

# Title: A Comparative Analysis of Bayesian Sampling Methodologies for Design Space Identification in Quality by Design (QbD) Approach

Deniz Akinc Abdulhayoglu<sup>(a,c)</sup>, Jean-François Michiels<sup>(b,c)</sup>

- (a) deniz.akinc@pharmalex.com, Senior Manager in Cencora Pharmalex, Global Statistics and Data Science (GSDS)  
Bio: She has been working as a statistical consultant in Cencora Pharmalex since 2021. She has a PhD degree in statistics (KULeuven, Belgium). She is mostly involved in CMC projects related to several areas of pharmaceutical research and industry including bioassay and process development/validation, drug and vaccine discovery and manufacturing, with the usage various statistical tools such as design of experiments and advanced statistical models.
- (b) jean-francois.michiels@pharmalex.com, Associate Director in Cencora Pharmalex, GSDS
- (c) Cencora Pharmalex GSDS Belgium, Rue Edouard Belin 5, 1435 Mont-St-Guibert, Belgium

## Abstract:

Quality by Design (QbD) has emerged as a pivotal framework in the pharmaceutical industry, emphasizing proactive approaches to ensure product quality. Central to QbD is the identification of a robust design space, encompassing the range of input variables and process parameters that guarantee pharmaceutical product quality. In this study, we present a comparative analysis of random walk sampling used in Markov Chain MCMC and nested sampling methodologies employed in design space identification within the QbD paradigm.

Random walk sampling, a key component of MCMC methods, has been widely used in Bayesian inference for its simplicity and effectiveness in exploring parameter space. However, the efficacy of random walk sampling can be affected by the presence of high-dimensional and multimodal distributions, which are common in complex pharmaceutical processes.

Nested sampling offers an alternative approach to sampling from complex probability distributions by focusing on the marginal likelihood. Unlike MCMC, nested sampling systematically improves the estimation of the evidence, enabling more efficient exploration of the parameter space, particularly in scenarios with a high number of design factors.

Drawing on insights from the work of Kusumo et al. (2019) on a Bayesian approach to probabilistic design space characterization using nested sampling strategy, our comparative analysis evaluates the strengths, limitations, and applicability of classical random walk sampling in MCMC and nested sampling in the context of design space identification. We consider factors such as computational efficiency, accuracy of parameter estimation, scalability to high-dimensional spaces, and robustness to multimodal distributions. Furthermore, we discuss practical considerations for selecting appropriate methodologies based on the specific characteristics of the pharmaceutical process under investigation.

Through case studies and theoretical discussions, we illustrate the strengths and limitations of random walk sampling in MCMC and nested sampling methodologies for design space identification within the QbD framework. Our findings aim to provide insights into the comparative performance of these methodologies and inform researchers, practitioners, and regulatory authorities involved in pharmaceutical development about their suitability for achieving the goals of QbD.

This study contributes to advancing the understanding of methodologies for design space identification in QbD and facilitates informed decision-making in pharmaceutical process optimization and quality assurance.

Ref: Kusumo, K. P., Gomoescu, L., Paulen, R., García Muñoz, S., Pantelides, C. C., Shah, N., & Chachuat, B. (2019). Bayesian approach to probabilistic design space characterization: A nested sampling strategy. *Industrial & Engineering Chemistry Research*, 59(6), 2396-2408.