# Article information:

Optimal sampling for design‐based estimators of regression models - Chen - 2022 - Statistics in Medicine - Wiley Online Library
<https://onlinelibrary.wiley.com/doi/10.1002/sim.9300>

# Article summary:

1. Electronic health records (EHR) are increasingly used to study research questions, but variables within these databases can be error-prone.

2. Design-based methods produce robust estimations for fitting regression models to two-phase stratified sampling designs.

3. This article derives a closed-form solution of the optimal design for analysis via generalized raking estimators and compares it with the optimal allocation for analysis by the IPW estimator and other commonly used sampling designs.

# Article rating:

Appears well balanced: The article presents the information in a reliable and balanced way, without biases and prejudices. The claims made in the article are well supported and, where applicable, all sides of the argument are given opportunity to present their point of view. The article appears trustworthy and reliable.

# Article analysis:

The article is written in an objective manner and provides a comprehensive overview of the current state of research on optimal sampling for design-based estimators of regression models. The authors provide a clear explanation of the different types of design-based estimators, as well as their advantages and disadvantages compared to other estimation methods. The authors also provide a detailed description of how to derive the optimal design for analysis via generalized raking estimators, which is supported by numerical studies and an interactive Shiny app.

The article does not appear to have any biases or one-sided reporting, as it presents both sides equally and provides evidence for its claims. Furthermore, all potential risks are noted and discussed in detail throughout the article. The only potential issue is that some counterarguments may have been unexplored or underrepresented in comparison to other points made in the article. However, this does not detract from its overall trustworthiness and reliability.

# Topics for further research:

* Design-based estimators of regression models
* Optimal sampling strategies
* Generalized raking estimators
* Design-based estimation methods
* Design-based estimator advantages and disadvantages
* Design-based estimator numerical studies

# Report location:

<https://www.fullpicture.app/item/045cb77bba6e7c0f9d7fc3151590249d>