárások jóságát is két mutatóval érdemes jellemezni. A szenzitivitás, azt mutatja meg, hogy a diagnosztikai teszt a betegek hány százalékát diagnosztizálja helyesen, a specificitás pedig azt, hogy a teszt a betegségben nem szenvedők hány százalékát diagnosztizálja helyesen (*Bailey–Handu, 2013*).

A mutatók konkrét példán történő érzékeltetéséhez érdemes megvizsgálni egy SARS-CoV-2-(Covid-)gyorsteszt klinikai vizsgálatban tapasztalt jellemzőit. *Jegerlehner és munkatársai (2021)* 1465 résztvevőt vizsgáltak, a „gold standard” RT-PCR-teszt segítségével meghatározott valódi prevalencia 9,6% volt (141 beteg az 1465 résztvevőből). A kutatásban használt Roche/SD Biosensor rapid antigénteszt szenzitivitása 65,2% (92 a 141-ből), specificitása pedig 99,8% volt (1319 az 1321-ből). A részletes eredmények összefoglalásából látható (1. táblázat), hogy a teszt használata esetén az alacsony szenzitivitás következtében a fals negatív eredmények száma jelentős, továbbá kitűnik az is, hogy amennyiben pusztán a rapid antigénteszt eredményei alapján szeretnénk becslést készíteni a prevalenciára, akkor a valódi 9,6% helyett mindössze 6,4%-ot kapnánk (94 pozitív teszteredmény 1462 tesztből, 3 értékelhetetlen eredmény mellett).

1. táblázat

**Teszteredmények *Jegerlehner et al. (2012)* alapján  
a Roche/SD Biosensor rapid antigénteszt vizsgálatában**  
*Test results of the Roche/SD Biosensor rapid antigen test from the study of  
Jegerlehner et al. (2012)*

Teszteredmény	Valódi betegségstátusz	
	ténylegesen beteg	ténylegesen nem beteg
Pozitív	92 fő valódi pozitív	2 fő fals pozitív
Negatív	49 fő fals negatív	1319 fő valódi negatív
Összesen	141 fő	1321 fő

Amikor egy betegség prevalenciáját egy diagnosztikai teszt segítségével szeretnénk megbecsülni, fontos, hogy a potenciális félreosztályozásokat is figye-

lembe vegyük. Rogan és Gladen (1978) javaslatot tett a tesztpozitívok megfigyelt arányának (az úgynevezett látszólagos prevalenciának) a korrekciójára a diagnosztikai teszt ismert szenzitivitása és specificitása esetén. A valódi prevalenciára normális közelítésen alapuló konfidenciaintervallumot (KI) is ajánlottak, amelynek lefedési valószínűsége azonban meglehetősen gyengének bizonyult (Greiner–Gardner, 2000). Reiczigel és munkatársai (2010) kidolgoztak egy egzakt konfidenciaintervallum-szerkesztési módszert a valódi prevalenciára a használt diagnosztikai teszt ismert szenzitivitásának és specificitásának esetére. Később Lang és Reiczigel (2014) rávilágítottak arra, hogy ha egy diagnosztikai teszt szenzitivitásának, illetve specificitásának becslését ismert értékeknek tekintjük, akkor figyelmen kívül hagyjuk a bennük rejlő bizonytalanságot, ami indokolatlanul rövid intervallumbecsléseket eredményezhet, különösen, ha a pontbecslések kis vagy közepes mintából történtek. Emiatt javasoltak egy olyan eljárást, amely figyelembe veszi a szenzitivitás és a specificitás becslésének bizonytalanságát a Rogan és Gladen által alkalmazott formula és az Agresti és Coull (1998) által használt, „adjunk hozzá két sikert és két kudarcot” módszerhez hasonló korrekció kombinálásával.

Flor és munkatársai (2020) egy bayesi kredibilis intervallumot, mégpedig a legnagyobb sűrűségű intervallumot (*highest density interval = HDI*) alkalmazták a probléma megoldására. A HDI azon lehetséges paraméterértékek tartománya, amelyeknek a legnagyobb az *a posteriori* sűrűsége. Így egy 95%-os HDI az a legrövidebb 95%-os kredibilis intervallum, amelyen kívüli bármely paraméterérték kevésbé valószínű, mint azok, amelyeket tartalmaz az intervallum. Ennek a módszernek a használata különösen akkor előnyös, ha a szenzitivitás és a specificitás becslésére használt minták kis elemszámúak, ami általában jellemző szűrővizsgálatok validációs elemzése esetén (Farnham et al., 2012).

Tanulmányunk célja egy új megközelítés, a profil-likelihood konfidenciaintervallum (*profile-likelihood confidence interval = PLCI*) bemutatása a prevalencia becslésére olyan esetekben, amikor a betegség szűrésére használt diagnosztikai teszt szenzitivitását és specificitását független validációs mintákból becsüljük. A továbbiakban megvizsgáljuk az intervallum jellemzőit, és összehasonlítjuk teljesítőképességét a hasonló problémára korábban javasolt konfidenciaintervallumokkal, a Lang és Reiczigel (2014), továbbá a Flor és munkatársai (2020) által bemutatottakkal.

## 1. Módszerek

A PLCI-t a likelihood-hányados teszt (*likelihood ratio test* = *LRT*) invertálásával kaphatjuk meg. A teszttel a  $H_0: \theta = \theta_0$  nullhipotézist vizsgáljuk, ahol  $\theta$  a vizsgált paraméter,  $\theta_0$  pedig annak feltételezett értéke. Jelölje továbbá  $x$  a megfigyelt mintát,  $\nu$  pedig egy kolonc paramétert (*nuisance parameter*). A  $\theta$  és a  $\nu$  lehet vektor is. A likelihood-hányados ( $Q$ ) két likelihood hányadosát jelöli. Az elsőt a likelihood-függvény maximalizálásával kapjuk a kolonc paraméter, a  $\nu$  változtatása, illetve a  $\theta$  parameter  $\theta_0$  értéken tartása mellett, míg a másik likelihoodhoz a maximalizálás a  $\theta$  és a  $\nu$  paraméterek változtatásával történik. A képlete:

$$Q = \frac{L_0}{L_s}, \quad (1)$$

$$\text{ahol } L_0 = \max_{\nu} L(\theta_0, \nu, x), \quad (2)$$

$$\text{és } L_s = \max_{\theta, \nu} L(\theta, \nu, x). \quad (3)$$

$$\text{A LRT-teszt próbastatisztikája: } -2 \ln Q = -2 \ln \left( \frac{L_0}{L_s} \right). \quad (4)$$

A fenti próbastatisztika közelítőleg khi-négyszet eloszlású a  $\theta$  paramétervektor dimenziószámával megegyező szabadsági fokkal. Így ha  $\theta$  skalár, akkor egy  $(1-\alpha)$ -szintű PLCI-intervallum azon paraméterértékeket tartalmazza, amelyeket  $x$  megfigyelése mellett az LRT  $\alpha$  szinten nem utasítja el:

$$\{\theta_0: -2 \ln Q < \chi_{1,1-\alpha}^2\} \quad (5)$$

ahol  $\chi_{1,1-\alpha}^2$  jelöli az 1 szabadsági fokú khi-négyszet-eloszlás  $1-\alpha$  kvantilisét (Millar, 2011).

Esetünkben a betegség prevalenciája ( $p$ ) a becsülendő paraméter, míg a szűréshez használt diagnosztikai teszt szenzitivitása ( $Se$ ) és specificitása ( $Sp$ ) a kolonc paraméterek. A rendelkezésünkre álló adatok három független ( $n_{Se}$ ,  $n_{Sp}$  és  $n$  elemű) mintából állnak  $Se$ ,  $Sp$  és  $p$  becsléséhez. Jelölje  $k_{Se}$  és  $k_{Sp}$  az  $n_{Se}$  és  $n_{Sp}$  elemű mintákban helyesen diagnosztizált alanyok számát, és jelölje  $k$  az  $n$  elemű minta pozitívként diagnosztizált egyedei számát.

Belátható, hogy  $k_{Se}$ ,  $k_{Sp}$  és  $k$  binomiális eloszlású a  $B(n_{Se}, Se)$ , a  $B(n_{Sp}, Sp)$  és a  $B(n, \tilde{p} = p \cdot Se + (1-p) \cdot (1-Sp))$  paraméterrel, a likelihood-függvény pedig az adatok ismeretében ( $n_{Se}$ ,  $n_{Sp}$ ,  $n$ ,  $k_{Se}$ ,  $k_{Sp}$ , és  $k$ ) a következő:

$$L(p, Se, Sp) = \binom{n_{Se}}{k_{Se}} Se^{k_{Se}} (1 - Se)^{n_{Se} - k_{Se}} \binom{n_{Sp}}{k_{Sp}} Sp^{k_{Sp}} (1 - Sp)^{n_{Sp} - k_{Sp}} \binom{n}{k} \tilde{p}^k (1 - \tilde{p})^{n-k}. \quad (6)$$

A korábban leírtaknak megfelelően egy  $p$  paraméterre szerkesztett  $(1-\alpha)$ -szintű PLCI-intervallum azon  $p_0$  paramétereket tartalmazza, amelyeket az LRT  $\alpha$  szinten nem utasítja el, tehát:

$$PLCI_{1-\alpha} = \left\{ p_0 : -2 \ln \left( \frac{\max_{Se, Sp} L(p_0, Se, Sp)}{\max_{p, Se, Sp} L(p, Se, Sp)} \right) < \chi_{1, 1-\alpha}^2 \right\}, \tag{7}$$

ahol  $\chi_{1, 1-\alpha}^2$  jelöli az 1 szabadsági fokú khi-négyzet-eloszlás  $1-\alpha$  kvantilisét.

A PLCI lefedési valószínűsége 95%-os nominális szint mellett akár 90% alatti is lehet, még akkor is, ha a szenzitivitás és a specificitás 99,5% feletti. Szerencsére az intervallum csak az extrém megfigyelések esetében teljesít gyengén (megfigyelt prevalencia = 0%, vagy 100%, megfigyelt Se vagy Sp = 100%). Extrém megfigyelés például, ha  $n_{Se} = 100$ ,  $n_{Sp} = 200$ ,  $n = 300$  mellett  $k_{Se} = 95$ ,  $k_{Sp} = 199$ ,  $k = 0$ . Az intervallum lefedési valószínűsége ezért jelentősen javítható egy heurisztikus korrekcióval. A korrekció abban nyilvánul meg, hogy extrém megfigyelés esetén meghatározzuk a konfidenciaintervallumot az eggyel kevésbé extrém megfigyelésre is, és e két intervallum végpontjainak számtani közepét véve kapjuk a korrigált intervallumot. Az előző példát folytatva, kiszámítjuk a PLCI-intervallumot a  $k = 1$  esetben is, majd a két esetben ( $k = 0$  és  $k = 1$ ) kapott intervallumok alsó és felső végpontjainak (*lower confidence limit, LCL* és *upper confidence limit, UCL*) átlagát véve kapjuk meg az új korrigált intervallumot. A képlet:

$$UCL_{\text{javított}} = (UCL_{k=0} + UCL_{k=1})/2. \tag{8}$$

$$LCL_{\text{javított}} = (LCL_{k=0} + LCL_{k=1})/2. \tag{9}$$

Ha egynél több megfigyelt paraméter ( $k$ ,  $k_{Se}$  és  $k_{Sp}$ ) értéke is extrém, a korrigált intervallumok unióját használjuk.

A módszerhez készítettünk egy R-függvényt, amelyben az optimalizáláshoz az `optim()` algoritmust használtuk. A PLCI végpontjait intervallumfelezéssel határoztuk meg, a paramétereknek pedig a logit transzformáltját használtuk. A függvény egy- és kétoldali intervallum készítésére is alkalmas a választott megbízhatósági szinten, és bárki számára elérhetővé tettük a GitHub oldalán az alábbi linken: <https://github.com/Ragnar0ss/>.

A konfidenciaintervallumok összehasonlításához a szakirodalomban leggyakrabban használt jellemzőiket vizsgáltuk (*Agresti-Coull, 1998; Gonçalves et al., 2012; Vollset, 1993*), a szimulált lefedési valószínűséget (*coverage probability, CP*) és az intervallum várható hosszát (*expected length, EL*). A lefedési valószínűség annak a valószínűsége, hogy egy intervallum tartalmazza a becsült paramétert, esetünkben a valódi prevalenciát. A várható hossz alatt az intervallum átlagos hosszát értjük:

$$EL(n, p) = \sum_{j=0}^n \binom{n}{j} p^j (1-p)^{n-j} (UCL(j) - LCL(j)), \tag{10}$$

ahol  $UCL(j) - LCL(j)$  az intervallum alsó és felső végpontja közötti különbség. Belátható, hogy mindkét jellemző függ a valódi paraméterértéktől, így az is le-

hetséges, hogy bizonyos paraméterértékek esetén az egyik, míg mások esetén egy másik intervallum teljesít jobban.

A könnyebb összehasonlíthatóság érdekében az új intervallumot mindkét másik vizsgált intervallummal összevetettük, mégpedig olyan körülmények között, ahogyan az eredeti cikkekben vizsgálták őket.

A Lang–Reiczigel-féle intervallum esetén az eredeti cikkben szereplő 5625 paraméterkombinációt alkalmaztuk, azaz a valódi szenzitivitásra és specifitásra 5-5 értéket használtunk (1, 0,99, 0,95, 0,90 és 0,70), a mintaelemszámra mindhárom mintánál (a szenzitivitás, a specifitás és a prevalencia becsléséhez) szintén 5-5 (30, 100, 300, 1000 és 3000), a valódi prevalenciára pedig 9 értéket (0,005, 0,01, 0,02, 0,03, 0,05, 0,1, 0,2, 0,3 és 0,5). A konfidenciaintervallumok szimmetriája miatt a 0,5-nél magasabb prevalenciaértékek vizsgálatára nem volt szükség, ugyanis a  $p = p$ ,  $Se = x$ ,  $Sp = y$ , illetve a  $p = 1 - p$ ,  $Se = y$ ,  $Sp = x$  paraméterek esetén adódó intervallumok egymás tükörképei.

A szimulációban 20 000 replikációval dolgoztunk, azaz a fenti paraméterek minden kombinációjához 20 000 számhármast generáltunk (a megfigyelt prevalenciát, szenzitivitást és specifitást), és kiszámítottuk az ezekre adódó intervallumok lefedési valószínűségét és várható hosszát az új módszerrel és a Lang és Reiczigel által leírt módszerrel is. A generált 20 000 replikáció esetében a lefedési valószínűség átlagának standard hibája mindössze 0,0015, azaz a szimuláció pontossága elfogadható szinten van.

Az intervallumok hosszának az összehasonlításához a relatív különbséget,  $100 \frac{EL_{PLCI} - EL_{Lang}}{EL_{Lang}} \%$  használtuk, amelynek értéke 0, ha a két intervallum várható hossza megegyezik, illetve negatív, ha a PLCI a rövidebb.

Flor az általa javasolt intervallumhoz az eredeti cikkben egy kisebb adathalmazt használt,  $n_{Se}$ ,  $n_{Sp}$ ,  $n$ ,  $Se$ ,  $Sp$ ,  $p$  paramétereknek 1000 darab véletlenszerűen választott kombinációját. Az  $n_{Se}$  és  $n_{Sp}$  lehetséges értékei 50, 100, 200, 500, 1000, 2000 és 5000 voltak,  $n$  egy 50 és 2000 közötti véletlen egész szám,  $Se$  és  $Sp$  a [0,6, 1] míg  $p$  a [0, 1] intervallumból választott véletlen érték. Minden egyes paraméterkombinációban 20 000 véletlen minta generálása után számítottuk ki az új intervallum és a Flor-féle intervallum esetén a lefedési valószínűség és a várható hossz értékeit.

A vizsgált Lang–Reiczigel- és Flor-intervallumok részletes bemutatása megtalálható a saját közleményünkben (Flor et al., 2020; Lang–Reiczigel, 2014).

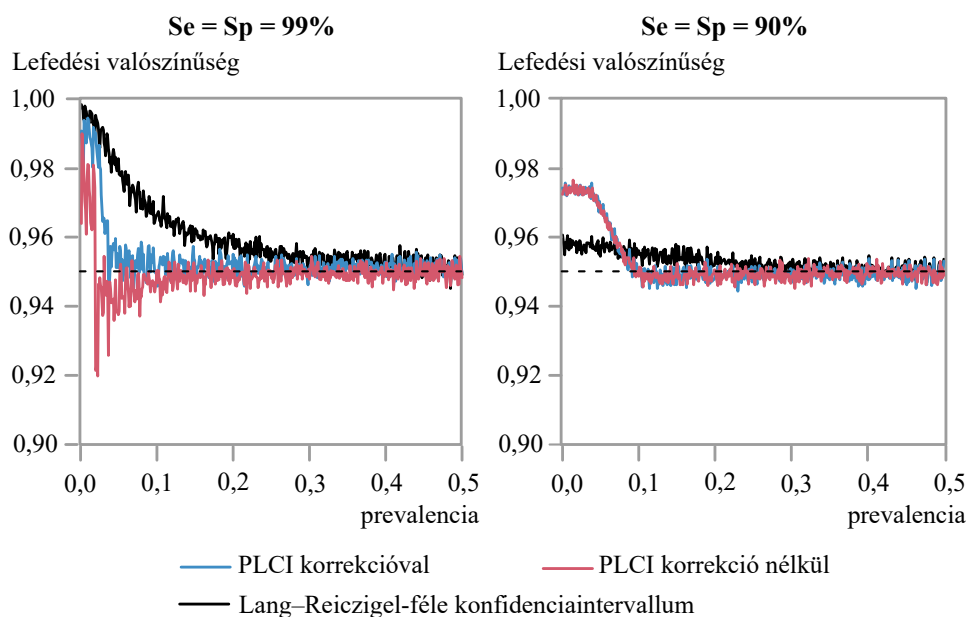
## 2. A PLCI- és a Lang–Reiczigel-intervallumok összehasonlítása

A korrigált PLCI lefedési valószínűsége hasonló a Lang–Reiczigel-intervalluméhoz, a vizsgált 5625 eset 82%-ánál (Lang) és 89%-ánál (PLCI) érték a nominális (95%) szintet. 94% alatt mindössze 12 és 27 alkalommal (0,21% és 0,48%) volt a lefedési valószínűségük. A PLCI esetében ezek az alacsony lefedések  $n = 30$  és  $100$  (24 és 3 alkalommal), a Lang–Reiczigel-intervallumoknál pedig  $n = 30$  és  $3000$  (3 és 9 alkalommal) esetén fordultak elő.

1. ábra

**95%-os konfidenciaintervallumok lefedési görbéi  $n_{Se} = n_{Sp} = n = 100$  esetén**

*Coverage curves of 95% CIs for  $n_{Se} = n_{Sp} = n = 100$*



Az 1. ábrán láthatók a Lang–Reiczigel-féle és a korrigált, valamint a korrekció nélküli PLCI-intervallumok lefedési görbéi  $n_{Se} = n_{Sp} = n = 100$ , 99%-os (bal oldali panel), illetve 90%-os (jobb oldali panel) szenzitivitás és specificitás mellett. A PLCI-intervallumnál alkalmazott korrekció előnye magas szenzitivitás-és/vagy specificitásértékek esetén a legélesebb: korrekció nélkül a lefedési valószínűség ezekben az esetekben akár 92% alá is eshet, míg a korrekció alkalmazása mellett 94,4% feletti marad. A görbéket csak a 0 és 0,5 közötti prevalenciatartományban ábráztuk, ugyanis a paramétertartomány másik felében (0,5 feletti



prevalenciaértékek) az ábra tükörképe a felcserélt szenzitivitás- és specificitásértékek esetén kapott görbének.

A PLCI-intervallum várható hossza átlagosan kisebb, mint a Lang–Reiczigel-féle intervallumé, amennyiben a valódi prevalencia 0,2 alatti, vagy 0,8 feletti. A 0,3 és 0,7 közötti prevalenciaértékek esetén viszont nagyobb (lásd 2. táblázat). Érdekes módon az új intervallum hosszabbnak bizonyult azokban, a gyakorlatban meglehetősen ritkán előforduló esetekben is, amikor a szenzitivitás és a specificitás becsléséhez használt mintanagyság 3000 volt. A részletes szimulációs eredmények elérhetők az [online mellékletben](#).

2. táblázat

**A 95%-os kétoldali PLCI- és Lang–Reiczigel-féle konfidenciaintervallumok várható hossza közötti relatív különbség**  
*Relative difference between the expected length of 95% two-sided PLCI and Lang–Reiczigel CI*

Megnevezés	Valódi prevalencia					
	≤0,02	0,03	0,05	0,1	0,2	0,3–0,5
Relatív hosszkülönbség az összes esetre	-7,8	-5,5	-4,0	-2,3	-0,0	1,3
ha $Se$ és $Sp \geq 0,9$	-9,2	-6,2	-4,4	-2,6	-0,1	1,0
ha $Se$ és $Sp \geq 0,95$	-11,0	-7,4	-5,1	-3,2	-0,2	1,1

\* A 95%-os kétoldali PLCI- és Lang–Reiczigel-féle konfidenciaintervallumok várható hossza közötti relatív különbség a valódi prevalenciától függően (az eredmények a prevalencia 0,5 feletti értékeire a szimmetria miatt ezekkel egyezők). A negatív értékek azt jelentik, hogy a PLCI rövidebb (a Lang–Reiczigel-féle konfidenciaintervallum hosszát tekintettük 100%-nak).

### 3. A PLCI- és a Flor-féle intervallumok összehasonlítása

A PLCI- és a Flor-féle intervallumok összehasonlítása ugyanazzal az 1000 véletlen  $n_{Se}$ ,  $n_{Sp}$ ,  $n$ ,  $Se$ ,  $Sp$ ,  $p$  paraméterkombináció segítségével történt, amelyet Flor is használt az intervalluma értékelésére. A Flor-féle intervallum lefedési valószínűsége alacsonyabb, mint a másik két vizsgált intervallumé. A vizsgált adathalmazon a 95%-os nominális szintű intervallum lefedésének minimuma 88,4% (a PLCI-nél ez 92,5, a Lang–Reiczigel-félénél 92,6%), és 93% alá a vizsgált 1000 eset 7,7%-a esetében csökken (a PLCI-nél ez 0,3, a Lang–Reiczigel-félénél pedig 0,1%-ban fordul csak elő).

A PLCI- és a Flor-féle intervallumok hossza nagyon kevésbé különbözik. A vizsgált 1000 paraméterkombinációból 512 alkalommal a Flor-, 488 alkalommal a PLCI-intervallum bizonyult rövidebbnek. A Flor-intervallum átlagosan 1%-kal rövidebb, de ez az előny magasabb szenzitivitás- és specificitásértékek esetén kisebb volt. Azokban az esetekben, amikor a szenzitivitás és a specificitás is legalább 75%-os, a Flor-intervallum átlagosan még mindig 0,6%-kal rövidebb, mint a PLCI, amellet, hogy az ilyen vizsgált esetek 54,8%-ában a PLCI a rövidebb.

#### 4. Epidemiológiai alkalmazás

Annak érdekében, hogy az intervallumok teljesítőképességét valós adatokon is szemléltessük, kiszámoltuk őket a Lang és Reiczigel 2014-es tanulmányában található első három példára.

##### *Első példa*

*Boelaert és munkatársai (2000)* a szarvasmarha-herpeszvírus 1-es típusának előfordulását vizsgálták szarvasmarhákban. A be nem oltott állományokban 11 284 állatból 4060 fertőzöttet azonosítottak. Felhasználva, hogy a vizsgálatban alkalmazott ELISA teszt szenzitivitása 0,99 (178 a 179-ből – *Kramps et al., 1994*) és specificitása 0,997 (358 a 359-ből – *De Wit et al., 1998*), a valódi prevalencia Rogan–Gladen-féle becslése 0,362. Mivel ebben a példában mindhárom paraméter becsléséhez használt minta meglehetősen nagy, mindhárom módszerrel hasonló intervallumot kapunk (PLCI: 0,349–0,371; Lang–Reiczigel: 0,349–0,372; Flor: 0,348–0,372).

##### *Második példa*

*Anderson és munkatársai (2009)* a trichomonas protozoonok előfordulását vizsgálták szabadon élő énekesmadarakban. Ezek közül mi a *Trichomonas gallinae* házipintyekben (*Carpodacus mexicanus*) való előfordulásának adatait használtuk. A 2971 madárból 51-ben volt kimutatható a parazita, ami 0,017-es látszólagos prevalenciát eredményezett. A használt diagnosztikai teszt érzékenysége és specificitása 0,97 (32/33) és 1 (20/20) volt. Ezen értékek felhasználásával a valódi prevalenciára vonatkozó pontbecslés 0,018, a 95%-os Lang–Reiczigel-féle konfidenciaintervallum pedig 0–0,053 volt. Ebben a példában mind a PLCI-

mind a Flor-módszerrel szerkesztett intervallumok ennél jelentősen rövidebbek (PLCI: 0–0,023 és Flor: 0–0,019).

### **Harmadik példa**

*Bernard és munkatársai* a tuberkulózis (tbc) előfordulását vizsgálták Ugandában a tejelő szarvasmarhák körében (*Bernard et al., 2005*). 11 862 állatot teszteltek tbc-re, a látszólagos prevalencia 0,06 volt. A használt diagnosztikai teszt szenzitivitása és specifitása *Quirin és munkatársai (2001)* szerint 0,80 (8/10) és 1 (12/12) volt. Ezen adatok felhasználásával a 95%-os Lang–Reiczigel-féle konfidenciaintervallum 0–0,147, a valódi prevalencia Rogen–Gladen-féle pontbecslése pedig 0,075. A PLCI- és a Flor-intervallumok itt is rövidebbek, nevezetesen 0–0,116 és 0–0,090 közöttiek.

## **5. Következtetések**

Az általunk javasolt új intervallum lefedési valószínűsége hasonló a Lang–Reiczigel-féle intervalluméhoz, miközben a várható hossza rövidebb, kivéve, ha a valódi prevalencia közel van 0,5-hez. A Flor-intervallum hosszát tekintve hasonló a PLCI-hez, viszont lefedési valószínűsége jelentősen kisebb.

## **Irodalom**

- Agresti, A. – Coull, B. A. (1998): Approximate is Better than “Exact” for Interval Estimation of Binomial Proportions. *The American Statistician*. Vol. 52. pp. 119–126.  
<https://doi.org/10.1080/00031305.1998.10480550>
- Anderson, N. L. – Grahn, R. A. – Van Hoosear, K. – BonDurant, R. H. (2009): Studies of trichomonad protozoa in free ranging songbirds: prevalence of *Trichomonas gallinae* in house finches (*Carpodacus mexicanus*) and corvids and a novel trichomonad in mockingbirds (*Mimus polyglottos*). *Veterinary Parasitology*. Vol. 161. pp. 178–186.
- Bailey, S. – Handu, D. (2013): *Introduction to epidemiologic research methods in public health practice*. Jones & Bartlett Publishers.
- Bernard, F. – Vincent, C. – Matthieu, L. – David, R. – James, D. (2005): Tuberculosis and brucellosis prevalence survey on dairy cattle in Mbarara milk basin (Uganda). *Preventive Veterinary Medicine*. Vol. 67. pp. 267–281.
- Boelaert, F. – Biront, P. – Soumare, B. – Dispas, M. – Vanopdenbosch, E. – Vermeersch, J. – Raskin, A. – Dufey, J. – Berkvens, D. – Kerkhofs, P. (2000): Prevalence of bovine herpesvirus-1 in the Belgian cattle population. *Preventive Veterinary Medicine*. Vol. 45. pp. 285–295.

- De Wit, J. – Hage, J. – Brinkhof, J. – Westenbrink, F. (1998): A comparative study of serological tests for use in the bovine herpesvirus 1 eradication programme in The Netherlands. *Veterinary Microbiology*. Vol. 61. pp. 153–163.
- Farnham, P. G. – Sansom, S. L. – Hutchinson, A. B. (2012): How much should we pay for a new HIV diagnosis? A mathematical model of HIV screening in US clinical settings. *Medical Decision Making*. Vol. 32. pp. 459–469.
- Flor, M. – Weiß, M. – Selhorst, T. – Müller-Graf, C. – Greiner, M. (2020): Comparison of Bayesian and frequentist methods for prevalence estimation under misclassification. *BMC Public Health*. Vol. 20. pp. 1–10.
- Gonçalves, L. – Oliveira, M. R. – Pascoal, C. – Pires, A. (2012): Sample size for estimating a binomial proportion: Comparison of different methods. *Journal of Applied Statistics*. Vol. 39. pp. 1–21. <https://doi.org/10.1080/02664763.2012.713919>
- Greiner, M. – Gardner, I. A. (2000): Application of diagnostic tests in Veterinary Epidemiologic studies. *Preventive Veterinary Medicine*. Vol. 45. pp. 43–59.
- Jegerlehner, S. – Suter-Riniker, F. – Jent, P. – Bittel, P. – Nagler, M. (2021): Diagnostic accuracy of a SARS-CoV-2 rapid antigen test in real-life clinical settings. *International Journal of Infectious Diseases*. Vol. 109. pp. 118–122.
- Kramps, J. – Magdalena, J. – Quak, J. – Weerdmeester, K. – Kaashoek, M. – Maris-Veldhuis, M. – Rijsewijk, F. – Keil, G. – Van Oirschot, J. (1994): A simple, specific, and highly sensitive blocking enzyme-linked immunosorbent assay for detection of antibodies to bovine herpesvirus 1. *Journal of Clinical Microbiology*. Vol. 32. pp. 2175–2181.
- Lang, Z. – Reiczigel, J. (2014): Confidence limits for prevalence of disease adjusted for estimated sensitivity and specificity. *Preventive Veterinary Medicine*. Vol. 113. pp. 13–22. <https://doi.org/10.1016/j.prevetmed.2013.09.015>
- Millar, R. B. (2011): Maximum likelihood estimation and inference: with examples in R. *SAS and ADMB*. John Wiley & Sons.
- Quirin, R. – Rasolofo, V. – Andriambololona, R. – Ramboasolo, A. – Rasolonavalona, T. – Raharisolo, C. – Rakotoaritahina, H. – Chanteau, S. – Boisier, P. (2001): Validity of intradermal tuberculin testing for the screening of bovine tuberculosis in Madagascar. *Onderstepoort Journal of Veterinary Research*. Vol. 68. No. 3. pp. 231–238.
- Reiczigel J. – Földi J. – Ózsvári L. (2010): Exact confidence limits for prevalence of a disease with an imperfect diagnostic test. *Epidemiology and Infection*. Vol. 138. pp. 1674–1678. <https://doi.org/10.1017/S0950268810000385>
- Rogan, W. – Gladen, B. (1978): Estimating prevalence from results of a Screening-test. *American Journal of Epidemiology*. Vol. 107. pp. 71–76. <https://doi.org/10.1093/oxfordjournals.aje.a112510>
- Vollset, S. (1993): Confidence intervals for a binomial proportion. *Statistics in Medicine*. Vol. 12. No. 9. pp. 809–824. <https://doi.org/10.1002/sim.4780120902>