



HAL
open science

Optimizing IFN Alpha Therapy against Myeloproliferative Neoplasms

Gurvan Hermange, Paul-Henry Cournède, Isabelle Plo

► **To cite this version:**

Gurvan Hermange, Paul-Henry Cournède, Isabelle Plo. Optimizing IFN Alpha Therapy against Myeloproliferative Neoplasms. *Journal of Pharmacology and Experimental Therapeutics*, 2023, 387 (1), pp.31-43. 10.1124/jpet.122.001561 . hal-04308358

HAL Id: hal-04308358

<https://centralesupelec.hal.science/hal-04308358v1>

Submitted on 5 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optimizing IFN alpha therapy against Myeloproliferative Neoplasms

Gurvan Hermange¹, Paul-Henry Cournède^{1*}, and Isabelle Plo^{2, 3, 4*}

¹Université Paris-Saclay, CentraleSupélec, Laboratory of Mathematics and Informatics (MICS), Gif-sur-Yvette, France.

²INSERM U1287 (INSERM, Gustave Roussy, Université Paris-Saclay), Villejuif, France

³Gustave Roussy, Villejuif, France

⁴Université Paris-Saclay, Villejuif, France

* corresponding authors: paul-henry.cournede@centralesupelec.fr ; isabelle.plo@gustaveroussy.fr

Abstract

Myeloproliferative Neoplasms (MPNs) are hematological malignancies that result from acquired driver mutations in hematopoietic stem cells (HSCs), causing overproduction of blood cells and an increased risk of thrombo-hemorrhagic events. The most common MPN driver mutation affects the *JAK2* gene (*JAK2*^{V617F}). Interferon alpha (IFN α) is a promising treatment against MPNs by inducing a hematological response and molecular remission for some patients. Mathematical models have been proposed to describe how IFN α targets mutated HSCs, indicating that a minimal dose is necessary for long-term remission. This study aims to determine a personalized treatment strategy. First, we show the capacity of an existing model to predict cell dynamics for new patients from data that can be easily obtained in clinic. Then, we study different treatment scenarios *in silico* for three patients, considering potential IFN α dose-toxicity relations. We assess when the treatment should be interrupted, depending on the response, the patient's age, and the inferred development of the malignant clone without IFN α . We find that an optimal strategy would be to treat the patients with a constant dose so that the treatment could be interrupted as fast as possible. Higher doses result in earlier discontinuation but also higher toxicity. Without knowledge of the dose-toxicity relationship, trade-off strategies can be found for each patient. A compromise strategy is to treat patients with medium doses (60-120 $\mu\text{g}/\text{week}$) for 10-15 years. Altogether, this work demonstrates how a mathematical model calibrated from real data can help build a clinical decision-support tool to optimize long-term IFN α therapy for MPN patients.

Significance statement

Myeloproliferative Neoplasms (MPNs) are chronic blood cancers. Interferon alpha (IFN α) is a promising treatment with the potential to induce a molecular response by targeting mutated hematopoietic stem cells. MPN patients are treated over several years, and there is a lack of knowledge concerning the posology strategy and the best timing for interrupting therapy. The study opens avenues for rationalizing how to treat MPN patients with IFN α over several years, promoting a more personalized approach to treatment.

1 Introduction

BCR-ABL negative myeloproliferative neoplasms (MPNs) are hematological malignancies that include essential thrombocythemia, polycythemia vera and primary myelofibrosis leading to overproduction of myeloid cells (red cells, platelets and granulocytes). These diseases are clonal due to the expansion over decades of a hematopoietic stem cell (HSC) that has undergone the acquisition of a genetic abnormality [Williams et al., 2022, Van Egeren et al., 2021, Hermange et al., 2022]. Following homologous recombination, homozygous malignant clones can develop in parallel to heterozygous clones. The mutations affect *JAK2*, calreticulin (*CALR*), and the thrombopoietin receptor (*MPL*) genes, with *JAK2*^{V617F} being the most prevalent in more than 95% of polycythemia vera and around half of essential thrombocythemia and primary myelofibrosis. All these mutations are gain-of-function leading to constitutive activation of the JAK2/STAT pathway [Vainchenker et al., 2011].

The thrombo-hemorrhagic complications are the main comorbidities even if the transformation into secondary leukemia of dismal prognosis occurs in a significant fraction of cases, between 5% to 20%. The clinicians generally treat MPN patients to normalize the blood parameters and/or to improve symptoms, mainly using aspirin, cytoreductive treatment (hydroxyurea) and JAK1/2 inhibitors (ruxolitinib and inrebic). However, neither of these treatments impacts the malignant clone selectively, and the only curative option is bone marrow allogeneic transplantation in the most severe myelofibrotic patients. In this context, Peg-IFN α 2a (IFN α) is a promising treatment. When used, IFN α is mainly a second-line therapy. It harbors high rates of hematological responses in *JAK2*^{V617F} MPN patients and some molecular responses, including high molecular responses (HMR), even after treatment arrest [Mascarenhas et al., 2022, Gisslinger et al., 2020, Kiladjian et al., 2008]. In clinic, the molecular response refers to measuring the variant allele frequency (VAF) in mature cells and quantifying how the VAF decreases over the therapy. Despite these positive effects, some *JAK2*^{V617F} MPN patients do not respond to IFN α , and HMR is only observed in $\approx 20\%$ of *JAK2*^{V617F} patients. Long-term treatment of 2-5 years is required to obtain HMR. Additionally, toxicity increases with the IFN α dose [Yamane et al., 2008], resulting in the choice of dose de-escalation in many patients after one year, potentially jeopardizing the success of long-term HMR [Mosca et al., 2021]; all of which result in decision-making challenges for clinicians. Therefore, this therapy could be improved and optimized to increase the molecular response while limiting toxicity.

Previously, it has been shown that the IFN α targets the *JAK2*^{V617F} HSCs in a preclinical mouse model [Hasan et al., 2013, Austin et al., 2020, Mullally et al., 2013]. Due to the difficult access to human HSCs by biological methods, the effect of IFN α on the mutated human HSC dynamics was investigated using a tailor-made compartmental mathematical model calibrated with data from a prospective, observational, and longitudinal cohort of treated patients. Our objective with this model was to quantify not only the molecular response among mature cells, but also the response at the stem cell level. We inferred that the IFN α slowly targets human *JAK2*^{V617F} HSCs, and the higher the dose, the better the response [Mosca et al., 2021]. The latter mathematical model was then extended to account for the dose variations during the treatment when the mutated HSCs were targeted [Hermange et al., 2021]. This model is used in this present work to develop a clinical decision-support tool for predicting the response of MPN patients to IFN α and optimizing their long-term therapy. First, we verify that we can get accurate predictions from the model, clinical observations from a new patient, and prior knowledge from a cohort of patients. Then, we explore several therapeutical strategies to find which one would minimize the IFN α toxicity until the end of the treatment. Finally, combining the previous mathematical model with another one describing how *JAK2*^{V617F} mutated HSCs - when not treated - expand over time [Hermange et al., 2022], we assess when the therapy could be interrupted depending on how the patient responds to IFN α , their age, and the proportion of homozygous and heterozygous *JAK2*^{V617F} HSCs (that is, the zygosity).

2 Materials and Methods

2.1 Experimental and clinical data

The experimental and clinical data used in this study are those of 19 $JAK2^{V617F}$ patients from [Mosca et al., 2021]. For each patient of this cohort, we have:

- Their age at the beginning of the treatment.
- Measurements of the VAF in mature cells (in granulocytes) at different time points during the treatment. We define them as clinical data since these observations are molecular biology data obtained during routine clinical exams.
- Observations of the clonal architecture (that is, measurement of the heterozygous and homozygous Clonal Fractions - CF) at different time points during the treatment. We define them as experimental data since they are obtained in biology research laboratories.
- The variations of the dose they received over time.

We display the experimental and clinical data of patient #32 and those of each patient (Fig. 1 and Supp. Fig. A.1-19). We denote by \mathcal{D} the dataset for all 19 patients (see Supplemental Methods A: Data and observation model).

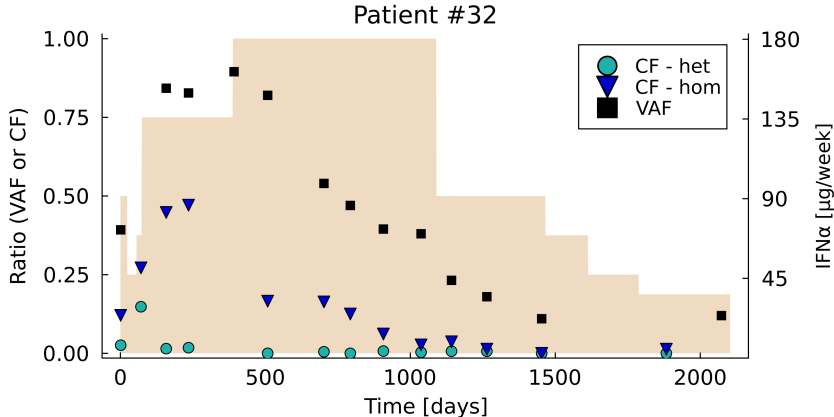


Figure 1: Longitudinal observations of patient #32. The x-axis corresponds to the time from the start of the $IFN\alpha$ therapy. Clinical data consist of measurements of the VAF in mature cells (black squares), experimental data consist of measurements of the heterozygous (green circles) and homozygous (blue triangles) CF in progenitor cells. The shaded beige areas correspond to the weekly dose of $IFN\alpha$ received over time.

2.2 Predicting the long-term dynamics

The model studied in [Mosca et al., 2021] (see Fig. 2 and Supplemental Methods B: Model) was calibrated from observations on 19 MPN patients having the mutation $JAK2^{V617F}$. A hierarchical Bayesian inference method was used to estimate the parameters of each individual in addition to a population effect, thus reducing the risk of overfitting. Then, the minimal $IFN\alpha$ dose $d_{inf}^{(i)}$ required for a patient i to reach long-term remission was estimated (Supp. Methods C.2.4). Our objective is now to demonstrate how the model could be used as a clinical decision-support tool. In clinical routines, trimestrial measurements of the clonal architecture of progenitor cells are not feasible. However, the clonal architecture of heterozygous and homozygous mutated immature cells (i.e. the measurement of CF) yields essential information which determines the response to the treatment as shown in [Tong et al., 2021]. In terms of cost, time, and human resources, it can be considered reasonable to measure the CF twice for each patient: one just before they start their therapy and one afterward (generally after about 300 days of therapy; we explore the impact of this choice in the Supplemental Methods F: Estimating the best timing for measuring the clonal architecture). The VAF in mature

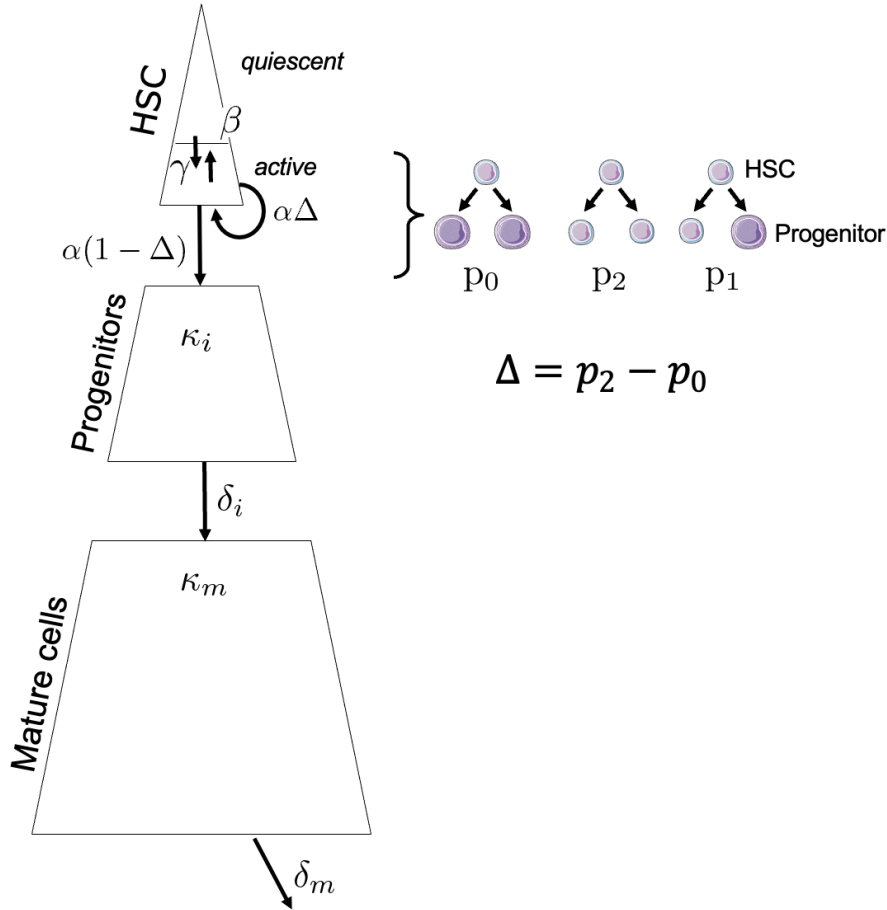


Figure 2: Schematic representation of the model for one cell population (for example, mutated $JAK2^{V617F}$ heterozygous cells). Quiescent HSCs can exit quiescence at a rate $\gamma N_q(t)$ with $N_q(t)$ the number of quiescent HSCs at time t . Active HSCs can become quiescent at a rate $\beta N_a(t)$ or be recruited to divide at a rate $\alpha N_a(t)$ with $N_a(t)$ the number of active HSCs at time t . In the latter case, an HSC could divide symmetrically - producing two HSCs - with probability (w.p.) p_2 , asymmetrically w.p. p_1 , or produce two immature cells (progenitors) w.p. p_0 . $\Delta = p_2 - p_0$ corresponds to the balance between symmetrical and differentiated division. Progenitor cells expand with a rate constant κ_i and become mature cells with a rate constant δ_i . Mature cells expand with a rate constant κ_m and die at a death rate δ_m .

cells can easily be measured during routine clinical follow-up.

If the model were to be used by physicians, then we would have to evaluate its capacity to predict cell dynamics for new patients for whom such data can be easily obtained in clinical routines. Yet, we do not have access to patients outside the cohort of Mosca et al. to study the predictive capabilities of the model and then to study different dose strategies. Therefore, we choose to consider our 19 $JAK2^{V617F}$ mutated MPN patients from which we will remove one individual (leave-one-out) and for whom we will perform data assimilation (Fig. 3-i and Supplemental Methods C: Parameter estimation).

We denote $\mathcal{D}_{-i} = \mathcal{D} \setminus \mathcal{D}^{(i)}$ the dataset for all patients except individual i (Table 1 summarizes the main notations).

In this article, we will consider three individuals who were each in turn excluded from the cohort: patients #12, 18, and 32. These three patients are chosen because several observations of their response before $t = 300$ days of treatment are available, which is only the case for a few patients in the cohort (see Supp. Methods A). Furthermore, each of the three patients exhibits a distinctive type of response to IFN α , making their study interesting: patient #12 has mostly homozygous cells, and they seem to be slowly but steadily targeted over the treatment; patient #18 presents both homozygous and heterozygous mutated cells initially, in the same proportions, and only the homozygous clone is targeted during the treatment; and patient #32 responds with a so-called "bell curve", meaning that their VAF

Notation	Meaning
\mathcal{D}	Dataset of the 19 considered $JAK2^{V617F}$ patients
i	Index of the patient
t	Time
T	Assimilation time (until which observations are considered for calibrating the model)
τ	Patient's age at which the treatment could be interrupted
$\mathcal{D}^{(i)}$	Dataset of patient i
\mathcal{D}_{-i}	Data of all patients except patient i
$\mathcal{D}_T^{(i)}$	Observations of patient i , before time T , used for calibrating the model
$\mathcal{D}_c^{(i)}$	Observations of patient i not used for the model calibration
d	Normalized dose of IFN α (between 0 and 1; 1 corresponding to 180 $\mu\text{g}/\text{week}$)
$z(d)$	IFN α toxicity (as a function of the dose d)
$M(\tau)$	Amount of IFN α administered over $[T, \tau]$ (when penalizing for high dose)

Table 1: List of the main notations.

and CF first strongly increases at the beginning of the treatment and then decreases.

We run three hierarchical estimation procedures (Supp. Methods C.1) corresponding respectively to the datasets \mathcal{D}_{-12} , \mathcal{D}_{-18} and \mathcal{D}_{-32} . Of these three calibrations, we will retain only the estimate of the population distribution (Fig. 3-ii), that will be used as prior distribution to predict the response of the left-apart patient i . The progressive inclusion of observations for this patient (through what is called in Bayesian statistics the likelihood) will allow the updating of the previous distribution (i.e., the prior). We refer to this statistical process as "data assimilation" in this article (Fig. 3-iii) (Supp. Methods C.2). Let $T \in \{300, 600, 1000\}$ [days] be the times (of assimilation) until which we consider the observations of patient i . The dataset of patient i used for the model calibration will be denoted by: $\mathcal{D}_T^{(i)}$. Note that this dataset includes all VAF measurements before time T and only two measurements of the clonal architecture. The other observations of the CF among progenitor cells will not be used for estimating the model parameters. Indeed, in clinical routines, we cannot have repeated measurements of the progenitor clonal architecture. Thus, we choose to use only some of the observations available for progenitor cells before the assimilation time to treat the patient's data as if they were obtained from realistic processes in clinical routine. We consider the observation of the progenitor CF at the start of the therapy and about 300 days afterward.

For patient i , for the assimilation time $T \in \{300, 600, 1000\}$, we will then estimate the *posterior* distribution of each parameter from the prior distribution and the observations $\mathcal{D}_T^{(i)}$. The parameters will then be used to infer the on-treatment dynamics of the VAF in the mature cells as well as the heterozygous and homozygous CF in the progenitors and the HSCs (Fig. 3-iii).

The confrontation of the predicted values with the observations not used for the model calibration will allow us to evaluate the quality of the predictions (Fig. 3-iv). The control dataset for assessing the quality of the predictions will be denoted:

$$\mathcal{D}_c^{(i)} = \mathcal{D}^{(i)} \setminus \mathcal{D}_T^{(i)}$$

This control dataset includes all observations made after time T as well as the observations of the CF among progenitors that were not used for estimating the model parameters.

To quantify how well the model fits the observations, either those used to estimate the model parameters ($\mathcal{D}_T^{(i)}$) or those used as control ($\mathcal{D}_c^{(i)}$), we will compute the mean square error (MSE) (See Supp. Methods C.2.5). In particular, we will check how MSE_{pred}^{VAF} - which confronts the predicted VAF values to the ones truly observed after the assimilation time T (i.e., not used for the model calibration) - evolves for higher T values.

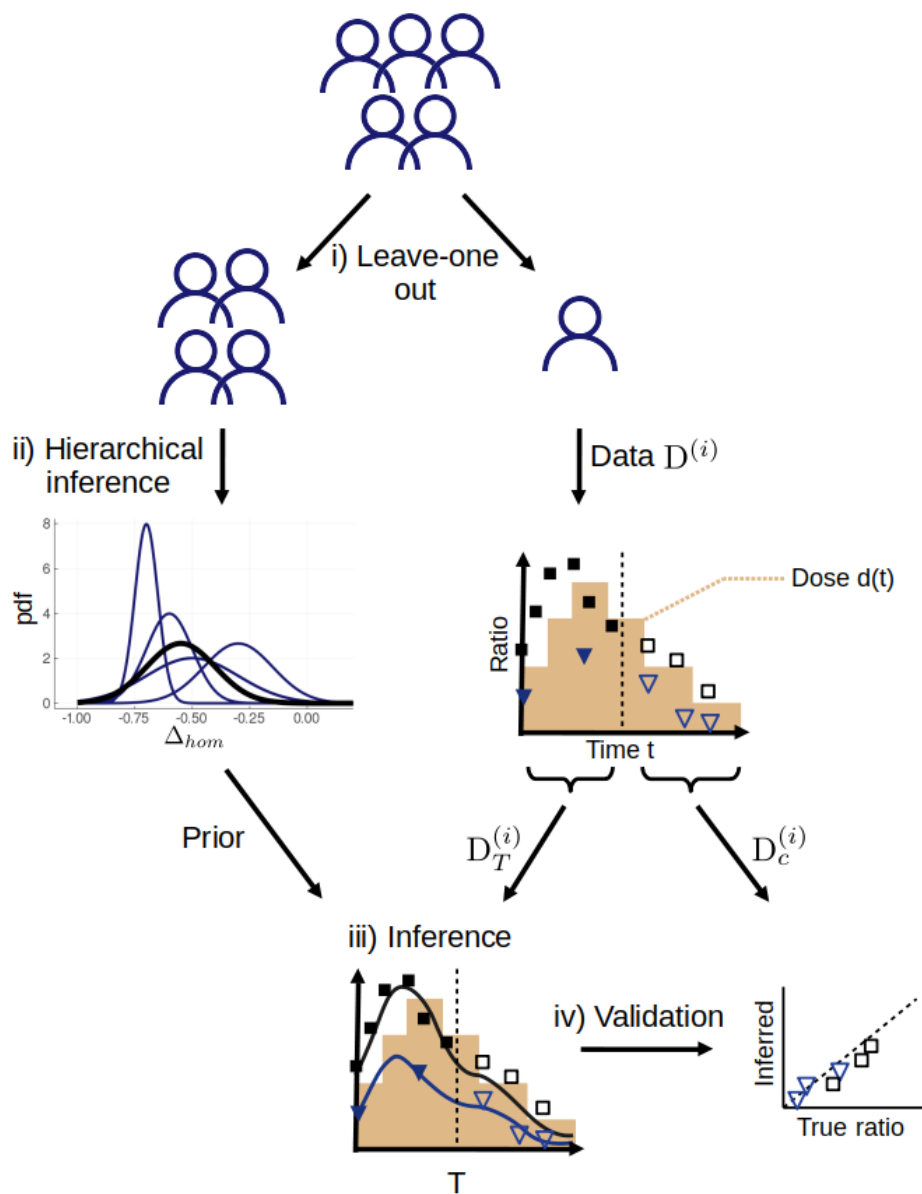


Figure 3: Illustration of the method for predicting the long-term cell dynamics over IFN α therapy. Removing one patient from the cohort (i), we run a hierarchical Bayesian inference procedure using the data of the remaining patients (ii). That is, for each model parameter (here, for example, Δ_{hom}), we estimate its probability density (pdf) for each patient (blue lines) as well as for the patients' population (black line). The latter will be retained as prior knowledge. The data $D^{(i)}$ of the excluded patient from the initial cohort consists of VAF measurements in mature cells (black squares) and CF in progenitor cells, here homozygous CF (blue triangles), over time t . We also have the information on the dose of IFN α $d(t)$ the patient received (in brown). We split their dataset into two parts, the first one $D_T^{(i)}$ corresponding to observations before a time T (that we call assimilation time), and the second one $D_c^{(i)}$ to the observations after that time. Based on the dataset $D_T^{(i)}$ and the prior obtained from the remaining cohort of patients, we infer the dynamics of the mutated cells on treatment (iii). The black line corresponds to the dynamics of the VAF, and the blue one to the dynamics of the homozygous CF. To evaluate the capacity of the model to predict the long-term dynamics correctly, we compare the model predictions and the control dataset $D_c^{(i)}$ that was not used for estimating the model parameters.

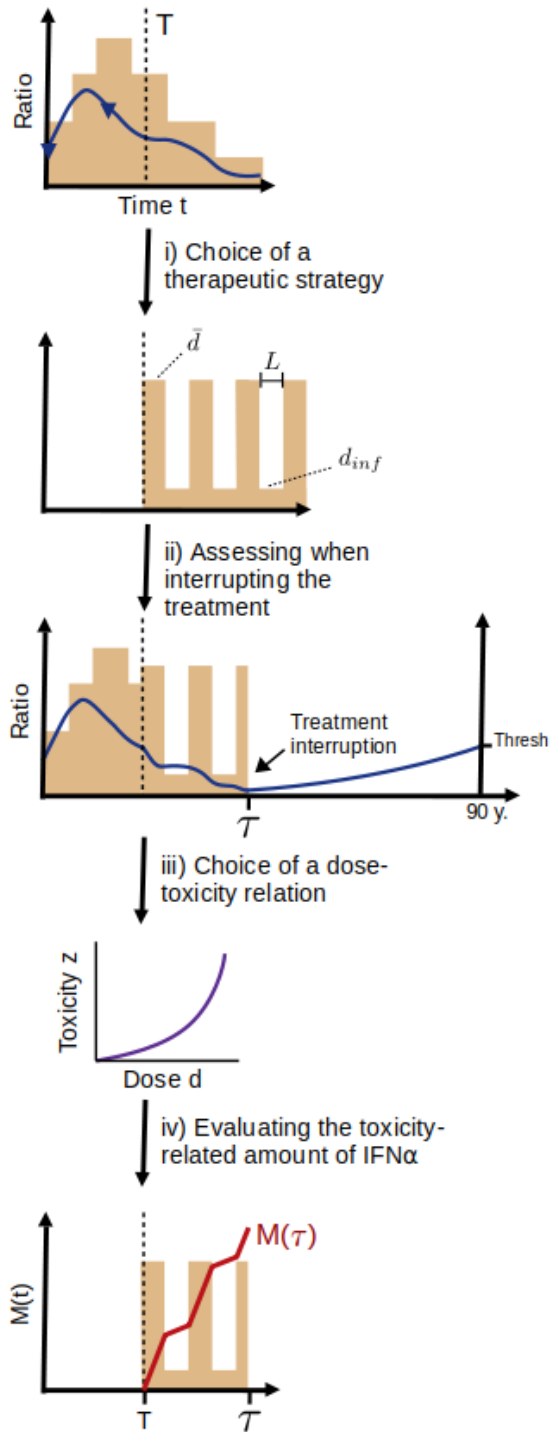


Figure 4: Illustration of the method for optimizing the IFN α therapy. Considering the observations for a patient until time T and their model parameters estimated, we can obtain the dynamics of the mutated cells (here, for example, the CF of homozygous progenitors in blue) over time t , for a given dose $d(t)$. We can then consider different therapeutic strategies, that is, different ways $d(t)$ will evolve for $t \geq T$ (i). Here, for example, we illustrate a therapeutic strategy that is periodic. We can estimate the dynamics of the mutated cells for this new posology and assess the time τ when the CF will be low enough so that, after treatment interruption, the further clonal expansion will not lead to the onset of MPN symptoms before the patient is 90 years old (ii). Note that the illustration is done considering progenitor cells but that in our calculations, we estimate τ based on the inferred CF among HSCs. We then consider one scenario for the increase of IFN α toxicity z with the dose d (here, convex) (iii). Finally, we can compute $M(t)$ for $T \leq t \leq \tau$, which describes the amount of IFN α administered over the time interval $[T, t]$ when penalizing for high doses (iv). The quantity of interest is $M(\tau)$, which we want to minimize across different therapeutic strategies and for which we want to evaluate the impact of the dose-toxicity relation.

2.3 Optimizing the therapy

Once we have evaluated the ability of the model to predict the long-term response from data potentially available in clinical routines, we can explore different therapeutic strategies (§ 2.3.2) and find the optimal one (Fig. 4). The optimality will be defined as the minimal dose of IFN α administered until treatment interruption (§ 2.3.1) when penalizing higher doses associated with higher risks of toxicity (§ 2.3.3). The complete method is presented in the Supplemental Methods E: Optimizing the therapy, while the focus here is on the most important aspects for understanding the article.

2.3.1 Assessing when to interrupt the treatment

IFN α therapy against MPN is a long-term therapy. It has been proven to induce, in some cases, an HMR, which is an undetectable level of $JAK2^{V617F}$ mutated cells in peripheral blood [Kiladjian et al., 2008]. However, we cannot clinically assess whether all mutated HSCs have been completely eradicated and whether a relapse might occur in the future. $JAK2^{V617F}$ positive MPN are hematological malignancies that develop over decades before the onset of the symptoms, as shown recently in different studies [Williams et al., 2022, Van Egeren et al., 2021, Hermange et al., 2022]. It was shown that heterozygous (het) $JAK2^{V617F}$ mutated HSCs, when escaping stochastic extinction, would expand on average according to the following dynamics [Hermange et al., 2022]:

$$\tilde{N}_{het}(t) = \tilde{N}_{0,hct} \exp(s_{het}(t - t_0)) \quad (1)$$

with a fitness estimated to be equal to $s_{het} = 20.4\%/year$, and $\tilde{N}_{het}(t)$ corresponding to the total number of mutated heterozygous HSCs over time t , during the MPN development in the absence of IFN α therapy.

The dynamic of the homozygous clonal expansion was not studied in [Hermange et al., 2022]. Here, we generalize previous findings and assume that the homozygous clone would grow exponentially like the heterozygous one, but with a higher fitness $s_{hom} > s_{het}$, as shown by [Williams et al., 2022]. We estimate $s_{hom} = 1.21 \times s_{het}$ (Supp. Methods D.3).

It is further assumed that the clonal expansion given by eq. (1), inferred from individuals before they received IFN α , will also be valid after they interrupt the treatment. Then, when inferring the CF of heterozygous and homozygous mutated HSCs from the model, at any time during the therapy, we could estimate using eq. (1) what would be the future development of the malignant clones after a treatment interruption at time τ . We can thus deduce the time $t > \tau$ at which the VAF among HSCs would exceed 7.5%. This value corresponds to the classical risk threshold above which it is considered that there could be a risk of thrombosis or cardiovascular events, that is to say, a reappearance of the disease [Dupont et al., 2007]. The later this time, the better for the patient. Ideally, the therapy should be stopped at a time τ so that the relapse does not occur during the patient’s life. In practice, considering a life expectancy at 65 years old, equal to 90 [Kontis et al., 2017], we will consider interrupting the treatment of a patient at an age τ such that the VAF among HSCs stays below 7.5% when the patient is under 90.

Thus, the criterion to interrupt the treatment depends on the age of the patients, the way they respond to the treatment, and their zigosity.

In this study, mutations associated with MPN - such as *DNMT3A*, *TET2*, *TP53* - which could potentially undergo a clonal expansion independent of the $JAK2^{V617F}$ clonal dynamics, are not considered. In the Supplemental Methods D: Interrupting the therapy, we detail the choice of the criterion for the treatment interruption (Supp. Methods D.1), the model used for describing the clonal expansion (Supp. Methods D.2), and the estimation of the homozygous $JAK2^{V617F}$ fitness (Supp. Methods D.3).

2.3.2 Therapeutic strategies

If we can correctly estimate the model parameters of a new patient (and thus correctly predict their mutated cell dynamics when they receive the dose $d(t)$ for $t \geq T$, with T the assimilation time), we could explore alternative therapeutic strategies to those currently used by the clinicians. In practice, clinicians generally increase the dose until they observe a hematological response and then de-escalate the dose, as it was observed on the cohort of Mosca et al. [Mosca et al., 2021]. However, this strategy might not be optimal since it can lead to early relapse.

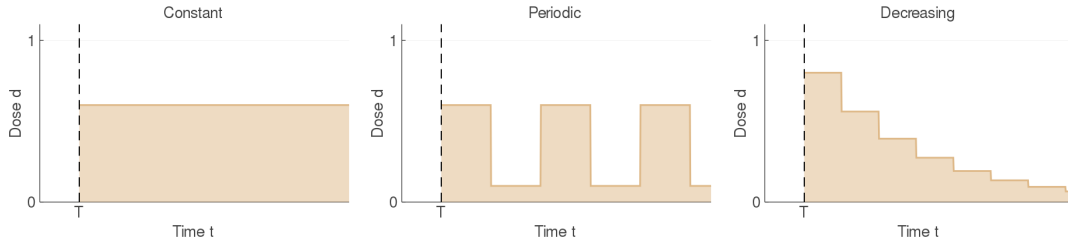


Figure 5: Illustration of the three different therapeutic strategies we consider. On the left, a constant dose over time (here, $\bar{d} = 0.6$). In the middle, a periodic strategy (here, $\bar{d} = 0.6$, $L = 4$ months, and d_{inf} assumed to be equal to 0.1). On the right, a decreasing strategy (here, $\bar{d} = 0.8$, $L = 3$ months, and $\lambda = 0.7$).

In this article, different treatment strategies are studied, all of which will involve some parameters to optimize (§ 2.3.3). Then, the strategies can be compared and the best one can be chosen (according to a criterion that we define later).

The list of strategies studied in this article is far from exhaustive, and we stick to simple ones that clinicians could employ easily. It is considered that the strategy will be set up from the assimilation time T to the interruption time τ . Before time T , the clinician might have treated the patient with a dose-escalation allowing an assessment of the patient's tolerance to the treatment. During the period from the start of the therapy to T , measurements are made and will allow the estimation of the model parameters.

The three different treatment strategies studied are the following (Fig. 5):

- constant:

$$d(t) = \bar{d} \text{ for } t \in [T, \tau]$$

- periodic:

$$d(t) = \begin{cases} \bar{d} & \text{for } t \in [T + 2kL, T + (2k + 1)L], k \in \mathbb{N} \\ d_{inf} & \text{for } t \in [T + (2k + 1)L, T + (2k + 2)L], k \in \mathbb{N} \end{cases}$$

- decreasing:

$$d(t) = \begin{cases} \bar{d} & \text{for } t \in [T, T + L], \\ \lambda d(t - L) & \text{for } t \in [T + L, \tau] \end{cases}$$

The first strategy involves only one parameter \bar{d} to optimize, the second strategy two parameters (\bar{d} and the period L), and the last, three parameters, namely \bar{d} , λ and the period L . This third strategy is the one often implemented in practice. Note that in the equation above, the expression of the decreasing function d is recursive (see Supp. Methods E.3 for another way to express the decreasing strategy).

d_{inf} corresponds to the minimal personalized dose that can be computed from the estimated parameter vector (Supp. Methods C.2.4). For all strategies, we consider $\bar{d} \geq d_{inf}$, $\lambda \in]0, 1[$. We also want $L \geq 3$ months to avoid too frequent changes of posologies (regarding the treatment duration of generally many years). Besides, for too-high values of L , the decreasing and periodic strategies are equivalent to the constant one. To avoid finding ourselves in that case, the following condition is set: $L \leq 2$ years. In addition, λ is constrained to be inferior to 0.95 to avoid the case where the decreasing strategy would approach the constant one when $\lambda \rightarrow 1$. The optimal parameters will be estimated by doing a grid search (Supp. Methods E.1).

To note, we voluntarily choose not to consider stop-and-go strategies. Indeed, patients might need some time to tolerate the treatment; therefore, it should not be stopped to be resumed later to avoid this adaptation phase unless, of course, the interruption is permanent.

The periodic strategy - as defined by imposing the minimal dose to be equal to d_{inf} - is then the strategy that is the closest to the stop-and-go strategy.

2.3.3 Minimizing drug-related toxicity during the treatment

The main objective of the study is to optimize the IFN α therapy against MPN. For that purpose, several potential therapeutic strategies were considered in § 2.3.2. We should define the clinically

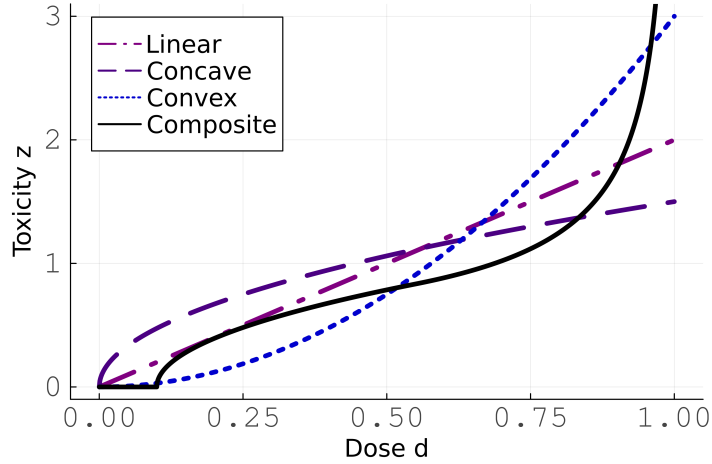


Figure 6: The four different scenarios we consider for modelling the way the IFN α toxicity may increase with the dose.

relevant quantity that has to be optimized. Intuitively, we want the treatment to be as short as possible. With the considered model, the corresponding optimization problem would be solved by choosing to give the maximal IFN α dose of $180\mu\text{g}/\text{week}$. However, the toxicity of IFN α is known to increase with the dose [Yamane et al., 2008], so optimizing only the duration of treatment would result in high-toxicity strategies. Thus, a penalty has to be applied for high doses. Toxicity includes everything harmful to the patient, either hematological or not. In particular, IFN α is known to be associated with side effects such as depression [Lotrich et al., 2007, Trask et al., 2000], flu symptoms, or thrombocytopenia. Since no data exist to quantify the dose-toxicity relation, we will consider several hypothetical scenarios and evaluate their impact on the results. We define $z(d)$ the drug toxicity as a function of the dose and consider four potential behaviours for the dose-toxicity relation (see Fig. 6):

- Linear: $z(d) = 2d$
- Convex: $z(d) = 3d^2$
- Concave: $z(d) = \frac{3}{2}\sqrt{d}$
- Composite : $z(d) = \frac{5\sqrt{5}}{9}\sqrt{d-0.1} \mathbb{1}_{d \in [0.1, 0.55]} + \frac{5}{4\sqrt{5}} \frac{1}{\sqrt{1-d}} \mathbb{1}_{d \in [0.55, 1]}$

For normalization, the average toxicity is set to one for the four formulations:

$$\int_0^1 z(d) \, dd = 1$$

For the composite relation, we assume that there is no toxicity below a low threshold of 0.1 (corresponding to $18 \mu\text{g}/\text{week}$), then a concave behavior for $d \in [0.1, 0.55]$ followed by a convex one where the toxicity tends to infinity for $d \rightarrow 1$. The value delimiting both behaviors - that is, $d = 0.55$ - is chosen so that the derivative of z is continuous over $]0.1, 1[$. More details concerning the construction of the composite relation are provided in the Supplemental Methods E.4.

$M(t)$ is defined as the toxicity-related amount of IFN α administered over $[T, t]$:

$$M(t) = \int_T^t z(d(u)) \, du \quad (2)$$

where d (as a function of time) is assumed to be either constant, periodic, or decreasing, and z (as a function of d) is either linear, convex, concave, or composite. Then, the quantity to minimize is $M(\tau)$. In the linear scenario, $M(\tau)$ would then correspond to the total amount of IFN α administered between

T and τ . To note that, in eq. (2), we do not consider the interval $[0, T]$ since we only want to optimize the treatment from $t = T$.

We study for patients #12, 18, and 32 three therapeutic strategies and four scenarios of dose toxicity, that is, twelve different conditions. For each of them, we will estimate the parameters related to the therapeutic strategy that minimize the quantity $M(\tau)$ (Supp. Methods E.1). The strategies are then compared based on this value, for a given hypothetical scenario of toxicity, to find the optimal one. Moreover, in the absence of prior knowledge on how IFN α toxicity increases with the dose, a trade-off strategy can be proposed (Supp. Methods E.2).

3 Results

3.1 Minimal observations are sufficient to predict the long-term mutated cell dynamics

First we investigated whether some measurements of the VAF and only two measurements of the progenitor clonal architecture were sufficient to predict the long-term response to the treatment. To present the results in this section, we will focus on patient #18 (All our results are details in the Supplemental Material G: Detailed Results; the results for patients #12 and #32 concerning the prediction part are detailed in Supp. Material G.1). The cell dynamics observed during IFN α therapy for patient #18 are interesting since they initially presented both heterozygous and homozygous clones (and in the same proportions), but the homozygous clone was targeted during the treatment when the heterozygous one continued to expand. Such dynamics might be challenging to predict when the parameter estimation is mainly based on VAF measurements.

For patient #18, we assess whether we could predict the evolution of their VAF correctly, but also both their heterozygous and homozygous CF in progenitors (Fig. 7), from:

1. only two measurements of the clonal architecture in immature cells (at the start of the treatment and 248 days afterward),
2. several measurements of the VAF before a time $T \in \{300, 600, 1000\}$ [days], and
3. the prior knowledge obtained from the remaining patients.

By progressively adding more information on the VAF dynamics (that is, increasing the assimilation time), we increase the confidence in our predictions. It is particularly true for the VAF and the homozygous CF. As early as $T = 300$ days, when only four VAF measurements are used for the parameter estimation, we observe a good agreement between the predicted values, both for the VAF and the CF, and the experimental values not used for estimating the model parameters.

MSE_{pred}^{VAF} - quantifying the error between the observed and predicted VAF for different assimilation times T - is equal to:

- $6.6 \cdot 10^{-3}$ for $T = 300$ (10 VAF measurements used for computing MSE_{pred}^{VAF})
- $5.8 \cdot 10^{-4}$ for $T = 600$ (6 measurements used for computing this MSE)
- $5.2 \cdot 10^{-4}$ for $T = 1,000$ (3 measurements used for computing this MSE)

We show that adding more observations for the VAF improves the quality of our prediction. However, the (median) inferred dynamics for the mutated heterozygous progenitors are closer to the observations for $T = 600$ days than $T = 1,000$, resulting for the latter to a higher MSE on the heterozygous CF measurements not used for the inference (Supp. Table G.2). For $T = 1,000$, we overestimate the heterozygous CF, illustrating the difficulty in correctly predicting both the dynamics of heterozygous and homozygous cells when the inference is mainly based on the VAF measurements.

For patient #12, we also observe a good agreement between the predicted VAF and the observed ones, with a MSE_{pred}^{VAF} of $1.1 \cdot 10^{-3}$ and $1.5 \cdot 10^{-3}$ for T equal to 300 and 600, respectively. However, we systematically underestimate the true homozygous CF (Supp. Material G.1.1, Supp. Fig. G.1, and Supp. Table G.1). For patient #32, we obtained good predictions, both for the VAF and the homozygous CF (patient #32 has almost no heterozygous clones) for $T = 300$ or 1,000 days (with

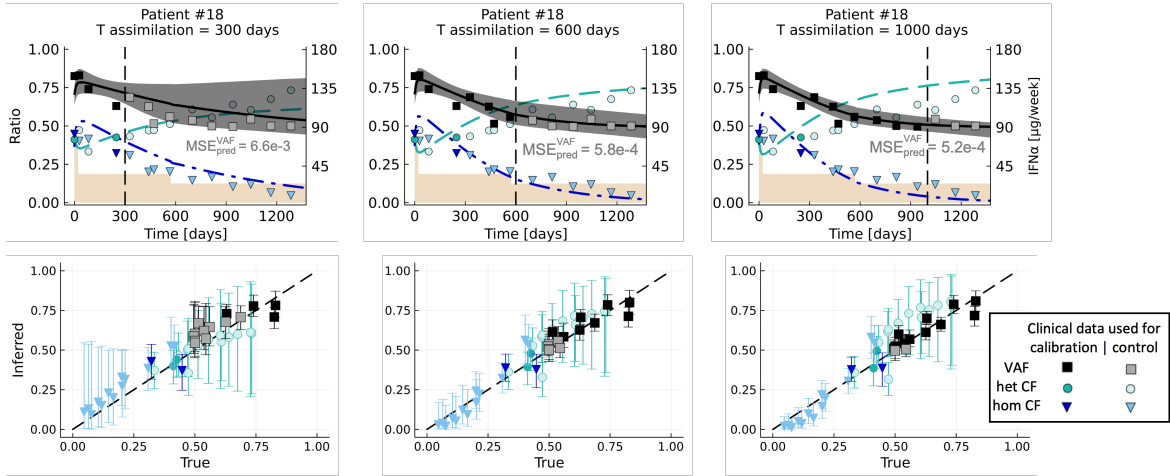


Figure 7: Data assimilation results for patient #18. From left to right, the assimilation times are $T = 300$, 600, and 1,000 days (vertical dash line). At the top, we present the predicted dynamics. In black are shown the inferred VAF (median value and 95% credibility interval) and in green and blue, the inferred CF (median values) of heterozygous and homozygous mutated progenitors, respectively. The squares correspond to the clinical data, and the circles and triangles to the experimental data, either used to calibrate the model (dark colors, $\mathcal{D}_T^{(i)}$) or to control the quality of the predictions (light colors - control dataset $\mathcal{D}_c^{(i)}$). We represent the IFN α dose variations during treatment (in $\mu\text{g}/\text{week}$) in brown. The MSE displayed just below the VAF dynamics, is computed on the VAF measurements not used for the calibration (i.e., the VAF measurements after the assimilation time T), such that MSE_{pred}^{VAF} quantifies the predicted error on the VAF. At the bottom, we compare the inferred values (median values, on the y-axis) with those observed (on the x-axis). The error bars correspond to the 95% credibility intervals.

MSE_{pred}^{VAF} respectively equal to $2.17 \cdot 10^{-2}$ and $2.35 \cdot 10^{-2}$), but, surprisingly, poor predictions for $T = 600$ days ($\text{MSE}_{pred}^{VAF} = 0.25$) (Supp. Material G.1.3, Supp. Fig. G.5, and Supp. Table G.3). But, if instead of considering the measurement of the clonal architecture at time $t = 392$, we chose to take the observation at $t = 508$, we observed that the predictions for $T = 600$ days were far better ($\text{MSE}_{pred}^{VAF} = 7.5 \cdot 10^{-3}$) (Supp. Material G.1.4, Supp. Fig. G.6, and Supp. Table G.4), illustrating the importance of the experimental design to make accurate predictions. We explore this question in the Supplemental Material G.3.

From the results on the three patients considered in this study, we conclude that it is possible to make predictions sensible of the mutated cell dynamics from VAF observations and two measurements of the CF among progenitor cells, but that the choice of the time point for the clonal architecture matters, raising the concern of how the observation time should be chosen. After 300 days of treatment, it is possible to get good predictions of the cell dynamics and, after 600 days, to have good confidence in our predictions.

3.2 The dose should be kept constant until the treatment is stopped

We studied whether other therapeutic strategies could lead to better results - faster molecular response while limiting the potential IFN α toxicity. To evaluate these strategies, we focused on patient #32 (results for patients #12 and #18 are detailed in the Supplemental Material G.2). For this patient, the model predicts that a decrease in the dose after 2,000 days of therapy might slow down the molecular response, and even induce a risk of relapse (Supp. Fig. G.6). The way the patient was treated, with a dose escalation over the first three years of therapy up to the maximal dose of $180\mu\text{g}/\text{week}$, then a de-escalation, is also characteristic of the strategies applied in clinical routine and already observed in [Mosca et al., 2021].

We consider $T = 600$ days. Using the observations before that time (5 VAF observations, two clonal architectures: one at $t = 0$, the second one at $t = 508$) and the prior knowledge obtained from the 18 remaining patients, we estimated the posterior distribution of the model parameters and showed

that we obtain good predictions (Supp. Material G.1.4).

We now consider that the actual parameter vector is the estimated one (mean of the posterior distribution) and study how different dose strategies impact the response of the treatment after $T = 600$ days. With the estimated parameter vector, the minimal dose - under which the treatment would not target sufficiently the mutated HSCs, resulting in a relapse - is estimated to $d_{inf} = 0.345$, that is, $62 \mu\text{g}/\text{week}$. We consider different scenarios for how the drug toxicity increases as a function of the dose, as presented in Fig. 6. For each scenario, we study the three therapeutic strategies - presented in § 2.3.2 - for which we find the parameters (e.g. the choice of the dose \bar{d}) that minimize the value of $M(\tau)$. This latter value corresponds to the toxicity-related amount of IFN α administered from T to the time τ when the therapy could be interrupted (§ 2.3.1). We can compare the three therapeutic strategies for a given drug-toxicity relation and select the one with the best (i.e., the lowest) value for $M(\tau)$. Results are presented in Fig. 8. It turns out that the constant strategy is the optimal one among the four studied scenarios, also for patients #12 (Supp. Fig. G.8) and #18 (Supp. Fig. G.9). With the two other strategies, namely the periodic and the decreasing one, we obtain at the optimum lower values of $M(\tau)/(\tau - T)$ (that we interpret as values of the yearly toxicity-related amount of IFN α) compared to the constant strategy. However, having periods with low doses or decreasing the dose over time also delays treatment interruption. Therefore, these two strategies are less effective in terms of treatment duration.

However, the choice of the optimal dose highly depends on how the IFN α toxicity increases with the dose. If there were a sharp increase in the toxicity for low doses, followed by a slower one (concave scenario), we would recommend using high doses (about $180 \mu\text{g}/\text{week}$ for patient #32). Indeed, only slightly decreasing the dose would not strongly impact the toxicity, while decreasing the dose to a too large extent would delay the treatment interruption or even induce a relapse if the dose is below a minimal value, as shown in [Hermange et al., 2021]. If, on the contrary, the toxicity would only sharply increase for high doses (convex scenario), then we would recommend using medium doses (about $90 \mu\text{g}/\text{week}$ for patient #32). Eventually, for the linear dose-toxicity relation (which is both concave and convex), or for the composite relation (which is first concave, then convex), we find that the recommended dose for patient #32 would be about $135 \mu\text{g}/\text{week}$. Given the lack of existing data quantifying the relationship between the dose and the toxicity of IFN α , we cannot conclude which dose level would be optimal. Our results highlight the need to estimate such relations since we demonstrate their importance in deciphering how MPN patients should be treated.

In the absence of prior knowledge on whether the dose toxicity would better be described by a linear, concave, convex, or composite relationship, we can determine the strategy which - even if not optimal - would be the best compromise. Finding a trade-off strategy is important since the best strategy under the hypothesis that the dose-toxicity relation would be concave would be detrimental under another hypothesis. We illustrate that point with patient #32. In the concave scenario, the best strategy is to treat patient #32 at a constant dose of $180 \mu\text{g}/\text{week}$ over about five years. This value would actually correspond to the highest dose considered in the study, that is, a very high dose. If the dose-toxicity relation were actually not concave but linear, such a strategy would still be acceptable, with a value for $M(\tau)$ equal to 4020, placing it in the top 3.4 percent and better than the best periodic strategy. But, in the case where the dose-toxicity relationship was convex, treating patient #32 with such a high dose would be harmful. Indeed, it turns out that the optimal strategy in the concave scenario is also the worst one in the convex scenario.

Therefore, without having assessed in the first two years of the therapy how the patient responds to the treatment or in the absence of data that would quantify the dose-toxicity relation for IFN α , the best choice would be to select the trade-off strategy.

This strategy is the one that gives good results under the four hypothetical dose-toxicity relationships. For patient #32, the trade-off would be to treat him with a constant dose of $115 \mu\text{g}/\text{week}$ over eight years, until the treatment could be interrupted, here at age 62 (i.e., after about 10 years of IFN α therapy since patient #32 began the treatment at age 52). This trade-off strategy would be:

- in the top 1.7 percent if the dose-toxicity relation were linear,
- in the top 2.3 percent if the dose-toxicity relation were convex,
- in the top 2.9 percent if the dose-toxicity relation were concave,

- and in the top 0.2 percent if the dose-toxicity relation were composite.

For patient #18, the trade-off strategy would also be the constant one, with $\bar{d} = 61 \mu\text{g}/\text{week}$, until treatment discontinuation at the age of 72 (i.e., after about 15 years of IFN α therapy). For patient #12, the trade-off strategy would be the decreasing one, with $\bar{d} = 86 \mu\text{g}/\text{week}$, $\lambda = 0.45$, and $L = 16$ months, until treatment discontinuation at the age of 77 (i.e., after about 14 years of IFN α therapy) (Supp. Material G.2).

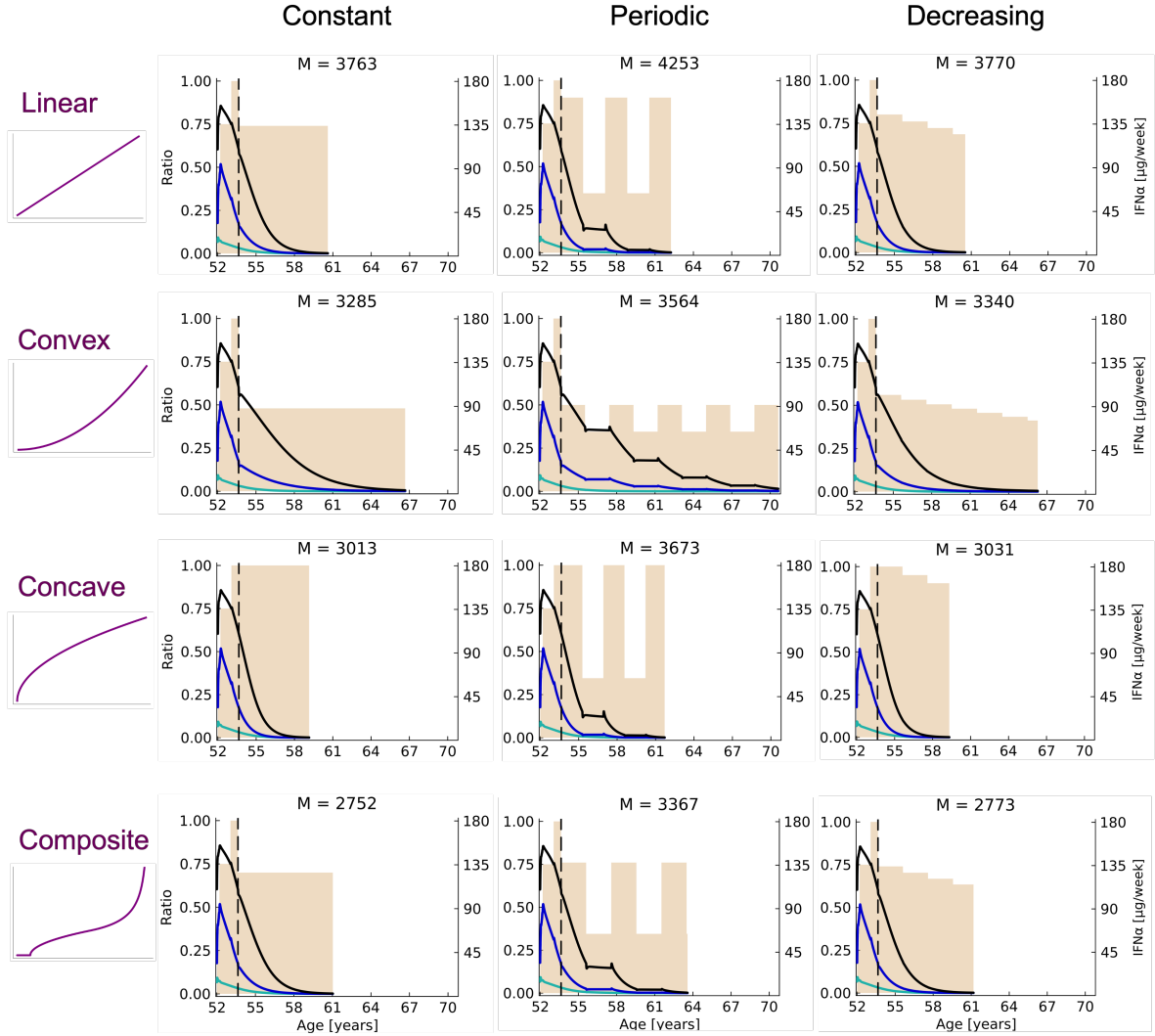


Figure 8: Results of the treatment optimization for patient #32. Four hypothetical relationships between IFN α toxicity and the dose are studied: linear (at the top), convex (on the second line), concave (third line), and composite (at the bottom). For each of them, three therapeutic strategies - constant on the left, periodic on the middle, and decreasing on the right - are optimized to minimize $M(\tau)$ (denoted by M on the figures), which is the toxicity-related amount of IFN α administrated from T (vertical dashed line) to the treatment interruption time τ (when the weekly dose of interferon - represented in beige - drops to zero). The dynamics of the VAF (black line), the heterozygous CF (green line), and the homozygous CF (blue line) will differ for $t \geq T$ according to how the doses vary (depending on the considered therapeutic strategy and its parameter values). M values are only comparable for a given dose-toxicity relation; the lower, the better.

If the actual dose-toxicity relation were linear, the constant strategy would be optimal for $\bar{d} = 135 \mu\text{g}/\text{week}$ (with the associated age of treatment interruption $\tau = 61$ years and an optimal value of $M(\tau) = 3763$), the periodic strategy would be optimal for $\bar{d} = 153 \mu\text{g}/\text{week}$ and $L = 22$ months, and the decreasing one for $\bar{d} = 144 \mu\text{g}/\text{week}$, $L = 2$ years, and $\lambda = 0.95$. Among these three strategies, the constant one is the best. We can see how the optimal parameters of the decreasing strategy actually tend to make it approach a constant strategy, which would be the limiting case when $\lambda \rightarrow 1$ and/or $L \rightarrow \infty$. It is precisely to avoid finding oneself in this limiting case that constraints were imposed on the parameters (Supp. Methods E.1).

If the actual dose-toxicity relation were convex, the constant strategy would be optimal for $\bar{d} = 86 \mu\text{g}/\text{week}$, the periodic strategy would be optimal for $\bar{d} = 90 \mu\text{g}/\text{week}$ and $L = 23$ months, and the decreasing one for $\bar{d} = 100 \mu\text{g}/\text{week}$, $L = 2$ years, and $\lambda = 0.95$. Among these three strategies, the constant one is the best.

If the actual dose-toxicity relation were concave, the constant strategy would be optimal for $\bar{d} = 180 \mu\text{g}/\text{week}$, the periodic strategy would be optimal for $\bar{d} = 180 \mu\text{g}/\text{week}$ and $L = 20$ months, and the decreasing one for $\bar{d} = 180 \mu\text{g}/\text{week}$, $L = 2$ years, and $\lambda = 0.95$. Among these three strategies, the constant one is the best.

If the actual dose-toxicity relation were composite, the constant strategy would be optimal for $\bar{d} = 126 \mu\text{g}/\text{week}$, the periodic strategy would be optimal for $\bar{d} = 137 \mu\text{g}/\text{week}$ and $L = 2$ years, and the decreasing one for $\bar{d} = 133 \mu\text{g}/\text{week}$, $L = 2$ years, and $\lambda = 0.95$. Among these three strategies, the constant one is the best.

4 Discussion

We proposed a mathematical approach, combining modeling, statistical inference, and optimization, to rationalize the manner in which we treat $JAK2^{V617F}$ MPN patients in the long term using $IFN\alpha$. Clinicians often proceed to a dose de-escalation when they observe a hematological response. The rationale behind this strategy is that $IFN\alpha$ is associated with general or hematological side effects, such as depression [Lotrich et al., 2007, Trask et al., 2000], flu symptoms, or thrombocytopenia, and the toxicity increases with the dose [Yamane et al., 2008]. In the absence of both data quantifying the dose-toxicity relation and information on the response of mutated HSCs to the treatment, their therapeutic strategy, albeit empirical, is relevant. However, as observed from the cohort of Mosca et al. [Mosca et al., 2021], clinicians sometimes have to increase the dose again - after a treatment interruption or a strong decrease - when a relapse is observed. To avoid such behaviors that could harm the patient, we explored which therapeutic strategy would be optimal given hypothetical dose-toxicity relations. In particular, we estimated which dose level should be administered. We found that treating the patients with a constant dose until the therapy is interrupted rather than decreasing or alternating a low and a higher dose periodically should be more relevant. Even if not considered initially as a selection criterion, the simplicity of the constant strategy - involving only one value to choose, that is, the $IFN\alpha$ dose to be kept constant until the treatment interruption - makes it easy to deploy in clinics. Such a simple strategy appears more optimal than more sophisticated ones. However, the number of strategies we explored was far from exhaustive; thus, we could not exclude that MPN patients could be treated more successfully with different strategies. In our mathematical approach, we do not guarantee that, after interrupting the treatment, there will not be a relapse. On the contrary, we quantified this risk and chose to interrupt the treatment so that the reappearance of the MPN disease would not occur before the age of 90. This age threshold is based on the estimated life expectancy of 65-year-old individuals in developed countries [Kontis et al., 2017]. Increasing this age, thus reducing the risk, would result in treating the patient for a longer time. It might also seem relevant to adjust this criterion regarding the sex of the patient, given that women have a higher life expectancy than men. Concerning the reappearance of the MPN disease, we considered that the relapse might occur when the VAF in HSCs exceeds 7.5% [Dupont et al., 2007]. However, some patients with essential thrombocythemia might exhibit lower VAF, while a higher VAF is necessary for others to get the MPN symptoms. The choice of this value particularly matters when considering the problem of early screening [Hermange et al., 2022]. In our case, however, the patients who have interrupted their treatment would still be followed so that clinicians might detect quickly if and when a relapse occurs. Given hypothetical dose-toxicity relations, we found that treating a patient with a constant dose is optimal. However, the value of the dose depends on how $IFN\alpha$ toxicity increases with the dose. High doses should be recommended if the patient tolerates the treatment well; medium or low doses otherwise. No data quantify, in the general case, what the dose-toxicity relation could be. Such a relation could be actually patient-dependent. In our approach, we looked for the best therapeutic strategy once the patient had already been treated for 600 days. Therefore, we could already get prior insight into how the patient responded to a dose escalation. In other words, the clinician might provide information on the potential dose-toxicity relation, so we could recommend the constant dose value to administer to the patient. Involving the clinicians in getting such prior information would benefit our method, which could, in return, provide more relevant guidelines on how it might be optimal to treat the patient. Without such dose-toxicity relation, we can find a trade-off strategy. This strategy is such that the dose-toxicity-related quantity of $IFN\alpha$ administered to the patient would be as little as possible for all the four scenarios of dose-toxicity we considered. It is worth noting that, for optimizing the therapy, we only considered minimizing a quantity related to the drug toxicity and did not account for economic criteria, for example. The issue of the economic cost of the therapy was studied by [Pedersen et al., 2020]. We also limit ourselves to four potential dose-toxicity relations, where the toxicity would strictly increase with the dose. It should not be excluded that more complex relations might exist and that the toxicity could, for example, reach a plateau. With the four dose-toxicity relations we considered, we found that the trade-off strategy would also be when the dose is maintained at a constant value for two of the three patients studied. For the third patient, the decreasing strategy was found to be the best compromise; our results are patient-dependent, and more patients should be studied to identify general patterns. In particular, the dose level is specific to each patient and, more precisely, to how they respond to the treatment. For the two patients whose trade-off strategy was the constant one, we found different optimal values of 61 and 115 $\mu\text{g}/\text{week}$.

Our study considered any potential dose between 0 and $180\mu\text{g}/\text{week}$. In practice, clinicians can inject 135 or $180\mu\text{g}$ of IFN α . By choosing the period between two medication intakes, we can have a range of potential doses, but some dose values would still not be clinically relevant. Furthermore, when the frequency of IFN α injection is low, below one every ten days [Xu et al., 1998], the modeling assumption that the dose input is a piece-wise constant function might not be justified. It would then be worth studying pharmacokinetic/pharmacodynamic (PK/PD) models (for a review, see, for example, [Gabrielsson and Green, 2009]).

Our mathematical approach focuses on the mutated $JAK2^{V617F}$ HSCs: how IFN α targets them and then how they expand after the treatment interruption. The fact that HSCs cannot be observed *in vivo* justifies using mathematical models to infer their dynamics. Here, we combined two models: one that describes the dynamics of mutated cells during the therapy [Mosca et al., 2021] and one that describes the clonal expansion of the mutated cells in the absence of IFN α [Hermange et al., 2022]. We assumed that the stochastic effect could be neglected. However, when the treatment lasts for a long time, the number of mutated cells might reach a low value, so deterministic models might not be appropriate anymore. It would then be relevant to propose a stochastic model for how IFN α would selectively target a few mutated HSCs within a large pool of wild-type (normal) HSCs. Concerning the wild-type HSCs, our model is based on the hypothesis that their number stays constant, even during disease development. Alternative models could also be proposed, either for the effect of the treatment [Pedersen et al., 2021, Ottesen et al., 2020] or the clonal expansion without IFN α [Van Egeren et al., 2021, Williams et al., 2022]. In particular, the model from Pedersen, Ottesen et al. [Pedersen et al., 2021, Ottesen et al., 2020] presents the advantage of having been initially designed for calibration from VAF measurements; thus, it might be more relevant for use in clinical routine. The counterpart is that they do not account for the zygosity, which is a determinant of the response to IFN α [Tong et al., 2021, Mosca et al., 2021]. A roundabout way exists to access information on zygosity in mature cells from a VAF measurement: using the 46/1 haplotype [Hasan et al., 2014] for patients heterozygous for this polymorphism (it would also be possible to determine other polymorphisms as informative). Indeed, it has been shown that generally, in the case of mitotic recombination by which a heterozygous $JAK2^{V617F}$ mutated cell gives a homozygous mutated cell, the cell will also become homozygous for the 46/1 haplotype. Thus, measuring the VAF for this haplotype will tell us the proportion of homozygous mutated cells. Even if the model from [Mosca et al., 2021] was not originally calibrated from data that could be easily obtained in clinical routine, we showed that having only two observations of the progenitor clonal architecture would be sufficient to get accurate predictions when having some prior information on the parameter distribution. This prior knowledge came from a hierarchical Bayesian inference when using all available information about the heterozygous and homozygous CF dynamics in progenitor cells (and not only two measurements of the clonal architecture) from other patients. Still, the prediction quality is uneven depending on the timing for measuring the clonal architecture. Poor predictions also mean that our recommendations for treating the patient could not be adapted. This is why we explored the issue of optimal experimental design: how to choose the best timing for measuring the clonal architecture, i.e., the one leading to the most accurate inferences. Intensive *in silico* investigations have shown that the time for the second CF measurement is important, but a search for any systematic was inconclusive (Supp. Materials F and G.3). It turned out that the best timing for a given patient might not be the best for another, preventing us from finding a general pattern on when it would be optimal to measure the CF among progenitor cells for a new patient. Thus, it might be worth adding more patients to the study. In addition, the experimental design question could be extended to the dates at which the VAF should be measured. We only superficially addressed the issue of experimental design, which deserves to be studied in more detail. However, if clinicians can easily use our model-based predictions and treatment recommendations, it would be more complicated to make recommendations on when they should collect patient blood samples. Indeed, the timing for collecting patient observations is subject to scheduling constraints over which we have no influence. It would be difficult to impose the dates when the patient should have an appointment with their clinicians. For that reason, we limited ourselves to the experimental design study and only considered the question of the best timing for the second measurement of the progenitor clonal architecture. Even if we could not conclude when it would be optimal for a new patient to get that measurement, we showed that not all choices were equal.

Finally, we demonstrated how a mathematical model calibrated from real data could predict MPN patients' response to IFN α and optimize their long-term therapy. The proposed methods were designed to be easily applied as a decision-support tool in clinical routine. Our mathematical approach can be extended to study other drugs against MPN, such as Ropeg-IFN α 2b [Gisslinger et al., 2020, Barbui et al., 2020, Mascarenhas et al., 2022], which is a different pegylated IFN α from the one studied in this article, or even its combination with ruxolitinib [Kiladjian et al., 2022]. Furthermore, our methods could be applied more broadly to studying other chronic hematological malignancies.

References

- [Austin et al., 2020] Austin, R. J., Straube, J., Bruedigam, C., Pali, G., Jacquelin, S., Vu, T., Green, J., Gräsel, J., Lansink, L., Cooper, L., et al. (2020). Distinct effects of ruxolitinib and interferon-alpha on murine jak2v617f myeloproliferative neoplasm hematopoietic stem cell populations. *Leukemia*, 34(4):1075–1089.
- [Barbui et al., 2020] Barbui, T., Vannucchi, A., De Stefano, V., Masciulli, A., Carobbio, A., Ghirardi, A., Ciceri, F., Bonifacio, M., Iurlo, A., Palandri, F., et al. (2020). Phase ii randomized clinical trial comparing ropeginterferon versus phlebotomy in low-risk patients with polycythemia vera. results of the pre-planned interim analysis. *Hemasphere*, 4:2602.
- [Dupont et al., 2007] Dupont, S., Massé, A., James, C., Teyssandier, I., Lécluse, Y., Larbret, F., Ugo, V., Saulnier, P., Koscielny, S., Le Couédic, J. P., et al. (2007). The jak2 617v_f mutation triggers erythropoietin hypersensitivity and terminal erythroid amplification in primary cells from patients with polycythemia vera. *Blood*, 110(3):1013–1021.
- [Gabrielsson and Green, 2009] Gabrielsson, J. and Green, A. R. (2009). Quantitative pharmacology or pharmacokinetic pharmacodynamic integration should be a vital component in integrative pharmacology. *J Pharmacol Exp Ther*, 331(3):767–774.
- [Gisslinger et al., 2020] Gisslinger, H., Klade, C., Georgiev, P., Krochmalczyk, D., Gercheva-Kyuchukova, L., Egyed, M., Rossiev, V., Dulicek, P., Illes, A., Pylypenko, H., et al. (2020). Ropeginterferon alfa-2b versus standard therapy for polycythaemia vera (proud-pv and continuation-pv): a randomised, non-inferiority, phase 3 trial and its extension study. *Lancet Haematol*, 7(3):e196–e208.
- [Hasan et al., 2014] Hasan, S., Cassinat, B., Droin, N., Le Couedic, J., Favale, F., Monte-Mor, B., Lacout, C., Fontenay, M., Dosquet, C., Chomienne, C., et al. (2014). Use of the 46/1 haplotype to model jak2v617f clonal architecture in pv patients: clonal evolution and impact of ifn α treatment. *Leukemia*, 28(2):460–463.
- [Hasan et al., 2013] Hasan, S., Lacout, C., Marty, C., Cuingnet, M., Solary, E., Vainchenker, W., and Villeval, J.-L. (2013). Jak2v617f expression in mice amplifies early hematopoietic cells and gives them a competitive advantage that is hampered by ifn α . *Blood*, 122(8):1464–1477.
- [Hermange et al., 2022] Hermange, G., Rakotonirainy, A., Bentriou, M., Tisserand, A., El-Khoury, M., Girodon, F., Marzac, C., Vainchenker, W., Plo, I., and Cournède, P.-H. (2022). Inferring the initiation and development of myeloproliferative neoplasms. *Proc Natl Acad Sci U S A*, 119(37):e2120374119.
- [Hermange et al., 2021] Hermange, G., Vainchenker, W., Plo, I., and Cournède, P.-H. (2021). Mathematical modelling, selection and hierarchical inference to determine the minimal dose in ifn alpha therapy against myeloproliferative neoplasms. *arXiv preprint arXiv:2112.10688*.
- [Kiladjian et al., 2008] Kiladjian, J.-J., Cassinat, B., Chevret, S., Turlure, P., Cambier, N., Roussel, M., Bellucci, S., Grandchamp, B., Chomienne, C., and Fenaux, P. (2008). Pegylated interferon-alfa-2a induces complete hematologic and molecular responses with low toxicity in polycythemia vera. *Blood*, 112(8):3065–3072.
- [Kiladjian et al., 2022] Kiladjian, J.-J., Ianotto, J.-C., Soret, J., Maslah, N., Chaffaut, C., Barraco, F., Dubruille, V., Capron, C., Tisserand, A., Rolland-Neyret, V., et al. (2022). Final results of ruxopeg, a phase 1/2 adaptive randomized trial of ruxolitinib (rux) and pegylated interferon alpha (ifna) 2a in patients with myelofibrosis (mf). *Blood*, 140(Supplement 1):577–578.
- [Kontis et al., 2017] Kontis, V., Bennett, J. E., Mathers, C. D., Li, G., Foreman, K., and Ezzati, M. (2017). Future life expectancy in 35 industrialised countries: projections with a bayesian model ensemble. *Lancet*, 389(10076):1323–1335.
- [Lotrich et al., 2007] Lotrich, F. E., Rabinovitz, M., Girona, P., and Pollock, B. G. (2007). Depression following pegylated interferon-alpha: characteristics and vulnerability. *J Psychosom Res*, 63(2):131–135.

- [Mascarenhas et al., 2022] Mascarenhas, J., Kosiorek, H. E., Prchal, J. T., Rambaldi, A., Berenzon, D., Yacoub, A., Harrison, C. N., McMullin, M. F., Vannucchi, A. M., Ewing, J., et al. (2022). A randomized phase 3 trial of interferon- α vs hydroxyurea in polycythemia vera and essential thrombocythemia. *Blood*, 139(19):2931–2941.
- [Mosca et al., 2021] Mosca, M., Hermange, G., Tisserand, A., Noble, R., Marzac, C., Marty, C., Le Sueur, C., Campario, H., Vertenoil, G., El-Khoury, M., et al. (2021). Inferring the dynamics of mutated hematopoietic stem and progenitor cells induced by ifn α in myeloproliferative neoplasms. *Blood*.
- [Mullally et al., 2013] Mullally, A., Bruedigam, C., Poveromo, L., Heidel, F. H., Purdon, A., Vu, T., Austin, R., Heckl, D., Breyfogle, L. J., Kuhn, C. P., et al. (2013). Depletion of jak2v617f myeloproliferative neoplasm-propagating stem cells by interferon- α in a murine model of polycythemia vera. *Blood*, 121(18):3692–3702.
- [Ottesen et al., 2020] Ottesen, J. T., Pedersen, R. K., Dam, M. J., Knudsen, T. A., Skov, V., Kjær, L., and Andersen, M. (2020). Mathematical modeling of mpns offers understanding and decision support for personalized treatment. *Cancers*, 12(8):2119.
- [Pedersen et al., 2020] Pedersen, R. K., Andersen, M., Knudsen, T. A., Sajid, Z., Gudmand-Hoeyer, J., Dam, M. J., Skov, V., Kjær, L., Ellervik, C., Larsen, T. S., et al. (2020). Data-driven analysis of jak2v617f kinetics during interferon-alpha2 treatment of patients with polycythemia vera and related neoplasms. *Cancer Med*, 9(6):2039–2051.
- [Pedersen et al., 2021] Pedersen, R. K., Andersen, M., Knudsen, T. A., Skov, V., Kjær, L., Hasselbalch, H. C., and Ottesen, J. T. (2021). Dose-dependent mathematical modeling of interferon- α -treatment for personalized treatment of myeloproliferative neoplasms. *Computational and Systems Oncology*, 1(4):e1030.
- [Tong et al., 2021] Tong, J., Sun, T., Ma, S., Zhao, Y., Ju, M., Gao, Y., Zhu, P., Tan, P., Fu, R., Zhang, A., et al. (2021). Hematopoietic stem cell heterogeneity is linked to the initiation and therapeutic response of myeloproliferative neoplasms. *Cell stem cell*, 28(3):502–513.
- [Trask et al., 2000] Trask, P. C., Esper, P., Riba, M., and Redman, B. (2000). Psychiatric side effects of interferon therapy: prevalence, proposed mechanisms, and future directions. *J Clin Oncol*, 18(11):2316–2326.
- [Vainchenker et al., 2011] Vainchenker, W., Delhommeau, F., Constantinescu, S. N., and Bernard, O. A. (2011). New mutations and pathogenesis of myeloproliferative neoplasms. *Blood*, 118(7):1723–1735.
- [Van Egeren et al., 2021] Van Egeren, D., Escabi, J., Nguyen, M., Liu, S., Reilly, C. R., Patel, S., Kamaz, B., Kalyva, M., DeAngelo, D. J., Galinsky, I., et al. (2021). Reconstructing the lineage histories and differentiation trajectories of individual cancer cells in myeloproliferative neoplasms. *Cell stem cell*, 28(3):514–523.
- [Williams et al., 2022] Williams, N., Lee, J., Mitchell, E., Moore, L., Baxter, E. J., Hewinson, J., Dawson, K. J., Menzies, A., Godfrey, A. L., Green, A. R., et al. (2022). Life histories of myeloproliferative neoplasms inferred from phylogenies. *Nature*, 602(7895):162–168.
- [Xu et al., 1998] Xu, Z., Hoffman, J., Patel, I., and Joubert, P. (1998). Hoffmann-la-roche in. single dose safety/tolerability and pharmacokinetic/pharmacodynamics (pk/pd) following administration of ascending subcutaneous doses of pegylated interferon (pegifn) and interferon alpha2a (ifn alpha2a) to healthy subjects. *Hepatology*, 28.
- [Yamane et al., 2008] Yamane, A., Nakamura, T., Suzuki, H., Ito, M., Ohnishi, Y., Ikeda, Y., and Miyakawa, Y. (2008). Interferon- α 2b-induced thrombocytopenia is caused by inhibition of platelet production but not proliferation and endomitosis in human megakaryocytes. *Blood*, 112(3):542–550.

Acknowledgements

P.-H.C. and I.P. are granted by the Prism project, funded by the Agence Nationale de la Recherche under grant number ANR-18-IBHU-0002. This work was also supported by grants from INCA Plbio2018, 2021 to IP and Ligue Nationale Contre le Cancer (équipe labellisée 2019) and from the Appel a Projets Pre-neoplasies 2021 (C21021LS) with financial support from Institut Thématique Multi-Organismes (ITMO) Cancer of Aviesan within the framework of the 2021-2030 Cancer Control Strategy, on funds administered by INSERM. We thank D. Madhavan for her help in editing the manuscript.

Data Availability Statement

The authors declare that all the data supporting the findings of this study are available within the paper and its Supplemental Data.

Author contributions

G.H., I.P., P.-H.C. participated in research design, G.H., P.-H.C. performed data analysis, G.H., I.P., P.-H.C. wrote the manuscript.

Conflict-of-interest disclosure

The authors declare no competing financial interests.