**University of Veterinary Medicine**

**Doctoral School of Veterinary Science**

# Conditional and marginal modelling in animal health using bovine paratuberculosis as an example


PhD thesis

**Katalin Veres**

**2025**

**Supervisors**

.....

Dr. Zsolt Lang, Senior research fellow, PhD

Department of Biostatistics, Institute of Economics and Biostatistics, University of Veterinary Medicine Budapest

.....

Prof. Dr. László Ózsvári, Full professor, PhD

Department of Veterinary Forensics and Economics, Institute of Economics and Biostatistics, University of Veterinary Medicine Budapest

**Committee members**

Dr. Andrea Harnos, Associate professor, PhD

Department of Biostatistics, Institute of Economics and Biostatistics, University of Veterinary Medicine Budapest

Dr. Jenő Reiczigel, Full professor, DSc

Department of Biostatistics, Institute of Economics and Biostatistics, University of Veterinary Medicine Budapest

.....

Katalin Veres

Copy Nr.:…..

## Table of contents

# 1 Introduction and aims of the study

To assess the prevalence of infectious diseases various diagnostic tests are used. These tests are used to calculate the apparent prevalence, i.e. the proportion of test-positive cases, which may differ greatly from the true prevalence due to the possible occurrences of false positive and false negative cases. Using the apparent prevalence (AP) as base for herd management decisions can be highly misleading. However, this measure can be corrected using the method of Rogan and Gladen to obtain the true underlying prevalence.

Animal groups as economic units imply a hierarchical data structure. Both conditional and marginal models are available to accommodate hierarchical data. Marginalized multilevel models (MMM) combine the advantages of these two approaches, while allowing for a likelihood-based estimation. The true prevalence of the investigated disease can be estimated using a Bayesian approach. To my knowledge there was no method that would have allowed researchers to model animal subgroups as correlated factors providing both conditional and marginal estimates. Thus, the aim of the research was to develop biostatistical models providing new insight into the interrelation of subgroups. I introduce a new framework based on legacy methods, the DMMM, and illustrate its merits by the application

to paratuberculosis (PTBC) prevalence modelling in Hungarian dairy cattle herds.

As illustration for the methodological novelties, I used data on PTBC infection of dairy cows. PTBC, caused by Mycobacterium avium subsp. paratuberculosis (MAP), is a globally widespread cattle disease. Both the clinical and subclinical stages of PTBC infection cause substantial economic losses, but the subclinical stage is economically more influential, as at this stage the animals do not show obvious symptoms, but the negative impact on production is already present. The latent and slow progression of MAP infection and the low sensitivity of diagnostic tests make diagnosis challenging, leading to a marked difference between estimated apparent and true prevalence levels. This affects the estimation of disease frequency and the assessment of biosecurity measures which can lead to poor herd health management decisions and MAP control measures.

The true prevalence of PTBC infection in Hungary was estimated only once on a limited dataset. The results indicated a high herd-level (89.1%) and a moderately high median animal-level true prevalence (primipara: 4.4%, multipara:10.3%).

The aims of this study were:

- to develop biostatistical models providing easily interpretable marginal and conditional estimates for subgroup means and new insight into the interrelation of subgroups,

- to illustrate the use of the new framework by modeling the prevalence of PTBC infection in Hungarian dairy cattle herds, and to consolidate previous findings using a recent, extensive dataset and historical priors,

- to model the PTBC prevalence of individual herds both in Hungary and internationally and providing rule of thumb for the simple estimation of true prevalence.

- to explore a conditional modelling approach in survival analysis from the field of human epidemiology which has not been previously applied in veterinary science and adapt it to analyze the impact of PTBC seropositivity on culling times in dairy cows.

# 2 Materials and methods

## 2.1 Datasets

Two datasets were used in the analyses. The *prevalence dataset* consisted of the cow-level data and the PTBC screening test results of 58,196 cows in 117 large (number of cows>=100) intensive dairy cattle herds in 2019, which represented 24.3% of all dairy cows in Hungary and 26.8% of the performance tested herds. This dataset was used in the prevalence estimation studies. The survival dataset comprised cow-level records from the PTBC screening test results of five herds, including 4,347 milking cows in 2018. The dataset was used in the study assessing the impact of seropositivity on the risk of culling.

Milk diagnostic tests were carried out by Paratuberculosis Screening Antibody Test (IDEXX Laboratories Inc., Westbrook, ME, USA) as per the manufacturer's instructions. To classify the test results, I used 0.2 and 0.3 as negativity and positivity thresholds for the S/P ratio, respectively. Results within the two thresholds were classified as inconclusive.

## 2.2 Distribution based Marginalized Multilevel Models

I developed a special type of marginalized multilevel model called Distribution based Marginalized Multilevel Models (DMMM) and explored its basic characteristics. The DMMM is a novel, innovative modelling technique appropriate to calculate the true, underlying population level and herd-level mean prevalences of relevant within-herd subgroups, taking into account the hierarchical structure of the data. This method sheds light on the interrelation of subgroup prevalences and enables us to assess whether the effect of covariates is more pronounced at the population or herd level.

## 2.3 Bayesian latent class modelling of true prevalence in animal subgroups with application to bovine paratuberculosis infection

The prevalence of infectious diseases in animals living in separate groups (e.g. herds) can be naturally analyzed using a Bayesian hierarchical latent class model. In this work I aimed to extend this methodology considering the specificities of distinct animal subgroups by modeling correlated conditional, within herd subgroup level prevalences. I illustrated the merits of our methodology with an application reassessing the prevalence of bovine paratuberculosis (PTBC) infection in Hungarian

commercial dairy farms. I aimed to provide easily interpretable subgroup-level estimates for population means, and to gain new insights into the interrelation of parity groups. Since the true prevalence of PTBC infection in Hungary was estimated only once on a limited dataset, my secondary goal was to consolidate previous findings using a large body of recent data and prior distributions derived from historical results.

I propose a generalization of the models of Hanson et al. and Branscum et al. to estimate the true infection prevalence. To model the subgroup level infection prevalences within animal groups, I considered correlated prevalences following beta distributions derived from independent normally distributed random herd effects. In the application, herd true infection status was handled as latent class, while multiparous and primiparous cows were considered as within-herd subgroups. The novel methodology allowed estimation of both the mean and median conditional within-herd true prevalence ($CWHP_1$ $and$ $CWHP_2$) related to each animal subgroup as well as of other measures characterizing the interrelation of subgroups.

## 2.4 Estimation of paratuberculosis infection in subgroups of individual dairy herds

The aim of this research was to review the main concepts of Bayesian methodology and to illustrate its practical applications. The goal of the practical application was to estimate the true prevalence of paratuberculosis infection in an individual dairy cattle herd taking into account the limitations of the diagnostic test. Model results were illustrated by analyzing two simulated herds.

Herd-level data was used from a Hungarian national survey. First, I fitted a Bayesian two-stage hierarchical model to the data of a single herd. ELISA tests results, the parity and the age of the cows were used to estimate the true prevalence. Informative priors were constructed for $CWHP_1$ (conditional within-herd prevalence among primiparous cows) and $CWHP_2$ (conditional within-herd prevalence among multiparous cows) based on the national-level results of the survey.

The model's accuracy and the impact of reducing the sample size were evaluated using computer simulation. Second, the model was run on individual PTBC ELISA milk test results from 116 Hungarian dairy cattle farms, and a linear regression was fitted to the results. My aim was to develop an easy-to-use approximation method. I considered the estimated true

prevalence as a dependent variable and the apparent prevalence calculated from the raw data as an explanatory variable.

I provided international context for the analysis by integrating prior information from diverse regional studies together with region specific synthetic data. To infer the infection probability of the herd, Bayes factors and the posterior probability of the model were used. Bayes factor is defined as the ratio of the probability of the observed data in one statistical model to the probability of the same data in another model. The Bayes factor quantifies how much more (or less) likely the data are in one model than in another, thus indicating which model provides a better fit to the observed data. The posterior probability of the model given the data is $\frac{BF \times \mu_{HTP}}{BF \times \mu_{HTP} + (1 - \mu_{HTP})}$, where BF is the Bayes factor, and $\mu_{HTP}$ is the expected value of the prior distribution of HTP.

## 2.5 The relationship between PTBC infection and survival in Hungarian dairy herds

Latent variables are not observed or measured directly; their values are inferred from other observed variables. Individual values of the latent variables are of secondary interest, and usually we model only their variance. The heterogeneity of the

data (represented by individual or cluster level latent variables) can be handled using marginal models, conditional models and the mixture of them (MMM). Although these modelling strategies serve a common purpose, they have different targets of inference thus the interpretation of their results is different. The choice between these options should be made according to the research question, on subject-matter grounds. I illustrated the marginal and conditional approach with case studies in the context of survival analysis.

I analyzed the *survival* dataset to assess the impact of PTBC positivity on the risk of culling of cows in Hungarian dairy herds. Cows were followed from the day of the ELISA test until culling or death, but at the latest until the end of the study on 31 July 2019, when administrative censoring was performed. I estimated the relative culling risk (hazard ratio) of the ELISA-positive group compared to the ELISA-negative group.

First, I fitted a marginal Cox proportional hazards regression model to the survival data to estimate the hazard ratio (HR) of the two groups. In this marginal model, the hazard is a function of the covariates, and the latent effect related to herd is averaged out.

Next, I fitted a frailty (conditional) Cox proportional hazards regression model. Frailties are survival analysis counterparts of

random effects. In this model, within-herd correlations were modelled with frailties. I assumed Gaussian distribution for the logarithmically scaled frailties.

Last, I fitted a stratified Cox proportional hazard regression model to the data. This model accounts for the heterogeneity between strata by allowing different baseline hazard functions for each stratum. I considered age group and herd as source of heterogeneity. These covariates were adjusted for by matching. For each ELISA positive cow, 5 control animals were chosen randomly, without replacement, from a pool of ELISA negative cows from the same herd and from the same age group. Strata were defined by matched animal groups.

# 3 Results

## 3.1 Bayesian latent class modelling of true prevalence in animal subgroups with application to bovine paratuberculosis infection

In total, 3,343 cows were classified as ELISA positive, resulting in an overall AP of 6.0%. While the apparent herd prevalence was 96.6% based on ELISA test result, mean HTP was estimated to a lower level of 92.5% (95% credible interval (CrI): 88.0%; 96.1%). Mean CWHP was estimated to 8.4% (6.6%; 10.4%) for primiparous and 15.8% (13.5%; 18.4%) for multiparous cows, in contrast with 3.3% and 8% of AP, respectively. The estimates for median CWHP within infected herds are 4.7% (3.2%; 6.4%) for primiparous and 12.4% (10%; 15%) for multiparous cows. The median and 95% CrI of within-herd prevalence ratio between primiparous and multiparous cows being infected is 2.6 (0.8; 56.7).

## 3.2 Estimation of paratuberculosis infection in subgroups of individual dairy herds

Model results were illustrated by analyzing two simulated herds, one with low, while the other with high level of infection. Based on the model, an easy-to-use downloadable software tool was

developed to estimate the true prevalence of a herd from cow data and ELISA test results.

In the simulation testing the validity of estimates, the true prevalence lied in the 95% credible interval for 90% and 87% of the herds, for primiparous and multiparous cows, respectively. At a lower sampling percentage, there was no substantial change in the proportion of cases in which the credible interval contained the true parameter. The proportion remained around 90% for both primiparous and multiparous cows. However, the width of the credible interval increased considerably, indicating the decreased accuracy of the estimate.

Based on the results of the regression, the true prevalence can be estimated as 1.6 times the apparent prevalence for primiparous cows and as 1.5 times the apparent prevalence for multiparous cows.

In the analysis of the infection status of the simulated herds, based on the value of Bayes factor (BF), the infection status was correctly identified in 66 of the 78 pseudo-herds that were not classified as “weakly refuted” ($0.33 \leq BF < 1$) or “weakly supported” ($1 \leq BF < 3$). When estimating the posterior probability (PP) of infection of the herds, the infection status was correctly identified for 69 out of the 72 pseudo-herds with explicit results ($0.25 > PP$ or $PP >= 0.75$).

## 3.3 The relationship between PTBC infection and survival in Hungarian dairy herds

The hazard ratio (HR) of the ELISA positive group was 2.02 (95% confidence interval (CI): 1.5-2.7) in the marginal model. ELISA positivity has a multiplicative effect on the hazard of culling of the whole cow population, with a factor of 2.02.

In the frailty model, the estimate of the HR of the ELISA positive group was 2.03 (95% CI: 1.7-2.4). ELISA positivity has a multiplicative effect with a factor of 2.03 on the hazard of culling taking into account the age and the herd of the cows.

From the stratified Cox regression with individual level strata the HR estimate of the ELISA positive group was 2.13 (95% CI: 1.75-2.60), meaning that ELISA positivity multiplies the hazard of culling with a factor of 2.13. This conditional estimate is adjusted for matching factors. Its interpretation is that a randomly selected ELISA positive cow has a 2.13 times higher hazard to be culled than a randomly selected ELISA negative cow from the same herd and from the same age group.

# 4 Discussion

## 4.1 Bayesian latent class modelling of true prevalence in animal subgroups with application to bovine paratuberculosis infection

I presented a novel Bayesian modeling approach for disease prevalence that accounts for the impact of individual-level covariates on the prevalence estimation. This methodological innovation fits well into the existing literature as generalization of the well-established Hanson's and Branscum's model. The correlated prevalences of the subgroups are modelled by beta distributions within the same framework. The use of beta distribution to model the true prevalence of infectious diseases is widely accepted in Bayesian methodology. In these studies, one common assumption is the homogeneity of disease prevalence within commercial units (herds). In my approach, subgroups are allowed to have different marginal means, and the association of the prevalence of the subgroups within the same herd is accounted for through correlated random herd effects. The correlation coefficient is estimated from the model. Modelling subgroup level correlated prevalence by beta distributions allows estimating subgroup means and gives new insights into their interrelation.

This method is designed to estimate the mean prevalence yet also allows for easy computing the median. For methods directly modeling the median, the calculation of the mean estimate is in general a lengthy, cumbersome procedure. Median and mean can highlight different aspects of the posterior distribution and, when used together, provide a natural measure of skewness. Subgroup means can be summarized into a meaningful weighted population average. The concept of mean is familiar to end-users; the interpretation of the results is seamless. The posterior distributions from the model offer the possibility to compute various additional measures from MCMC iterates and via simulation techniques.

The evaluation of PTBC situation in Hungary was updated using a robust methodology that allowed the inclusion of historical prior information. The results of the application aligned with the findings of the former PTBC study, while the more recent and considerably larger dataset and the use of historical priors increased the reliability of the results.

As for the results of the application, the positive difference between the mean and the median estimate for both $CWHP_1$ and $CWHP_2$ indicate right skewness of the posterior distributions. This means that there is a high probability that $CWHP_1$ or $CWHP_2$ is unusually large.

The high maximal correlation of $CWHP_1$ and $CWHP_2$ demonstrates a strong link between the conditional within-herd prevalence of primi- and multiparous cows from the same the herd. Apart from the correlation between the prevalence of primiparous and multiparous cows, we can quantify the differences using prevalence ratio. This ratio is a relative risk, showing how much more common the infection is among multiparous cows than among primiparous cows. The high prevalence ratio of infection found between multiparous and primiparous is a measure of the effect of the parity and of the age of the cows.

The random herd and parity effects correspond to the heterogeneity of the prevalence between herds. Furthermore, they highlight the source of correlation between parity groups in herds. The estimation of its standard deviation (SD) cannot be directly interpreted. Instead, to assess the heterogeneity of the prevalence between herds, we can derive the standard deviation of CWHP based on the SD of the random effects. In the application, the SD of $CWHP_1$ and $CWHP_2$ are of the same magnitude as their means indicating strong heterogeneity of the disease prevalence between herds.

## 4.2 Estimation of paratuberculosis infection in subgroups of individual dairy herds

According to Rogan-Gladen formula, in single herds, for specificities close to 1, the ratio of apparent to true prevalence depends primarily on the sensitivity of the test. To examine this ratio, we also need to take into account the age dependence of sensitivity as described by the Meyer formula. The vast majority of primiparous cows in the study were 2-3 years old and the age variance is low, so the sensitivity in this category is relatively stable. Although age variance is higher for multiparous cows, sensitivity at higher ages is less dependent on age variation and heterogeneity according to the Meyer formula. Thus, the relationship between true and apparent prevalence is linear and stable and can be well estimated. Fitting a linear regression is theoretically appropriate. The TP/AP ratios of 1.5 and 1.6 of single herds found for primi- and multiparous cows are comparable to the ratios of 1.4, 1.5 and 1.5 published in the Danish study by Verdugo et al. from three consecutive years, 2011, 2012 and 2013 in overall.

It is important to note that the international model used synthetic datasets to illustrate the approach, and therefore, regional prevalence estimates should not be interpreted as measures of PTBC prevalence in these areas. The results for non-Hungarian regions, in the absence of local validation, demonstrate the

flexibility of the model rather than its accuracy. Yet, the international model can be customized for any region using real local data and prior information.

## 4.3 The relationship between PTBC infection and survival in Hungarian dairy herds

The three estimates are close to each other, however, the results have different interpretations. The marginal estimate is adequate to support country or region level infection-control decisions but cannot be used to draw conclusions for herds or for individual animals. The frailty model estimates herd- and individual-level relative risks and allows comparison of the risk for culling of randomly selected ELISA-positive versus ELISA-negative cows from the same herd.

The stratified Cox model with matched exposed and control patients provides a HR estimate which is conditional on the strata. This is an effective methodology to fit conditional, subject specific survival models to time-to-event data. Although a few electronic databases of time-to-event veterinary data exist, to my knowledge, this approach has not yet been applied in veterinary science.

# 5 Conclusions

The awareness of the difference between the apparent and the true prevalence is key in the control of infectious diseases. Bayesian methods can be used to estimate the true prevalence, helping decision-makers assess the production losses caused by the infection and develop appropriate prevention and control measures. The complex hierarchical structure of the data requires a new modeling approach that allows inference of correlated subgroup means within economic units. In response to this need, DMMMs are a useful novel tool for prevalence estimation providing simple and flexible control over the marginal and conditional means.

The synergy of DMMM and Bayesian methodology is innovative and forward-looking while at the same time fitting well into the proven framework led by Hanson, Branscum and McAloon. Incorporating correlated within-herd subgroup prevalences allows us to estimate various measures characterizing the interrelation of the prevalences in important subgroups. This is a powerful tool providing easily interpretable subgroup-level mean estimates. Allowing for a direct, distributional representation of subgroup level prevalences, this approach is also convenient to use as basis for the calculation of further measures. For instance, the conditional estimates can be summarized into population means with simple weighting.

Median and other characteristics of interest are easy to infer from posterior distributions.

The applications of the method to estimate the prevalence of PTBC in Hungary illustrate the implementation and the merits of the novel methodology. The use of the new methodology along with a more recent and uniquely large dataset and historical priors confirm and refine the results of the latest PTBC screening in Hungary.

The second Bayesian model provides an estimate of the true within-herd prevalence of infection in primiparous and multiparous cows for individual herds. To the best of my knowledge, this is the first Bayesian model for estimating the true PTBC infection at the level of correlated subgroups within an individual herd in Hungary using priors from a national PTBC survey. Based on the results of the model, I also described a straightforward method for estimating the true prevalence for situations in which practitioners do not have the opportunity to run the model.

The above-described frameworks could provide a methodological basis for a regular assessment of PTBC prevalence in Hungary. Future prevalence studies could consider incorporating other locally relevant factors such as climate, production type, breed, age group, genetic markers, or

cohabiting species into the model, to improve fit and to better reflect real-world observations.

Combined with high-quality local data and prior information, the international model serves as a robust and adaptable framework to support evidence-based decision making at the herd level. The novelty of the method is the estimation of both the probability and the Bayes factor of infection of the target herd, combined with the estimation of infection prevalence within homogeneous subgroups of the herd, specifically parity groups.

With technological advances, tracking data is now also available in veterinary medicine. Survival analysis is the sine qua non of the adequate analysis of time-to event data. It is a versatile tool that can be used to analyze mortality, assess the risk of an event, predict the prognosis of a disease and identify factors that influence the above. Thus, it is important to promote the use of its various methods in veterinary medicine.

## 6 New and novel scientific results

**1.** I introduced a new class of models, the *Distribution based marginalized multilevel models*. This framework provides easily interpretable marginal and conditional estimates and gives new insight into the interrelation of subgroups.

**2.** I illustrated the use of the new framework by modeling the prevalence of PTBC infection in Hungarian dairy cattle herds. I consolidated previous findings using a recent, extensive dataset and historical priors. While the apparent HP was 96.6% based on ELISA test result, mean HTP was estimated to a lower level of 92.5% (95% CrI: 88.0%; 96.1%). Mean CWHP was estimated to 8.4% (6.6%; 10.4%) for primiparous and 15.8% (13.5%; 18.4%) for multiparous cows, in contrast with 3.3% and 8% of AP, respectively.

**3.** I modelled PTBC prevalence in parity groups within individual herds and developed a rule of thumb for simple estimation of true prevalence. The true prevalence can be expressed as 1.6 times the apparent prevalence for primiparous cows and 1.5 times for multiparous cows. The model can be customized for any international region using local data and prior information.

4. I fitted a special stratified Cox regression model to human data to examine the effect of herpes zoster on dementia. Building on the lessons learned from that study, I adapted this technique to the veterinary field and performed a survival analysis to assess the impact of PTBC seropositivity on the longevity of Hungarian dairy cows.

# 7 Scientific publications

## 7.1 Related scientific publications in peer-reviewed scientific journals with impact factor


**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). Application of survival analysis in veterinary practice – The relationship between PTBC infection and culling in dairy cattle herds (Túléléselemzés alkalmazása az állatorvosi gyakorlatban – A paratuberkulózis fertőzöttség és a selejtezés kapcsolata tejelő szarvasmarha állományokban). 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) Bayesian modelling in practice. Estimation of within-herd paratuberculosis prevalence in dairy cattle herds (Bayes-i modellezés a gyakorlatban – tejelő tehénállományok állományon belüli paratuberkulózisérintettségének becslése). Magy Allatorvosok Lapja 146:323–337. https://doi.org/10.56385/magyallorv.2024.06.323-337

## 7.2 Related conference presentations

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 Scientific publications not related to the topic of the doctoral research

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