Journal of Data Science 23 (1), 62–63
January 2025
DOI: 10.6339/25-JDS1161B
Discussion

Discussion of "Power Priors for Leveraging Historical Data: Looking Back and Looking Forward"☆

Panpan Zhang[1,2,*]

[1]Department of Biostatistics, Vanderbilt University Medical Center, Nashville, TN, U.S.A.
[2]Vanderbilt Memory & Alzheimer's Center, Nashville, TN, U.S.A.

In the rapidly evolving world of healthcare, the ability to effectively analyze and leverage the vast amounts of available patient data is becoming increasingly important. The accelerated growth of electronic health records and other historical data sources has provided researchers and clinicians with access to valuable longitudinal information about patients and new opportunities to enhance understanding of disease progression, risk factors, and the effectiveness of various treatments.

Power priors have emerged as a powerful tool for incorporating historical data into Bayesian analyses, offering a flexible and principled way to leverage relevant prior information (Ibrahim and Chen, 2000; Chen and Ibrahim, 2006; Ibrahim et al., 2015). In practice, the judicious use of power priors can significantly enhance the statistical power of statistical analysis, allowing researchers to achieve more robust and reliable conclusions (Ibrahim et al., 2015).

One of the major advantages of power priors is their ability to adaptively weight the historical information based on its relevance. This is particularly valuable in the context of clinical trials, where data from previous studies can provide unique insights, but may not be directly comparable due to differences in patient populations, study designs, or other factors. By incorporating power priors, researchers can strike a balance between the information contained in the current data and the historical evidence. This approach can lead to more accurate parameter estimates, increased statistical power, and more informed decision-making, ultimately leading to improved patient outcomes.

Highlights We congratulate Chen et al. (2025) for recently completing a high-quality, comprehensive review of power priors. Since the last systematic review in Ibrahim and Chen (2000), especially after the release of the FDA (Food and Drug Administration) regulations about guidance for the use of Bayesian statistics in medical device clinical trials in 2010, there have been tremendous methodological advances and novel applications in using power priors, which were summarized in a large body of literature given in Chen et al. (2025). This succinct, but complete, summary enables researchers interested in developing new methods and theories on power priors to use Chen et al. (2025) as a valuable resource for literature search.

Different from most review articles that were primarily based on text summaries, Chen et al. (2025) leveraged two case studies, the Kochiba and National Toxicology Program (NTP) and the Alzheimer's Disease Neuroimaging Initiative (ADNI), as entry points to introduce the use of power priors in major classes of regression models: binomial regression and normal linear regression. This presentation scheme not only took into account the practical applicability of power priors, but also helped new researchers to this burgeoning field better understand the underlying concepts.

☆Main article: https://doi.org/10.6339/24-JDS1161.
* Email: panpan.zhang@vumc.org.

In addition to the normalized power priors and their extension to multiple historical data sets, Chen et al. (2025) introduced and compared four variants of power priors; namely, they are partial borrowing partial priors, borrowing-by-parts power priors, partial borrowing-by-parts power priors, and propensity-score-based power priors. Mathematical formulas for these variants of power priors were provided, and the evaluations and comparisons were clearly demonstrated through applications to the Knochiba and NTP and ADNI data. What was particularly valuable was that Chen et al. (2025) summarized available `R` packages and `SAS` procedures for model fitting that incorporated historical data with power priors, and shared `R` and `SAS` codes for empirical studies on GitHub for reproducibility.

Discussion Although Chen et al. (2025) provided a unique and valuable review of power priors, some important issues regarding the appropriate use of power priors were not discussed extensively in depth in the paper. One of the most important questions is how to optimally choose the discounting parameter α_0 such that the influence (or over-borrowing) of historical data is well controlled. This was briefly discussed in the last section of Chen et al. (2025), but a systematic sensitivity analysis of α_0 may be needed to further improve understanding of how to use this powerful tool correctly. Another practical issue is how to deal with heterogeneity when multiple sources of historic data from extremely diverse populations or study designs are available to use. While Chen et al. (2025) briefly introduced the extension of normalized power priors for multiple historical data sets, an extended, in-depth discussion alongside numerical studies would be helpful. Besides, developing rigorous statistical methods to select useful historical data in analyses may be critical for properly using power priors and reducing the influences of data noise and multi-source heterogeneity. Other potential challenges regarding result interpretation and computational complexity in power priors research also warrant further investigation.

References

Chen MH, Guan Z, Lin M, Sun M (2025). Power priors for leveraging historical data: Looking back and looking forward. *Journal of Data Science*, 23(1): 1–30. https://doi.org/10.6339/24-JDS1161

Chen MH, Ibrahim JG (2006). The relationship between the power prior and hierarchical models. *Bayesian Analysis*, 1(3): 551–574. https://doi.org/10.1214/06-BA118

Ibrahim JG, Chen MH (2000). Power prior distributions for regression models. *Statistical Science*, 15(1): 46–60. https://doi.org/10.1214/ss/1009212673

Ibrahim JG, Chen MH, Gwon Y, Chen F (2015). The power prior: Theory and application. *Statistics in Medicine*, 34(28): 3724–3749. https://doi.org/10.1002/sim.6728