Bayesian Framework to Augment Tumor Board Decision Making

Stefano Pasetto, PhD[1]; Robert A. Gatenby, MD[1,2]; and Heiko Enderling, PhD[1,3]

original reports

abstract

PURPOSE Ideally, specific treatment for a cancer patient is decided by a multidisciplinary tumor board, integrating prior clinical experience, published data, and patient-specific factors to develop a consensus on an optimal therapeutic strategy. However, many oncologists lack access to a tumor board, and many patients have incomplete data descriptions so that tumor boards must act on imprecise criteria. We propose these limitations to be addressed through a flexible but rigorous mathematical tool that can define the probability of success of given therapies and be made readily available to the oncology community.

METHODS We present a Bayesian approach to tumor forecasting using a multimodel framework to predict patient-specific response to different targeted therapies even when historical data are incomplete.

RESULTS We demonstrate that the Bayesian decision theory's integrative power permits the simultaneous assessment of a range of therapeutic options.

CONCLUSION This methodology proposed, built upon a robust and well-established mathematical framework, can play a crucial role in supporting patient-specific clinical decisions by individual oncologists and multi-specialty tumor boards.

INTRODUCTION

The treatment or treatment combinations for individual patients with cancer are often determined by a tumor board of physicians from different specialties such as surgery, pathology, medical oncology, and radiation oncology. The doctors' knowledge and experience, available published studies, and facilities accessibility in the treatment center or hospital or clinic guide the decisional process. Expertise and opinions converge to form, in a collective decisional effort, the *optimal treatment*. Although the combined clinical and empirical knowledge of tumor board members yields improved outcomes, the decision-making process is often imprecise, particularly when a patient's status does not match cohorts in prior clinical investigations. Furthermore, many physicians do not have access to the multidisciplinary expertise of a tumor board.

With the growing amount of data collected for individual patients and cancer populations, a general and robust mathematical framework may contribute to a reproducible clinical decision using a reliable decisional algorithm. Ideally, such an algorithm would systematically and rigorously integrate patient-specific data with published cohort studies and large-scale population data from multiple institutions to predict treatment response with potentially adverse effects from all available clinical options. Such algorithms are not built to replace the oncologists and medical expertise; instead, they are proposed to help integrate and rigorously analyze the ever-increasing amount of data on highly heterogeneous diseases with significant interperson and intraperson heterogeneity for informed clinical decision making.

Historical examples in this direction are available already since late 1980s.[1,2] Nowadays, artificial intelligence is used in evidence-based learning to support the decision-making process.[3] The most notable example of this approach is probably the IBM Watson Health Program,[4] although less complicated online applications based on statistical indicators inclusive or not of past or modern genomic tests (eg, Oncotype DX, but see also Mamma/BluePrint) such as Adjuvant! or PREDICT became available much earlier,[5-7] not without skepticism.[8] However, mechanistic mathematical modeling akin to clinical decision making involves many degrees of freedom, or variables and parameters, such that models with different biologic assumptions can simulate the selfsame data sets. When applied to patient-specific clinical data, this

CONTEXT

Key Objective

How to decide on the optimal treatment for a patient? Our approach aims to rate tumor-board-preselected optimal treatments with a Bayesian statistical tool rather than determine optimal treatment through externally specific indexes.

Knowledge Generated

We highlight the importance of the Bayesian model comparison as a useful statistical framework in the tumor board decisional process. Its ability to naturally weigh a patient's state and medical doctors' knowledge represents critical complementary support to oncology work.

Relevance

The tool is designed to support a tumor board decisional process. Throughout the access to the proposed cloud-computing system, an oncologist will be able to insert the patient's information and receive the most successful therapeutic path that has already been applied in the literature.

assortment of models and model predictions complicates decision making.

We propose to exploit Bayesian statistics to provide well-developed principles and frameworks to recapitulate the tumor board decisional process in terms of probability. The tumor board discussions can be formalized as an optimization process, acting on a suitably defined fitness function for the patient. Here, we propose a flexible decisional framework inclusive of several clinical solutions from both the literature and the clinical tumor board expertise such that by having a fully comprehensive view on the possibilities of outcomes of cancer therapy, ie, a panoptic view on the problem, it can attempt to rank available solutions by the likelihood of success and therefore suggest a best one within the uncertainties.

In this approach, each patient is a set of clinical data points in multidimensional parameter space, including demographics, clinical diagnosis, laboratory values, histologic features, comorbid conditions, current medications, and so on ([Fig 1](#)) upon which the best (combination of) therapies need to be identified. The model examines available treatments, including surgeries, radiotherapy, chemotherapy, immunotherapy, or psychologic support in the context of the desired outcome (tumor control, palliation, etc). For example, the model formulates the probability that a treatment X_1 or perhaps a series of treatments X_1 combined with X_2, X_3, ..., will produce some outcome Y, eg, the tumor burden, relapse-free survival, tumor control for 12+ months, and similar. Mathematically and clinically, both the existence and uniqueness of a successful solution are not always available, and the tumor board needs to identify suboptimal solutions. In [Figure 1](#), the optimal solution shifts the life expectation line intercept with the time axis to the right as much as possible with additional output on that prolonged life quality.

METHODS

Bayesian Decision Making

The attempt to model decisional processes starting from logic deductions finds its natural setting in the Bayesian framework.[9] We refer to S as a clinical hypothesis of interest (eg, S = radiotherapy can control tumor burden, or the S = drug X will increase time to progression compared with drug Y) and I as the proposition representing prior or previously acquired information (eg, I = the tumor is an early-stage breast cancer without lymph node or distant metastases). The plausibility of the sentence S given (conditional on) the truth of the information I is called prior probability and labeled as Pr(S|I). What is the patient-specific likelihood that the tumor burden is controlled, eg, by radiation therapy? The patient-specific probability is obtained once patient-specific data D are acquired (eg, D = the patient tumor is 3 cm in diameter), and we will label this probability as Pr(D|S, I). Then, the posterior probability of interest, ie, the probability that the tumor burden can be controlled by radiotherapy provided that the tumor is early-stage and positive for a molecular biomarker, is given by the Bayes theorem[9]:

$$\Pr(S|D, I) = \frac{\Pr(S|I)\Pr(D|S, I)}{\Pr(D|I)}, \qquad (1)$$

where Pr(D|I) is the normalization constant.

To identify the treatment with the highest likelihood of success requires the ability to grade different treatment models, frequently dealing with non-Gaussian or skewed error likelihood, which can interpret the same data rigorously in a patient-specific way. Inherent to the Bayesian framework is a natural way to rank diverse model solutions. Here, we discuss Bayesian decision making on this model of models (MoM) for oncologic decision theory.

Bayes theorem, Equation 1, encodes previous knowledge that can influence an outcome. In the Bayesian interpretation, the probability is a (real number) measure of a proposition or hypothesis plausibility, given the truth of the patient-specific information acquired. Most textbooks on Bayesian statistics introduce a comparison between
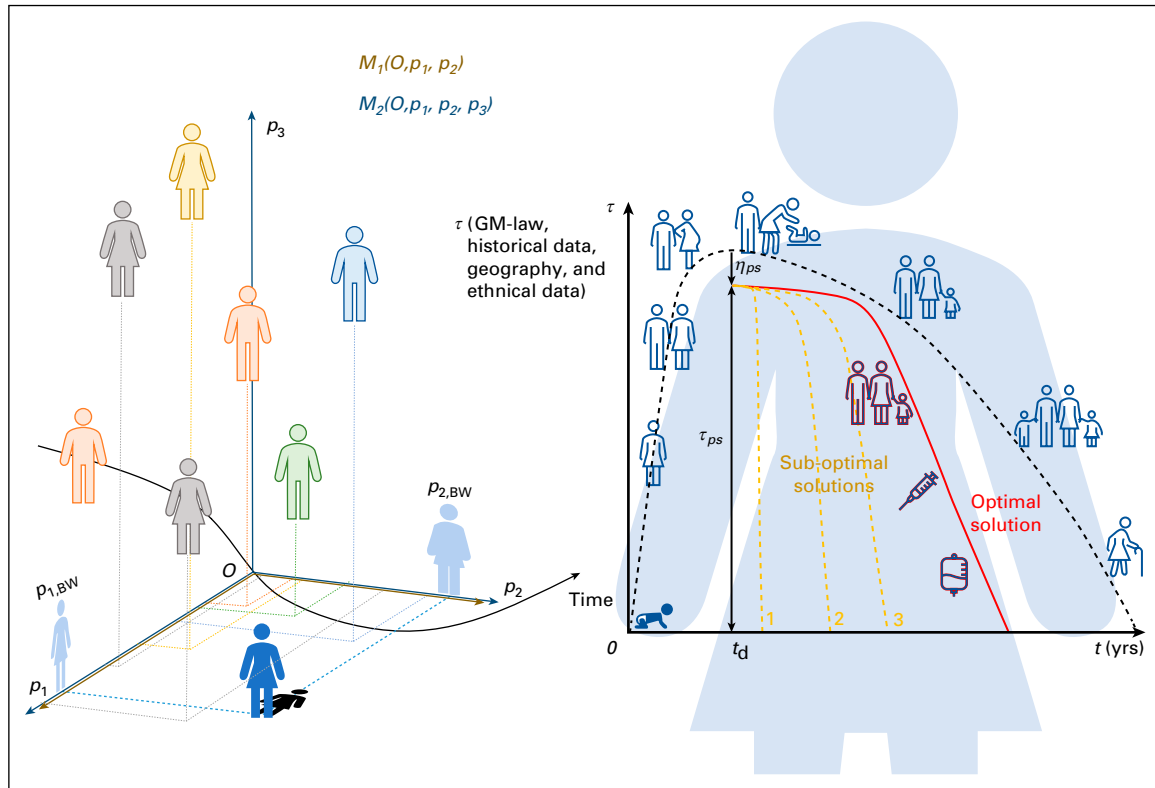
FIG 1. Every patient of a trial is located in a point in the space of parameters of the model considered. For example, the patient under consideration, a BW, has coordinates $\{p_1 = p_{1,BW}, p_2 = p_{2,BW}\}$ in model $M_1(O, p_1, p_2)$, and coordinates $\{p_1 = p_{1,BW}, p_2 = p_{2,BW}, p_{3,BW} = 0\}$ in model $M_2(O, p_1, p_2, p_3)$. The common origin O of the defining set of parameters travels on the timeline: we are considering dynamical systems of equations; hence, the only parameter common to all the models is the time t (here represented ideally with a black curve with a direction passing through O). In the figure on the right, for each patient, the BW-specific life expectancy function is considered, $\tau = \tau(t)$. The Gompertz-Makeham law of mortality (GM-law, black dashed line) is sketched. At t_d, the BW receives the diagnosis of cancer. We assume a negative slope for τ at t_d, and we assume the patient-specific life expectation $\tau_{ps}(t_d)$ to be penalized under the GM-law by a penalizing factor η_{ps} because of geographic, ethnic, or social factors, eg, the BW patient is a former smoker. The red curve is the optimal trajectory of life expectancy as a function of time for a patient predicted by the optimal patient-specific treatment identified by the tumor board. The suboptimal trajectories, ie, the temporal evolution of a decisional curve, eg, yellow-dashed curves labeled 1, 2, or 3, are located below the optimal path with curve 3 to be preferred over 2 and 2 to be preferred over 1 because its intercept with the y-axis (ie, the death of the individual) is farther on the right (ie, the life is longer). The evolutionary tumor board is the taskforce act to choose the optimal red curve. BW, blue woman.

models. We refer the interested reader to the many excellent extended reviews on this topic.[10-12]

We assume that a set of models is available to work in synergy to achieve the optimization problem introduced above. Let different treatments represent the set of models, eg, a combination of different chemotherapeutics, surgical procedures, radiation therapies, or immunotherapies. For simplicity, we assume that they are all available in the hospital, but see considerations on Discussion section. We identify a patient as a point in a multidimensional space, $\mathbf{p} = \{p_1, \ldots, p_N\}$, where each value $p_1, p_2, \ldots, p_N$ is some clinically available measurement (eg, age, sex, tumor burden size, prostate-specific antigen [PSA] level, WBC count, etc) at the time of the diagnosis $t = t_d$ (Fig 1). The decision support framework must then establish each variable's role, according to its clinical significance for the

cancer treatment, based on historical literature and other available clinical trial outcomes. For example, in prostate cancer treatment, metastatic sites and initial Gleason scores are relevant, but the sex is fixed, and specific blood cell counts are probably not prognostic unless abnormal. Once the treatment response model is identified, patient-specific disease trajectories can be simulated to optimize and adapt therapy following the model forecasting.[13] Such a framework would then have to dynamically analyze and switch between different solutions (ie, treatment approaches or protocols) when new data (such as clinical response measurements) or treatments become available.

RESULTS

Bayesian Oncology MoM

We require all the models to be able to be compared in some output, ie, in some clinically relevant metric such as a

tumor marker (eg, PSA), or the tumor volume V_i, or survival τ_i. We then score different therapies on that scale, eg, the effects of radiation therapy or chemical or immunological treatment with or without surgery on overall survival. Thus, highly different therapeutic strategies can be fit into the same patient-specific set of data.

An essential ability of a probabilistic descriptive or predictive framework is its ability to deal with continuous (eg, PSA values) and discrete variables (eg, disease-free or disease progressed). Furthermore, since virtually all clinical data are collected at discrete intervals (eg, computerized tomography scan every 3 months), the model accommodates discontinuous neoplasia volume reduction or, as in surgical resection, with simple step-functions. Finally, if we consider each model with its prior distribution over a joint likelihood, the model with the highest probability (global-likelihood or evidence) naturally leads to model selection.

For each model considered, M_i, we start by encoding the prior state of knowledge, I, into a prior probability distribution $Pr(\mathbf{p}_i|I)$, with $\mathbf{p}_i = \{p_1, ..., p_n\}_i$ being the set of parameters for the model M_i. The first step is to establish a prior distribution of credibility for the ith-model parameter values $\mathbf{p}_i$. By training, validating, and testing over many clinical cases, we build a library of examples that shape the prior distribution and increase the MoM prediction efficiency. In this way, when new drugs or techniques or data become available, they can be added to the library of models or model parameters, reshaping priors' predictive power (after eventual retraining of MoM). For illustration purposes, we can assume little or null prior knowledge of the success rate that a specific model (ie, treatment; model M_i) has on a particular type of cancer. Thus, a model descriptive of a brand-new drug or illustrative of a new theoretical framework to explain the disease and weak knowledge of the parameter value has a broad probability that would span a wide range of the parameter space (Fig 2B).

The Bayesian analysis provides a precise redistribution of the probability over the model parameter range once data, say the jth-data set considered D_j, become available through the likelihood terms $Pr(D_j|\mathbf{p}_i, I)$. In the assumption of identical and independent distributed errors for D_j, we can advocate the central limit theorem[10] or the maximum entropy principle[9] to combine testable information I with Shannon's entropy (or Shannon-Jaynes or Kullback entropies) and measure the uncertainty in a unique posterior distribution function through the use of the likelihood $L_{ij}(I) = Pr(D_j|\mathbf{p}_i, I)$. Under quite a broad general hypothesis, these principles assert that unless some information justifies the use of other sampling distributions, Gaussian likelihood for the error distribution makes the fewest assumptions possible about unavailable information on the collected data. Hence, this approach yields the most conservative estimate because no model is assumed a

priori to be better than another. Instead, all models M_i are considered correct, and each single data value d is related to a model value m through an error e, which represents the unknown error counterpart in the measurement of the data d. Here, a Gaussian distribution describes the source of errors (a noise with finite variance) for the error e. A new patient at the beginning of treatment, ie, with a few data constraining his or her treatment or model, can be encoded with much fewer specificities, ie, an extremely broad likelihood (Fig 2A).

Once the posterior is obtained from the previous two steps for each of the models M_i, $i = 1, ..., n_N$, with $N = N(t)$ the number of models considered (not necessarily constant), the patient-specific-fitness function

$$Pr(\mathbf{p}_i|D, I) = \frac{Pr(\mathbf{p}_i|I)Pr(D|\mathbf{p}_i, I)}{Pr(D|I)} \forall i, \qquad (2)$$

needs to be maximized. The topology of a nonlinear model posterior can be very intricate with many hills and valleys. Fortunately, the past 20 years have seen considerable advancement in algorithms to perform Bayesian calculations, although there is no general solution available to the global optimization problems.[10,11,14,15] Roughly speaking, the most common search approach is based on asymptotic normal (ie, truncated) approximations for a small number of parameters as the Bayesian Information Criterion approximation to the log-marginal $Pr(D|\mathbf{p}_i, I)$, or on the Laplace approximation to the posterior $Pr(\mathbf{p}_i|D, I)$ around the mode, together with more numerical approaches based on random search techniques such as Monte Carlo, Simulated annealing, genetic algorithms for a more substantial number of parameters, or a combination of the above. In most cases, each model is built upon a large number of parameters with only a subset being of interest for clinical decision making or, likely more prominently, several parameters that have to be included cannot be validated by data. Still, these parameters must be quantitatively accounted for without knowing their hidden probability distribution function because of their influence on fitting the model to the available data. These so-called nuisance parameters can be integrated out or marginalized.

As introduced above, a beneficial aspect of Bayesian statistics is the ability to efficiently couple discrete and continuous variables. This feature stands at the basis of the model comparison. When two different models explain the same data set equally well (eg, by producing comparable χ^2 values or comparable log-evidence in a fitting procedure[16,17]), a rigorous and reproducible approach needs to select one model, or treatment, over the other. The possibility of labeling the models lets us consider the model index itself as an independent parameter. The selection process, hence, results from an inference problem on the (discrete) model number. For example, the celebrated odd ratio of the probability of M_1 over M_2 simplifies as

FIG 2. Schematic for two significantly different cases where the generic model achieves similar conditional probability distribution (posterior, lower row) M_i encoding the information I, with very different prior and likelihood probability distribution function (upper row). The multiple parameter space is represented with a single axis $\mathbf{p} = \{p_1, p_2, p_3, \ldots, p_n\}$. (A) The likelihood distribution (light blue) is flat, and the prior probability distribution (gray) is highly peaked. The resulting posterior distribution function is sharply peaked because of the prior distribution (lower left panel). (B) The likelihood distribution (light blue) is peaked, and a flat prior distribution (gray) is assumed. The low likelihood reflects missing patient-specification of the therapy as in a not yet calibrated model. The resulting posterior probability might still result in a very peaked but not patient-specific shape.
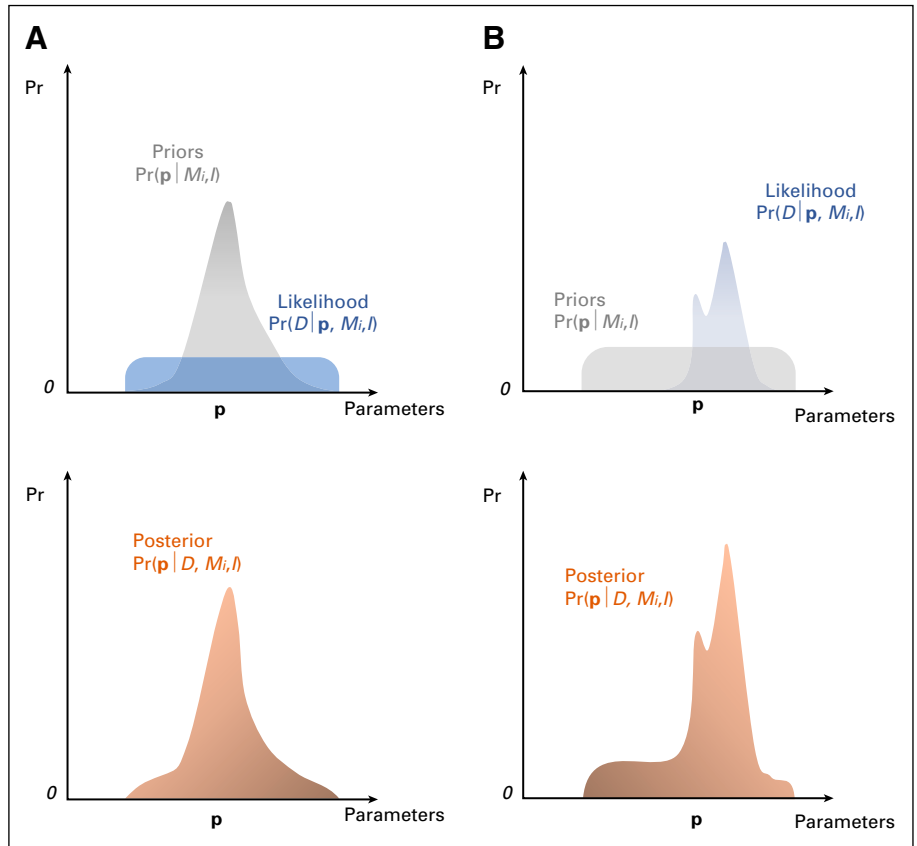
$$O_{12} = \frac{\Pr(\mathbf{p}_1|I)\Pr(D|\mathbf{p}_1, I)}{\Pr(\mathbf{p}_2|I)\Pr(D|\mathbf{p}_2, I)} = \frac{\Pr(\mathbf{p}_1|I)}{\Pr(\mathbf{p}_2|I)}B_{12}, \qquad (3)$$

with B_{12} the Bayes factor of model 1 over model 2. Note that the model-index probability is sensitive to the entire parameter space, not only to the single model's prior distribution at its best-fitting parameters position. More peaked prior distribution on well-fitting data will result in a higher probability density function (PDF) and, vice versa, when the prior distribution of a model flattens the PDF over a more extensive parameter range that does not fit the data well, the posterior PDF will tend to be small. This characteristic is advantageous in the MoM approach, where models of different complexities may be simultaneously considered. The more complex models will always be able to fit data better than restricted models. MoM balances data fit and model complexity (ie, degrees of freedom provided by the number of parameters) and can select simpler models with fewer degrees of freedom over more complex models. By diluting the prior probability over larger areas, the more complex model assigns a lower chance for any parameter value that fits the data, resulting in a downweighed PDF. However, a more complex model will be selected if parameter values and model dynamics that are not accessible by a more restricted model provide a sufficiently better fit to the data.

If new data or a full new data set D_{j+1} is acquired (eg, a new value of a biomarker for the patient or clinical trial results published elsewhere), MoM does not need to be (re) trained, including the original data set. (We do not exclude that there are obvious situations where retraining is unavoidable, eg, because some clinical options are not available because of geographical constraints, economical constraints, psychologic constraints, availability of a new vaccine etc. We will comment briefly on this point and its relation with Bayesian priors in the Discussion section.) Bayes theorem is applied to compute each model's new posterior to redistribute the latest knowledge state. The most recent prior, I', is the posterior derived from D_j, I, ie, $I' = D_j$, I. Then, the new posterior is

$$\Pr(\mathbf{p}_i|D_{j+1}, I') \propto \Pr(\mathbf{p}_i|I')\Pr(D_{j+1}|\mathbf{p}_i, I') \qquad (4)$$

This iterative process can shift the weights, and thus optimal selection, from a model to another (Data Supplement and Fig 3). Vice versa, the inclusion of a significant new paradigm, treatment, or approach (eg, the discovery of a vaccine) might, of course, have such an impact to require MoM retraining to include availability factors (eg, distribution factors and geographic factors).

Finally, superseding the global optimization problem mentioned above, it is worth stress that, a priori, we do not always expect MoM to provide a usable or meaningful full
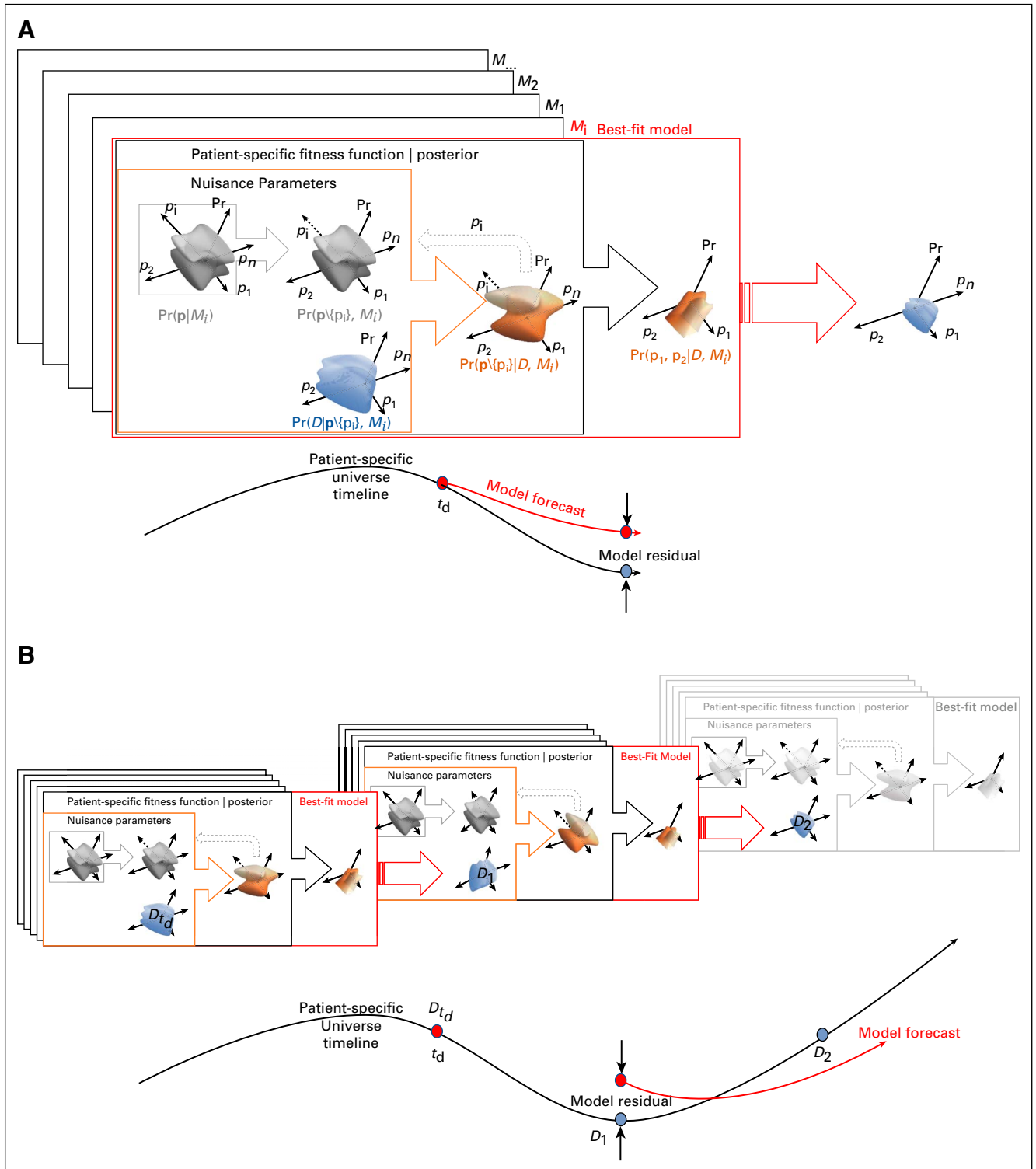
FIG 3. (A) Schematic representation of the MoM approach. One of the models considered, say M_i, represented here as a prior distribution (gray manifold) in the system of reference with the axis p_1, p_2, p_3, ..., p_n (Fig 1), outputs a probability distribution function of interest, Pr. A generic parameter p_i results in marginal benefit from the fit hence marginalized as a nuisance parameter from the PDF (dashed corresponding axis). The likelihood accounting for the set of data, $\text{Pr}(D|\mathbf{p} - p_i, M_i)$, is represented with a blue manifold. A combination of prior and likelihood is used to obtain the model posterior distribution function (ie, a patient-specific fitness function, orange manifold). A fit to the data might evidence the relevance of the parameters previously discarded, eg, p_i, thus suggesting reconsidering the model (gray dashed arrow). Fit-dependence and the marginalization of nuisance parameters lead to the PDF that we consider in the model comparison. The optimal posterior distribution function from the best fitting model (named generically M_i in the figure) is used to forecast the evolution of a patient-specific treatment. The bottom part of the panel shows the patient-specific timeline. (continued on following page)

answer to the therapy selection problem. The approach proposed is a data-driven approach, both in the use of the priors, built on literature or trial results, and in patient-specific data. Data are provided with errors that inevitably propagate on the model selection process. The global-likelihood or evidence of models M_1 and M_2 are determined at the best of uncertainties that affect the Bayes factor B_{12} determination. Therefore, MoM might determine the best clinical path to follow but not outside any reasonable doubt, ie, not outside the errors because of the available data quality. For this reason, the instrument we propose to introduce in this oncologic contest (the tumor board) should be considered as a suggestion in the hands, and under the control, of the medical oncologist in charge.

MoM Decision Making

Before presenting an example of decision-theory applied to tumor forecasting, we first address the mathematical mechanism that leads to the decision. In the Data Supplement, we detail the principle-of-operation of this mechanism with the help of an example, ie, a hypothetical situation where the Bayesian decisional theory introduced above is crucial in helping a tumor board deciding which clinical path to follow. Despite the case being elementary and meant to match a situation with a well-known decisional output, it is pedagogical in its attempt to show how a tumor board opinion is mathematically coded and treated in the present formalism. The example is inspired by the Laplace approximation mentioned above. Still, it does not require the use of Gaussian approximation or the knowledge of information theory. Instead, it focuses on transmitting how the decisional process happens, ie, it proposes an exemplification of the MoM model selection.

Model Forecasting and Decision Theory

What happens after data analysis has suggested the best model available outside any reasonable doubt? How can inference be used in making a final decision? One of the statistical analysis' aim is undoubtedly to aid a decision process. We assumed implicitly above that our goal is to estimate the probability of medical treatment to prevent or delay disease progression and death.

Leading health informatics and medical informatics journals cover the Bayesian approach to model forecasting (eg, *Journal of the American Medical Informatics Association*, *Journal of Medical Internet Research*, and *Medical*

Decision Making) together with numerous bestseller books.[14,15,18-20] Nevertheless, autoregressive moving average,[21] vector autoregressive models,[22] together with the broad class of regressive neural network,[23] eg, the long-short time memory,[24] are mainly focused on predicting the future data of a time series from the sequence of data collected rather than from a model comprehensive of the biologic mechanisms involved as we assume for MoM.

Here, we aim to advance a framework able to engage the cancer description over its biologic multiscale, robust in forecasting the evolution of the disease, and readily available to give us a biologic interpretation of the results. In the following section and the Data Supplement, we will limit ourselves to review the basics of such methods by extending the context of the example sketched above with simple Bayesian considerations, but focusing on the forecasting problem, whence medical decisions depend on data gathered after the first decision at t_d has already been taken.

Tumor Board Evolution: MoM Decision Process

Classically, from the posterior probability of the best model, the expectation value $E[*]$ of a suitable defined function can be determined as the life expectancy τ from the evidenced best model at the medical screening time, conditional to the decision we take d, $E[\tau|d]$. With additional data and tumor dynamics becoming available from a patient on the clinical response to therapy, MoM exploits its flexibility to relocate the probability over the entire parameter space, including the model index, to evaluate treatment adaptations. If a patient does not respond to a drug or drug change, the new information provides new prior data to recalculate posteriors for the remaining treatment options. The MoM approach automatically proceeds to this modal PDF's reallocation, suggesting (when existing) the best combination of treatments available within the errors (Data Supplement and Fig 3).

In the section above and the dedicated clinical example, the Data Supplement, we dug deeper into where and *how* the decisional process happens. In the Data Supplement, we focus on a more classical result of the Bayesian framework: the frame's *forecasting* character and its ability to *include new information*. Again, in the Data Supplement, we will develop the concepts elaborating over an example of clinical interest. Many self-similar exercises are available in

---

FIG 3. (Continued). The patient-specific set of data represents the status of the patient in their timeline. Once new data are acquired (blue dot) establishing the patient's location in the parameter space, the difference between the above-predicted model forecast and the newly acquired data represents the residual difference between the model prediction and the current state the patient. (B) The residual is used to reiterate the model procedure to obtain the best model representing the patient. In the upper part of the panel, the scheme of (A) is reiterated several times with new incoming data. Starting from the patient-specific data set available at the time of the diagnosis D_{t_d}, the model selection is performed (see the bottom part of the panel). Once a new data set is obtained, D_1, a new model selection is performed, iterating the MoM procedure. The sequence proceeds similarly once new data points or a new set of data, graphically represented by D_2, D_3, …, are collected. Note how the best fit model never generally matches all the data sets available entirely: a residual is never null. Nevertheless, we assume to have a library of models that contains at least a model that the more data are acquired, the better approximate the patient-specific data, the smaller the residuals are as time passes. MoM, model of models; PDF, probability density function.

FIG 4. Schematic logic flow of a clinical treatment augmented by the MoM framework. MoM is offered as a cloud computing service: it does not store patient generalities; instead, it outputs the optimal solution available to an inputted patient data point under discussion in a tumor board. The same service can be made accessible to a qualified medical doctor as a consultation-dynamical-library of treatment outcomes. Afterward, they might eventually defer the patient to a clinical structure where the suggested treatment is available. MoM, model of models.
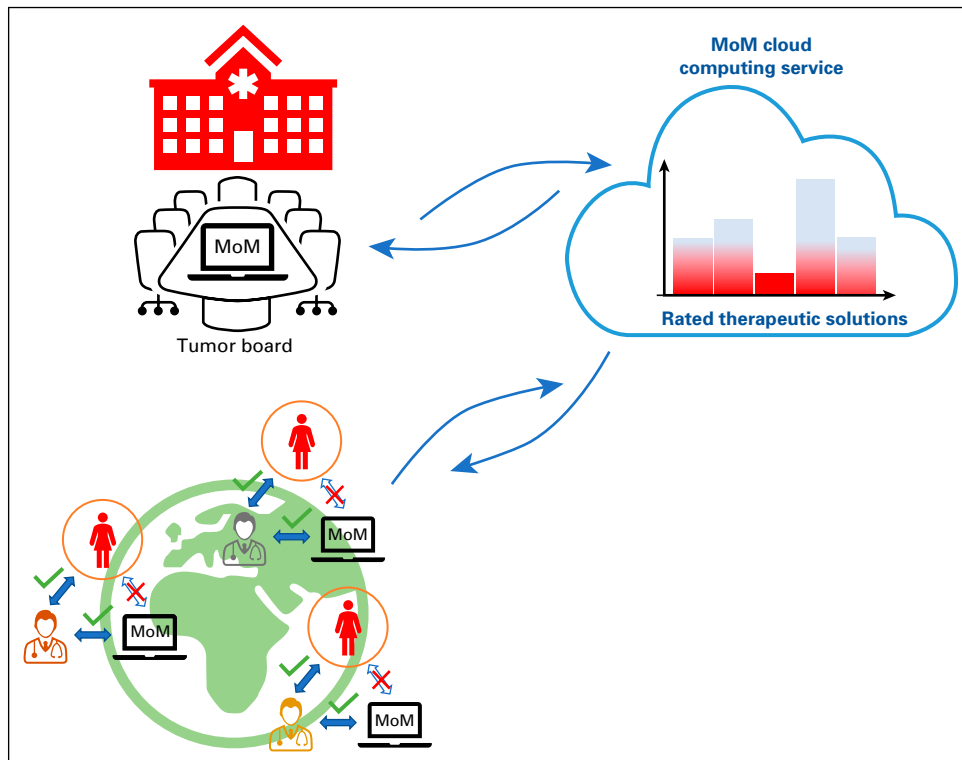
the literature or on the web, especially concerning medical research.[25]

DISCUSSION

Here, we advanced the idea of a framework to support the decisional theory modeled on tumor boards' function to identify patient-specific clinical pathways. Cancers are highly heterogeneous diseases with significant interperson and intraperson heterogeneity and high variability in clinical categorizations, definitions, and delivery of treatments, and outcome determination. For any clinical decision making, it is essential to rely on medical doctor (MD) experiences and the most accurate data and account for uncertainty and probabilities—for which Bayesian approaches are strongly suited. The proposed MoM is a fully comprehensive ecosystem able to account for patient-specific data, data uncertainty, different data-driven biologic models, various treatment approaches, and, most importantly, it is a way to include human expertise represented by the tumor board. MoM's strength is its panoptic view on the different aspects contributing to optimal therapies with reproducible uncertainties and confidence measurements. Exploiting the coexistence of opinions, equations, and techniques resulting from diverse expertise (chemistry, physics, and biology), MoM aims to offer a reproducible framework to compare and upgrade knowledge on cancer therapy.

Despite being transparent to the clinician's perspective, MoM is intended to be a freely accessible library of posterior probabilities of already-actioned (both successfully and unsuccessfully) clinical path on specific tumors. Any tumor

board, or clinician-oncologist, might want to access it, eg, through a webpage. Once the clinical parameter of a patient-specific case of interest is inserted, MoM will rate the relative merits of the therapeutics paths. One of the strengths of the MoM approach is that the largest the number of points is in the database (or inputted by oncologists spread worldwide), the more useful and efficient this instrument turns out to be, not only in a tumor board setting (ie, in an in-person meeting) but also, outscoping, as rapid informative clinician instrument. Nevertheless, this very same approach might also hide a weakness. If MoM builds up its prior on a larger and larger database, its response might be closer and closer to the optimal clinical pathways to follow. Influential priors might be extremely sensitive to the locoregional constraints: many therapeutic solutions immediately available in large cities might vice-versa require patients living in smaller towns to face long trips that might not be possible because of their clinical conditions. We need to evaluate software design that might include retraining options if we see this become necessary to make MoM useful. A workflow of the MoM concept can be evicted from Figure 4. Note how MoM is a logic-probabilistic tool for *informative* purposes only. It is *not* intended by any means to indicate the treatment pathway: it is only intended to rate the therapeutic options whose selection is left first and only to the patients through their MDs. This concept is represented in the figure as a connection between MoM patients and MDs. Furthermore, MoM's development as a computational tool (eg, in a cloud computing service) instead of a patient database is aimed

to stress that no patient generalities need to be stored and only their correspondent data-point needs to be inputted.

In conclusion, the MoM framework is conceptually designed to identify optimal treatments based on patient-specific data-points in the context of information from published studies and the institutional (or multi-institutional) databases. Here, we present MoM's concept as a decision support tool and provide the initial clinical translation step. To be used in tumor boards and by individual oncologists, it must be trained, tested, and prospectively validated for specific cancers and purposely developed mathematical and statistical models of cancer progression and treatment response. Once implemented, the Bayesian MoM approach will have to be compared with other decision support tools to evaluate its clinical applicability and to ethically integrate the framework into clinical practice so that no harm is done.

AFFILIATIONS

[1]Department of Integrated Mathematical Oncology, H. Lee Moffitt Cancer & Research Institute, Tampa, FL

[2]Department of Radiology, H. Lee Moffitt Cancer & Research Institute, Tampa, FL

[3]Department of Radiation Oncology, H. Lee Moffitt Cancer & Research Institute, Tampa, FL

CORRESPONDING AUTHOR

Stefano Pasetto, PhD, Department of Integrated Mathematical Oncology, H. Lee Moffitt Cancer & Research Institute, 12902 Magnolia Dr, Tampa, FL 33612; e-mail: stefano.pasetto@moffitt.org.

AUTHOR CONTRIBUTIONS

Conception and design: Stefano Pasetto, Heiko Enderling
Financial support: Heiko Enderling
Data analysis and interpretation: Robert A. Gatenby
Manuscript writing: All authors
Final approval of manuscript: All authors
Accountable for all aspects of the work: All authors

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by the authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/cci/author-center.

Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians (Open Payments).

Heiko Enderling

Patents, Royalties, Other Intellectual Property: US Patent 62/900,003: Modeling variability in radiosensitivity and tumor immune contexture to personalize radiotherapy (provisional), US Patent 62/944,804: Methods for prostate cancer intermittent adaptive therapy (provisional), US Patent 63/010,327: Forecasting individual patient response to radiotherapy with a dynamic carrying capacity model (provisional), US Patent 62/040,579: Predicting glioma treatment response (provisional)

No other potential conflicts of interest were reported.

REFERENCES

1. Lawless JF, Singhal K: ISMOD: An all-subsets regression program for generalized linear models. I. Statistical and computational background. Comput Methods Programs Biomed 24:117-124, 1987

2. Lawless JF, Singhal K: ISMOD: An all-subsets regression program for generalized linear models. II. Program guide and examples. Comput Methods Programs Biomed 24:125-134, 1987

3. Somashekhar SP, Sepúlveda MJ, Shortliffe EH, et al: A prospective blinded study of 1000 cases analyzing the role of artificial intelligence: Watson for oncology and change in decision making of a Multidisciplinary Tumor Board (MDT) from a tertiary care cancer center. J Clin Oncol 37, 2019 (suppl; abstr 6533)

4. IBM Watson Health. Wikipedia, 2020. https://en.wikipedia.org/w/index.php?title=IBM_Watson_Health&oldid=951313570

5. Candido dos Reis FJ, Wishart GC, Dicks EM, et al: An updated PREDICT breast cancer prognostication and treatment benefit prediction model with independent validation. Breast Cancer Res 19:58, 2017

6. Olivotto IA: Population-based validation of the prognostic model ADJUVANT! For early breast cancer. J Clin Oncol 23:2716-2725, 2005

7. Wishart GC, Azzato EM, Greenberg DC, et al: PREDICT: A new UK prognostic model that predicts survival following surgery for invasive breast cancer. Breast Cancer Res 12:R1, 2010

8. de Glas NA, van de Water W, Engelhardt EG, et al: Validity of Adjuvant! Online program in older patients with breast cancer: A population-based study. Lancet Oncol 15:722-729, 2014

9. Jaynes E: Probability Theory: The Logic of Science. Volume 27. Cambridge University Press, 2003

10. Theodoridis S: Machine Learning: A Bayesian and Optimization Perspective (ed 1). Academic Press, 2015

11. Katz D, Azen SP, Schumitzky A: Bayesian approach to the analysis of nonlinear models: Implementation and evaluation. Biometrics 37:137-142, 1981

12. Murphy KP: Machine Learning: A Probabilistic Perspective. The MIT Press, 2012

516

13. Enderling H, Alfonso JCL, Moros E, et al: Integrating mathematical modeling into the roadmap for personalized adaptive radiation therapy. Trends Cancer 5:467-474, 2019

14. Pole A, West M, Harrison J: Applied Bayesian Forecasting and Time Series Analysis. CRC Press, 1994. https://www.crcpress.com/Applied-Bayesian-Forecasting-and-Time-Series-Analysis/Pole-West-Harrison/p/book/9780412044014

15. West M, Harrison J: Bayesian Forecasting and Dynamic Models (ed 2). Springer-Verlag, 1997

16. Hosseini R, Zarepour M: A Bayesian Nonparametric Chi-Squared Goodness-of-Fit Test. 2016. http://arxiv.org/abs/1602.00197

17. Johnson VE: A Bayesian χ^2 test for goodness-of-fit. Ann Stat 32:2361-2384, 2004

18. Parmigiani G: Modeling in Medical Decision Making: A Bayesian Approach. Wiley, 2002. https://www.wiley.com/en-us/Modeling+in+Medical+Decision+Making%3A+A+Bayesian+Approach-p-9780471986089

19. Owens DK, Sox HC: Medical decision-making: Probabilistic medical reasoning, in ShortliffeEH, PerreaultLE (eds): Medical Informatics: Computer Applications in Health Care and Biomedicine. Health Informatics. Springer, 2001, pp 76-131

20. Degroot MH: Optimal Statistical Decisions. Wiley, 2004. https://www.wiley.com/en-us/Optimal+Statistical+Decisions-p-9780471680291

21. Whittle P: Hypothesis Testing in Time Series Analysis. Almqvist & Wiksell, 1951

22. Sims CA: Macroeconomics and reality. Econometrica 48:1-48, 1980

23. Rumelhart DE, Hinton GE, Williams RJ: Learning representations by back-propagating errors. Nature 323:533-536, 1986

24. Hochreiter S, Schmidhuber J: Long short-term memory. Neural Comput 9:1735-1780, 1997

25. Bayes' Theorem. https://en.wikipedia.org/wiki/Bayes%27_theorem

■■■