



# Intermittent Hormone Therapy Models Analysis and Bayesian Model Comparison for Prostate Cancer

S. Pasetto<sup>1</sup>  · H. Enderling<sup>1,2,3</sup> · R. A. Gatenby<sup>1,4</sup> · R. Brady-Nicholls<sup>1</sup>

Received: 1 March 2021 / Accepted: 27 September 2021 / Published online: 19 November 2021  
© The Author(s) 2021

## Abstract

The prostate is an exocrine gland of the male reproductive system dependent on androgens (testosterone and dihydrotestosterone) for development and maintenance. First-line therapy for prostate cancer includes androgen deprivation therapy (ADT), depriving both the normal and malignant prostate cells of androgens required for proliferation and survival. A significant problem with continuous ADT at the maximum tolerable dose is the insurgence of cancer cell resistance. In recent years, intermittent ADT has been proposed as an alternative to continuous ADT, limiting toxicities and delaying time-to-progression. Several mathematical models with different biological resistance mechanisms have been considered to simulate intermittent ADT response dynamics. We present a comparison between 13 of these intermittent dynamical models and assess their ability to describe prostate-specific antigen (PSA) dynamics. The models are calibrated to longitudinal PSA data from the Canadian Prospective Phase II Trial of intermittent ADT for locally advanced prostate cancer. We perform Bayesian inference and model analysis over the models' space of parameters on- and off-treatment to determine each model's strength and weakness in describing the patient-specific PSA dynamics. Additionally, we carry out a classical Bayesian model comparison on the models' evidence to determine the models with the highest

---

✉ S. Pasetto  
stefano.pasetto@moffitt.org

✉ R. Brady-Nicholls  
renee.brady@moffitt.org

<sup>1</sup> Department of Integrated Mathematical Oncology, H. Lee Moffitt Cancer and Research Institute, 12902 Magnolia Drive, Tampa, FL 33612, USA

<sup>2</sup> Department of Radiation Oncology, H. Lee Moffitt Cancer and Research Institute, 12902 Magnolia Drive, Tampa, FL 33612, USA

<sup>3</sup> Department of Genitourinary Oncology, H. Lee Moffitt Cancer Center and Research Institute, 12902 Magnolia Drive, Tampa, FL 33612, USA

<sup>4</sup> Department of Radiology, H. Lee Moffitt Cancer and Research Institute, 12902 Magnolia Drive, Tampa, FL 33612, USA

likelihood to simulate the clinically observed dynamics. Our analysis identifies several models with critical abilities to disentangle between relapsing and not relapsing patients, together with parameter intervals where the critical points' basin of attraction might be exploited for clinical purposes. Finally, within the Bayesian model comparison framework, we identify the most compelling models in the description of the clinical data.

**Keywords** Prostate cancer · Intermittent hormone therapy models

## 1 Introduction

The prostate is an exocrine gland of most mammals' male reproductive system. The normal prostate is dependent on androgens, specifically testosterone and 5 $\alpha$ -dihydrotestosterone (DHT), for development and maintenance (Feldman and Feldman 2001). Prostate carcinoma (PCa) results from the abnormal growth of tissue from the prostate's epithelial cells, which might induce metastasis in bones and lymph nodes. PCa is the second most common cancer in the USA and the second leading cause of cancer-related death after lung cancer (Siegel et al. 2021). The average male is 70 years of age at the time of diagnosis, with a strong of the distribution asymmetry biased toward older ages. PCa risk is often influenced by genetics. Men with a first-degree relative with PCa are twice as likely to develop it themselves; men with high blood pressure are also at higher risk of PCa. Treatment options typically include surgery, radiotherapy, high-intensity focused ultrasound, chemotherapy, and hormonal therapy.

Screening for PCa is commonly performed through rectal examination or the noninvasive blood biomarker prostate-specific antigen (PSA), although its efficiency remains controversial (Lin et al. 2008). Today, more robust marker indicators, such as the overexpression of prostate cancer gene 3 (PCA3) obtained from the messenger-RNA (mRNA) in the urines, are considered more suited to monitoring the cancer evolution (Bussemakers et al. 1999, p. 3; Laxman et al. 2008; Neves et al. 2008; Hesses and Schalken 2009, p. 3; Borros 2009; Qin et al. 2020). PSA is a measure of a hematic enzyme produced by the prostate. PSA levels between 4.0 and 6.5  $\mu\text{h L}^{-1}$  are generally considered normal (with a strong dependence on age). PSA is naturally present in the serum, and usually, only a small amount of PSA of the prostate leaks into the blood. Hence, high levels are an indication of prostatic hyperplasia or cancer. Since prostate cells and their malignant counterparts require androgen stimulation to grow, prostate cancer can be treated by androgen deprivation therapy (ADT), a type of hormone therapy. This therapy reduces androgen-dependent (AD) cancer cells by preventing their growth and inducing cellular apoptosis.

Unfortunately, treating with ADT often results in a relapse in the form of hormone-refractory PCa due to the selection for the androgen-independent (AI) cells. Intermittent androgen deprivation (IAD) therapy, whereby treatment is cycled on and off, is often used as an alternative to ADT to delay treatment resistance. In IAD, androgen deprivation therapy is administered until a patient experiences a remission and then is withheld until the disease progresses up to a certain level. Clinical studies have

shown that patients are responsive to multiple hormone therapy cycles, eventually delaying the androgen independence insurgence (Klotz et al. 1986; Larry Goldenberg et al. 1995; Bruchofsky et al. 2006).

We consider models of intermittent therapy due to clinical interest and solve the inference problem using longitudinal PSA data from the Canadian Prospective Phase II Trial of IAD for locally advanced prostate cancer. This work aims to present the first systematic comparative study of IAD models, emphasizing their ability to disentangle relapsing and not relapsing patients and compare the models in the Bayesian framework. The goal is to detect the single model (or the group of models) that best represent the information in the considered dataset and, therefore, if possible, the most promising biological frame representing them. A general and historical review of the available prostate cancer models can be found elsewhere (Phan et al. 2020).

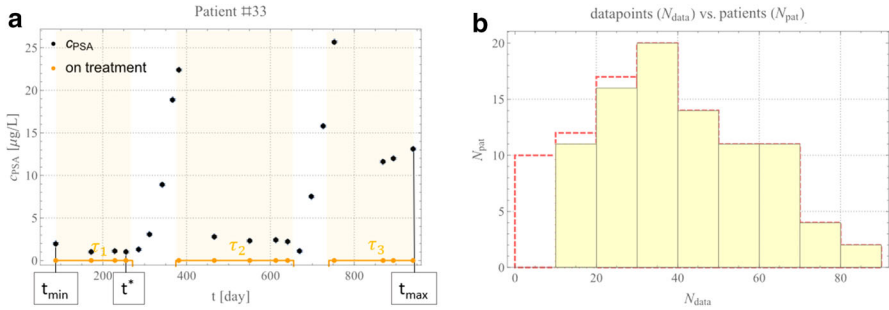
In Sect. 2, we present the data included in our analysis. In Sect. 3, we introduce the statistical framework used to analyze the data. Section 4 presents an analysis of the models and their performance over the dataset utilizing the framework considered. Section 5 compares the performance of all the models, and Sect. 6 concludes and discusses the paper's findings and future developments.

## 2 Dataset

### 2.1 Data Cohort

We consider data from the Canadian Prospective Phase II Trial of intermittent ADT for biochemically recurrent prostate cancer (Bruchofsky et al. 2006, 2008). The total patient number is  $N_{\text{pat}} = 101$ . Their median pretreatment serum testosterone is  $13.0 \mu\text{g L}^{-1}$ , ranging between 0.4 and  $23.0 \mu\text{g L}^{-1}$ . Over a maximum of  $n = 5$  intermittent ADT cycles, a median of 35.1–36.0 weeks is spent on-treatment (depending on  $n$ ), and 25.6–53.7 weeks (e.g.,  $n = 5$  and  $n = 1$  respectively) are off-treatment during the 6-year study. An example of a PSA profile for an individual patient is shown in Fig. 1a. This patient responded to treatment during the first two treatment cycles ( $\tau_1$  and  $\tau_2$ ) and progressed in his third cycle of treatment ( $\tau_3$ ). The oscillatory dynamics demonstrate the effect of the intermittent treatment, with a decrease in PSA during treatment and an increase once treatment is turned off. Each data point is assigned with an error of 1 day in time (i.e., the time resolution of the dataset) and a maximal PSA error value  $e_{\text{max}}$  assigned of  $e_{\text{max}} = 0.1 \mu\text{g L}^{-1}$  assumed from the literature (Borros 2009).

We set the minimal PSA detection threshold equal to  $\Delta_1 = 0.1 \mu\text{g L}^{-1}$ , i.e., any patient data below this threshold is set to  $0.1 \mu\text{g L}^{-1}$ . Patients with a minimal per-day fluctuation below  $2.0 \mu\text{g L}^{-1}$ , i.e., a minimal per-day fluctuation of the KLK3 glycoprotein enzyme of PSA of a typical man (Morgentaler and Connors 2015), are excluded because such small fluctuations are considered natural and not pathological. To only consider PSA concentrations above Poisson noise, patients with less than  $\Delta_2 = \sqrt{N_{\text{pat}}}$  (i.e., the sample shot/Poisson noise) data points are also excluded. These exclusions result in our analysis considering data from 89 ( $N_{\text{pat}} = 89$ ) rather than 101



**Fig. 1** Model data. **a** PSA data for patient #33 from  $t_{\min} = 88$  [day] to  $t_{\max} = 941$  [day]. Black dots indicate PSA values (error bars are omitted due to little variability), orange points indicate where PSA was collected, and graphically represented as an orange continuous box function, evidenced only in this example panel by yellow shaded areas. Treatment intervals are labeled  $\tau_1$ ,  $\tau_2$  and  $\tau_3$ .  $t^*$  is the first minimum of PSA in  $\tau_1$ . **b** Distribution of the number of data points per patient. The original data are shown by the red dashed lines, while the selected subset of patients used in this analysis is shown in the yellow shaded region (Color figure online)

patients. The patients’ distribution per number of data points used after the selection process is shown in Fig. 1b compared to the original distribution.

### 2.2 Data Interpretation

The PSA trend shown in Fig. 1a is based on the interplay between cellular populations, i.e., with a compartment modeling approach. An androgen-dependent set of  $N_D \geq 1$  cell populations (each with a concentration  $n_{D,k} = n_{D,k}(t)$ ,  $k = 1, \dots, N_D$  representing the compartment concentration [ $\mu\text{g}L^{-1}$ ],  $t \in \mathbb{R}$  time [day]) is assumed to contribute to the oscillatory behavior of PSA. Additionally, a set of time-dependent androgen-independent cellular populations of  $N_I \geq 0$  cell populations  $n_{I,l} = n_{I,l}(t)$ ,  $l = 0, \dots, N_I$ , also contributes to the PSA profile, such that PSA concentration  $c_{\text{PSA}}$  is given by  $c_{\text{PSA}} = f(n_{D,k}, n_{I,l})$ , where  $f \in C^0$  is a function belonging to the class of continuous solution  $C^0$  (not necessarily smooth) of a suitably designed ODEs system. Any further dependence on space, temperature, and pressure is generally neglected in the IAD models’ compartment approach. Furthermore,  $f$  is often assumed to be a linear combination of the  $N_D$  and  $N_I$  compartments, e.g.,  $c_{\text{PSA}} = \sum_k w_k n_{D,k} + \sum_l w_l n_{I,l}$  for some weights  $w_l$  and  $w_k$ .

By assuming ADT to be highly effective in the first treatment interval  $\tau_1$ , we can set  $n_D(t \in \tau_1) \cong 0$  for some  $t$  as an initial condition (hereafter i.c.). This approach does not necessarily hold for  $\tau_i$  with  $i > 1$ : Generally,  $\forall i$  where  $c_{\text{PSA}} \cong 0$ , we can equally well assume this setting for the i.c. of the  $n_{D,k} = n_{D,k}(t)$  equations. Equivalently we can assume that a non-holonomic (i.e., with inequalities) condition for the fitting procedure holds at the beginning of the patient time series  $n_D(t) \leq n_I(t)$  for some  $t \in \tau_1$ . Furthermore, in most of the models that we accounted for, these considerations are articulated with the addition of a few extra equations that interpret, at a local or global level in the parameter space, the contribution to  $c_{\text{PSA}}(t)$  by the androgen quota, cellular plasticity, staminal cells populations, or other model specificities.

Finally, we know from biological arguments that, under treatment, the models' equations are designed to permit, at least for  $c_{PSA}$ , to tend asymptotically to the value  $c_{PSA} = 0$ . Any model that does not permit the phase state  $c_{PSA} = c_{PSA}(t)$  to reach approximately null values for any  $t$ , i.e.,  $\nexists t \in \mathbb{R}: c_{PSA}(t) \cong 0$ , would fail to reproduce the patients whose first treatment is always successful (see Fig. 1a). We elaborate more on this in Sect. 4. Therefore, it is worth investigating if the models allow for stationary equilibria outside the treatment intervals and then if any of the patient best fit values have fallen close to those equilibria (when they exist). This behavior would imply a stationary or recurrent solution for the dynamics and, therefore, a constrained PSA's evolution if this "basin of attraction" is achievable in a biological time of interest. We stress that this mathematical behavior does not imply that the patient can effectively reach the point of equilibrium on the biological timescale of interest or a plausible point regarding toxicity levels.

### 3 Bayesian Inference

The Bayesian regression approach stems from the concept of probability as a measure of the plausibility of a model *given* the truth of the information in the data presented above. First, we encode the prior state of knowledge about the parameters considered  $\mathbf{p} = \{p_1, p_2, \dots\}$  into a prior distribution function  $\Pr(\mathbf{p}|I)$ , where  $I$  represents any available information. Typically, this can be achieved with a flat, uniform, and not informative prior at the beginning or with a sharper prior when the model is better trained. We return to this point in Sect. 3.1. Secondly, we consider the dataset,  $D$ , through the likelihood  $L(\mathbf{p}, D) = \Pr(D|\mathbf{p}, I)$ . Finally, the inference problem is solved, studying the probability distribution function encoding the knowledge of the prior and the information encoded in the likelihood of the data  $\Pr(\mathbf{p}|D, I) \propto \Pr(\mathbf{p}|I)L(\mathbf{p}, D)$ .

Standard techniques to achieve this result are fully analytical (e.g., for some linear regression), approximated (e.g., asymptotic approximation, Laplacian approximation, Gaussian approximation, etc.), iterative (e.g., Levenberg–Marquardt), or fully numerical (e.g., simulated annealing genetic algorithms). The choice between these techniques depends on the nature of the problem. Here, we start using Laplacian approximation with hyperparameters (Hutter et al. 2011; Murphy 2012; Theodoridis 2015), as a few of the mathematical models that we consider herein are nested, to solve the inference problem (i.e., to search for the optimal set of parameters  $\mathbf{p}$  that best represent the data). In order to confirm the inference results and to perform the Bayesian model comparison numerically, we test the results both against the nested sampling approach to the global likelihood (hereafter *evidence*) (Skilling 2004; Mukherjee et al. 2006; Feroz and Hobson 2008) and with the differential evolution search (Feoktistov 2006; Goode and Annin 2015) with up to aggressive scaling factors ( $\leq 0.9$ ) and cross probabilities ( $\geq 0.1$ ). For the Bayesian model comparison part of our work, see Sect. 5, the nested sampling-based approach will embed the results in a natural framework.

Finally, we note that substantial limitations in the fitting procedure came from the sparse and irregular temporal sampling in the clinical data. This irregularity impacts the parameter space exploration due to the lack of condition on the PSA trend's derivative.

The partial derivative  $\partial_t c_{PSA}(\mathbf{p}; t)$  is not smooth, thus inhibiting using some straightforward optimization techniques based on the PSA curves' gradients or convexity (Theodoridis 2015).

### 3.1 The Priors $\Pr(\mathbf{p}|I)$

While robust approximations or numerical tools have been adopted for the Bayesian framework, special attention is paid to the use of priors. As mentioned, Bayesian inference requires the use of the priors,  $\Pr(\mathbf{p}|I)$ , for parameter estimation. With initially unknown priors, we implement uniform priors over the parameters' full ranges (Fig. 2a). By requiring all model parameters to be positive, we can assume the Heaviside step function  $\theta = \theta(\mathbf{p})$  as (unnormalized) prior, this approach is generally referred to as "improper prior" as it is unbounded above, it cannot be normalized and therefore it does not have a mean, standard deviation, median, or quantiles. We set an upper bound for each parameter to be  $p < p_{\max}$  with a max value  $p_{\max} < +\infty \forall p$  strictly. An alternative functional tested is the non-informative Jeffreys prior,  $\Pr(\mathbf{p}|I) \propto \sqrt{\det(\mathbf{F}(\mathbf{p}))}$  with  $\mathbf{F}$  symbol referring to the Fisher Information matrix (Jeffreys 1946) and "det" to the matrix determinant.

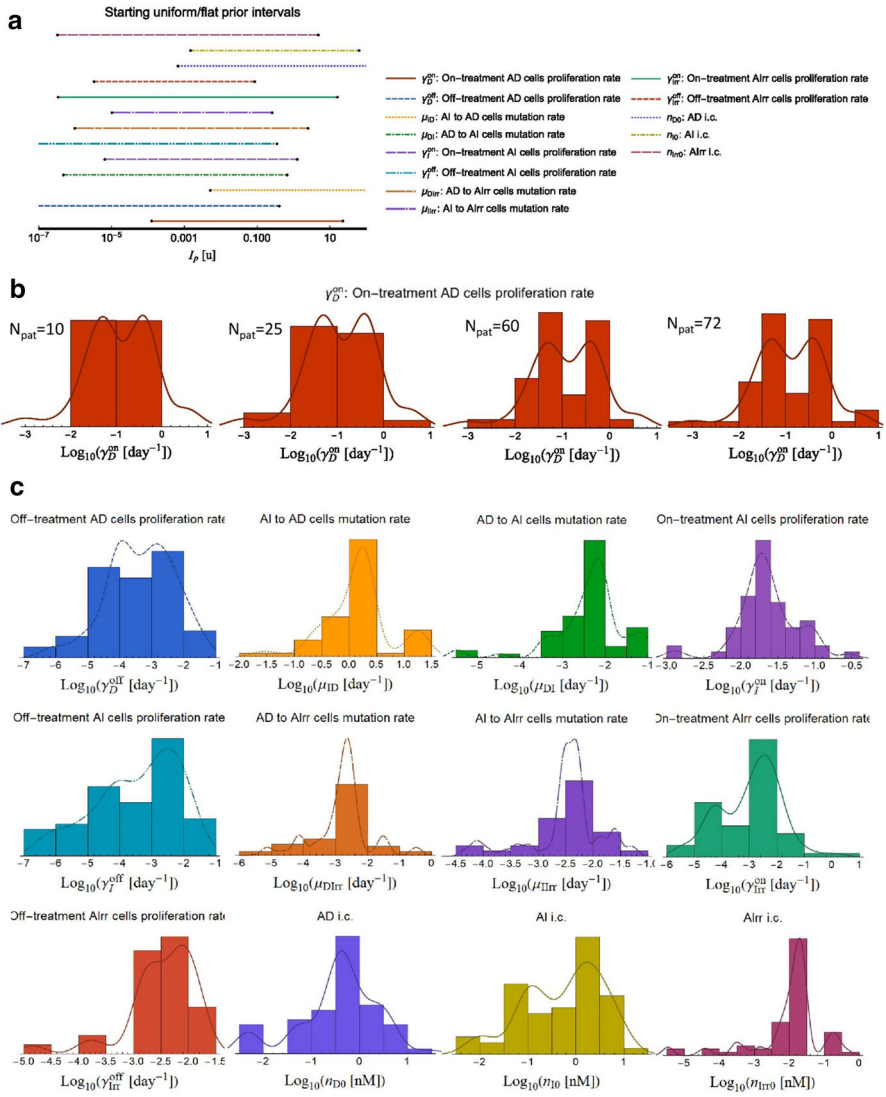
Extra than testing with flat/Jeffreys priors, in the numerical nested sampling approach, we explore the parameter space logarithmically to avoid divergences, and once we reach a statistically significant sample, i.e., above the Poisson noise fluctuation ( $\sim \sqrt{N_{\text{pat}}}$ ) shaping the posterior PDF, we proceed to implement the posterior as a prior for the patients analyzed in the dataset; finally, we reiterate by implementing a recursive determination of the prior (Fig. 2b, c). Further details can be found in Pasetto et al. (2021), where we discuss Bayesian analysis of retrospective data to guide clinical decisions.

## 4 Analysis of IADT Mathematical Models

Here we consider only IAD models due to current clinical interest. Each model is presented and justified in a biological and mathematical sense in the original paper where the model was first presented, and we refer the reader to that papers for detailed model derivations. Similarly, the sensitivity analysis of the model parameters is presented in each paper individually, and we elaborate on it here only where necessary. We will refer to the relapsing patient set as  $\Omega$ -set and not relapsing to as relapse  $\neg\Omega$ -set.

We parametrize the individual IAD data with a patient-specific control function  $T_{ps}$  defined as follows:  $T_{ps}(t) = \sum_{i=1}^n \mathbf{1}_{\tau_i}(t)$ ,  $0 < t \in [t_{\min}, t_{\max}]$ <sup>1</sup> with  $t_{\min}$  and  $t_{\max}$  minimum and maximum patient-specific treatment under consideration (e.g., Fig. 1a) and  $t_{\min}$  generally after the first treatment drop;  $n \geq 1$  is the number of intervals  $\tau_i$  considered.  $\tau_i \subseteq ]t_{\min}, t_{\max}[ \forall i$  is referred to as the "i<sup>th</sup> treatment cycle,"

<sup>1</sup> We use compact set notation here, e.g.,  $0 < t \in [t_{\min}, t_{\max}]$  means all the possible values of  $t$ , positive, between  $t_{\min}$  and  $t_{\max}$ , i.e.,  $0 < t_{\min} \leq t \leq t_{\max}$ . Open brackets will exclude the borders, e.g., soon after  $\tau_i \subseteq ]t_{\min}, t_{\max}[$  means that the interval  $\tau_i$ , e.g.,  $\tau = [a, b]$  is properly included between  $t_{\min}$  and  $t_{\max}$ , but the limits of  $t_{\min}$  and  $t_{\max}$  are excluded:  $t_{\min} < \tau_i < t_{\max}$ . This allows us to work with the domain of existence of the indicator functions but arbitrarily truncate it at  $t_{\min}$  or  $t_{\max}$ .



**Fig. 2** Model prior development. This example refers to the model by Hirata et al. 2010 and its 13 defining parameters. A similar technique is adopted for the other models. a. Initial bounded flat prior. b. Evolution of prior development for  $\gamma_D^{on}$  [day<sup>-1</sup>] as the number of patients analyzed is increased ( $N_{pat} = \{10, 25, 60, 72\}$ , respectively). c Final priors for the remaining 12 parameters (colors correspond to those shown in panel a) (Color figure online)

and  $\mathbf{1}_{\tau_i}$  is the indicator function<sup>2</sup> for the interval  $\tau_i$  (defined as  $\mathbf{1}_{\tau_i} = 1$  for  $t \in \tau_i$ , 0 otherwise). For modeling purposes, the weights/errors,  $e_i$  for each data  $i$ , have been assigned either uniformly  $e_i = \text{cnst} \cdot \forall i$  or with a linear decreasing relevance from the last PSA concentration  $c_{PSA}$  peak, say  $\hat{c}_{PSA}$  (e.g., in  $\tau_3$  of Fig. 1) with  $e_i = |c_{PSAi} - \hat{c}_{PSA}|_{t=\hat{t}}$ , i.e. at  $t = \hat{t}$ , and  $e_i = |\hat{c}_{PSA} - |\hat{t} - t| + |c_{PSAi}|$  for  $t \neq \hat{t}$ . Finally, we performed sensitivity analysis on all the models included here. Comments on the technique adopted are technical and left to Supplement A.

### 4.1 Ideta et al. (2008)

The model by Jackson (2004) can be considered the continuous ADT model prototype. Its extension to IAD therapy of interest was presented by Ideta et al. (2008). In this model (hereafter, I08), the authors drop the dependence of Jackson’s model on the spatial distribution, which is only of theoretical interest but not resolved in clinical PSA data. Model simulations predict that intermittent ADT can only prevent progression if normal androgen levels decrease the growth rate of AI cells, which may be biologically unlikely since AI cells have androgen receptors with increased sensitivity (Grossmann et al. 2001). We consider the I08 model in the following form:

$$\begin{aligned} \frac{dn_D}{dt} &= (\gamma_D - \delta_D - \mu_{DI})n_D, \\ \frac{dn_I}{dt} &= \mu_{DI}n_D + (\gamma_I - \delta_I)n_I, \end{aligned} \tag{1}$$

with initial conditions  $n_D(t_{0D}) = n_{D0}$ ,  $n_I(t_{0I}) = n_{I0}$ .<sup>3</sup> As previously mentioned,  $n_D$  and  $n_I$  are the androgen-dependent and -independent population number of cells (or concentration).  $\gamma_i$  and  $\delta_i$ ,  $i \in \{D, I\}$  are growth and apoptosis rates for AD and AI cells, given, respectively, by:

$$\begin{aligned} \gamma_D &= \gamma_{D\max} \left( \gamma_{DA} + (1 - \gamma_{DA}) \frac{c_A}{c_A + k_{DA\gamma}} \right), \quad \gamma_I = 1 - \left( 1 - \frac{\delta_{IA}}{\gamma_{IA}} \right) \frac{c_A}{c_{A0}}, \\ \delta_D &= \delta_{D\max} \left( \delta_{DA} + (1 - \delta_{DA}) \frac{c_A}{c_A + k_{DA\delta}} \right), \quad \delta_I = 1. \end{aligned} \tag{2}$$

In Eq. (2)  $\gamma_{D\max}$  and  $\delta_{D\max}$  are the maximal AD proliferation and apoptosis rates,  $\delta_{DA}$  is a control parameter on the effect of low androgen levels on the AD apoptosis rate,  $k_{DA\gamma} \neq 0$  is the AD half-saturation rate,  $k_{DA\delta} \neq 0$  is the AD apoptosis rate dependence on androgen. Finally,  $\delta_{IA}$  and  $\gamma_{IA} \neq 0$  modulate hormonally patient failing death and growth. Mutation from AD to AI cells are allowed at a mutation rate:

$$\mu_{DI} = \mu_{DI\max} \left( 1 - \frac{c_A}{c_{A0}} \right), \tag{3}$$

thus, the mutation rate decreases as the androgen (here normalized at its homeostatic level  $c_{A0} \neq 0$ ) approaches its max value  $\mu_{DI\max}$ . A decoupled ODE model of the

<sup>2</sup> Note that in our case the indicator function  $\mathbf{1}_{\tau_i}$  is not-continuous scalar function, but traditionally indicated with bold characters even if it is not a matrix or a vector.

<sup>3</sup> Note how, in general  $t_{0D} \neq t_{0I}$  for what is seen in Sect. 2.2.



serum androgen concentration under treatment  $c_A$  is given by:

$$\frac{dc_A}{dt} = \delta_{c_A}(c_{A0} - c_A) - \delta_{c_A}c_{A0}T_{ps}, \tag{4}$$

with initial condition  $c_A(t_{0A}) = c_{A0} \neq 0$ , where  $\delta_{c_A}$  is the androgen clearance rate. Here  $T_{ps} = T_{ps}(t)$  is the patient treatment-specific function as defined in Sect. 2.1. Finally, the PSA density concentration of interest to us,  $c_{PSA}$ , is a linear combination with weight  $w_i$  of the population densities

$$c_{PSA} = \sum_{i \in \{D, I\}} w_i n_i. \tag{5}$$

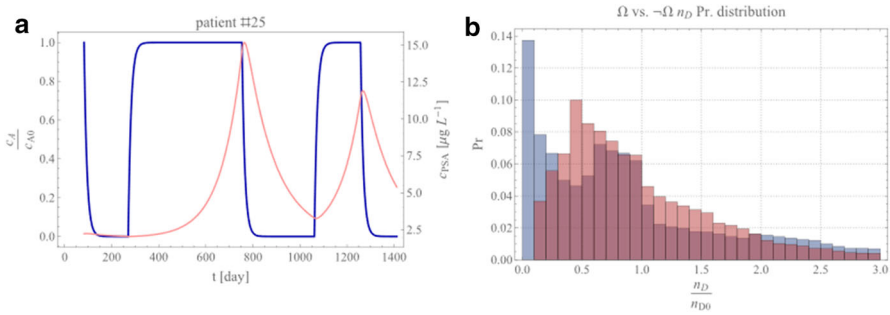
Based on the original analysis of Ideta et al. and the available dataset, we explored two versions of this model, namely, where  $\delta_{IA} = \gamma_{IA}$  in Eq. (2), i.e.,  $\gamma_I = \text{cnst}$ . (hereafter model I08A) and the original form of the equations ( $\delta_{IA} \neq \gamma_{IA}$  hereafter model I08B).

### 4.1.1 I08A in the Context of the Data

We noted that the system of equations (hereafter SoE) composed by Eqs. (1)–(4) decouples in the androgen concentration  $c_A$ . The analysis of the system results in a line of infinite equilibria on the intersection of the plane  $n_D = 0$  with the plane  $c_A = c_{A0} - c_{A0}T_{ps}$  in the space of phase-state variables  $(n_D, n_I, c_A)$ . Thus,  $c_A = c_{A0}$  off-treatment and  $c_A = 0$  on-treatment. Standard linear stability analysis (Wiggins 2003) shows that the Jacobian of the system produces a null generalized eigenvalue  $\lambda_1 = 0$ , a negative one  $\lambda_2 = -\delta_A$ , and a more complicate third generalized eigenvalue that takes, off-treatment, the elegant form:  $\lambda_3^{\text{off}} = \gamma_{D\text{max}} + \frac{(\gamma_{DA} - 1)\gamma_{D\text{max}}k_{D\gamma/2}}{c_{A0} + k_{D\gamma/2}} - \delta_{D\text{max}} - \frac{(\delta_{DA} - 1)\delta_{D\text{max}}k_{D\delta/2}}{c_{A0} + k_{D\delta/2}}$ . The sign of  $\lambda_3^{\text{off}}$  can be evaluated for the best-fit parameter values that result from the inference works of Sect. 3 in the patients’ cohort considered here (Sect. 2), resulting in being always positive for all the patients. Therefore, the above-found equilibria lines represent a 1D nonstable manifold, and further investigations (e.g., in the context of the central manifold theory) are not of additional interest to us.

We are indeed more interested to further exploit the characteristics of the present dataset in the context of this model by using the decoupled nature of the serum androgen concentration  $c_A$ . All the patients are considered from their first cycle of treatment, starting with  $T_{ps}(t) = 1$  for  $t \in \tau_1$ . Hence, we can emulate with a Heaviside step function  $T_{ps} = \theta(-t)$  a cycle of treatment followed by the off-treatment period for a suitable cyclic interval (on–off, on–off, on–off, and so forth) around the off-treatment start, set at  $t = 0$ . Within this approach, the general solution of Eq. (4) is algebraic and reads:

$$c_A(t) = c_{A0}e^{-\delta_A(t+1)}(e^{\delta_A}\theta(e^{\delta_A t} - 1) + 1). \tag{6}$$



**Fig. 3** Ideta et al. model analysis results. **a** Evolution of the normalized androgen concentration  $c_A$  (normalized to its homeostatic value  $c_{A0}$ , (left y-axis) as a function of time (blue curve) for a representative patient. For completeness, the PSA profile is also reported in light red (right y-axis). **b** Normalized androgen-dependent probability distribution functions on the cohort of best-fit parameters for progressive ( $\Omega$ , red) and responsive ( $-\Omega$ , blue) sets (Color figure online)

This equation is monotonic on the two phases on/off-treatment because the derivative  $dc_A/dt = c_{A0}\delta_A e^{-\delta_A(t+1)}(e^{\delta_A\theta} - 1)$  is never null neither for  $t < 0$ , i.e., on-treatment nor for  $t \geq 0$ , off-treatment. By splitting the treatment in on/off-time, we can always reverse the bilinear map  $c_A = c_A(t)$  in  $t = t(c_A)$ . For example, in our case, it reads  $t = -\frac{1}{\delta_A} \log\left(\frac{c_A}{c_{A0}}\right) + \delta_A$  on-treatment and  $t = \frac{1}{\delta_A} \log\left(\frac{c_{A0}(e^{-\delta_A}-1)}{c_A - c_{A0}}\right)$  off-treatment<sup>4</sup> for  $c_A \neq c_{A0}$ , and  $c_A \neq 0$  and  $\delta_A \neq 0$ . We exploit Eq. (6) to obtain the probability distribution function (PDF) of the orbits over all the sets of patients remapping each cycle over the phase space section  $(O, n_D, n_I)$ . We take advantage by the sharp  $c_A$  passage from its homeostasis value  $c_{A0}$  to null and vice versa in conjunction with the bijection map just found. Figure 3a shows the  $c_A$  profile for a representative patient. While time is a monotonic increasing function, the map we considering is one-to-one only over the treatment cycle  $T_{ps} = 1$  and the off-cycle  $T_{ps} = 0$ , respectively, and in these two tracks we can write the SoE as  $n_D = n_D(t(c_A)) = n_D(c_A)$  and  $n_I = n_I(c_A)$ . As  $c_A$  sharply switches from  $c_A = c_{A0}$  and  $c_A = 0$ , we can limit ourselves to a first-order solution of the SoE. After simple algebra, we arrive at the approximate solution of the SoE in the form:

$$n_D \simeq n_{D0} - \frac{1}{c_{A0}\delta_A} n_{D0}(c_A - c_{A0}) \left( \frac{\gamma_D \max(c_{A0} + \gamma_A k_{D\gamma/2})}{c_{A0} + k_{D\gamma/2}} - \frac{\delta_D \max(c_{A0} + \delta_D k_{D\delta/2})}{c_{A0} + k_{D\delta/2}} \right), \quad (7)$$

$$n_I \simeq n_{I0},$$

to the first order in  $c_A$  (and where  $\simeq$  means asymptotic-to). As evident, the second equation remains close to its initial value  $n_{I0}$ , while the first is perturbed away, suggesting that we can first sample the PDF of the dataset for fixed values in  $n_I$  around  $n_{I0}$  and then investigate the PDF as sampled from the best fit obtained by the patient in the trial with Eq. (7). The results are shown in Fig. 3b. The trend of the two distributions

<sup>4</sup> Note how we could consider the resulting SoE as function of the variable  $c_A$  to reach a fully algebraic solution of the system by taking the ration of  $\frac{dn_D}{dc_A} / \frac{dn_I}{dc_A}$ . Nevertheless, it is more fruitful to look at the trend of Eq. (6) as obtained by the best fit procedure introduced in the next section.

for the development of resistance and continuing response patients is comparable as above the starting value  $n_D = n_{D0}$ , while the trend diverges for smaller values  $n_D$ . Because we assume  $n_D$  is a proxy for  $c_{PSA}$  at small values of  $n_I$ , as evicted from Eqs. (5) and (7), if the model correctly interprets the data, then a *patient with an initial PSA-drop below 10% of its initial value* is highly likely to be a continuous responder. The risk of resistance development grows to about 50% when the initial drop in PSA is around 30%.

### 4.1.2 I08B in the Context of the Data

For I08B, where  $\delta_{IA} \neq \gamma_{IA}$ , the ratio presented in Eq. (3) evidences the structural non-identifiability of the SoE. The treatment of the equilibria and their stability is more straightforward in this model form than in I08A. The only equilibrium point is given by  $\{n_D, n_I, c_A\}_{eq} = \{0, 0, c_{A0} - c_{A0}T_{ps}\}$  with generalized eigenvalues  $\lambda_i$  of the Jacobian at the equilibrium given by  $\lambda_1 = (T_{ps} - 1)(\gamma_{IA} - \delta_{IA})\gamma_{IA}^{-1}$  with  $\gamma_{IA} \neq 0$ ,  $\lambda_2 = \lambda_2^{I08A}$  and  $\lambda_3 = \lambda_3^{I08A}$ . Following the I08A assumptions, we investigate the model under the conditions  $\delta_{Dmax} > \gamma_{Dmax}$ ,  $\delta_{DA} > 1$ , and by requiring that  $\mu_{max} < \gamma_I - \delta_I$  to avoid the annihilation of the populations. Under these conditions, we can prove that  $\lambda_1 \leq 0$  and  $\lambda_2 \leq 0$ , and that for  $\lambda_3$  it holds the same consideration as for I08A due to the non-stable nature of the resulting equilibrium manifold.

Analogous consideration on the non/relapsing treatment holds for I08B as for I08A, but with more straightforward treatment for I08B than for I08A: the two equilibria at homeostasis  $c_A = c_{A0}$  and at null androgen concentration,  $c_A = 0$ , attract the dynamics as for I08A and self-explain the orbit profiles. Therefore, the identical results from the inference of I08B on patients' trials can be obtained for the PDF but are not depicted again.

### 4.2 Eikenberry et al. (2010)

The model developed by Eikenberry et al. (2010, hereafter E10) was an attempt to describe the interaction between testosterone (T, the primary androgen in the serum), its enzyme 5a-reductase to dihydrotestosterone (DHT), and their binding (T:AR and DHT:AR) with the androgen receptors (AR) in the prostate. Because of model E10's versatility, we have included it in the IAD treatment model comparison. Of note, the authors have not proposed the model to fit data, and here we reinterpret E10 beyond the scope of the original paper. The modulation due to intermittent IAD is assumed in testosterone time modulation. While a linear relation might not be readily available from the literature between testosterone and PSA level (Elzanaty et al. 2017), we recode the testosterone concentration  $n_T$ , in E10 as follows:

$$\frac{dn_T}{dt} = n_T \left( \delta_T - \frac{\mu_{cat}n_{5\alpha}}{k_M + n_T} - \kappa_{T:R}n_R \right) + \delta_{T:R}q_{T:R} - (T_{ps} - 1)\Upsilon(n_S), \quad (8)$$

which we couple with the original system of equations:

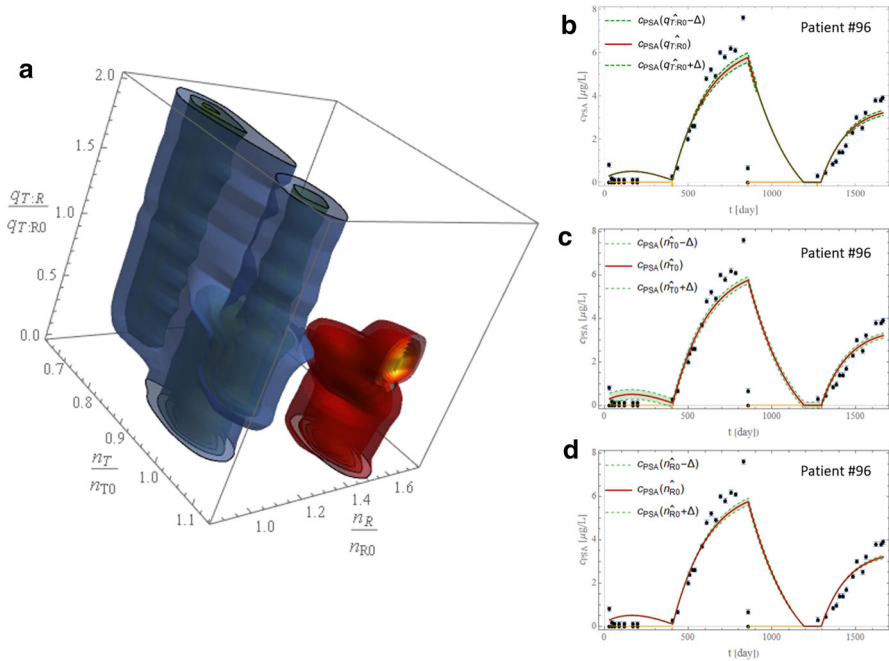
$$\begin{aligned}
 \frac{dn_R}{dt} &= n_R(\gamma_R - \delta_R - \kappa_{DHT}n_{DHT} - \kappa_{T:R}n_T) + \delta_{DHT:R}q_{DHT:R} + \delta_{T:R}q_{T:R}, \\
 \frac{dn_{DHT}}{dt} &= \frac{\mu_{cat}n_{5\alpha}n_T}{k_M + n_T} - n_{DHT}(\delta_{DHT} + \kappa_{DHT}n_R) + \delta_{DHT:R}q_{DHT:R}, \\
 \frac{dq_{T:R}}{dt} &= \kappa_{T:R}n_Rn_T - \delta_{T:R}q_{T:R}, \\
 \frac{dq_{DHT:R}}{dt} &= \kappa_{DHT}n_{DHT}n_R - \delta_{DHT:R}q_{DHT:R},
 \end{aligned}
 \tag{9}$$

with five nominals initial conditions:  $n_{R0} = n_R(t_{0R})$ ,  $n_{T0} = n_T(t_{0T})$ ,  $n_{DHT0} = n_{DHT}(t_{0DHT})$ ,  $q_{T:R0} = q_{T:R}(t_{0T:R})$  and  $q_{DHT:R0} = q_{DHT:R}(t_{0DHT:R})$ . Here, the treatment function  $T_{ps}$  modulates testosterone influx into the prostate-function  $\Upsilon(n_S)$  original in E10 and that we are going to adopt here, where  $n_S$  is the testosterone serum concentration. Furthermore, we consider the androgen receptor concentration  $n_R$  and the dihydrotestosterone concentration  $n_{DHT}$  together with two quota concentrations  $q_{T:R}$  and  $q_{DHT:R}$  (Droop 1968), here, taken to be the T:AR complex and the DHT:AR complex concentration, respectively.  $\gamma_R$  is the AR production rate,  $\delta_R$  is the AR degradation rate,  $\delta_T$  is the testosterone-specific degradation rate, and  $\delta_{DHT}$  is the dihydrotestosterone degradation rate. The mass-action constants for the androgen-dependent component (testosterone) and dihydrotestosterone binding the AR are  $\{\kappa_a^T, \kappa_d^T, \kappa_a^{DHT}, \kappa_d^{DHT}\}$ , and the  $5\alpha$  reductase converts T to DHT by Michaelis–Menten enzyme kinetics with concentration  $n_{5\alpha}$ , turnover number  $\mu_{cat}$  and constant  $k_M \neq 0$ .

### 4.2.1 The Model in the Context of the Data

If we set  $a \equiv \mu_{cat}n_{5\alpha} - \delta_Tk_M$ ,  $b \equiv (1 - T_{ps})\Upsilon(n_S)$ , and  $\Delta \equiv \sqrt{(a + b)^2 - 4b\delta_Tk_M}$ , then two critical points can be isolated at the intersection of the nullclines hyperplanes of the phase space. On-treatment, the first point  $\{n_R, n_T, n_{DHT}, q_{T:R}, q_{DHT:R}\}_{eq}^{(1,2)} = \left\{0, 0, 0, \frac{-a-b \mp \Delta}{2\delta_T}, \frac{-a+b \pm \Delta}{2\delta_{DHT}}\right\}$  holds as soon as  $\mp a + b + \Delta \leq 0 \wedge \mp a + \Delta \leq b$ . While only the second of these equilibria is of biological interest, it is not a stable equilibrium. Obtaining the complete set of generalized eigenvalues requires a cumbersome solution of three cubic equations, yet the check for the stability requires much less effort once we realize that one of the generalized eigenvalues from the characteristic equations reads simply  $\delta_T - \frac{4\delta_T^2\mu_{cat}k_Mn_{5\alpha}}{(a+b-\Delta-2\delta_Tk_M)^2}$  where  $a + b - \Delta - 2\delta_Tk_M \neq 0$  and that it proves to be always positive for all the inference results in the trial patients.

Finally, we note how the model could represent an essential instrument for investigating the relapsing mechanism evidenced in some patients, which remains one of the goals of this work for its potential clinical implications. We identify three over five state variables by inspecting the phase-state space with a striking separation between  $\Omega$  and  $-\Omega$ . Figure 4a shows the 3D probability distribution function of  $n_T, n_R$ , and  $q_{T:R}$ . The density map of the temporal evolution of  $\Omega$  and  $-\Omega$  sets clusters (over the orbital evolution spanned by the patients analyzed) on a well distinct area of the phase



**Fig. 4** Eikenberry et al. model analysis results. **a** Probability distribution function of the  $q_{T:R}$ ,  $n_T$ , and  $n_R$  space. The isocontours for the Pr of  $\Omega$  and  $\neg\Omega$  sets are shown in blue and red, respectively. A few isocontours are shown at the border-slicing-planes for  $\text{Pr} = \{0.1, 0.68, 0.95\}$ . **b–d** Sensitivity of  $c_{PSA}$  in response to changes in the normalized values of  $q_{T:R}$ ,  $n_T$  and  $n_R$  for a representative patient. The optimal fit  $c_{PSA}$  dynamics (i.e., for optimal parameters  $\hat{\mathbf{p}}$ ) and corresponding data are shown by the red curve and black dots with error bars, respectively; dashed green curves and the corresponding shadows show the sensitivity of  $c_{PSA}$  when the parameters are increased/decreased by  $\Delta$  (Color figure online)

space, splitting in the  $n_T$  vs.  $n_R$  space and at least partially in the orthogonal  $q_{T:R}$  space.

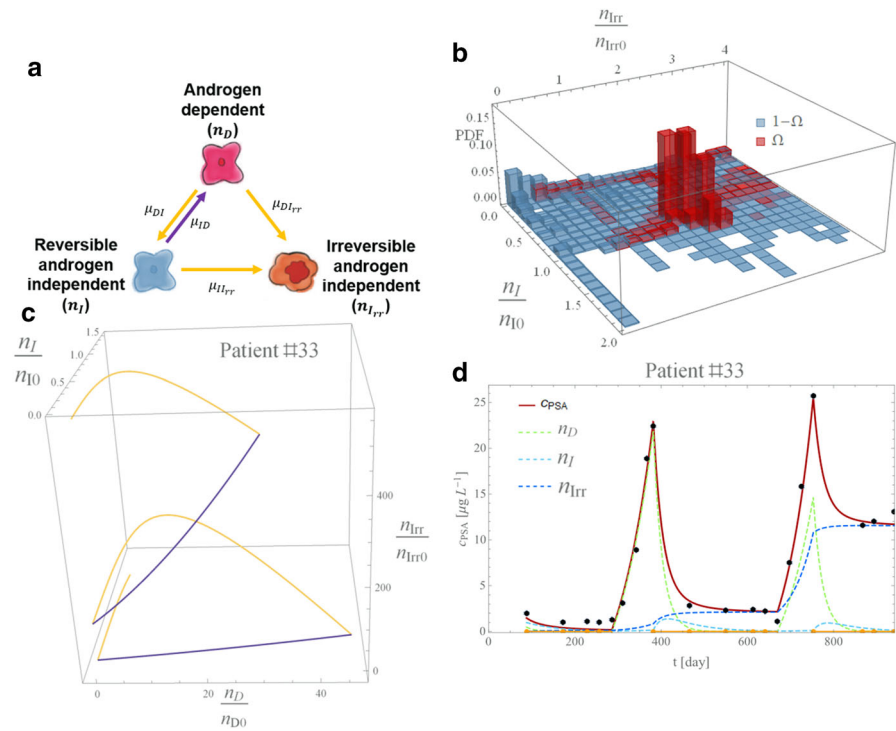
In Fig. 4, panels b, c, and d, we exploited the Direct Differential Method (DDM) for sensitivity analysis to track the time dependence of the sensitivity  $S_{c_{PSA},j} \equiv \frac{\partial c_{PSA}(t, \hat{\mathbf{p}})}{\partial p_j}$  computed at the best fit parameter values  $\hat{\mathbf{p}}$ , where  $\mathbf{p} = \{n_{T0}, n_{R0}, q_{T:R0}\}$ , respectively. As shown in Fig. 4b–d, a slight variation of the parameters does not dramatically affect the trend of  $c_{PSA}$ . Thus, there is minimal sensitivity of  $c_{PSA}$  to the parameters. This result shows that *the PDF of the combination of parameters investigated might be an excellent tool to explore the origin of the resistance with the E10 model.*

The sensitivities were computed using the Direct Differential Method (DDM), as mentioned at the beginning of Sect. 4 and it is reported more in details in Supplement A. As evident from Fig. 4b–d, different parameters have different sensitivity on a different phase orbit with  $n_{T0}$  more sensitive under treatment and  $n_R$  or  $q_{T:R}$  more sensitive out of treatment. DDM not only demonstrates the stability of the results obtained but also

adds extra information on when a model is sensitive to a parameter change. This result is significant when dealing with models with varying behavior on- and off-treatment.

### 4.3 Hirata et al. (2010)

A series of studies (Tanaka et al. 2010; Hirata et al. 2012; Hirata and Aihara 2015) motivated the model by Hirata et al. 2010 (hereafter model H10) to capture intermittent ADT dynamics. The model is based on the coupled AD-AI population cells, supplemented with a population of irreversible AI cells, AI-Irr representing the first three-compartment model in the literature (Fig. 5a). Here we report the mathematical formulation in the proposed framework’s formalism and refer to the original paper for a detailed model description. The SoE reads with our generalized notation:



**Fig. 5** Hirata et al. model analysis results. **a**. Model schematic: under treatment (yellow arrows) and off-treatment (violet arrow). **b** PDF on the  $n_I$  and  $n_D$  space as phase space density histograms for resistant ( $\Omega$ , red) and responsive ( $-\Omega$ , blue) patients. **c**. Flex in the PSA profile under treatment and off-treatment in the  $n_D$ ,  $n_I$  and  $n_{Irr}$  phase space for patient #33. Colors match the sketch of panel **a**. **d**. PSA density profile (red curve), with data points with error bar (black dots). The  $n_I$ ,  $n_D$  and  $n_{Irr}$  populations are shown by the dashed blue, dashed green, and dashed cyan curves, respectively. Yellow lines along the x-axis show the intervals of treatment (Color figure online)

$$\begin{aligned}
 \frac{dn_D}{dt} &= n_D(T_{ps}(\gamma_D^{on} - \gamma_D^{off}) + \gamma_D^{off}) + \mu_{ID}(1 - T_{ps})n_I, \\
 \frac{dn_I}{dt} &= \mu_{DI}T_{ps}n_D + n_I(T_{ps}(\gamma_I^{on} - \gamma_I^{off}) + \gamma_I^{off}), \\
 \frac{dn_{Irr}}{dt} &= \mu_{DIrr}T_{ps}n_D + \mu_{IIrr}T_{ps}n_I + n_{Irr}(T_{ps}(\gamma_{Irr}^{on} - \gamma_{Irr}^{off}) + \gamma_{Irr}^{off}),
 \end{aligned}
 \tag{10}$$

$n_D(t_{0D}) = n_{D0}, n_I(t_{0I}) = n_{I0}, n_{Irr}(t_{0Irr}) = n_{Irr0}$ , where terms retain the identical biological meaning as previously described and the two irreversible, and reversible changes in the AI cell population are considered with the relative growth rate  $\gamma_i^{on/off}$  on- and off-treatment with  $i \in \{D, I, Irr\}$ . The serum concentration is computed as in Eq. (5) for  $i \in \{D, I, Irr\}$ .

### 4.3.1 The Model in the Context of the Data

During both on- and off-treatment cycles, nullclines analysis leads to  $\{n_D, n_I, n_{Irr}\}_{eq} = \{0, 0, 0\}$  as the only equilibrium point. By setting  $a \equiv \gamma_D^{off} + \gamma_I^{off}$ ,  $b \equiv \gamma_D^{on} + \gamma_I^{on}$ ,  $c \equiv \gamma_D^{off} - \gamma_I^{off}$  and  $d \equiv \gamma_D^{on} - \gamma_I^{on}$ , with the discriminant  $\Delta$  implicitly defined by the relation  $\Delta^2 = c^2 + T_{ps}(T_{ps}((c - d)^2 - 4\mu_{DI}\mu_{ID}) - 2c(c - d) + 4\mu_{DI}\mu_{ID})$ , we can write the generalized eigenvalues of the Jacobian at the equilibrium as  $\lambda_1 = \gamma_{Irr}^{off} + T_{ps}(\gamma_{Irr}^{on} - \gamma_{Irr}^{off})$  and the other two  $\lambda_{2,3}$  in a compact form as  $\lambda_{2,3} = \frac{1}{2}(T_{ps}(b - a) + a \pm \Delta)$ . This result implies that the equilibrium is stable on-treatment and unstable off-treatment.

The phase space shows that responsive and resistant patients cluster differently on the phase-state variables. Figure 5b shows that the probability density function for the best-fit patient groups around the initial value for  $n_I \cong n_{I0}$  and  $n_{Irr} \cong 2.1n_{Irr0}$ . Thus, the irreversible component of the model offers a potential tool to disentangle patient responses from the model fitting. As the resistant patients are expected to increase their irreversible cell component (i.e., asymptotically  $n_{Irr} > n_{Irr0}$  with “>” meaning asymptotic greater), we note that  $n_I \ll n_{I0}$  in responsive patients.

The model structure allows for the simulation of various PSA profiles thanks to the introduction of a new degree of freedom carried out with the third-compartment equations. Figure 5c shows the phase space plane for an example taken from the  $\Omega$  set of patients (Patient #33), while Fig. 5d shows the quality of the captured PSA concentration  $c_{PSA}$  profile achieved by this model.

### 4.4 Portz et al. (2012)

The Portz et al. (2012) model is based on the cell quota concept (Droop 1968), which is modeled as:

$$\frac{dq_i}{dt} = \frac{v_{max}(q_{max} - q_i)(1 - T_{ps})}{(q_{max} - q_{imin})(k_{q/2} - T_{ps} + 1)} - \delta_q q_i + \gamma_{max}(q_{imin} - q_D), \tag{11}$$

with  $q(t_{0i}) = q_{0i}$  for  $i \in \{D, I\}$ . The cell quota can grow to the maximum cell quota rate  $\gamma_{max}$  and degrades at a constant rate  $\delta_q$ , with  $q_{max}$  representing the shared max

cell quota,  $v_{\max}$  the maximum cell quota uptake rate,  $q_{i\min} < q_{\max}$  the minimum cell quota for androgen, and  $1 \neq k_{q/2} > 0$  the uptake rate half-saturation level (Packer et al. 2011). The authors allow mutation between both cell populations, from AD to AI and vice versa, at rates  $\mu_{DI}$  and  $\mu_{ID}$  given, respectively, by the Hill's equations of index  $m = 2$ :

$$\mu_{DI}(q) = \mu_{DI\max} \frac{k_{DI/2}^m}{q^m + k_{DI/2}^m}, \mu_{ID}(q) = \mu_{ID\max} \frac{q^m}{q^m + k_{ID/2}^m}, \tag{12}$$

where  $\mu_{DI\max}$  is the maximum AD to AI mutation rate,  $\mu_{ID\max}$  is the maximum AI to AD mutation rate, and  $k_{DI/2}^m$  and  $k_{ID/2}^m$  are the cells mutation rate half-saturation level. The model follows the evolution of AD/AI cell populations,  $n_D$  and  $n_I$ , respectively, with the following equations:

$$\begin{aligned} \frac{dn_D}{dt} &= n_D \left( -\delta_D - \mu_{DI\max} \frac{k_{DI/2}^2}{k_{DI/2}^2 + q_D^2} + \gamma_{\max} \left( 1 - \frac{q_{D\min}}{q_D} \right) \right) + \mu_{ID\max} n_I \frac{q_I^2}{k_{ID/2}^2 + q_I^2}, \\ \frac{dn_I}{dt} &= n_I \left( -\delta_I - \mu_{ID\max} \frac{q_I^2}{k_{ID/2}^2 + q_I^2} + \gamma_{\max} \left( 1 - \frac{q_{I\min}}{q_I} \right) \right) + \mu_{DI\max} n_D \frac{k_{DI/2}^2}{k_{DI/2}^2 + q_D^2}, \end{aligned} \tag{13}$$

for  $q_i(t) \neq 0 \forall t$  and i.c.  $n_D(t_{0D}) = n_{D0}$  and  $n_I(t_{0I}) = n_{I0}$ . The cell apoptosis and proliferation rates are, respectively, given by  $\delta_i$  and  $\gamma_i$  for  $i = \{D, I\}$ . The authors model the quota for both AD and AI cell populations independently. In general, we assume  $q_{I\min} < q_{D\min}$  to ensure that AI cells have a greater proliferation capacity in low androgen environments and  $n_D(t_{0D}) \cong 0$  with  $t_{0D}$  soon after treatment, as well as  $n_I(t_{0I}) \cong 0$  at  $t_{0I}$  at the beginning of the first treatment. Furthermore, a communal maximum proliferation rate  $\gamma_{\max}$  between the two populations is assumed. Both AD and AI cells produce PSA at a baseline rate  $\gamma_{PSA0}$  under the androgen dependence specified by:

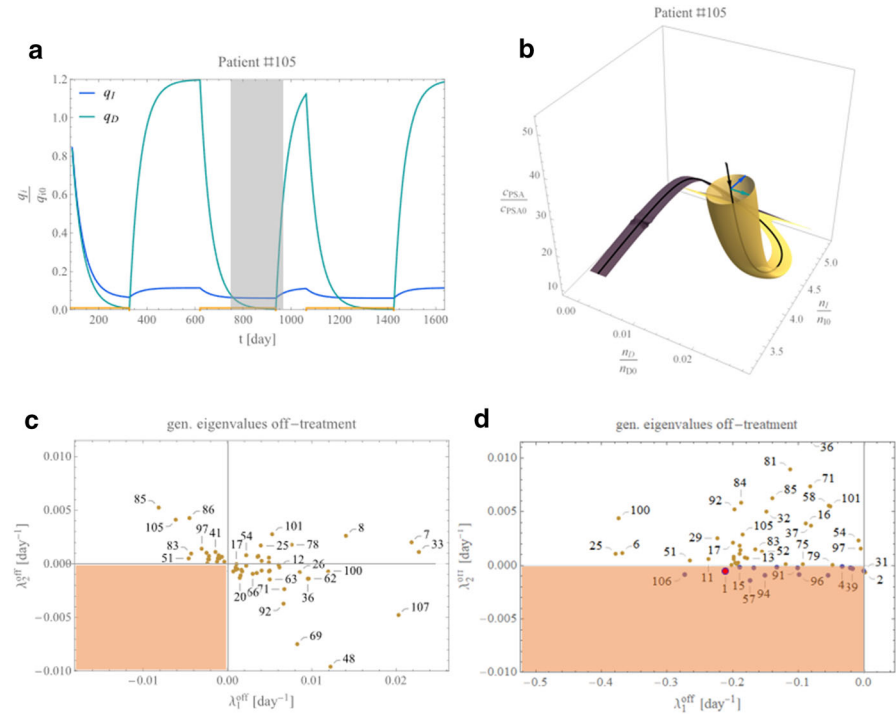
$$\frac{dc_{PSA}}{dt} = n_D \left( \gamma_{PSA0} + \frac{\gamma_{PSA,D} q_D^2}{k_{PSA,D/2}^2 + q_D^2} \right) - c_{PSA} \delta_{PSA} + n_I \left( \gamma_{PSA0} + \frac{\gamma_{PSA,I} q_I^2}{k_{PSA,I/2}^2 + q_I^2} \right), \tag{14}$$

with  $c_{PSA}(t_{0PSA}) = c_{PSA0}$ , and where  $k_{PSA,i/2}$  are the half-saturation rates and  $\gamma_{PSA,i}$  the growth rates, for  $i = \{D, I\}$ . Several variants of this quota model can be found in the literature. In the present work, we consider only a couple of them. A detailed comparison between (Hirata et al. 2010) and (Portz et al. 2012) can be found elsewhere (Everett et al. 2014). The model's complexity is demonstrated with a tube plot (Fig. 6a, b).

#### 4.4.1 P12A in the Context of the Data

The model is an extension of the models by Ideta et al., shown in Sect. 4.1, where the equation of the quota decouples from the two cell populations behavior. Nevertheless, the quota evolution  $q = q(t)$ , common to  $n_D$  and  $n_I$ , is generally





**Fig. 6** Portz et al. model analysis results. **a** Quotas for dependent and independent cells shaping the tube plot of panel **b** for a particular patient. The yellow dot-lines represent the on-treatment periods, and blue and green are the model’s independent and dependent quota levels, respectively. **b** The quota profiles are represented as a cross section. The orbital evolution corresponds to the gray box time interval in panel **a**. The blue and green arrows represent the independent and dependent quota levels of the model, respectively. On the orbit section, the tube cross section has been computed considering  $q_D$  along the normal  $N$  and  $q_I$  along the binormal  $B$  and by solving the Frenet–Serret formulas (assuming the vector field of the system under consideration being along the tangent  $T$ ). **c**, **d** The distribution of the eigenvalues  $\lambda_1^{off}$  and  $\lambda_2^{off}$  off-treatment for models P12B and P12A, respectively (Color figure online)

smoother than  $c_A = c_A(t)$  in I08A or I08B hence not justifying the approximations worked out in those models. In P12A, the only equilibrium point is at  $\{n_D, n_I\}_{eq}^{on/off} = \{0, 0\}$  and for the decoupled quota equation at  $q_{eq}^{on} = \frac{\gamma_{Dmax} q_{min}}{\gamma_{Dmax} + \delta_q}$  and  $q_{eq}^{off} = \frac{\gamma_{Dmax} (k_q/2 + 1) q_{min} (q_{max} - q_{min}) + q_{max} v_{max}}{(k_q/2 + 1) (\gamma_{Dmax} + \delta_q) (q_{max} - q_{min}) + v_{max}}$  for the on/off-treatments, respectively. The eigenvalues at this equilibrium point are real and negative along the direction of  $n_D$  and  $n_I$ :  $\lambda_i \in \mathbb{R}_0^-$ , for  $i = 1, 2$  both on- and off-treatment. In the decoupled  $q$  direction, the generalized eigenvalues  $\lambda_3^{on} = -(\gamma_{Dmax} + \delta_q)$  and  $\lambda_3^{off} = \lambda_3^{on} - \frac{v_{max}}{(k_q/2 + 1) (q_{max} - q_{min})}$  are always negative, leading to a node (attractor).

Nevertheless, we note that from the plot in Fig. 6c, how the best-fit solutions obtained from our inference work for all patients with this model falls in the area where  $\lambda_i > 0$ , for both  $i = 1$  and  $i = 2$ , i.e., we are never in the presence of an attractor (off-treatment). Therefore, patient dynamics never intercept an area of the

parameters' space defined by the hyperplane that would (eventually asymptotically) lead to the annihilation of the  $n_D$  and  $n_I$  cell population, i.e., a steady state or a reduction of the disease present under the detection threshold. This plot is compared with the companion model in the next section, which simplifies P12A.

#### 4.4.2 P12B in the Context of the Data

In this model, the authors extend the use of the quota concept to both  $n_D$  and  $n_I$  individually, i.e., fully exploiting Eq. (11), but retaining the same proliferation rate  $\gamma_{\max}$ . The large number of parameters required by the model makes the posterior maximization time-consuming and computationally expensive in the Bayesian framework, especially in a fully numerical nested sample approach (Skilling 2004) or using differential evolution optimization tool (Feoktistov 2006; Goode and Annin 2015). For this reason, a first inference approach has been performed within Laplace approximation and followed up at the patient-specific level where judged necessary.<sup>5</sup>

As in P12A, the P12B critical points are  $\{n_D, n_I, c_{PSA}\}_{\text{eq}} = \{0, 0, 0\}$  both on- and off-treatment, while for the decoupled quota equation-stability points are found at  $q_{i,\text{eq}}^{\text{on}} = \frac{\gamma_{\max}}{\delta_q + \gamma_{\max}} q_{i\text{min}}$  and  $q_{i,\text{eq}}^{\text{off}} = a^{-1} \mu_{\max} q_{i\text{min}} (k_{q/2} + 1) (q_{\max} - q_{i\text{min}}) + q_{\max} v_{\max}$  with  $a \equiv v_{\max} - (k_{q/2} + 1) (\delta_q + \gamma_{\max}) (q_{i\text{min}} - q_{\max}) \neq 0$ ,  $\delta_q \neq 0$  and  $\gamma_{\max} \neq 0$  and for  $i \in I, D$ . As in P12A, the three generalized eigenvalues of the Jacobian at the equilibrium are always negative. Equations for the generalized eigenvalues  $\lambda_i^{\text{on/off}}$  for  $i = 1, 2$  along the quota directions are analytically available but slightly cumbersome; vice versa, more interesting is the plot of  $\lambda_i^{\text{off}}$  for  $i = 1, 2$  shown in Fig. 6d. The P12B solutions distribute a small number of patients in the P12A inaccessible area of double negative generalized eigenvalues (orange square in Fig. 6c, d). *In this zone of the P12B parameter space off-treatment, the model predicts a constrained (or asymptotically constrainable) tumor cell population.* Finally, we note how P12A is nested in P12B. Thus, P12B always obtains a better score in the same data representation but suffers from overfitting. We investigate this problem and offer a solution in Sect. 5 in the context of the Bayesian model comparison.

#### 4.5 Morken et al. (2014)

In Morken et al. (2014), the authors extend model P12B by adding ADT-induced apoptosis of prostate cancer cells in addition to the inhibition of their growth and proliferation. Therefore, the model (hereafter M14) implements the per capita mortality of androgen-dependent and independent populations introduced in the previous section with the equation:

$$\delta_i(q_i) = \delta_i \max \frac{k_{i/2}^2}{q_i^2 + k_{i/2}^2}, \tag{15}$$

<sup>5</sup> Note on the statistical analysis: Details of statistical implementation are often omitted to focus the present work on the mathematical biology aspects pertinent to the work and the journal, but available upon request to the authors.

where  $k_{i/2}$  for  $i \in \{D, I\}$  are the apoptosis and half-saturation levels for the dependent and independent populations, respectively. We consider the SoE in the form of:

$$\begin{aligned} \frac{dn_D}{dt} &= n_D \left( -\delta_D - \frac{\delta_{D\max} k_{D\delta/2}^2}{k_{D\delta/2}^2 + q_D^2} - \frac{k_{DI/2}^2 \mu_{DI\max}}{k_{DI/2}^2 + q_D^2} + \gamma_{\max} \left( 1 - \frac{q_{D\min}}{q_D} \right) \right) + \frac{\mu_{ID\max} n_I q_I^2}{k_{ID/2}^2 + q_I^2}, \\ \frac{dn_I}{dt} &= \frac{k_{DI/2}^2 \mu_{DI\max} n_D}{k_{DI/2}^2 + q_D^2} + n_I \left( -\delta_I - \frac{\delta_{I\max} k_{I\delta/2}^2}{k_{I\delta/2}^2 + q_I^2} - \frac{\mu_{ID\max} q_I^2}{k_{ID/2}^2 + q_I^2} + \gamma_{\max} \left( 1 - \frac{q_{I\min}}{q_I} \right) \right), \end{aligned} \tag{16}$$

for  $q_i(t) \neq 0 \forall t$  and i.c.  $n_D(t_{0D}) = n_{D0}$  and  $n_I(t_{0I}) = n_{I0}$ , together with the equivalent of Eq. (14):

$$\frac{dc_{PSA}}{dt} = -c_{PSA} \delta_{PSA} + n_D \left( \gamma_{PSA0} + \frac{\gamma_{PSA,D} q_D^2}{k_{PSA,D/2}^2 + q_D^2} \right) + n_I \left( \gamma_{PSA0} + \frac{\gamma_{PSA,I} q_I^2}{k_{PSA,I/2}^2 + q_I^2} \right), \tag{17}$$

with i.c.  $c_{PSA}(t_{0PSA}) = c_{PSA0}$ . Furthermore, the same notation as in the models by Portz et al. is followed and not repeated here.

### 4.5.1 The Model in the Context of the Data

The analytical treatment is analogous to P12B but enriched in the dynamics variety for the extra parameters introduced in Eq. (15), although without changing equilibrium points.<sup>6</sup> Our model analysis did not report other notable features.

### 4.6 Baez and Kuang (2016)

The model by Baez and Kuang (2016) presents a variant of the P12A model that is able to fit PSA and androgen dynamics, thus improving PSA trend forecasting. Two models are presented in the authors' work and considered here. The first (hereafter B16A) is a single population model of cellular concentration  $n$ , and two equations are coupled with it, for  $\delta_{\max}$  the time-dependent (over a timescale  $\tau_{\delta_{\max}}$ ) maximum baseline cell death rate and  $c_{PSA}$  the PSA concentration that are modeled as:

$$\begin{aligned} \frac{dn}{dt} &= n \left( -n\delta - \frac{k_{n/2} \delta_{\max}}{q + k_{n/2}} - \frac{\gamma_{\max} q_{\min}}{q} + \gamma_{\max} \right), \\ \frac{d\delta_{\max}}{dt} &= -\tau_{\delta_{\max}} \delta_{\max}, \\ \frac{dc_{PSA}}{dt} &= q(\gamma_{PSA1} n + \gamma_{PSA0}) - \delta_{PSA} c_{PSA}, \end{aligned} \tag{18}$$

and a decoupled equation for androgen level:

$$\frac{dq}{dt} = \gamma(q_{\max} - q) - \gamma_{\max}(q - q_{\min}), \tag{19}$$

<sup>6</sup> Due to the complexity of the model, analogous inference approximations to P12B have been used in this analysis.

with  $n(t_0) = n_0$ ,  $c_{PSA}(t_{0PSA}) = c_{PSA0}$ ,  $\delta_{\max}(t_{0\delta_{\max}}) = \delta_{0\max}$ , and  $q(t_{0q}) = q_0 > 0$  strictly. The quota  $q \neq 0 \forall t$  is produced at a rate  $\gamma = \gamma_1 T_{ps} + \gamma_2$ .

In the same work, the authors also presented a two-populations model tracking both sensitive  $n_D$  and independent  $n_I$  cell evolution (hereafter B16B). By implementing their SoE within the approximation that all the cells have, on average, the same mass and density, we can recast their SoE in the form:

$$\begin{aligned} \frac{dn_D}{dt} &= n_D \left( -\frac{\delta_{D\max} k_{D/2}}{q + k_{D/2}} - \frac{k_{DI/2} \mu_{DI\max}}{q + k_{DI/2}} + \gamma_{\max} \left( 1 - \frac{q_{D\min}}{q} \right) \right) - \delta_D n_D^2, \\ \frac{dn_I}{dt} &= \frac{k_{DI/2} \mu_{DI\max} n_D}{q + k_{DI/2}} + n_I \left( \gamma_{\max} \left( 1 - \frac{q_{I\min}}{q} \right) - \frac{\delta_{I\max} k_{I/2}}{q + k_{I/2}} \right) - \delta_I n_I^2, \\ \frac{dq}{dt} &= -q(\gamma_2 + \gamma_{\max} + \gamma_1 T_{ps}) + \frac{\gamma_{\max}(q_{D\min} n_D + q_{I\min} n_I)}{n_D + n_I} + q_{\max}(\gamma_2 + \gamma_1 T_{ps}), \\ \frac{dc_{PSA}}{dt} &= q(\gamma_{PSA0} + \gamma_{PSA1}(n_D + n_I)) - \delta_{PSA} c_{PSA}, \end{aligned} \tag{20}$$

for  $n_i, i \in \{D, I\}$  never contemporaneously null, with initial conditions  $n_D(t_{0D}) = n_{D0}$ ,  $n_I(t_{0I}) = n_{I0}$ ,  $q(t_{0q}) = q_0$ , and  $c_{PSA}(t_{0PSA}) = c_{PSA0}$ . The maximum AD to AI mutation rate is given by  $\mu_{DI\max}$ . Furthermore, because AI cells,  $n_I$ , proliferate at lower androgen level it is assumed that  $q_{I\min} < q_{D\min}$ , and  $\delta_{D\max} > \delta_{I\max}$  because independent cells are less susceptible to apoptosis by androgen deprivation than sensitive cells.

### 4.6.1 B16A in the Context of the Data

The decoupled quota equation presents an equilibrium at  $q_{eq} = \frac{\gamma_{\max}(q_{\min} - q_{\max})}{\gamma_{\max} + (\gamma_2 + \gamma_1 T_{ps})} + q_{\max}$  when  $\gamma_{\max} + (\gamma_2 + \gamma_1 T_{ps}) \neq 0$ , belonging to the positive hyper-quadrant of the phase space (i.e., it is of biological interest). The remaining set in Eq. (18) shows two equilibria at  $\{n, \delta_{\max}, c_{PSA}\}_{eq}^{(1)} = \left\{ 0, 0, \frac{\gamma_{PSA,0} q_{eq}}{\delta_{PSA}} \right\}$ , which are always in the positive quadrant of the phase space of interest and  $\{n, \delta_{\max}, c_{PSA}\}_{eq}^{(2)} = \left\{ \frac{\gamma_{\max}(q_{eq} - q_{\min})}{q_{eq} \delta}, 0, \frac{\delta \gamma_{PSA,0} q_{eq} + \gamma_{\max} \gamma_{PSA,1}(q_{eq} - q_{\min})}{\delta_{PSA} \delta} \right\}$  with  $q_{eq} \neq 0$ ,  $\delta_{PSA} \neq 0$  and  $\delta \neq 0$ , which is also biologically meaningful. By studying the generalized eigenvalues, we see that the first of the equilibrium presents three negative generalized eigenvalues, one of which is always positive (i.e., it is a saddle point); the second equilibrium point produces the eigenvalues  $\lambda_1^{(2)} = \gamma_{\max} \left( \frac{q_{\min}}{q_{eq}} - 1 \right)$ ,  $\lambda_2^{(2)} = -\delta_{PSA}$ ,  $\lambda_3^{(2)} = -\tau_{\delta_{\max}}$  and  $\lambda_4^{(2)} = -\gamma_{\max} - (\gamma_2 + \gamma_1 T_{ps})$  which are all always negative, thus representing a stable point of attraction.

Due to the stability of the second equilibrium (on- and off-treatment), it is worth investigating the proximity of the patients' orbits to the equilibria on the Poincare sections involving the PSA concentration  $c_{PSA}$  we obtained from Eq. (18). Nevertheless, the low quality of the likelihood,  $L(\mathbf{p}) = \Pr(D|\mathbf{p}, I)$ , see Sect. 3) in the  $\Omega$  set of patients, demotivates further analysis. A single population  $n$  seems to not adequately capture disease progression, which remain the primary focus of our work, making the model less attractive for clinical implications and therefore not pushed forward here.

#### 4.6.2 B16B in the Context of the Data

The model presents cubic dependence on  $q$  and quadratic on  $n_D$ .<sup>7</sup> We select to investigate only the null-equilibrium point of independent and dependent cells. It is evident that  $n_D = 0$  is an equilibrium for the first of Eq. (20). Therefore, by assuming  $n_D = 0$  (and  $n_I > 0$  strictly), we can confirm the existence of two equilibria, the first located at  $\{n_I, q, c_{PSA}\}_{eq}^{(1)} = \left\{0, q_{max} + \frac{\gamma_{max}(q_{Imin} - q_{max})}{\gamma_{max} + (\gamma_2 + \gamma_1 T_{ps})}, \frac{\gamma_{PSA0}}{\delta_{PSA}} q_{eq}\right\}$ , for  $\gamma_{max} + (\gamma_2 + \gamma_1 T_{ps}) \neq 0$  and  $\delta_{PSA} \neq 0$ , which is of biological interest. The second, algebraically more cumbersome, reduces its nonnegativity condition to the simple one  $\delta_{I_{max}} + \frac{\gamma\gamma_{max}(q_{Imin} - q_{max})}{\gamma_{max}q_{Imin} + \gamma q_{max}} + \frac{\gamma\gamma_{max}(q_{Imin} - q_{max})}{k_{I/2}(\gamma + \gamma_{max})} \leq 0$ , that is verified over all studied patients.

Again, as explored in previous models, we are interested in the existence of negative generalized eigenvalues of the Jacobian at the equilibria off-treatment, i.e., a point of equilibrium with an asymptotic constrained expansion of the tumoral cell population. Despite the model complexity, it is easy to prove numerically that the Jacobian for both equilibrium points has at least one positive generalized eigenvalue, making these points saddle points that are not of interest to us.

#### 4.7 Elishmereni et al. (2016)

The Elishmereni et al. (Elishmereni et al. 2016) model accounts for two dynamics: disease dynamics represented by PSA used as a proxy for tumor volume and the pharmacology dynamics combined with the emergence of resistant cells from androgen receptor-independent  $n_I$  and testosterone androgen receptor-dependent  $n_{IAR}$  mechanism. The PSA concentration  $c_{PSA}$  of interest to us, is governed by the following numerically highly complex SoE<sup>8</sup>:

<sup>7</sup> Throughout algebraic manipulators as Mathematica or MAPLE it is possible to show that the system characterizing equation for  $c_{PSA}$  is algebraic of order 12, whose complete numeric roots investigation is beyond the scope of the present paper.

<sup>8</sup> Note on the integrators: details of numerical integration are often omitted to focus the present work on the mathematical biology aspects pertinent to the work and the journal, but available upon request to the authors.

$$\begin{aligned}
 \frac{dc_{PSA}}{dt} &= \widehat{c}_{PSA} \gamma_{T,PSA} \min\left(\gamma_{PSA} c_{PSA}^K, \frac{\log 2}{\gamma_{PSA \max}}\right) \\
 &\quad + \eta_{T,PSA} \left(c_{PSA} - \frac{\widehat{c}_{PSA}}{2}\right)^+ (\eta_{I,T} R_I \widehat{c}_{PSA} + n_T - 1), \\
 \frac{dn_T}{dt} &= \frac{\gamma_T (1 - T_{ps})}{\eta_{H,T} H + 1} - \gamma_T n_T, \\
 \frac{dH}{dt} &= T - \frac{\delta_T l_H^{\max} H e^{R_{T:AR}}}{e^{R_{T:AR}} + l_H^{\max}}, \\
 \frac{dR_{T:AR}}{dt} &= \gamma_{T:AR} T \widehat{R}_{T:AR}, \\
 \frac{dR_I}{dt} &= \gamma_I T \widehat{R}_I, \\
 \frac{dK}{dt} &= -\rho_K, \\
 \frac{dT}{dt} &= -\delta_T T
 \end{aligned} \tag{21}$$

with  $c_{PSA}(t_{0PSA}) = c_{PSA0}$ ,  $n_T(t_{0n_T}) = n_{T0}$ ,  $H(t_{0H}) = H_0$ ,  $R_{T:AR}(t_{0T:AR}) = R_{T:AR0}$ ,  $R_I(t_{0I}) = R_{I0}$ ,  $K(t_{0K}) = K_0$ , and  $T(t_{0T}) = T_0$  with  $(x)^+ = x\theta(x)$  ramp/positive function of the generic  $x$ ,  $\theta$  the previously introduced Heaviside step function. In the above equation  $\gamma_{PSA \max}$  is the limit to the PSA growth rate,  $\rho_K$  the K growth rate,  $\eta_{T,PSA}$  the testosterone,  $T$ , effect on the PSA growth,  $\gamma_T$  the instantaneous rate of change in  $T$ ,  $\eta_{H,T}$  the effect of intermediate components  $H$ , e.g., bound androgen receptor AR, on  $T$ , with same clearance rate  $\delta_T$ .  $\gamma_{T:AR}$  is the increase resistance rate,  $\gamma_I$  the increase resistance rate for testosterone-AR-independent paths  $R_I$ , and  $\eta_{I,T}$  rules the effect of  $R_I$  on the PSA growth. The growth rate of  $c_{PSA}$  is given by

$$\gamma_{PSA} = \begin{cases} 1 & c_{PSA} > c_{tPSA} \\ \sigma_{PSA} + (1 - \sigma_{PSA}) \frac{c_{PSA}}{c_{tPSA}} & c_{PSA} \leq c_{tPSA}, \end{cases} \tag{22}$$

where  $\sigma_{PSA}$  rules the steepness on the linear grown relation,  $c_{tPSA}$  the PSA threshold to switch in quiescent mode. Finally, control limits  $l_i$   $i \in \{PSA, H, n_I, n_{IAR}\}$  are added by hand to handle system divergences with a “manual”-bounding scheme ( $\widehat{f}_i \equiv \frac{(l_{i \max} - f_i)^+}{l_{i \max}}$  for the generic function  $f_i$ ).

In the practice the dynamics of the system is designed so that the instantaneous androgen rate of change  $\gamma_T$  is saturated by a control coefficient  $\eta_{T,PSA}$  through an intermediary delaying effect ruled by a delay modeling function  $H$  over the ADT therapy,  $T$  therapy function with scale factor  $\delta_{ADT}$  and a double mechanism for androgen independence cell population depending on  $\eta_{I,T}$ , and not depending on  $n_I$ , the androgen receptor (with the respective scale factor  $\gamma_I$  and  $\gamma_{T:AR}$ ).

### 4.7.1 The Model in the Context of the Data

The system has no equilibria influencing its dynamics, as evident from the 6th of Eq. (21). Further analysis is done in Sect. 5 to determine how well the model performs in the Bayesian model comparison.

### 4.8 Zhang et al. (2017)

Zhang et al. (2017) present a three-population competition model, based on Lotka—Volterra (LV) dynamics, where androgen-dependent  $n_D$ , androgen producing  $n_P$ , and androgen-independent cells  $n_I$ , are considered. Basing the approach on game theory, the authors derive a competition matrix  $\alpha = \alpha_{ij}$   $i, j \in \{D, I, P\}$  based on the parametrization of growth rates  $\gamma_i$  and carrying capacities  $K_i$  with  $i \in \{D, I, P\}$  resulting in this set of algebraic-differential equations:

$$\begin{aligned} \frac{dn_D}{dt} &= \gamma_D n_D \left( 1 - \frac{\alpha_{11}n_D + \alpha_{12}n_P + \alpha_{13}n_I}{n_P(\beta - T_{ps} + 1)} \right), \\ \frac{dn_P}{dt} &= \gamma_P n_P \left( 1 - \frac{\alpha_{21}n_D + \alpha_{22}n_P + \alpha_{23}n_I}{K_P} \right), \\ \frac{dn_I}{dt} &= \gamma_I n_I \left( 1 - \frac{\alpha_{31}n_D + \alpha_{32}n_P + \alpha_{33}n_I}{K_I} \right), \end{aligned} \tag{23}$$

where ADT is modeled by the decreasing carrying capacity with  $\beta < 1$  or supporting androgen-dependent cells with  $\beta > 1$ . The authors considered several constraints derived from the literature and researchers' experience to shape the model parameter influence:  $\alpha_{ii} = 1 \forall i, \alpha_{31} > \alpha_{21}, \alpha_{32} > \alpha_{12}, \alpha_{13} > \alpha_{23}, \alpha_{13} > \alpha_{21}, \alpha_{32} > \alpha_{31}$ , and  $\alpha_{ij} \in ]0, 1[ \forall i \neq j$ . Finally, the PSA dynamics is governed by:

$$\frac{dc_{PSA}}{dt} = \sum_{i \in \{D, P, I\}} n_i - \delta c_{PSA}, \tag{24}$$

with  $\delta$  the PSA clearance rate.

### 4.8.1 The Model in the Context of the Data

With the coupling of Eq. (24), the system presents four equilibria, but only two are of biological interest:  $\{n_D, n_P, n_I, c_{PSA}\}_{eq}^{(1)} =$

$$\left\{ 0, k_P, 0, \frac{k_P}{\delta} \right\} \in \mathbb{R}_0^{4+} \text{ and } \{n_D, n_P, n_I, c_{PSA}\}_{eq}^{(2)} = \left\{ \frac{k_P(\beta - \alpha_{12} - T_{ps} + 1)}{\alpha_{21}(\beta - \alpha_{12} - T_{ps} + 1) + 1}, \right.$$

$$\left. \frac{k_P}{\alpha_{21}(\beta - \alpha_{12} - T_{ps} + 1) + 1}, 0, \frac{k_P(\alpha - \alpha_{12} - T_{ps} + 2)}{\delta + \delta\alpha_{21}(\beta - \alpha_{12} - T_{ps} + 1)} \right\} \text{ where these ratios exist. For the first}$$

equilibrium, the eigenvalues of the Jacobian are positive in the  $n_D, n_P$  and  $c_{PSA}$  phase space and therefore of marginal interest. Vice versa, by setting  $a \equiv 1 + \beta, b \equiv \beta - \alpha_{12} + 1, d \equiv \alpha_{21}(\beta - \alpha_{12} + 1) + 1$  and  $e \equiv \beta + \alpha_{21}(\beta - \alpha_{12} + 1)^2 - \alpha_{12} + 1$  together with the discriminant squared  $\Delta^2 = (e\gamma_D + \beta\gamma_P + \gamma_P)^2 - 4ade\gamma_D\gamma_P,$

we can write the four eigenvalues of the Jacobian for the second equilibrium off-treatment as:  $\lambda_1^{(2)\text{off}} = -\delta$ ,  $\lambda_2^{(2)\text{off}} = \gamma_I - \frac{\gamma_I k_P (b\alpha_{31} + \alpha_{32})}{dk_I}$ ,  $\lambda_3^{(2)\text{off}} = -\frac{a\gamma_P + \Delta + e\gamma_D}{2ad}$  and  $\lambda_4^{(2)\text{off}} = \frac{\Delta - a\gamma_P - e\gamma_D}{2ad}$ , where the ratios exist, which are always negative for the fitted parameters, hence representing a stable equilibrium and opening the possibility to achieve an equilibrium off-treatment.

### 4.9 Phan et al. (2019)

The model (hereafter P19) presented by Phan et al. (Phan et al. 2019) is a variant of the work of Sect. 4.6 (Baez and Kuang 2016) in which the third population of weakly dependent cells,  $n_{wD}$ , is added to investigate the influence of extra degrees of freedom added by the new population. The death term is also adapted from Eq. (16). Retaining the notation used in Sect. 4.6, we can recast the model in the following form:

$$\begin{aligned} \frac{dn_D}{dt} &= n_D \left( -\frac{\delta_{D\max} k_{D/2}}{q + k_{D/2}} - \frac{2k_{DI/2} \mu_{DI\max}}{q + k_{DI/2}} + \gamma_{\max} \left( 1 - \frac{q_{D\min}}{q} \right) \right) \\ &\quad + \frac{k_{DI/2} \mu_{DI\max} n_{wD}}{q + k_{DI/2}} - \delta_D n_D^2 \\ \frac{dn_{wD}}{dt} &= n_{wD} \left( -\frac{2k_{DI/2} \mu_{DI\max}}{q + k_{DI/2}} - \frac{\delta_{wD\max} k_{wD/2}}{q + k_{wD/2}} + \gamma_{\max} \left( 1 - \frac{q_{wD\min}}{q} \right) \right) \\ &\quad + \frac{k_{DI/2} \mu_{DI\max} n_D}{q + k_{DI/2}} - \delta_{wD} n_{wD}^2 \\ \frac{dn_I}{dt} &= \frac{k_{DI/2} \mu_{DI\max} (n_D + n_{wD})}{q + k_{DI/2}} + n_I \left( \gamma_{\max} \left( 1 - \frac{q_{I\min}}{q} \right) - \frac{\delta_{I\max} k_{I/2}}{q + k_{I/2}} \right) - \delta_I n_I^2 \\ \frac{dq}{dt} &= q \left( -(\gamma_2 + \gamma_1 T_{ps}) - \gamma_{\max} \right) + \frac{\gamma_{\max} (q_{D\min} n_D + q_{I\min} n_I + q_{wD\min} n_{wD})}{n_D + n_I + n_{wD}} \\ &\quad + q_{\max} (\gamma_2 + \gamma_1 T_{ps}) \\ \frac{dc_{PSA}}{dt} &= q (\gamma_{PSA0} + \gamma_{PSA1} (n_D + n_I + n_{wD})) - \delta_{PSA} c_{PSA} \end{aligned} \tag{25}$$

with initial conditions  $n_D(t_{0D}) = n_{D0}$ ,  $n_{wD}(t_{0wD}) = n_{wD0}$ ,  $n_I(t_{0I}) = n_{I0}$ ,  $q(t_{0q}) = q_0$ ,  $c_{PSA}(t_{0PSA}) = c_{PSA0}$  together with the required biological inequalities  $q_{D\min} > q_{wD\min}$  and  $q_{D\min} > q_{I\min}$ .

#### 4.9.1 P19 in the Context of the Data

The idea of a third population is not new and already advanced with success in the model by Hirata et al. (2010). Nevertheless, the structure of the equations in Eq. (10) is very different from the Hirata et al. model in Eq. (10), with significantly more parameters not readily justifiable within the present dataset quality. Similar considerations were already worked out by Phan et al. We remark only that the complexity of the analysis, already evident in Sect. 4.6.2, is pushed further in this context, where only numerical investigation is available for equilibria



and stability. The only off-treatment equilibrium accessible by the orbits is the one for  $\{n_D, n_{wD}, n_I, q, c_{PSA}\}_{eq}^{off} = \left\{0, 0, 0, \frac{\gamma_{max}qI_{min} + \gamma_2q_{max}}{\gamma_2 + \gamma_{max}}, \frac{\gamma_{PSA0}(\gamma_{max}qI_{min} + \gamma_2q_{max})}{\delta_{PSA}(\gamma_2 + \gamma_{max})}\right\}$  and  $\delta_{PSA} \neq 0$  which is always positive with always negative eigenvalues  $\lambda_1^{off} = -\gamma_2 - \gamma_{max}$  and  $\gamma_2^{off} = -\delta_{PSA}$ . This is of limited biological interest as it is not compatible with the irreversibility nature of  $n_I$ , if not by surgical castration.

### 4.10 Brady-Nicholls et al. (2020)

The Brady-Nicholls et al. (2020) model (hereafter B20) is based on the hypothesis that prostate cancer stem cells' enrichment induces resistance. The model correlates stem cell proliferation with serum PSA through SoE for the prostate cancer stem cells  $n_S$ , the non-stem (differentiated) cells  $n_D$ , and for PSA serum concentration  $c_{PSA}$ . We report the system in the following way:

$$\begin{aligned} \frac{dn_S}{dt} &= \frac{p_S \log(2)n_S^2}{n_D + n_S}, \\ \frac{dn_D}{dt} &= \log(2)n_S \left(1 - \frac{p_S n_S}{n_D + n_S}\right) - \delta_D T_{p_S} n_D, \\ \frac{dc_{PSA}}{dt} &= \gamma_{PSA} n_D - \delta_{PSA} c_{PSA}, \end{aligned} \tag{26}$$

with initial conditions  $n_S(t_{0S}) = n_{S0}$ ,  $n_D(t_{0D}) = n_{D0}$  and  $c_{PSA}(t_{0PSA}) = c_{PSA0}$ . It is assumed that stem cells divide at rate  $\log(2)$ , and the division is either symmetric yielding two stem cells (Enderling 2015) or asymmetric, where the stem cell produces one stem and one differentiated cell. The parameter that governs this effect is  $p_S$ . The PSA differentiated cell production rate and PSA clearance rate are given by  $\gamma_{PSA}$  and  $\delta_{PSA}$ , respectively, and  $T_{p_S}$  is the patient-specific treatment function (see Sect. 2.1).

#### 4.10.1 The Model in the Context of the Data

The SoE presents an infinite set of equilibrium points when off-treatment  $T_{p_S}(t) = 0$  in the intersection of the plane  $n_S(t) = 0$  with the plane given by  $c_{PSA}(t) = \frac{\gamma_{PSA} n_D(t)}{\delta_{PSA}}$  conditional to  $n_D \neq 0$  and  $\delta_{PSA} \neq 0$  and the generalized eigenvalues of the Jacobian results in a double-zero generalized eigenvalue  $\lambda_1 = 0, \lambda_2 = 0$  and a third negative eigenvalue  $\lambda_3 = -\delta_{PSA}$ . Standard center manifold computation (Wiggins 2003) shows slow-2D-manifold dynamics that can be integrated to prove that the equilibria are unstable, and therefore not of interest.

## 5 Bayesian Model Comparison

Maybe the most vital point of the Bayesian framework, and the reason for its increasing popularity, is its innate model comparison ability, based on logic as an instrument for selection. We exploit this feature here using the Bayesian factor to compare the different models in their ability to simulate the data. It should be noted that this framework innate penalizes models based on the number of parameters required. This

phenomenon is sometimes referred to as the Occam's razor factor (Jefferys and Berger 1992).

Starting from the classical Bayesian theorem, the Bayes factor  $\beta_{ij}$  for PSA model  $M_i$  over the PSA model  $M_j$  is computed as a ratio of the probabilities of the two models (the odd-ratio,  $O_{ij}$ )

$$O_{ij} = \frac{\Pr(M_i, I) \Pr(D|M_i, I)}{\Pr(M_j, I) \Pr(D|M_j, I)} = \frac{\Pr(M_i, I)}{\Pr(M_j, I)} \beta_{ij}, \quad (27)$$

such that, because  $\sum_{i=1}^{N_m} \Pr(M_i|D, I) = 1$  (with  $N_m$  number of models to compare) if we are interested in how a model, say  $M_1$ , compares to the other models  $M_j$ , we arrive at

$$\Pr(M_1|D, I) = \frac{O_{i1}}{\sum_{j=1}^{N_m} O_{j1}}. \quad (28)$$

We implement Eq. (28) to compare one patient at a time in one model against all the other models individually. For example, we are going to implement the comparison between  $M_1$ , and every other  $M_2$  as  $\Pr(M_2|D, I) = \frac{1}{1+O_{21}}$ , and we proceed iteratively.

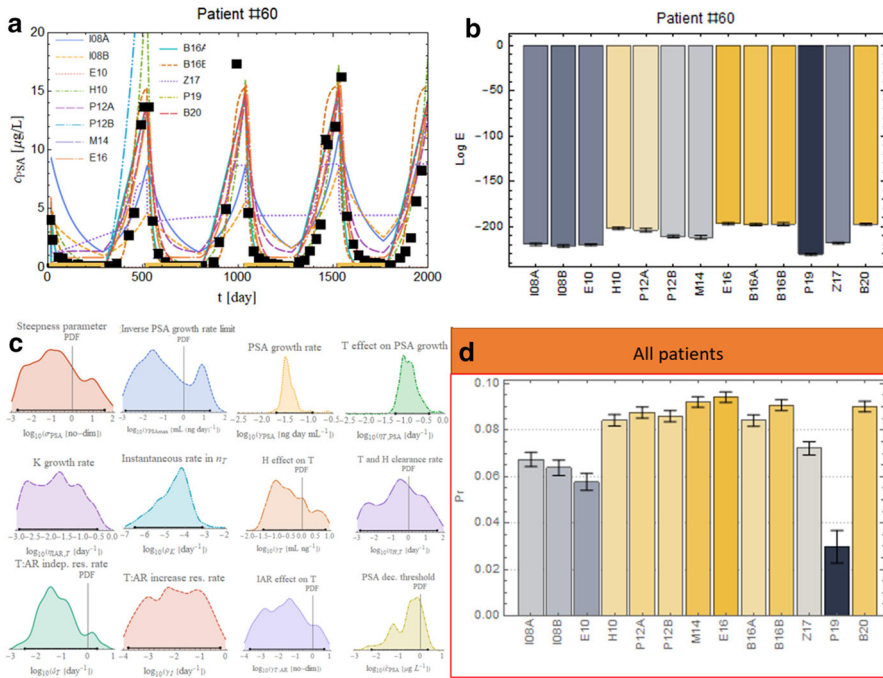
We first explore the Laplace approximation framework under the assumption of equally-prioritized models, i.e., assuming that no previous preference can be accorded to any of the PSA models considered. We can exploit the asymptotic approximation (Murphy 2012; Theodoridis 2015) to the global likelihood, i.e., the evidence of the  $i^{th}$  model,  $\Pr(D|M_i)$ , writing

$$\begin{aligned} \Pr(D|M_i) &= \int d\mathbf{p} \Pr(\mathbf{p}|M, I) L(\mathbf{p}|I) \\ &\cong \Pr(\hat{\mathbf{p}}|M_i) L(\hat{\mathbf{p}}) \sqrt{\det(F(\mathbf{p}))}, \end{aligned} \quad (29)$$

with  $F$  being the information matrix introduced in Sect. 3.1. A classical result of Bayesian analysis is to consider the limit of the previous expression, but for an increased number of data points ( $N_p \rightarrow \infty$ ) and flat priors, i.e., to compute the popular BIC index against AIC (Akaike 1974; Schwarz 1978). As the number of patient data points is often limited  $N_p \ll \infty$  and we make explicit use of priors, BIC or AIC indices are not justifiable for model comparison. Instead, we build up a model-of-model function (Pasetto et al. 2021) to encode prior information as soon as available. Furthermore, as introduced above, we verified the Laplace approximation with fully numerical integration based on nested sampling algorithms (Skilling 2004; Mukherjee et al. 2006; Feroz and Hobson 2008), i.e., a numerical technique designed explicitly to compute the global likelihood of models with different degrees of freedom.

## 5.1 Single Patient Comparison Results

Figure 7a shows an example of the quality of the model calibration achieved by Bayesian posterior inference introduced in Sect. 3 applied to the parameter inference



**Fig. 7** Bayesian model comparison results. **a** Best fits for the 13 models analyzed for a representative patient. The black squares are the error bars enhanced with a more prominent marker for visibility reasons; yellow lines along the x-axis represent the on-treatment periods. **b** Model log-evidence comparison on patient #60 with error bars as obtained by nest-sampling technique. The color range shows the best-performing model (yellow) and fades to the worst-performing model (gray). **c** Unnormalized posterior PDF for the best performing model, E16, and credible interval as black segment over the x-axis. **d** Comparison between the normalized log-evidence overall patient data. (The color scheme is consistent with panel b) (Color figure online)

problem to all the models. The T and H related disease dynamics vary significantly between the different models, and discrepancies between different models and patient data may indicate likely or unlikely biological mechanisms driving individual patients' resistance.

Model evidence (Fig. 7b) demonstrates that no single model represents all patient data accurately, suggesting that different biology drive individual patients' responses or that no model correctly faces the PSA problem. It may also imply that the PSA dynamics alone may be insufficient to discriminate between the different biological models. For some patients, model selection identifies models with a higher probability than others, but selection varies on a per-patient basis. As a classical proof-of-concept of the Bayesian technology employed, we report for the best performing model, E16, for patient #60 the unnormalized posterior marginalized PDF for each parameter in Fig. 7c. The PDFs are almost unimodal (but not for all parameters, see next Sect. 6), suggesting that this model represents fairly the patient and that the Laplace approximation could be justified. The credible intervals for the log parameters are also plotted and superimposed to the x-axis.

## 5.2 Overall Model Selection

We calculate the Bayesian maximum a posteriori performance for all the patients for each model (Fig. 7d), resulting in the Elishmereni et al. (2016) model marginally performing better on most patients. This result does not surprise us, as it is a model designed on clinical necessities, i.e., it was crafted with careful handling of the medical treatment. Nevertheless, as mentioned before, in the case of model comparison on a patient-to-patient basis, we could not identify a model that performed statistically better than the others thus eventually indicating the correct biological mechanics governing PSA dynamics. Figure 7d shows that E16 is preferred only on 10% of the patients, and eight of the 13 models have scores above the 8%.

## 6 Conclusion and Discussion

This work considers several mathematical models (Table 1) to simulate PSA dynamics of prostate cancer response to IADT in a prospective clinical trial. We exploit Bayesian continuous and discrete inference to interpret the data and identify the model with the highest likelihood of simulating the clinically observed dynamics. Using the PSA biomarker and the comparison between the different models, we 1) identify several models that can separate responding patients and patients that develop resistance to intermittent ADT through the model fitting, 2) performed the Bayesian model comparison and demonstrated that the model by Elishmereni et al. (2016) performed slightly better than the others, i.e., as a better representative of most patients in the trial. Nevertheless, as evidenced in the example of Fig. 7c, the marginalized posterior PDF is often not all optimally single-peaked, casting shadows in an attempt to use this model to solve forecast problems. While we have focused on the models' inference to evaluate the possible connection with their underpinned biology, we will explore the potentiality and limitation of the models' forecasting ability to predict clinical PSA trends in a follow-up paper (Pasetto et al. 2021, in preparation).

The models analyzed herein synonymously use longitudinal PSA data to infer biological mechanisms underlying the observed PSA dynamics. PSA alone limited the potentiality of the presented approach and did not identify a single dominant model. Further information is necessary to simulate accurately and ultimately predict patient-specific PSA trajectories and the corresponding biological drivers of resistance. PSA alone might not be a helpful biomarker due to several dominant environmental factors outside the models' scopes that influence its evolution under treatment. The use of PSA as a surrogate marker for prostate cancer burden is indeed controversial. Over-expression of the PCA3 gene obtained from the mRNA in urine samples is proposed to be more suited to monitoring the cancer evolution (Bussemakers et al. 1999, p. 3; Laxman et al. 2008; Neves et al. 2008; Hessels and Schalken 2009, p. 3; Borros 2009).

Two alternative directions might to improve our understanding of the PSA as a prostate cancer monitor biomarker. From one side, a deeper understanding of the connection between PSA and tumor burden throughout model investigation might present the opportunity for a new class of models. Recently, the role of immature blood vessels formed under angiogenesis cues has been investigated to decrypt the

**Table 1** Model compartment sketches with phase variables and parameters modeled in the inference process

Model	Reference paper and critical additions	Compartment draw	
I08A	Ideta et al. (2008): the AI-AD population paradigm with mutation rate	<p>The diagram for model I08A shows a network of compartments. At the top is Prostate-specific antigen (c<sub>PSA</sub>), represented by a grey cluster. Below it are Androgen dependent (n<sub>D</sub>, pink) and Androgen independent (n<sub>I</sub>, orange) cells. To the left is Androgen (c<sub>A</sub>, yellow). Arrows indicate transitions: c<sub>A</sub> to n<sub>D</sub> (rate γ<sub>D</sub>), n<sub>D</sub> to c<sub>PSA</sub> (rate μ<sub>D</sub>), n<sub>D</sub> to n<sub>I</sub> (rate μ<sub>DI</sub>), and n<sub>I</sub> to c<sub>PSA</sub> (rate γ<sub>I</sub>). There are also self-loops on n<sub>D</sub> and n<sub>I</sub>. Inhibitory arrows (red) point from n<sub>D</sub> to c<sub>A</sub> and from n<sub>I</sub> to c<sub>A</sub>. Parameters are indicated by δ symbols: δ<sub>D</sub> for n<sub>D</sub>, δ<sub>I</sub> for n<sub>I</sub>, δ<sub>A</sub> for c<sub>A</sub>, and δ<sub>PSA</sub> for c<sub>PSA</sub>.</p>	<p>Phase variable <math>n_D, n_I, c_A</math></p> <p>Parameters i.c. + <math>\gamma_{Dmax}, k_{DY/2}, \gamma_{DA}, \delta_{Dmax}, k_{D\delta/2}, \delta_{DA}, \mu_{DImax}, \delta_{cA}</math></p>
I08B	Ideta et al. (2008): AI cell proliferation is negatively regulated by androgens	<p>The diagram for model I08B is identical to I08A but includes an additional red inhibitory arrow from n<sub>I</sub> to n<sub>D</sub>, representing negative regulation of androgen-dependent cells by androgen-independent cells.</p>	<p>Phase variables <math>n_D, n_I, c_A</math></p> <p>Parameters i.c. <math>\cup \{I08A\} \cup \{\gamma_{IA}, \delta_{IA}\}</math></p>
E10	Eikenberry et al. (2010): interaction between testosterone, dihydrotestosterone, and their bindings	<p>The diagram for model E10 shows Testosterone (n<sub>T</sub>, blue), Dihydrotestosterone (n<sub>DHT</sub>, dark blue), Androgen receptor (n<sub>R</sub>, yellow), and Testosterone:AR Complex (q<sub>T:R</sub>, green). Testosterone (n<sub>T</sub>) can be converted to Dihydrotestosterone (n<sub>DHT</sub>) (rate μ<sub>cat</sub>) or bind to the Androgen receptor (n<sub>R</sub>) (rate γ<sub>R</sub>) to form the Testosterone:AR Complex (q<sub>T:R</sub>). Dihydrotestosterone (n<sub>DHT</sub>) can also bind to the Androgen receptor (n<sub>R</sub>) (rate γ<sub>R</sub>) to form the Testosterone:AR Complex (q<sub>T:R</sub>). The complex can then bind to the Androgen receptor (n<sub>R</sub>) (rate δ<sub>T:R</sub>) or be converted to Dihydrotestosterone (n<sub>DHT</sub>) (rate δ<sub>DHT:R</sub>). There are also self-loops on n<sub>T</sub> and n<sub>DHT</sub>. Inhibitory arrows (red) point from n<sub>T</sub> to n<sub>R</sub> and from n<sub>DHT</sub> to n<sub>R</sub>. Parameters are indicated by δ symbols: δ<sub>T</sub> for n<sub>T</sub>, δ<sub>DHT</sub> for n<sub>DHT</sub>, δ<sub>T:R</sub> for q<sub>T:R</sub>, and δ<sub>DHT:R</sub> for q<sub>T:R</sub>.</p>	<p>Phase variables <math>n_T, n_R, n_{DHT}, n_{T:R}, n_{DHT:R}</math></p> <p>Parameters i.c. <math>\cup \{n_S, \delta_R, \kappa_{T:R}, \delta_{T:R}, \kappa_{DHT:R}, \delta_{DHT:R}, n_{5\alpha}, \mu_{cat}, \kappa_M, \delta_{DHT}, \delta_T\}</math></p>
H10	Hirata et al. (2010): AD-AI population		<p>Phase variables <math>n_D, n_I, n_{Ir}</math></p>

relation between an increased tumor burden contemporaneous to decreasing PSA concentration (Barnaby et al. 2021). Additionally, models that include both PSA and androgen concentrations might present some advantages in the future. The modest but significant evidence of the E16 model over the other models might indicate a

**Table 1** (continued)

	<p>cells, supplemented with a population of irreversible AI cells</p>		<p>Parameters i.c. <math>\cup\{\gamma_D^{on}, \gamma_D^{off}, \mu_{DI}, \mu_{DI}, \gamma_I^{on}, \gamma_I^{off}, \mu_{DIrr}, \mu_{Irr}, \gamma_{Irr}^{on}, \gamma_{Irr}^{off}\}</math></p>
<p>P12A</p>	<p>Portz et al. (2012): the concept of quota, <math>q</math>, common to <math>n_D</math> and <math>n_I</math> population.</p>		<p>Phase variables <math>n_D, n_I, q</math></p> <p>Parameters i.c. <math>\cup\{k_{ID}/2, \mu_{IDmax}, \mu_{DImax}, k_{DI}/2, \gamma_{Dmax}, q_{min}, \delta_D, \gamma_I, v_{max}, q_{max}, k_{q}/2, \delta_q\}</math></p>
<p>P12B</p>	<p>Portz et al. (2012): separate quota levels for each population.</p>		<p>Phase variables <math>n_D, n_I, q_D, q_I, c_{PSA}</math></p> <p>Parameters i.c. <math>\cup\{\alpha \cup\{k_{ID}/, k_{PSAD}/2, k_{PSAI}/2, k_{q}/2, v_{max}, \gamma_{PSAD}, \gamma_{PSAI}, \delta_q, \mu_{IDmax}, \gamma\}</math></p>
<p>M14</p>	<p>Morken et al. (2014): introduce ADT-</p>		<p>Phase variables <math>n_D, n_I, q_D, q_I, c_{PSA}</math></p>

more critical relevance of the dormancy whose biology and mathematics are worth undoubtedly deeper understanding.

Exploring PSA model probability distributions to disentangle responsive and resistant patient cohorts in a clinical setting could be investigated through cross-correlations with PCA3 biomarkers. Such cross-correlation would provide independent verification of the analytical findings herein that remain, for the moment, data-driven and, therefore, entirely dependent on the one dataset utilized in all discussed models.

Alternatively, PSA could be a perfect biomarker, but inter-patient heterogeneity in resistance mechanisms may disallow identifying a single model for all patients. Additionally, different resistance mechanisms may evolve in an individual patient, with their

**Table 1** (continued)

	<p>induced apoptosis of prostate cancer cells in addition to inhibition of their growth and proliferation.</p>		<p>Parameters i.c. <math>\cup P12B \cup \{\delta_D, \delta_I, \delta_{Dmax}, \delta_{Imax}, k_{D\delta/2}, k_{I\delta/2}\}</math></p>
<p>B16A</p>	<p>Baez et al. (2016) is a single population model</p>		<p>Phase variables <math>n, \delta_{max}, c_{PSA}</math></p> <p>Parameters i.c. <math>\cup \{\gamma_1, \gamma_2, \gamma_{max}, q_{min}, k_{n/2}, \delta, \tau, \delta_{max}, q_{max}, \gamma_{PSA0}, \gamma_{PSA1}, \delta_{PSA}\}</math></p>
<p>B16B</p>	<p>Baez et al. (2016), pair the PSA with testosterone dynamics</p>		<p>Phase variables <math>n_D, n_I, q_D, q_I, c_{PSA}</math></p> <p>Parameters i.c. <math>\cup \alpha \cup \{k_{D/2}, k_{I/2}, \gamma_1, \gamma_2, \gamma_{PSA1}, \delta_{Dm}, \delta_{Imax}\}</math></p>
<p>E16</p>	<p>Elishmereni et al. (2016) introduce PCa dormancy</p>		<p>Phase variables <math>c_{PSA}, K, n_T, T, H, R_{T,AR}, R_I</math></p> <p>Parameters i.c. <math>\cup \alpha \cup \{k_{D/2}, k_{I/2}, \gamma_1, \gamma_2, \gamma_{PSA1}, \delta_{Dmax}, \delta_{Imax}\}</math></p>
<p>Z17</p>	<p>Zhang et al. (2017): game</p>		<p>Phase variables <math>n_D, n_P, n_I</math></p>

respective contribution to the observed response dynamics changing during therapy. More complex models and dynamic adaptive weighting of different variables, terms, and parameters may be necessary. Such models, however, would be non-identifiable with the presently available data. A close dialogue between biologists, statisticians, and mathematical and genitourinary oncologists may help identify which data should be collected in future clinical studies to help detangle the complex prostate cancer response dynamics to intermittent ADT.

**Table 1** (continued)

	<p>theory-based approach</p>	<p>The diagram shows three cell populations: Androgen dependent (<math>n_D</math>), Testosterone producing (<math>n_P</math>), and Androgen independent (<math>n_I</math>). They interact with each other and with Prostate-specific antigen (<math>c_{PSA}</math>). Self-loops are shown for <math>n_D</math> and <math>n_I</math>. Arrows indicate interactions with rates <math>\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}, \alpha_{23}, \alpha_{32}, \alpha_{33}</math>. Production rates are <math>\gamma_D</math> and <math>\gamma_I</math>. The antigen <math>c_{PSA}</math> is produced at rate <math>\delta</math>.</p>	<p>Parameters i.c. <math>\cup \{\gamma_D, \gamma_P, \gamma_I, K_P, K_I, \alpha_{12}, \alpha_{13}, \alpha_{21}, \alpha_{23}, \alpha_{31}, \alpha_{32}, \alpha, \delta\}</math></p>
<p>P19</p>	<p>Phan et al. (2019): three population model with weakly interacting population</p>	<p>The diagram shows four populations: Androgen dependent (<math>n_D</math>), Weakly Androgen dependent (<math>n_{WD}</math>), Androgen independent (<math>n_I</math>), and Androgen quota (<math>q</math>). It includes self-loops and interactions with rates <math>\gamma_{P143}, \delta_{Dmax}, \mu_{Dmax}, \mu_{IDmax}, \delta_I, \delta_{Imin}, \delta_{WDmax}, \delta_{WDmin}, \delta_{PSA}, \gamma_{PSA}, \gamma_{Dmax}, \gamma_{Imin}, \gamma_{WDmax}, \gamma_{WDmin}, \gamma_{qmax}, \gamma_{qmin}, \gamma_{Dmin}, \gamma_{Imax}, \gamma_{WDmin}, \gamma_{WDmax}</math>. The Androgen quota <math>q</math> promotes or inhibits other populations.</p>	<p>Phase variables <math>n_D, n_{WD}, n_I, q_D, q_{WD}, q_I, c_{PSA}</math></p> <p>Parameters i.c. <math>\cup \{UB16B \cup \{k_{WD}/2, q_{WDmin}, \delta_{WD}, \delta_{WDmax}\}\}</math></p>
<p>B20</p>	<p>Brady-Nicholls et al. (2020): stem cell model</p>	<p>The diagram shows PCa stem cells (<math>n_S</math>), Differentiate d cells (<math>n_D</math>), and Prostate-specific antigen (<math>c_{PSA}</math>). Stem cells differentiate into d cells at rate <math>\log(2)</math>. D cells produce PSA at rate <math>\gamma_{PSA}</math>. Self-loops are shown for <math>n_S</math> and <math>n_D</math>. Death rates are <math>\delta_D</math> and <math>\delta_{PSA}</math>.</p>	<p>Phase variables <math>n_S, n_D, c_{PSA}</math></p> <p>Parameters i.c. <math>\cup \{p_S, \gamma_{PSA}, \delta_{PSA}, \delta_D\}</math></p>

Initial condition (i.c.) are not reported, but phase variables are.  $\alpha = \{k_{D/2}, k_{ID/2}, k_{q/2}, q_{max}, v_{max}, \delta_D, \delta_q, \mu_{Dmax}, \mu_{IDmax}\}$ . Symbols used for the compartments are self-explanatory

While the Bayesian framework is an invaluable tool to estimate model parameters and fit model dynamics to clinical measurements, the goodness of a fit informs neither the reliability of the estimated parameters nor the likelihood of a model representing the data chosen for the valid biological reason. Relatively invariant PSA profiles can be obtained for a significant range in each parameter, as it is the case of a weakly sensitive—highly non-identifiable parameter. This fact is often omitted in the modeling literature, where the results are often presented without structural or practical identifiability analysis. Many of the herein discussed models have not demonstrated structural identifiability, hence jeopardizing the attempt to claim the inference’s practical identifiability herein. Nevertheless, we stress that a model’s value may also be found in its interpretative role (Enderling and Wolkenhauer 2021). The complexity of the mechanism involved in the biological responses to intermittent ADT can be captured correctly for a single patient but fail for others. *Therefore, the model comparison is not intended to provide an absolute ranking; instead, it provides an instrument to*



explore the different biological mechanisms implemented in mathematical models in clinically observed treatment response and progression dynamics.

## 7 Supplement

We have performed a sensitivity analysis for all the models included in the paper. However, as this analysis overlaps with the original papers' work, we do not include those results here. Our sensitivity is motivated by: (1) to understand the dependence of our results on the parameters. For example, if we can claim the possibility to split between relapsing ( $\Omega$ ) and not to relapse ( $\neg\Omega$ ) patients by exploiting some specific model parameter combination, then the robustness of our result worth be investigated on the same parameter sensitivity to assign it the correct relevance and to evaluate its possibility to be applied to clinical tumor forecasting. (2) The technique implemented for the sensitivity analysis investigates the parameter's sensitivity and the best-fit orbital integration, i.e., over the available longitudinal data. This approach enhances our understanding of when a particular  $\Omega/\neg\Omega$  segregating technique is more useful during or off-treatment, with consequent indications on the role that a model splitting potentiality might or might not have (and when) on a per-patient base. (3) Continuous but not differentiable functions might need particular attention in the computation of the sensitivities because of their definition as the Jacobian matrix's function. This approach represents a current research field often omitted in the mathematical oncology literature and is worth being brought to light.

Therefore, in what follows, we exploit the direct differential method (DDM) for sensitivity analysis (Gu and Wang 2013) to track the time dependence of the sensitivity  $S_{ij} = \frac{\partial x_i(t, \hat{\mathbf{p}})}{\partial p_j}$ , where in general it is  $x_i = c_{\text{PSA}}$  and  $p_j$  the generic parameter dependent on the particular model in the exam. For a generic vector field  $\frac{\partial \mathbf{x}(\mathbf{p}; t)}{\partial t} = f(\mathbf{x}, \mathbf{p})$  with  $x(t_0, \mathbf{p}) = \mathbf{x}_0$  we couple the integration of the SoE defining the model with:

$$\frac{\partial S_{ij}(t, \hat{\mathbf{p}})}{\partial t} = \frac{\partial f_i(x_k(t, \hat{\mathbf{p}}), \hat{\mathbf{p}})}{\partial x_k(t, \hat{\mathbf{p}})} S_{kj}(t, \hat{\mathbf{p}}) + \frac{\partial f_i}{\partial p_j}. \quad (30)$$

Generalized sensitivity (Stechlinski et al. 2018), based on the concept of generalized derivative for non-smooth  $c_{\text{PSA}}$  profiles (Clarke 1990) and used because of the loss of differentiability at the bifurcation points  $T_{ps} = \{0, 1\}$  on the treatment parameter, has also been considered. We will not report the DDM analysis if not relevant to strength our specific results and we refer to the original model paper for general sensitivity analysis of the presented models.

**Acknowledgements** Dr. S. Pasetto thanks Prof. D. Crnojevic, Prof. S. Ross, and Dr. S. Smit for stimulating discussions, technical support, or careful reading of this manuscript's early version. The research reported in this publication was supported by the Ocala Royal Dames for Cancer Research, Inc., The JAYNE KOSKINAS TED GIOVANIS FOUNDATION FOR HEALTH AND POLICY, a Maryland private foundation dedicated to effecting change in health care for the public good, and the National Cancer Institute of the National Institutes of Health under Award Numbers R21CA234787 and U54CA143970. The content is solely the authors' responsibility and does not necessarily represent the official views of the National Institutes of Health. The opinions, findings, and conclusions or recommendations expressed in this material are

those of the authors and not necessarily those of the JAYNE KOSKINAS TED GIOVANIS FOUNDATION FOR HEALTH AND POLICY, its directors, officers, or staff.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Autom Control* 19:716–723. <https://doi.org/10.1109/TAC.1974.1100705>
- Baez J, Kuang Y (2016) Mathematical models of androgen resistance in prostate cancer patients under intermittent androgen suppression therapy. *Appl Sci* 6:352. <https://doi.org/10.3390/app6110352>
- Barnaby JP, Sorribes IC, Jain HV (2021) Relating prostate-specific antigen leakage with vascular tumor growth in a mathematical model of prostate cancer response to androgen deprivation. *Comput Syst Oncol* 1:e1014. <https://doi.org/10.1002/cso2.1014>
- Borros A (2009) Clinical significance of measuring prostate-specific antigen. *Lab Med* 40:487–491. <https://doi.org/10.1309/LMEGGGLZ2EDWRXUK>
- Brady-Nicholls R, Nagy JD, Gerke TA et al (2020) Prostate-specific antigen dynamics predict individual responses to intermittent androgen deprivation. *Nat Commun* 11:1750. <https://doi.org/10.1038/s41467-020-15424-4>
- Bruchovsky N, Klotz L, Crook J et al (2006) Final results of the Canadian prospective phase II trial of intermittent androgen suppression for men in biochemical recurrence after radiotherapy for locally advanced prostate cancer. *Cancer* 107:389–395. <https://doi.org/10.1002/cncr.21989>
- Bruchovsky N, Klotz L, Crook J, Phillips N, Abersbach J, Goldenberg SL (2008) Quality of life, morbidity, and mortality results of a prospective phase II study of intermittent androgen suppression for men with evidence of prostate-specific antigen relapse after radiation therapy for locally advanced prostate cancer. *Clin Genitourin Cancer* 6(1):46–52. <https://doi.org/10.3816/CGC.2008.n.008>
- Bussemakers MJ, van Bokhoven A, Verhaegh GW et al (1999) DD3: a new prostate-specific gene, highly overexpressed in prostate cancer. *Cancer Res* 59:5975–5979
- Clarke FH (1990) Optimization and nonsmooth analysis. Society for Industrial and Applied Mathematics, Philadelphia
- Droop MR (1968) Vitamin B12 and marine ecology. IV. The kinetics of uptake, growth and inhibition in *Monochrysis Lutheri*. *J Mar Biol Assoc UK* 48:689–733. <https://doi.org/10.1017/S0025315400019238>
- Eikensberry SE, Nagy JD, Kuang Y (2010) The evolutionary impact of androgen levels on prostate cancer in a multi-scale mathematical model. *Biol Direct* 5:24. <https://doi.org/10.1186/1745-6150-5-24>
- Elishmereni M, Kheifetz Y, Shukrun I et al (2016) Predicting time to castration resistance in hormone sensitive prostate cancer by a personalization algorithm based on a mechanistic model integrating patient data. *Prostate* 76:48–57. <https://doi.org/10.1002/pros.23099>
- Elzanaty S, Rezanezhad B, Dohle G (2017) Association between serum testosterone and PSA levels in middle-aged healthy men from the general population. *Curr Urol* 10:40–44. <https://doi.org/10.1159/000447149>
- Enderling H (2015) Cancer stem cells: small subpopulation or evolving fraction? *Int Bio (cam)* 7:14–23. <https://doi.org/10.1039/c4ib00191e>
- Enderling H, Wolkenhauer O (2021) Are all models wrong? *Comput Syst Oncol* 1:e1008. <https://doi.org/10.1002/cso2.1008>
- Everett RA, Packer AM, Kuang Y (2014) Can mathematical models predict the outcomes of prostate cancer patients undergoing intermittent androgen deprivation therapy? *Biophys Rev Lett* 09:173–191. <https://doi.org/10.1142/S1793048014300023>

- Feldman BJ, Feldman D (2001) The development of androgen-independent prostate cancer. *Nat Rev Cancer* 1:34–45. <https://doi.org/10.1038/35094009>
- Feoktistov V (2006) *Differential evolution: in search of solutions*. Springer, New York
- Feroz F, Hobson MP (2008) Multimodal nested sampling: an efficient and robust alternative to Markov Chain Monte Carlo methods for astronomical data analyses. *Month Notices R Astronom Soc* 384:449–463. <https://doi.org/10.1111/j.1365-2966.2007.12353.x>
- Goode SW, Annin SA (2015) *Differential equations and linear algebra*. Prentice Hall, Upper Saddle River
- Grossmann ME, Huang H, Tindall DJ (2001) Androgen receptor signaling in androgen-refractory prostate cancer. *J Natl Cancer Inst* 93:1687–1697. <https://doi.org/10.1093/jnci/93.22.1687>
- Gu Q, Wang G (2013) Direct differentiation method for response sensitivity analysis of a bounding surface plasticity soil model. *Soil Dyn Earthq Eng* 49:135–145. <https://doi.org/10.1016/j.soildyn.2013.01.028>
- Hessels D, Schalken JA (2009) The use of PCA3 in the diagnosis of prostate cancer. *Nat Rev Urol* 6:255–261. <https://doi.org/10.1038/nrur.2009.40>
- Hirata Y, Aihara K (2015) Ability of intermittent androgen suppression to selectively create a non-trivial periodic orbit for a type of prostate cancer patients. *J Theor Biol* 384:147–152. <https://doi.org/10.1016/j.jtbi.2015.08.010>
- Hirata Y, Bruchovsky N, Aihara K (2010) Development of a mathematical model that predicts the outcome of hormone therapy for prostate cancer. *J Theor Biol* 264:517–527. <https://doi.org/10.1016/j.jtbi.2010.02.027>
- Hirata Y, Tanaka G, Bruchovsky N, Aihara K (2012) Mathematically modelling and controlling prostate cancer under intermittent hormone therapy. *Asian J Androl* 14:270–277. <https://doi.org/10.1038/aja.2011.155>
- Hutter F, Hoos HH, Leyton-Brown K (2011) Sequential model-based optimization for general algorithm configuration. In: Coello CAC (ed) *Learning and intelligent optimization*. Springer, Berlin, pp 507–523
- Ideta AM, Tanaka G, Takeuchi T, Aihara K (2008) A mathematical model of intermittent androgen suppression for prostate cancer. *J Nonlinear Sci* 18:593–614. <https://doi.org/10.1007/s00332-008-9031-0>
- Jackson TL (2004) A mathematical investigation of the multiple pathways to recurrent prostate cancer: comparison with experimental data. *Neoplasia* 6:697–704. <https://doi.org/10.1593/neo.04259>
- Jefferys WH, Berger JO (1992) Ockham's razor and Bayesian analysis. *Am Sci* 80:64–72
- Jefferys H (1946) An invariant form for the prior probability in estimation problems. *Proc R Soc Lond Ser A Math Phys Sci* 186:453–461. <https://doi.org/10.1098/rspa.1946.0056>
- Klotz LH, Herr HW, Morse MJ, Whitmore WF (1986) Intermittent endocrine therapy for advanced prostate cancer. *Cancer* 58:2546–2550. [https://doi.org/10.1002/1097-0142\(19861201\)58:11%3c2546::AID-CNCR2820581131%3e3.0.CO;2-N](https://doi.org/10.1002/1097-0142(19861201)58:11%3c2546::AID-CNCR2820581131%3e3.0.CO;2-N)
- Larry Goldenberg S, Bruchovsky N, Gleave ME et al (1995) Intermittent androgen suppression in the treatment of prostate cancer: a preliminary report. *Urology* 45:839–845. [https://doi.org/10.1016/S0090-4295\(99\)80092-2](https://doi.org/10.1016/S0090-4295(99)80092-2)
- Laxman B, Morris DS, Yu J et al (2008) A first-generation multiplex biomarker analysis of urine for the early detection of prostate cancer. *Cancer Res* 68:645–649. <https://doi.org/10.1158/0008-5472.CAN-07-3224>
- Lin K, Lipsitz R, Miller T, Janakiraman S (2008) Benefits and harms of prostate-specific antigen screening for prostate cancer: an evidence update for the U.S. preventive services task force. *Ann Intern Med* 149:192. <https://doi.org/10.7326/0003-4819-149-3-200808050-00009>
- Morgentaler A, Connors W (2015) Testosterone therapy in men with prostate cancer: literature review, clinical experience, and recommendations. <http://www.ajandrology.com/article.asp?issn=1008-682X;year=2015;volume=17;issue=2;spage=206;epage=211;aulast=Morgentaler>. Accessed 8 Jun 2020
- Morken JD, Packer A, Everett RA et al (2014) Mechanisms of resistance to intermittent androgen deprivation in patients with prostate cancer identified by a novel computational method. *Cancer Res* 74:3673–3683. <https://doi.org/10.1158/0008-5472.CAN-13-3162>
- Mukherjee P, Parkinson D, Liddle AR (2006) A nested sampling algorithm for cosmological model selection. *ApJ* 638:L51. <https://doi.org/10.1086/501068>
- Murphy KP (2012) *Machine learning: a probabilistic perspective*. The MIT Press, Cambridge
- Neves AF, Araújo TG, Biase WKFS et al (2008) Combined analysis of multiple mRNA markers by RT-PCR assay for prostate cancer diagnosis. *Clin Biochem* 41:1191–1198. <https://doi.org/10.1016/j.clinbiochem.2008.06.013>
- Packer A, Li Y, Andersen T et al (2011) Growth and neutral lipid synthesis in green microalgae: a mathematical model. *Bioresour Technol* 102:111–117. <https://doi.org/10.1016/j.biortech.2010.06.029>

- Phan T, Crook SM, Bryce AH et al (2020) Review: mathematical modeling of prostate cancer and clinical application. *Appl Sci* 10:2721. <https://doi.org/10.3390/app10082721>
- Phan T, He C, Martinez A et al (2019) Dynamics and implications of models for intermittent androgen suppression therapy. *MBE* 16:187–204. <https://doi.org/10.3934/mbe.2019010>
- Portz T, Kuang Y, Nagy JD (2012) A clinical data validated mathematical model of prostate cancer growth under intermittent androgen suppression therapy. *AIP Adv* 2:011002. <https://doi.org/10.1063/1.3697848>
- Qin Z, Yao J, Xu L et al (2020) Diagnosis accuracy of PCA3 level in patients with prostate cancer: a systematic review with meta-analysis. *Int Braz J Urol* 46:691–704. <https://doi.org/10.1590/S1677-5538.IBJU.2019.0360>
- Schwarz G (1978) Estimating the dimension of a model. *Ann Stat* 6:461–464. <https://doi.org/10.1214/aos/1176344136>
- Siegel et al. (2021) Cancer statistics, 2021. <https://acsjournals.onlinelibrary.wiley.com/doi/https://doi.org/10.3322/caac.21654>. Accessed 26 Jan 2021
- Skilling J (2004) Nested Sampling *AIP Conf Proc* 735:395–405. <https://doi.org/10.1063/1.1835238>
- Stechlinski P, Khan KA, Barton PI (2018) Generalized sensitivity analysis of nonlinear programs. *SIAM J Optim* 28:272–301. <https://doi.org/10.1137/17M1120385>
- Tanaka G, Hirata Y, Goldenberg SL et al (2010) Mathematical modelling of prostate cancer growth and its application to hormone therapy. *Proc R Soc A* 368:5029–5044. <https://doi.org/10.1098/rsta.2010.0221>
- Theodoridis S (2015) *Machine learning: a bayesian and optimization perspective*, 1st edn. Academic Press Inc, Orlando
- Wiggins S (2003) *Introduction to applied nonlinear dynamical systems and chaos*, 2nd edn. Springer, New York
- Zhang J, Cunningham JJ, Brown JS, Gatenby RA (2017) Integrating evolutionary dynamics into treatment of metastatic castrate-resistant prostate cancer. *Nat Commun* 8:1–9. <https://doi.org/10.1038/s41467-017-01968-5>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.