Found 9 Records

CONTROL ID: 3279328
PRESENTER: Collins Okoyo
PRESENTER (INSTITUTION ONLY): University of Nairobi, Kenya Medical Research Institute
TITLE: Mathematical Modeling of the Interruption of the Transmission of Soil Transmitted Helminths Infections in Kenya
ABSTRACT BODY:
Abstract Body: BACKGROUND
Kenya, just like other soil transmitted helminthes (STH) endemic countries, has been conducting regular treatment program for the last five years among school aged children as a way to reduce STH infections burden in the country. However, the point of interruption of transmission of these infections still remains unclear. In the current study, we analyzed an age structured mathematical model to predict the point of interruption of these infections in Kenya. The main objective was to develop and analyze an age structured model of the STH population dynamics under a regular STH treatment program to determine infection transmission rate, the point of infection interruption, and the optimal interpulse treatment interval sufficient to achieve STH infections elimination in Kenya.
METHODS
The study utilized age structured model of the STH population dynamics under a regular treatment program. The model was applied to two main age groups: school age children (5-14 years) and adult populations (>15 years) and investigated the potential for STH elimination with finite rounds of treatment while allowing the STH distribution to change dynamically as a function of treatment frequency and treatment coverage. The model was verified using a five year field data from the National School Based Deworming Program (NSBDP) for all the three main STHs; Ascaris lumbricoides, Trichuris trichiura and hookworms.
RESULTS
The model behaviour demonstrated convincingly an accurate predictions of prevalence and mean intensities of infections during and after treatment rounds in each of the age groups. The model indicated that the benefit derived from the regular treatment increases non-linearly with the treatment rounds and coverage. Additionally, it depicted that for elimination to be achieved within a shorter time period in the general population and within each age group, higher treatment coverage and biannual treatment rounds are more effective.
CONCLUSION
The model captured the dynamics of the STH burdens in vulnerable populations under regular treatment program as elimination is approached. It aided in examining the role of age structure to the persistent STH infections in Kenya. As a result of these findings, we aim to advise the STH control programs on the right mix of strategies needed to achieve faster elimination of the STH infections in Kenya.
AUTHORS/INSTITUTIONS: C. Okoyo, N. Owuor, School of Mathematics, University of Nairobi, Nairobi, KENYA|C. Okoyo, C. Mwandawiro, Eastern and Southern Africa Centre of International Parasite Control, Kenya Medical Research Institute, Nairobi, KENYA|G. Medley, Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, London, UNITED KINGDOM|

CONTROL ID: 3280392
PRESENTER: Takashi Yanagawa
PRESENTER (INSTITUTION ONLY): Kurume University
TITLE: Reviving Fisher's statistical inference by the concept of reproducibility
ABSTRACT BODY:
Abstract Body: Neyman-Pearson’s theory of testing statistical hypotheses is mathematically beautiful and exclusively illustrated in most statistical text books. But its misuse in practice is not negligible. In 2016, the American Statistical Association released a statement against the misuse of statistical significance and P-values, and in March, 2019, Amrhein, V., Greenland, S., McShane, S., and more than 800 signatories called for the entire concept of statistical significance to be abandoned. We appreciate it, in particular, in Biostatistics, the calling for giving up of the teaching of Neyman- Pearson type testing statistical hypotheses; but coming back to Fisher’s statistical inference that instructs to consider whether the observed effect is practically meaningful if P-value ≦0.05 and if meaningful, then decide it significance; and not to consider the data anymore if P-value > 0.5. If this is the case, the two important topics in statistics, one is the sample size determination in designing studies and the other is the optimality of test statistics used to compute P-values could drift on the air, since those topics were developed in the framework of Neyman-Pearson type tests of testing statistical hypothesis.

We show in this paper that the two topics can be settled without using the theory of Neyman-Pearson type test, but by introducing the concept of reproducibility of statistical test results based on P-value. Here, the reproducibility is defined as follows; suppose that the same test is repeated again as the first test, independently to the first test, if and only if the first test results significance. If the same result is obtained as the first test, we call the statistical test result has reproducibility. The reproducibility is natural and directly appeals to scientists whose biggest concern is the reproducibility of their findings.
AUTHORS/INSTITUTIONS: T. Yanagawa, Biostatistics Center, Kurume University, Kurume, JAPAN|K. Omae, Department of Clinical Biostatistics, Kyoto University, Kyoto, JAPAN|

CONTROL ID: 3285731

PRESENTER: Dennis Görlich

PRESENTER (INSTITUTION ONLY): University of Münster

TITLE: Personalized mechanistic mathematical models of acute myeloid leukemia

ABSTRACT BODY:

Abstract Body: Medical research now tends to a more personalized approaches including diagnosis and treatment. This shift of perspective is often not reflected in the data analysis step. We here present an approach based on a mechanistic mathematical model of acute myeloid leukemia (AML). Our primary aim is to develop and validate personalized models to improve therapy decision making and prediction of treatment success. The basic ODE model was published by Stiehl et al. (2014) and a further analysis with respect to combination chemotherapy treatment was published by Banck and Görlich (2019), recently. For our analysis we use individual patient data from the AMLCG2008 trial (NCT01382147, Braess et al. 2018). Patients were treated with a combination chemotherapy, i.e. the S-HAM protocol. The ODE model was adapted to implement the S-HAM treatment course. Personalization is introduced into the model by fitting a subset of model parameters on individual patient data. The main outcome is the leukemic burden i.e. percentage of blast cells. The parametrization comprises the leukemia's proliferation (p_l) and self-renewal (a_l) rates, as well as the effectiveness of the treatment components (k_1,k_2) on a molecular level. We present two personalization approaches: (1) in the leukemia parameters (p_l,a_l) and (2) additionally also in chemotherapy effectiveness (p_l,a_l,k_1,k_2). The ODE model and parameter fitting was implemented in R using the deSolve and DEoptim packages. The combination of model-fitted disease characteristics and clinical (study) data allows to investigate AML in more detail. Overall, our approach contributes to a more personalized analysis strategy in clinical research and opens the path to a model-supported decision making in clinical practice.

References

Banck JC and Görlich D (2019). In-silico comparison of two induction regimens (7+3 vs 7+3 plus additional bone marrow evaluation) in acute myeloid leukemia treatment. BMC Systems Biology 13:18.

Braess J, Amler S, Kreuzer KA et al. (2018) . Sequential High-Dose Cytarabine and Mitoxantrone (S-HAM) versus Standard Double Induction in Acute Myeloid Leukemia—a Phase 3 Study. Leukemia, 32:2558–2571.

Stiehl T, Baran N, Ho AD and Marciniak-Czochra A (2014). Clonal Selection and Therapy Resistance in Acute Leukaemias : Mathematical Modelling Explains Different Proliferation Patterns at Diagnosis and Relapse. J Roy Soc Int 11:20140079.

AUTHORS/INSTITUTIONS: D. Görlich, A. Faldum, Institute of Biostatistics and Clinical Research, University of Münster, Münster, NRW, GERMANY|J.C. Banck, Department of Internal Medicine III, University Hospital Munich, Munich, GERMANY|J. Braess, Department of Oncology and Hematology, Hospital Barmherzige Brüder, Regensburg, GERMANY|

CONTROL ID: 3327359
PRESENTER: Thomas Debray
PRESENTER (INSTITUTION ONLY): University Medical Center Utrecht
TITLE: Propensity-based standardization methods for prediction model research
ABSTRACT BODY:
Abstract Body:
Many prediction models perform worse when applied to new individuals, which may be caused by the use of invalid model parameters or by differences in case-mix distributions. Because it is increasingly common to develop and validate prediction models using data from different studies, settings and populations, the validity of model predictions may be affected by variation in the samples' representativeness of the targeted population.

We present propensity-based standardized measures of discrimination and calibration performance, and discuss how these can be used during external validation to assess model transportability and decide upon updating strategies. Further, when developing a new prediction model, we discuss how standardization of available samples improves the estimation of model parameters and subsequent performance across the included settings and populations.

We evaluate the proposed propensity-based standardization methods in an extensive simulation study where we explore under what circumstances standardization of multiple development samples improves prediction model performance. We also explore to what extent it remains possible to assess model performance in particular settings and populations if all data are used during model development. Results demonstrate that combining and standardizing all available samples for development purposes yields more favorable c-statistics and Brier scores even when some of the included studies have case-mix distributions or predictor-outcome associations that do not properly reflect the target population. When no samples are reserved for external validation, performance assessment requires standardized bootstrap procedures to avoid bias and over-optimism. Finally, we illustrate our methods in a motivating example where data from 13 studies were used to develop and externally validate a prediction model to diagnose deep vein thrombosis in patients suspected of deep vein thrombosis.

In conclusion, propensity score-based standardization might help (i) improve the interpretation of external validation studies of existing prediction models, and (ii) enhance the reproduciblity of newly derived prediction models across different settings and populations.
AUTHORS/INSTITUTIONS: V. de Jong, J. Hoogland, T. Debray, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, NETHERLANDS|L. Nguyen, Department of Public Health, University of Copenhagen, Copenhagen, DENMARK|

CONTROL ID: 3338737

PRESENTER: Christine Wallisch

PRESENTER (INSTITUTION ONLY): Medical University of Vienna, Charité - Universitätsmedizin Berlin

TITLE: Statistical modelling and machine learning: interpretable versus flexible effect estimation?

ABSTRACT BODY:

Abstract Body: While machine learning methods may have the potential for higher accuracy in predictions than classical statistical methods due to their flexibility, their results usually cannot be presented by a transparent formula. Hence, it is often not clear how specific values of the predictors lead to the predictions. We set up a competition between machine learning and classical statistical methods to model risk for cardiovascular events within one year from a health screening. In this competition, we compared these methods in terms of interpretability, variability of predictions, and prediction performance varying the size of the training set from small to very large.

The training set comprised well established risk factors for cardiovascular disease and additional potential risk factors from one million participants of the Austrian health screening program. We applied different modelling approaches to this training set: a logistic regression model with splines and interactions using R, neuronal networks using Matlab, and extreme gradient boosting with regression trees using Python. Subsequently, we divided the training set into subsets (2, 10, 25 and 100 random splits), adapted the modelling strategy of each approach to the sample size, and repeated the adapted modelling procedure in each subset. Using individual conditional expectation and partial dependence plots, we assessed the functional forms of the association of prognostic factors with predictions and their variability across the subsets. Model performance (c-index and calibration) was assessed in a test set comprising 500,000 individuals.

We found that functional forms were relatively similar across the modelling approaches. Extreme gradient boosting using regression trees showed the wiggliest functional forms that could not be rationally explained. The variability of predictions was higher for machine learning methods than for the logistic regression model when the sample size was lowered, pointing to overfitting. However, with large samples machine learning methods slightly outperformed logistic regression in terms of model test performance.

In a setting with large samples and few potential predictors, carefully parameterized machine learning and flexible statistical modelling lead to similar conclusions on the role of predictors. In small samples, however, naively applied machine learning can lead to overfit.

AUTHORS/INSTITUTIONS: C. Wallisch, A. Agibetov, G. Dorffner, D. Dunkler, G. Heinze, Medical University of Vienna, Vienna, AUSTRIA|C. Wallisch, Institute of Biometry and Clinical Epidemiology, Charité - Universitätsmedizin Berlin, Berlin, GERMANY|

CONTROL ID: 3349050

PRESENTER: Kévin Jean

PRESENTER (INSTITUTION ONLY): Conservatoire national des Arts et Métiers, Institut Pasteur, Conservatoire national des Arts et Métiers, Imperial College London

TITLE:

Applying the Self-Controlled Case Series method at the population level: Assessment of vaccination campaigns impact on yellow fever outbreaks in Africa

ABSTRACT BODY:

Abstract Body: Introduction: The self-controlled case series (SCCS) method is a case-only epidemiological study design for which individuals are used as their own control. As all time-invariant confounding are implicitly controlled for, this method is an alternative to classical cohort or case-control study designs when the risk of residual confounding is high. The SCCS method has successfully been applied at the individual level, but never at the population level. Here, we illustrate the use of the SCCS method at the population level by assessing the association between the occurrence of yellow fever (YF) outbreaks and the implementation of preventive mass vaccination campaigns (PMVCs) at the province level in Africa between 2005 and 2018.

Methods: Localization and date of YF outbreaks were identified from international epidemiological records, and information on PMVCs from international coordinators of vaccination activities. Using data collected in case provinces only, the incidence rate ratio (IRR) of YF outbreak between PMVCs-exposed and non-exposed time periods was estimated using a Poisson regression. The sensitivity to a range of assumptions was explored, and the results of the SCCS method were compared to those obtained based on a retrospective cohort study design. We further derived the number of YF outbreaks that have been prevented by PMVCs.

Results: The study sample consisted in 481 provinces in the 34 African countries endemic or at risk for YF. Over the 14-year observation period, we documented the implementations of PMVCs in 125 provinces; and the occurrence of 96 outbreaks in 81 provinces. Irrespectively of coverage achieved, exposure to PMVCs was found to reduce the risk for a province to experience an outbreak by 82% (95% CI: 63%-91%). This estimate was robust across a range of sensitivity analyses. We further conservatively estimated that PMVCs prevented 16.7 provinces (95% CI: 12.2-19.1) from experiencing YF outbreaks. Limitations of our study arise from the sparseness in available data on time-varying risk factors and surveillance quality.

Conclusion: Our estimates provide new evidence supporting PMVCs to prevent the risk of YF outbreak. To our knowledge, our study represents the first application of the SCCS method at the population level for a public health intervention evaluation.

AUTHORS/INSTITUTIONS: H. Raad, K. Jean, Laboratoire MESuRS, Conservatoire national des Arts et Métiers, Paris, FRANCE|H. Raad, M. Hocine, K. Jean, Unité PACRI, Institut Pasteur, Conservatoire national des Arts et Métiers, Paris, FRANCE|K. Gaythorp, A. Hamlet, T. Garske, K. Jean, MRC Centre for Global Infectious Disease Analysis, Department of Infectious Disease Epidemiology, Imperial College London, London, UNITED KINGDOM|H. Whitaker, Statistics, Modelling and Economics Department, National Infection Service, Public Health England, London, UNITED KINGDOM|J. Mueller, EHESP French School of Public Health, Paris/Rennes, FRANCE|M. Hocine, Paris, Conservatoire National des Arts et Métiers, Paris, FRANCE|

CONTROL ID: 3368027

PRESENTER: Sayantee Jana

PRESENTER (INSTITUTION ONLY): Indian Institute of Management, Nagpur

TITLE: Application of multiple testing procedures for identifying multi-morbidities preceding a condition of interest in big health-administrative data

ABSTRACT BODY:

Abstract Body: Background: Multiple testing procedures (MTP) are gaining increasing popularity in various fields of biostatistics, especially in statistical genetics. However, in epidemiologic literature utilizing health-administrative data, there are few studies that applied MTP and discussed their applications and challenges. The objective of our study is to introduce and apply the statistical methods of MTP in association studies and predictive modelling in the context of multi-morbidities using health-administrative data by using MTP to find a subset of variables that have a high likelihood of being associated with the outcome of interest.

Methods: 2600 International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) codes were used to assess the associations between the condition of interest and multi-morbidities on a matched dataset. For an illustration of the statistical tools outlined in this paper, we used a dataset of more than two hundred thousand patients with traumatic brain injury (TBI). McNemar tests were conducted on each of the 2600 ICD-10 codes and appropriate multiple testing adjustments were applied using the Benjamini-Yekutieli procedure. To study the direction of association between the ICD-10 codes and TBI event, odds ratios with 95% confidence intervals were constructed.

Results: The Benjamini-Yekutieli procedure captured 684 ICD-10 codes, out of the 2600, as codes relevant to a TBI event.

Conclusions: Our results illustrate the use of innovative statistical tools on big health-administrative datasets for data mining and dimension reduction.

AUTHORS/INSTITUTIONS: S. Jana, Indian Institute of Management, Nagpur, Nagpur, Maharashtra, INDIA|M. Sutton, A. Colantonio, M. Escobar, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, CANADA|T. Mollayeva, V. Chan, A. Colantonio, Toronto Rehabilitation Institute, University Health Network, Toronto, Ontario, CANADA|A. Colantonio, Rehabilitation Sciences Institute, Faculty of Medicine, University of Toronto, Toronto, Ontario, CANADA|

CONTROL ID: 3372258

PRESENTER: Cornelia Fütterer

PRESENTER (INSTITUTION ONLY): LMU Munich

TITLE: Target Adaptive Individual Loss LASSO (tailLASSO) - Incorporating discriminative power of genes into regularization-based variable selection

ABSTRACT BODY:

Abstract Body:

In personalized medicine, the identification of specific genes that are decisive for a particular cancer type is crucial, but the well-known n<<p issue in high dimensional gene expression data poses a severe challenge for standard classification methods. Common approaches to address this problem are either reducing the number of genes considered to be the best genes with regard to some univariate selection criteria, based on p-values or variance based measures, or to aim at estimation and variable selection simultaneously, in particular by relying on L1-regularized generalized linear regression models via the Least Absolute Shrinkage And Selection Operator (LASSO) (Tibshirani (1996, Royal Statistical Society)). LASSO often leads to very sparse solutions, setting most of the coefficients to exactly zero. A drawback of the LASSO is its over-shrinkage, forcing important covariates to have smaller coefficients. To alleviate this, the adaptive LASSO (Zou (2012, Journal of the American Statistical Association)) allows weights with different penalties for each covariate. This reduces over-shrinkage of important variables and more efficiently shrinks away the noise variables.

We propose Target Adaptive Individual Loss LASSO (tailLASSO), which synthesizes both ideas. In a first step, weights are determined proportional to their univariate performance, where we advocate for considering different clustering evaluation metrics on each gene individually, using the target variable as a grouping factor. The better a variable decomposes into distinct clusters with respect to the target, the higher weight it will receive. This reflects the assumption, motivated by the genetic background, that decisive genes with discriminative power, express differently in each target group.

In a second step, we fit the adaptive Lasso model utilizing these group-sensitive univariate weights. This can be seen as a soft filtering as we do not exclude any covariates before performing regression. Instead, it gives potentially unimportant covariates a higher penalty, while reducing the penalty on the important covariates, mitigating the over-shrinkage.

Applying our method to five well-known microarray benchmark datasets shows promising results: tailLASSO leads in average to sparser solutions while outperforming classical LASSO in terms of prediction accuracy.

AUTHORS/INSTITUTIONS: T. Augustin, Statistics, LMU Munich, Munich, GERMANY|C. Fütterer, M. Nalenz, Statistics, LMU Munich, Munich, GERMANY|

CONTROL ID: 3378960

PRESENTER: Hongwei Sun

PRESENTER (INSTITUTION ONLY): Binzhou Medical University, Shanxi Medical University

TITLE: Trimmed Cox Regression with Elastic Net Penalty for Robust High-dimensional Variables Selection and Outliers Identification

ABSTRACT BODY:

Abstract Body: Penalized Cox regression with Elastic Net penalty (EN) is effective to solve the problem of prognostic factor screening and prediction model building for high-dimensional data. However, experimental or recoding error, sample heterogeneity, cause outliers in the data. These outliers may distort the estimation of EN. If these outliers are not the result of the experimental or recoding errors, that probably means the survival times of these patients relative to its covariates have different response patterns. Identification and analysis of these outliers are likely for us to find new prognosis factor and individualized treatment for them. Outliers are challenging to be detected if they located near to each other, called masking. Robust methods can solve this problem well through trimming. In this article, a penalized maximum trimmed likelihood estimator for EN-type penalized Cox regression is proposed (MPTL-COX). An improved concentration step algorithm was adopted to find the solution of MPTL-COX). Simulation studies illustrated that robust MPTL-COX performed better than EN-type penalized Cox regression in variables selection, outliers detection and prediction in high-dimensional datasets with outliers. A gene expression data from glioma patients were analyzed to illustrate its application.

AUTHORS/INSTITUTIONS: H. Sun, J. Wang, Z. Zhang, School of Public Health and Management, Binzhou Medical University, Yantai, Shandong, CHINA|H. Sun, Q. Gao, T. Wang, Department of Health Statistics, School of Public Health, Shanxi Medical University, Taiyuan, Shanxi, CHINA|