Bayesian Model Selection for Determining the Biokinetics of Zirconium in the Human Body after Ingestion

Sabine Hug[1], Daniel Schmidl[1,2], Wei Bo Li[3], Matthias Greiter[3] & Fabian J. Theis[1,2]

[1] Institute for Bioinformatics and Systems Biology, Helmholtz Zentrum München
Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany

[2] Institute for Mathematical Sciences, Technische Universität München, Munich, Germany

[3] Research Unit Medical Radiation Physics and Diagnostics, Helmholtz Zentrum München
Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany

Extended Abstract

Multi-compartment models based on ordinary differential equations (ODEs) are well established tools for simulating complex systems evolving over time [5]. However, although they are very straightforward and easily interpretable, determining the compartment structure and interaction mechanisms can be a very daunting task. In this context, Bayesian model selection methods are an extremely useful tool for evaluating different kinds of models. In contrast to much of the frequentists' methodology, which is generally based on asymptotic approximations for large sample sizes and single best parameter value evaluations [1,10], the Bayesian framework provides a fully probabilistic approach that easily incorporates both model and parameter uncertainty [3]. Bayesian analysis is well-grounded on the so-called posterior distribution, i.e. the probability distribution of a problem specific parameter space conditioned on given data. This specifies a measure of belief for all possible parameter values. The posterior distribution is proportional to the product of the data likelihood and the parameter prior distributions, allowing to easily include a priori information into the modeling process. We here present a method for computing Bayes Factors between pairs of models for performing the task of model selection, i.e. we calculate the marginal likelihood for each of the models. This is particularly useful when there are just a few models to choose from. Due to the evaluation of each model on the whole parameter space, the Bayes factor also naturally penalizes model complexity and thus prevents overfitting issues [16,17]. The selection process is based on the inference of the model's parameters that represent the reaction rates governing the mass transport between the compartments. As it is computationally very difficult to compute the marginal likelihood of the model based on the model's parameters, we split the computation into several intermediate steps where not the full marginal likelihood has to be evaluated, but we only have to draw independent samples from the so-called power posterior [2,10]. The power posterior is similar to the general posterior density except for the introduction of an auxiliary "temperature" parameter that governs the influence of the parameter likelihood. This approach is known as thermodynamic integration [4,11]. It stabilizes the numerical evaluation of the marginal likelihood for each model by computing the expectation with respect to the temperature based power posterior. This expectation was calculated via Monte Carlo sampling using a copula based Metropolis-Hastings algorithm [18].

With the help of our method, we could use existent prior knowledge as well as measurements and sampling results to extend a model put forward by the Helmholtz Zentrum München (HMGU) for the processing of Zirconium in the human body after ingestion [12,13,14,15]. In radiation sciences, models like these are of crucial importance for dose prediction for workers who are exposed to radioactive substances. They predict limiting values for detrimental effects. Since the problem at hand is of great relevance, quite a large number of analyses such as mice experiments exist. These yielded excessive prior knowledge for our Bayesian modeling approach. Moreover, we compared the HMGU

as well as the extended HMGU model with the well-established standard model put forward by the International Commission on Radiological Protection (ICRP) [7,8,9].

We found that the classical HMGU model failed to adequately capture part of the data. Using Bayesian model inference, we were able to alleviate the problem by adding a new compartment, which also made the extended HMGU model physiologically more plausible. We challenged both versions of the HMGU model with the ICRP model by calculating the Bayes factors based on data from sixteen healthy human subjects obtained through the double tracer technique [6,13,14]. Since these sixteen datasets display quite diverse behavior, a Bayes factor was computed separately for every subject. This is also in agreement with the biological knowledge that the reaction rates can differ between humans. We found that the extended HMGU model best fits the data and is favored by the Bayes factors both over the original HMGU model and over the ICRP model. Our new model can readily be used to improve detrimental effect limit predictions.

Figures

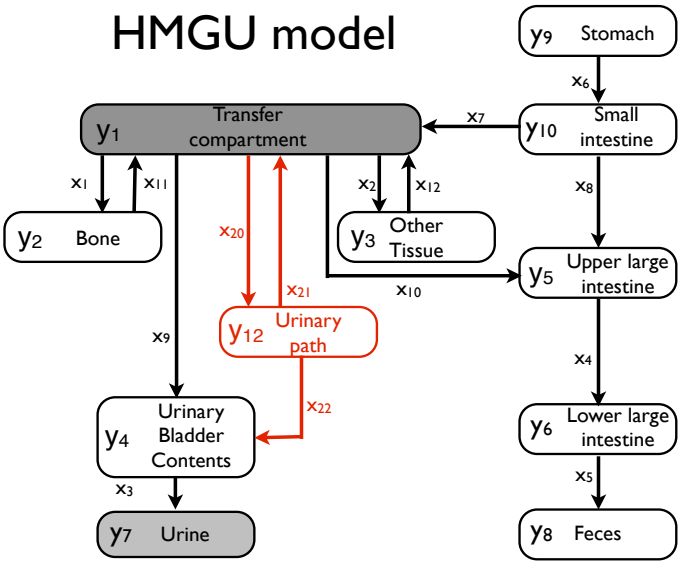

Figure 1: Schematic representation of the HMGU model for the biokinetics of zirconium consisting of ten compartments $y_1, \ldots, y_{10}$ and twelve rate flows $x_1, \ldots, x_{12}$. The extended HMGU model additionally contains the compartment y_{12} and the flow rates x_{20}, x_{21} and x_{22} (shown in red). In either model zirconium enters the body in the stomach compartment y_9 and diffuses through the system until it reaches either one of the two final compartments urine, y_7, or feces, y_8. The gray compartments y_1 and y_7 are directly related to the data points measured.

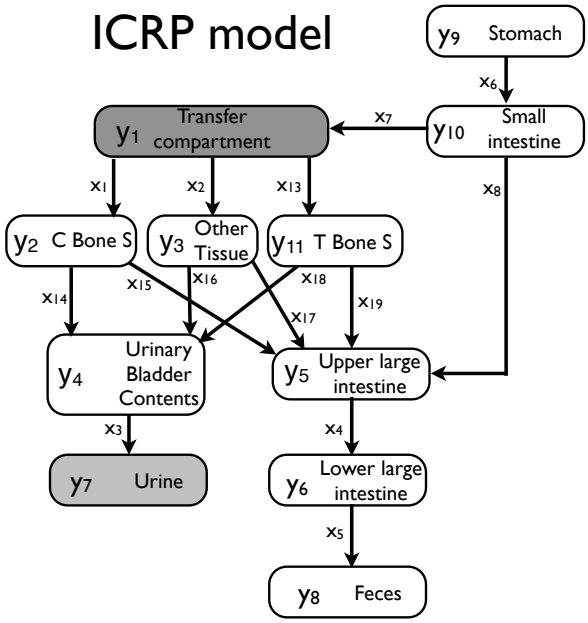

Figure 2: The ICRP model for the biokinetics of zirconium consists of eleven compartments $y_1, \ldots, y_{11}$ and 15 time independent transfer rates $x_1,\ldots,x_8,x_{13},\ldots,x_{19}$. Zirconium enters the body in the stomach compartment y_9 and diffuses through the system until it reaches either one of the two final compartments urine, y_7, or feces, y_8. The gray compartments y_1 and y_7 are directly related to the data points measured.

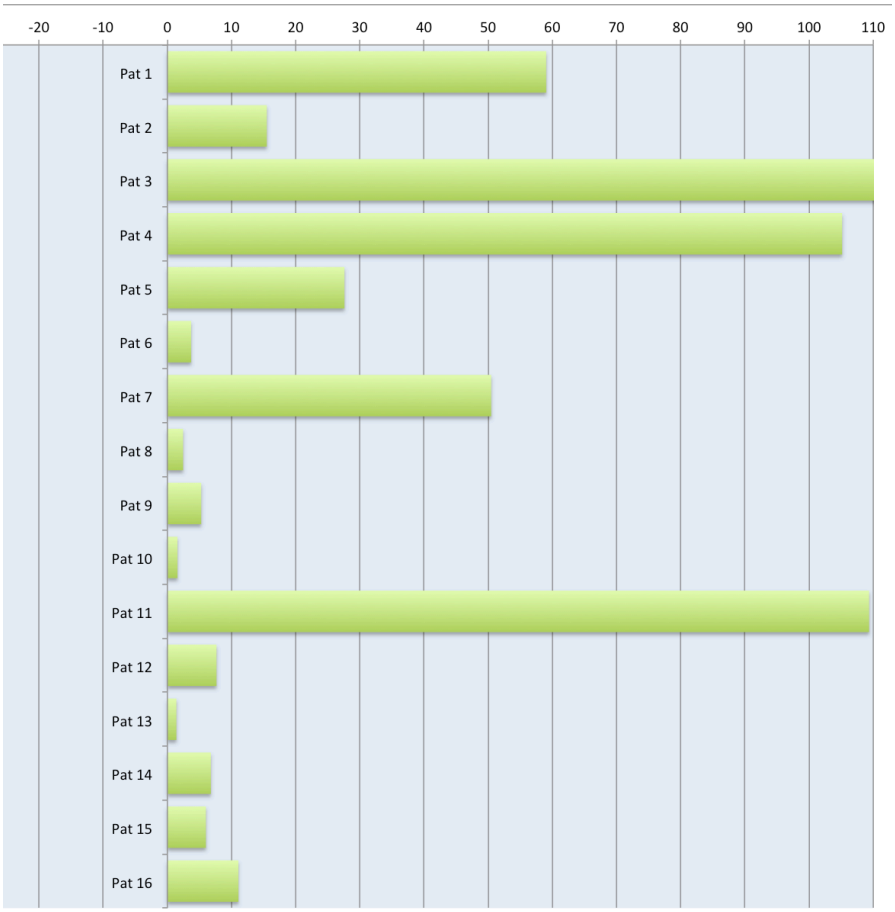

Figure 3: Bayes Factors for the extended HMGU vs. the ICRP model. On the vertical axis the 16 patients are depicted from Patient 1 (top) to Patient 16 (bottom). Positive values denote a preference for the HMGU model, negative for the ICRP model. A value bigger than 3 indicates substantial preference of the improved HMGU model.

References

[1] J. O. Berger and M. J. Bayarri, "The Interplay of Bayesian and Frequentist Analysis," *Statistical Science*, vol. 19, no. 1, pp. 58-80, Feb. 2004.

[2] B. Calderhead and M. Girolami, "Estimating Bayes factors via thermodynamic integration and population MCMC," *Computational Statistics & Data Analysis*, vol. 53, no. 12, pp. 4028-4045, Oct. 2009.

[3] M. Clyde and E. I. George, "Model uncertainty," *Statistical Science*, vol. 19, no. 1, pp. 81-94, Feb. 2004.

[4] N. Friel and a N. Pettitt, "Marginal likelihood estimation via power posteriors," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 70, no. 3, pp. 589-607, Jul. 2008.

[5] K. Godfrey, *Compartmental Models and their Applications*. 1983.

[6] M. B. Greiter, A. Giussani, V. Höllriegl, W. B. Li, and U. Oeh, "Human biokinetic data and a new compartmental model of zirconium - A tracer study with enriched stable isotopes.," *The Science of the total environment*, vol. 409, no. 19, pp. 3701-3710, Jul. 2011.

[7] ICRP, "Age-dependent doses to members of the public from intake of radionuclides: Part 1," 1989.

[8] ICRP, "Limits for intakes of radionuclides by workers," 1990.

[9] ICRP, "Report of the task group on reference man," 1975.

[10] R. E. Kass and A. Raftery, "Bayes Factors," *Journal of the American Statistical Association*, 1995.

[11] N. Lartillot and H. Philippe, "Computing Bayes factors using thermodynamic integration.," *Systematic biology*, vol. 55, no. 2, pp. 195-207, Apr. 2006.

[12] W. B. Li and C. Hoeschen, "Uncertainty and sensitivity analysis of biokinetic models for radiopharmaceuticals used in nuclear medicine.," *Radiation protection dosimetry*, vol. 139, no. 1-3, pp. 228-31, 2010.

[13] W. B. Li, M. Greiter, U. Oeh, and C. Hoeschen, "Reliability of a New Biokinetic Model of Zirconium in Internal Dosimetry. Part I. Parameter Uncertainty Analysis," *Health Physics*, 2011.

[14] W. B. Li, M. Greiter, U. Oeh, and C. Hoeschen, "Reliability of a New Biokinetic Model of Zirconium in Internal Dosimetry. Part II. Parameter Sensitivity Analysis," *Health Physics*, 2011.

[15] W. B. Li, V. Höllriegl, P. Roth, and U. Oeh, "Influence of human biokinetics of strontium on internal ingestion dose of 90Sr and absorbed dose of 89Sr to organs and metastases.," *Radiation and environmental biophysics*, vol. 47, no. 2, pp. 225-39, Apr. 2008.

[16] T. Lodewyckx, W. Kim, M. D. Lee, F. Tuerlinckx, and P. Kuppens, "A tutorial on Bayes factor estimation with the product space method," *Journal of Mathematical Psychology*, 2011.

[17] I. J. Myung and M. A. Pitt, "Applying Occam ' s razor in modeling cognition : A Bayesian approach," *Psychonomic Bulletin & Review*, vol. 4, no. 1, pp. 79–95, 1997.

[18] D. Schmidl, C. Czado, and F. J. Theis, "A novel copula based MCMC sampler for highly dependent dynamical systems," *in preparation*.