Received: 8 March 2021 | Revised: 30 August 2021 | Accepted: 2 October 2021

DOI: 10.1002/pst.2186

WILEY

MAIN PAPER

Empirical likelihood confidence interval for sensitivity to the early disease stage

Husneara Rahman | **Yichuan Zhao** | **for the Alzheimer's Disease Neuroimaging Initiative**

Department of Mathematics and Statistics, Georgia State University, Atlanta, Georgia, USA

Correspondence
Yichuan Zhao, Department of Mathematics and Statistics, Georgia State University, Atlanta, GA 30303, USA.
Email: yichuan@gsu.edu

Funding information
National Institute of Biomedical Imaging and Bioengineering; National Institute on Aging; Department of Defense, Grant/Award Number: W81XWH-12-2-0012; National Institutes of Health, Grant/Award Number: U01 AG024904; Alzheimer's Disease Neuroimaging Initiative (ADNI); Simons Foundation, Grant/Award Number: 638679; National Security Agency (NSA), Grant/Award Number: H89230-20-1-0020

Abstract

Disease status can naturally be classified into three or more ordinal stages rather than just being binary stages. Many works have been done for the estimation and inference procedure regarding three ordinal disease stages, which are non-disease, early disease, and full disease stages. The early disease stage can be very important for therapeutic intervention and prevention potentiality. As a diagnostic measure, sensitivity to the early disease stage is often used. In this article, we propose confidence intervals for the sensitivity to early disease stage based on given target specificity for non-disease stage and target sensitivity to full disease stage using both empirical likelihood (EL) and adjusted EL procedures. We compare the performance of the proposed EL confidence intervals with other procedures in our simulation study. The proposed procedures are further applied to Alzheimer's Disease Neuroimaging Initiative data set.

KEYWORDS

confidence interval, early disease stage, empirical likelihood, ordinal disease stages, sensitivity

1 | INTRODUCTION

The performance of diagnostic tests in identifying different diseases is frequently studied in statistics. Usually the considered diseases are of binary class. Many statistical procedures have been developed related to binary diseases such as sensitivity, specificity, positive predictive values, negative predictive values, receiver operating characteristic (ROC) curve, Lorenz curve, Gini index, kappa statistic, and so forth. Among them, ROC curves are studied extensively. Summary measures for ROC curves are area under the curve (AUC), partial AUC and Youden index, and so forth.[1,2,3] Ordinal disease stages with more than two classes are studied for some diseases. Increased awareness on the adversity of different diseases makes the studies of these disease stages, especially the early disease stage, of more interest.

Early disease stage plays a significant role in prevention and therapeutic intervention for the diseases with multiple stages. Being able to correctly identify diseases at early stages increases the chance of an effective treatment. In some situations taking extra caution, therapeutic measures and dietary or behavioral controls help to prevent or to delay the

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

onset of the diseases. In this way, the early detection can prevent extensive complications of diseases and reduce expenditures on expensive medications.

There are several diseases where the studies of early disease stage play a vital role. For example, chronic kidney disease (CKD) has an increased prevalence worldwide. There are five stages to CKD, whereas stage 1 and stage 2 are considered as early stages. Taking some cautious measures like treatment of high-blood pressure, glycaemic control, cessation of smoking, diatery control and exercise, and so forth, during the early stages of CKD can slow down the progression of the disease and reduce the associated risks of cardiovascular events, kidney failure, and death.[4,5] Another disease with high prevalence is type I diabetes. It is a chronic disease, which causes insulin dependence. Over time, the disease can cause lethal complications. The early stage of type I diabetes is being emphasized by researchers for the early diagnosis and prevention purposes. Effective interventions at presymptomatic stages of type I diabetes may delay the progression to symptomatic type I diabetes.[6] Alzheimer's disease (AD) is one of the most common cause of dementia. Detection at an early disease stage appears to be very important to take precautions and to lessen the adverse effect of the disease.[7,8,9] Among other diseases, for which early disease stages are studied are multiple sclerosis, heart disease, different cancers like breast cancer, ovarian cancer or lung cancer, and so forth. Multiple sclerosis develops gradually and starting of treatment in early stage is highly recommended.[10] Early detection increases the survival rate in cancer patient and can save lives for heart diseased patients as well.

An extension of ROC curve for ordinal three stage diseases is ROC surface, which is used for observing the performance of biomarkers in identifying three stages of the disease. Volume under the surface (VUS) is studied as a summary measure. There have been several developments on VUS or ROC surface. For example, nonparametric Mann–Whitney *U*-statistic based variance estimation for the VUS,[11] generalization of three-class classification and ROC hyper-surface construction for ordered multi-class classification problem,[12] kernel smoothing estimation for VUS with an application to liver cancer data,[13] nonparametric calculation of VUS and parametric three-way ROC surface analysis using MATHEMATICA,[14] and so forth. Luo and Xiong[15] developed a useful R package to calculate and analyze three-group classification problem. An application to neuropsychological markers for AD is also illustrated by them. SAS can also be used to estimate VUS with different parametric and nonparametric approaches.[16] Xiong et al.[17] proposed a general linear mixed model for the clustered ordinal diagnostic group. The model incorporates the dependency or correlation on the diagnostic marker. It also allows the use of covariates and missing data.

In the analysis of diagnostic test accuracy, trade-offs exist between the sensitivity and the specificity. By changing cut-off values, we can find the varying combination of sensitivity and specificity. One of the convenient procedures to choose the cut-off or the combination of sensitivity and specificity is by considering a target specificity. Thus, the interval estimation for the sensitivity for given specificity level is of interest. Zhou and Qin[18] and Qin et al.[19] incorporated empirical likelihood (EL) and bootstrap procedures to improve the accuracy of the confidence interval estimation for the sensitivity, whereas Tian[20] proposed interval estimation for the sensitivity for a combination of markers at fixed level of specificity. Dong et al.,[21] Dong and Tian[22] proposed different parametric and nonparametric approaches along with an EL approach to the estimation and the construction of confidence intervals for the sensitivity to the early disease stage at the given specificity level. More nonparametric developments like jackknife EL confidence interval in the related area such as ROC curve, AUC, ordinal dominance curve, VUS and Youden index are observed.[23,24,25,26,27,28]

Since the development of EL by Owen,[29] it has been playing an important role in statistical inference procedures. EL is a nonparametric approach that allows us to use likelihood methods, where no assumption is required for the family of underlying distributions. The empirical distribution of the data plays the central role in such inference procedures. EL has the flexibility of performing different types of statistical analysis and inference procedures, whereas different other nonparametric methods fail to have the flexibility. More details can be found in Owen.[30,31] In this article, we propose two EL approaches for the sensitivity to early disease stage for given target specificity and target sensitivity to full disease stage.

Along with the difficulty of satisfying the underlying distribution assumption, parametric approaches to confidence interval estimation have some other disadvantages. It requires formulation and estimation of variance of the parameter of interest, which becomes complicated for different situations. Existing methods for the construction of confidence interval for the sensitivity to early disease stage either utilize parametric approach, EL approach with scaled chi-square distribution of the test statistic or bootstrap procedures. While the EL approaches with scaled chi-square approximation still requires the calculation of variance of the parameter. The proposed EL approach, which utilizes profiling out the nuisance parameters, does not have such complications. As a data driven and computationally intensive procedure, the profile EL may face difficulty in attaining the convergence in some situations. An improvement to the procedure, the adjusted EL approach was proposed by Chen et al.[32] The procedure confirms to attain the convergence. We propose

an adjusted EL confidence interval for the sensitivity to the early disease stage. We compared the performance of our proposed approaches with an existing EL approach, which use bootstrap variance estimation (ELB), and with percentile bootstrap approach (PB) using simulation studies and real data applications.

The organization of the rest of the article is as follows. We discuss the formulation of a sensitivity to early disease stage, nonparametric approaches to estimate the sensitivity and existing approaches to construct confidence intervals in Section 2. The new method for the proposed confidence interval using profile empirical likelihood (PEL) is described in Section 3. In Section 4, the proposed adjusted EL approach is illustrated. Our extensive simulation studies are carried out in Section 5. Application of the procedures to ADNI data set is given in Section 6. Concluding remarks and discussions are stated in Section 7. The proofs of theorems are provided in the Appendix A.

2 | SENSITIVITY TO THE EARLY DISEASE STAGE

Let X, Y, Z be observations of diagnostic tests from non-diseased, early diseased and diseased individuals, respectively, with corresponding distribution functions $F(X)$, $G(Y)$, and $H(Z)$. For the given thresholds c_1 and c_2, where $c_1 < c_2$, we assume that a subject is considered as diseased if the corresponding test value is greater than c_2, is considered as non-diseased if corresponding test value is smaller than c_1 and as early diseased otherwise (i.e., test value between c_1 and c_2). Let P_1, P_2, and P_3 be correct classification rates for non-disease, early disease and full disease stages, respectively. Then, we define the specificity and the sensitivities as

$$\begin{aligned} P_1 &= P(X < c_1) = F(c_1) = 1 - S_F(c_1), \\ P_2 &= P(c_1 < Y < c_2) = G(c_2) - G(c_1) = S_G(c_1) - S_G(c_2), \\ P_3 &= P(Z > c_2) = 1 - H(c_2) = S_H(c_2). \end{aligned} \tag{1}$$

Here S_F, S_G and S_H are corresponding survival functions of non-diseased, early diseased and fully diseased observations, respectively. Thus we can define the thresholds as $c_1 = F^{-1}(P_1) = S_F^{-1}(1-P_1)$ and $c_2 = H^{-1}(1-P_3) = S_H^{-1}(P_3)$. P_2, the sensitivity to the early disease stage, can be expressed as a function of P_1, the specificity and of P_3, the sensitivity for the full disease stage.

$$P_2 = P_2(P_1, P_3) = G(c_2) - G(c_1) = G\big(H^{-1}(1-P_3)\big) - G\big(F^{-1}(P_1)\big). \tag{2}$$

For the given P_1, and P_3, P_2 can be viewed as the ROC surface on three-dimensional space (P_1, P_3, P_2).

2.1 | Nonparametric estimation

We assume that there are n_1 observations from the non-diseased population as $\{X_i;\ i = 1, ..., n_1\}$ with corresponding empirical distribution function $F_{n_1}(\cdot)$ and empirical survival function $S_{F,n_1}(\cdot)$, respectively. From the early diseased population, there are n_2 observations as $\{Y_j;\ j = 1, ..., n_2\}$ with $G_{n_2}(\cdot)$ as the empirical distribution function and $S_{G,n_2}(\cdot)$ as the empirical survival function, respectively. From the fully diseased population, there are n_3 observations as $\{Z_k;\ j = 1, ..., n_3\}$ with $H_{n_3}(\cdot)$ as the empirical distribution function and $S_{H,n_3}(\cdot)$ as the empirical survival function respectively. We define $S^{-1}_{F,n_1}(r) = F^{-1}_{n_1}(1-r) := X_{((1-r)n_1)}$, $S^{-1}_{G,n_2}(r) = G^{-1}_{n_2}(1-r) := Y_{((1-r)n_2)}$, $S^{-1}_{H,n_3}(r) = H^{-1}_{n_3}(1-r) := Z_{((1-r)n_3)}$, and where $X_{(i)}$, $Y_{(j)}$, $Z_{(k)}$ are corresponding order statistic and a denotes the largest integer smaller than a for any real number a. The above definitions will provide an estimate of c_1 as $F^{-1}_{n_1}(P_1) = S^{-1}_{F,n_1}(1-P_1)$ and an estimate of c_2 as $H^{-1}_{n_3}(1-P_3) = S^{-1}_{H,n_3}(P_3)$. The estimator of P_2 is defined as follows,

$$\widehat{P}_2(P_1, P_3) = G_{n_2}\Big(H^{-1}_{n_3}(1-P_3)\Big) - G_{n_2}\Big(F^{-1}_{n_1}(P_1)\Big). \tag{3}$$

From a probability perspective, an alternative trimmed Mann–Whitney U-statistic is given as follows,

$$\widehat{P}_2(P_1,P_3)=\frac{1}{n_2}\sum_{j=1}^{n_2}V_j(P_1,P_2), \tag{4}$$

where $V_j(P_1,P_2):=I\left\{F_{n_1}^{-1}(P_1)\le Y_j\le H_{n_3}^{-1}(1-P_3)\right\}$.

Reference 33 proposed an adjustment to the estimation by

$$\widehat{P}_2(P_1,P_3)=\frac{\sum_{j=1}^{n_2}V_j(P_1,P_2)+z_{1-\alpha/2}^2/2}{n_2+z_{1-\alpha/2}^2}. \tag{5}$$

2.2 | Existing approaches

An existing approach to construct $100(1-\alpha)\%$ confidence interval is the PB approach. The PB confidence interval is as follows,

$$\left(\widehat{P}_2^{\,b}(\alpha/2),\widehat{P}_2^{\,b}(1-\alpha/2)\right), \tag{6}$$

where $\widehat{P}_2^{\,b}(\alpha)$ is the $(100\alpha)^{th}$ percentile of bootstrap resample estimates.

Dong and Tian[22] proposed another nonparametric approach that utilizes bootstrap variance estimation to the EL approach, where the EL ratio test statistic approximately follows scaled chi-square distribution. Dong and Tian[22] established the following result

$$r_{P_1,P_2,P_3}l(P_2)\to\chi_1^2,$$

where $l(P_2)=-2\log r(P_2)$, $r(P_2)$ is the EL ratio and

$$r_{P_1,P_2,P_3}=\frac{\widehat{\sigma^2_{U_i}}}{n_2\widehat{\sigma^2_{P_2}}}.$$

Then, the $100(1-\alpha)\%$ ELB confidence interval can be constructed as follows

$$I(\alpha)=\left\{P_2(P_1,P_3):r^*_{P_1,P_2,P_3}l(P_2)\le\chi_1^2(\alpha)\right\}, \tag{7}$$

where $\chi_1^2(\alpha)$ is the upper α–quantile of χ_1^2 distribution and an estimate of scale constant is, $r^*_{P_1,P_2,P_3}=\widehat{P}_2\left(1-\widehat{P}_2\right)/\left(n_2\widehat{\sigma}^2_{P_2}\right)$.[22] For the ELB method, $\widehat{\sigma}^2_{P_2}$ is estimated from $b=500$ bootstrap resamples.

3 | EMPIRICAL LIKELIHOOD METHOD FOR THE SENSITIVITY

Let $X_1,\ldots,X_{n_1}$, $Y_1,\ldots,Y_{n_2}$, and $Z_1,\ldots,Z_{n_3}$ be three independent samples from non-diseased, early diseased and fully diseased individuals' test results with distribution functions F, G, and H, respectively. For given values of P_1, P_3, the hypotheses are

$$H_0:\begin{cases}F^{-1}(P_1)=c_1,\\ H^{-1}(1-P_3)=c_2,\\ G(H^{-1}(1-P_3))-G(F^{-1}(P_1))=P_2(P_1,P_3),\end{cases}$$

which are equivalent to

$$H_0: \begin{cases} F(c_1) = P_1, \\ H(c_2) = 1 - P_3, \\ P(c_1 < Y < c_2) = P_2(P_1, P_3). \end{cases}$$

To generalize the hypotheses, we denote

$$\begin{cases} h_1(X, c_1) = h_1(X) = I(X < c_1), \\ h_2(Z, c_2) = h_2(Z) = I(Z > c_2), \\ h_3(Y, c_1, c_2) = h_3(Y) = I(c_1 < Y < c_2). \end{cases}$$

Suppose that $p_1, ..., p_{n_1}$ are the probabilities at $X_1, ..., X_{n_1}$ with all $p_i > 0$ and $\sum_{i=1}^{n_1} p_i = 1$. Similarly, let $q_1, ..., q_{n_2}$ be the probabilities at $Y_1, ..., Y_{n_2}$ for all $q_j > 0$, $\sum_{j=1}^{n_2} q_j = 1$. We also assume $r_1, ..., r_{n_3}$ be the probabilities at $Z_1, ..., Z_{n_3}$ for all $r_k > 0$, $\sum_{k=1}^{n_3} r_k = 1$. The above hypotheses can be expressed as

$$\begin{cases} \sum_{i=1}^{n_1} (h_1(X_i) - P_1) p_i = 0, \\ \sum_{k=1}^{n_3} (h_2(Z_k) - P_3) r_k = 0, \\ \sum_{j=1}^{n_2} \left(h_3\left(Y_j\right) - P_2(P_1, P_3)\right) q_j = 0. \end{cases}$$

We denote $\boldsymbol{h}^T(X, Y, Z) = (h_1(X), h_2(Z), h_3(Y))$, $\boldsymbol{\Theta}^T = (P_1, P_3, P_2(P_1, P_3))$, and $\boldsymbol{O}^T = (0, 0, 0)$. We express the above hypotheses using the vector notation as follows,

$$\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} \left(\mathbf{h}\left(X_i, Y_j, Z_k\right) - \boldsymbol{\Theta}\right) p_i q_j r_k = \boldsymbol{O}, \tag{8}$$

where $(c_1, c_2, P_2(P_1, P_3))$ is the unknown parameter and c_1, c_2 are contained in the function $\boldsymbol{h}(X, Y, Z)$. Then the EL is given by

$$\begin{aligned} L(c_1, c_2, P_2(P_1, P_3)) = \sup_{p_i, q_j, r_k} \Big\{ & \prod_{i=1}^{n_1} \prod_{j=1}^{n_2} \prod_{j=1}^{n_3} \left(p_i q_j r_k\right) : \ p_i > 0, q_j > 0, r_k > 0 \\ & \sum_{i=1}^{n_1} p_i = 1, \sum_{j=1}^{n_2} q_j = 1, \sum_{k=1}^{n_3} r_k = 1, \\ & \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} \left(h\left(X_i, Y_j, Z_k\right) - \Theta\right) p_i q_j r_k = O \Big\}. \end{aligned} \tag{9}$$

The likelihood ratio can be constructed by taking into account of the fact that unconstrained maximum of the likelihood is attained at $p_i = 1/n_1; i = 1, ..., n_1, q_j = 1/n_2; j = 1, ..., n_2, r_k = 1/n_3; k = 1, ..., n_3$. The EL ratio can be expressed as follows,

$$\begin{aligned} R(c_1, c_2, P_2(P_1, P_3)) = \sup_{p_i, q_j, r_k} \Big\{ & \prod_{i=1}^{n_1} \prod_{j=1}^{n_2} \prod_{j=1}^{n_3} \left(n_1 n_2 n_3 p_i q_j r_k\right) : p_i > 0, q_j > 0, r_k > 0 \\ & \sum_{i=1}^{n_1} p_i = 1, \sum_{j=1}^{n_2} q_j = 1, \sum_{k=1}^{n_3} r_k = 1 \\ & \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} p_i q_j r_k \left(h\left(X_i, Y_j, Z_k\right) - \Theta\right) = O \Big\}. \end{aligned} \tag{10}$$

The logarithm of the EL ratio is re-expressed as follows

$$\log R(c_1,c_2,P_2(P_1,P_3)) = \sup_{p_i,q_j,r_k}\Big\{\Big(\sum_{i=1}^{n_1}\log(n_1p_i)+\sum_{j=1}^{n_2}\log\Big(n_2q_j\Big)+\sum_{k=1}^{n_3}\log(n_3r_k)\Big): \\ p_i>0, q_j>0, r_k>0, \sum_{i=1}^{n_1}p_i=1, \sum_{j=1}^{n_2}q_j=1, \sum_{k=1}^{n_3}r_k=1, \\ \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3}p_iq_jr_k\big(h(X_i,Y_j,Z_k)-\Theta\big)=O\Big\}. \tag{11}$$

The Lagrangian expression of the constrained optimization problem is

$$\mathcal{L}\Big(p_i,q_j,r_k\gamma,\eta,\lambda\Big) = \sum_{i=1}^{n_1}\log(n_1p_i)+\sum_{j=1}^{n_2}\log\Big(n_2q_j\Big)+\sum_{k=1}^{n_3}\log(n_3r_k)+\gamma\Big(\sum_{i=1}^{n_1}p_i-1\Big)+\eta\Big(\sum_{j=1}^{n_2}q_j-1\Big)+\zeta\Big(\sum_{k=1}^{n_3}r_k-1\Big) \\ -\lambda^T\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3}\big(h(X_i,Y_j,Z_k)-\Theta\big)p_iq_jr_k.$$

In the above constrained optimization problem, γ, η, ζ, and λ are Lagrange multipliers and $\boldsymbol{\lambda}^T=(\lambda_1,\lambda_2,\lambda_3)$. From the standard Lagrange multipliers method, we have optimal values as

$$p_i=\frac{1}{n_1+\lambda_1(h_1(X_i)-P_1)};\sum_{i=1}^{n_1}p_i=1,$$
$$q_j=\frac{1}{n_2+\lambda_3\big(h_3(Y_j)-P_2\big)};\sum_{j=1}^{n_2}q_j=1,$$
$$r_k=\frac{1}{n_3+\lambda_2(h_2(Z_k)-P_3)};\sum_{k=1}^{n_3}r_k=1.$$

The values of $\boldsymbol{\lambda}$ can be obtained using numerical search procedures. c_1 and c_2 are implicit in the function $\boldsymbol{h}(\cdot)$. The EL ratio function in Equation (11) contains c_1 and c_2, which are not of our interest. We can profile out the two nuisance parameters simultaneously, and obtain the optimal value. Using the profile EL over c_1 and c_2, we have

$$-2\log R(P_2(P_1,P_3)) = -2\max_{c_1,c_2}[\log R(c_1,c_2,P_2(P_1,P_3))] \\ = -2\log R(\widehat{c}_1,\widehat{c}_2,P_2(P_1,P_3)).$$

We establish Wilk's theorem for the EL as follows,

Theorem 1. *We assume that $n_1/n_2\to\rho_1$, $0<\rho_1<\infty$ and $n_3/n_2\to\rho_2$, $0<\rho_2<\infty$. The density functions of F, G, and H are positive and continuous at c_1 and c_2. At the true value $P_{20}(P_1,P_3)$ of $P_2(P_1,P_3)$, $-2\log R(P_{20}(P_1,P_3))\xrightarrow{D}\chi_1^2$ as $\min(n_1,n_2,n_3)\to\infty$,*

where χ_1^2 is chi-square distribution with one degree of freedom.
An EL confidence interval with $100(1-\alpha)\%$ nominal level is constructed as follows

$$I(\alpha)=\big\{P_2(P_1,P_3):-2\log R(P_2(P_1,P_3))\le\chi_1^2(\alpha)\big\},$$

where $\chi_1^2(\alpha)$ is the upper α-quantile of χ_1^2.

4 | ADJUSTED EMPIRICAL LIKELIHOOD METHOD

We have denoted, $\boldsymbol{h}^T(X,Y,Z)=(h_1(X),h_2(Z),h_3(Y))$, we also have denoted that $\boldsymbol{\Theta}^T=(\theta_1,\theta_2,\theta_3)=(P_1,P_3,P_2(P_1,P_3))$ and $\boldsymbol{O}^T=(0,0,0)$. Our hypotheses using vector notation are expressed as follows,

$$\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3}\left(h(X_i,Y_j,Z_k)-\Theta\right)p_iq_jr_k=O. \tag{12}$$

For the adjusted EL, let us denote.
$H_{1i}=H_1(X_i)=h_1(X_i)-\theta_1;i=1,...,n_1$ and $H_{1(n_1+1)}=-a_1\sum_{i=1}^{n_1}(H_{1i}/n_1)=-a_1\overline{H}_{1n_1}$,
$H_{2k}=H_2(Z_k)=h_2(Z_k)-\theta_2;k=1,...,n_3$ and $H_{2(n_3+1)}=-a_3\sum_{k=1}^{n_3}(H_{2k}/n_3)=-a_3\overline{H}_{2n_3}$,
$H_{3j}=H_3(Y_j)=h_3(Y_j)-\theta_3;j=1,...,n_2$ and $H_{3(n_2+1)}=-a_2\sum_{j=1}^{n_2}(H_{3j}/n_2)=-a_2\overline{H}_{3n_2}$.
Here $a_s=\max(1,\log(n_s)/2);s=1,2,3$. The adjusted empirical log likelihood ratio function is as follows,

$$\begin{aligned}\log R^a(c_1,c_2,P_2(P_1,P_3))=\sup_{p_i,q_j,r_k}\Big\{&\Big(\sum\nolimits_{i=1}^{n_1+1}\log((n_1+1)p_i)+\sum\nolimits_{j=1}^{n_2+1}\log\Big((n_2+1)q_j\Big)\\&+\sum\nolimits_{k=1}^{n_3+1}\log((n_3+1)r_k)\Big):p_i>0,q_j>0,r_k>0,\sum\nolimits_{i=1}^{n_1+1}p_i=1,\sum\nolimits_{j=1}^{n_2+1}q_j=1,\\&\sum\nolimits_{k=1}^{n_3+1}r_k=1\sum\nolimits_{i=1}^{n_1+1}\sum\nolimits_{j=1}^{n_2+1}\sum\nolimits_{k=1}^{n_3+1}p_iq_jr_k\left(H_{1i},H_{2k},H_{3j}\right)^T=O\Big\}.\end{aligned} \tag{13}$$

We profile out nuisance parameters c_1 and c_2 and obtain adjusted log-likelihood ratio as follows:

$$\begin{aligned}-2\log R^a(P_2(P_1,P_3))&=-2\max_{c_1,c_2}[\log R^a(c_1,c_2,P_2(P_1,P_3))]\\&=-2\log R^a(\widehat{c}_1,\widehat{c}_2,P_2(P_1,P_3)).\end{aligned}$$

We follow the similar derivations like $-2\log R(P_2(P_1,P_3))$.
The Wilk's theorem for the adjusted EL is established as follows,

Theorem 2. *We assume that* $n_1/n_2\to\rho_1,\ 0<\rho_1<\infty$ *and* $n_3/n_2\to\rho_2,\ 0<\rho_2<\infty$. *The density functions of* F, G, *and* H *are positive and continuous at* c_1 *and* c_2. *At the true value* $P_{20}(P_1,P_3)$ *of* $P_2(P_1,P_3)$, $-2\log R^a(P_{20}(P_1,P_3))\xrightarrow{D}\chi_1^2$ *as* $\min(n_1,n_2,n_3)\to\infty$.

An adjusted EL (AEL) confidence interval with $100(1-\alpha)\%$ nominal level is constructed as follows

$$I^a(\alpha)=\left\{P_2(P_1,P_3):-2\log R^a(P_2(P_1,P_3))\le\chi_1^2(\alpha)\right\}.$$

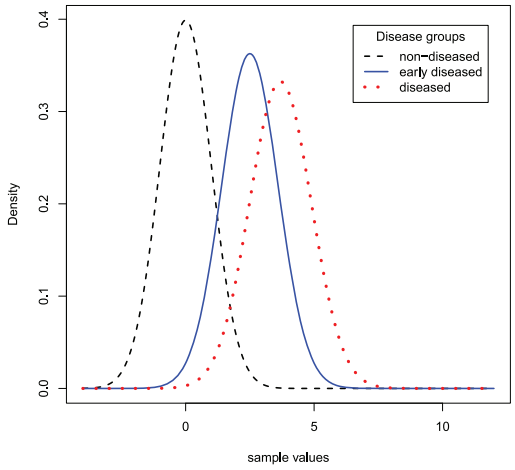
5 | SIMULATION STUDY

To examine the performance of the proposed confidence intervals, we generated data from normal, beta and combination of gamma, log-normal and Weibull distributions with specificity as 0.8 and sensitivity to the full disease stage as 0.8. We calculated confidence intervals using profile empirical likelihood procedure (PEL), the EL procedure using scaled chi-squared distribution with variance estimation from the bootstrap procedure (ELB), the adjusted empirical likelihood procedure (AEL) and the PB procedure. Each simulation was repeated 1000 times and average values of the results were reported. We used 500 bootstrap times for each of the bootstrap procedures. (Figures 1–4)

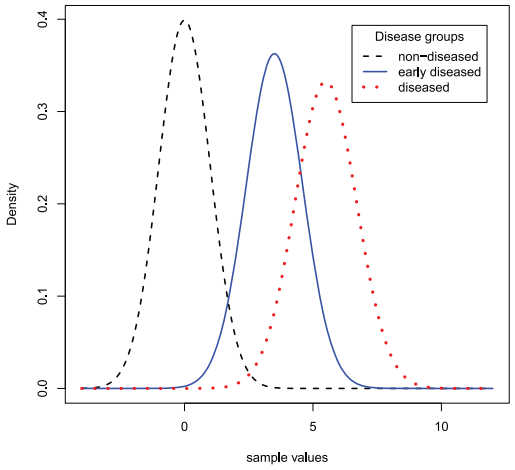
The usage of a smoothing function can improve the coverage accuracy of EL confidence intervals. We used the following function to smooth the indicator function:

$$I_\varepsilon(x,x^*)=I(x\le x^*)=\begin{cases}1, & \text{if } x\le x^*-\varepsilon\\ 0.5-\dfrac{3(x-x^*)}{4\varepsilon}+\dfrac{(x-x^*)^3}{4\varepsilon^3}, & \text{if } x^*-\varepsilon<x\le x^*+\varepsilon\\ 0, & \text{if } x>x^*+\varepsilon,\end{cases}$$

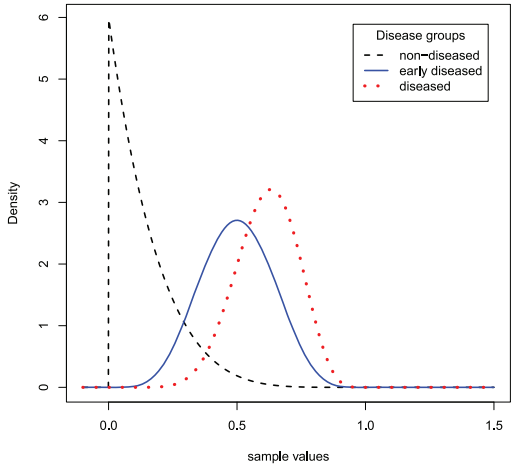
where $\varepsilon>0$. This is the second order kernel smoothing function. More details can be found in Chen and Hall.[34]



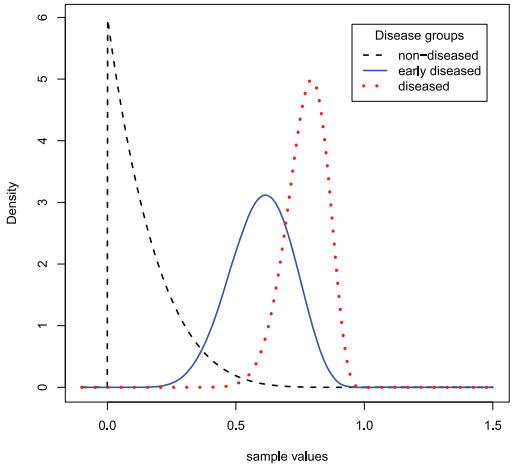
(A) $N(0, 1)$ for non-diseased, $N(2.5, 1.1^2)$ for early diseased, $N(3.69, 1.2^2)$ for diseased, $P_2 = 0.5$.



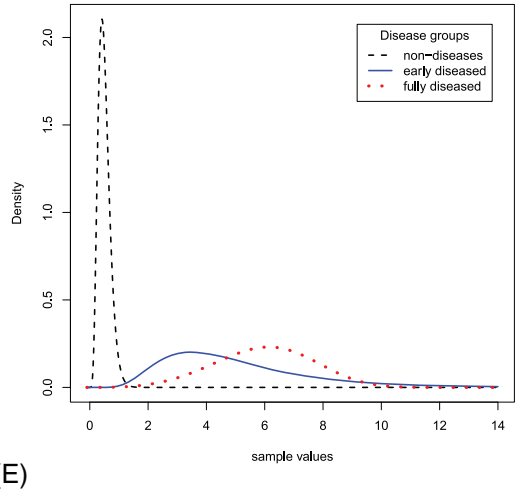
(B) $N(0, 1)$ for non-diseased, $N(3.5, 1.1^2)$ for early diseased, $N(5.5, 1.2^2)$ for diseased, $P_2 = 0.8$.



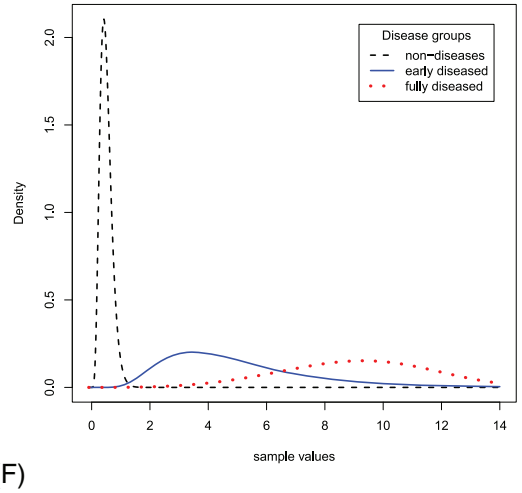
(C) $\beta(1, 6)$ for non-diseased, $\beta(6, 6)$ for early diseased, $\beta(9.6, 6)$ for diseased, $P_2 = 0.5$.



(D) $\beta(1, 6)$ for non-diseased, $\beta(9, 6)$ for early diseased, $\beta(20.4, 6)$ for diseased, $P_2 = 0.8$.



(E) Gamma (6, 12) for non-diseased, log-normal (1.5, 0.5) for early diseased, Weibull (4, 6.6) for diseased, $P_2 = 0.5$.



(F) Gamma (6, 12) for non-diseased, log-normal (1.5, 0.5) for early diseased, Weibull (4, 10) for diseased, $P_2 = 0.8$.

FIGURE 1 Density plots of the biomarkers from different disease groups for the simulation studies

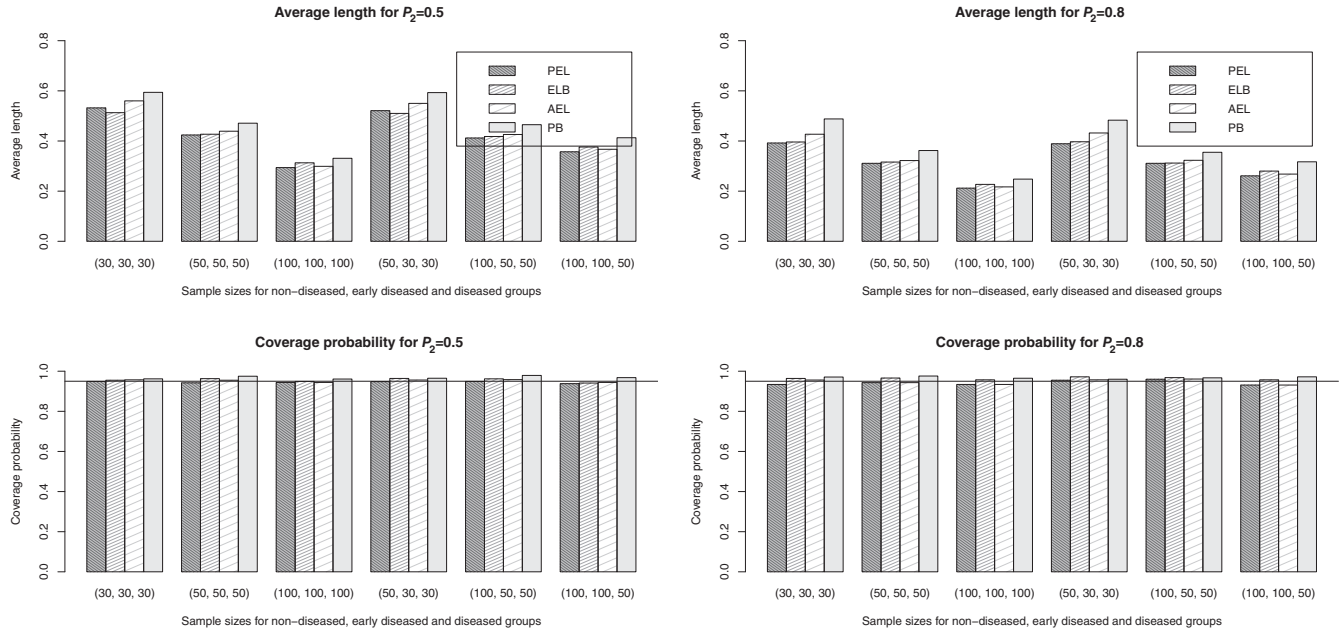


FIGURE 2 Plots of average length (smaller is better) and coverage probability (closer to 0.95 is better) for simulation results from normal distribution (corresponding to Table 1)

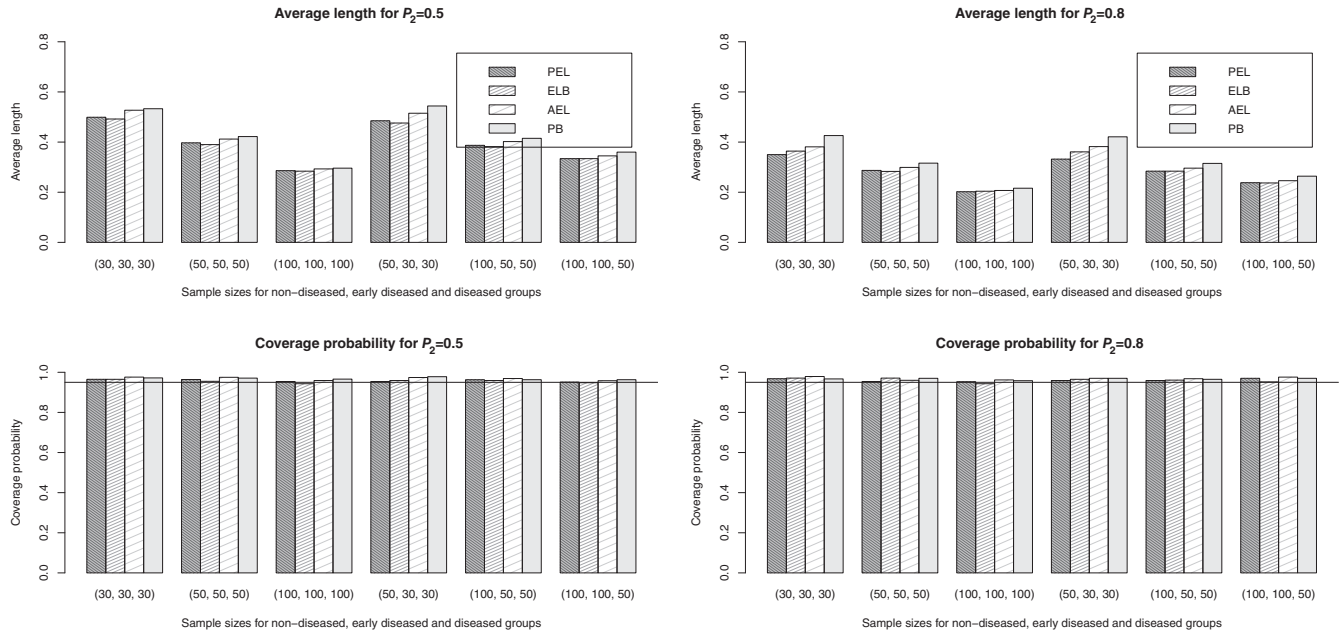


FIGURE 3 Plots of average length (smaller is better) and coverage probability (closer to 0.95 is better) for simulation results from beta distribution (corresponding to Table 2)

For the first set of simulations, we considered using normal distribution (simulation 1). When we considered true value of $P_2 = 0.5$, non-diseased observations were generated from $N(0,1)$, early diseased observations were generated from $N(2.5, 1.1^2)$ and the diseased observations were generated from $N(3.69, 1.2^2)$. For the second set of simulations, we considered the true value of $P_2 = 0.8$. Non-diseased observations were generated from $N(0,1)$, early diseased observations were generated from $N(3.5, 1.1^2)$ and the diseased observations were generated from $N(5.5, 1.2^2)$. The results are displayed in Table 1. The PEL procedure provides coverage probabilities closer to 0.95 level with few instances of under coverage probabilities and no significant over coverage probability. The PEL procedure results in the smallest average length of the interval in most of the situations. The ELB procedure performs better than other procedures in few situations. The procedure provides no considerable under coverage probability, but frequently leads to over

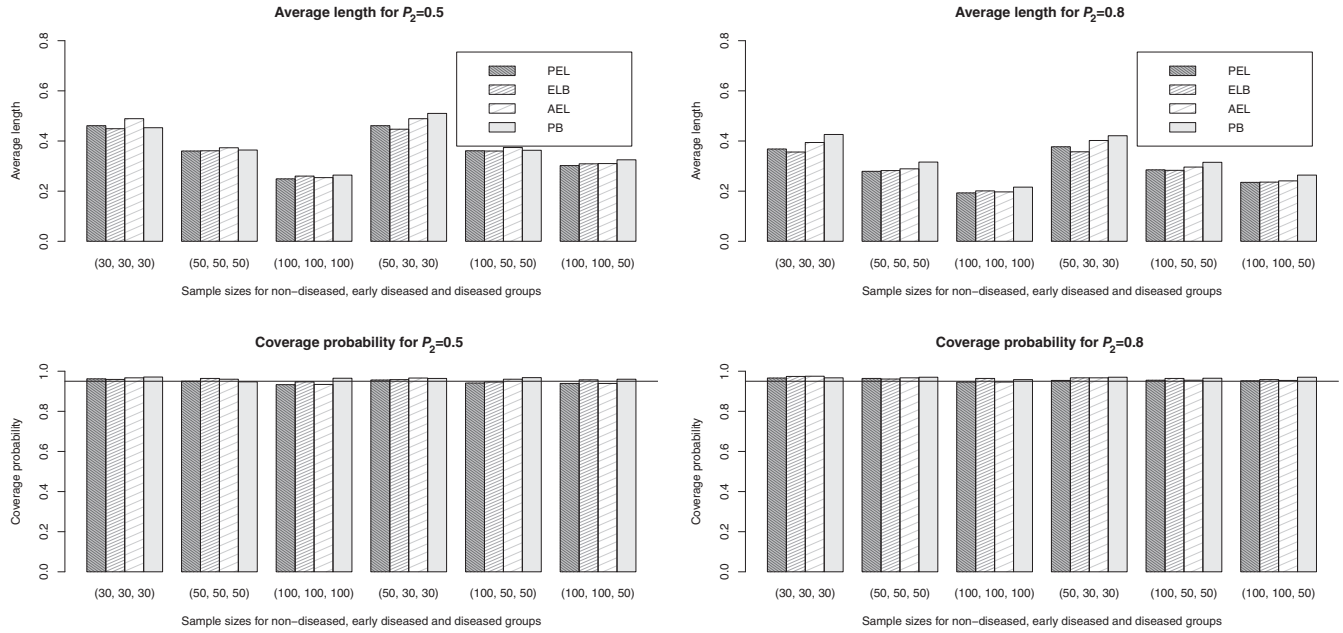


FIGURE 4 Plots of average length (smaller is better) and coverage probability (closer to 0.95 is better) for simulation results from mixture of gamma (non-disease), log-normal (early disease), and Weibull (disease) distributions (corresponding to Table 3)

coverage probability and wider length compared to the PEL procedure. The AEL procedure produces good coverage probability but wider average length and the PB procedure provides over coverage probability and wider average length.

We performed two sets of simulations considering data from beta distribution (simulation 2). For true value of $P_2 = 0.5$, we generated data from $\beta(1,6)$, $\beta(6,6)$, and $\beta(9.6,6)$ for non-diseased, early diseased and fully diseased individuals, respectively. For the true value of $P_2 = 0.80$, we generated data from $\beta(1,6)$, $\beta(9,6)$, and $\beta(20.4,6)$ for non-diseased, early diseased and fully diseased individuals, respectively. The results are presented in Table 2. In terms of coverage probability, the proposed PEL method provides closer to 0.95 level compared to other procedures in most of the situations. When comparing the average length, we observe that the PEL procedure results in wider average length compared to the ELB procedure for $P_2 = 0.5$, whereas for $P_2 = 0.8$, the PEL procedure has better performance. Both the AEL and the PB procedures result in over coverage probability and wider length.

We also considered three different distributions for the three-disease groups (simulation 3). For this set of simulations, the non-diseased individuals were generated from gamma distribution $(G(6,12))$, the early diseased individuals were generated from log-normal distribution $(LN(1.5,0.5))$. For $P_2 = 0.5$, the diseased observations were generated from Weibull distribution with parameters $(a,b) = (4,6.6)$ and for $P_2 = 0.8$, the diseased observations were generated from Weibull distribution with parameters $(a,b) = (4,10)$. The results are displayed in Table 3. We observe that the PEL procedure performs very closely to or better than the ELB procedure in terms of coverage probability in most of the situations for $P_2 = 0.5$, whereas the PEL procedure leads to better coverage probability in all the situations for $P_2 = 0.8$. In terms of average length, we observe some instances where the ELB procedure performs better and some instances where the PEL procedure performs better for $P_2 = 0.5$, whereas PEL performs better in the majority of the situations for $P_2 = 0.8$. For smaller P_2, both ELB and PEL procedures are competitive whereas for higher P_2, the PEL procedure performs better. The AEL procedure performs better than the PB procedure in terms of average length. Both the AEL and the PB procedures lead to over coverage probability in most of the considered situations.

6 | APPLICATION TO REAL DATA

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://adni.loni.usc.edu). The ADNI was launched in 2003 as a public–private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic

TABLE 1 Coverage probability (and average length) of 95% confidence intervals for normally distributed samples

(n_1, n_2, n_3)	PEL[a]	ELB[b]	AEL[c]	PB[d]
$P_2 = 0.5$				
(30, 30, 30)	0.949 (0.532)	0.955 (0.513)	0.957 (0.560)	0.962 (0.594)
(50, 50, 50)	0.942 (0.424)	0.963 (0.427)	0.955 (0.439)	0.975 (0.471)
(100, 100, 100)	0.944 (0.294)	0.949 (0.313)	0.944 (0.299)	0.961 (0.331)
(50, 30, 30)	0.948 (0.521)	0.964 (0.510)	0.956 (0.550)	0.965 (0.593)
(100, 50, 50)	0.949 (0.412)	0.962 (0.418)	0.958 (0.426)	0.979 (0.465)
(100, 100, 50)	0.938 (0.357)	0.941 (0.376)	0.944 (0.367)	0.968 (0.413)
$P_2 = 0.8$				
(30, 30, 30)	0.934 (0.392)	0.964 (0.396)	0.956 (0.427)	0.971 (0.488)
(50, 50, 50)	0.943 (0.311)	0.966 (0.316)	0.943 (0.322)	0.976 (0.362)
(100, 100, 100)	0.934 (0.212)	0.957 (0.227)	0.934 (0.217)	0.965 (0.248)
(50, 30, 30)	0.955 (0.389)	0.972 (0.397)	0.957 (0.432)	0.960 (0.483)
(100, 50, 50)	0.960 (0.311)	0.968 (0.312)	0.961 (0.323)	0.967 (0.355)
(100, 100, 50)	0.931 (0.261)	0.957 (0.280)	0.931 (0.268)	0.972 (0.317)

Note: 95% confidence intervals using.
[a]Empirical likelihood procedure.
[b]Scaled chi-square procedure using bootstrap variance estimation.
[c]Adjusted empirical likelihood procedure.
[d]Bootstrap procedure using percentiles.

resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early AD.

AD is one of the leading causes of dementia. It is a degenerative disease of brain. Three stages of the disease were considered as non-diseased or healthy control, early stage and advanced stage of AD. CDR and MMSE scores along with clinical measures were used to classify the subjects into different disease groups. Subjects with MMSE scores 24–30, CDR of 0, nondepressed, non-MCI, and non-demented were classified as normal. MCI subjects were identified with MMSE scores between 24 and 30, memory complaint, objective memory loss measured by education adjusted scores on Wechsler Memory Scale Logical Memory II, CDR of 0.5, the absence of significant levels of impairment in other cognitive domains, essentially preserved activities of daily living, and an absence of dementia. Patients were classified with mild AD if they have MMSE scores between 20 and 26, CDR of 0.5 or 1.0, and meets NINCDS/ADRDA criteria for probable AD. The details can be found in the study protocol (http://adni.loni.usc.edu/wp-content/uploads/2010/09/ADNI_GeneralProceduresManual.pdf). We considered three prospective biomarkers or diagnostic procedures such as ADAS13 score, FDG and the ratio Tau/AB42 (at 24 month visit). Box plots of the markesr for the three disease stages can be found in Figure 5. We considered the specificity to non-disease stage and sensitivity to the full disease stage as 0.8 and 0.8, respectively. We calculated sensitivities to the early disease stage and constructed confidence intervals for the sensitivities. The results are displayed in Table 4. We observe that ADAS13 score has better sensitivity, which is higher than that of other two biomarkers. The confidence intervals for ADAS13 are significantly higher than zero. Hence, ADAS13 score is a better diagnostic procedure compared to other two biomarkers in identifying the early stage of AD. There is a shorter confidence interval using proposed PEL approach compared to other approaches for ADAS13 score. For other two biomarkers, we observe very close lengths of the confidence intervals using different approaches. Thus, the proposed procedure performs satisfactorily in real data application.

To observe the performance of ADAS13 score in different situations, we considered other two different values for the target specificity to the non-disease stage (P_1) and target sensitivity to the full disease stage (P_3). For the target P_1 as 0.83 and P_3 as 0.83, estimated sensitivity to the early disease stage (P_2) for ADAS13 is 0.416 with the PEL CI as $(0.350, 0.466)$, for FDG, estimated P_2 is 0.007 with 95% PEL CI as $(0.006, 0.081)$, for the ratio Tau/AB42 estimated P_2 is 0.054 with 95% PEL CI $(0.002, 0.151)$. For the target $P_1 = 0.7$ and $P_3 = 0.7$, the estimated P_2 for ADAS13 is 0.628 with 95% PEL CI as (0.582, 0.658), for FDG the estimated sensitivity is 0.290 with 95% PEL CI as (0.216, 0.375), for the ratio Tau/AB42, the estimated sensitivity is 0.352 with 95% PEL CI as (0.202, 0.476). Hence, ADAS13 performs better

TABLE 2 Coverage probability (and average length) of 95% confidence intervals for beta distributed samples

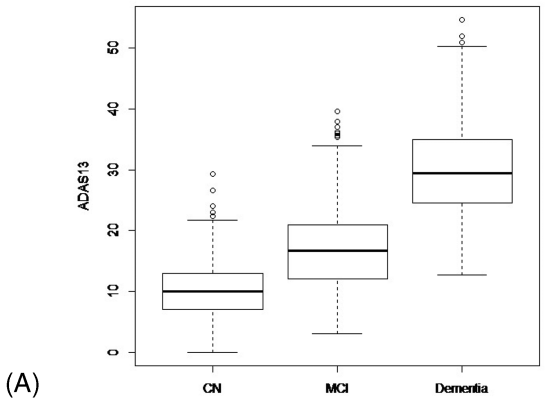
(n_1, n_2, n_3)	PEL[a]	ELB[b]	AEL[c]	PB[d]
$P_2 = 0.5$				
(30, 30, 30)	0.965 (0.499)	0.965 (0.492)	0.976 (0.527)	0.972 (0.533)
(50, 50, 50)	0.964 (0.397)	0.955 (0.390)	0.975 (0.412)	0.971 (0.422)
(100, 100, 100)	0.954 (0.286)	0.943 (0.284)	0.959 (0.293)	0.966 (0.296)
(50, 30, 30)	0.954 (0.485)	0.959 (0.476)	0.974 (0.515)	0.978 (0.544)
(100, 50, 50)	0.963 (0.387)	0.959 (0.382)	0.969 (0.402)	0.963 (0.415)
(100, 100, 50)	0.951 (0.334)	0.948 (0.334)	0.958 (0.345)	0.963 (0.360)
$P_2 = 0.8$				
(30, 30, 30)	0.968 (0.350)	0.971 (0.364)	0.979 (0.381)	0.967 (0.426)
(50, 50, 50)	0.954 (0.287)	0.971 (0.283)	0.960 (0.299)	0.970 (0.316)
(100, 100, 100)	0.953 (0.202)	0.943 (0.204)	0.962 (0.207)	0.958 (0.216)
(50, 30, 30)	0.959 (0.332)	0.965 (0.361)	0.970 (0.382)	0.970 (0.421)
(100, 50, 50)	0.959 (0.284)	0.961 (0.284)	0.968 (0.296)	0.965 (0.315)
(100, 100, 50)	0.970 (0.238)	0.952 (0.237)	0.976 (0.246)	0.970 (0.264)

Note: 95% confidence intervals using.
[a]Empirical likelihood procedure.
[b]Scaled chi-square procedure using Bootstrap variance estimation.
[c]Adjusted empirical likelihood procedure.
[d]Bootstrap procedure using percentiles.

TABLE 3 Coverage probability of (and average length) 95% confidence intervals for samples from mixture of gamma (non-disease), log-normal (early disease), and Weibull (disease) distributions

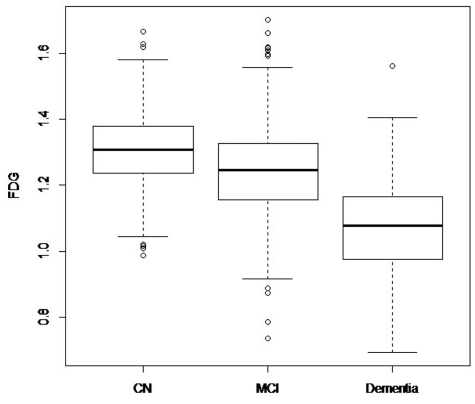
(n_1, n_2, n_3)	PEL[a]	ELB[b]	AEL[c]	PB[d]
$P_2 = 0.5$				
(30, 30, 30)	0.962 (0.461)	0.959 (0.449)	0.967 (0.489)	0.971 (0.453)
(50, 50, 50)	0.950 (0.360)	0.964 (0.361)	0.960 (0.373)	0.947 (0.364)
(100, 100, 100)	0.933 (0.249)	0.947 (0.260)	0.934 (0.254)	0.965 (0.264)
(50, 30, 30)	0.956 (0.461)	0.958 (0.447)	0.966 (0.489)	0.964 (0.510)
(100, 50, 50)	0.941 (0.361)	0.946 (0.360)	0.960 (0.374)	0.968 (0.363)
(100, 100, 50)	0.939 (0.302)	0.957 (0.309)	0.939 (0.310)	0.960 (0.325)
$P_2 = 0.8$				
(30, 30, 30)	0.966 (0.368)	0.974 (0.356)	0.975 (0.394)	0.967 (0.426)
(50, 50, 50)	0.964 (0.279)	0.961 (0.282)	0.967 (0.289)	0.970 (0.316)
(100, 100, 100)	0.946 (0.193)	0.964 (0.201)	0.946 (0.197)	0.958 (0.216)
(50, 30, 30)	0.954 (0.377)	0.967 (0.357)	0.967 (0.402)	0.970 (0.421)
(100, 50, 50)	0.955 (0.285)	0.964 (0.283)	0.955 (0.296)	0.965 (0.315)
(100, 100, 50)	0.952 (0.235)	0.958 (0.236)	0.954 (0.241)	0.970 (0.264)

Note: 95% confidence intervals using.
[a]Empirical likelihood procedure.
[b]Scaled chi-square procedure using Bootstrap variance estimation.
[c]Adjusted empirical likelihood procedure.
[d]Bootstrap procedure using percentiles.



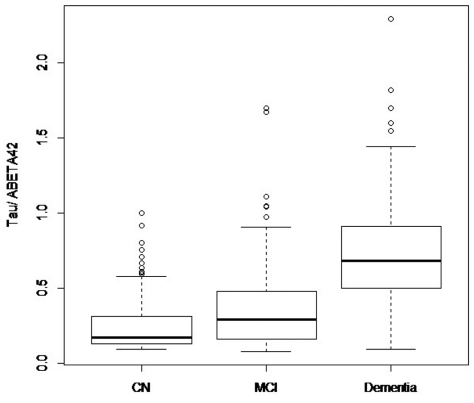
(A)

Box plot of observed ADAS13 scores for the non-disease, early disease and full disease stages of dementia.



(B)

Box plot of observed FDG values for the non-disease, early disease and full disease stages of dementia.



(C)

Box plot of observed Tau and ABeta42 ratio values for the non-disease, early disease and full disease stages of dementia.

FIGURE 5 Box plots for considered biomarkers from three disease groups

TABLE 4 Estimated sensitivities to the early disease stage and 95% confidence intervals for ADNI data

Diagnostic test	Estimated P_2	*PEL*[a]	*ELB*[b]	*AEL*[c]	*PB*[d]
ADAS13	0.479	(0.411, 0.524)	(0.421, 0.537)	(0.411, 0.524)	(0.405, 0.528)
FDG	0.078	(0.015, 0.160)	(0.026, 0.170)	(0.015, 0.160)	(0.005, 0.158)
Tau/AB42	0.097	(0.002, 0.251)	(0.021, 0.252)	(0.002, 0.254)	(0.000, 0.234)

Note: 95% confidence intervals using.
[a]Empirical likelihood procedure.
[b]Scaled chi-square procedure using Bootstrap variance estimation.
[c]Adjusted empirical likelihood procedure.
[d]Bootstrap procedure using percentiles.

compared to other two considered diagnostic procedures to detect the early stage of AD in terms of sensitivity for different target specificities to the non-disease stage and sensitivities to the early disease stage.

7 | DISCUSSION

We propose two EL confidence intervals for sensitivity to the early disease stage. The performance of the proposed PEL and AEL confidence intervals is investigated using simulation studies in different settings. We compare the performance with two other existing nonparametric approaches like EL confidence interval with scaled chi-square

distribution using bootstrap variance estimation (ELB), PB procedure, and so forth. The proposed EL confidence intervals perform well in terms of coverage probability and average length of the intervals in different situations. The PB procedure results in over coverage probability with wider average lengths of the intervals. We compared PB procedure with other EL procedures and the findings support previous conclusions for similar situations. Dong et al.[21] show that the PB procedure performs better in some situations when compared with other parametric or nonparametric procedures. When the EL is considered, Dong et al.[35] have found that the ELB procedure outperforms other existing procedures including the PB method.

The proposed PEL procedure performs better in most of the situations. The proposed AEL procedure provides over coverage probability for beta distribution (simulation 2). For a normal distribution (simulation 1) and for a mixture of gamma, log-normal and Weibull distributions (simulation 3), the AEL procedure performs well. The ELB procedure outperforms other procedures in some instances but provides over coverage probabilities in many situations. We observe better performance for both PEL and ELB procedures compared to other two procedures, whereas, the ELB procedure shows a tendency to provide over coverage probability. The PEL procedure shows very few instances of over coverage probability, some instances of under coverage probability and closer to 0.95 in the majority of the situations. More simulation studies may be carried out to acquire additional evidences.

There are scopes of improving the computational algorithm and reducing the computational cost. More simulation studies considering different scenarios and other distributions can be of future interest. The EL inference for the difference between two sensitivities is also our interest. In the future, we will explore the Bayesian approach to improve coverage probabilities.[36]

ACKNOWLEDGMENTS

We would like to thank the editor Dr. Lilly Yue, an associate editor and two reviewers for their helpful comments and insightful suggestions, which improve the article significantly. Dr. Yichuan Zhao acknowledges the support from the National Security Agency (NSA) Grant (H89230-20-1-0020) and the Simons Foundation (grant number: 638679). Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for NeuroImaging at the University of Southern California.

CONFLICT OF INTEREST

The authors declare there is no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy restrictions.

ORCID

Yichuan Zhao https://orcid.org/0000-0002-5069-1968

REFERENCES

1. Pepe MS. *The Statistical Evaluation of Medical Tests for Classification and Prediction*. Oxford University Press Inc; 2003.
2. Zhou XH, Obuchowski NA, McClish DK. *Statistical Methods in Diagnostic Medicine*. John Wiley & Sons; 2002.

3. Zou KH, Liu A, Bandos AI, Ohno-Machado L, Rockette HE. *Statistical Evaluation of Diagnostic Performance: Topics in ROC Analysis*. Chapman & Hall/CRC; 2011.
4. James MT, Hemmelgarn BR, Tonelli M. Early recognition and prevention of chronic kidney disease. *The Lancet*. 2010;375(9722):1296-1309.
5. Levey AS, Coresh J, Balk E, et al. National kidney foundation practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Ann Intern Med*. 2003;139(2):137-147.
6. Insel RA, Dunne JL, Atkinson MA, et al. Staging presymptomatic type 1 diabetes: a scientific statement of JDRF, the Endocrine Society, and the American Diabetes Association. *Diabetes Care*. 2015;38(10):1964-1974.
7. Sharma N, Singh AN. Exploring biomarkers for Alzheimer's disease. *J Clin Diagn Res*. 2016;10(7):KE01-KE06.
8. Nestor PJ, Scheltens P, Hodges JR. Advances in the early detection of Alzheimer's disease. *Nat Med*. 2004;10(7s):S34-s41.
9. Caroli A, Frisoni G, Initiative ADN. The dynamics of Alzheimer's disease biomarkers in the Alzheimer's disease neuroimaging initiative cohort. *Neurobiol Aging*. 2010;31(8):1263-1274.
10. Kuhlmann T, Lingfeld G, Bitsch A, Schuchardt J, Brück W. Acute axonal damage in multiple sclerosis is most extensive in early disease stages and decreases over time. *Brain*. 2002;125(10):2202-2212.
11. Dreiseitl S, Ohno-Machado L, Binder M. Comparing three-class diagnostic tests by three-way ROC analysis. *Med Decis Making*. 2000; 20(3):323-331.
12. Nakas CT, Yiannoutsos CT. Ordered multiple-class ROC analysis with continuous measurements. *Stat Med*. 2004;23(22):3437-3449.
13. Kang L, Tian L. Estimation of the volume under the ROC surface with three ordinal diagnostic categories. *Comput Stat Data Anal*. 2013;62:39-51.
14. Heckerling PS. Parametric three-way receiver operating characteristic surface analysis using mathematica. *Med Decis Making*. 2001; 21(5):409-417.
15. Luo J, Xiong C. DiagTest3Grp: an R package for analyzing diagnostic tests with three ordinal groups. *J Stat Softw*. 2012;51(3):1-24.
16. Kapasnỳ J, Řezáč M. Three-way ROC analysis using SAS software. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*. 2013;61(7):2269-2275.
17. Xiong C, Luo J, Chen L, et al. Estimating diagnostic accuracy for clustered ordinal diagnostic groups in the three-class case-application to the early diagnosis of Alzheimer disease. *Stat Methods Med Res*. 2018;27(3):701-714.
18. Zhou XH, Qin G. Improved confidence intervals for the sensitivity at a fixed level of specificity of a continuous-scale diagnostic test. *Stat Med*. 2005;24(3):465-477.
19. Qin G, Davis AE, Jing BY. Empirical likelihood-based confidence intervals for the sensitivity of a continuous-scale diagnostic test at a fixed level of specificity. *Stat Methods Med Res*. 2011;20(3):217-231.
20. Tian L. Confidence interval estimation for sensitivity at a fixed level of specificity for combined biomarkers. *J Biopharm Stat*. 2013;23(3):499-512.
21. Dong T, Tian L, Hutson A, Xiong C. Parametric and non-parametric confidence intervals of the probability of identifying early disease stage given sensitivity to full disease and specificity with three ordinal diagnostic groups. *Stat Med*. 2011;30(30):3532-3545.
22. Dong T, Tian L. Confidence interval estimation for sensitivity to the early diseased stage based on empirical likelihood. *J Biopharm Stat*. 2015;25(6):1215-1233.
23. Liu X, Zhao Y. Semi-empirical likelihood inference for the ROC curve with missing data. *J Stat Plan Inference*. 2012;142(12):3123-3133.
24. Yang H, Zhao Y. Smoothed empirical likelihood for ROC curves with censored data. *J Multivar Anal*. 2012;109:254-263.
25. Yang H, Zhao Y. Smoothed jackknife empirical likelihood inference for the difference of ROC curves. *J Multivar Anal*. 2013;115: 270-284.
26. Yang H, Zhao Y. Smoothed jackknife empirical likelihood inference for ROC curves with missing data. *J Multivar Anal*. 2015;140: 123-138.
27. Yang H, Lu K, Zhao Y. A nonparametric approach for partial areas under ROC curves and ordinal dominance curves. *Stat Sinica*. 2017; 27:357-371.
28. Wang D, Tian L, Zhao Y. Smoothed empirical likelihood for the Youden index. *Comput Stat Data Anal*. 2017;115:1-10.
29. Owen AB. Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*. 1988;75(2):237-249.
30. Owen AB. Empirical likelihood ratio confidence regions. *Ann Stat*. 1990;18(1):90-120.
31. Owen AB. *Empirical Likelihood*. Chapman and Hall/CRC; 2001.
32. Chen J, Variyath AM, Abraham B. Adjusted empirical likelihood and its properties. *J Comput Graph Stat*. 2008;17(2):426-443.
33. Agresti A, Coull BA. Approximate is better than "exact" for interval estimation of binomial proportions. *Am Stat*. 1998;52(2):119-126.
34. Chen SX, Hall P. Smoothed empirical likelihood confidence intervals for quantiles. *Ann Stat*. 1993;21(3):1166-1181.
35. Dong T, Kang L, Hutson A, Xiong C, Tian L. Confidence interval estimation of the difference between two sensitivities to the early disease stage. *Biom J*. 2014;56(2):270-286.
36. Cheng Y, Zhao Y. Bayesian jackknife empirical likelihood. *Biometrika*. 2019;106(4):981-988.
37. Zhao Y. Statistical inference on trimmed means, Lorenz curves, and partial area under ROC curves by empirical likelihood method. Theses and Dissertations–Statistics, University of Kentucky 2016; 24.
38. Qin J, Lawless J. Empirical likelihood and general estimating equations. *Ann Stat*. 1994;22(1):300-325.

How to cite this article: Rahman H, Zhao Y, for the Alzheimer's Disease Neuroimaging Initiative. Empirical likelihood confidence interval for sensitivity to the early disease stage. *Pharmaceutical Statistics*. 2022;21(3): 566-583. doi:10.1002/pst.2186

APPENDIX A

A.1 | PROOF OF THEOREM 1

Proof of the theorem follows from Zhao[37] and Owen.[31] At Chapter 11.4, of the book,[31] it is shown that for multi-sample case, $-2\log R$ converges to a quadratic form after utilizing Taylor's approximation. The quadratic form follows a χ^2 distribution with one degree of freedom when there is only one dimensional parameter. Using this generalization and Theorem 3.2 of Owen,[31] when there are k dimensional parameter of interest (θ), $-2\log R \approx Q$ converges to χ^2 distribution with k degrees of freedom under $\theta = \theta_0$.

We first assume a general case and later we discuss particular situation for this article. For the general case, we want to estimate $\theta = (\theta_1, \cdots, \theta_k)$ with hypothesized values being $\theta_0 = (\theta_{10}, \cdots, \theta_{k0})$ and the sample estimate from $(n_1 + n_2 + n_3)$ observations being $\widehat{\theta}_1, \cdots, \widehat{\theta}_k$. We partition first l elements and the remaining $k - l$ elements of these vectors such that $Z = (Z_1, Z_2)$, where $Z_1 = \left(\widehat{\theta}_1 - \theta_1, \cdots, \widehat{\theta}_l - \theta_l\right)$ and $Z_2 = \left(\widehat{\theta}_{l+1} - \theta_{l+1}, \cdots, \widehat{\theta}_k - \theta_k\right)$. We made similar partitioning for variance–covariance matrix of θ. The variance covariance matrix is described as follows:

$$V = \frac{1}{n_1 + n_2 + n_3} \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{k1} & \sigma_{p2} & \cdots & \sigma_{kk} \end{pmatrix} = \frac{1}{n_1 + n_2 + n_3} \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & V_{22} \end{pmatrix},$$

where $\sigma_{i,j}/(n_1 + n_2 + n_3) = \sigma_{j,i}/(n_1 + n_2 + n_3)$ is the covariance of θ_i, θ_j for $i, j = 1, 2, 3$. Similarly, $(n_1 + n_2 + n_3)^{-1} V_{11}$ and $(n_1 + n_2 + n_3)^{-1} V_{22}$ are variance–covariance matrices of Z_1 and Z_2, respectively. Let B be inverse matrix of V and assume,

$$B = (n_1 + n_2 + n_3) \begin{pmatrix} B_{11} & B_{12} \\ B_{12}^T & B_{22} \end{pmatrix}.$$

Then the quadratic form is,

$$\begin{aligned} Q &= (n_1 + n_2 + n_3)(Z_1, Z_2) \begin{pmatrix} B_{11} & B_{12} \\ B_{12}^T & B_{22} \end{pmatrix} \begin{pmatrix} Z_1^T \\ Z_2^T \end{pmatrix} \\ &= (n_1 + n_2 + n_3)\left(Z_1 B_{11} Z_1^T + Z_2 B_{12}^T Z_1^T + Z_1 B_{12} Z_2^T + Z_2 B_{22} Z_2^T\right). \end{aligned} \tag{A1}$$

The quadratic form follows χ_k^2 under $\theta = \theta_0$. When our interest is on a part of the parameters, new constraint becomes $\theta^2 = \theta_0^2$, where $\theta = \left(\theta^1, \theta^2\right)$, and $\theta^1 = (\theta_1, \cdots, \theta_l)$, $\theta^2 = (\theta_{l+1}, \cdots, \theta_k)$. We can optimize the likelihood ratio by profiling out l nuisance parameters. When we use the profile empirical likelihood, we minimize $-2\log R$ for l parameters. Thus to minimize the quadratic form over l nuisance parameters, we may take partial derivatives of Q for the parameters in θ^1 and set that equal to zero,

$$B_{11} Z_1^T + B_{12} Z_2^T = 0.$$

Solving for Z_1, one obtains

$$Z_1^T = -B_{11}^{-1} B_{12} Z_2^T.$$

Plugging this into Equation (A1), we have,

$$\min_{\theta_1,\cdots,\theta_l} Q = (n_1+n_2+n_3)Z_2\left(B_{22}-B_{12}^T B_{11}^{-1}B_{12}\right)Z_2^T.$$

Now V_{22} is the covariance matrix of $\sqrt{(n_1+n_2+n_3)}Z_2$. From the property of the inverse of a square matrix, we have, $V_{22}=\left(B_{22}-B_{12}^T B_{11}^{-1}B_{12}\right)^{-1}$ and it is a $(k-l)$ dimensional matrix. Using the Central Limit Theorem, $\sqrt{(n_1+n_2+n_3)}Z_2$ converges to the multivariate normal distribution with covariance matrix V_{22} under $\theta^2=\theta_0^2$. Thus, $\min_{\theta_1,\cdots,\theta_l} Q \xrightarrow{D} \chi^2_{k-l}$..

Under the conditions of the theorem, we consider $\theta=(c_1,c_2,P_2)$. We profile out two nuisance parameters (c_1,c_2).

$$-2\log R(P_2)=\min_{\theta_1,\theta_2}\{-2\log R(\theta_1,\theta_2,\theta_3)\}=\min_{c_1,c_2}\{-2\log R(c_1,c_2,P_2)\}.$$

Hence, according to previous result and Corollary 5 of Qin and Lawless,[38] under $P_2=P_{20}$ we have.

$$-2\log R(P_{20}) \xrightarrow{D} \chi^2_1.$$

A.2 | PROOF OF THEOREM 2

To sketch the proof of the theorem, we first need to show that the $-2\log R^a(\theta_0)$ approximates to χ^2_3 distribution. Then using the similar argument as in the proof of Theorem 1, we can show that, when we profile out the nuisance parameters, the EL converges to χ^2_1 distribution. Let us denote $l^a(\theta_0)=-2\log R^a(\theta_0)$. For the adjusted empirical likelihood procedure, we have,

$$\begin{aligned} l^a(\theta_0) &= -2\sum_{i=1}^{n_1+1}\log((n_1+1)p_i)+(-2)\sum_{j=1}^{n_2+1}\log\left((n_2+1)q_j\right)+(-2)\sum_{k=1}^{n_3+1}\log((n_3+1)r_k) \\ &= 2\sum_{i=1}^{n_1+1}\log\left(1+\lambda_1^a H_1(X_i)\right)+2\sum_{j=1}^{n_2+1}\log\left(1+\lambda_3^a H_3\left(Y_j\right)\right)+2\sum_{k=1}^{n_3+1}\log\left(1+\lambda_2^a H_2(Z_k)\right) \\ &= Q_1+Q_2+Q_3. \end{aligned}$$

Now we can show that the expression converges to a sum of three independent χ^2_1. Like Chen et al.,[32] we can show that $\lambda_1^a=O_p\left(n_1^{-1/2}\right)$. Here λ_1^a is the solution of the following equation

$$\begin{aligned} 0 &= \frac{1}{n_1}\sum_{i=1}^{n_1+1}\frac{H_{1i}}{1+\lambda_1^a H_{1i}} \\ &= \overline{H}_{1n_1}-\lambda_1^a\widehat{V}_{1n_1}+o_p\left(n_1^{-1/2}\right), \end{aligned}$$

where $\widehat{V}_{1n_1}=n_1^{-1}\sum_{i=1}^{n_1}H_{1i}^2$. Thus $\lambda_1^a\approx\widehat{V}_{1n_1}^{-1}\overline{H}_{1n_1}$ for $n_1\to\infty$.

$$\begin{aligned} Q_1 &= 2\sum_{i=1}^{n_1+1}\left\{\lambda_1^a H_1(X_i)-\left(\lambda_1^a H_1(X_i)\right)^2/2\right\}+o_p(1) \\ &= \sum_{i=1}^{n_1+1}\lambda_1^a H_1(X_i)+o_p(1) \\ &= n_1\overline{H}_{1n_1}\widehat{V}_{1n_1}^{-1}\overline{H}_{1n_1}+o_p(1). \end{aligned}$$

Therefore, Q_1 converges to χ_1^2. Similarly, we can show that Q_2 and Q_3 independently converge to χ_1^2. Thus, $l^a(\theta_0)$ converges to a sum of three independent χ_1^2 variables, that is, χ_3^2 when $\min(n_1, n_2, n_3) \to \infty$. Then, following the same procedure like the proof of Theorem 1, we can show that, $-2\log R^a(\theta) \approx Q$, where Q is the quadratic form of rank three. After profiling out two nuisance parameters, $-2\log R^a(P_{20}(P_1, P_3))$ converges to χ_1^2.