Technical report "On Combining in-silico and in-vivo Experiments"

Report is on the proposal under acronym STriTuVaD
In Silico Trial for Tuberculosis Vaccine Development
Call: H2020-SC1-2016-2017 (Personalized Medicine)
Topic: SC1-PM-16-2017
Type of action: Research and Innovation action
Proposal number: 777123
Deadline Id: H2020-SC1-2017-CNECT-2
Digital acknowledgement of receipt is 2017.03.14

Grant proposed by the line manager Miguel Juarez from the University of Sheffield, UK

Author of the report: Dr Ksenia N. Kyzyurova
School of Mathematics and Statistics,
The University of Sheffield, UK
kseniak.ucoz.net, ksenia.kyzyurova@gmail.com

October 17, 2019

Abstract

An idea of augmenting data obtained from real patients (more generally, in-vivo experiments on living beings) diagnosed or expected to have a particular medical condition with the data from in-silico trials, that is, computer generated experiments of this medical condition, is considered. The mathematically coherent way of the augmentation is described in this report. The proposed mathematical solution belongs to the framework of the Bayesian inversion of the computer model. Discussion of the applicability of the proposal to the treatment of Tuberculosis appears in the end of the report.

1 Introduction

This document reports on the research during the author's work as a Research Associate in Bayesian Models for Synthetic Trials at the University of Sheffield, UK. The report focuses on a project of developing mathematical and computational modelling techniques for in-silico-augmented clinical trials for Tuberculosis vaccination development. The next section provides the mathematical description of the problem in this project with the proposed solution described in section 3.

2 Mathematical description of the project idea

Two groups of patients are imagined to "participate" in the medical study. First group consists of real patients or living organisms from an in-vivo experiment.[1] An in-vivo experiment is assumed to be performed under some

[1] "In-vivo experiment" is an experiment on a living organism.

conditions that are peculiar to a patient and, therefore, are measurable parameters $\boldsymbol{\phi}$ of this patient; and some conditions that are known to exist, but are not measurable on a patient. Let such conditions be described by parameters $\boldsymbol{\theta}$. One example of such a condition, relevant to the discussion of the Tuberculosis vaccination development, is the dose of a specific drug.

We imagine that there exist a type of cells which are responsible for a patient's medical condition; and that along with the measurements of the known conditions ϕ, the measurements of the number of particular cells responsible for the medical condition are taken.

A real patient is imagined to be observed at several time points, say $t_1, t_2, \ldots, t_n$ during some time period T; that is, $\gamma(\boldsymbol{\phi}(t), \boldsymbol{\theta}(t))$ is the number of cells of a patient observed at time t together with the corresponding parameters $\boldsymbol{\phi}(t)$ and $\boldsymbol{\theta}(t)$.

Suppose now that a mathematical computer model of a virtual living being exists, such that for any set of fixed conditions $\boldsymbol{\phi}(t)$ and $\boldsymbol{\theta}(t)$ at time t the model produces the number of cells $\varkappa(\boldsymbol{\phi}(t), \boldsymbol{\theta}(t))$ for this living being.

The research objective in this project is to find meaningful purposes and ways of how the information from the computer model of the virtual living being (or a virtual patient), that is, the in-silico experiment may be used to augment data from the in-vivo experiment on a living being.

Consider first the in-silico experiments. Both sets of conditions described by parameters $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ are known, and therefore parameters $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ are fixed for every virtual living being at any time point t. Therefore, the distribution over these parameters is considered to be uniform. This means that for the computer model any fixed set of these parameters is equally likely to occur.

This is in contrast to the distribution of $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ in a set of real living beings. The distribution of these conditions in living organisms is not known but is not expected to be uniform either. Therefore, we assume that this distribution is not available to us and that this assumption is the correct reflection of the reality. Arguably, this distribution of parameters $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ in living beings is not possible to know, because, for example, it may change with time and may be difficult to extract.

Additionally, there always exist some other conditions $\boldsymbol{\eta}$ in a group of real living beings that are not captured by researchers and therefore bring unknown *noise* in the measurements. However, we hope that trustworthy researchers and scientists try to minimize the amount and the uncertainty from the rest of uncontrolled conditions $\boldsymbol{\eta}$.

One meaningful purpose is to use information from the in-silico experiments is for future prognosis of a given living organism with respect to his current state given by parameters $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ and measurements γ, possibly obtained at several time points $1:T$.

Of interest is to find parameters $\boldsymbol{\theta}_{1:T}$ of a real patient which must be plugged-in to the computer model to find the number of cells $\varkappa$ achieved by a corresponding cohort of in-silico patients and to learn possible future outcomes of a real patient. The tangible way is to perform *calibration* of the model with respect to information from one or more of the previous or current conditions of the patient. The methodological description of the calibration within the Bayesian inverse framework has been laid out in [1,2,3]. The next section provides an illustration of the calibration of a particular model which generates a certain number of cells given some inputs — parameters — to the model.

Note that augmenting in-vitro experiments (that is, experiments in the laboratory environment) with the in-silico trials is mathematically absolutely the same as those for the augmenting of in-vivo experiments with in-silico trials, which are described in this report.

3 Calibration of in-silico experiment with respect to an in-vivo experiment

This section provides the mathematical solution to augmenting in-vivo experiments with in-silico experiments. For illustrative purpose of how calibration of an in-silico experiment may be done, the model of division of cells

described in the monograph "Stochastic processes" (see [4]) is considered. Analytical description of the model and many properties of the model is given in the monograph.

In this report the model, which takes two parameters $0 < a < 1$ and $0 < b < 1$ which are subject to the constraint $a + b < 1$ and one more parameter $\psi \in \mathbb{N}_{\geq 1} = \{1, 2, \ldots, \infty\}$, and produces the number of cells $\varkappa(a, b, \psi)$. Without loss of generality the parameters $\boldsymbol{\phi} \in \varnothing$ are assumed to form an empty set of conditions peculiar to a living being; $\boldsymbol{\theta} = (a, b, \psi)$ are the parameters of interest, such that $\theta(t) = \theta$ for any time point $t \geq 0$. That is, parameters θ are constant in time.

$$P(\boldsymbol{\theta} \mid \gamma_{1:T}) = \frac{P_{\varkappa}(\gamma_{1:T} \mid \boldsymbol{\theta}) P(\boldsymbol{\theta})}{\sum_{\boldsymbol{\theta} \in \Theta} P_{\varkappa}(\gamma_{1:T} \mid \boldsymbol{\theta}) P(\boldsymbol{\theta})} \overset{\text{if } P(\boldsymbol{\theta}) = \text{const}}{=} \frac{P_{\varkappa}(\gamma_{1:T} \mid \boldsymbol{\theta})}{\sum_{\boldsymbol{\theta}} P_{\varkappa}(\gamma_{1:T} \mid \boldsymbol{\theta})}, \tag{1}$$

where $P_{\varkappa} = P_{\varkappa|\boldsymbol{\theta}}$ denotes the distribution of the number of cells given parameters $\boldsymbol{\theta}$ provided by a corresponding computer model. $P_{\varkappa}(\gamma_{1:T}) = \sum_{\boldsymbol{\theta} \in \Theta} P_{\varkappa}(\gamma_{1:T} \mid \boldsymbol{\theta}) P(\boldsymbol{\theta})$, where $\boldsymbol{\theta} \in \Theta$, discrete space of parameters, is the marginal distribution of measurements $\gamma_{1:T}$.[2]

4 Conclusion

This report has presented the mathematically coherent framework for augmenting in-vivo experiments with in-silico experiments. Namely, calibration of the model as the mathematical solution to the research problem has been given.

4.1 Discussion

The computer model proposed by the collaborators of Miguel Juarez generates B cells (an immunological term) as its output. Performing the literature review on the immunology, the author found out that B cells do not participate in treatment of Tuberculosis (see, e.g. [5]). The author is not aware of scientists or anyone interested in developing a biologically coherent model of B cells production as well. The author herself lacks the required biological knowledge for developing such a model.

The author acknowledges that the most difficult part of data analysis, including analysis of data from in-vivo and in-silico experiments, is to obtain adequate and trustworthy data. These are not available for the project of Tuberculosis vaccination development. Due to lack of a meaningful scientific model and valid data, the implementation of the mathematical solution for the treatment of Tuberculosis on real patients is not possible to perform. The conclusion of this report is in agreement with the objective of the call H2020-SC1-2016-2017 (Personalized Medicine) to decrease the size and length of clinical trials.

5 Bibliography

[1] Kseniia N. Kyziurova, 2016, Calibration of To-be Coupled Computer Models via a Linked Emulator. *Los Alamos database*, LA-UR-16-26622.

[2] Ksenia Kyzyurova, 2017, On Uncertainty Quantification for Systems of Computer Models, *Philosophy doctorate thesis*, Duke University.

[3] Ksenia Kyzyurova, 2019, Calibration of mathematical computer models. `http://kseniak.ucoz.net`.

[4] Ksenia Kyzyurova, 2019, Stochastic processes. `http://kseniak.ucoz.net`.

[5] R.A. Goldsby, T.J. Kindt, B.A. Osborne, J. Kuby "Immunology" — 5th ed.

[2] While we may consider Θ to be a continuous space of parameters, in practice, for computationally challenging computer models, the space of parameters Θ is discrete.