Costing ‘the’ MTD

David C. Norris

Precision Methodologies, LLC, Seattle WA, USA

Corresponding author:
David C. Norris

Email address: david@precisionmethods.guru

ABSTRACT

**Background:** Absent adaptive, individualized dose-finding in early-phase oncology trials, subsequent registration trials risk suboptimal dosing that compromises statistical power and lowers the probability of technical success (PTS) for the investigational drug. While much methodological progress has been made toward adaptive dose-finding, and quantitative modeling of dose-response relationships, most such work continues to be organized around a concept of ‘the’ maximum tolerated dose (MTD). But a new methodology, Dose Titration Algorithm Tuning (DTAT), now holds forth the promise of individualized ‘MTD’ dosing. Relative to such individualized dosing, current ‘one-size-fits-all’ dosing practices amount to a constraint that imposes costs on society. This paper estimates the magnitude of these costs.

**Methods:** Simulated dose titration as in (Norris 2017) is extended to 1000 subjects, yielding an empirical MTD distribution to which a gamma density is fitted. Individual-level efficacy, in terms of the probability of achieving remission, is assumed to be an E\textsuperscript{max}-type function of dose relative to MTD\textsubscript{i}, scaled (arbitrarily) to identify MTD\textsubscript{i} with the LD\textsubscript{50} of the individual’s tumor. (Thus, a criterion 50% of the population achieve remission under individualized dosing in this analysis.) Current practice is modeled such that all patients receive a first-cycle dose at ‘the’ MTD, and those for whom MTD\textsubscript{i} < MTD\textsubscript{the} experience a ‘dose-limiting toxicity’ (DLT) that aborts subsequent cycles. Therapy thus terminated is assumed to confer no benefit. Individuals for whom MTD\textsubscript{i} ≥ MTD\textsubscript{the} tolerate a full treatment course, and achieve remission with probability determined by the E\textsubscript{max} curve evaluated at MTD\textsubscript{the}/MTD\textsubscript{i}. A closed-form expression is obtained for the population remission rate, and maximized numerically over MTD\textsubscript{the} as a free parameter, thus identifying the best result achievable under one-size-fits-all dosing. A sensitivity analysis is performed, using both a perturbation of the assumed E\textsubscript{max} function, and an antipodal alternative specification.

**Results:** Simulated MTD\textsubscript{i} follow a gamma distribution with shape parameter α ≈ 1.75. The population remission rate under one-size-fits-all dosing at the maximizing value of MTD\textsubscript{the} proves to be a function of the shape parameter—and thus the coefficient of variation (CV)—of the gamma distribution of MTD\textsubscript{i}. Within a plausible range of CV(MTD\textsubscript{i}), one-size-fits-all dosing wastes approximately half of the drug’s population-level efficacy. In the sensitivity analysis, sensitivity to the perturbation proves to be of second order. The alternative exposure-efficacy specification likewise leaves all results intact.

**Conclusions:** The CV of MTD\textsubscript{i} determines the efficacy lost under one-size-fits-all dosing at ‘the’ MTD. Within plausible ranges for this CV, failure to individualize dosing can effectively halve a drug’s value to society. In a competitive environment dominated by regulatory hurdles, this may reduce the value of shareholders’ investment in the drug to zero.

**Keywords**

Economics of drug development, dose-finding studies, oncology, Phase I clinical trial, individualized dose-finding, precision medicine
INTRODUCTION

Dose Titration Algorithm Tuning (DTAT), a new methodology for individualized dose-finding in early-phase oncology studies, holds forth a promise of individualized dosing from the earliest stages of oncology drug development (Norris 2017). Most immediately and obviously, such individualized dosing serves the imperative of individual ethics in seeking to optimize the care of each person who enrolls in a Phase I study. But by increasing the efficiency of drug development overall, individualized dosing also serves wider social aims. Less effective, ‘one-size-fits-all’ dosing may condemn valuable drugs to failure in later registration trials. More efficacious, individualized dosing may therefore avert financial losses to shareholders in pharmaceutical innovation, while preserving innovations valuable to society at large. This brief technical note estimates the magnitude of the social costs incurred by one-size-fits-all dose-finding studies. The argument should be of interest to shareholders in pharmaceutical innovation, and to executives having fiduciary responsibilities to them.

THE DISTRIBUTION OF MTD_i

In (Norris 2017), DTAT was demonstrated by simulated dose titration in 25 simulated subjects drawn randomly from a population model of the pharmacokinetics and myelosuppressive dynamics of docetaxel. By extending this simulation to 1000 subjects, we obtain the empirical distribution of individualized maximum tolerated dose (MTD_i) shown in Figure 1.

![Figure 1. MTD_i is approximately Gamma distributed.](image-url)

Whether the fitted Gamma density in Figure 1 represents a true distribution in any actual human population matters less for what follows than establishing the basic plausibility of a Gamma-distributed MTD_i generally.

DOSE-RESPONSE MODEL

To estimate the cost of sub-MTD_i dosing, one must model individual-level efficacy as a function of dose. A traditional approach in this context is to posit a dose-effect model of a standard ‘E_max’ type. Taking the tumor’s point of view, we may write in fact a ‘toxicology’ form of the model:
\[ P_r(D) = \frac{D}{D + LD_{50}}, \]

where \( P_r(D) \) is the probability of achieving remission as a function of \( D \), the dose received, and \( LD_{50} \) is the dose that would be ‘lethal’ to the tumor in 50% of patients—that is, the dose that would achieve remission with probability 0.5. By supposing further that MTD\(_i\) is the LD\(_{50}\) for the tumor in individual \( i \), we obtain:

\[ P_r = \frac{D}{D + MTD_i} = \left(1 + \frac{MTD_i}{D}\right)^{-1} = \left(1 + \frac{1}{\theta_i}\right)^{-1}. \]

Thus, identifying MTD\(_i\) with the LD\(_{50}\) of the tumor yields a modeled remission probability that is a function of \( \theta_i = D/MTD_i \), the fraction of MTD\(_i\) received. The reasonableness of this identification will be explored in the Discussion below. As it turns out, a slightly different functional form for \( P_r(\theta) \) supports obtaining an intermediate result in terms of standard functions:

\[ P_r(\theta) = \frac{1}{2} \theta^{\frac{1}{2}}. \quad (1) \]

The reader suspicious of this departure from tradition should take reassurance in noting that this revised functional form is uniformly more forgiving of suboptimal dosing than the standard form:

\[ \frac{1}{2} \theta^{\frac{1}{2}} \geq \left(1 + \frac{1}{\theta}\right)^{-1} \text{ for } \theta \geq 0. \]

**THE DISTRIBUTION OF \( \theta_i = MTD_{\text{the}}/MTD_i \)**

If MTD\(_i\) \( \sim \) Gamma\((\alpha, \beta)\), then \( \frac{1}{MTD_i} \sim \text{Inv-Gamma}(\alpha, \beta) \) and consequently

\[ \theta_i = \frac{MTD_{\text{the}}}{MTD_i} \sim \text{Inv-Gamma}(\alpha, \beta \cdot MTD_{\text{the}}). \quad (2) \]

**THE TWO COSTS OF ONE-SIZE-FITS-ALL DOsing**

Under the prevailing practice of one-size-fits-all dosing at ‘the’ MTD, we take the following to occur: (1) Those individuals \( i \) for whom MTD\(_i\) > MTD\(_{\text{the}}\) will receive suboptimal dosing at a fraction \( \theta_i < 1 \) of their optimal dose; (2) those for whom MTD\(_i\) < MTD\(_{\text{the}}\) will experience intolerable adverse effects with a first dose, and will not receive subsequent cycles of therapy. (Those rare individuals for whom MTD\(_i\) = MTD\(_{\text{the}}\) holds exactly will receive their optimal \( \theta_i = 1 \) dose, and enjoy the full benefit of the drug.) Thus, dosing everyone at MTD\(_{\text{the}}\) has two social costs: individuals who cannot tolerate ‘the’ MTD derive no benefit from the drug, while those who could have tolerated higher doses derive suboptimal benefit. This latter cost is well described in a literature stretching back 2 decades, documenting (across many types of cancer) that patients who experience milder adverse effects from chemotherapy tend to have worse outcomes (Saarto et al. 1997, Cameron et al. (2003), Di Maio et al. (2005), Yamanaka et al. (2007), Y. H. Kim et al. (2009), Lee et al. (2011), Shitara et al. (2011), McTiernan et al. (2012), Liu, Zhang, and Li (2013), Shiozawa et al. (2014), Su et al. (2015), Osorio et al. (2017)).
Figure 2 depicts the balance of these costs under 3 different choices of MTD_{the} \in \{100, 200, 300\} mg. If MTD_{i} \sim \text{Gamma}(\alpha = 1.75, \beta = 1/200), then \(\theta_{i} = \frac{\text{MTD}_{\text{the}}}{\text{MTD}_{i}}\) will follow the inverse gamma distribution:

\[\theta_{i} = \frac{\text{MTD}_{\text{the}}}{\text{MTD}_{i}} \sim \text{Inv-Gamma}(\alpha = 1.75, \beta \in \{\frac{100}{200}, \frac{200}{200}, \frac{300}{200}\})\]

These 3 densities are plotted in green in Figure 2, superimposed on the dose-response relationship of Equation 1. Here, it is readily seen that setting MTD_{the} = 100mg causes most individuals to receive doses below half of their MTD_{i}'s (\(\theta < 0.5\)). Conversely, setting MTD_{the} = 300mg causes few individuals to be dosed at \(\theta \leq 0.5\), but excludes a large fraction of the population from treatment—as indicated by the large area under the dashed curve to the right of \(\theta = 1\).

**Figure 2. Social costs of one-size-fits-all dosing at 3 different choices of ‘the’ MTD.** Against the purple dose-response function, the distribution of \(\theta_{i} = \frac{\text{MTD}_{\text{the}}}{\text{MTD}_{i}}\) is plotted for 3 different values of MTD_{the}. When ‘the’ MTD is set low (100 mg), few individuals are excluded from treatment (area under dashed curves), but most are treated at a low fraction (\(\theta < 0.5\)) of their MTD_{i}'s. Conversely, when ‘the’ MTD is set high (300 mg), fewer individuals are dosed so low, but many (large area under dashed curve) cannot tolerate the drug and do not receive a full course of treatment.

**POPULATION-LEVEL EFFICACY OF ONE-SIZE-FITS-ALL DOSING**

Given that \(\theta_{i}\) is distributed as in Equation 2, and that the individual-level probability of remission is as given by Equation 1, then the population rate \(P_{r}\) of achieving remission may be calculated by integrating \(P_{r}(\theta_{i})\) over the treated population \(0 \leq \theta_{i} \leq 1\). Normalizing \(\tilde{\beta} = \beta \cdot \text{MTD}_{\text{the}}\), we can calculate:
\[ \hat{P}_r = \int_0^1 P_r(\theta) \cdot \text{Inv-Gamma}(\theta; \alpha, \hat{\beta}) \, d\theta \]

\[ = \int_0^1 \frac{1}{2} \theta^{1/2} \cdot \frac{\hat{\beta}^\alpha}{\Gamma(\alpha)} \theta^{-\alpha-1} \exp \left( -\frac{\hat{\beta}}{\theta} \right) \, d\theta \]

\[ = \frac{1}{2} \frac{\hat{\beta}^{\frac{\alpha}{2}} \Gamma(\alpha - \frac{1}{2})}{\Gamma(\alpha)} \int_0^1 \frac{\theta^{\alpha - \frac{1}{2}}}{\Gamma(\alpha - \frac{1}{2})} \exp \left( -\frac{\hat{\beta}}{\theta} \right) \, d\theta \]

\[ = \frac{1}{2} \frac{\hat{\beta}^{\frac{\alpha}{2}} \Gamma(\alpha - \frac{1}{2})}{\Gamma(\alpha)} \int_0^1 \text{Inv-Gamma}(\theta; \alpha - \frac{1}{2}, \hat{\beta}) \, d\theta \]

\[ = \frac{1}{2} \frac{\hat{\beta}^{\frac{\alpha}{2}} \Gamma(\alpha - \frac{1}{2})}{\Gamma(\alpha)} Q(\alpha - \frac{1}{2}, \hat{\beta}), \]

where \( Q \) denotes the regularized gamma function.

The best-case population rate of remission is obtained by choosing \( \text{MTD}_{\text{the}} \) optimally:

\[ \hat{P}_r(\alpha) = \max_{\hat{\beta}} \left[ \frac{1}{2} \frac{\hat{\beta}^{\frac{\alpha}{2}} \Gamma(\alpha - \frac{1}{2})}{\Gamma(\alpha)} Q(\alpha - \frac{1}{2}, \hat{\beta}) \right] = \frac{1}{2} \frac{\Gamma(\alpha - \frac{1}{2})}{\Gamma(\alpha)} \max_{\hat{\beta}} \left[ \hat{\beta}^{\frac{\alpha}{2}} Q(\alpha - \frac{1}{2}, \hat{\beta}) \right], \] (3)

in which it should be noted particularly that \( \hat{P}_r \) is a function of the ‘shape parameter’ \( \alpha \), which determines the coefficient of variation (CV) of our gamma-distributed \( \text{MTD}_i \) via \( \text{CV} = \alpha^{-1/2} \).

The maximand on the right-hand side of Equation 3 is readily evaluated using the implementation of the regularized gamma function \( Q \) provided in R package \texttt{zipfR} (Evert and Baroni 2007), and the maximum obtained numerically. The dependence of \( \hat{P}_r \) on \( \text{CV}(\text{MTD}_i) \) is plotted in Figure 3.

![Figure 3](http://example.com/figure3.png)

**Figure 3.** Estimated minimum cost of one-size-fits-all dosing, as a function of the coefficient of variation (CV) of MTD$_i$ in the population. It is assumed that ‘the’ MTD is chosen to maximize the population-level remission rate, under the constraint of one-size-fits-all dosing. The cost of the one-size-fits-all constraint is calculated relative to a reference remission probability of 50% for optimal individualized dosing at each patient’s MTD$_i$. The more MTD$_i$ varies within the population, the more untenable one-size-fits-all dosing becomes.
SENSITIVITY ANALYSIS

One assumption essential to the development of my argument thus far was that individual-level outcomes are a function of $\theta_i = D/\text{MTD}_i$, the fraction of MTD$_i$ received by individual $i$. This assumption would hold in the limiting case where inter-individual variation in MTD$_i$ was driven entirely by pharmacokinetic heterogeneity. (Consider the particularly simple example of an oral drug for which otherwise-identical individuals differed only regarding bioavailability.) Nevertheless, the sensitivity of my results to this assumption does seem to warrant further examination.

Consider the following perturbation of Equation 1:

$$P_r(D, \text{MTD}_i) = \frac{1}{2} \left( \frac{D}{\text{MTD}_i} \right)^{\frac{1}{2}} \left[ 1 + \delta \left( \frac{\text{MTD}_i}{\alpha/\beta} - 1 \right) \right], \quad (4)$$

where the factor in brackets induces a dependence of $P_r$ on MTD$_i$ that is not accounted for by $\theta_i = D/\text{MTD}_i$. This factor is motivated as a first-order Taylor expansion of a general functional dependence, centered on the population mean $E[\text{MTD}_i] = \alpha/\beta$. Not only do we recover Equation 1 in the limit as $\delta \to 0$, but we also preserve independently of $\delta$ the same population-average remission rate of 1/2 under individualized ‘MTD$_i$’ dosing. (To appreciate this latter point, set $D = \text{MTD}_i$ in Equation 4 then take expectations on both sides.)

To obtain $\bar{P}_r$ as previously, we rearrange Equation 4 as follows:

$$\bar{P}_r(D, \text{MTD}_i) = \frac{1}{2} \left\{ (1 - \delta) \left( \frac{D}{\text{MTD}_i} \right)^{\frac{1}{2}} + \delta \frac{D}{\alpha/\beta} \left( \frac{\text{MTD}_i}{D} \right)^{\frac{1}{2}} \right\},$$

and then integrate as before:

$$\bar{P}_r(\text{MTD}_{\text{the}}) = \frac{1}{2} \int_{0}^{\text{MTD}_{\text{the}}} P_r(\text{MTD}_{\text{the}}, \text{MTD}_i) \cdot \text{Inv-Gamma}(\text{MTD}_i; \alpha, \beta) \, d\text{MTD}_i$$

$$= \frac{1}{2} \int_{0}^{1} \left[ (1 - \delta) \theta^{\frac{1}{2}} + \delta \frac{\text{MTD}_{\text{the}}}{\alpha/\beta} \theta^{-\frac{1}{2}} \right] \cdot \text{Inv-Gamma}(\theta; \alpha, \beta) \, d\theta$$

$$= \frac{1}{2} \int_{0}^{1} \left[ (1 - \delta) \theta^{\frac{1}{2}} + \delta \beta^{\frac{1}{2}} \theta^{-\frac{1}{2}} \right] \cdot \text{Inv-Gamma}(\theta; \alpha, \beta) \, d\theta$$

$$= \frac{1}{2} (1 - \delta) \beta^{\frac{3}{2}} \frac{\Gamma(\alpha - \frac{1}{2})}{\Gamma(\alpha)} Q(\alpha - \frac{1}{2}, \bar{\beta}) + \frac{1}{2} \delta \beta^{\frac{3}{2}} \frac{\Gamma(\alpha + \frac{1}{2})}{\Gamma(\alpha)} Q(\alpha + \frac{1}{2}, \bar{\beta})$$

$$= \frac{1}{2} \frac{\beta^{\frac{3}{2}} \Gamma(\alpha - \frac{1}{2})}{\Gamma(\alpha)} Q(\alpha - \frac{1}{2}, \bar{\beta}) \left[ 1 - \frac{\delta}{2\alpha} \right] + \frac{\beta^{\frac{3}{2}} e^{-\beta}}{\Gamma(\alpha)} \left( \frac{\delta}{2\alpha} \right),$$

in which the final step requires applications of the recurrence relations for $\Gamma$ and Q.

The best-case population rate of remission is obtained, as before, by choosing $\text{MTD}_{\text{the}}$ optimally:

$$\bar{P}_r(\alpha) = \max_\beta \left[ \frac{1}{2} \frac{\beta^{\frac{3}{2}} \Gamma(\alpha - \frac{1}{2})}{\Gamma(\alpha)} Q(\alpha - \frac{1}{2}, \bar{\beta}) \left[ 1 - \frac{\delta}{2\alpha} \right] + \frac{\beta^{\frac{3}{2}} e^{-\beta}}{\Gamma(\alpha)} \left( \frac{\delta}{2\alpha} \right) \right]. \quad (5)$$

Equation 5 generalizes Equation 3, and reduces to it on setting $\delta = 0$.  

6/10
Without attempting to develop an interpretation of $\delta$, we plot in Figure 4 the relative efficacy of one-size-fits-all dosing at selected values of $CV$, for $\delta \in [-1, 1]$. Evidently, the perturbation is of second order. Moreover, absolute deviations from the picture of Figure 3 prove small over the range of what seems from Equation 4 to be the natural scale for $\delta$.

![Figure 4. Sensitivity of efficiency loss estimate to perturbation as in Equation 4.](image)

**Sensitivity under an interpretable alternative**

As an alternative to the foregoing perturbation analysis, we might instead posit a single, readily interpretable alternative to Equation 1. A possibility that immediately presents itself as a ‘diametrically opposed’ alternative is to suppose $P_r(D, MTD_i)$ independent of $MTD_i$, as in:

$$P_r^*(D, MTD_i) \equiv P_r^*(D) = \gamma \left( \frac{D}{\alpha/\beta} \right)^{\frac{1}{2}} = \gamma \left( \frac{D}{E[MTD_i]} \right)^{\frac{1}{2}}. \quad (6)$$

Under one-size-fits-all dosing, Equation 6 yields a population-level efficacy of

$$\bar{P}_r^* = \int_D^\infty P_r^*(D) \cdot \text{Gamma}(x; \alpha, \beta) \, dx = \gamma \left( \frac{D}{\alpha/\beta} \right)^{\frac{1}{2}} Q(\alpha, \beta \cdot D). \quad (7)$$

In contrast to the cases considered above, here the integration extends into a region where $P_r > 1$; but, by taking $\gamma \to 0$, we can push this region as far as desired into the tail of our Gamma-distributed $MTD_i$. (As will be seen presently, $\gamma$ drops out of our analysis of relative efficacy.) Differentiating with respect to $D$, and employing the same normalization $\hat{\beta} = \beta \cdot D$ introduced previously, we find that Equation 7 is maximized at $\hat{\beta} = \hat{\beta}$ defined by

$$Q(\alpha, \hat{\beta}) = 2 \frac{\hat{\beta}^\alpha e^{-\hat{\beta}}}{\Gamma(\alpha)}. \quad (8)$$

Substituting (8) into (7), we find that this maximum value is:

$$\hat{P}_r^* = \gamma \left( \frac{\hat{\beta}}{\alpha} \right)^{\frac{1}{2}} 2 \frac{\hat{\beta}^\alpha e^{-\hat{\beta}}}{\Gamma(\alpha)} = 2 \gamma \frac{\hat{\beta}^\alpha + \frac{1}{2} e^{-\hat{\beta}}}{\sqrt{\alpha} \Gamma(\alpha)}. \quad (9)$$
Under the MTD\textsubscript{i}-independent remission probability (6), optimal individualized dosing does not reduce trivially to a constant $1/2$. Rather, we must calculate as follows:

$$
\bar{P}^{*,i}_r = \int_0^\infty P^{*}_r(D) \cdot \text{Gamma}(D; \alpha, \beta) \, dD = \frac{\Gamma(\alpha + \frac{1}{2})}{\sqrt{\pi \Gamma(\alpha)}}.
$$

(10)

Dividing (9) by (10), we obtain the following expression for relative efficacy of one-size-fits-all dosing under Equation 6:

$$
\frac{\bar{P}^{*,i}_r}{P^{*,i}_r} = 2 \hat{\beta}^{\alpha + \frac{1}{2}} e^{-\hat{\beta}}.
$$

(11)

Solving Equation 8 numerically for $\hat{\beta} = \hat{\beta}(\alpha)$, and substituting this into Equation 11, we obtain the following counterpart to Figure 3:

![Graph](image)

---

**DISCUSSION**

At two points in this argument, I have adopted modeling assumptions that are relatively forgiving of one-size-fits-all dosing, and therefore would tend to underestimate its costs. Firstly, my highly concave square-root $E_{\text{max}}$ model (Equation 1) regards under-dosing more favorably than does a typical $E_{\text{max}}$ model, such as appears in the dotted purple curve in Figure 2. Secondly, the optimization itself in Equation 3 surely overestimates the population-level outcomes achieved by one-size-fits-all dosing as implemented in current Phase I designs. Indeed, these designs tend to target DLT rates without explicit reference or regard to outcomes.

Dose reduction protocols, as seen both in trials and in clinical practice, do somewhat relax the extreme form of one-size-fits-all dosing constraint that I have modeled in this paper. Clearly, such protocols exist precisely to recover some part of the lost value I calculate here. But given that these protocols are readily interpreted as a (very) poor man’s DTAT, their existence only underscores the urgent need for rational dose individualization in oncology. This conclusion fully withstands the rather vigorous sensitivity analysis performed above.
CONCLUSIONS

Taking population-level efficacy as a proxy, I have estimated the social cost of one-size-fits-all dosing organized around a concept of ‘the’ maximum tolerated dose (MTD) in oncology. The magnitude of this cost is seen to depend primarily on the coefficient of variation (CV) of individually optimal MTD$_i$ doses in the population. Within plausible ranges for this CV, the failure to individualize dosing can effectively halve a drug’s value to society. Notably, in a competitive environment dominated by regulatory hurdles, this may reduce the value of shareholders’ investment in a drug to zero.

DATA AVAILABILITY

Open Science Framework: Data for Figure 1 may be found in R package DTAT (v0.1-1), available together with code for reproducing all of this paper’s Figures and analyses, at doi: 10.17605/osf.io/vtxwq.

Competing interests

The author operates a scientific and statistical consultancy focused on precision-medicine methodologies such as those advanced in this article.

Grant information

The author declared that no grants were involved in supporting this work.
REFERENCES


