MATHEMATICAL BIOSCIENCES
AND ENGINEERING
Volume **7**, Number **4**, October **2010**

doi:10.3934/mbe.2010.7.923

pp. **923–933**

ANTIBIOTIC CYCLING VERSUS MIXING: THE DIFFICULTY OF USING MATHEMATICAL MODELS TO DEFINITIVELY QUANTIFY THEIR RELATIVE MERITS

Robert E. Beardmore and Rafael Peña-Miller

Department of Mathematics, Imperial College London
SW7 2AZ, London, UK

(Communicated by Yang Kuang)

Abstract. We ask the question *Which antibiotic deployment protocols select best against drug-resistant microbes: mixing or periodic cycling?* and demonstrate that the statistical distribution of the performances of both sets of protocols, mixing and periodic cycling, must have overlapping supports. In other words, it is a general, mathematical result that there must be mixing policies that outperform cycling policies and vice versa.

As a result, we agree with the tenet of *Bonhoefer et al.* [1] that one should not apply the results of [2] to conclude that an antibiotic cycling policy that implements cycles of drug restriction and prioritisation on an *ad-hoc* basis can select against drug-resistant microbial pathogens in a clinical setting any better than random drug use. However, nor should we conclude that a random, per-patient drug-assignment protocol is the *de facto* optimal method for allocating antibiotics to patients in any general sense.

1. **Background.** The last decade or so has produced a number of theoretical approaches towards increasing our understanding of how to best deploy antibiotics in order to select against resistant pathogens in clinical settings and [3, 4] are two of the most important papers in this field. We too are interested in this important problem from a theoretical perspective. In a previous paper [2], motivated by [3, 4], we sought to address the search for optimal drug deployment protocols from the perspective of optimal control theory.

This article is an attempt to understand a possible source of the discrepancy between what is claimed in the letter [1] and in our article [2]. It is not our intention here or elsewhere to advocate on behalf of cycling antibiotics in clinical contexts or even in theory, rather it is to highlight the difficulties of definitively answering the question of which protocol class selects best against resistance, cycling or mixing.

The purpose of [2] was to show that by relaxing the periodicity constraint of antibiotic cycling, it is *theoretically* possible to design drug rotation protocols that are arbitrarily close to the optimum. Therefore, we agree with [1] that there is absolutely no theoretical basis to support the claim that periodic cycling is optimal in any general sense and we too are concerned that a reader might be led to that conclusion. We also argue in [2], however, that strategies based on the random,

2000 *Mathematics Subject Classification.* 93C15, 92B05.
Key words and phrases. Epidemiology, antibiotic cycling, antibiotic mixing, drug resistance.
REB is supported by an MRC fellowship and RPM by CONACYT-Mexico.

per-patient allocation of drugs can only be optimal in numerically rare, non-generic theoretical cases.

To illustrate that 50-50 mixing is not *always* superior to antibiotic cycling as claimed in [3, 4], in [2] we constructed counterexamples whereby a specific period of rotation between drugs from different functional classes and with different costs of resistance outperformed even the best mixing protocol. We did not claim that *any* periodic cycling protocol could outperform *all* antibiotic mixing protocols.

What we can assert, and prove in this paper, is a general property of theoretical models of drug deployment as proposed in [3, 4] that, mathematically speaking, the best periodic cycling protocols are as good as the best mixing strategies. More precisely, we prove that the statistical support of the distributions of performances of both classes of protocol must overlap and believe that this property might be invoked as a tentative explanation of the inconclusiveness of recent clinical trials designed to evaluate the efficacy of cycling [5].

1.1. **Scientific context.** Consider the following clinical scenario. An intensive care unit or hospital ward has a limitless supply of two antibiotic drugs each of which may be used to treat a pathogenic infection, the pathogen can acquire resistance to either drug, but not both simultaneously and the drugs cannot be given as part of a combination treatment. Let us choose a period of observation of the ward, we call this time T and in the absence of any synergy between model and data, we refrain from defining its units. It is our goal to maintain and indeed maximise the efficacy of the drugs in our possession over this period and within this scenario we ask the following question.

> *Do the protocols that select optimally against drug-resistant pathogens 'rotate' between different antibiotics, never treating two patients simultaneously with different drugs*?

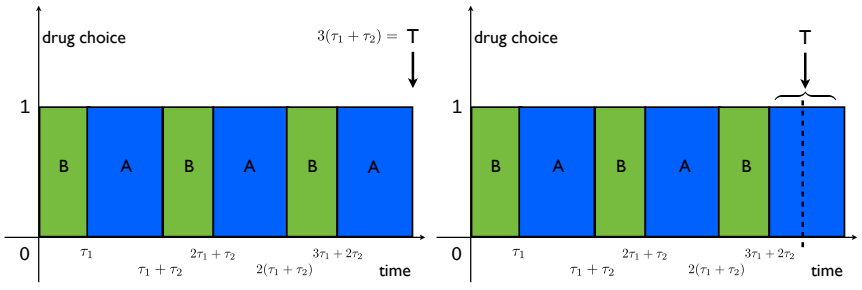
This question is close in spirit to one posed by Niederman [6] in the wake of a clinical trial [11] where the outcome of a drug rotation policy, at least in that trial, had been positive. An inspection of the clinical trials literature since that time reports the body of studies on the performance of antibiotic rotation as *inconclusive* [5].

1.2. **Definition of rotation.** For the purposes of this paper we define a *cycling protocol* between the two antibiotics, labelled 'drug A' and 'drug B', as follows. Two parameters, τ_1 and τ_2 are used to define the periodic cycling protocol, a function $A(\tau_1, \tau_2)(t)$, defined for all t within the time interval $[0, T]$ such that

$$A(\tau_1,\tau_2)(t) = \begin{cases} 1 & : \quad 0 \le t \le \tau_1, \\ 0 & : \quad \tau_1 < t \le \tau_1 + \tau_2, \end{cases}$$

and extended so that $A(\tau_1, \tau_2)$ has period $\tau_1 + \tau_2$. The indicator function $A(\tau_1, \tau_2)$ so-defined tells us when to treat *everyone* with drug A and hence the number '1' in the definition of $A(\tau_1, \tau_2)$. If $A(\tau_1, \tau_2)$ is zero, everyone is treated with drug B.

This, we believe, is different to the definition of what constitutes a cycling protocol in [1] whereby it appears that parameters are restricted so that, in our notation, $T/(\tau_1 + \tau_2)$ and $\tau_1 + \tau_2$ are both integers. As a result, we shall further restrict the definition of $A(\tau_1, \tau_2)$ and attach the adjective *congruent* to describe cycling protocols such that $\tau_1 + \tau_2$ is the (real-valued) period of cycling and $T/(\tau_1 + \tau_2)$ is an integer. The set of all congruent cycling protocols defined in this way will be denoted Cyc, a larger set than considered in [1].



(A) (B)

FIGURE 1. (A) A congruent cycling protocol whereby $T = 3(\tau_1 + \tau_2)$. (B) A non-congruent cycling protocol: it is possible that the observation time, T, could fall mid-cycle for the best protocols; this possibility is not analysed here.

It is our contention that non-congruent cycles should also be analysed. When a clinical trial is designed and a period of observation, T, chosen, there is no reason to expect that the optimal cycling protocol will turn out be related to the value of T in a manner imposed *a priori* by the observer. To artificially constrain the period of cycling to the interval of observation might well miss the best cycling cadences; how might we ascertain, *a priori*, that the correct cadence for the exchange of different antibiotics over an observation period of 365 days is not 27 days? Nevertheless, it is sufficient for our purposes to restrict attention to congruent cycles in the remainder.

For brevity we do not consider the effect of clinician compliance on our results, although this is done in [3]; this would require a straightforward re-definition of $A(\tau_1, \tau_2)$ to permit it to take on values above 0% and below 100%.

1.3. **Optimal cycling and mixing: Two impractical protocols.** The criticism in [1] that protocols determined using optimal control theory cannot be implemented in practise is valid. While the randomised allocation of drugs in a population (or 50-50 mixing) described in [3, 4] are practicable (also see [16]), we argue that the optimal mixing protocol defined in [1] as *optimally dividing the drugs between appropriately sized patient groups* is also an impractical, theoretical concept.

Moreover, even if it may be possible in theory to determine the best period of cycling, the best mixing proportion or even the optimal protocol, doing so certainly does not answer the much harder question

> *What is the likelihood that a given rotation will outperform a given mixing protocol?*

The difficulties in providing a meaningful answer to this are not to be underestimated. One would need a mathematical model of the epidemiological dynamics at play in the hospital capable of making hindcasts and predictions calibrated against known data. This would be akin to predicting the weather, only potentially harder because of the absence of physical laws. Attempts have been made at using simple mathematical models to produce forecasts of antibiotic resistance evolution over the coming decades [9], but criticism of the lack of biological depth followed [8]. The best one can ever hope for is to provide an ensemble of possible futures, as in [7].

As a result, it would be churlish to assert, on the basis of any one epidemiological model, that a specific protocol achieves optimality in real, clinical contexts or even just for a large class of mathematical models.

The purpose of [2] is to provide general results, really little more than a reformulation of known mathematical ideas, which are subsequently applied to two specific mathematical models taken from [3, 4]. These applications illustrate that a wide variety of conditions are met whereby the best non-periodic rotations can outperform the best mixing protocol and such conditions are not rare but arise when so-called *drug symmetries* are broken. The negative answer to Brown and Nathwani's question posed in the clinical review [5], *Does resistance to different antibiotics develop at the same rate?* provides one example of such an asymmetry.

The difficulty of determining optimal treatment protocols in practice is real and so we proposed adaptive rotation protocols as a practicable strategy that relies only on observations of local patterns of susceptibility and resistance to decide which drug, A or B, to give [2]. Note, we *do not* claim that our adaptive feedback rule is anywhere near to being optimal nor a good approximation for such. It is a merely a tool for determining rotational protocols that outperform optimal mixing and we find, by example, that even this simple heuristic can achieve this goal for particular exemplar models. We do claim that the adaptive protocols defined in [2] that seek to exploit more information by sampling the patient cohort for drug-resistant pathogens more frequently, in turn, perform better. We give examples to show that such protocols can, given enough information, outperform optimal mixing; [1] does not provide any evidence that this statement is false. Furthermore, it is absolutely true that the difference in performance between optimal rotation and optimal mixing found in models may be marginal, but this difference will be specific to the parameters implemented within any particular model and specific to the model itself.

We would like to emphasize that the analysis provided in [2] is valid for a large class of epidemiological models[1] and independent of any particular parameter values or initial conditions. In particular, our only assumption over the interval of observation, T, is that it has to be finite, for our arguments to work it cannot be infinite. For this reason, we disagree with [1] that the arguments presented in [2] are only valid in general for short intervals of observation and are solely due to 'transients'. One must also bear in mind that as T is finite it is also a parameter in the problem and that optimal mixing, optimal cycling and the theoretically optimal protocol may well all depend on T itself. Hence there is no general rationale to support the idea that an optimal protocol found at $T = 50$, say, will also be optimal when extended to $T = 51$, let alone to $T = \infty$.

Instead of seeking for a particular set of parameters where periodic cycling outperforms mixing in specific models, in the following section we will provide statistical arguments that may help understand the difficulties that arise when trying to quantify the differences between cycling and mixing, and we will demonstrate that it is always possible to find cycling protocols that outperform mixing, and vice versa.

[1] There are many antibiotic deployment problems that one could pose outside of the mathematical class that we analyse in this article, but this class does encompass some of the models presented in [4, 3].

2. **Cycling versus mixing: A theorem concerning their statistics.** Rather than delve into the details of any particular epidemiological scenario or mathematical model, we shall keep this analysis general and work for the moment in an abstract framework. First of all, we shall need the vector of all the different patient classes. Following previous approaches we could write

$$\boldsymbol{y}(t) = (S(t), R_A(t), R_B(t), X(t)),$$

say, where each entry represents a time-series of the proportion of patients in different classes: drug-susceptible bacterial infections, drug-A resistant infections, drug-B resistant infections and uncolonised patients, respectively.

In principle we would need a mathematical model to describe how the vector $\boldsymbol{y}(t)$ changes in time. Let us assume that the hospital will operate at its maximal capacity at all times so that the sum of elements in $\boldsymbol{y}(t)$ must be a constant value. We may assume without the loss of any generality that this value is unity in the remainder, hence the entries of $\boldsymbol{y}(t)$ will contain the relative proportions of each of the different patient classes.

Let us now suppose the existence of a dynamical model, specifying none of its detail, of the form

$$\frac{d}{dt}\boldsymbol{y} = F(\boldsymbol{y}) + A(t)G_a(\boldsymbol{y}) + B(t)G_b(\boldsymbol{y}); \tag{1}$$

the models in [4, 3] are of this form. In (1) the functions G_a and G_b describe how the dynamics of the ICU unit or hospital ward are affected when we use one of the drugs, drug A or drug B. For example the use of drug A might well be positively correlated with selection for drug A-resistant pathogens, thus increasing the value of $\frac{d}{dt}R_A(t)$. For equation (1) to make any sense we also need an initial condition. We could say $\boldsymbol{y}(0) = (1, 0, 0, 0)$ so that all patients are infected with the drug-susceptible pathogen to begin with, but with no drug-resistant pathogens. Clearly there are many other choices that could also be used and we will not restrict attention to any one in particular. Let us also suppose that policy ensures that everyone in the hospital is treated, according to our assumptions this means that $A(t) + B(t) = 1$ for all times.

Under mild conditions, the resulting equation

$$\frac{d}{dt}\boldsymbol{y} = F(\boldsymbol{y}) + A(t)G_a(\boldsymbol{y}) + (1 - A(t))G_b(\boldsymbol{y}). \tag{2}$$

can be shown to define a mathematical dependancy or mapping in the sense that to each function $A(t)$ representing a drug-deployment policy, we can associate a solution $\boldsymbol{y}(t)$ that depends on A; we will write $\boldsymbol{y}(A)$ or $\boldsymbol{y}(A)(t)$ for this dependancy. Provided a natural technical assumption given in [2] holds, namely that there exists a constant K independent of A such that for any solution of (2) there results

$$\limsup_{t\to\infty} \|\boldsymbol{y}(t)\| \leq K, \tag{3}$$

the vector-valued function $\boldsymbol{y}(A)(t)$ can be extended to the entire interval $[0, T]$. As we tacitly assumed earlier in this discussion that the entries of $\boldsymbol{y}(A)(t)$ sum to unity for all t and for any bounded function A between 0 and 1, we may assume (3) holds throughout our discussion (see appendix for notation).

We want to decide upon a drug deployment policy, $A(t)$, that minimises some performance criterion. So, we now define a weight vector $\boldsymbol{w}$ that gives the relative importance of each component of $\boldsymbol{y}(A)(t)$. A mathematical optimisation procedure can determine the optimal policies by locating a function or family of functions $A(t)$

between 0 and 1 such that the measure of the performance of the protocol A defined by

$$\mathcal{P}(A) := \frac{1}{T}\int_0^T \boldsymbol{w}\cdot\boldsymbol{y}(A)(t)dt \tag{4}$$

is minimal.

So, for example, if $\boldsymbol{w} = (0,1,1,0)$, then the performance of our policy is measured by the number of observed drug-resistant infections per unit time: $\boldsymbol{w}\cdot\boldsymbol{y}(t) = R_A(t) + R_B(t)$; throughout the remainder, the latter will be our performance criterion. Likewise, if $\boldsymbol{w} = (0,0,0,-1)$ then $\boldsymbol{w}\cdot\boldsymbol{y}(t) = -X(t)$ and by minimising $\mathcal{P}(A)$ we can *maximise* the number of uncolonised patients seen over the observation period.

2.1. **Which is optimal: Mixing or cycling?** Notice that the performance measure can be thought of as a functional: to each protocol $A(t)$, a function defined on the interval $[0,T]$ taking values in the interval $[0,1]$, we can associate the value $\mathcal{P}(A)$. Our performance measure $\mathcal{P}(\cdot)$ makes sense as a mapping when we supply it with a domain, so let us define the following spaces of functions:

Let Per denote the set of all periodic functions bounded between 0 and 1, defined on the observation interval $[0,T]$ and with period T/n for some integer $n \geq 1$; these functions are so constrained because we cannot possibly treat more than all the patients in the hospital or ward. Let BB be the set of functions bounded between 0 and 1 such that for all t between 0 and T one of either $\alpha(t) = 0$ or $\alpha(t) = 1$ must be true, where α represents any function in BB; thus BB represents the set of protocols that only deploy one drug at a time. Now define $\mathsf{Cyc} := \mathsf{BB} \cap \mathsf{Per}$ which is the set of all congruent, periodic cycling protocols. The set of mixing protocols, Mix, is much smaller and defined by the constant functions, so that a mixing protocol α satisfies $\alpha(t) = c$ for $0 \leq t \leq T$ and some constant c between 0 and 1; the value $c = 1/2$ corresponds to the so-called '50-50' or random mixing protocol.

The following mathematical results are key to our argument.

Corollary 1. *The best possible performance of the set of congruent antibiotic cycling protocols is at least as good as the best performance of the mixings in the sense that*

$$\inf_{A\in\mathsf{Cyc}} \mathcal{P}(A) \leq \inf_{A\in\mathsf{Mix}} \mathcal{P}(A).$$

Proof. Suppose, seeking a contradiction, that $\inf_{A\in\mathsf{Cyc}} \mathcal{P}(A) > \inf_{A\in\mathsf{Mix}} \mathcal{P}(A) = \mathcal{P}(m)$ where $m \in [0,1]$ is a constant. However, by Theorem B.1 of the appendix there is a sequence $c_n \overset{*}{\rightharpoonup} m$ as $n \to \infty$ and therefore

$$\mathcal{P}(m) = \lim_{n\to\infty} \mathcal{P}(c_n) \geq \inf_{A\in\mathsf{Cyc}} \mathcal{P}(A) > \inf_{A\in\mathsf{Mix}} \mathcal{P}(A) = \mathcal{P}(m)$$

because $\mathcal{P}$ is continuous with respect to weak* convergence. This is a contradiction and the result follows. □

Let us examine more closely what this mathematical result means theoretically for our ambiguously phrased question of whether 'cycling or mixing is best'. Firstly, because of the form of $\mathcal{P}$ (it is continuous with respect to weak* convergence) a type of robustness with respect to the cycles follows: if there is at least one cycling protocol that outperforms optimal mixing, there are infinitely many nearby cycling and acyclical drug rotation protocols that also outperform optimal mixing.

Secondly, if the optimal mixing protocol, m say, performs better than all the congruent cycling protocols and yet $\epsilon > 0$ is any number giving a measure of sub-optimality, then there is a congruent cycling policy A_ϵ^{cyc} and a corresponding state response $\boldsymbol{y}(A_\epsilon^{cyc})$ such that

$$\mathcal{P}(m) + \epsilon \ > \ \mathcal{P}(\boldsymbol{y}(A_\epsilon^{cyc})).$$

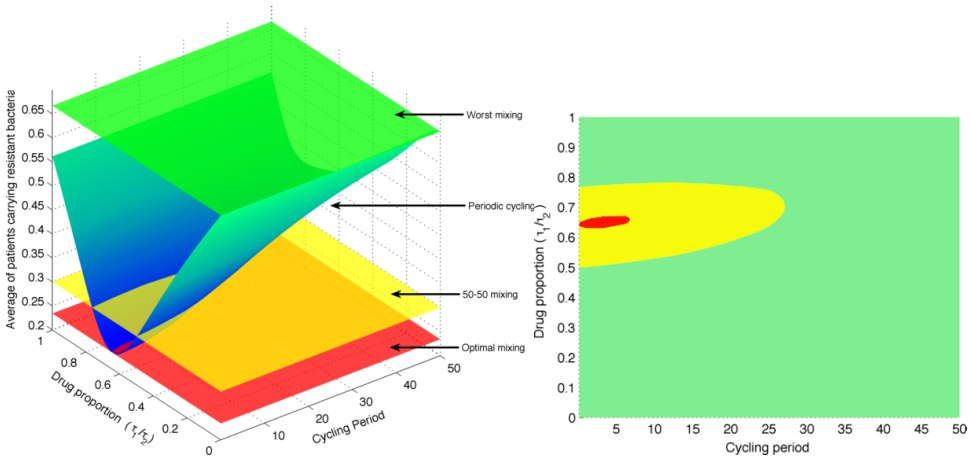
This tells us that if the optimal antibiotic mixing performs better than all the cycling protocols in our model, if we want a cycling policy that performs within 99% or 99.9% of the performance of the best mixing, there are cycling policies in Cyc that will achieve this.

From the perspective of seeking theoretical support for antibiotic cycling this appears to represent a positive outcome but, particularly in light of the accompanying critique [1], we will discuss the following result. We believe it may have important practical implications.

Corollary 2. *The worst possible performance of the set of congruent antibiotic cycling protocols is at least as bad as the worst performance of the mixings in the sense that*

$$\sup_{A \in \mathsf{Cyc}} \mathcal{P}(A) \geq \sup_{A \in \mathsf{Mix}} \mathcal{P}(A).$$

Proof. The proof is almost identical to Corollary 1. □



(A) (B)

FIGURE 2. (A) The performance measure $\mathcal{P}$ defines a performance surface formed from a sample of congruent cycling protocols (shown as a blue-green surface); red shows the performance of optimal mixing ($\underline{p}$), yellow of 50-50 mixing and green the worst possible mixing ($\overline{p}$). (B) The red region shows the neighbourhood from (A) where the best cycling protocols outperform the best mixing, the yellow region shows where cycling outperforms 50-50 mixing. In both (A) and (B) we used $T = 50$.

Let us now define the performances of the best and worst mixing protocols:

$$\underline{p} := \min_{0 \leq m \leq 1} \mathcal{P}(m) \qquad \text{and} \qquad \overline{p} := \max_{0 \leq m \leq 1} \mathcal{P}(m).$$

If we draw the graph of the performance surface $\mathcal{P}(A(\tau_1, \tau_2))$ as a function of (τ_1, τ_2) then Corollaries 1 and 2 predict that this two-dimensional graph will either just touch, or lie above $\overline{p}$ in one region of the (τ_1, τ_2) plane. The same theory predicts it will also either just touch, or lie below $\underline{p}$ in another region of the plane. Figure 2 illustrates a particular instance of this idea, it was obtained using a model from [3] whereby the performance surface of the cycling protocols does in fact pass below the performance of the optimal mixing protocol (by around 1% or less); Figure 2(b) illustrates where this property is satisfied in the (τ_1, τ_2)-plane.

Let us phrase this slightly differently: for each performance measure $p \in (\underline{p}, \overline{p})$ lying anywhere between the extreme performances of the best and worst mixing protocols, there is a congruent cycling protocol defined by (τ_1, τ_2) such that $p = \mathcal{P}(A(\tau_1, \tau_2))$. From this we deduce what might be called an interlacing property: for any model of the form (2) there exist congruent cycling protocols, labelled (τ_{11}, τ_{12}) and (τ_{21}, τ_{22}), and two mixing protocols, m_1 and m_2, such that

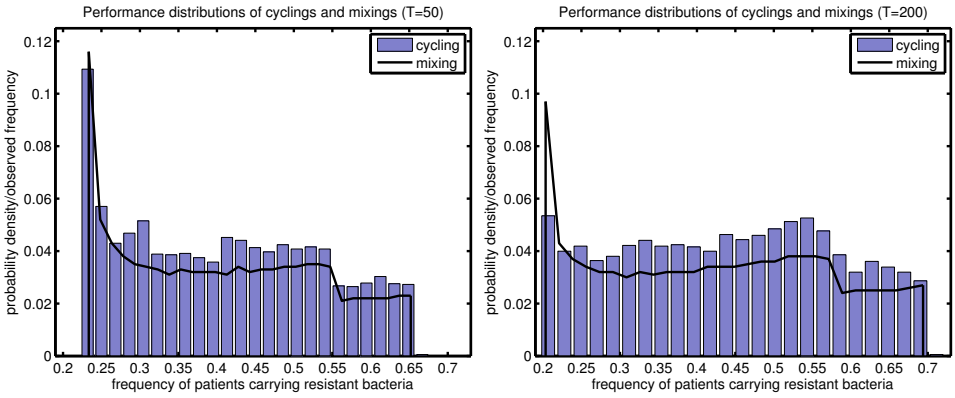
$$\mathcal{P}(m_1) < \mathcal{P}(A(\tau_{11}, \tau_{12})) < \mathcal{P}(m_2) < \mathcal{P}(A(\tau_{21}, \tau_{22})). \tag{5}$$

In light of this, when we ask *Which is optimal: mixing or cycling?* it is impossible to answer unless further clarification is given as to what meaning is really intended by that question. In more statistical language, (5) means that the support of the two distributions of all possible performances of the mixing protocols and the congruent cycling protocols *must* overlap, with the distribution of performances of the cyclings having the potentially larger support because of Corollaries 1 and 2.

This universal result within the model class (2) points to one possible reason for the inconclusive nature of clinical trials performed over the last decade when evaluating the efficacy of cycling. If such a trial were to be mimicked computationally by simulating a model such as (2), if model parameters are sampled sufficiently widely one must find an inequality like (5) within the simulated data. Indeed, this property is observed in Figure 3 computed using a model from [3] where one can clearly see that the empirical distributions of performances of cycling and mixing protocols have almost identical supports.

3. **Conclusion.** It is our contention that the search for effective antibiotic usage strategies should not be reduced to a straightforward comparison between mixing and rotation protocols. More recent treatment paradigms place the individual patient at the core of the treatment and, for example, seek to maximise the appropriateness of the drugs prescribed while minimising the duration of empirical therapy [13]. With these goals in mind, the development of DNA-based rapid sequencing tools that aim to reduce the delay in obtaining microbiology results, as discussed in [14], may prove to be a successful strategy. However, prior models on which this article is based have not accounted for treatment rationales focused at the scale of the individual host and say nothing about how the evolution and spread of pathogens is mediated by those treatments.

In the interests of completeness we highlight the fact that prior mathematical models designed to compare mixing and cycling strategies do not predict, in a general sense that goes beyond particular exemplar models, that protocols allocating different antibiotics randomly on a per-patient basis is the best method for deploying those drugs. Despite this, clinical trials designed to evaluate the efficacy of protocols that increase drug heterogeneity across both hospitals and surgical wards, a property associated with random mixing, have had both beneficial [16] and insignificant outcomes [15] in relation to prior protocols.



(A) (B)

FIGURE 3. Two histograms comparing the distribution of performances of mixing protocols and a sample of the congruent cycling protocols determined using the mathematical model given in [3]; the same parameters are used in both plots except that (A) T = 50 and (B) T= 200 time units. Note, consistent with the theoretical results given in the text, the support of the mixing distribution lies either on or within the cycling distribution in both cases.

We certainly agree with *Bonhoeffer et al.* that the last word on the nature of the theoretically optimal antibiotic deployment protocol has not been spoken. This is a difficult theoretical and practical problem and we hope that no avenue of investigation will be set aside when searching for measures to combat the scourge of antibiotic resistance.

REFERENCES

[1] S. Bonhoeffer, P. S. Zur Wiesch and R. D. Kouyos, *Rotating antibiotics does not minimize selection for resistance*, MBE, **7** (2010), 919–922.

[2] R. E. Beardmore and R. Peña-Miller, *Rotating antibiotics selects optimally against antibiotic resistance, in theory*, MBE, **7** (2010), 527–552.

[3] C. T. Bergstrom, M. Lo and M. Lipsitch, *Ecological theory suggests that antimicrobial cycling will not reduce antimicrobial resistance in hospitals*, PNAS, **101** (2004), 13285–13290.

[4] S. Bonhoeffer, M. Lipsitch and B. R. Levin, *Evaluating treatment protocols to prevent antibiotic resistance*, PNAS, **94** (1997), 12106–12111.

[5] E. M. Brown and D. Nathwani, *Antibiotic cycling or rotation: A systematic review of the evidence of efficacy*, J. Antimicrob. Chemother., **55** (2005), 6–9.

[6] M. Niederman, *Is 'crop rotation' of antibiotics the solution to a 'resistant' problem in the ICU?*, Am. J. Respir. Crit. Care Med., **156** (1997), 1029–1031.

[7] T. Cohen and M. Murray, *Modeling epidemics of multidrug-resistant M. tuberculosis of heterogeneous fitness*, Nat. Med., **10** (2004), 1117–1121.

[8] P. Huovinen, *Mathematical model tell us the future!*, J. Antimicrob. Chemother., **56** (2005), 257–258.

[9] J. T. Magee, *The resistance ratchet: Theoretical implications of cyclic selection pressure*, J. Antimicrob. Chemother., **56** (2005), 427–430.

[10] J. A. Martinez, J. M. Nicolas, F. Marco, J. P. Horcajada, G. Garcia-Segarra, A. Trilla, C. Codina, A. Torres and J. Mensa, *Comparison of antimicrobial cycling and mixing strategies in two medical intensive care units*, Crit. Care Med., **34** (2006), 329–336.

[11] M. H. Kollef, J. Vlasnik, L. Sharpless, C. Pasque, D. Murphy and V. Fraser, *Scheduled change of antibiotic classes: A strategy to decrease the incidence of ventilator-associated pneumonia*, Am. J. Respir. Crit. Care Med., **156** (1997), 1040–1048.
[12] D. Lukkassen and P. Wall, *On weak convergence of locally periodic functions*, J. Nonlinear Math. Phys, **9** (2002), 47–57, arXiv:math/0210293v1.
[13] R. G. Masterton, *The new treatment paradigm and the role of carbapenems*, Int. J. Antimicrob. Agents, **33** (2009), 105–110.
[14] M. G. Bergeron, *Revolutionizing the practice of medicine through rapid (¡ 1h) DNA-based diagnostics*, Clin. Invest. Med., **31** (2008), E265–E271.
[15] Y. Takesue, H. Ohge, M. Sakashita, T. Sudo, Y. Murakami, K. Uemura and T. Sueda, *Effect of antibiotic heterogeneity on the development of infections with antibiotic-resistant gram-negative organisms in a non-intensive care unit surgical ward*, World J. Surg., **30** (2006), 1269–1276.
[16] Y. Takesue, K. Nakajima, K. Ichiki, M. Ishihara, Y. Wada, Y. Takahashi, T. Tsuchida, T and H. Ikeuchi, *Impact of a hospital-wide programme of heterogeneous antibiotic use on the development of antibiotic-resistant Gram-negative bacteria*, J. Hosp. Infect., **75** (2010), 28–32.

Received September 1, 2010; Accepted September 12, 2010.

Appendix A. **Notation.** Throughout, a dot $(\cdot)$ denotes the usual Euclidean dot product and we refer throughout to the following spaces, defined informally within the text. Here $T > 0$ is a fixed constant and, as is standard, *a.e.* means 'Lebesgue almost everywhere'. If we begin with the space of T-periodic functions

$$\mathsf{Per}_T := \{u \in L^\infty((-\infty,\infty),\mathbb{R}^n) : \ u(t+T) = u(t) \text{ for } a.e.\ t \in (-\infty,\infty)\},$$

we can define a space of T/n-periodic functions, for integers $n \geq 1$,

$$\mathsf{Per}_{n,T} := \{v(t) \in L^\infty([0,T],\mathbb{R}^n) : v(t) = u(nt) \text{ for some } u \in \mathsf{Per}_T, \text{ some } n \geq 1 \text{ and for all } t \in [0,T]\},$$

and finally $\mathsf{Per} := \bigcup_{n\geq 1} \mathsf{Per}_{n,T}$. Now define

- $\mathsf{BB} := \{u \in L^\infty([0,T],\mathbb{R}^n) : \text{such that } u(t) \in \{0,1\} \text{ for } a.e.\ t \in [0,T]\}$,
- $\mathsf{Cyc} := \mathsf{BB} \cap \mathsf{Per}$ which is the set of congruent cycling protocols and
- $\mathsf{Mix} := \{u \in L^\infty([0,T],\mathbb{R}^n) : \exists c \in [0,1] \text{ such that } u(t) = c \text{ for } a.e.\ t \in [0,T]\}$.

Note a slight oddity according to these definitions: the functions $A(t) = 0$ and $A(t) = 1$ whereby only one of the drugs are used are both designated as cycling *and* mixing protocols.

Appendix B. **An essential mathematical result.** In common with standard terminology, we say that a sequence of functions $(u_n) \subset L^\infty([0,T],\mathbb{R}^n)$ converges weak* to u if

$$\lim_{n\to\infty} \int_0^T \varphi(t)\cdot u_n(t)dt = \int_0^T \varphi(t)\cdot u(t)dt$$

for all $\varphi \in L^1([0,T],\mathbb{R}^n)$ and we write $u_n \overset{*}{\rightharpoonup} u$ as $n \to \infty$.

Assume that $F(\cdot), G_a(\cdot)$ and $G_b(\cdot)$ are smooth functions and that there is a K independent of A such that for any solution of (2), there results (in any finite dimensional norm) $\limsup_{t\to\infty} \|\boldsymbol{y}(t)\| \leq K$.

Theorem B.1. *The mapping $\mathcal{P} : L^\infty([0,T],\mathbb{R}^n) \to \mathbb{R}$ defined in (4) is a continuous functional with respect to weak* convergence and for each $m \in \mathsf{Mix}$ there is*

a sequence $(c_n) \subset \mathsf{Cyc}$ *such that* $c_n \overset{*}{\rightharpoonup} m$ *as* $n \to \infty$. *There is an optimal mixing protocol and a worst mixing protocol, namely values* $\underline{m}, \overline{m} \in [0,1]$ *such that*

$$\mathcal{P}(\underline{m}) = \min_{0 \le m \le 1} \mathcal{P}(m) \qquad \text{and} \qquad \mathcal{P}(\overline{m}) = \max_{0 \le m \le 1} \mathcal{P}(m).$$

Proof. The first part follows almost immediately from, for example, [12, Theorem 1], the weak* continuity of $\mathcal{P}(\cdot)$ with respect to $A \in L^\infty([0,T], \mathbb{R}^n)$ stemming from the form of (2) as a smooth differential equation defined affinely with respect to A. The last part follows because the map $m \mapsto \mathcal{P}(m)$ is a continuous function defined on the interval $[0,1]$ and so achieves its extreme values. □

Appendix C. **Construction of Figure 3.** To generate Figure 3 we constructed an approximation of the probability density function of the performances of mixing protocols numerically directly from the performance locus $\{\mathcal{P}(m) : 0 \le m \le 1\}$. For the cycling protocols we first defined a set of *periods* $\rho := T/N$ where $N = 1, ..., 30$ are integers, we then defined $\tau_1 := t\rho$ and $\tau_2 := (1-t)\rho$ where $t = 0/M, 1/M, 2/M, ..., M/M = 1$, so that $\tau_1 + \tau_2 = \rho$, we finally set $M = 100$. The performances of the 3,030 different congruent cycling protocols: $\mathcal{D} := \{\mathcal{P}(A(\tau_1, \tau_2)) : \tau_1 = t\rho, \tau_2 = (1-t)\rho,\ \rho = T/N, N = 1, ..., 30, t = 0, 1/100, ..., 1\}$ were then binned (using 25 bins) with the `histc` command in Matlab. If the choice of time units as *days* and other parameter values given in [3, 1] could be justified from empirical data, some of the cycling protocols used to form $\mathcal{D}$ could be eliminated as clinically impractical.

Received September 1, 2010; Accepted September 12, 2010.

E-mail address: r.beardmore@imperial.ac.uk
E-mail address: r.penamiller07@imperial.ac.uk