Állatorvostudományi Egyetem

Állatorvostudományi Doktori Iskola

Kondicionális és marginális modellek az állategészségügyben a szarvasmarha paratuberkulózis példáján bemutatva

PhD értekezés tézisei

Veres Katalin

2025

Témavezetők

.....

Dr. Lang Zsolt, tudományos főmunkatárs

Állatorvostudományi Egyetem, Gazdaságtudományi és Biostatisztikai Intézet, Biostatisztika Tanszék

.....

Prof. Dr. Ózsvári László, egyetemi tanár

Állatorvostudományi Egyetem, Gazdaságtudományi és Biostatisztikai Intézet, Törvényszéki Állatorvostani és Gazdaságtudományi Tanszék

Témabizottsági tagok

Dr. Harnos Andrea, docens, PhD

Állatorvostudományi Egyetem, Gazdaságtudományi és Biostatisztikai Intézet, Biostatisztika Tanszék

Dr. Reiczigel Jenő, egyetemi tanár, az MTA doktora

Állatorvostudományi Egyetem, Gazdaságtudományi és Biostatisztikai Intézet, Biostatisztika Tanszék

.....

Veres Katalin

Ez a(z) … számú példány

Tartalomjegyzék

Tartalomjegyzék **3**

1 Bevezetés és a kutatás céljai **5**

2 Anyag és módszer **9**

2.1 Adatbázisok 9

2.2 Eloszlásalapú marginalizált többszintű modellek 10

2.3 Valódi prevalencia becslése állatok alcsoportjaiban Bayes-i látens osztálymodellel és a módszer alkalmazása szarvasmarha paratuberkulózis fertőzés esetében 10

2.4 A paratuberkulózis fertőzöttség becslése tejelő tehénállományok alcsoportjaiban 12

2.5 A PTBC-fertőzöttség és a selejtezés összefüggése a hazai tejelő szarvasmarha-állományokban 13

3 Eredmények **16**

3.1 Valódi prevalencia becslése állatok alcsoportjaiban Bayes-i látens osztálymodellel és a módszer alkalmazása szarvasmarha paratuberkulózis fertőzés esetében 16

3.2 A paratuberkulózis fertőzöttség becslése tejelő tehénállományok alcsoportjaiban 17

3.3 A PTBC-fertőzöttség és a selejtezés összefüggése a hazai tejelő szarvasmarha-állományokban 18

4 Megbeszélés **20**

4.1 Valódi prevalencia becslése állatok alcsoportjaiban Bayes-i látens osztálymodellel és a módszer alkalmazása szarvasmarha paratuberkulózis fertőzés esetében 20

4.2 A paratuberkulózis fertőzöttség becslése tejelő tehénállományok alcsoportjaiban 23

4.3 A PTBC-fertőzöttség és a selejtezés összefüggése a hazai tejelő szarvasmarha-állományokban 24

5 Következtetések és javaslatok 26

6 Új és újszerű eredmények 29

7 Publikációk 31

7.1 A disszertáció alapjául szolgáló közlemények 31

7.2 Konferencia kiadványokban megjelent publikációk ..32

7.3 Egyéb tudományos közlemények 32

1 Bevezetés és a kutatás céljai

A fertőző betegségek prevalenciájának felmérése diagnosztikai tesztek segítségével történik. Ezekkel a tesztekkel az úgynevezett „látszólagos prevalencia” (AP), vagyis a pozitív esetek aránya határozható meg. A látszólagos és a valódi prevalencia között markáns különbségek lehetnek a téves pozitív és téves negatív teszteredmények miatt. Kizárólag a látszólagos prevalenciára hagyatkozni hibás döntésekhez vezethet a telepi menedzsmentben. Ez a nyers mérőszám azonban korrigálható a Rogan-Gladen módszer segítségével, és így megbecsülhető a valódi, mögöttes prevalencia.

Az állatállományok, mint gazdálkodási egységek, hierarchikus adatszerkezettel ragadhatóak meg. Hierarchikus adatok modellezése feltételes és marginális modellekkel is történhet. A marginalizált többszintű modellek (MMM) ötvözik e két eltérő megközelítés előnyeit, robusztus és könnyen értelmezhető becslést adnak, emellett becsülhetőek likelihood alapú módszerekkel. A vizsgált betegség valódi prevalenciája Bayes-i módszerrel becsülhető. Tudomásom szerint nem létezett korábban olyan módszer, amely lehetővé tette volna, hogy állatállományok alcsoportjait a közöttük lévő korreláció figyelembevételével úgy modellezzünk, hogy feltételes és marginális becsléseket is kapjunk a prevalenciára. A dolgozat célja olyan biostatisztikai modellek kifejlesztése és alkalmazása

volt, amelyek pótolják ezt a hiányt, betekintést nyújtva az alcsoportok közötti összefüggésekbe. A jelen munkában bevezettem egy, a már bevett módszereken alapuló, mégis újszerű modellkeretet, az eloszlásalapú marginalizált többszintű modelleket (DMMM). A módszer előnyeit a magyar tejelő szarvasmarha-állományok paratuberkulózis (PTBC)-prevalenciájának modellezésén keresztül mutatom be.

A módszertani újdonságok illusztrálására a tejelő tehenek PTBC-fertőzésére vonatkozó adatokat használtam. A PTBC, amelyet a *Mycobacterium avium subsp. paratuberculosis* (MAP) okoz, világszerte széles körben elterjedt szarvasmarha betegség. A PTBC-fertőzés klinikai és szubklinikai stádiuma egyaránt jelentős gazdasági károkat okoz, de a szubklinikai stádium hatása jelentősebb, mivel ebben a szakaszban az állatok nem mutatnak nyilvánvaló tüneteket, de a termelésre gyakorolt negatív hatás már jelen van. A MAP-fertőzés látenciája és lassú lefolyása, valamint a diagnosztikai tesztek alacsony érzékenysége megnehezítik a betegség diagnosztizálását, így a látszólagos és a valódi prevalencia között jelentős különbségek mutatkozhatnak. Pontatlanná válhat mind a betegség prevalenciájának becslése, mind a járványvédelmi intézkedések hatásának kiértékelése, ami hibás állomány-egészségügyi döntésekhez és nem megfelelő intézkedésekhez vezethet.

A PTBC fertőzés valódi prevalenciáját Magyarországon korábban csak egyszer becsülték meg kis méretű adathalmazt használva. Az eredmények magas állományszintű (89,1%) és közepesen magas medián egyedszintű valódi prevalenciát jeleztek (egyszer ellett tehenek: 4,4%, többször ellett tehenek: 10,3%).

A dolgozat céljai a következők:

- olyan biostatisztikai modellek kidolgozása, melyek könnyen értelmezhető, kondicionális és marginális becslése adnak az alcsoport átlagokra, és mélyebb betekintést nyújtanak az köztük lévő kapcsolatba,

- az új modellkeret használatának szemléltetése a magyarországi tejelő szarvasmarha-állományok PTBC-fertőzés prevalenciájának modellezésén keresztül, továbbá a korábbi eredmények megerősítése nagy mennyiségű új adat és historikus priorok felhasználásával,

- az egyedi állományok PTBC-fertőzés prevalenciájának modellezése Magyarországon és nemzetközileg, valamint egyszerű hüvelykujjszabály megfogalmazása a valódi prevalencia becslésére,

- egy újszerű - a humán epidemiológia területén már használt - feltételes túléléselemzési modellezési megközelítés alkalmazásának vizsgálata az állatorvos-tudományban, ennek adaptálása a PTBC-fertőzés és a tehenek selejtezési idejére közötti összefüggés elemzésére.

2 Anyag és módszer

2.1 Adatbázisok

Az elemzésekben két adatsort dolgoztam fel. A 2019-es *prevalencia adatsor* a tehénszintű adatokat, valamint a PTBC-szűrővizsgálatok eredményeit tartalmazza 117 nagy (≥100 tehén) intenzív tejelő szarvasmarha-állomány 58.196 tehénéről. Az adatállomány lefedi az összes magyarországi tejelő tehén 24,3%-át, és a teljesítményvizsgálatban szereplő állományok 26,8%-át. Ezt az adatsort a prevalencia-becslési vizsgálatokban használtam. Az *élettartam adatsor* öt állomány 4.347 tejelő tehenének egyedi szintű adatait és PTBC-szűrővizsgálati eredményeit tartalmazza a 2018-as évből. Ezt az adatállományt a szeropozitivitás selejtezési kockázatra gyakorolt hatásának vizsgálatára használtam.

A tejminták diagnosztikai vizsgálatait a gyártó utasításainak megfelelően a Paratuberculosis Screening Antibody Test (IDEXX Laboratories Inc., Westbrook, ME, USA) ELISA teszt segítségével végeztük el. Az eredmények kategorizálásánál a negatív ill. pozitív küszöbérték a 0,2-es ill. 0,3-as S/P arány volt. A két küszöbérték közé eső eredményeket bizonytalannak minősítettem.

2.2 Eloszlásalapú marginalizált többszintű modellek

Kidolgoztam egy speciális, marginalizált többszintű modellt, az eloszlás-alapú marginalizált többszintű modellt (Distribution based Marginalized Multilevel Models, DMMM) és feltártam ennek alapvető jellemzőit. A DMMM egy új, innovatív modellezési technika, amely alkalmas a populációs és állományszintű valódi, mögöttes átlagos prevalenciák becslésére a releváns állományon belüli alcsoportok esetében, figyelembe véve az adatok hierarchikus szerkezetét. Ez a módszer rávilágít az alcsoport-prevalenciák közötti összefüggésekre, és megmutatja, hogy a kovariánsok hatása populációs vagy az állományszinten érvényesül erősebben.

2.3 Valódi prevalencia becslése állatok alcsoportjaiban Bayes-i látens osztálymodellel és a módszer alkalmazása szarvasmarha paratuberkulózis fertőzés esetében

A csoportokban (pl. tenyésztelepeken) élő állatok fertőző betegségeinek prevalenciája természetes módon elemezhető Bayes-i hierarchikus látens osztálymodellel. Munkámban e módszertan kibővítését tűztem ki célul oly módon, hogy az állatok alcsoportjainak sajátosságait a korrelált, alcsoport-szintű prevalenciák modellezésén keresztül vettük figyelembe. Módszertanomat a magyarországi nagyüzemi intenzív tejelő

szarvasmarha telepek paratuberkulózis (PTBC) fertőzés prevalenciájának újraértékelésével illusztráltam. Az volt a célom, hogy könnyen értelmezhető alcsoport-szintű becsléseket adjak a populációs átlagokra, és betekintést nyújtsak a paritás-csoportok fertőzöttsége közötti összefüggésekbe. A PTBC-fertőzés valódi prevalenciáját Magyarországon korábban csak egyszer becsülték, kis méretű adathalmazt használva, ezért másodlagos célom az volt, hogy nagy mennyiségű új adat és historikus prior eloszlások felhasználásával megerősítsem a korábbi eredményeket.

A valódi fertőzési prevalencia becslésére a Hanson és mtsai. és Branscum és mtsai. modelljeinek általánosítását használtam. Az állatcsoportokon belüli fertőzés prevalencia becsléséhez a korrelált, telepen belüli alcsoport prevalenciákat független, normális eloszlású telepi random hatásokból származtatott béta eloszlással modelleztem. Az alkalmazásban a telepek fertőzöttségi státuszát látens osztályokként kezeltem, a tehenek paritás szerinti csoportjait (egyszer- illetve többször ellett tehenek) pedig állományon belüli alcsoportokként. Az új módszertan lehetővé tette, hogy minden alcsoportra vonatkozóan egyaránt megbecsüljem a feltételes állományon belüli valódi prevalencia ($CWHP_1$ $és$ $CWHP_2$) átlagát és mediánját, valamint, hogy további, az alcsoportok közötti összefüggéseket jellemző mérőszámokat származtathassak.

2.4 A paratuberkulózis fertőzöttség becslése tejelő tehénállományok alcsoportjaiban

Jelen kutatás célja a Bayes-i módszertan legfontosabb fogalmainak, működésének áttekintése és gyakorlati alkalmazásának szemléltetése volt. A gyakorlati alkalmazás célja az volt, hogy a diagnosztikai teszt korlátait figyelembe véve megbecsüljem a paratuberkulózis fertőzés valódi prevalenciáját egy tejelő szarvasmarha állományban. A modell eredményeit két számítógéppel szimulált telepen szemléltettem.

A kutatás során a korábbi magyarországi felmérés adatait használtam fel. Elsőként Bayes-i kétlépcsős hierarchikus modellt illesztettem az egyes telepek adataira. A valódi prevalencia megbecsléséhez az ELISA teszteredményeket, a paritáscsoportot és a tehenek életkorát használtam fel. A felmérés országos szintű eredményei alapján informatív priorokat konstruáltam a $CWHP_1$ (feltételes telepen belüli prevalencia az egyszer ellett tehenek között) és $CWHP_2$ (feltételes telepen belüli prevalencia a többször ellett tehenek között) eloszlásához.

A modell pontosságát és a mintavételi százalék csökkentésének hatását számítógépes szimulációval vizsgáltam. Ezt követően a modellt lefuttattam 116 magyar,

tejelő szarvasmarha telep tejalapú PTBC ELISA teszteredményeire. Célom egyszerűen használható közelítő módszer kidolgozása volt. A modell által becsült valódi prevalenciát függő változónak, a nyers adatokból számított látszólagos prevalenciát magyarázó változónak tekintve lineáris regressziót illesztettem.

Az elemzést nemzetközi kontextusba helyeztem: A különböző régiók prior információt régióspecifikus szintetikus adatokkal ötvöztem.

2.5 A PTBC-fertőzöttség és a selejtezés összefüggése a hazai tejelő szarvasmarha-állományokban

A látens változók olyan jellemzők, melyeket nem tudunk közvetlenül megfigyelni vagy mérni, így értékeikre egyéb megfigyelt változókból következtethetünk. A látens változók egyedi megfigyelési egységekhez tartozó értékei többnyire kevéssé érdeklik a kutatókat, általában csak a varianciájukat modellezik. Az adatok heterogenitása (ami az egyéni vagy klaszterszintű látens változók szintjén jelenik meg) kezelhető marginális modellekkel, feltételes modellekkel és ezek társításával (MMM). Bár ezek a modellezési stratégiák mind a heterogenitás kezelésére szolgálnak, más következtetésekre alkalmasak és eredményeik értelmezése is eltérő. A megfelelő

módszertant mindig a konkrét kutatási kérdésnek megfelelően szakmai alapon kell megválasztani. A kondicionális és marginális megközelítés használatát a túléléselemzés területén esettanulmányokkal illusztráltam.

A PTBC-pozitivitás hatását a tehenek selejtezési kockázatára az *élettartam* adatsor elemzésével vizsgáltam a hazai tejelő tehénállományokban. A teheneket az ELISA-teszt elvégzésének napjától a selejtezésig vagy elhullásig, de legkésőbb a vizsgálat végéig (2019. július 31.) követtem, amikor adminisztratív cenzorálás történt. Megbecsültem az ELISA-pozitív csoport relatív selejtezési kockázatát (hazard ratio) az ELISA-negatív csoporthoz képest.

Először Cox-féle marginális arányos hazárd regressziós modellt illesztettem a túlélési adatokra, hogy megbecsüljem a két csoport relatív kockázatát (HR). Ebben a marginális modellben a hazárd a kovariánsok függvénye, és a telepi látens hatás ki van átlagolva.

Ezt követően kondicionális, Cox-féle arányos hazárd frailty regressziós modellt illesztettem az adatokra. A frailty hatások a random hatások túlélés-elemzésbeli megfelelői. Ebben a modellben a klaszteren belüli korrelációkat frailty hatásokkal modelleztem. A logaritmikus skálán lévő frailty hatásokra normális eloszlást feltételeztem (Gaussian frailty).

Végezetül rétegzett Cox-féle arányos hazárd regressziós modellt is illesztettem az adatokra. Ez a modell úgy veszi figyelembe a rétegek közötti heterogenitást, hogy a különböző rétegekben különböző alaphazárd-függvényt enged meg. A korcsoportot és az állományi hovatartozást tekintettem a heterogenitás forrásának. Annak érdekében, hogy ezek torzító hatását eltávolítsam a pozitív és negatív esetek kockázatának összehasonlításából, minden ELISA-pozitív tehénhez visszatevés nélkül 5 véletlen kontrollállatot választottam ki az azonos az állományból és korcsoportból származó ELISA-negatív tehenek közül. A rétegeket ezek az illesztett állatcsoportok határozták meg.

3 Eredmények

3.1 Valódi prevalencia becslése állatok alcsoportjaiban Bayes-i látens osztálymodellel és a módszer alkalmazása szarvasmarha paratuberkulózis fertőzés esetében

A prevalencia adatbázisban összesen 3.343 tehén bizonyult ELISA pozitívnak, ami összességében 6,0%-os látszólagos prevalenciát jelent. Míg az ELISA teszteredmény alapján számított látszólagos állomány-szintű prevalencia 96,6% volt, a becsült átlagos valódi állomány-szintű prevalencia alacsonyabb szintre, 92,5%-ra (95% kredibilis intervallum (KI): 88,0%; 96,1%) becsülhető. A feltételes állományon belüli prevalenciát (CWHP) 8,4%-ra (6,6%; 10,4%) becsültük az egyszer ellett és 15,8%-ra (13,5%; 18,4%) a többször ellett tehenek esetében, szemben a 3,3% és 8% látszólagos prevalenciával. A fertőzött állományok CWHP-jának mediánjára az egyszer ellett csoportban 4,7%-ot (3,2%; 6,4%), míg a többször ellett csoportban 12,4%-ot (10%; 15%) kaptam. A telepen belüli, egyszer és többször ellett tehenek közötti prevalencia-arány mediánja 2,6 (0,8; 56,7) volt.

3.2 A paratuberkulózis fertőzöttség becslése tejelő tehénállományok alcsoportjaiban

A modell eredményeit két számítógéppel szimulált telep elemzésével szemléltettem. Az egyik vizsgált telepen a fertőzés valódi prevalenciája magas volt, míg a másikon alacsony. A modelleredmények felhasználásával egy könnyen használható, internetről letölthető eszközt fejlesztettem, mely a tehenek adatai és ELISA teszteredmény alapján megbecsüli a valódi prevalenciát a megadott állomány alcsoportjaiban.

A becslések ellenőrzésére végzett szimulációban az először ellett tehenek 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 tapasztaltam, 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 tehenek esetében 90% körül maradt. Azonban a becslés pontossága romlott, a kredibilis intervallum szélessége jelentősen megnőtt.

A regresszió eredménye alapján egyszer ellett tehenek esetében a valódi prevalencia becsülhető a látszólagos

prevalencia 1,6-szorosával, többször ellett teheneknél pedig 1,5-szörösével.

3.3 A PTBC-fertőzöttség és a selejtezés összefüggése a hazai tejelő szarvasmarha-állományokban

Az ELISA pozitív csoport hazárdarány (HR) értéke 2,02 (95% konfidencia intervallum (CI): 1,5-2,7) volt a marginális modellben. Ez azt jelenti, hogy az ELISA-pozitivitás a teljes tehénpopulációban 2,02-szeresére növeli a selejtezés kockázatát.

A feltételes frailty modellben az ELISA-pozitív csoport HR becslése 2,03 (95% CI: 1,7-2,4) volt. Ez azt jelzi, hogy az ELISA-pozitív tehenek selejtezési kockázata a tehenek életkorára és az állományra való korrekció után 2,03-szoros az ELISA-negatívokhoz képest.

A rétegzett Cox modell alapján az ELISA-pozitív csoport HR értéke 2,13 (95% CI: 1,75-2,60) volt, ami azt jelenti, hogy az ELISA-pozitivitás mellett 2,13-szorosára nő a selejtezés kockázata. Ez a feltételes becslés korrigálva van az illesztéshez használt kovariánsokra, azaz egy véletlenszerűen kiválasztott ELISA-pozitív tehén esetében 2,13-szor nagyobb a selejtezés

kockázata, mint az ugyanabból az állományból és korcsoportból véletlenszerűen kiválasztott ELISA-negatív tehénnek.

4 Megbeszélés

4.1 Valódi prevalencia becslése állatok alcsoportjaiban Bayes-i látens osztálymodellel és a módszer alkalmazása szarvasmarha paratuberkulózis fertőzés esetében

Egy új, Bayes-i modellezési megközelítést mutattam be betegségek prevalenciájának modellezésére, amely figyelembe veszi az egyéni szintű kovariánsok hatását. Ez a módszertani innováció jól illeszkedik a meglévő szakirodalomba, hiszen az elterjedt Hanson- és Branscum-modell általánosítása. Az alcsoportok korrelált prevalenciáit béta eloszlásokkal modelleztem. A béta eloszlás használata a fertőző betegségek valódi prevalenciájának modellezésére széles körben elfogadott Bayes-i módszer. Ezen tanulmányok közös feltételezése a betegség prevalenciájának telepen belüli homogenitása. Az én megközelítésemben az alcsoportokhoz különböző marginális átlagok tartoznak és az azonos állományon belüli alcsoportok prevalenciájának összefüggését a korrelált random telepi hatásokkal modellezzük. A korrelációs együtthatót a modellből becsüljük. Az alcsoportok korrelált prevalenciáinak modellezése lehetővé teszi az alcsoport átlagok becslését és betekintést ad ezek kapcsolatába.

A módszer elsődlegesen az átlagos prevalenciát becsüli, de lehetővé teszi a medián egyszerű kiszámítását is. A közvetlenül a mediánt modellező módszerek esetében az átlag becslésének kiszámítása általában nehézkes. Az átlag és a medián a poszterior eloszlás különböző aspektusait ragadják meg, és együttesen természetes mértéket adnak az eloszlás ferdeségre. Az alcsoportátlagok könnyen értelmezhető súlyozott populációs átlaggá összegezhetők. Az átlag fogalma ismerős a végfelhasználók számára, az eredmények értelmezése egyszerű. A modell poszterior eloszlásaiból kiindulva további mérőszámok számíthatók MCMC iterációk és szimulációs technikák segítségével.

A robusztus módszertan segítségével és historikus prior információk felhasználásával újraértékeltem a magyarországi PTBC helyzetet. Az alkalmazás eredményei összhangban vannak a korábbi PTBC tanulmány megállapításaival. Az újabb és jelentősen nagyobb adathalmaz, valamint a historikus priorok használata megerősítette a korábbi tanulmány eredményeit.

Az alkalmazás eredményei tekintetében a $CWHP_1$ és $CWHP_2$ átlag- és mediánbecslés közötti pozitív különbség a telepek közötti poszterior eloszlások jobbra ferdeségét jelzi. Ez azt jelenti, hogy nagy a valószínűsége annak, hogy bizonyos

telepek esetében akár a $CWHP_1$, akár a $CWHP_2$ szokatlanul magas.

A $CWHP_1$ és $CWHP_2$ közötti magas maximálkorreláció erős kapcsolatot mutat az ugyanazon állományból származó egyszer és többször ellett tehenek feltételes állományon belüli prevalenciája között. Az egyszer és többször ellett tehenek prevalenciája közötti korreláción felül a különbségeket a prevalencia-aránnyal számszerűsíthetjük. Ez az arányszám relatív kockázatként értelmezhető, amely megmutatja, hányszor gyakoribb a fertőzés a többször ellett tehenek körében az egyszer ellett tehenekhez képest. A magas prevalencia-arány a paritás és a tehenek korának erős hatását mutatja.

A random állomány- és paritás hatások segítségével a prevalencia állományok közötti heterogenitását mérhetjük fel, továbbá rávilágíthatunk a paritáscsoportok prevalenciája közötti korreláció forrására. Becsült szórásuk nem értelmezhető közvetlenül, a heterogenitás értékeléséhez a CWHP szórását kell meghatároznunk a random hatások szórásából kiindulva. Az alkalmazásban a $CWHP_1$ és $CWHP_2$ szórásának nagyságrendje az átlagokkal megegyező, ami erős heterogenitást jelez az állományok közötti prevalencia értékekben.

4.2 A paratuberkulózis fertőzöttség becslése tejelő tehénállományok alcsoportjaiban

Az egyes telepeken a látszólagos és valódi prevalencia aránya a Rogan-Gladen képlet alapján 1-hez közeli specificitás esetén elsősorban a teszt szenzitivitásától függ. Az arány vizsgálatához ezen felül számításba kell venni a szenzitivitás Meyer-képlet által leírt életkorfüggését. Az egyszer ellett tehenek túlnyomó többsége 2-3 éves, az életkor szórása alacsony, ezért a szenzitivitás ebben a kategóriában viszonylag stabil. Bár többször ellett tehenek esetében nagyobb az életkor szerinti szórás, a Meyer-képlet alapján nagyobb életkor mellett a szenzitivitás kevésbé függ az életkor változásától, heterogenitásától. A fentiekben leírtak szerint a valódi és a látszólagos prevalencia összefüggése lineáris és stabil, jól becsülhető, a lineáris regresszió illesztése elméletileg megalapozott. Az általam talált 1,5, ill. 1,6-os TP/AP arány nagyon hasonló Verdugo és mtsai dán kutatásában publikált egymást követő három évből, 2011, 2012 és 2013-ből származó, az összes tehénre vonatkozó 1,4, 1,5 illetve 1,5-os arányhoz.

Fontos hangsúlyozni, hogy a nemzetközi modellben szintetikus adatokat használtam a módszer szemléltetésére, ezért a külföldi régiókra vonatkozó prevalencia becslések nem tükrözik az adott területek PTBC-prevalenciáját. A külföldi régiók

elemzésének eredményei helyi validálás hiányában csupán a modell rugalmasságát mutatják, nem pedig annak pontosságát. Ugyanakkor valós adatok és prior információk felhasználásával a nemzetközi modell alkalmas bármely régió fertőzés-prevalenciájának modellezésére.

4.3 A PTBC-fertőzöttség és a selejtezés összefüggése a hazai tejelő szarvasmarha-állományokban

A háromféle becslés hasonló eredményt ad, azonban különbözőképpen értelmezendőek. A marginális becslés az országos vagy regionális szintű mentesítési döntések támogatására használható, de nem alkalmas állományokra vagy egyes állatokra vonatkozó következtetések levonására. A frailty modell a klaszter- vagy egyedszintű relatív kockázatot becsüli, így lehetővé teszi, hogy összehasonlítsuk az ugyanabból az állományból véletlenszerűen kiválasztott ELISA-pozitív és negatív tehenek selejtezési kockázatát.

A rétegzett Cox-modell az esetek és kontrollok párosításával olyan becslést ad a HR-re, amely kondicionális az elemzés rétegeire. Ez egy hatékony módszertan feltételes, egyedspecifikus túlélési modellek illesztésére. Bár ma már szép számmal léteznek követési időket tartalmazó elektronikus

állatorvosi regiszterek, tudomásom szerint az állatorvos-tudományban eddig még nem alkalmazták ezt a megközelítést.

5 Következtetések és javaslatok

A látszólagos és a valódi prevalencia közötti különbség tudatosítása kulcsfontosságú a fertőző betegségek elleni védekezésben. A Bayes-i módszerek alkalmasak a valódi prevalencia becslésére, ezzel segítve a döntéshozókat a fertőzés okozta termelési veszteségek felmérésében és a megfelelő megelőzési és kontrollintézkedések kidolgozásában. Az adatok összetett hierarchikus struktúrája új modellezési megközelítést igényel, amely lehetővé teszi a gazdasági egységek korrelált alcsoport-átlagainak megbecsülését. A DMMM kielégíti ezt az igényt és hasznos, új eszköz a prevalencia becslésében. A marginális és feltételes átlagok egyszerű és rugalmas kezelését teszi lehetővé.

A DMMM és a Bayes-i módszertan szinergiája innovatív és előremutató, ugyanakkor jól illeszkedik a Hanson, Branscum és McAloon által bevezetett, már elfogadott keretrendszerbe. A korrelált, állományon belüli alcsoport prevalenciák beépítése lehetővé teszi az alcsoportok prevalenciái közötti összefüggéseket jellemző mérőszámok becslését. A DMMM hatékony eszköz, amely könnyen értelmezhető becsléseket biztosít az alcsoportátlagokra. Az alcsoportok prevalenciáinak közvetlen, eloszlás-alapú reprezentációjával jó alapot nyújt további mérőszámok kiszámításához is. A feltételes átlagok becsléseiből például egyszerű súlyozással megkaphatjuk a

teljes populáció átlagát. A poszterior eloszlásokból kiindulva pedig könnyen meghatározható a medián és más érdekes jellemzők is.

Az új módszertan implementációját és előnyeit a magyarországi PTBC prevalencia becslésével szemléltettük. A friss és egyedülállóan nagy adatbázis és a historikus priorok használata a modellezés során megerősítette és finomította a legutóbbi magyarországi PTBC felmérés eredményeit.

A második Bayes-i modell egyedi telepeken ad becslést a fertőzés valódi prevalenciájára az egyszer-, illetve többször ellett tehenek csoportjában. A szerzők tudomása szerint ez az első olyan Bayes-i modell, mely egyedi telepek esetén historikus priorok felhasználásával a korrelált alcsoportok szintjén becsüli a valódi PTBC prevalenciát. A modell eredményei alapján egy egyszerűen használható módszert is adtam a valódi prevalencia becslésére, ha a gyakorló szakembereknek nem lenne lehetőségük a modell futtatására.

A fent leírt modellkeretek módszertani alapot biztosíthatnának a PTBC prevalencia rendszeres értékeléséhez Magyarországon. Alkalmazásukkal további, helyileg releváns tényezők is beépíthetőek a modellbe, például az éghajlat, termelési típus, fajta, korcsoport, genetikai markerek vagy együtt élő fajok, javítva ezzel a modell illeszkedését.

Jó minőségű helyi adatokkal és prior információkkal kombinálva a nemzetközi modell robusztus és rugalmas keretrendszerként szolgálhat a bizonyítékokon alapuló döntéshozatal támogatására állomány-szinten. A módszer újdonsága abban rejlik, hogy nemcsak az állomány fertőzöttségének valószínűségét és a hozzá tartozó Bayes-faktort becsüli, hanem az állomány homogén alcsoportjainak (specifikusan paritási csoportjainak) fertőzés prevalenciáját is.

A technológiai fejlődésnek köszönhetően ma már az állatorvoslásban is rendelkezésre állnak követési időket tartalmazó adatsorok. A túléléselemzés az ilyen típusú adatok megfelelő elemzésének „sine qua non”-ja. Sokoldalú eszköz, amely egyaránt alkalmas mortalitáselemzésre, események bekövetkezési kockázatának megbecsülésére, betegségek prognózisának előrejelzésére, illetve a fentieket befolyásoló tényezők azonosítására. A fentiek miatt fontos, hogy az állatorvosi kutatásba is mielőbb bekerüljenek a túléléselemzésben használt különféle módszerek.

6 Új és újszerű eredmények

1. Bevezettem az eloszlásalapú marginalizált többszintű modellt. Ez a keretrendszer könnyen értelmezhető marginális és feltételes becsléseket szolgáltat, valamint betekintést ad az állományon belüli alcsoportok közötti összefüggésekbe.

2. A fenti keretrendszer használatát a magyar tejelő szarvasmarha állományok paratuberkulózis szeroprevalenciájának felmérésén keresztül mutattam be. A nagyobb adathalmaz használata és a historikus priorok beépítése megerősítette a korábbi eredményeket. Míg az ELISA teszteredmény alapján számított látszólagos állomány-szintű prevalencia 96,6% volt, a becsült átlagos valódi állomány-szintű prevalencia alacsonyabb szintre, 92,5%-ra (95% KI: 88,0%; 96,1%) becsülhető. A CWHP-t 8,4%-ra (6,6%; 10,4%) becsültük az egyszer ellett és 15,8%-ra (13,5%; 18,4%) a többször ellett tehenek esetében, szemben a 3,3% és 8% látszólagos prevalenciával.

3. Modelleztem az egyedi telepek különböző paritású tehéncsoportjainak PTBC prevalenciáját és a valódi prevalencia becslésére hüvelykujjszabályt állítottam fel. Egyszer ellett tehenek esetében a valódi prevalencia

kifejezhető a látszólagos prevalencia 1,6-szorosaként, míg többször ellett tehenek esetében a látszólagos prevalencia 1,5-szöröse. A modell bármely nemzetközi régióra adaptálható, helyi adatok és prior információk használatával.

4. Speciális rétegzett Cox-regressziót illesztettem humán adatokra a herpes zoster demenciára gyakorolt hatásának vizsgálatára. Az ebből a vizsgálatból levont tanulságokra építve adaptáltam ezt a módszert állatorvosi területre, és elemeztem a PTBC szeropozitivitás hazai tejelő tehenek selejtezési idejére gyakorolt hatását.

7 Publikációk

7.1 A disszertáció alapjául szolgáló közlemények

Veres K, Lang Zs, Ózsvári L (2025) Bayesian Assessment of True Prevalence of Paratuberculosis Infection in Dairy Herds and Their Parity Subgroups. Pathogens 14:900. https://doi.org/10.3390/pathogens14090900

Veres K, Lang Zs, Monostori A, Kostoulas P, Ózsvári L (2024) Bayesian latent class modelling of true prevalence in animal subgroups with application to bovine paratuberculosis infection. Prev Vet Med 224:106133. https://doi.org/10.1016/j.prevetmed.2024.106133

Schmidt SAJ, **Veres K,** Sørensen HT, Obel N, Henderson VW (2022) Incident Herpes Zoster and Risk of Dementia. Neurology 99:e660–e668. https://doi.org/10.1212/WNL.0000000000200709

Veres K, Lang, Zs, Ózsvári, L. (2025). Túléléselemzés alkalmazása az állatorvosi gyakorlatban – A paratuberkulózis fertőzöttség és a selejtezés kapcsolata tejelő szarvasmarha állományokban (Application of survival analysis in veterinary practice – The relationship between PTBC infection and culling in dairy cattle herds). Magy Allatorvosok Lapja 147:573-595. https://doi.org/10.56385/magyallorv.2025.10.579-595

Veres K, Lang Zs, Monostori A, Ózsvári L (2024) Bayes-i modellezés a gyakorlatban – tejelő tehénállományok állományon belüli paratuberkulózis érintettségének becslése (Bayesian modelling in practice. Estimation of within-herd paratuberculosis prevalence in dairy cattle herds). Magy Állatorvosok Lapja 146:323–337. https://doi.org/10.56385/magyallorv.2024.06.323-337

7.2 Konferencia kiadványokban megjelent publikációk

Bárdos K, Máté M, Lang Zs, **Veres K**, Czopowicz M, Kaba J, Ózsvári L (2023): Main keeping conditions and CAE infection of Hungarian goat herds. 31st International Congress of the Hungarian Association of Buiatrics, 11-14. October 2023, Keszthely, Hungary, 152-160. pp. ISBN: 978-615-81413-3-8

Lang Zs, Ózsvári L, **Veres K** (2023): Comparison of conditional and marginal means in distribution based marginalized multilevel models. 44th Annual Conference of the International Society for Clinical Biostatistics, 27-31 August 2023, Milan, Italy, Abstract book, 151.

7.3 Egyéb tudományos közlemények

Kurt G, Troelsen FS, **Veres K,** Sørensen HT, Erichsen R (2025) Risk of venous thromboembolism following colorectal cancer surgery in patients with and without inflammatory bowel disease. Am J Gastroenterol. https://doi.org/10.14309/ajg.0000000000003761

Sundbøll J, **Veres K,** Troelsen FS, Würtz M, Bøtker HE, Sørensen HT (2025) Myocardial infarction and short- or long-term risk of a subsequent cancer diagnosis: a Danish Nationwide Cohort Study. Journal of Thrombosis and Haemostasis 23:1023–1033. https://doi.org/10.1016/j.jtha.2024.12.008

Sun Y, **Veres K**, Hasselbalch HC, Frederiksen H, Østgård LSG, Horváth-Puhó E, Henderson VW, Sørensen HT (2024) Myeloproliferative Neoplasms and Dementia Risk: A Population-Based Cohort Study. Eur J Haematol. https://doi.org/10.1111/ejh.14297

Gadgaard NR, **Veres K,** Henderson VW, Pedersen AB (2024) Frozen Shoulder and the Risk of Parkinson's Disease: A Danish Registry-Based Cohort Study. Clin Epidemiol 16:447–459. https://doi.org/10.2147/CLEP.S463571

Korsgaard S, Troelsen FS, **Veres K**, Fuglsang CH, Sørensen HT (2024) Evaluation of Left Truncation and Censoring When Changing the Use of the International Classification of Diseases Eighth Revision Codes to Tenth Revision Codes in the Danish National Patient Registry. Clin Epidemiol 16:319–327. https://doi.org/10.2147/CLEP.S456171

Adelborg K, **Veres K,** Horváth-Puhó E, Clouser M, Saad H, Sørensen HT (2024) Risk and adverse clinical outcomes of thrombocytopenia among patients with solid tumors—a Danish population-based cohort study. Br J Cancer 1–8. https://doi.org/10.1038/s41416-024-02630-w

Laugesen K, **Veres K,** Hernandez-Diaz S, Chiu Y-H, Oberg AS, Hsu J, Rinaudo P, Spaan M, van Leeuwen F, Sørensen HT (2023) Overweight or obesity in children born after assisted reproductive technologies in Denmark: A population-based cohort study. PLoS Med 20:e1004324. https://doi.org/10.1371/journal.pmed.1004324

Collin LJ, **Veres K,** Gradus JL, Ahern TP, Lash TL, Sørensen HT (2022) Pre-existing stress-related diagnoses and mortality: a Danish cancer cohort study. Cancer 128:1312–1320. https://doi.org/10.1002/cncr.34036

Nørgaard M, **Veres K,** Sellebjerg FT, Svingel LS, Foch C, Boutmy E, Sabidó M, Magyari M (2021) Incidence of malignancy in multiple sclerosis: A cohort study in the Danish Multiple Sclerosis Registry. Mult Scler J Exp Transl Clin 7:20552173211053939. https://doi.org/10.1177/20552173211053939

Arnesen CAL, **Veres K,** Horváth-Puhó E, Hansen J-B, Sørensen HT, Brækkan SK (2021) Estimated lifetime risk of venous thromboembolism in men and women in a Danish nationwide cohort: impact of competing risk of death. Eur J Epidemiol. https://doi.org/10.1007/s10654-021-00813-w

Egenvall M, Martling A, **Veres K,** Horváth-Puhó E, Wille-Jørgensen P, Høirup Petersen S, Laurberg S, Sørensen HT, Syk I, Andersen PV, Christensen H, Nielsen JD, Jess P, Madsen

MR, Pedersen AG, Østergaard E, Andersson PH, Bengtsson J, Bragmark M, Buchwald P, Egenvall M, Farahnak P, Folkesson J, Påhlman L, Goldinger M, Heuman R, Lindberg K, Martling A, Näsvall P, Ottosson J, Sandzén B, Barberousse C, Horváth-Puhó E, Laurberg S, Påhlman L, Renehan A, Smedh K, Syk I, Sørensen HT, Wille-Jørgensen P (2021) No benefit of more intense follow-up after surgery for colorectal cancer in the risk group with elevated CEA levels – An analysis within the COLOFOL randomized clinical trial. European Journal of Surgical Oncology 47:2053–2059. https://doi.org/10.1016/j.ejso.2021.03.235

Mark-Christensen A, Erichsen R, **Veres K**, Laurberg S, Sørensen HT (2021) Rectal Cancer Risk and Survival After Total Colectomy for IBD: A Population-Based Study. Diseases of the Colon & Rectum 64:583–591. https://doi.org/10.1097/DCR.0000000000001810

Søgaard KK, **Veres K**, Vandenbroucke-Grauls CMJE, Vandenbroucke JP, Sørensen HT, Schønheyder HC (2020) Community-Acquired Escherichia coli Bacteremia after Age 50 and Subsequent Incidence of a Cancer Diagnosis: A Danish Population-Based Cohort Study. Cancer Epidemiol Biomarkers Prev 29:2626–2632. https://doi.org/10.1158/1055-9965.EPI-20-0705

Ording AG, Horváth-Puhó E, **Veres K,** Glymour MM, Rørth M, Sørensen HT, Henderson VW (2020) Cancer and risk of Alzheimer's disease: Small association in a nationwide cohort study. Alzheimers Dement. https://doi.org/10.1002/alz.12090

Gribsholt SB, Cronin-Fenton D, **Veres K,** Thomsen RW, Ording AG, Richelsen B, Sørensen HT (2020) Hospital-diagnosed overweight and obesity related to cancer risk: a 40-year Danish cohort study. J Intern Med 287:435–447. https://doi.org/10.1111/joim.13013

Bengtsen MB, **Veres K,** Nørgaard M (2020) First-time postmenopausal bleeding as a clinical marker of long-term cancer risk: A Danish Nationwide Cohort Study. Br J Cancer 122:445–451. https://doi.org/10.1038/s41416-019-0668-2

Adelborg K, **Veres K,** Sundbøll J, Gregersen H, Sørensen HT (2019) Risk of cancer in patients with fecal incontinence. Cancer Med 8:6449–6457. https://doi.org/10.1002/cam4.2509

Nørgaard M, **Veres K,** Didden EM, Wormser D, Magyari M (2019) Multiple sclerosis and cancer incidence: A Danish nationwide cohort study. Mult Scler Relat Disord 28:81–85. https://doi.org/10.1016/j.msard.2018.12.014

Sundbøll J, Thygesen SK, **Veres K,** Liao D, Zhao J, Gregersen H, Sørensen HT (2019) Risk of cancer in patients with constipation. Clin Epidemiol 11:299–310. https://doi.org/10.2147/CLEP.S205957

Sundbøll J, Larsen AP, **Veres K,** Adelborg K, Sørensen HT (2019) Cardiovascular event rates and trajectories of LDL-cholesterol levels and lipid-lowering therapy in patients with atherosclerotic cardiovascular disease: A population-based cohort study. Thromb Res 183:124–130. https://doi.org/10.1016/j.thromres.2019.09.034

Søgaard KK, **Veres K,** Nørgaard M, Djurhuus JC, Sørensen HT (2019) Pyelonephritis in persons after age 50 as a clinical marker of urogenital cancer. Clin Microbiol Infect 25:87–91. https://doi.org/10.1016/j.cmi.2018.04.003

Ording AG, **Veres K,** Horváth-Puhó E, Glymour MM, Rørth M, Henderson VW, Sørensen HT (2019) Alzheimer's and Parkinson's Diseases and the Risk of Cancer: A Cohort Study. J Alzheimers Dis 72:1269–1277. https://doi.org/10.3233/JAD-190867

Mark-Christensen A, Erichsen R, **Veres K,** Laurberg S, Sørensen HT (2019) Extracolonic cancer risk after total colectomy for inflammatory bowel disease: A population-based cohort study. J Crohns Colitis. https://doi.org/10.1093/ecco-jcc/jjz199

Ahern TP, **Veres K,** Jiang T, Farkas DK, Lash TL, Sørensen HT, Gradus JL (2018) Adjustment disorder and type-specific cancer incidence: a Danish cohort study. Acta Oncol 57:1367–1372. https://doi.org/10.1080/0284186X.2018.1465586

Nørgaard M, **Veres K,** Ording AG, Djurhuus JC, Jensen JB, Sørensen HT (2018) Evaluation of Hospital-Based Hematuria Diagnosis and Subsequent Cancer Risk Among Adults in Denmark. JAMA Netw Open 1:e184909. https://doi.org/10.1001/jamanetworkopen.2018.4909

Hansen AT, **Veres K,** Prandoni P, Adelborg K, Sørensen HT (2018) Retinal vein thrombosis and risk of occult cancer: A nationwide cohort study. Cancer Med 7:5789–5795. https://doi.org/10.1002/cam4.1803

Sundbøll J, **Veres K,** Horváth-Puhó E, Adelborg K, Sørensen HT (2018) Risk and Prognosis of Cancer After Lower Limb Arterial Thrombosis. Circulation 138:669–677. https://doi.org/10.1161/CIRCULATIONAHA.117.032617

Ording AG, **Veres K,** Farkas DK, Adelborg K, Sørensen HT (2018) Risk of cancer in patients with epistaxis and haemoptysis. Br J Cancer 118:913–919. https://doi.org/10.1038/bjc.2017.494

Jiang T, **Veres K,** Körmendiné Farkas D, Lash TL, Toft Sørensen H, Gradus JL (2018) Post-traumatic stress disorder and incident fractures in the Danish population. Osteoporos Int 29:2487–2493. https://doi.org/10.1007/s00198-018-4644-0

Christensen DH, **Veres K,** Ording AG, Jørgensen JOL, Cannegieter SC, Thomsen RW, Sørensen HT (2018) Risk of cancer in patients with thyroid disease and venous thromboembolism. Clin Epidemiol 10:907–915. https://doi.org/10.2147/CLEP.S158869

Hansen AT, **Veres K,** Horváth-Puhó E, Ehrenstein V, Prandoni P, Sørensen HT (2017) Pregnancy-related venous thromboembolism and risk of occult cancer. Blood Adv 1:2059–2062. https://doi.org/10.1182/bloodadvances.2017010231

Schmidt SAJ, **Veres K,** Ording AG, Farkas DK, Fogh K, Sørensen HT (2016) Venous ulcers and risk of occult hematological or other cancers: a nationwide cohort study. Blood 128:874–877. https://doi.org/10.1182/blood-2016-03-705038