

**Резюмета на публикации  
на Деница Григорова  
на английски език и превод на български език  
за участие в конкурс за доцент по  
професионално направление 4.5. Математика  
(Вероятности и статистика)  
обявен в Държавен вестник бр. 5/17.01.2025 г.**

**Velizar Shivarov, Denitsa Grigorova, and Angel Yordanov.  
Relative Risk of Death in Bulgarian Cancer Patients during  
the Initial Waves of the COVID-19 Pandemic. *Healthcare*,  
11(18), 2023**

**Abstract (English)**

Background: The COVID-19 pandemic has led to millions of documented deaths worldwide, with diverse distribution among countries. Surprisingly, Bulgaria, a middle-income European Union member state, ranked highest in COVID-19 mortality. This study aims to assess whether Bulgarian cancer patients experienced a higher relative risk (RR) of death compared to the general Bulgarian population during the pandemic.

Materials and Methods: Data from the Bulgarian National Statistical Institute and the Bulgarian National Cancer Registry were analyzed to estimate monthly RR of death in cancer patients compared to the general population before and during the first two years of the pandemic. The impact of the COVID-19 waves and predominant SARS-CoV-2 variants on RR was evaluated on various cancer types and age groups using a multiple linear regression approach.

Results: During the COVID-19 waves, both the general population and cancer patients experienced a significant increase in mortality rates. Surprisingly, the RR of death in cancer patients was lower during pandemic waves. The results from the statistical modeling revealed a significant association between the COVID-19 waves and reduced RR for all cancer patients. Notably, the effect was more pronounced during waves associated with the Alpha and Delta variants. The results also showed varying impacts of the COVID-19 waves on RR when we analyzed subsamples of data grouped depending on the cancer type, age and sex.

Conclusions: Despite increased overall mortality in Bulgarian cancer patients during the pandemic, the RR of death was lower compared to the Bulgarian general population, indicating that protective measures were relatively effective in this vulnerable group. This study underscores the importance of implementing

and encouraging preventive measures, especially in cancer patients, to mitigate the impact of future viral pandemics and reduce excess mortality.

### **Резюме (Български)**

**Обосновка:** Пандемията от COVID-19 доведе до милиони документираны смъртни случаи по света с различни относителни честоти в различните страни. Изненадващо, България, страна-членка на Европейския съюз със средни доходи, се класира на първо място по смъртност от COVID-19. Това проучване има за цел да оцени дали пациентите с рак в България са имали по-висок относителен риск (ОР) за смърт в сравнение с общото българско население по време на пандемията.

**Материали и методи:** Анализирани са данни от Национален статистически институт и Национален раков регистър на България, за да се оцени месечният ОР за смърт при пациенти с рак в сравнение с общата популация преди и през първите две години от пандемията. Беше оценено въздействието на вълните на COVID-19 и преобладаващите варианти на SARS-CoV-2 върху ОР за различни видове рак и възрастови групи, като се използва подход на линейна регресия с множество предиктори.

**Резултати:** По време на вълните на COVID-19, както общото население, така и раково болните пациенти са имали значително увеличение на смъртността. Изненадващо, ОР за смърт при пациенти с рак е по-нисък по време на пандемични вълни. Резултатите от статистическото моделиране разкриха значителна връзка между вълните на COVID-19 и намаления ОР при всички ракови пациенти. Трябва да се отбележи, че ефектът е по-изразен по време на вълни, свързани с Алфа и Делта вариантите. Резултатите също показаха различни въздействия на вълните на COVID-19 върху ОР, когато анализирахме подизвадки от данни, групирани в зависимост от вида на рака, възрастта и пола.

**Изводи:** Въпреки повишената обща смъртност при раковите пациенти в България по време на пандемията, ОР за смъртността е по-нисък в сравнение с българската обща популация, което показва, че защитните мерки са били относително ефективни в тази уязвима група. Това проучване подчертава значението на прилагането и насърчаването на превантивни мерки, особено при пациенти с рак, за смекчаване на въздействието на бъдещи вирусни пандемии и намаляване на свръхсмъртността.

Sophia Lazarova, Denitsa Grigorova, Dessislava Petrova-Antonova, and for the Alzheimer's Disease Neuroimaging Initiative. Detection of Alzheimer's Disease Using Logistic Regression and Clock Drawing Errors. *Brain Sciences*, 13(8), 2023

**Abstract (English)**

Alzheimer's disease is an incurable disorder that accounts for up to 70% of all dementia cases. While the prevalence of Alzheimer's disease and other types of dementia has increased by more than 160% in the last 30 years, the rates of undetected cases remain critically high. The present work aims to address the underdetection of Alzheimer's disease by proposing four logistic regression models that can be used as a foundation for community-based screening tools that do not require the participation of medical professionals. Our models make use of individual clock drawing errors as well as complementary patient data that is highly available and easily collectible. All models were controlled for age, education, and gender. The discriminative ability of the models was evaluated by area under the receiver operating characteristic curve (AUC), the Hosmer-Lemeshow test, and calibration plots were used to assess calibration. Finally, decision curve analysis was used to quantify clinical utility. We found that among 10 possible CDT errors, only 3 were informative for the detection of Alzheimer's disease. Our base regression model, containing only control variables and clock drawing errors, produced an AUC of 0.825. The other three models were built as extensions of the base model with the step-wise addition of three groups of complementary data, namely cognitive features (semantic fluency score), genetic predisposition (family history of dementia), and cardiovascular features (BMI, blood pressure). The addition of verbal fluency scores significantly improved the AUC compared to the base model (0.91 AUC). However, further additions did not make a notable difference in discriminatory power. All models showed good calibration. In terms of clinical utility, the derived models scored similarly and greatly outperformed the base model. Our results suggest that the combination of clock symmetry and clock time errors plus verbal fluency scores may be a suitable candidate for developing accessible screening tools for Alzheimer's disease. However, future work should validate our findings in larger and more diverse datasets.

**Резюме (Български)**

Болестта на Алцхаймер е нелечимо заболяване, което представлява до 70% от всички случаи на деменция. Въпреки че разпространението на болестта на Алцхаймер и други видове деменция се е увеличило с повече от 160% през последните 30 години, нивата на нерегистрираните случаи остават критично високи. Настоящата работа има за цел да обърне внимание на недостатъчното откриване на болестта на Алцхаймер, като предложи четири логистични регресионни модела, които могат да се използват като основа за

инструменти за скрининг в общността, които не изискват участието на медицински специалисти. Нашите модели използват индивидуалните грешки при чертане на часовник, както и допълнителни данни за пациенти, които са лесно достъпни и събираеми. Във всички модели бяха включени контролиращи променливи за възраст, образование и пол. Дискриминационната способност на моделите беше оценена чрез площта под работната характеристична крива на приемника (area under the receiver operating characteristic curve (AUC)), а тестът на Hosmer-Lemeshow и калибровъчни графики бяха използвани за оценка на калибрирането. За финал анализът на кривата за взимане на решения беше използван за количествено определяне на клиничната полезност. Установихме, че сред 10 възможни грешки на теста с чертане на часовник само 3 са информативни за откриване на болестта на Алцхаймер. Нашият основен регресионен модел, съдържащ само контролни променливи и грешки в чертането на часовник, дава AUC от 0.825. Другите три модела са изградени като разширения на основния модел с поетапно добавяне на три групи допълващи се характеристики, а именно когнитивни характеристики (оценка за семантична плавност при говор), генетично предразположение (фамилна анамнеза за деменция) и сърдечно-съдови характеристики (индекс на телесна маса, кръвно налягане). Добавянето на оценки за вербална плавност значително подобри AUC в сравнение с основния модел (0.91 AUC). По-нататъшните допълнения обаче не доведоха до забележима разлика в дискриминационната мощност на модела. Всички модели показаха добро калибриране. По отношение на клиничната полезност, надградените модели имат сходни резултати, сравнени помежду си, но значително надминават базовия модел. Нашите резултати предполагат, че комбинацията от две грешки при изчертаване на часовник - грешка при симетрия и грешка за време плюс резултати за вербална плавност може да бъде подходяща разработка за достъпни инструменти за скрининг за болестта на Алцхаймер. Бъдещата работа обаче трябва да потвърди нашите открития, анализирайки по-големи и по-разнообразни набори от данни.

V. Shivarov and D. Grigorova. Dynamic trends in life expectancy and life years lost over five decades in patients from the SEER database with chronic myeloid leukaemia. *Br J Haematol*, 198(2):e23–e25, 07 2022

#### Abstract (English)

This study examines the trends in life expectancy (LE) and life years lost (LYLs) among patients with chronic myeloid leukemia (CML) over five decades using data from the SEER database. The introduction of tyrosine kinase inhibitors (TKIs) after 2000 significantly increased LE, from 3.2 years in the 1970s to 15.9 years in the 2010s, and reduced LYLs. Disparities were observed based on age, sex, race, and insurance status, with younger patients and those with inadequate insurance experiencing worse outcomes. Despite improvements, LE for CML patients remains below that of the general population, highlighting the need for better treatment accessibility and adherence strategies.

#### Резюме (Български)

Това изследване анализира тенденциите в продължителността на живота (LE) и загубените години живот (LYLs) при пациенти с хронична миелоидна левкемия (CML) в продължение на пет десетилетия, използвайки данни от базата данни SEER. Въвеждането на тирозин-киназни инхибитори (TKIs) след 2000 г. значително увеличава LE – средно от 3.2 години през 70-те до 15.9 години през 2010-те, и намалява LYLs. Наблюдавани са различия в зависимост от възраст, пол, раса и застрахователен статус, като по-младите пациенти и тези с неадекватна здравна застраховка показват по-лоши резултати. Въпреки подобренията, LE на пациентите с CML остава по-ниска от тази на общата популация, което подчертава нуждата от по-добър достъп до лечение и стратегии за подобряване на придържането към терапията.

**Hristo Y Ivanov, Denitsa Grigorova, Volker M Lauschke, Branimir Velinov, Kaloyan Stoychev, Gergana Kyosovska, and Peter Shopov. CYP2C19 and CYP2D6 genotypes and metabolizer status distribution in a bulgarian psychiatric cohort. *J Pers Med*, 12(7), July 2022**

**Abstract (English)**

CYP2D6 and CYP2C19 are enzymes of essential significance for the pharmacokinetics of a multitude of commonly used antidepressants, antipsychotics, antiemetics,  $\beta$ -blockers, opioids, antiestrogen, antacids, etc. Polymorphisms in the respective genes are well established as resulting in functional differences, which in turn can impact safety and efficacy. Importantly, the prevalence of genetic CYP2D6 and CYP2C19 variability differs drastically between populations. Drawing on the limited information concerning genotype frequencies in Bulgaria, we here analyzed 742 Bulgarian psychiatric patients predominantly diagnosed with depression and/or anxiety. Specifically, we analyzed frequencies of CYP2C19\*2, \*4 and \*17, as well as of CYP2D6\*2, \*3, \*4, \*5, \*6, \*10 and \*41. In total, 571 out of 742 patients (77%) carried at least one variant which impacts metabolizer status. Overall, 48.6% of the studied individuals were classified as non-normal metabolizers of CYP2D6 with most exhibiting reduced function (38.2% intermediate metabolizers and 6.6% poor metabolizers). In contrast, for CYP2C19, the majority of non-normal metabolizers showed increased functionality (28.9% rapid and 5.5% ultrarapid metabolizers), while reduced activity metabolizer status accounted for 25.6% (23.8% intermediate and 1.8% poor metabolizers). These results provide an important resource to assess the genetically encoded functional variability of CYP2D6 and CYP2C19 which may have significant implications for precision medicine in Bulgarian psychiatry practice.

**Резюме (Български)**

CYP2D6 и CYP2C19 са ензими от съществено значение за фармакокинетиката на множество често използвани антидепресанти, антипсихотици, антиеметици,  $\beta$ -блокери, опиоиди, антиестрогени, антиациди и др. Полиморфизми в съответните гени са добре изучени, с установени функционални различия, които могат от своя страна да повлияят на ефикасността и безопасността. Важно е да се отбележи, че генетичната вариабилност/полиморфност на CYP2D6 и CYP2C19 се различава драстично между популациите. Въз основа на ограничената информация относно генотипните честоти в България, анализирахме 742 български психиатрични пациенти, предимно диагностицирани с депресия и/или тревожност. По-конкретно, ние анализирахме честотите на CYP2C19\*2, \*4 и \*17, както и на CYP2D6\*2, \*3, \*4, \*5, \*6, \*10 и \*41. Общо 571 от 742 пациенти (77%) бяха установени като носители на поне един вариант, който оказва влияние върху състоянието на метаболизатора. Като цяло 48.6% от изследваните лица са класифицирани като аномал-

ни метаболитори по CYP2D6 с повечето проявяващи намалена функция (38.2% междинни метаболитори и 6.6% бавни метаболитори). За разлика от това, за CYP2C19, по-голямата част от аномалните метаболитори показват повишена функционалност (28.9% бързи и 5.5% ултрабързи метаболитори), докато метаболиторите със статус на намалена активност представляват 25.6% (23.8% междинни и 1.8% бавни метаболитори). Тези резултати предоставят важен ресурс за оценка на генетично кодираната функционална разнообразност на CYP2D6 и CYP2C19, която може да има значителни последици за прецизната медицина в българската психиатрична практика.

**Mila Dragomirova, Albena Antonova, Slavena Stoykova, Gergana Mihova, and Denitsa Grigorova. Myopia in Bulgarian school children: prevalence, risk factors, and health care coverage. *BMC Ophthalmology*, 22(1):248, Jun 2022**

#### **Abstract (English)**

**Background:** The prevalence of myopia has increased in recent years, with changes being dynamic and uneven in different regions. The purpose of this study is to evaluate the prevalence of visual impairment caused by myopia in Bulgarian school children, associated risk factors, and health care coverage.

**Methods:** A cross-sectional study among 1401 children (mean age 10.38, standard deviation 2.70) is conducted in three locations in Bulgaria from 2016 to 2020. Refractive error is measured with an auto-refractor in the absence of cycloplegia, the visual acuity is assessed without refractive error correction. A paper-based preliminary questionnaire is used to collect data on previous eye examinations, prescribed optical vision correction, regularity of wearing corrective glasses and risk factors.

**Results:** Children with myopic objective refraction  $\leq -0.75$  D and decimal visual acuity  $\leq 0.8$  on at least one of the eyes are 236 out of 1401 or 16.85%. The prevalence of myopia varies depending on age, geographical location, and school profile. The rate of myopic children in age group 6–10 is 14.2% compared to 19.9% in age group 11–15. The prevalence of myopic children in the urban populations is 31.4% (capital) and 19.9% (medium-sized town) respectively, and only 8.4% in the rural population. Our results show 53% increase of the prevalence of myopia in the age group 11–15 compared to a 2009 report. The analysis of data associated with health care coverage factors of all myopic pupils shows that 71.6% had a previous eye examination, 43.2% have prescription for corrective glasses, 27.5% wear their glasses regularly. Risk factors for higher odds of myopia are gender (female), age (adolescence), and parents with impaired vision. Residence in a small town and daily sport activities correspond to lower odds for myopia. The screen time (time in front of the screen calculated in hours per day) is self-reported and is not associated with increased odds of myopia when accounted for the other risk factors.

**Conclusions:** The prevalence of myopia in this study is higher compared to previous studies in Bulgaria. Additional studies would be helpful in planning adequate prevention and vision care.

#### **Резюме (Български)**

**Обосновка:** Разпространението на миопията се увеличава през последните години, като промените са динамични и неравномерни в различните региони. Целта на това проучване е да се оцени разпространението на зрителните увреждания, причинени от късогледство, при българските деца в училищна възраст, свързаните рискови фактори и покритието на здравните грижи.

Методи: Проведено е проучване с едно наблюдение над обектите във времето сред 1401 деца (средна възраст 10.38 г., стандартно отклонение 2.70) на три местоположения в България в периода от 2016 до 2020 г. Рефракционната грешка се измерва с авторефрактор при липса на циклоплегия, зрителната острота се оценява без корекция на рефракционната грешка. Чрез предварителен въпросник на хартиен носител се събират данни за предишни очни прегледи, предписана оптична корекция на зрението, редовност на носене на коригиращи очила и рискови фактори.

Резултати: Децата с миопична рефракция на обекта  $\leq -0.75D$  и десетична зрителна острота  $\leq 0.8$  на поне едно от очите са 236 от 1401 или 16.85%. Разпространението на миопията варира в зависимост от възрастта, географското местоположение и училищния профил. Дялът на децата с късогледство във възрастовата група 6-10 години е 14.2% в сравнение с 19.9% във възрастовата група 11-15 години. Децата с късогледство съставляват по отношение на градовете съответно 31.4% (столица) и 19.9% (средно голям град) и само 8.4% в селското население. Нашите резултати показват 53% увеличение на разпространението на миопия във възрастовата група 11-15 в сравнение с доклад от 2009 г. Анализът на данните, свързани с факторите на здравно покритие на всички ученици с късогледство, показва, че 71.6% са имали предишен очен преглед, 43.2% имат предписание за коригиращи очила, 27.5% носят очилата си редовно. Рискови фактори за по-високи шансове за късогледство са пол (жена), възраст (юношество) и родители с увредено зрение. Местоживеенето в малък град и ежедневните спортни дейности отговарят на по-ниски шансове за късогледство. Времето пред екрана (времето пред екрана, изчислено в часове на ден) се отчита от самите потребители и не е свързано с повишени шансове за миопия, когато се отчитат другите рискови фактори.

Изводи: Разпространението на късогледство в това проучване е по-високо в сравнение с предишни проучвания в България. Допълнителни проучвания биха били полезни при планирането на адекватна профилактика и грижа за зрението.

**Hristo Y. Ivanov, Branimir Velinov, Gergana Kyosovksa, Denitsa Grigorova, and Peter Shopov. Exploring pharmacogenetic variation in a bulgarian psychiatric cohort. *Folia Medica*, 63(3):365–371, 2021**

#### **Abstract (English)**

**Introduction:** Pharmacogenetics in psychiatry is currently gaining momentum. The efficiency of antipsychotic therapy is often limited by the lack of response and the presence of side effects. Pharmacogenetic variation is probably one of the causative factors for the observed interindividual differences in the response to and the side effects of antipsychotics, which could be addressed and whose negative effects could be avoided or mitigated.

**Aim:** The present study aimed to conduct a comprehensive analysis of the frequency of DRD2 rs1799732, COMT rs4680, MC4R rs489693, and HTR2C rs3813929 in Bulgarian psychiatric patients.

**Materials and methods:** The frequency of genotypes and the alleles of variants DRD2 rs1799732, COMT rs4680, MC4R rs489693, and HTR2C rs3813929 were studied in a cohort of 515 Bulgarian psychiatric patients using the polymerase chain reaction (PCR) method.

**Results:** We found no significant difference between our cohort and the dataset of the 1000 Genomes Project. Moreover, we found that 433 out of 515 patients carried at least one, and 191 out of 515 carried at least two variants which, based on multiple scientific sources with consistent findings, could potentially alter the expected response rate, time to respond and/or risk of side effects to antipsychotic medications.

**Conclusions:** Considering the consistent data about the frequency of these pharmacogenetic variants, testing these genetic variants may prove useful in clinical practice. Further studies regarding the clinical interpretation and frequency distribution in larger cohorts and different populations are warranted.

#### **Резюме (Български)**

**Въведение:** Фармакогенетиката в психиатрията намира все повече приложение в момента. Ефективността на антипсихотичната терапия често е ограничена от липсата на клиничен отговор и наличието на странични ефекти. Фармакогенетичните вариации вероятно са една от причините за наблюдаваните различия между индивидите в отговора и страничните ефекти от антипсихотиците, които могат да бъдат адресирани и чиито отрицателни ефекти могат да бъдат избегнати или смекчени.

**Цел:** Настоящото проучване има за цел да извърши цялостен анализ на честотата на DRD2 rs1799732, COMT rs4680, MC4R rs489693 и HTR2C rs3813929 при български психиатрични пациенти.

**Материали и методи:** Честотите на генотипите и алелите на вариантите DRD2 rs1799732, COMT rs4680, MC4R rs489693 и HTR2C rs3813929 в група от 515 български психиатрични пациенти бяха изследвани чрез метода на

полимеразна верижна реакция.

Резултати: Не открихме значима разлика между нашата кохорта и набора от данни от проекта 1000 генома. Освен това открихме, че 433 от 515 пациенти са носители на поне един, а 191 от 515 са носители на поне два варианта, които, въз основа на множество научни източници с консистентни изводи, имат потенциала да изменят очакваната честота на клинична ефективност, времето за клиничен отговор и/или риска от странични ефекти на антипсихотични лекарства.

Заключения: Имайки предвид консистентните данни за честотата на тези фармакогенетични варианти, тестването им може да се окаже полезно в клиничната практика. Необходими са допълнителни проучвания относно клиничната интерпретация и разпределението на честотите в по-големи кохорти и различни популации.

Miroslava Ivanova, Lilko Dospatliev, Denitsa Grigorova, and Penko Papazov. Evaluation of the Human Health Risk of Wild Edible Mushrooms Consumption from Batak Mountain, Bulgaria. *Iranian Journal of Public Health*, 49(8):1573–1575, Jul. 2020

#### Abstract (English)

This study evaluates the human health risks associated with consuming wild edible mushrooms from Batak Mountain, Bulgaria. The concentrations of heavy metals (Pb, Cd, Ni, Cr, Mn, Co, Cu, and Zn) in nine mushroom species were analyzed, and the health risk index (HRI) was calculated. Samples were collected in 2014 and 2018, and statistical analysis was performed using R 3.4.4. Results indicated that for children aged 1-3 years, HRI values exceeded the safety threshold of 1 for *Amanita caesarea*, *Cantharellus aurora*, and *C. tubaeformis*, primarily due to elevated copper (Cu) and cadmium (Cd) levels. Principal Component Analysis (PCA) demonstrated distinct metal accumulation patterns among species, with two primary groups identified. Hierarchical clustering further revealed similarities in element accumulation among different mushrooms. While Cd levels remained below EU regulatory limits, consumption of certain species by young children may pose health risks.

#### Резюме (Български)

Това изследване оценява рисковете за човешкото здраве, свързани с консумацията на диви ядливи гъби от Баташката планина, България. Анализирани са концентрациите на тежки метали (Pb, Cd, Ni, Cr, Mn, Co, Cu и Zn) в девет вида гъби и е изчислен здравният рисков индекс (HRI). Пробите са събрани през 2014 и 2018 г., а статистическият анализ е извършен с помощта на R 3.4.4. Резултатите показват, че при деца на възраст 1-3 години стойностите на HRI надвишават безопасния праг от 1 за *Amanita caesarea*, *Cantharellus aurora* и *C. tubaeformis* главно поради високи нива на мед (Cu) и кадмий (Cd). Анализът на главните компоненти (principal component analysis) разкрива различни модели на натрупване на метали сред видовете, като се идентифицират две основни групи. Йерархичното клъстериране допълнително показва сходства в натрупването на елементи сред различните гъби. Въпреки че нивата на Cd остават под регулаторните граници на ЕС, консумацията на определени видове от малки деца може да представлява здравен риск.

**D. Grigorova. EM algorithm for maximum likelihood estimation of correlated probit model for two longitudinal ordinal outcomes. *Ann. Sofia Univ., Fac. Math and Inf.*, 104:217–232, 2017**

**Abstract (English)**

Correlated probit models (CPMs) are widely used for modeling of ordinal data or joint analyses of ordinal and continuous data which are common outcomes in medical studies. When we have clustered or longitudinal data CPMs with random effects are used to take into account the dependence between clustered measurements. When the dimension of the random effects is large, finding of the maximum likelihood estimates (MLEs) of the model parameters via standard numerical approximations is computationally cumbersome or in some cases impossible. EM algorithms for CPM for one ordinal longitudinal variable [1] and a joint CPM for one ordinal and one continuous longitudinal variable [2] are recently developed. ECM algorithm for ML estimation of the parameters of a joint CPM for two longitudinal ordinal variables will be presented. The algorithm is applied to estimation of CPM for the longitudinal ordinal outcomes self-rated health and categorized body mass index from the Health and Retirement Study (<http://hrsonline.isr.umich.edu/>, HRS). Results from fitting the model to the data and also results from some simulation studies will be reported.

**Резюме (Български)**

Корелираните probit модели (correlated probit model, CPM) се използват широко за моделиране на категорни данни с наредба на нивата или съвместни анализи на категорни данни с наредба на нивата и непрекъснати данни, които са често срещани променливи от основен интерес в медицинските изследвания. Когато имаме клъстерирани/групирани данни или данни от дългосрочни проучвания, корелираните probit модели със случайни ефекти се използват, за да се вземе предвид зависимостта между клъстерирани/групирани наблюдения. Когато размерността на случайните ефекти е голяма, намирането на оценките на параметрите на модела по метода на максималното правдоподобие чрез стандартни числени приближения е тромаво от изчислителна гледна точка или в някои случаи дори невъзможно. Наскоро са разработени ЕМ алгоритми за корелиран probit модел за една наредена категорна променлива от дългосрочни проучвания [1] и съвместен корелиран probit модел за една категорна променлива с наредба на нивата и една непрекъсната променлива от дългосрочни проучвания [2]. Ще бъде представен ЕСМ алгоритъм за намиране на максимално правдоподобни оценки на параметрите на съвместен корелиран probit модел за две наредени категорни променливи от дългосрочни проучвания. Алгоритъмът се прилага за оценка на корелиран probit модел за две ординални променливи - самооценка на здравето и категоризиран индекс на телесна маса от

Здравно и пенсионно проучване (<http://hrsonline.isr.umich.edu/>). Ще бъдат докладвани резултатите от прилагането на модела към данните, както и резултати от някои симулационни изследвания.

**D. Grigorova and R. Gueorguieva. Correlated probit analysis of repeatedly measured ordinal and continuous outcomes with application to the Health and Retirement Study. *Statistics in Medicine*, 35(23):4202–4225, 2016. sim.6982**

**Abstract (English)**

The Health and Retirement Study was designed to evaluate changes in health and labor force participation during and after the transition from working to retirement. Every 2 years, participants provided information about their self-rated health (SRH), body mass index (BMI), smoking status, and other characteristics. Our goal was to assess the effects of smoking and gender on trajectories of change in BMI and SRH over time. Joint longitudinal analysis of outcome measures is preferable to separate analyses because it allows to account for the correlation between the measures, to test the effects of predictors while controlling type I error, and potentially to improve efficiency. However, because SRH is an ordinal measure while BMI is continuous, formulating a joint model and parameter estimation is challenging. A joint correlated probit model allowed us to seamlessly account for the correlations between the measures over time. Established estimating procedures for such models are based on quasi-likelihood or numerical approximations that may be biased or fail to converge. Therefore, we proposed a novel expectation–maximization algorithm for parameter estimation and a Monte Carlo bootstrap approach for standard errors approximation. Expectation–maximization algorithms have been previously considered for combinations of binary and/or continuous repeated measures; however, modifications were needed to handle combinations of ordinal and continuous responses. A simulation study demonstrated that the algorithm converged and provided approximately unbiased estimates with sufficiently large sample sizes. In the Health and Retirement Study, male gender and smoking were independently associated with steeper deterioration in self-rated health and with lower average BMI.

**Резюме (Български)**

Здравното и пенсионно проучване е предназначено да оцени промените в здравето и работната сила по време и след прехода от работа към пенсиониране. На всеки 2 години участниците предоставят информация за здравната си самооценка (ЗСО), индекс на телесна маса (ИТМ), дали пушат и други характеристики. Нашата цел беше да оценим ефектите от тютюнопушенето и пола върху траекториите на промяна в ИТМ и ЗСО с течение на времето. Съвместният анализ на променливите от дългосрочни проучвания е за предпочитане пред анализа на всяка променлива поотделно, тъй като позволява да се отчете корелацията между характеристиките от основен интерес, да се тестват ефектите на предикторите, докато се контролира грешка от първи тип, и потенциално да се подобри ефективността. Въпреки това, тъй

като ЗСО е наредена категорна променлива, докато ИТМ е непрекъсната, формулирането на съвместен модел и оценка на параметрите е предизвикателство. Съвместният корелиран probit модел ни позволи безпроблемно да отчетем корелациите между променливите във времето. Установените процедури за оценка на такива модели се основават на quasi-правдоподобие или числени приближения, които може да дадат изместени оценки или да не постигнат сходимост. Ето защо ние предлагаме нов ЕМ алгоритъм за оценка на параметрите и Монте Карло bootstrap метод за приближение на стандартните грешки. ЕМ алгоритми са били разглеждани за комбинации от бинарни и/или непрекъснати повторни наблюдения над променливи, въпреки това бяха необходими модификации, за да се приложат към комбинации от наредени категорни и непрекъснати характеристики. Симулационно проучване показва, че алгоритъмът има сходимост и предоставя приблизително неизместени оценки при достатъчно големи размери на извадката. В данните от Здравно и пенсионно проучване мъжкят пол и тютюнопушенето са независимо свързани с по-рязко влошаване на самооценката на здравето и с по-нисък среден ИТМ.

## Литература

- [1] D. Grigorova and R. Gueorguieva. Implementation of the EM algorithm for maximum likelihood estimation of a random effects model for one longitudinal ordinal outcome. *Pliska Stud. Math. Bulgar.*, 22:41–56, 2013.
- [2] D. Grigorova and R. Gueorguieva. Correlated probit analysis of repeatedly measured ordinal and continuous outcomes with application to the Health and Retirement Study. *Statistics in Medicine*, 35(23):4202–4225, 2016. sim.6982.