



# BAYES-I MODELLEZÉS A GYAKORLATBAN

## TEJELŐ TEHÉNÁLLOMÁNYOK ÁLLOMÁNYON BELÜLI PARATUBERKULÓZIS-ÉRINTETTSÉGÉNEK BECSLÉSE I.

A közlemény a Magyar Állatorvosok Lapja 2024. júniusi számában megjelent cikk másodközlése

Veres Katalin<sup>1\*</sup>  
Lang Zsolt<sup>1</sup>  
Monostori Attila<sup>2</sup>  
Ózsvári László<sup>3</sup>

<sup>1</sup> ÁTE, Gazdaságtudományi és Biostatistikai Intézet, Biostatistika Tanszék

<sup>2</sup> Állattenyésztési Teljesítményvizsgáló Kft.

<sup>3</sup> ÁTE, Gazdaságtudományi és Biostatistikai Intézet, Törvényszéki Állatorvostani és Gazdaságtudományi Tanszék

### ÖSSZEFOGLALÁS

A Bayes-i módszertan széles körben elterjedt fertőző betegségek prevalenciájának modellezésére az állatorvosi irodalomban, hiszen ezzel a megközelítéssel a korábbi ismeretek és az új adatok egyszerre építhetők be a becslésekbe. A szerzők áttekintik a Bayes-i modellezés alapvető fogalmait és működési elveit, továbbá bemutatják, hogyan alkalmazható ez a módszer a gyakorlatban a

paratuberkulózis valódi prevalenciájának telepi szintű becslésére tejelő szarvasmarhák esetében (<https://github.com/VeresKatalin/PTBC>). A modell eredményei alapján egyszer ellett tehenek esetében a valódi prevalencia a látszólagos prevalencia 1,6-szorosaként, többször ellett teheneknél pedig 1,5-szöröseként becsülhető.

### Bayes-i modellezés

A Bayes-i módszertan széles körben elterjedt fertőző betegségek prevalenciájának modellezésére az állatorvosi irodalomban. A HANSON és mtsai által javasolt keretrendszer felhasználásával számos, e módszertant alkalmazó tanulmány született az utóbbi évtizedben. A módszer gyors terjedésének fő oka az, hogy a Bayes-i megközelítés segítségével a korábbi ismeretek és az új adatok egyszerre építhetők be a becslésekbe. Fertőző betegségek esetén pl. az országos prevalencia vagy a korábbi

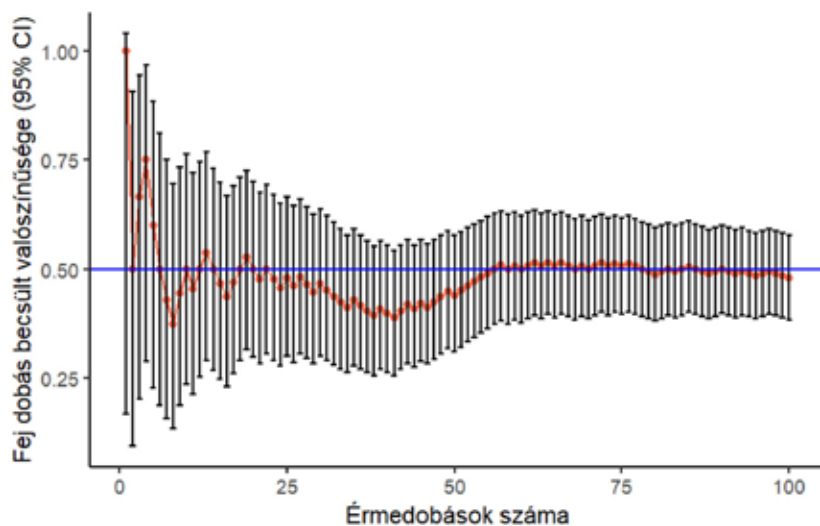
évek felméréseiből származó adatok egyaránt olyan hasznos információt hordoznak, amelyeket hiba lenne figyelmen kívül hagyni. Sokan mégis félve közelítenek ehhez a módszertanhoz, hiszen hagyományosan az ún. frekventista ( $p$ -értékekre támaszkodó) statisztikát alkalmazzuk. A közleményünk célja ezért a Bayes-i módszertan legfontosabb fogalmainak, működésének áttekintése és gyakorlati alkalmazásának illusztrálása tejelő szarvasmarha állományok paratuberkulózis- (PTBC) érintettség becslésének bemutatásával.



A Bayes-i statisztika valószínűségfelfogása alapjaiban különbözik a hagyományosan alkalmazott frekventista statisztikától. A frekventista statisztika egy esemény valószínűségét sok kísérlet alapján kiszámított előfordulási részarány, azaz *relatív gyakoriság* alapján határozza meg. Minél többször végezzük el a

kísérletet, annál jobban megközelíti az esemény relatív gyakorisága az elméleti valószínűséget. Pl. minél többször dobunk fel egy szabályos pénzérmét, annál jobban megközelíti a fejek (és az írások) részaránya a várt 50%-ot (**1. ábra**).

1. ábra: Érmédobás szabályos érmével



Piros: fej dobások aránya, fekete: 95%-os konfidenciaintervallum

E keretrendszerben a valószínűség objektív mennyiség, független minden korábbi tapasztalattól. Több egymás utáni fej dobása után továbbra is azt feltételezzük, hogy a következő dobásnál egyaránt 50–50% a fej, ill. az írás dobásának esélye. Egy modell paramétereinek frekventista becslésénél ezt a rögzített valószínűséget szeretnénk meghatározni.

A Bayes-i módszertanban a valószínűség megadásakor az adatok mellett az előzetes tapasztalatokat és ismereteket is figyelembe vesszük. Ez a módszer közel áll hétköznapi gondolkodásunkhoz. Az érmefeldobás példájánál maradva, ha a sokadik dobásra is fejet kapunk, gyanakodni kezdünk, hogy az érme nem szabályos, tapasztalataink alapján „újrászámoljuk” magunkban az esélyeket, és a következő dobásoknál nagyobb eséllyel várunk fejet.

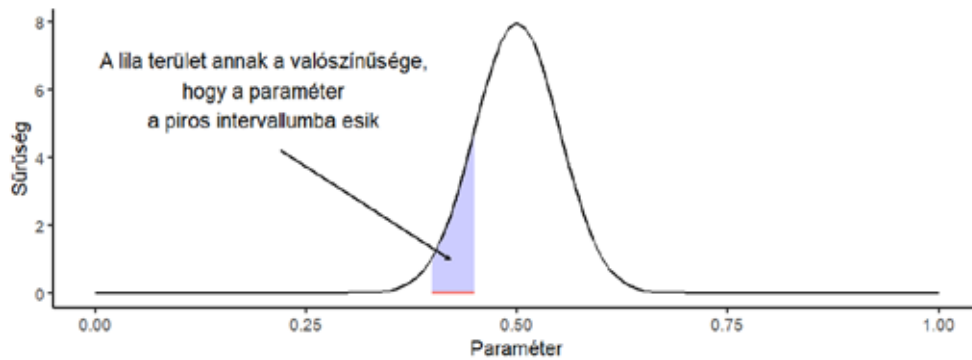
A Bayes-i modellezéshez három alapvető összetevőre van szükségünk: a modellre, az adatokra és az ún. *prior* ismeretekre. A prior információ sűríti magába a folyamattal kapcsolatos előzetes várakozásokat és ismeretanyagot. A prior három fő típusát különböztetjük meg. Ha nem áll rendelkezésünkre olyan információ, amelyet hasznos lenne beépíteni a modellbe, pl. egy új, eddig ismeretlen tudományterület feltérképezésekor, akkor ún. *nem informatív priort* választunk. Ez a prior nem hordoz információt előzetes várakozásokról, egyedül a rendelkezésre álló adatok alapján becsüljük meg a modell paramétereit. Abban az esetben is választhatunk nem informatív priort, ha

nagy mennyiségű adattal rendelkezünk, amelyből további feltételezések nélkül is hatékony becslés várható.

A második lehetőség az ún. *gyengén informatív prior* használata. Ennek szerepe csupán a paraméterek tartományának korlátozása, ill. a modell stabilizálása. Egy diagnosztikai eljárás érzékenységéről legtöbbször feltehető, hogy 50%-nál nem kisebb. Gyógyíthatatlan betegség modellezésekor kizárhatjuk azokat a paraméterértékeket, amelyek megengednék a gyógyulást. Az adatok heterogenitását jellemző varianciaparaméterek általában nem lehetnek nulla közelében és túl nagyok se, különben instabil és értékelhetetlen lehet az illesztett Bayes-i modell. A prior eloszlások harmadik típusa az *informatív prior*, mely lehetővé teszi, hogy megbízható előzetes ismereteket építsünk be a modellbe. Ilyenkor szakértők bevonásával, korábbi tanulmányok eredményei, tapasztalataink, ismereteink, várakozásaink alapján megadjuk, hogy a paraméterek nagy valószínűséggel milyen értéktartományba eshetnek. A paraméterek ún. *prior eloszlását* úgy szerkesztjük meg, hogy annak valószínűsége, hogy a paraméter egy adott intervallumba esik, a prior eloszlás adott intervallum fölé eső görbe alatti területével egyezzen meg. A függvénygörbe értékeit *sűrűségnek* nevezzük; nagy sűrűség a paraméterérték nagy valószínűségét jelzi (**2. ábra**).



## 2. ábra: Prior és poszterior eloszlás értelmezése

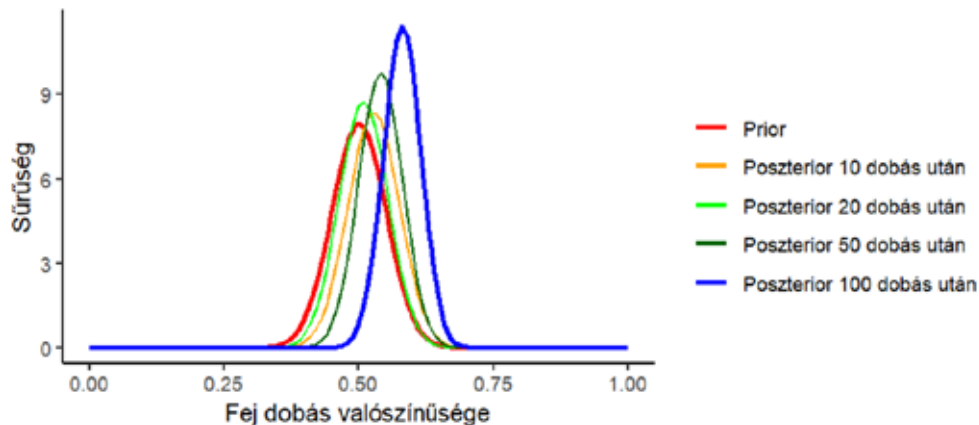


A prior eloszlást a Bayes-i modellben az adatok segítségével pontosítjuk, aktualizáljuk, így jutunk el az ún. *poszterior eloszláshoz*.

Érmefeldobás esetében, ha az az előfeltételezésünk, hogy a pénzérme szabályos, de persze előfordulhat-

nak benne apró gyártási hibák, használati kopások, válasszunk priornak egy olyan eloszlást (értéktartományt), mely szerint a fej-írás esélye kb. 50–50%, ettől el is térhet némiképp, de a nagyobb eltérés esélye kicsi (**3. ábra**).

## 3. ábra: Fej dobás valószínűségének Bayes-i becslése érmedobás során – szimuláció, cinkelt érme, fej valószínűsége: 0,6 (60%)



### Prior feltételezés: szabályos érme

A keresett paramétert a prior eloszlásból kiindulva, az adatokban rejlő többletinformáció felhasználásával becsüljük meg. Becslésünk nem egy konkrét szám, hanem az előbbiekkal összhangban egy ún. *poszterior eloszlás* lesz. Annak a valószínűsége, hogy a valódi paraméter egy adott intervallumba esik, a poszterior eloszlás adott intervallum fölé eső görbe alatti területének kiszámításával adható meg (**2. ábra**).

Érmedobás esetében pl. azt mondja meg a poszterior eloszlás, hogy milyen eséllyel szabályos az érme, ill. milyen eséllyel tér el ettől egyik vagy másik irányban adott mértékkel.

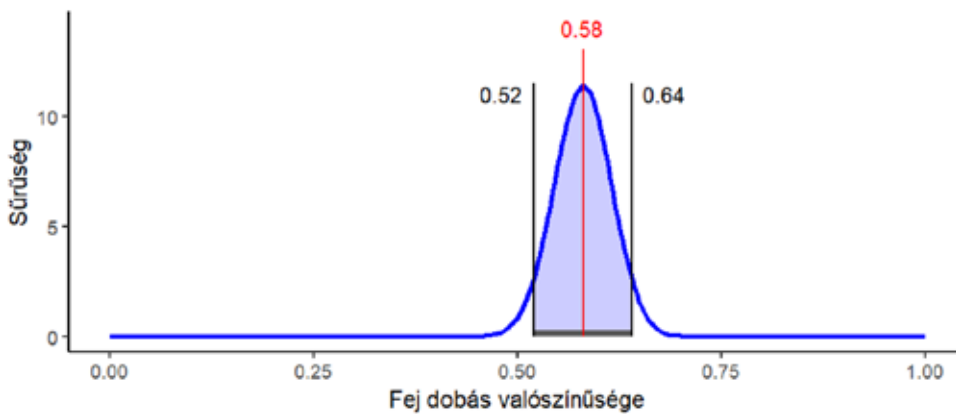
A **3. ábrán** láthatjuk, hogy az előfeltételezéseinket reprezentáló prior görbéből kiindulva a tapasztalatok felhalmozódásával (dobásszám növelése) egyre keskenyebb lesz a poszterior eloszlásunk, azaz nő a becslésünkbe vetett bizalom. Ezzel párhuzamosan a poszterior görbe a valódi paraméterérték felé tolódik el, azaz a valódi paraméter környékén található értékeknek lesz a legnagyobb a valószínűsége a



poszterior eloszlás alapján. A paraméterek poszterior eloszlását az adatok és a prior felhasználásával számítjuk ki. A könnyebb értelmezhetőség kedvéért a poszterior eloszlást egyetlen számba is sűrítjük, kiszámítva pl. az átlagos értékét vagy mediánját és az ehhez tartozó adott lefedettségű *kredibilis intervallumot* ( $K_{r1}$ ). A *kredibilis intervallum lefedettsége* annak a valószínűsége, hogy a paraméter az adott intervallumba esik (**4. ábra**).



4. ábra: Poszterior eloszlás, becsült átlag és 95%-os kredibilis intervallum fej dobás valószínűségére 100 érmedobás után



Prior feltételezés: szabályos érme, valóság: cinkelt érme, fej valószínűsége 0.6 (60%), becsült átlag és kredibilis intervallum: 0.58 (95% Kr: 0.52-0.64)

A becslés eredményében megjelenik az adatokban rejlő többletinformáció, ismereteink gyarapodnak. A frissített poszterior ismerethalmaz felhasználható egy következő modell prior eloszlásaként. Ez a munkamódszer jól illeszkedik a tudományos megismerés természetes folyamatába. Előzetes ismereteinkből és az aktuális kutatásban tapasztaltakból születik meg az új elmélet, amelyet az újabb adatok tükrében ismét újra és újra felülvizsgálunk, ezzel finomítva lépésről-lépésre a vizsgált témakörrel alkotott képünket.

Az orvosi diagnosztika gondolkodásmódja is természetes módon Bayes-i szellemű. A diagnosztika fő kérdése, hogy adott tünetegyüttes ismeretében mi a valószínűsége az egyes kiváltó okoknak. Az esélyek latolgatásánál figyelembe vesszük a páciens jellemzőit, adottságait és a környezeti tényezőket is. Attól függően pl., hogy milyen idős, milyen korábbi betegségei voltak, továbbá milyen környezetből származik, más és más lehet egy-egy mögöttes ok valószínűsége. Pl. lázas, hasmenéses, bágyadt és levert, 6 hónapos kölyökutya esetében, amely egyébként rendszeres állategészségügyi ellátásban és az életkornak megfelelő ajánlott vakcinázásban

is részesült, első körben nem parvovírusos fertőzésre gyanakodunk. Azonban ugyanezekkel a tünetekkel egy tenyésztőtől most vásárolt, még semmilyen vakcinázásban nem részesülő, állatorvosnál először járó, 3 hónapos kutyakölyök esetében a parvovírus okozta megbetegedés valószínűségét jóval nagyobbra fogja értékelni a klinikus állatorvos.

Fertőző betegségek populációs részarányának, prevalenciájának becslésében a Bayes-i megközelítés használatát többnyire két fő érv motiválja. Egyrészt szeretnénk beépíteni a korábbi ismereteinket a becslés pontosításához. Ha ismerjük pl. a legutóbbi szűrésnél becsült prevalenciát, azt feltételezhetjük, hogy újra a korábbihoz hasonló értéket kapunk, de akár nagy változás is előfordulhat, kis eséllyel. Ezeket az ismereteket, várakozásokat építjük be priorként a modellbe. Másrészt a diagnosztikai tesztek jellemzőit, a szenzitivitást (érzékenységet) és a specifitást (fajlagosságot) is általában bizonytalanság terheli, hiszen azokat is kísérleti adatokból becsülik meg. Egy Bayes-i modell a diagnosztikai paraméterek bizonytalanságát is képes prior eloszlás formájában beépíteni.





# BAYES-I MODELLEZÉS A GYAKORLATBAN

## TEJELŐ TEHÉNÁLLOMÁNYOK ÁLLOMÁNYON BELÜLI PARATUBERKULÓZIS-ÉRINTETTSÉGÉNEK BECSLÉSE II.

A közlemény a Magyar Állatorvosok Lapja 2024. júniusi számában megjelent cikk másodközlése

Veres Katalin<sup>1</sup>  
Lang Zsolt<sup>1</sup>  
Monostori Attila<sup>2</sup>  
Ózsvári László<sup>3</sup>

<sup>1</sup> ÁTE, Gazdaságtudományi és Biostatistikai Intézet, Biostatistika Tanszék

<sup>2</sup> Állattenyésztési Teljesítményvizsgáló Kft.

<sup>3</sup> ÁTE, Gazdaságtudományi és Biostatistikai Intézet, Törvényszéki Állatorvostani és Gazdaságtudományi Tanszék

### Tejelő tehenek PTBC-érintettségének modellezése

A PTBC világszerte és Magyarországon is széles körben elterjedt szarvasmarha-betegség, amit a *Mycobacterium avium* subsp. *paratuberculosis* (MAP) okoz. Az MAP-fertőzöttség klinikai és szubklinikai stádiumban egyaránt jelentős gazdasági károkat okoz, ám a két stádium közül gazdasági szempontból a szubklinikai a jelentősebb, hiszen ekkor az egyedek szembetűnő tüneteket nem mutatnak, ugyanakkor a termelésre gyakorolt káros hatás már fellép. Az MAP-fertőzöttség, ill. annak különböző stádiumai negatívan befolyásolják a szarvasmarhák vágási súlyát és – a hús minőségének romlásán keresztül is – a vágási értékét. Az MAP-pozitív szarvasmarhák selejtezési és a mortalitási aránya is jelentősen nagyobb a nem fertőzött egyedekhez képest. Ezen túlmenően az MAP-fertőzöttség káros hatással van a tejhozamra és jelentősen emeli a szomatikus sejtszámot. Az MAP-pozitív szarvasmarhák reprodukciós adatai is számottevően rosszabbak az MAP-negatív társaikhoz képest: jelentősen nő a teheneknek az elléstől termékenyülésig eltelt ideje, a két ellés közötti ideje és számottevően romlik a termékenyítési indexe.

Hazai nagylétszámú tehenészetben az MAP-pozitív egyedek termékenyítési indexe 2-vel nőtt az MAP-negatív egyedek mutatójához képest, amely eltérés – a megnövekedett felhasznált spermamennyiség következtében – milliós többletköltséget okozott a telepnek éves szinten.

Az MAP-fertőzöttség szűrésére nem létezik „gold-standard” eljárás, ezért a fertőzés azonosítása és a telepen belüli fertőzött állatok részarányának meghatározása kihívást jelent. A diagnosztikai tesztek által szolgáltatott eredmények alapján kalkulált *látszólagos prevalencia* (AP – apparent prevalence, a pozitív esetek részaránya) és a *valódi prevalencia* (TP – true prevalence, a valóban fertőzött esetek részaránya) között markáns különbségek lehetnek a téves negatív és téves pozitív teszteredmények miatt. Kizárólag a diagnosztikai tesztekre hagyatkozni félrevezető, és hibás telepi menedzsment döntésekhez vezethet. A fertőzöttségi szint pontosabb meghatározásához a látszólagos prevalenciát korrigálni kell a teszt szenzitivitása, specificitása és egyéb telepi adatok figyelembevételével.



Magyarországon az MAP-fertőzés országos elterjedését a tejelő tehenek körében elsőként 2018-ban mérték fel. Az eredményeket 2019-ben a korábbi mérés tapasztalatait beépítve, átfogó, országos adatokon újraértékeltük. Jelen dolgozatban az országos eredmények priorként való felhasználásával egyedi tejelőszarvasmarha-telepek valódi PTBC-prevalencia becslésére szolgáló Bayes-i modellt mutatunk be. A modell eredményeit két, számítógépes eljárással

szimulált teleppel illusztráljuk. A kapott eredményekre építve olyan egyszerű módszert is ajánlunk, amelyek segítségével a fertőzés prevalenciája a modell használata nélkül is könnyen, gyorsan közelíthető. Emellett megvizsgáljuk, hogy hogyan változik a becslés hatékonysága, ha nem szűrjük le az összes tejelő tehenet, csak egy adott százalékukat teszteljük a költségtakarékosság érdekében.

## ANYAG ÉS MÓDSZER

### FELHASZNÁLT ADATOK

Magyarországon 2018 februárjában országos önkéntes PTBC-felmérő program indult. Az Állattenyésztési Teljesítményvizsgáló Kft. (ÁT Kft.) tejből végezte a PTBC ELISA-vizsgálatokat végezte a Paratuberculosis Screening Antibody Test (IDEXX Laboratories Inc., Westbrook, ME, USA) segítségével a gyártó utasításai szerint. A kapott optikai sűrűség (OD) értékeket S/P hányadosá alakítottuk a következők szerint:

$$\frac{S}{P} = \frac{\text{minta OD értéke} - \text{negatív kontroll OD értéke}}{\text{pozitív kontroll OD értéke} - \text{negatív kontroll OD értéke}}$$

A vizsgálat eredményét negatívnak tekintettük azon tejminták esetében, ahol az S/P hányados 0,2-nél kisebb volt, kétesnek, ha 0,2–0,3 közé esett, pozitívnak, ha meghaladta a 0,3-at. A Bayes-i modellezés során csak a 100, vagy annál több tehenet tartó tejelőszarvasmarha-telepek 2019-ben végzett PTBC-vizsgálati eredményeit használtuk fel, így a Bayes-i modell összesen 116 tehenészet 55 594 tehenének egyedi tejmintáján végzett PTBC-vizsgálati eredményein alapul. A kétes teszteredményt adó teheneket ( $n = 573$ ) kizártuk az elemzésből.

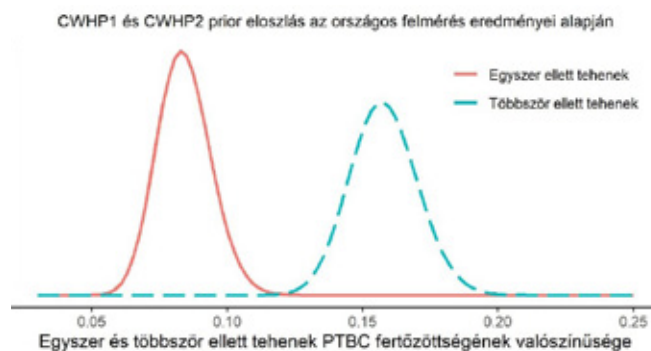
### A STATISZTIKAI MODELL

Célunk egy konkrét tehenészet adatai és az országos prior tapasztalatok alapján a tejelőszarvasmarha-telep valódi fertőzöttségének becslése. Mivel az országos adatok alapján becsült valódi prevalencia 92%, a továbbiakban azt feltételezzük, hogy fertőzött teleppel dolgozunk. Bayes-i kétlépcsős hierarchikus modellt illesztünk egy fertőzött telep adataira az országos felmérés eredményeit prior eloszlásként használva. A modell eredményeit két szimulált telep elemzésével illusztráljuk.

A modellezés során annak a valószínűségét határoztuk meg, hogy egy, a telepről véletlenszerűen kiválasztott egyszerrellett, ill. többszörrellettegyedfertőzött-e ( $CWHP_1$ : conditional within-herd prevalence, prevalencia az egyszer ellett tehenek körében,  $CWHP_2$ : prevalencia a többször ellett tehenek körében). Feltételezzük, hogy az

egyed telepek valódi prevalenciáját jól reprezentálják az országos adatok, ezért az országos felmérés tapasztalatai alapján informatív priori konstruáltunk a  $CWHP_1$  és  $CWHP_2$  paraméterekhez (1. ábra).

**1. ábra:** Prior eloszlások az egyszer, ill. többször ellett tehenek valódi PTBC-prevalenciájának Bayes-i becsléséhez az országos adatokból származó poszterior eloszlás



A priorokat az R 4.1.0 epiR csomagjához tartozó epi.betabuster eljárással határoztuk meg, oly módon, hogy minél jobban illeszkedjenek az országos modellből kapott megfelelő poszterior eloszlásokhoz.



A becslés kiszámításához a prior mellett a telepi ELISA-tesztek eredményét, a paritást és a tehenek életkorát használtuk. A nyers adatokból kiszámítható a pozitív esetek előfordulási aránya (AP), azonban a diagnosztikai tesztek téves pozitív és téves negatív eredményei miatt a valódi prevalencia a gyakorlatban rendre eltér ettől. A tanulmányban használt ELISA-teszt érzékenysége, szenzitivitása ( $Se$ ) - annak valószínűsége, hogy a fertőzött egyedekhez pozitív teszteredmény tartozik - az életkor ismeretében



meghatározható MEYER és mtsai képlete alapján. Az összefüggést a PTBC-szűrt állományra alkalmazva az állatok életkorától függően a tesztek 24–67%-a tévesen negatív eredményt adott. A teszt fajlagossága, specificitása ( $Sp$ ) – annak valószínűsége, hogy a fertőzéstől mentes egyedekhez negatív teszteredmény tartozik – 99,5%. Ezt azt jelenti, hogy csupán az esetek fél százalékában kapunk tévesen pozitív eredményt. Modellünk arra szolgál, hogy a diagnosztikai teszt teljesítmény korlátait figyelembe véve becslést adjon a fertőzés valódi prevalenciájára.

Egy fertőzött telepen egy tehén teszteredménye AP valószínűséggel lesz pozitív. A látszólagos prevalenciát megkaphatjuk a valódi pozitív és a téves pozitív teszteredmények valószínűségeinek összegeként:

$$AP = Se \times CWHP + (1 - Sp) \times (1 - CWHP) \quad (1)$$

A fenti képlet azt írja le, hogy ha a vizsgált paritáscsoportban a fertőzött állatok részaránya  $CWHP$ , akkor a tehenek  $Se \times CWHP$  részére kapunk valódi pozitív teszteredményt, míg a tehenek  $(1 - Sp) \times (1 - CWHP)$  része téves pozitív eredményt produkál.

A  $CHWP_1$  és  $CHWP_2$  poszterior eloszlását a telepi adatokra illesztett Bayes-i modellel számítottuk ki. A  $CHWP_1$ , ill. a  $CHWP_2$  poszterior eloszlás a konkrét telep egyszer ellett ill. többször ellett tehenekhez tartozó PTBC prevalenciájának Bayes-i becslése.

A Bayes-i modellt az R 4.1.3 statisztikai szoftver rstan csomagjának segítségével állítottuk fel és futtattuk (<https://github.com/VeresKatalin/PTBC>).

### KÖZELÍTŐ MÓDSZER A TELEP FERTŐZÖTTSÉGÉNEK BECSLÉSÉRE

A Bayes-i modell mellett egyszerűen használható közelítő módszert is kidolgoztunk a valódi állomány szintű prevalencia becslésére. A modellt lefuttattuk az elemzésünkben használt 116 db, legalább 100 tehenet tartó tejelő szarvasmarha telep egyedi PTBC ELISA tejvizsgálati eredményeire, amelyre lineáris regressziót illesztettünk. A becsült valódi prevalenciát függő változónak, a nyers adatok alapján számított látszólagos prevalenciát magyarázó változónak tekintettük. A regressziós egyenes egyenlete alapján a nulla közeli tengelymetszet miatt a valódi állomány szintű prevalencia egyszerű szorzással becsülhető a látszólagos prevalenciából.

### SZIMULÁCIÓS VIZSGÁLAT

A Bayes-i modellünk pontosságát számítógépes szimulációval vizsgáltuk. Összesen 500 telep adatait szimuláltuk, azaz számítógéppel véletlen értékeket generáltunk, megtartva a valódi telepek adatainak nagyságrendjeit, arányait, PTBC fertőzési mintázatait. A szimulált adatokra lefuttattuk a modellt és ellésszám szerint megvizsgáltuk, hogy az esetek mekkora hányadában esik a valódi prevalencia a modell által szolgáltatott 95%-os kredibilis intervallumba. Ezt követően megnéztük, hogy hogyan változik a becslés pontossága, ha nem tesztlünk le minden tehenet, csupán az egyszer ellett, ill. többször ellett tehenek egy adott hányadát.



A szimulációt a következőképpen végeztük. Elsőként a telepi alapadatokat generáltuk. A teleplétszámot 450 átlagú, 500 szórású negatív binominális eloszlásból kaptuk meg. Ehhez hozzáadtunk 50-et, így a minimális létszám 50 tehén és az átlagos létszám 500 tehén lett. Az országos jellemzőket figyelembe véve, minden tehén esetében 40%-os eséllyel választottunk „először ellett” státuszt. A tehenek életkorát gamma eloszlás szerint adtuk meg, a paritáscsoportokhoz tartozó hozzáadott minimum életkort, az átlagos életkort, ill. a szórást az országos adatok alapján állítottuk be.

Ezt követően minden egyes virtuális tehenhez egy diagnosztikai teszteredményt generáltunk a következő módon. Az ellésszám szerinti csoportokhoz (egyszer ellett vs. többször ellett) tartozó valódi prevalenciák között pozitív a korreláció, nem függetlenek egymástól (olyan telepen, ahol az egyszer ellett állatok az átlagnál fertőzöttebbek, általában a többször ellett állatok is fertőzöttebbek az átlagnál), így ezeket az országos adatokra épülő tanulmányban kiszámított együttes poszterior eloszlás szerint, a korrelációt figyelembe véve határoztuk meg. A  $CWHP$  és az adott életkorhoz tartozó  $Se$  és az  $Sp$  ismeretében az (1) képlet alapján kiszámítottuk az  $AP$ -t. A tehenek teszteredményét  $AP$  valószínűséggel generáltuk pozitívnak.





# BAYES-I MODELLEZÉS A GYAKORLATBAN

## TEJELŐ TEHÉNÁLLOMÁNYOK ÁLLOMÁNYON BELÜLI PARATUBERKULÓZIS-ÉRINTETTSÉGÉNEK BECSLÉSE III.

A közlemény a Magyar Állatorvosok Lapja 2024. júniusi számában megjelent cikk másodközlése

Veres Katalin<sup>1\*</sup>  
Lang Zsolt<sup>1</sup>  
Monostori Attila<sup>2</sup>  
Ózsvári László<sup>3</sup>

<sup>1</sup> ÁTE, Gazdaságtudományi és Biostatistikai Intézet, Biostatistika Tanszék

<sup>2</sup> Állattenyésztési Teljesítményvizsgáló Kft.

<sup>3</sup> ÁTE, Gazdaságtudományi és Biostatistikai Intézet, Törvényszéki Állatorvostani és Gazdaságtudományi Tanszék

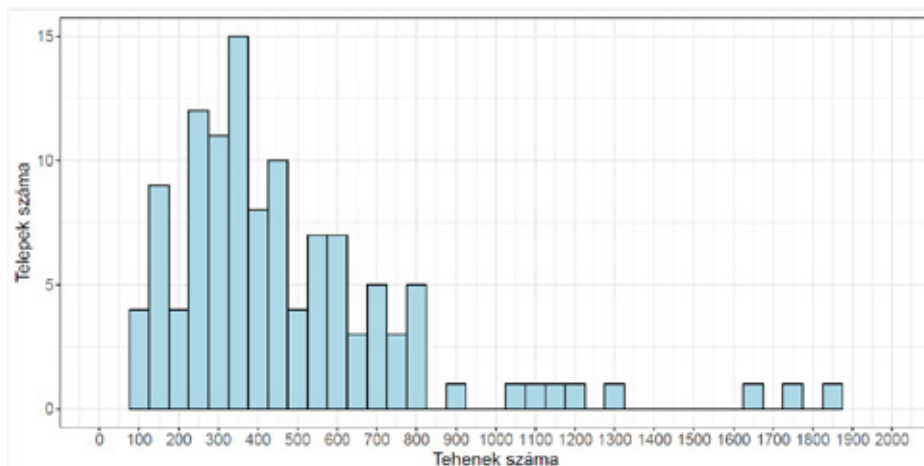
### EREDMÉNYEK

#### LEÍRÓ STATISZTIKA

A tanulmányban priorként felhasznált információ 116 magyarországi tejelő tehenészet 55 594 tehenének tej alapú ELISA teszt eredményéből származik. A medián

állományméret 479 tehen (interkvartilis tartomány: 274–602). Az állomány méretének megoszlását a **1. ábra** mutatja.

**1. ábra:** Állományméret megoszlása



A látszólagos egyedi PTBC prevalencia 6,01% és az állományok 96,6%-ában található pozitív teszteredményű állat. A látszólagos prevalencia a

telepek 92,2%, 78,4%, 53,4%, ill. 19%-ában volt 1%, 2%, 5%, ill. 10%-nál nagyobb (**1. táblázat**).





**1. táblázat:** Tejló tehén ELISA-teszt eredményei és látszólagos paratuberkulózis prevalencia (AP) Magyarországon 2019-ben

Telepméret	Egyszer ellett tehenek			Többször ellett tehenek			Összesen		
	N	+	AP (%)	N	+	AP (%)	N	+	AP (%)
100–300 tehén	2 590	138	5,3	4 073	444	10,9	6 663	582	8,7
301–600 tehén	9 361	271	2,9	13 472	1 081	8	22 833	1 352	5,9
>600 tehén	11 457	357	3,1	14 641	1 052	7,2	26 098	1 409	5,4
Overall	23 408	766	3,3	32 186	2 577	8	55 594	3 343	6

ELISA: Paratuberculosis Screening Antibody Test (IDEXX Laboratories Inc., Westbrook, ME, USA), AP: látszólagos prevalencia, N: tesztek száma, +: pozitív tesztek száma

## A MODELL EREDMÉNYEI

A modell eredményeit két szimulált telep elemzésével illusztráljuk. Az első telepen a fertőzés valódi szintje kicsi, a második telepen nagy (**2-3. táblázat**).

**2. táblázat:** Szimulált telepi adatok – 1. telep

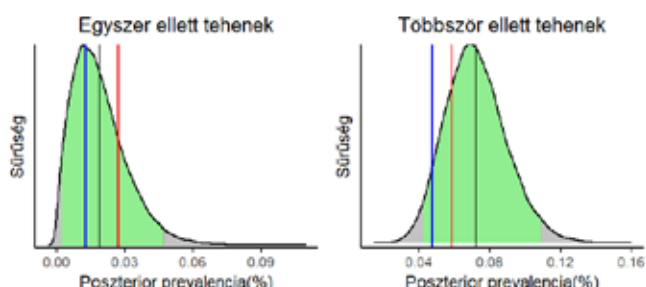
1. telep	Létszám	Résarány (%)	Látszólagos prevalencia (%)	Valódi prevalencia (%)
Egyszer ellett	310	44	1,3	2,7
Többször ellett	400	56	4,8	5,8
Összesen	710	100	3,3	4,4

**3. táblázat:** Szimulált telepi adatok – 2. telep

2. telep	Létszám	Résarány (%)	Látszólagos prevalencia (%)	Valódi prevalencia (%)
Egyszer ellett	307	43	8,5	16,9
Többször ellett	475	67	19,6	28,3
Összesen	782	110	15,2	23,8

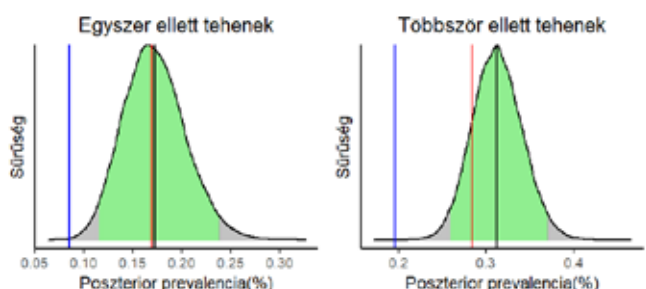
A két telepre kapott poszterior becslést a **2. és 3. ábrán** láthatjuk.

**2. ábra:** Poszterior prevalencia becslés az 1. telepre



Piros függőleges: valódi prevalencia, fekete függőleges: poszterior átlag, kék függőleges: látszólagos prevalencia, fekete görbe: poszterior eloszlás, zöld értéktartomány: 95% kredibilis intervallum a valódi prevalenciára.

**3. ábra:** Poszterior prevalencia becslés a 2. telepre



Piros függőleges: valódi prevalencia, fekete függőleges: poszterior átlag, kék függőleges: látszólagos prevalencia, fekete görbe: poszterior eloszlás, zöld értéktartomány: 95% kredibilis intervallum a valódi prevalenciára.

A **2. és 3. ábrán** a Bayes-i modell becslését, a poszterior eloszlást láthatjuk. A vízszintes tengelyen

a betegség prevalenciája, a függőleges tengelyen az adott paraméterértékhez tartozó sűrűség látható. Annak a valószínűsége, hogy a valódi prevalencia egy adott intervallumba esik, megegyezik az adott intervallumban a görbe alatti területtel. A fekete függőleges vonal jelzi a poszterior eloszlás átlagát, a zöld értéktartomány a hozzá tartozó 95%-os kredibilis intervallum. Az ábrán késsel jelöltük a látszólagos, pirossal a valódi prevalenciát.

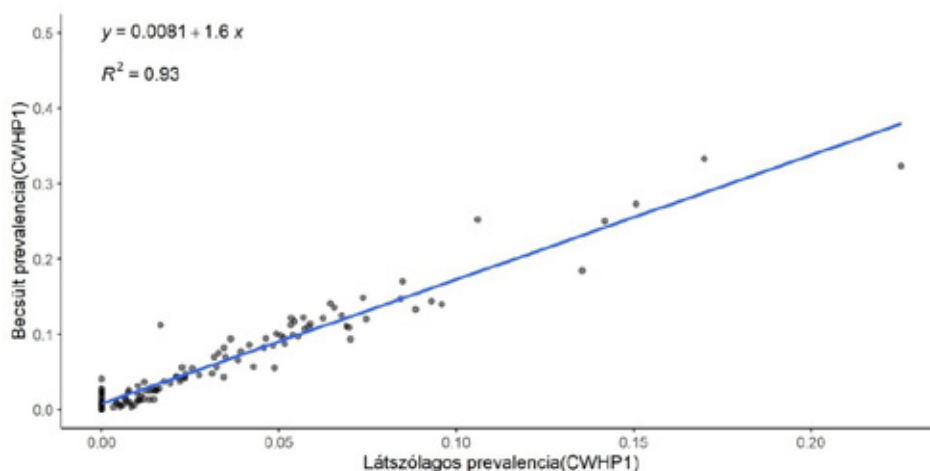


## KÖZELÍTŐ MÓDSZER EGY ÁLLOMÁNY FERTŐZÖTTSÉGÉNEK BECSLÉSÉRE

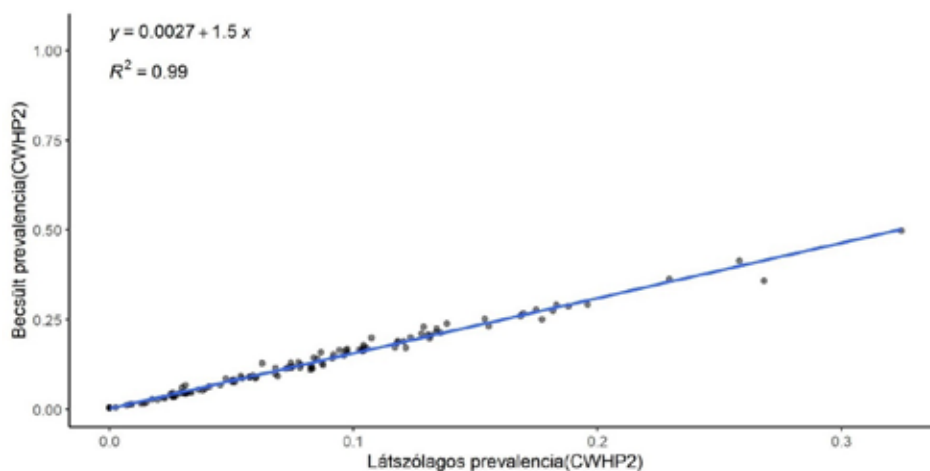
A valódi telepek vizsgálatához a modellt lefuttattuk egyenként az összes tesztelt telepre és az eredményekre lineáris regressziót illesztettünk. A becslött valódi prevalenciát függő változónak, a nyers adatok alapján számított látszólagos prevalenciát magyarázó változónak tekintettük. A látszólagos (tesztek által mért) és a valódi prevalencia összefüggéseit a 4. és 5. ábra szemlélteti:



4. ábra: Lineáris regresszió az egyszer ellett tehének látszólagos és becsült valódi PTBC prevalenciájára a Bayes-i modell eredményei alapján



5. ábra: Lineáris regresszió a többször ellett tehének látszólagos és becsült valódi PTBC prevalenciájára a Bayes-i modell eredményei alapján



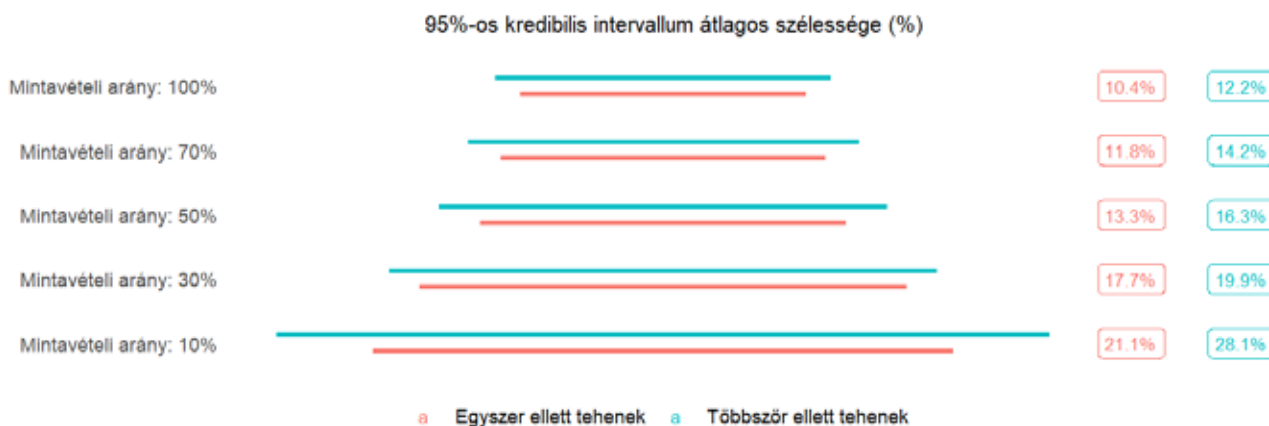
A 4. és 5. ábrákról leolvasható a regressziós egyenes egyenlete. A tengelymetszet mindkét esetben nullához közeli, ezért a látszólagos prevalenciából egyszerű szorzással megkapható a valódi prevalencia becslése. Egyszer ellett tehének esetében a valódi prevalencia becsülhető a látszólagos prevalencia 1,6-szorosával, többször ellett tehéneknél pedig 1,5-szörösével.

### SZIMULÁCIÓS VIZSGÁLAT

Az 500 szimulált telepen lefuttatott modell eredmények alapján az először ellett tehének esetében a 95%-os

kredibilis intervallum a telepek 90%-ára, míg többször ellettek esetében a telepek 87%-ára tartalmazta a valódi prevalenciát. A mintavételi százalék csökkentésének szimulációs vizsgálata során azt tapasztaltuk, hogy a csökkentés nem volt hatással arra, hogy a kredibilis intervallum az esetek hány százalékában tartalmazza a valódi paramétert. Ez a hányad mind az egyszer, mind a többször ellett tehének esetében 90% körül maradt. Azonban a becslés pontossága romlik, a kredibilis intervallum szélessége a 6. ábrának megfelelően jelentősen megnő.

6. ábra: A Bayes-i modell becsléseinek megbízhatósága a mintavételi arány változásával



A jobb oldalon feltüntetett százalékok a 95%-os kredibilis intervallum hosszát mutatják.



## MEGVITATÁS

A különböző országok szarvasmarha-állományainak MAP-fertőzöttségét vizsgáló korábbi tanulmányok CWHP-becslései nem mutatnak egységes képet. Az éghajlat, a telepi menedzsment, az állomány mérete, az egyes fajták genetikai fogékonysága és a vizsgált populáció kormegoszlása mind befolyásolják a fertőzöttség mértékét. A becslések közötti eltérést tovább növelheti a modellezni kívánt állapot (fertőzött, fertőző vagy beteg), a mintavételi stratégiák, a biológiai minta (szövet, szérum, tej vagy környezeti ürülék), és a diagnosztikai teszt (ELISA, AGID, CFT, PCR) megválasztása is.

NIELSEN és TOFT átfogó tanulmánya szerint az állományokon belül a pozitívat tesztelő tehenek részaránya 2,7–15% között mozog. Az általunk prior információként használt CWHP-becslések összhangban vannak ezekkel az eredményekkel. A prior információkat szolgáltató korábbi tanulmányunkban a CWHP átlagát 8,4%-ra (6,6%; 10,4%) becsültük az egyszer ellett és 15,8%-ra (13,5%; 18,4%) a többször ellett tehenek esetében. A medián CWHP 4,7% (3,2; 6,4%) volt az egyszer ellett és 12,4% (10%; 15%) a többször ellett tehenek esetében.

Mind az állatok MAP-fertőzésre való fogékonysága, mind pedig a fertőzés kimutathatósága életkorfüggő. A modellben al csoportként használt ellésszám az állomány természetes kategorizálása, és erősen összefügg az állatok életkorával. Az állomány

ellésszám szerinti bontása olyan többé-kevésbé homogén csoportokat eredményez, amelyekben jelentősen különböző fertőzési prevalenciát találunk. A pozitív ELISA-teszteredmény valószínűsége az egyszer ellett tehenek esetében két-háromszor kisebb, mint a többször ellett tehenek esetében.

A Bayes-i modellünk látszólagos prevalenciát meghatározó képlete alapján a látszólagos és valódi prevalencia aránya az 1-et majdnem elérő specificitás esetén elsősorban a teszt szenzitivitásától függ. A látszólagos és valódi prevalencia arányának vizsgálatához ezen felül számításba kell vennünk a szenzitivitás Meyer-képlet által leírt életkorfüggését. A prior adatokat szolgáltató tanulmányban az egyszer ellett tehenek túlnyomó többsége 2–3 éves, az életkor szórása kicsi, ezért a szenzitivitás ebben a kategóriában viszonylag stabil. Többször ellett tehenek esetében ugyan nagyobb az életkor szerinti szórás, viszont nagyobb életkor mellett a szenzitivitás a Meyer-képlet alapján kevésbé függ az életkor változásától, heterogenitásától. A fentiekben leírtak szerint a valódi és a látszólagos prevalencia összefüggése lineáris és stabil, ezért jól becsülhető, a lineáris regresszió illesztése elméletileg megalapozott. Az általunk talált 1,5, ill. 1,6-os TP/AP arány nagyon hasonló VERDUGO és mtsai dániai kutatásában publikált egymást követő három évből, 2011, 2012 és 2013-ból származó, az összes tehenre vonatkozó 1,4, 1,5 valamint 1,5-ös arányhoz.

## KÖVETKEZTETÉSEK ÉS JAVASLATOK

A bemutatott Bayes-i modellt az adott telepek adataira illesztve becslés kapható a fertőzés valódi állományon belüli prevalenciájára az egyszer ellett és a többször ellett tehenek körében. Amennyiben a Bayes-i modell alkalmazása nehézségekbe ütközik, a valódi prevalenciák hüvelykujjszabállyal becsülhetőek a látszólagos és a valódi prevalencia lineáris összefüggése és az országos felmérés eredményei alapján. Egyszer ellett tehenek esetében a valódi prevalencia becslése a látszólagos prevalencia 1,6-szorosa, többször ellett teheneknél pedig 1,5-szöröse. A telepi menedzsment számára meglepő és tanulságos lehet, hogy a tesztek által szolgáltatott nyers adathoz képes a valódi fertőzöttség jóval nagyobb. Ennek elsődleges oka a tehenek paratuberkulózisának kimutatására alkalmazott tejalapú ELISA-tesztek gyenge érzékenysége.

A Bayes-i modell által a prevalencia becslésére szolgáltatott alsó és felső értékhatár, az ún. intervallumbecslés abban az esetben is nagy valószínűséggel lefedi a valódi prevalenciát, ha csak a tehenek egy adott százalékát szűrjük, ám pontossága romlik, az intervallum hossza jelentősen nő a

mintavételi százalék csökkenésével. A becslés kívánt pontosságának, vagyis az intervallum hosszának előre megadásával meghatározható, hogy a telepről az egyszer, ill. többször ellett tehenek hány százalékából kell mintát venni és diagnosztikailag megvizsgálni.

A fertőző betegségek elleni küzdelemben kulcsfontosságú a látszólagos és a tényleges prevalencia közötti különbség tudatosítása.

A Bayes-i módszerek alkalmasak a valódi előfordulás becslésére, segítve a telepi döntéshozókat a fertőzés okozta károk felmérésében és a megfelelő megelőző, ill. védekezési intézkedések kidolgozásában. A szerzők tudomása szerint ez az első magyarországi, egyedi telepekre illeszthető Bayes-i modell az állományon belüli (egyedi szintű) valódi MAP-fertőzés becslésére, amely során hazai, országos PTBC-felmérésen alapuló priorokat használtunk. A modell eredményeire építve a valódi prevalencia becslésére egyszerűen használható módszert adunk azok számára is, akiknek nincs módjukban a modellt futtatni.

*A felhasznált irodalom a Szerzőknél rendelkezésre áll.*

