Original Article
http://mjiri.iums.ac.ir
Medical Journal of the Islamic Republic of Iran (MJIRI)

Med J Islam Repub Iran. 2026 (30 Sep);40.106. https://doi.org/10.47176/mjiri.40.106

Hierarchical Bootstrapping for Quadratic Exponential Models in Genetic Epidemiology: Familial Aggregation Analysis with Small Samples and Unequal Family Sizes

Mohammad Hassan Rakhshani[1], Hojjat Zeraati[2], Hassan Hashemi[3], Akbar Fotouhi[2]*

Received: 5 May 2026 Revised: 21 Sep 2026 Accepted: 26 Sep 2026 Published: 30 Sep 2026

Abstract

Background: Familial aggregation analysis evaluates whether disease clustering within families exceeds chance expectations. A key challenge is irreproducibility, which arises when a model cannot be consistently applied to families with different family sizes. Although the quadratic exponential model is widely used for correlated binary outcomes, its performance is affected by irreproducibility and small-sample bias.

Methods: This study compares the marginal approach, designed to address irreproducibility, with a hierarchical bootstrap procedure. Simulations were conducted under exchangeable and non exchangeable dependence structures using the quadratic exponential framework, focusing on samples with unequal family sizes. Performance was evaluated using RMSE and empirical coverage. The two methods were also applied to data from the Tehran Eye Study.

Results: In exchangeable settings, both methods produced reasonable estimates for baseline disease odds; however, the hierarchical bootstrap provided more stable inference. It also performed better in estimating the incremental effect associated with an additional affected relative (γ) in samples of 200 families (coverage: 94.8% vs 92.5%). In non-exchangeable scenarios, the bootstrap approach showed better performance for key parameters, including conditional child–parent odds (95.1% vs 93.5%) and γ (93.9% vs 93%). Results from the Tehran Eye Study were consistent with these findings (93.9% vs 93.6% for γ and 94.7% vs 92.6% for γ_M).

Conclusion: These findings suggest that hierarchical bootstrapping may provide a useful alternative for inference in familial aggregation studies with heterogeneous family sizes.

Keywords: Familial aggregation, Hierarchical bootstrap, Quadratic exponential model, Irreproducibility, Tehran Eye Study

This work has been published under CC BY-NC-SA 4.0 license.

Copyright© Iran University of Medical Sciences

Cite this article as: Rakhshani MH, Zeraati H, Hashemi H, Fotouhi A. Hierarchical Bootstrapping for Quadratic Exponential Models in Genetic Epidemiology: Familial Aggregation Analysis with Small Samples and Unequal Family Sizes. *Med J Islam Repub Iran.* 2026 (30 Sep);40:106. https://doi.org/10.47176/mjiri.40.106

Introduction

Familial aggregation, defined as the increased prevalence of a disease among relatives of an affected individual compared with the general population, is often considered an initial step in identifying potential genetic determinants of disease (1). In recent years, familial aggregation has been investigated for a wide range of diseases (2–6), reflecting its importance in genetic epidemiology. Despite this growing interest, the statistical methodologies used in familial aggregation studies have advanced more slowly than those in other areas of genetic epidemiology (7, 8). A major methodological challenge arises when modeling disease odds conditional on the disease status of other family members. Because family sizes vary, a single statistical model cannot fully represent the dependence

Corresponding author: Dr Akbar Fotouhi, afotouhi@tums.ac.ir

1. Department of Biostatistics and Epidemiology, School of Health, Sabzevar University of Medical Sciences, Sabzevar, Iran
2. Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran
3. Noor Research Center for Ophthalmic Epidemiology, Noor Eye Hospital, Tehran, Iran

Key Messages:

↑What is "already known" in this topic:

Familial aggregation—defined as the increased prevalence of a disease among relatives of an affected individual compared with the general population. However, variation in family size complicates statistical inference, because a single statistical model may not fully capture the dependence structure across all families, a phenomenon sometimes referred to as "irreproducibility".

→What this article adds:

This study introduces a hierarchical bootstrap approach for analyzing familial aggregation. The method provides more reliable empirical estimates of standard errors and confidence intervals while remaining easy to implement using standard statistical software, offering a practical tool for researchers working with family based observational data.

[Downloaded from mjiri.iums.ac.ir on 2026-11-15]

[DOI: 10.47176/mjiri.40.106]

structure across all families, a phenomenon referred to as "irreproducibility" (9). For example, in a three-person family, a child's odds of disease depend on parental status, whereas in a four-person family they may also depend on a sibling's status. Moreover, family history is inherently related to family size, since larger families have a higher probability of containing at least one additional affected relative (10). Ignoring this informative characteristic may lead to misleading conclusions regarding familial aggregation (11). To address these issues, Matthews et al. introduced the Quadratic Exponential Model (QEM) (9), which allows familial aggregation to be quantified through odds ratios and can be implemented using standard statistical software (12, 13). However, QEM is also affected by the irreproducibility problem. Several strategies have been proposed to mitigate this limitation, including the naive approach, the restricted approach, the missing-data approach, the marginal approach, and hybrid methods combining the latter two strategies. Among these approaches, the missing-data and hybrid methods appear conceptually problematic because they treat smaller families as if they were larger families with missing observations. In contrast, the marginal approach has generally been regarded as the most logically coherent solution. Under this strategy, the analysis is based on the smallest family size and all corresponding subsets of larger families (e.g., all two-person combinations when the minimum family size is two). The marginal approach typically relies on generalized estimating equations (GEE) to account for within-family correlation across these subsets (14, 15). Although GEE models are appropriate for correlated outcomes, their performance may deteriorate when cluster sizes are small or when intra-cluster correlation is high (16, 17). In such situations, misspecification of the working correlation structure may lead to biased estimates and inadequate coverage probabilities, particularly when the number of clusters is limited (18). Because comprehensive sampling of all family members is often logistically and financially challenging, studies with relatively small sample sizes are common in practice. Given these challenges, the present study pursued two main objectives. First, we sought to propose an analytical approach that is simpler than the marginal method and therefore more accessible to researchers with limited statistical programming expertise. Second, we examined whether familial aggregation can be assessed reliably using fewer than the 500 simulated families considered by Matthews et al. (9), while accounting for small family sizes (2–6 persons) and a limited number of clusters. To address these objectives, we investigated the hierarchical bootstrap method as a potential alternative to generalized estimating equations. In this framework, families are treated as the bootstrap sampling units, allowing estimation of regression parameters without requiring explicit assumptions about the within-family correlation structure. The performance of the hierarchical bootstrap method was evaluated in comparison with the marginal approach using both simulation studies and real-world data from the Tehran Eye Study (19), which was designed to investigate the prevalence and causes of visual impairment in the population of Tehran.

Methods

Study design

Quadratic exponential model (QEM) is one of the models used in familial aggregation analysis, where all three-way associations and beyond are considered as zero (20-22). Let y_i represent the binary disease status of the ith family member and $(y_1, \dots, y_n)$ are the disease status for all family members. In its simplest form, the QEM for a family with n binary infection statuses is:

$$P(y_1, \dots, y_n) = \Delta_n^{-1} exp\left(\delta \sum_{i=1}^{n} y_i + \gamma \sum_{i<j} y_i y_j\right) \qquad (1)$$

Where Δ_n is the normalization constant. The parameters of this model are interpreted as conditional measures of response and association. In particular

$$exp(\delta) = \frac{P(y_i = 1|\boldsymbol{y_{-i}} = \mathbf{0})}{P(y_i = 0|\boldsymbol{y_{-i}} = \mathbf{0})} \qquad (2)$$

$$exp(\gamma) = \frac{P(y_i = 1, y_j = 1|\boldsymbol{y_{-ij}})P(y_i = 0, y_j = 0|\boldsymbol{y_{-ij}})}{P(y_i = 1, y_j = 0|\boldsymbol{y_{-ij}})P(y_i = 0, y_j = 1|\boldsymbol{y_{-ij}})} \qquad (3)$$

where $\boldsymbol{y_{-i}}$ in equation (2) and $\boldsymbol{y_{-ij}}$ in equation (3) are the vectors of all outcomes within a family except for the i-th individual and both the i-th and j-th individuals.

A simple and efficient way to fit the QEM using common statistical software is to transform it into a logistic regression model. With a simple transformation, the equation becomes:

$$logit\{P(y_i = 1|\boldsymbol{y_{-i}})\} = \delta + \gamma \sum_{j \neq i} y_j \qquad (4)$$

In this model, δ denotes the log-odds of disease for an individual when no other family member is affected, and γ represents the conditional odds ratio for disease associated with having affected relatives. To fit the model, we included a family identifier, coded each member's disease status as affected (1) or unaffected (0), and defined, as a third variable, the number of affected family members excluding the individual. Equation (4) assumes exchangeability; that is, the odds of disease depend only on relatives' disease status and not on specific pedigree relationships. Under non-exchangeable settings, the model was extended by adding relationship-specific covariates (e.g., γ_M) to distinguish different familial ties. In both settings, generalized estimating equations (GEE) were used to model the association between covariates and the response. Instead of GEE, we employed a nonparametric bootstrap procedure following Davison and Hinkley (23). Specifically, we adopted their first strategy: in Stage 1, families were sampled with replacement to retain parents; in Stage 2, for the families selected in Stage 1, offspring were randomly sampled within each family without replacement. Under this design, the family served as the bootstrap resampling unit. We applied this strategy to the

[Downloaded from mjiri.iums.ac.ir on 2026-11-15]

[DOI: 10.47176/mjiri.40.106]

naive approach. The complex within-family correlation structure associated with this strategy and the second strategy is presented in the Appendix.

In order to facilitate comparison with the study by Matthews et al., we restricted family size to 2–6 members in both the simulations and the real-data analysis. We restricted the family size to 2–6 members to ensure comparability with previous studies and consistency with local demographic data. Matthews et al. conducted their simulation study with 500 families (9). In contrast, we performed simulations using sample sizes of 25, 50, 100, and 200 families.

The Simulation Study

We simulated the data according to the first scenario described by Matthews et al. (9). The parameter values were set to$(\delta, \gamma, \gamma_M) = (-2\,, 0.5, 0.59)$, representing a fairly common disease with moderate familial aggregation (9). This scenario is of particular interest because the different approximations proposed by Matthews et al. (9) yield discrepant results. To generate the simulated datasets, we examined the structure of the Tehran Eye Study data, a population-based cross-sectional study (19), and the simulated samples followed the same household-size distribution as the study population (25%, 28%, 32%, 12%, and 3% for sizes 2–6, respectively).

Two datasets were generated: one assuming an exchangeable correlation structure and the other assuming a non-exchangeable correlation structure. Each dataset comprised 100,000 families with 2–6 members. In the exchangeable dataset, the conditional odds ratio for disease given at least one other affected family member was set to 1.65 (i.e., exp(0.5) = 1.65), consistent with settings such as twin or multiple-birth studies. In the non-exchangeable setting, we added an additional parameter to capture parent-to-child transmission; this parent–child conditional odds ratio was set to 1.80 (i.e., exp(0.59)). Specifically, we generated 1,000 replicates with sample sizes of 25, 50, 100, and 200 families, sampling family sizes according to the observed distribution of 2–6-member families in the parents file.

To simulate the data, the free software R can be used (24). For simulation of correlated binary data, the cquad package can also be used (25). A custom R script was developed to simulate Equation (4) and to randomly select families according to the specified structure. The code is available upon request. Also, for the marginal approach analysis a program in R software was used, which can be obtained via email from the third author of the Matthews' article (9). For the hierarchical bootstrap analysis, the free ClusterBootstrap package in R software was used (26).

To compare the two methods, we evaluated several performance measures across simulation replicates. In particular, the ("true") variance of $\hat{\beta}$ was estimated as the empirical variance of the estimator across simulation runs:

$$var(\hat{\beta}) = \sum_{i=1}^{m} \frac{\left(\hat{\beta}_i - \bar{\hat{\beta}}\right)^2}{m-1} \qquad (5)$$

Here, $\hat{\beta}_i$ denotes the estimate obtained from the i-th successful replicate (i.e., a replicate for which the model-fitting procedure converged), $\bar{\hat{\beta}} = \sum_{i=1}^{m} \frac{\hat{\beta}_i}{m}$ is the average estimate across replicates, and m is the number of successful replicates. Replicates for which the model-fitting procedure failed to converge were excluded from the analysis. All quantities were computed separately for the bootstrap and marginalized approaches.

Performance was assessed using two measures: the root mean square error (RMSE) and the coverage probability. Bias was calculated as the difference between the true parameter value and the mean of the estimates obtained under each method (the hierarchical bootstrap and the marginal approach). The RMSE was defined as the square root of the mean squared error (MSE). For an estimator $\hat{\beta}$, the MSE is defined as follows:

$$\text{M.S.E}(\hat{\beta}) = var(\hat{\beta}) + Bias^2(\hat{\beta}) \qquad (6)$$

Smaller values of RMSE indicate smaller variance of the estimator, smaller bias, or both. Coverage probability was defined as the proportion of confidence intervals constructed using the bootstrap and marginal approaches that contained the true parameter value. For the bootstrap method, we used bias-corrected and accelerated (BCa) confidence intervals with B=2000 resamples. For the marginal approach, Wald confidence intervals were used. Finally, departures of the empirical coverage from the nominal 95% level were evaluated using an exact binomial test for a single proportion, testing H_0:p =0.95 versus H_1:p≠0.95.

Results

Table 1 presents the results of the comparison between the two methods under the exchangeable correlation structure. For the estimation of δ, at sample sizes of 25 and 50 families, the hierarchical bootstrap demonstrated relatively better performance than the marginal approach; however, neither method achieved the nominal 95% coverage level. When the sample size increased to 100 and 200 families, both methods attained 95% coverage, and the difference in coverage between the two approaches was negligible. Nevertheless, the hierarchical bootstrap produced mean estimates closer to the true parameter value and smaller RMSE values.

For the estimation of γ, the marginal approach failed to achieve the nominal 95% coverage level, whereas the hierarchical bootstrap performed better and achieved this level of coverage.

For the non-exchangeable case, Table 2 shows that the mean estimates obtained from the bootstrap method are closer to the true value than those from the marginal approach across all scenarios. Comparisons of RMSE did not uniformly favor the bootstrap method, although it yielded lower RMSE values in most cases. The bootstrap method achieved 95% coverage only with 200 families, whereas the marginal approach reached the nominal 95%

[Downloaded from mjiri.iums.ac.ir on 2026-11-15]

[DOI: 10.47176/mjiri.40.106]

Table 1. Estimation of coefficients, root mean square error, coverage level (%) and *P*-value in hierarchical bootstrap model and marginal approach (exchangeable correlation structure).

Number of families	Coefficient estimation (H.B/M.A)	R.M.S.E (H.B/M.A)	Coverage% (H.B/M.A)	*P*-Value (H.B/M.A)
		$\delta = -2$		
25	-1.84/-1.96	0.42/0.54	90.6/87.1	<0.001/<0.001
50	-1.88/-2.13	0.39/0.54	93.4/87.1	0.037/<0.001
100	-1.89/-2.14	0.41/0.43	95.3/95	0.717/0.988
200	-1.90/-2.49	0.24/0.26	95.4/95	0.596/1
		$\gamma = 0.5$		
25	0.3/0.35	0.30/0.24	91/90	<0.001/<0.001
50	0.37/0.69	0.24/0.30	92.5/91.1	<0.001/<0.001
100	0.43/0.75	0.24/0.24	93.5/90.5	0.051/<0.001
200	0.45/0.76	0.26/0.20	94.8/92.5	0.823/0.002

H.B: Hierarchical Bootstrap, M.A: Marginal Approach, P-value based on two-sided binomial exact test

Table 2. Estimation of coefficients, Root mean square error, Coverage level (%) and *P*-value in hierarchical bootstrap model and marginal approach (non-exchangeable correlation structure).

Number of families	Coefficient estimation (H.B/M.A)	R.M.S.E (H.B/M.A)	Coverage% (H.B/M.A)	*P*-Value (H.B/M.A)
		$\delta = -2$		
25	-1.85/-1.79	0.50/0.54	88.6/87.1	<0.001/<0.001
50	-1.85/-1.79	0.28/0.51	89/88.5	<0.001/<0.001
100	-1.89/-1.86	0.26/0.24	93/90	<0.001/<0.001
200	-1.9/-1.86	0.26/0.36	94/91.7	0.166/<0.001
		$\gamma = 0.5$		
25	0.29/0.25	0.52/0.52	88/86.5	<0.001/<0.001
50	0.35/0.28	0.42/0.50	90.6/89	<0.001/<0.001
100	0.40/0.34	0.24/0.26	93/92.2	0.008/<0.001
200	0.43/0.4	0.22/0.26	93.9/93	0.166/0.01
		$\gamma_M = 0.59$		
25	0.39/0.30	0.37/0.44	89.7/73.4	<0.001/<0.001
50	0.41/0.35	0.28/0.55	93.4/87	0.036/<0.001
100	0.41/0.38	0.26/0.20	93.5/92	0.043/<0.001
200	0.45/0.40	0.26/0.28	95.1/93.5	0.946/0.058

H.B: Hierarchical Bootstrap, M.A: Marginal Approach, P-value based on two-sided binomial exact test

coverage level only at a sample size of 200 families and only for the estimation of γ_M, with a weak *P*-value.

Overall, across the two scenarios examined, the hierarchical bootstrap showed some advantage over the marginal approximation, particularly at the larger sample size of 200 families.

Following the determination of a 200-household minimum sample size, we compared our methods against a naïve approach. In the exchangeable scenario, results (estimate, RMSE, coverage, *P*-value) for δ=−2 and γ=0.5 were (−1.98, 0.3, 92.9, <0.001) and (0.27, 1.34, 88.4, <0.001), respectively. Corresponding values for the non-exchangeable case were (−1.99, 0.37, 91.5,<0.001) for δ=−2, (0.42, 0.38, 89.6,<0.001) for γ=0.5, (0.58, 0.49, 93.4, 0.02) and (0.58, 0.49, 93.4, 0.02) for $\gamma_M = 0.59$. Overall, the naïve model consistently underperformed relative to our methods, except in point estimation.

Application to the Tehran Eye Study

Refractive errors are common visual disorders, with an overall prevalence of about 43%, and astigmatism is the most frequent type, reported in 8% to 62% of individuals (27). Astigmatism can reduce visual quality and is associated with difficulties in daily activities, decreased productivity, and increased healthcare costs (27–29).

Among the three types of astigmatism— with-the-rule (WTR), against-the-rule, and oblique— WTR astigmatism was selected for comparison of the two analytical methods because it is more common in younger individuals (28). Astigmatism was defined as cylinder power ≤ −0.5 D, and WTR astigmatism as an axis of 0 ± 19^0. In this study, 33.6% of participants had WTR astigmatism.

Data were obtained from the Tehran Eye Study, which included 160 clusters randomly selected from the 22 urban districts of Tehran. From the full dataset, 1,261 families with 2–6 members were selected. To compare the two analytical methods in this real-data setting, parameter values were first approximated using the full dataset. Because the true parameter values are unknown in observational data, a logistic regression model was fitted to the full dataset under the assumption of independence, that is, without accounting for within-family correlation—to obtain point estimates of the model parameters. As noted by Agresti, "Under the independent structure, the estimates of parameters are identical to the maximum likelihood estimates obtained by treating all observations within and between clusters as independent" (30). These estimates were therefore used as reference values for evaluating the performance of the two methods. Using sampling with replacement, 1,000 datasets were generated, each containing 200 families. The model was fitted to each dataset using both the bootstrap and marginal approaches. Method performance was evaluated using the root mean squared error (RMSE) of parameter estimates relative to the reference values and the coverage probability of the corresponding 95% confidence intervals.

[Downloaded from mjiri.iums.ac.ir on 2026-11-15]

[DOI: 10.47176/mjiri.40.106]

Table 3. Estimation of model coefficients under the assumption of independence, and Coefficients, Root mean square error, Coverage level (%), and *P*-value in hierarchical bootstrap and marginal approach in the Tehran Eye Study

Number of families	Coefficient estimation (H.B/M.A)	R.M.S.E (H.B/M.A)	Coverage% (H.B/M.A)	*P*-Value (H.B/M.A)
		$\hat{\delta} = -2.7$		
200	-2.51/-2.38	0.24/0.40	94.8/93.4	0823/0.041
		$\hat{\gamma} = 0.315$		
200	0.31/0.41	0.30/0.32	93.9/93.6	0.168/0.078
		$\hat{\gamma}_M = 0.324$		
200	0.31/0.30	0.28/0.20	94.7/92.6	0.711/0.002

H.B.: Hierarchical Bootstrap, M.A.: Marginal Approach, P-value based on two-sided binomial exact test

As shown in Table 3, the results showed that the bootstrap method generally produced more accurate estimates and better coverage probabilities than the marginal approach. The bootstrap method achieved the nominal 95% coverage for all parameters, whereas in the marginal approach only the confidence interval for γ reached this level.

Discussion

The findings suggest that the hierarchical bootstrap tended to provide somewhat improved coverage compared with the marginal approach across most sample sizes and under both exchangeable and non-exchangeable correlation structures.

RMSE values were frequently lower for the bootstrap estimates, indicating a tendency toward reduced bias, variance, or both. In the exchangeable setting, slight underestimation was observed in very small samples, with estimates gradually approaching the true values as the sample size increased. In contrast, the marginal approach showed a tendency toward overestimation once the sample size exceeded 50 families. In the non-exchangeable setting, estimation patterns were broadly similar, and although differences between methods were less pronounced, bootstrap estimates generally remained closer to the true values in many scenarios.

Given the family-based design and relatively small cluster sizes, a sample size of at least 200 families appears to be a reasonable practical choice. Accordingly, the Tehran Eye Study data were analyzed using 200 families. The empirical analysis aligned with the simulation results, indicating slightly more favorable performance of the bootstrap approach in terms of coverage and estimation accuracy.

The quadratic exponential model remains a useful framework for analyzing correlated binary outcomes. Nevertheless, concerns regarding irreproducibility have led some investigators to restrict analyses to equal cluster sizes (31), while others have either disregarded the issue (12, 13) or adopted adjustment strategies such as that proposed by Matthews et al. (9). In this study, even without explicitly addressing irreproducibility, the hierarchical bootstrap produced results broadly comparable to those obtained using Matthews' marginal approach for family sizes of 2 to 6 individuals, which are common in many populations.

A practical advantage of the hierarchical bootstrap is its simplicity. It can be implemented in standard statistical software with minimal additional programming, potentially enhancing accessibility for applied researchers. In our setting, it also remained computationally feasible at somewhat smaller sample sizes than the marginal approach. Furthermore, unlike previous work based largely on subjectively specified demographic scenarios (9), our simulation design was informed by the demographic structure of a real population, namely Tehran.

Several limitations should be acknowledged. A potential limitation in complex statistical models is the occurrence of non-convergence during parameter estimation. This issue may arise due to factors such as small sample size, sparse data, model over-parameterization, or strong correlations among parameters. Several methodological strategies have been proposed to address this problem. Penalized likelihood methods, such as ridge-type penalties or Firth's penalized likelihood, can stabilize parameter estimation and reduce bias in the presence of sparse data or separation (32, 33). Alternatively, Bayesian approaches incorporate prior distributions that can regularize parameter estimates and improve the stability of the estimation process, particularly in settings with limited or imbalanced data (34). Although these approaches were not implemented in the present study, they represent useful alternatives that could be considered in future research when convergence difficulties arise.

Bias, RMSE, and coverage probabilities were calculated conditional on successful convergence, as non-convergent replicates were excluded. If non-convergence occurs more frequently in more challenging datasets, this exclusion may introduce bias into the reported performance measures. Notably, unconditional performance (i.e., averaged over both convergent and non-convergent runs) may differ, and the higher non-convergence observed for the marginal method in the smaller-sample setting indicates reduced numerical stability. This limitation should be carefully considered in the interpretation of the results. The highest observed non-convergence rates were 6% for the bootstrap approach and 9% for the marginal approach in one scenario based on 1,000 simulated datasets. It is also important to note that we did not use the bootstrap procedure as a formal theoretical correction for irreproducibility, but rather as a robust inferential strategy that performs well empirically even when a single working model may not fully capture all aspects of the dependence structure.

In addition, our conclusions are confined to the scenarios examined; further evaluation under a wider range of

[Downloaded from mjiri.iums.ac.ir on 2026-11-15]

[DOI: 10.47176/mjiri.40.106]

dependence structures and cluster-size distributions would enhance generalizability. Future research could also explore alternative resampling strategies, such as the parametric bootstrap, although this may reduce simplicity and increase implementation complexity.

Finally, irreproducibility is an inherent logical property of the quadratic exponential model. Methodological refinements should therefore be regarded as practical adjustments aimed at mitigating its impact rather than as a definitive resolution.

Conclusion

In this study, we compared the hierarchical bootstrap and Matthews' marginal approach for inference in correlated family data under both exchangeable and non-exchangeable dependence structures. Overall, the hierarchical bootstrap tended to yield slightly more reliable interval estimation, with coverage probabilities generally closer to the nominal level and, in many scenarios, lower RMSE than the marginal approach.

However, in very small samples, neither method consistently achieved the nominal 95% coverage, and the bootstrap estimates occasionally exhibited modest underestimation. Given the typically small cluster sizes encountered in family-based studies, our findings suggest that analyses based on at least 200 families may represent a reasonable practical choice. The application to the Tehran Eye Study data was broadly consistent with the simulation findings.

Acknowledgment

The authors would like to thank the Research Council and the Ethics Committee of Sabzevar University of Medical Sciences for their support and for the approval of this study.

Conflict of Interests

The authors declare that they have no competing interests.

Authors' Contributions

M.H.R., A.F. and H.Z. contributed to conceptualization, methodology, conducting the research, and writing the original draft of the manuscript.

A.F. and H.H. also contributed to data collection. M.H.R. analyzed the data.

All authors contributed to supervision, validation, and reviewing and editing the manuscript.

All authors have read and approved the final version of the manuscript.

Ethical Considerations

Ethical approval was obtained from the Ethics Committee of Sabzevar University of Medical Sciences (IR.MEDSAB.REC.1403.117). In line with the Committee's stipulations, Tehran Eye Study data were released to the principal investigator in an anonymized format, with all direct and indirect identifiers removed, including names, national identification numbers, and any other information that could enable identification of individuals or their families.

Funding Support

This research received no external funding.

Data Availability

The R scripts and the data that support the findings of this study are available from the first author upon reasonable request. However, due to privacy and ethical considerations, the data are not publicly available.

AI Use Statement

The first author used ChatGPT 5.4 AI to assist with grammar correction, clarity, and language editing after the first round of peer review of the manuscript. The authors reviewed and edited all AI generated suggestions and take full responsibility for the final content of the manuscript.

References

1. Susser E, Susser M. Familial aggregation studies. A note on their epidemiologic properties. Am J Epidemiol. 1989;129(1):23-30.
2. Jang S, Kim ST, Kim Y-K, Song YH. Association of blood pressure and hypertension between parents and offspring: The Korea National Health and Nutrition Examination Survey. Hypertens Res. 2023;46(2):368-76.
3. John M, Orathel SP, Varghese J, Pradeep M, Jacob QM, Thomas R, et al. Familial Aggregation and ABO Blood Groups and COVID-19 Severity Among Hospitalized Patients. Infect Dis Clin Pract. 2024;32(2):e1335.
4. Kandasamy S, Chanchlani R. Familial Aggregation of Blood Pressure and the Heritability of Hypertension. In: Flynn JT, Ingelfinger JR, Brady TM, editors. Pediatric Hypertension. Cham: Springer International Publishing; 2023. p. 159-67.
5. Wei Y, Liu S, Andersson T, Feychting M, Kuja-Halkola R, Carlsson S. Familial aggregation and heritability of childhood-onset and adult-onset type 1 diabetes: a Swedish register-based cohort study. Lancet Diabetes Endocrinol. 2024;12(5):320-9.
6. Zöller B, Pirouzifard M, Holmquist B, Sundquist J, Halling A, Sundquist K. Familial aggregation of multimorbidity in Sweden: national explorative family study. BMJ Med. 2023;2(1):e000070.
7. Duggal P, Ladd-Acosta C, Ray D, Beaty TH. The Evolving Field of Genetic Epidemiology: From Familial Aggregation to Genomic Sequencing. Am J Epidemiol. 2019;188(12):2069-77.
8. Elston RC, Anne Spence M. Advances in statistical human genetics over the last 25 years. Stat Med. 2006;25(18):3049-80.
9. Matthews AG, Finkelstein DM, Betensky RA. Analysis of Familial Aggregation in the Presence of Varying Family Sizes. J R Stat Soc C Appl Stat. 2005;54(5):847-62.
10. Khoury MJ, Beaty TH, Cohen BH. Fundamentals of Genetic Epidemiology: Oxford University Press; 1993.
11. Austin MA, Beaty TH, Dotson WD. Genetic Epidemiology: Methods and Applications: CABI; 2013.
12. Hudson JI, Laird NM, Betensky RA. Multivariate logistic regression for familial aggregation of two disorders. I. Development of models and methods. Am J Epidemiol. 2001;153(5):500-5.
13. Hudson JI, Laird NM, Betensky RA, Keck PE, Jr., Pope HG, Jr. Multivariate logistic regression for familial aggregation of two disorders. II. Analysis of studies of eating and mood disorders. Am J Epidemiol. 2001;153(5):506-14.
14. Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G. Longitudinal Data Analysis (1st ed.): Chapman and Hall/CRC; 2008.
15. Liang K-Y, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika. 1986;73(1):13-22.
16. McNeish DM, Harring JR. Clustered data with small sample sizes: Comparing the performance of model-based and design-based approaches. Commun Stat B Simul Comput.. 2017;46(2):855-69.
17. Yu H-T, de Rooij M. Model Selection for the Trend Vector Model. J Classif. 2013;30(3):338-69.
18. Trégouët D-A, Ducimetiere P, Tiret L. Testing association between candidate-gene markers and phenotype in related individuals, by use of estimating equations. Am J Hum Genet. 1997;61(1):189-99.

[Downloaded from mjiri.iums.ac.ir on 2026-11-15]

[DOI: 10.47176/mjiri.40.106]

19. Hashemi H, Fotouhi A, Mohammad K. The Tehran Eye Study: research design and eye examination protocol. BMC Ophthalmol. 2003;3(1):8.
20. Betensky RA, Whittemore AS. An Analysis of Correlated Multivariate Binary Data: Application to Familial Cancers of the Ovary and Breast. J R Stat Soc C Appl Stat. 1996;45(4):411-29.
21. Rabbee N, Betensky RA. Power calculations for familial aggregation studies. Genet Epidemiol. 2004;26(4):316-27.
22. Zhao LP, Prentice RL. Correlated binary regression using a quadratic exponential model. Biometrika. 1990;77(3):642-8.
23. Davison AC, Hinkley DV. Bootstrap Methods and Their Application: Cambridge University Press; 1997.
24. Team RC. R: A Language and Environment for Statistical Computing 2023 [Available from: https://www.R-project.org/.
25. Bartolucci F, Pigini C. cquad: An R and Stata Package for Conditional Maximum Likelihood Estimation of Dynamic Binary Panel Data Models. J Stat Softw. 2017;78(7):1 - 26.
26. Deen M, de Rooij M. ClusterBootstrap: An R package for the analysis of hierarchical data using generalized linear models with the cluster bootstrap. Behav Res Methods. 2020;52(2):572-90.
27. Hashemi H, Fotouhi A, Yekta A, Pakzad R, Ostadimoghaddam H, Khabazkhoob M. Global and regional estimates of prevalence of refractive errors: Systematic review and meta-analysis. J Curr Ophthalmol. 2018;30(1):3-22.
28. Zhang J, Wu Y, Sharma B, Gupta R, Jawla S, Bullimore MA. Epidemiology and Burden of Astigmatism: A Systematic Literature Review. Optom Vis Sci. 2023;100(3):218-31.
29. Hashemi H, Hatef E, Fotouhi A, Mohammad K. Astigmatism and its determinants in the Tehran population: the Tehran eye study. Ophthalmic Epidemiol. 2005;12(6):373-81.
30. Agresti A. Categorical Data Analysis: Wiley; 2013.
31. Bartolucci F, Nigro V, Pigini C. Testing for state dependence in binary panel data with individual covariates by a modified quadratic exponential model. Econom Rev. 2018;37(1):61-88.
32. Firth D. Bias reduction of maximum likelihood estimates. Biometrika. 1993;80(1):27–38.
33. Heinze G, Schemper M. A solution to the problem of separation in logistic regression. Stat Med. 2002;21(16):2409–2419.
34. Gelman A, Jakulin A, Pittau MG, Su YS. A weakly informative default prior distribution for logistic and other regression models. Ann Appl Stat. 2008;2(4):1360–1383.

[Downloaded from mjiri.iums.ac.ir on 2026-11-15]

[DOI: 10.47176/mjiri.40.106]

Appendix

In this section, we focus on calculating the covariance—and consequently the correlation—between observations of family members under the two resampling strategies proposed by Davison and Hinkley (1).

Suppose a general linear model for a two-stage nested design

$y_{ij} = x_i + z_{ij}$ i=1,2,…,m families

j=1,…,n_i individuals in family i.

The variables $x_i$represent cluster-specific characteristics unique to the i-th cluster and are assumed to be independently drawn from a distribution denoted by F_X.

In this model :

$E(x_i) = 0$, $Var(x_i) = \sigma_x^2$

The differences among the $x_i$values arise from variability between families ($Var(x_i) = \sigma_x^2$)

The z_{ij}values represent characteristics specific to the j-th member of the i-th family and are assumed to be drawn from another distribution, F_Z.

$E(z_{ij}) = 0$, $Var(z_{ij}) = \sigma_z^2$

$Var(y_{ij}) = \sigma_x^2 + \sigma_z^2$ and $Cov(y_{ij}, y_{ik}) = \sigma_x^2$ forj ≠ k

Therefore $Corr(y_{ij}, y_{ik}) = \frac{\sigma_x^2}{(\sigma_x^2 + \sigma_z^2)}$

Since this correlation structure among family members complicates the resampling process, they adopted two different strategies. In the first step, a random sample of families is drawn with replacement. The parents are kept fixed. In the next step, a random sample is drawn from the children either without replacement (Strategy 1) or with replacement (Strategy 2).

Shin (2) extended their calculations to the case of varying family sizes. In this case

$$Cov(y_{ij}, y_{ik}) = \begin{cases} \frac{SS_B}{m} - \frac{1}{m}\sum_{i=1}^{m} \frac{SS_{w(i)}}{n_i(1-n_i)} & \text{Strategy1} \\ \frac{SS_B}{m} & \text{Strategy 2} \end{cases}$$

Where $SS_B = \sum(\bar{y}_i - \bar{y}_.)^2$ and $SS_{W(i)} = \sum_i \sum_j (y_{ij} - \bar{y}_i^{\,2})^2$

And finally, the expected value of this covariance is given by

$$E\left(Cov(y_{ij}, y_{ik})\right) = \begin{cases} \sigma_x^2\left[1 - \frac{\sum_i n_i^2}{(\sum_i n_i)^2}\right] - \frac{\sigma_z^2}{\sum_i n_i} & \text{Strategy 1} \\ \sigma_x^2\left[1 - \frac{\sum_i n_i^2}{(\sum_i n_i)^2}\right] + \frac{\sigma_z^2}{m}\left[\sum \frac{1}{n_i} - \frac{m}{\sum_i n_i}\right] & \text{Strategy 2} \end{cases}$$

1. Davison AC, Hinkley DV. Bootstrap Methods and Their Application: Cambridge University Press; 1997.
2. Shin J. Evaluation of Candidate Genes in Family Studies [master's thesis]. Toronto: University of Toronto; 1998.

[Downloaded from mjiri.iums.ac.ir on 2026-11-15]

[DOI: 10.47176/mjiri.40.106]