Published in final edited form as:
Stat Med. 2015 February 10; 34(3): 525–538. doi:10.1002/sim.6363.

A likelihood ratio test for nested proportions

Yi-Fan Chen[a,*,†], **Jonathan G. Yabes**[b,c], **Maria M. Brooks**[b,d], **Sonia Singh**[e], and **Lisa A. Weissfeld**[f]

[a]Center for Clinical and Translational Science, University of Illinois at Chicago, Chicago, IL, U.S.A

[b]Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA, U.S.A

[c]Department of Medicine, University of Pittsburgh, Pittsburgh, PA, U.S.A

[d]Department of Epidemiology, University of Pittsburgh, Pittsburgh, PA, U.S.A

[e]Royal Children's Hospital, Brisbane, Australia

[f]Statistics Collaborative Inc., Washington, DC, USA

Abstract

For policy and medical issues, it is important to know if the proportion of an event changes after an intervention is administered. When the later proportion can only be calculated in a portion of the sample used to compute the previous proportion, the two proportions are nested. The motivating example for this work comes from the need to test whether admission rates in emergency departments are different between the first and a return visit. Here, subjects who contribute to the admission rate at the return visit must be included in the first rate and also return, but not vice versa. This conditionality means that existing methods; including the basic test of equality of two proportions, longitudinal data analysis methods, and recurrent event approaches are not directly applicable. Currently, researchers can only explore this question by the use of descriptive statistics. We propose a likelihood ratio test to compare two nested proportions by using the product of conditional probabilities. This test accommodates the conditionality, subject dependencies and cluster effects and can be implemented in SAS PROC NLMIXED allowing for the proposed method to be readily used in an applied setting. Simulation studies showed that our approach provides unbiased estimates and reasonable power. Moreover, it generally outperforms the two-sample proportion z-test, in the presence of heterogeneity, and the Cochran-Mantel-Haenszel test. An example based on readmission rates through an emergency department is used to illustrate the proposed method.

Keywords

likelihood ratio test; nested/conditional proportions; non-linear mixed model

[*]Correspondence to: Yi-Fan Chen, Center for Clinical and Translational Science, University of Illinois at Chicago, Chicago, IL60612, U.S.A.
[†]Email: yfchen2@uic.edu

1. Introduction

The idea of comparing two nested proportions is not uncommon and can be useful for making effective medical and health and public policy decisions. Many questions in the health care setting involve the use of nested proportions, since the population of interest is the group who failed the system in some way. In health services research, this group is often defined as a readmission or second visit group as these subjects tend to use more resources. Examples also occur in other areas, including the comparison of the divorce rates between first marriages and subsequent marriages [1, 2]. In the area of treatment evaluation, it is critical to understand the rates of successful remission for patients who undergo initial chemotherapy and the same chemotherapy undertaken as a second line option, or initial revascularization for significant coronary lesions and repeat revascularization for the same lesion. With greater numbers of registry and large administrative databases available for analysis, similar questions involving an outcome comparison of a group that experiences a defined situation and the subset that again faces that same situation will arise. However, based on our knowledge, a valid statistical testing method is not available for this circumstance when only a small proportion of the original subjects contributes data to the second observation period. This work was motivated by an example comparing the proportion of hospital admissions in the same cohort at two time points, with the first time point being the initial emergency department (ED) visit and the second time point being the second ED visit encounter within 48 hours. Since a large proportion of subjects will not have a second ED visit, it is difficult to compare the hospital admission rate as these proportions will be nested and only those subjects that came to the ED for a second visit are of interest.

When two proportions are collected successively from the same group of subjects, the statistical method should incorporate the dependence within a given subject. A simple two-sample z-test for proportions does not incorporate this dependence. If we frame this as an analysis from a cross-over design, a longitudinal repeated measures analysis, or a recurrent event analysis, we encounter a large number of missing values for the second time point, because many or even most subjects have only one event. Moreover, it is unnecessary to utilize an incomplete data analysis approach, because we are not interested in estimating the outcomes for the second time point for those individuals without any subsequent events (e.g. marriages, diagnoses, recurrences). Thus, repeated measures analyses, including the paired t-test as a special case, are unable to appropriately capture the data structure of the conditionality or nesting in a single model. Although study designs, such as the outcome-dependent sampling [3] or the case-cohort study design [4], collect only a subset of data from a whole cohort, the conditionality is based on a manipulated sampling scheme. This predefined sampling probability is then used in the statistical analysis to draw inferences for the entire population. The subset data in our study, however, is observed, based on the occurrence of the event of interest, and recovering the information for the whole population is not of interest. Without an appropriate and direct statistical method, most researchers explore these issues by presenting descriptive statistics and a simple t-test or stratification and lack the necessary methodology for statistical modeling or testing [1, 2, 5].

The motivating example for this work comes from a clinical and hospital policy question about hospital admission procedures for emergency department visits. In particular, the investigators wanted to compare the hospital admission rates between the first/index emergency department visit and any ED visits that occurred within 48 hours of the first visit in 30 hospitals. Three hypotheses were of interest (1) the hospital admission rate at the index ED visit is different from the hospital admission rate at the 48-hour return ED visit for patients who were admitted at the first visit; (2) the hospital admission rate at the index ED visit is different from the hospital admission rate at the 48-hour return ED visit for patients who were discharged at the first visit; (3) the hospital admission rate at the index ED visit is different from the hospital admission rate among all return visits within 48 hours. Figure 1 depicts this process of coming to the ED for care and admission to the hospital. Considering these hypotheses, three statistical issues should be considered. First, these two rates are nested, because only patients who returned to the ED contribute to the second rate. Second, for those patients who came to the ED twice, the dependence within a subject needs to be taken into account in any analyses. Finally, the dataset consists of 30 hospitals, and so the method should also account for the cluster effect within the hospitals.

The goal of this study is to propose a method for performing the comparison of nested proportions, while simultaneously accounting for the conditional structure, the within-subject dependence, and the hospital cluster effect. This method is based on the conditional probability and the likelihood ratio test. It can be easily extended to more than two layers, to individual patient data, and to alternative outcome variables through the use of other distribution functions. Standard statistical software is available for this type of modeling and we will demonstrate how it can be used to implement the proposed method. The organization of this article is as follow. We describe the notation and our proposed method in Section 2 and evaluate its performance via simulation in Section 3. The ED data described previously is used as an example to demonstrate the method in Section 4. Finally, the discussion is presented in Section 5.

2. Method

Let $j = 1, \ldots, M$ denote the hospital with the total number of hospitals being M, and $i = 1, \ldots, n_j$ be the subject with n_j as the number of subjects within j^{th} hospital. The total number of subjects is $N=\sum_{j=1}^{M} n_j$. We assume a Bernoulli distribution for each admission and return with corresponding proportion parameters described below, and that the hospital cluster effect follows a normal distribution, $b_j \sim N(0, \sigma^2)$. For simplicity, we only illustrate the proposed method in the framework of the first hypothesis, where we compare the admission rates between the index visit and the return visit for patients admitted at the index visit. The same idea is applied to the other two hypotheses and other more scenarios. To generalize the notation, we number the visits, instead of using the index visit and the return visit, to accommodate cases with more than two visits. We denote these visits as follows:

$$A_{1ij}=I\{\text{Admission at visit1 for } i\ th \text{ patient in hospital } j\} \sim Bernoulli(p_1)$$
$$R_{ij}=I\{\text{Return visit for } i\ th \text{ patient in hospital } j\} \sim Bernoulli(p_r)$$
$$A_{2ij}=I\{\text{Admission at visit2 for } i\ th \text{ patient in hospital } j\} \sim Bernoulli(p_2)$$
$$R_{ij}^{(1)}=I\{\text{Return visit for } i\ th \text{ patient in hospital } j|A_{1ij}=1\} \sim Bernoulli\left(p_r^{(1)}\right)$$
$$A_{2ij}^{(1)}=I\{\text{Admission at visit2 for } i\ th \text{ patient in hospital } j|R_{ij}^{(1)}=1\} \sim Bernoulli\left(p_2^{(1)}\right)$$
$$R_{ij}^{(0)}=I\{\text{Return visit for } i\ th \text{ patient in hospital } j|A_{1ij}=0\} \sim Bernoulli\left(p_r^{(0)}\right)$$
$$A_{2ij}^{(0)}=I\{\text{Admission at visit2 for } i\ th \text{ patient in hospital } j|R_{ij}^{(0)}=1\} \sim Bernoulli\left(p_2^{(0)}\right)$$

The corresponding schematic plot depicting the quantities listed above is shown in Figure 2.

2.1. Likelihood function

To incorporate the nested structure, we construct the likelihood function for each subject as the product of conditional probabilities, which correspond to the process of coming to the emergency department for care and admission to the hospital. The likelihood function for each patient is

$$f(A_{1ij}, R_{ij}, A_{2ij}, b_j;\sigma^2)=f(b_j;\sigma^2)f(A_{1ij}|\sigma^2, b_j)f(R_{ij}|\sigma^2, b_j, A_{1ij})f(A_{2ij}|\sigma^2, b_j, A_{1ij}, R_{ij})$$

At the right hand side of this function, the first term represents the likelihood of the hospital cluster effect, and second one is the likelihood of the first visit admission rate for a subject given the cluster effect. Likewise, the third term describes the likelihood of return after knowing patients admission status at the first visit, and the last term is the likelihood of admission at the second visit for a subject given the admission status at the first visit, and the return. This function accounts for both the within-patient fixed-effect dependence between the ED visits and the cluster random-effect within hospital. For the hypotheses $H_0{:}p_1{=}p_2^{(1)}$ vs. $H_A{:}p_1 \neq p_2^{(1)}$, we make two conditional independence assumptions, denoted as $\perp$ as the following:

- Assumption 1. $R|\ (A_1 = 1) \perp R|\ (A_1 = 0)$ and $A_2|\ (A_1 = 1 \text{ and } R = 1) \perp R|\ (A_1 = 0)$
- Assumption 2. $R|\ (A_1 = 1) \perp A_2|\ (A_1 = 0 \text{ and } R = 1)$ and $A_2|\ (A_1 = 1 \text{ and } R = 1) \perp A_2|\ (A_1 = 0 \text{ and } R = 1)$

The first part of assumption (1) refers to the conditional independence of two return status random variables: the return status conditional on being admitted, and the return status conditional on not being admitted (i.e., discharged). The second part of assumption (1), likewise, indicates the conditional independence of two other random variables: the second admission status conditional on being admitted and returning and the return status conditional on not being admitted (i.e., discharged). Overall, assumption (1) states that return status for those who were discharged at the first visit does not affect the return status and readmission for those who were admitted at the first visit. A scenario that could potentially lead to violation of this assumption occurs when there is substantial interaction between admitted and discharged patients. This happens in practice, but is uncommon.

Applying a similar idea, the second assumption relies on the conditional independence of the return/readmission status conditional on being admitted and the readmission status conditional on being discharged. In other words, these assumptions assume that two patients who took different paths are independent and simplify the likelihood function. Therefore, the likelihood function for the target sample related to $p1$ and $p_2^{(1)}$, i.e. paths (1)-(3), can be simply written as

$$\begin{aligned}
&L\left(p_1, p_r^{(1)}, p_2^{(1)}, b_j, \sigma^2\right) \\
&= \prod_{j=1}^{M}\prod_{i=1}^{n_j} f(A_{1ij}, R_{ij}, A_{2ij}, b_j; \sigma^2) \\
&= \prod_{j=1}^{M}\prod_{i=1}^{n_j} f(b_j; \sigma^2) f(A_{1ij} = a_{1ij} | \sigma^2, b_j) f(R_{ij} \\
&= r_{ij} | \sigma^2, b_j, A_{1ij}) f(A_{2ij} \\
&= a_{2ij} | \sigma^2, b_j, A_{1ij}, R_{ij}) \propto \prod_{j=1}^{M}\prod_{i=1}^{n_j} \left\{ \left\{ \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\right.\right.\right. \\
&\qquad \left.\left.\left. - \frac{b_j^2}{2\sigma^2} \right) \right\} \left\{ p_1^{a_{1ij}} (1 - p_1)^{(1 - a_{1ij})} \right\} \\
&\qquad \times \left\{ p_r^{(1) r_{ij}} (1 - p_r^{(1)})^{(1 - r_{ij})} \right\}^{a_{1ij}} \left\{ p_2^{(1) a_{2ij}} (1 - p_2^{(1)})^{(1 - a_{2ij})} \right\}^{a_{1ij} r_{ij}} \Bigg\}
\end{aligned}$$

which only uses the information from the target sample by the restriction of indicators. The hospital random effect is incorporated in the probabilities as

$$\begin{gathered}
p_1 = 1/\{1 + \exp\{-(\beta_{0,1} + b_j)\}\}, \\
p_r^{(1)} = 1/\left\{1 + \exp\left\{-\left(\beta_{0,r}^{(1)} + b_j\right)\right\}\right\}, \text{and} \\
p_2^{(1)} = 1/\left\{1 + \exp\left\{-\left(\beta_{0,2}^{(1)} + b_j\right)\right\}\right\}
\end{gathered}$$

based on a generalized mixed model with a random intercept for a binary outcome when no covariate is included. A direct extension to include covariates for adjusted analyses will be discussed later.

Based on our experience, using the full data could be more efficient than the using a restricted sample when estimating the cluster effect. Therefore, the following likelihood function for the entire sample comprising paths (1)-(6) is preferred.

$$L(p_1, p_r^{(1)}, p_2^{(1)}, p_r^{(0)}, p_2^{(0)}, b_j, \sigma^2) = \prod_{j=1}^{M}\prod_{i=1}^{n_j} \left\{ \left\{ \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(- \frac{b_j^2}{2\sigma^2} \right) \right\} \left\{ p_1^{a_{1ij}} (1-p_1)^{(1-a_{1ij})} \right\} \times \left\{ p_r^{(1) r_{ij}} \left(1 - p_r^{(1)}\right)^{(1-r_{ij})} \right\}^{a_{1ij}} \left\{ p_2^{(1) a_{2ij}} \left(1 - p_2^{(1)}\right)^{(1-a_{2ij})} \right\}^{a_{1ij} r_{ij}} \times \left\{ p_r^{(0) r_{ij}} \left(1 - p_r^{(0)}\right)^{(1-r_{ij})} \right\}^{(1-a_{1ij})} \left\{ p_2^{(0) a_{2ij}} \left(1 - p_2^{(0)}\right)^{(1-a_{2ij})} \right\}^{(1-a_{1ij}) r_{ij}} \right\}$$

2.2. Likelihood ratio test

To test the hypothesis, $H_0{:}p_1{=}p_2^{(1)}$ vs. $H_A{:}p_1 \neq p_2^{(1)}$, the likelihood ratio statistic is

$$\lambda = \frac{L\left(p_1{=}p_2^{(1)}{=}\hat{p}, \hat{p}_r^{(1)}, \hat{p}_r^{(0)}, \hat{p}_2^{(0)}, \hat{b}_j, \hat{\sigma}^2\right)}{L\left(\hat{p}_1, \hat{p}_2^{(1)}, \hat{p}_r^{(1)}, \hat{p}_r^{(0)}, \hat{p}_2^{(0)}, \hat{b}_j, \hat{\sigma}^2\right)} = \frac{\prod_{j=1}^{M}\prod_{i=1}^{n_j} \left\{ \hat{p}^{\{a_{1ij} + a_{2ij} a_{1ij} r_{ij}\}} (1-\hat{p})^{\{1 - a_{1ij} + (1-a_{2ij}) a_{1ij} r_{ij}\}} \right\}}{\prod_{j=1}^{M}\prod_{i=1}^{n_j} \{\hat{p}_1^{a_{1ij}} (1-\hat{p}_1)^{(1-a_{1ij})}\} \left\{ \hat{p}_2^{(1) a_{2ij}} \left(1 - \hat{p}_2^{(1)}\right)^{(1-a_{2ij})} \right\}^{a_{1ij} r_{ij}}}$$

with $-2\log\lambda \sim \chi_1^2$. The test statistic is the result after simplifying the following equation.

$$\lambda = \frac{L\left(p_1{=}p_2^{(1)}{=}\hat{p}, \hat{p}_r^{(1)}, \hat{p}_r^{(0)}, \hat{p}_2^{(0)}, \hat{b}_j, \hat{\sigma}^2\right)}{L\left(\hat{p}_1, \hat{p}_2^{(1)}, \hat{p}_r^{(1)}, \hat{p}_r^{(0)}, \hat{p}_2^{(0)}, \hat{b}_j, \hat{\sigma}^2\right)} = \frac{\prod_{j=1}^{M}\prod_{i=1}^{n_j} \left\{ \begin{array}{c} \left\{ \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp\left(- \frac{b_j^2}{2\hat{\sigma}^2} \right) \right\} \left\{ \hat{p}^{a_{1ij}} (1-\hat{p})^{(1-a_{1ij})} \right\} \\ \left\{ \hat{p}_r^{(1) r_{ij}} \left(1 - \hat{p}_r^{(1)}\right)^{(1-r_{ij})} \right\}^{a_{1ij}} \left\{ \hat{p}^{a_{2ij}} (1-\hat{p})^{(1-a_{2ij})} \right\}^{a_{1ij} r_{ij}} \\ \left\{ \hat{p}_r^{(0) r_{ij}} \left(1 - \hat{p}_r^{(0)}\right)^{(1-r_{ij})} \right\}^{(1-a_{1ij})} \left\{ \hat{p}_2^{(0) a_{2ij}} \left(1 - \hat{p}_2^{(0)}\right)^{(1-a_{2ij})} \right\}^{(1-a_{1ij}) r_{ij}} \end{array} \right\}}{\prod_{j=1}^{M}\prod_{i=1}^{n_j} \left\{ \begin{array}{c} \left\{ \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp\left(- \frac{b_j^2}{2\hat{\sigma}^2} \right) \right\} \left\{ \hat{p}_1^{a_{1ij}} (1-\hat{p}_1)^{(1-a_{1ij})} \right\} \\ \left\{ \hat{p}_r^{(1) r_{ij}} \left(1 - \hat{p}_r^{(1)}\right)^{(1-r_{ij})} \right\}^{a_{1ij}} \left\{ \hat{p}_2^{(1) a_{2ij}} \left(1 - \hat{p}_2^{(1)}\right)^{(1-a_{2ij})} \right\}^{a_{1ij} r_{ij}} \\ \left\{ \hat{p}_r^{(0) r_{ij}} \left(1 - \hat{p}_r^{(0)}\right)^{(1-r_{ij})} \right\}^{(1-a_{1ij})} \left\{ \hat{p}_2^{(0) a_{2ij}} \left(1 - \hat{p}_2^{(0)}\right)^{(1-a_{2ij})} \right\}^{(1-a_{1ij}) r_{ij}} \end{array} \right\}}$$

2.3. Likelihood function and likelihood ratio test in terms of relative risk

From the prospective of researchers, a relative risk may be more interpretable than two separate probabilities. Let $RR^{(1)}{=}p_2^{(1)}/p_1$. Reparameterizing the likelihood equation above gives

$$L\left(p_1, RR^{(1)}, p_r^{(1)}, p_r^{(0)}, p_2^{(0)}, b_j, \sigma^2\right) = \prod_{j=1}^{M}\prod_{i=1}^{n_j}\left\{\left\{\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{b_j^2}{2\sigma^2}\right)\right\}\left\{p_1^{a_{1ij}}(1-p_1)^{(1-a_{1ij})}\right\}\right.$$
$$\times\left\{p_r^{(1)r_{ij}}\left(1-p_r^{(1)}\right)^{(1-r_{ij})}\right\}^{a_{1ij}}\left\{p_1 RR^{(1)a_{2ij}}\left(1-p_1 RR^{(1)}\right)^{(1-a_{2ij})}\right\}^{a_{1ij}r_{ij}}$$
$$\left.\times\left\{p_r^{(0)r_{ij}}(1-p_r^{(0)})^{(1-r_{ij})}\right\}^{(1-a_{1ij})}\left\{p_2^{(0)a_{2ij}}\left(1-p_2^{(0)}\right)^{(1-a_{2ij})}\right\}^{(1-a_{1ij})r_{ij}}\right\}$$

The likelihood ratio test for $H_0 : RR^{(1)} = 1$ vs. $H_A : RR^{(1)} \neq 1$ is

$$\lambda = \frac{L\left(p_1 = \hat{p}, \hat{RR}^{(1)} = 1, \hat{p}_r^{(1)}, \hat{p}_r^{(0)}, \hat{p}_2^{(0)}, \hat{b}_j, \hat{\sigma}^2\right)}{L\left(\hat{p}_1, \hat{RR}^{(1)}, \hat{p}_r^{(1)}, \hat{p}_r^{(0)}, \hat{p}_2^{(0)}, \hat{b}_j, \hat{\sigma}^2\right)}$$
$$= \frac{\prod_{j=1}^{M}\prod_{i=1}^{n_j}\left\{\hat{p}^{\{a_{1ij}+a_{2ij}a_{1ij}r_{ij}\}}(1-\hat{p})^{\{1-a_{1ij}+(1-a_{2ij})a_{1ij}r_{ij}\}}\right\}}{\prod_{j=1}^{M}\prod_{i=1}^{n_j}\left\{\hat{p}_1^{a_{1ij}}(1-\hat{p}_1)^{(1-a_{1ij})}\right\}\left(\hat{p}_1\hat{RR}^{(1)}\right)^{a_{2ij}}\left\{\left(1-\left(\hat{p}_1\hat{RR}^{(1)}\right)\right)^{(1-a_{2ij})}\right\}^{a_{1ij}r_{ij}}}$$

and $-2\log\lambda \sim \chi_1^2$, which gives exactly the same result as using the probabilities.

2.4. Estimation in and properties of the proposed method

The basis of the proposed method is the conventional likelihood ratio test. Our method therefore inherits all the properties of likelihood ratio test. When no random effects are identified, the maximum likelihood estimator (MLE) is valid for estimation. When a random effect is included, the MLE with adaptive Gauss-Hermite quadrature is used. These estimators have all of the nice statistical properties of the MLE and can be easily calculated in SAS (SAS Institute Inc., Cary, NC, USA) by using PROC NLMIXED. Sample code is provided in the Appendix.

3. Simulation study

We evaluated the performance of the proposed method via simulation in comparison to the two-sample z-test for proportion [6] and the Cochran-Mantel-Haenszel test [7, 8]. As to the two-sample z-test for proportion, the admitted proportions at the first and at the return visits are directly compared and the relative risk is simply the ratio of two proportions [9]. For the Cochran-Mantel-Haenszel test, the relative risk is a ratio of expected proportions summarized across strata [7, 8] and the relative risk is compared to the null value of 1. These two naive tests ignore the complexity and dependence inherent in the data structure, although the Cochran-Mantel-Haenszel test considers the stratification by hospital. We investigated the bias of the estimated parameters of interest and the corresponding standard error estimators. We also examined the power and type I error of the associated test. We

study the properties of the proposed method under varying sample sizes, effect sizes, and clustering effects. Sample sizes were chosen to be 45,000 and 500 to represent, respectively, large and moderate-size datasets. Effect sizes were varied by specifying the true difference between p_1 and $p_2^{(1)}$ at 3 levels: severe, mild, and none. We studied 6 levels cluster effects: 0, 0.001, 0.01, 0.1, 0.5 and 1. These numbers were chosen as they resemble the data features in the example that will be described in detail later. The total number of hospitals, *M*, was set to 30 and 1000 datasets were simulated for each scenario. Details of the simulation settings considered are shown in Table 1.

Tables 2 and 3 show the simulation results for $N = 45{,}000$ and $N = 500$ respectively. The estimators using the proposed method were unbiased in most of the probability parameters considered. The estimated standard errors were close to standard deviations, which increased from the first visit to return and to the second visit due to the declining sample size. Non-convergence, a common problem when using PROC NLMIXED to fit models with random effects, was also encountered in some of the datasets. Note that when the cluster correlation is very small (e.g., under 0.001 with a large sample size or 0.01 with a moderate sample size), the proposed method tended to be unstable and provided biased estimation for relative risks (results not shown). One solution for this is to remove the cluster effect from the likelihood function. The results for these cases were specifically presented in the tables in the boldface. In contrast to the proposed method, the two-sample proportion z-test gave biased relative risks when a cluster effect existed and the severity of the bias expanded as the effect increased from 0.1. When the cluster effect is more than 0.5 for a large sample or 1 for a moderate sample, the standard error was smaller than the estimated standard deviation. For the Cochran-Mantel-Haenszel test, the bias soared with increasing sample size, cluster effect and true difference. The deflated standard error was only observed as the sample size is large and the difference exists, and the cluster effect is more than 0.5. The estimated standard deviations and standard errors under theses ordinary tests were for the relative risks in the natural log scale which approximately follow the normal distribution. However, the standard deviations and standard errors from the proposed method were directly for the relative risks following the normal assumption under MLE. Hence, they are not directly comparable between the proposed methods and two ordinary tests.

With respect to the power, the two-sample proportion z-test exhibited inflated type I error when no difference truly existed. The type I error was 33% to 100% when the cluster effect was more than 0.1 and the sample size was large, and was 44% to 84% when the cluster effect was bigger than 0.5 and the sample size was moderate. In the same scenarios, the proposed method provided reasonable error rates, around 4-6%, regardless of the sample size. Meanwhile, the powers from the Cochran-Mantel-Haenszel test are slightly lower and higher than 5% for large and moderate samples respectively. When the difference increased to a mild or large level in a large dataset, the proposed method and two ordinary tests performed equally well with almost 100% power. However, as the sample size decreased, the proposed method only had a 29% to 32% and 78% to 81% power to discover a mild and severe difference respectively, while the z-test had a stronger power when the random effect is large. Nonetheless, the proportion z-test still gave highly biased estimated relative risks.

In the same scenarios, the Cochran-Mantel-Haenszel test had much lower power compared to the other two methods.

4. Example

Our example is a retrospective cohort study of 30 children's hospital emergency departments using the Pediatric Health Information System (PHIS) 2009 database. The data contains inpatient and ED data for patients < 18 years of age, and we studied the number of index visits, return visits, and admitted and discharged rate for each visit. To be specific, the index visit is defined as a visit without a prior ED visit within 48-hrs, while the return visit is a visit within 48 hours after the index visit. The admission rate is the proportion of subjects admitted to the hospital among the total ED visits for the index visit. The readmission rate is the proportion of the subjects admitted to the hospital among those who return. Both admission rates are either for patients who are discharged or admitted at the index ED visit. The clinical question of interest is to investigate the quality of ED and inpatient care by comparing the proportions of admission rates between the index and the return visits.

Considering the relative risk by using the index visit as the reference, i.e., $RR^{(1)}=p_2^{(1)}/p_1$, the following three hypotheses are of interest: (1) $H_0 : RR^{(1)} = 1$ vs. $H_A : RR^{(1)} \neq 1$, (2) $H_0 : RR^{(0)} = 1$ vs. $H_A : RR^{(0)} \neq 1$, and (3) $H_0 : RR = 1$ vs. $H_A : RR \neq 1$. Two rationales are behind these comparisons to answer the clinical questions. Firstly, the admission rate to a hospital at the index visit, i.e., p_1, is treated as the general admitted probability for any person in the population. Secondly, it is assumed that patients are fully recovered once they leave the hospital from their index visit. However, this assumption may not be appropriate for patients with severe and chronic diseases; we ignored this issue in the analysis due to the lack of disease information in the data. Nevertheless, given these two conditions, the admission rate at the second visit should be similar to the first one as these two visits were assumed to be independent. Any significant differences between these two sequential admission rates might indicate certain quality issues regarding to the caring in the hospital.

In the data, there were 1,847,465 total index ED visits, with an 11.69% (n=215,906) admission rate. For patients admitted and discharged at the index visit, the return rates were 2.22% (n=4,792), and 3.42% (n=55,745) respectively, and the overall return rate was 3.28% (n=60,537). Among those who returned, the admission rates were 46.39% (n=2,223), 20.20% (n=11,263), and 22.28% (n=13,486) for patients admitted, discharged at the index visit, and pooled patients respectively. The admission at the 48-hr return visit was more likely for both patients discharged and admitted at the index visit. To test if this difference was statistically significant, we used the proposed method, the two-sample proportion z-test and the Cochran-Mantel-Haenszel test. Because the cluster effect (<0.0001) in the data set was extremely small, we did not include any random effects in the likelihood function of the proposed method, as what the simulation study suggested. In Table 4, both of the proposed method and the two-sample proportion z-test provided similar estimations, while the Cochran-Mantel-Haenszel test gave slightly different results. This echoes the simulation study with a large sample and a severe difference and results in similar conclusions that the difference in rates was statistically significant. The significant differences observed deserve further evaluation to identify a multitude of underlying quality of care issues including

physician behavior, incomplete medical treatment, missed diagnoses, or failure of adequate discharge planning.

5. Discussion

This study proposed a likelihood ratio test for comparing nested proportions. Statistically, the method could accommodate a data structure with conditionality, within-subject dependence and between cluster heterogeneity. It can be easily extended to individual patient data with covariate adjustment, other distributions of outcomes, and more than two time points of visits. When focusing on one likelihood function with covariate adjustment, the model works just like a conventional model; however, with a more generalized and flexible format to accommodate more than one distribution and a more complicated data structure. This model can be implemented in SAS PROC NLMIXED without complex programming requirements. Compared to the naive two-sample proportion z-test, it preserves the type-I error level when no difference exists, and provides less bias estimates when the cluster effect is large. The inflated bias and type I error for the two-sample proportion z-test are related to the magnitude of the true difference and the ignorance of the cluster effect in the data. In contrast to the Cochran-Mantel-Haenszel test, the proposed method provides less bias estimates as the sample size, cluster effect and true difference increases and is much more powerful with a moderate sample size. Also, the performance of the Cochran-Mantel-Haenszel test might be adversely affected by the heterogeneity of differences across the hospitals. Generally, the proposed method performs well when the sample size is large and does require a large sample size to detect a mild to severe difference. Non-convergence of the estimates can be an issue and a model without random effects is preferred when the cluster effect is close to 0, which can be roughly determined by the estimated cluster effect from the data. As to covariate adjustment, it is not clear if the covariates effects should be fixed across all the distributions in the likelihood function in terms of interpretation. Also, further studies could relax the assumptions of independent paths, and include the random effects at the subject level as well as the hospital level. Clinically, the proposed method could provide more precise results and avoid false-positive findings in most scenarios studied here. Moreover, it could be applied in many areas, when one distribution cannot capture the conditional structure in the data.

Acknowledgments

The authors would like to thank Drs. Marnie Bertolet, Chung-Chou H. Chang and reviewers for insightful comments. The project described was supported by the National Institutes of Health through Grant Numbers UL1 RR024153, UL1TR000005 and UL1TR000050.

Appendix A

```
************************** SAMPLE CODE **************************;
* Definition:
*       1. t.visitsWeight: the name of your dataset of group form
*       2. beta0.p1, beta0.pr_0, beta0.pr_1, beta0.p2_0, beta0.p2_1:
*                       random intercepts for probabilities in all paths
*       3. var: variance of random effect
*       4. Weight: number of subjects for each path
*****************************************************************;

*-- calculate log likelihood;
proc nlmixed data=t.VisitsWeight itdetails tech=trureg maxiter=1000 maxfunc=2000;
        parms beta0.p1=0.1 beta0.pr_0=0.1 beta0.pr_1=0.1
        beta0.p2_0=0.1 beta0.p2_1=0.1 var=1;

        p1=min(max(1/(1+exp(-(beta0.p1+bj))),1E-10),1-1E-10);
        l_1=p1**A1 * (1-p1)**(1-A1);

        if A1=1 then do;
                pr_1=min(max(1/(1+exp(-(beta0.pr_1+bj))),1E-10),1-1E-10);
                l_2=pr_1**R * (1-pr_1)**(1-R);

                if R=1 then do;
                        p2_1=min(max(1/(1+exp(-(beta0.p2_1+bj))),1E-10),1-1E-10);
                        l_3=p2_1**A2 * (1-p2_1)**(1-A2); end; end;
        else if A1=0 then do;
                pr_0=min(max(1/(1+exp(-(beta0.pr_0+bj))),1E-10),1-1E-10);
                l_2=pr_0**R * (1-pr_0)**(1-R);

                if R=1 then do;
                        p2_0=min(max(1/(1+exp(-(beta0.p2_0+bj))),1E-10),1-1E-10);
                        l_3=p2_0**A2 * (1-p2_0)**(1-A2); end; end;
        if R=0 then ll=Weight*log(l_1*l_2);
        else if R=1 then ll=Weight*log(l_1*l_2*l_3);

        model A2 ~ general(ll);
        random bj ~ normal(0,log(exp(var))) subject=hospID out=re;
        estimate "p1" 1/(1+exp(-(beta0.p1)));
        estimate "pr_1" 1/(1+exp(-(beta0.pr_1)));
        estimate "p2_1" 1/(1+exp(-(beta0.p2_1)));
        estimate "RR_11" (1/(1+exp(-(beta0.p2_1))))/(1/(1+exp(-(beta0.p1))));
        predict p1 out=p1;
        predict pr_1 out=pr_1;
        predict p2_1 out=p2_1;


        ods output FitStatistics=lld;
        ods output AdditionalEstimates=Est;
run;

proc nlmixed data=t.VisitsWeight itdetails tech=trureg maxiter= 1000 maxfunc=2000;
        parms beta0.p=0.1 beta0.pr_0=0.1 beta0.pr_1=0.1 beta0.p2_0=0.1 var=1;

        p=min(max(1/(1+exp(-(beta0.p+bj))),1E-10),1-1E-10);
        l_1=p**A1 * (1-p)**(1-A1);

        if A1=1 then do;
                pr_1=min(max(1/(1+exp(-(beta0.pr_1+bj))),1E-10),1-1E-10);
                l_2=pr_1**R * (1-pr_1)**(1-R);

                if R=1 then do;
                        p=min(max(1/(1+exp(-(beta0.p+bj))),1E-10),1-1E-10);
                        l_3=p**A2 * (1-p)**(1-A2); end; end;
        else if A1=0 then do;
                pr_0=min(max(1/(1+exp(-(beta0.pr_0+bj))),1E-10),1-1E-10);
                l_2=pr_0**R * (1-pr_0)**(1-R);

                if R=1 then do;
                        p2_0=min(max(1/(1+exp(-(beta0.p2_0+bj))),1E-10),1-1E-10);
                        l_3=p2_0**A2 * (1-p2_0)**(1-A2); end; end;
        if R=0 then ll=Weight*log(l_1*l_2);
        else if R=1 then ll=Weight*log(l_1*l_2*l_3);

        model A2 ~ general(ll);
        random bj ~ normal(0,log(exp(var))) subject=hospID;
        ods output FitStatistics=lln;
run;

*-- likelihood ratio test;
data LRT;
        retain descr lln lld lrt pvalue;
        merge lln (rename=(value=lln)) lld(rename=(value=lld)); by descr;
        lrt=lln-lld;
        pvalue=1 - probchi(lrt, 1);
        if descr='-2 Log Likelihood' then delete;
run;

*-- Output;
proc print data=Est; proc print data=LRT; run;
```

References

1. Clarke SC, Wilson BF. The relative stability of remarriages: A cohort approach using vital statistics. Family Relations. 1994; 43(3):305–310.
2. Chadwick, BA.; Heaton, TB. Statistical handbook on the American family. Oryx Press; Phoenix, AZ: 1999.
3. Zhou H, Chen J, Rissanen TH, Korrick SA, Hu H, Salonen JT, Longnecker MP. An efficient sampling and inference procedure for studies with a continuous outcome. Epidemiology (Cambridge, Mass). 2007; 18(4):461.
4. Prentice RL. A case-cohort design for epidemiologic cohort studies and disease prevention trials. Biometrika. 1986; 73(1):1–11.
5. Alessandrini EA, Lavelle JM, Grenfell SM, Jacobstein CR, Shaw KN. Return visits to a pediatric emergency department. Pediatric emergency care. 2004; 20(3):166–171. [PubMed: 15094574]
6. Bain, LJ.; Engelhardt, M. Introduction to probability and mathematical statistics. Vol. 4. Duxbury Press; Belmont, CA: 1992.
7. Mantel N, Haenszel W. Statical aspects of the analysis of data from retrospective studies of disease. Journal of the National cancer Institute. 1959; 22(4):v.
8. Agresti, A. Categorical data analysis. John Wiley & Sons; Hoboken, New Jersey: 2002.
9. Sistrom CL, Garvan CW. Proportions, odds, and risk. Radiology-Oak Brook Il. 2004; 230(1):12–19.

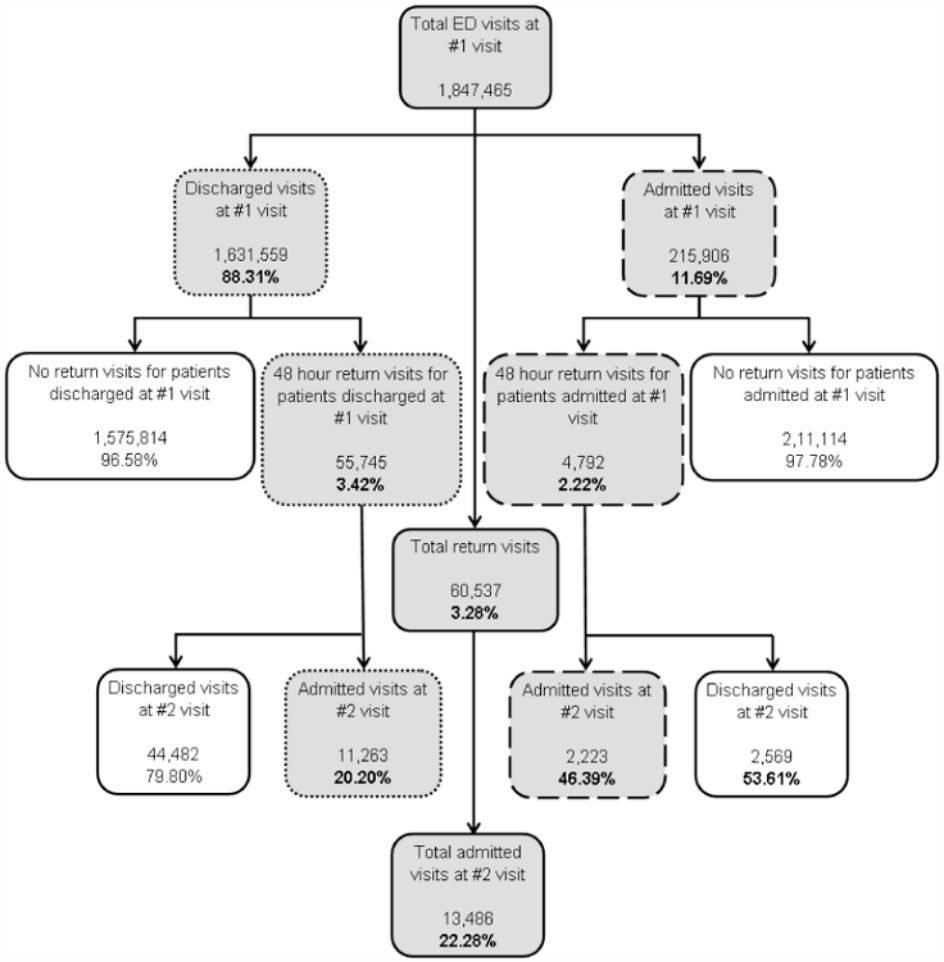

Figure 1. Schematic diagram of the ED data

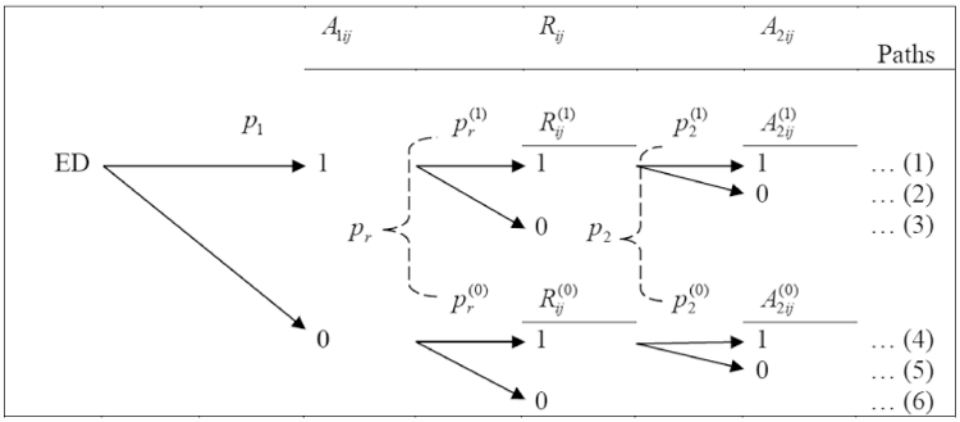

Figure 2. Sample paths associated with the ED Data

Table 1

Simulation setting

Items	Values	
Hospital size: M	30	
Sample size: N	45,000	500
True difference: (p_1, $p_2^{(1)}$)	(0.46, 0.46) (0.30, 0.46) (0.12, 0.46)	(0.46, 0.46) (0.38, 0.46) (0.30, 0.46)
Cluster effect: σ^2	0/0.001/0.01/0.1/0.5/1	
($p_r^{(1)}$, $p_r^{(0)}$, $p_2^{(0)}$)	(0.02, 0.03, 0.20)	(0.50, 0.55, 0.20)

Table 2

Simulation results comparing the proposed method to the ordinary tests for proportions under different scenarios with N=45,000.[ab]

	Proposed method														Two-sample proportion z-test					Cochran-Mantel-Haenszel test				
	p_1				$p_r^{(1)}$			$p_2^{(1)}$			$RR^{(1)}$				$RR^{(1)}$					$RR^{(1)}$				
Difference (p_1, $p_2^{(1)}$)	**Runs Est/Power[c]**	**Bias**	**SD**	**SE**	**Bias**	**SD**	**SE**	**Bias**	**SD**	**SE**	**Bias**	**SD**	**SE**	**Power[d]**	**Runs**	**Bias**	**Est SD[e]**	**Est SE[f]**	**Power[d]**	**Runs**	**Bias**	**Est SD[e]**	**Est SE[f]**	**Power[d]**
None: (0.46, 0.46)																								
Cluster effect:																								
0	**981**	**-0.01**	**0.23**	**0.23**	**<0.01**	**0.10**	**0.10**	**0.02**	**2.44**	**2.45**	**0.06**	**5.32**	**5.35**	**0.044**	1000	0.09	5.37	5.36	0.044	1000	0.02	5.37	5.36	0.041
0.001	**978**	**-0.01**	**0.27**	**0.23**	**<0.01**	**0.10**	**0.10**	**0.12**	**2.47**	**2.45**	**0.27**	**5.38**	**5.35**	**0.047**	1000	0.22	5.41	5.35	0.044	1000	0.06	5.41	5.35	0.046
0.01	1000	-0.01	0.51	0.50	<0.01	0.11	0.10	-0.01	2.51	2.48	-0.01	5.40	5.33	0.053	1000	0.77	5.38	5.30	0.055	1000	-0.09	5.37	5.29	0.051
0.1	1000	-0.01	1.44	1.41	<0.01	0.15	0.14	<0.01	2.71	2.74	0.03	5.05	5.17	0.040	1000	7.53	5.05	4.78	0.330	1000	-0.06	4.64	4.73	0.039
0.5	959	-0.10	3.17	3.10	<0.01	0.27	0.26	-0.09	3.72	3.76	0.03	4.58	4.71	0.043	1000	30.71	6.76	3.30	0.995	1000	-0.02	3.14	3.19	0.039
1	889	-0.08	4.39	4.37	0.02	0.36	0.36	-0.02	4.74	4.79	0.14	4.36	4.39	0.045	1000	49.75	8.47	2.34	1.000	1000	0.06	2.27	2.24	0.037
Mild: (0.30, 0.46)																								
Cluster effect:																								
0	**999**	**<0.01**	**0.22**	**0.22**	**<0.01**	**0.13**	**0.12**	**0.01**	**3.01**	**3.03**	**0.04**	**10.11**	**10.17**	**1.000**	1000	0.02	6.63	6.66	1.000	1000	-0.20	6.64	6.67	1.000
0.001	**999**	**<0.01**	**0.25**	**0.22**	**<0.01**	**0.12**	**0.12**	**0.05**	**3.03**	**3.03**	**0.14**	**10.16**	**10.16**	**0.999**	1000	0.13	6.65	6.65	0.999	1000	-0.27	6.65	6.65	1.000
0.01	1000	<0.01	0.44	0.43	<0.01	0.13	0.12	-0.02	3.05	3.05	-0.06	10.20	10.12	1.000	1000	1.08	6.62	6.56	1.000	1000	-0.86	6.62	6.56	1.000
0.1	1000	<0.01	1.22	1.20	<0.01	0.16	0.16	-0.03	3.16	3.20	-0.04	9.67	9.80	1.000	1000	10.80	6.12	5.82	1.000	1000	-5.54	5.84	5.75	1.000
0.5	931/930	-0.01	2.72	2.62	<0.01	0.27	0.26	-0.08	3.96	3.99	0.11	8.98	9.13	1.000	1000	41.86	7.59	3.80	1.000	1000	-19.54	5.18	3.64	1.000
1	882	-0.11	3.72	3.71	0.01	0.36	0.36	-0.17	4.88	4.94	0.57	9.34	9.04	1.000	1000	63.61	9.48	2.60	1.000	1000	-28.73	5.20	2.44	1.000
Severe: (0.12, 0.46)																								
Cluster effect:																								
0	**982**	**<0.01**	**0.16**	**0.15**	**<0.01**	**0.19**	**0.19**	**-0.19**	**4.85**	**4.79**	**-1.53**	**40.84**	**40.23**	**1.000**	1000	-1.17	10.84	10.64	1.000	1000	-3.00	10.86	10.66	1.000
0.001	**977**	**0.01**	**0.17**	**0.15**	**<0.01**	**0.19**	**0.19**	**-0.14**	**4.83**	**4.79**	**-1.24**	**40.57**	**40.19**	**1.000**	1000	-0.76	10.69	10.61	1.000	1000	-3.11	10.67	10.63	1.000
0.01	998	0.01	0.25	0.24	<0.01	0.19	0.19	-0.09	4.69	4.78	-0.89	39.26	40.00	1.000	1000	1.50	10.31	10.44	1.000	1000	-6.16	10.27	10.44	1.000
0.1	1000	0.02	0.63	0.61	<0.01	0.21	0.21	-0.03	4.68	4.69	-0.56	38.23	38.49	1.000	1000	22.52	9.44	9.04	1.000	1000	-34.56	9.47	8.91	1.000

	Proposed method																						
	p_1				$p_r^{(1)}$			$p_2^{(1)}$			$RR^{(1)}$				Two-sample proportion z-test					Cochran-Mantel-Haenszel test				
															$RR^{(1)}$					$RR^{(1)}$				
Difference (p_1, $p_2^{(1)}$)	**Runs Est/Power[c]**	**Bias**	**SD**	**SE**	**Bias**	**SD**	**SE**	**Bias**	**SD**	**SE**	**Bias**	**SD**	**SE**	**Power[d]**	**Runs**	**Bias**	**Est SD[e]**	**Est SE[f]**	**Power[d]**	**Runs**	**Bias**	**Est SD[e]**	**Est SE[f]**	**Power[d]**
0.5	979	0.05	1.37	1.34	0.01	0.29	0.28	0.01	4.86	4.83	0.20	34.38	35.78	1.000	1000	76.56	9.40	5.39	1.000	1000	-118.81	12.82	5.13	1.000
1	913	0.08	1.92	1.89	0.02	0.37	0.37	0.06	5.44	5.42	1.93	37.66	36.59	1.000	1000	95.45	11.72	3.44	1.000	1000	-171.12	16.01	3.22	1.000

[a] Bias, standard deviation (SD) and standard error (SE) are in the scale of 100 times.

[b] Cells in bold are from likelihood functions without random effects.

[c] Different numbers of available simulated datasets for estimation and power because of negative likelihood ratio statistics.

[d] This is the type I error under the scenario of no difference.

[e] Estimated standard deviation of $ln(RR)$

[f] Estimated standard error of $ln(RR)$

Est, estimate

Table 3

Simulation results comparing the proposed method to the ordinary tests for proportions under different scenarios with N=500.[ab]

	Proposed method														Two-sample proportion z-test					Cochran-Mantel-Haenszel test				
	p_1				$p_r^{(1)}$			$p_2^{(1)}$			$RR^{(1)}$				$RR^{(1)}$					$RR^{(1)}$				
Difference (p_1, $p_2^{(1)}$)	**Runs Est/Power[c]**	**Bias**	**SD**	**SE**	**Bias**	**SD**	**SE**	**Bias**	**SD**	**SE**	**Bias**	**SD**	**SE**	**Power[d]**	**Runs**	**Bias**	**Est SD[e]**	**Est SE[f]**	**Power[d]**	**Runs**	**Bias**	**Est SD[e]**	**Est SE[f]**	**Power[d]**
None: (0.46, 0.46)																								
Cluster effect:																								
0	**1000**	**0.12**	**2.24**	**2.23**	**0.01**	**3.19**	**3.29**	**0.16**	**4.53**	**4.63**	**0.33**	**11.10**	**11.21**	**0.056**	1000	0.33	11.07	11.23	0.055	1000	-5.07	11.01	11.20	0.063
0.001	**1000**	**0.13**	**2.26**	**2.23**	**0.01**	**3.17**	**3.29**	**0.13**	**4.47**	**4.63**	**0.25**	**11.01**	**11.20**	**0.050**	1000	0.25	10.99	11.24	0.050	1000	-5.32	10.96	11.19	0.077
0.01	**999**	**0.14**	**2.25**	**2.23**	**0.19**	**3.20**	**3.29**	**0.35**	**4.49**	**4.63**	**0.70**	**10.92**	**11.19**	**0.042**	1000	0.71	10.86	11.18	0.042	1000	-5.26	10.85	11.13	0.075
0.1	950	0.11	2.67	2.66	0.04	3.61	3.65	0.16	4.70	4.90	0.37	11.13	11.35	0.042	1000	5.25	10.40	10.67	0.064	1000	-4.77	10.34	10.59	0.076
0.5	961	-0.08	3.83	3.88	-0.06	4.38	4.71	-0.06	5.50	5.66	0.31	11.20	11.60	0.043	1000	19.84	9.90	9.32	0.440	1000	-3.89	8.86	9.06	0.060
1	913	-0.19	4.88	4.98	-0.06	5.52	5.72	0.06	6.30	6.49	0.83	11.74	11.93	0.048	1000	32.06	9.97	8.35	0.836	1000	-3.10	7.83	7.89	0.064
Mild: (0.38, 0.46)																								
Cluster effect:																								
0	**999**	**0.10**	**2.15**	**2.17**	**0.08**	**3.50**	**3.62**	**0.27**	**5.00**	**5.09**	**0.80**	**15.08**	**15.15**	**0.286**	1000	0.80	12.34	12.52	0.292	1000	-8.46	12.36	12.49	0.129
0.001	**1000/999**	**0.10**	**2.15**	**2.17**	**0.08**	**3.50**	**3.62**	**0.25**	**4.91**	**5.09**	**0.74**	**14.89**	**15.15**	**0.303**	1000	0.74	12.20	12.52	0.311	1000	-8.60	12.17	12.48	0.129
0.01	**1000**	**0.17**	**2.20**	**2.17**	**0.25**	**3.56**	**3.61**	**0.47**	**4.95**	**5.08**	**1.09**	**14.96**	**15.11**	**0.318**	1000	1.09	12.20	12.45	0.323	1000	-8.78	12.21	12.40	0.133
0.1	949/948	0.13	2.54	2.57	0.06	3.94	3.95	0.18	5.09	5.33	0.50	14.99	15.26	0.296	1000	6.43	11.73	11.84	0.460	1000	-9.17	11.60	11.75	0.131
0.5	936/934	0.06	3.71	3.72	0.11	4.68	4.96	0.04	5.78	5.99	0.41	14.91	15.41	0.296	1000	22.91	10.87	10.19	0.886	1000	-10.98	9.76	9.86	0.141
1	833/831	0.23	4.63	4.72	0.31	5.68	5.88	0.35	6.43	6.72	0.81	15.13	15.66	0.303	1000	35.91	10.87	9.05	0.986	1000	-12.14	8.46	8.47	0.152
Severe: (0.30, 0.46)																								
Cluster effect:																								
0	**1000**	**0.09**	**2.07**	**2.05**	**0.07**	**3.98**	**4.07**	**0.24**	**5.66**	**5.73**	**1.04**	**21.71**	**21.93**	**0.780**	1000	1.04	14.06	14.32	0.788	1000	-15.68	14.01	14.32	0.531
0.001	**1000**	**0.09**	**2.05**	**2.05**	**0.07**	**3.99**	**4.07**	**0.27**	**5.55**	**5.73**	**1.14**	**21.45**	**21.94**	**0.777**	1000	1.14	13.85	14.31	0.790	1000	-15.51	13.73	14.31	0.525
0.01	**1000**	**0.14**	**2.09**	**2.05**	**0.27**	**4.07**	**4.07**	**0.51**	**5.65**	**5.72**	**1.72**	**21.88**	**21.89**	**0.789**	1000	1.72	14.07	14.23	0.798	1000	-15.65	14.01	14.22	0.529
0.1	948/947	0.09	2.45	2.39	0.01	4.39	4.38	0.18	5.84	5.93	1.00	22.35	22.03	0.793	1000	8.14	13.55	13.47	0.895	1000	-17.45	13.40	13.38	0.565

	Proposed method														Two-sample proportion z-test					Cochran-Mantel-Haenszel test				
	p_1				$p_r^{(1)}$			$p_2^{(1)}$			$RR^{(1)}$				$RR^{(1)}$					$RR^{(1)}$				
Difference (p_1, $p_2^{(1)}$)	**Runs Est/Power[c]**	**Bias**	**SD**	**SE**	**Bias**	**SD**	**SE**	**Bias**	**SD**	**SE**	**Bias**	**SD**	**SE**	**Power[d]**	**Runs**	**Bias**	**Est SD[e]**	**Est SE[f]**	**Power[d]**	**Runs**	**Bias**	**Est SD[e]**	**Est SE[f]**	**Power[d]**
0.5	885	0.17	3.39	3.38	0.17	5.14	5.29	0.42	6.17	6.46	1.65	21.90	22.07	0.807	1000	26.99	12.19	11.38	0.996	1000	-23.26	11.08	10.99	0.591
1	711	0.53	4.44	4.25	0.51	6.27	6.12	0.55	6.87	7.06	0.66	22.47	22.14	0.792	1000	40.04	12.31	9.99	1.000	1000	-27.96	9.74	9.31	0.622

[a] Bias, standard deviation (SD) and standard error (SE) are in the scale of 100 times.

[b] Cells in boldface are from likelihood functions without random effects.

[c] Different numbers of available simulated datasets for estimation and power because of negative likelihood ratio statistics.

[d] This is the type I error under the scenario of no difference.

[e] Estimated standard deviation of $ln(RR)$

[f] Estimated standard error of $ln(RR)$

Est, estimate

Table 4

Relative risk for admission at a 48-hr return visit compared to an index visit

		Two proportion z-test	Cochran-Mantel-Haenszel test	Proposed method[a]		
	Crude rate % (SE)	Relative risk (95% CI)	Relative risk (95% CI)	Estimated rate % (SE)	Relative risk (95% CI)	*p*-value
Admission rate at the index visit	11.687 (0.024)	Reference	Reference	11.690 (0.024)	Reference	–
Overall admission rate at the return visit	22.277 (0.169)	1.906 (1.877, 1.936)	1.923 (1.894, 1.953)	22.280 (0.169)	1.906 (1.877, 1.936)	<0.0001
Admission rate at the return visit for patients discharged at the index visit	20.205 (0.170)	1.729 (1.700, 1.758)	1.760 (1.730, 1.789)	20.200 (0.170)	1.729 (1.699, 1.758)	<0.0001
Admission rate at the return visit for patients admitted at the index visit	46.390 (0.720)	3.969 (3.850, 4.093)	3.627 (3.518, 3.739)	46.390 (0.720)	3.970 (3.847, 4.092)	<0.0001

[a] Approach without random effects was used because of the extremely small cluster effect in the data.