Dear Recommender and Reviewers,

first of all I apologize for the delay in resubmitting this manuscript, which is due to the fact that my team has been working intensively on the COVID-19 epidemics over the last months (http://covid-ete.ouvaton.org/index_en.html).

Based the reviewers' numerous suggestions, I was encouraged to perform additional analyses and resubmit the manuscript. In particular, I extended the adaptive dynamics section to further explore long-term effects of having treatment rate depend on infection virulence. I also tried to improve the general clarity of the writing and added a summary table to make the mathematical notations more explicit.

Finally, I tried to address the somehow philosophical issue regarding the applicability of the Price equation approach. This approach is more recent in the field than adaptative dynamics. However, I do think that this approach is as useful to study short term dynamics, as adaptative dynamics is to study long-term dynamics and hope this is now better explained in the manuscript.

I thank you in advance for your time and for your consideration.

Best regards,
- Samuel

Response to Recommender (Ludek Berec)

Three reviewers evaluated the manuscript and all agree that the study is interesting. On the other hand, all raise a number of comments that require a revision of the manuscript. These comments vary and include a request to (1) provide a more detailed biological ground for the model, (2) perform some sensitivity analysis with respect to model parameters as well as to motivate the selected parameter values, (3) provide a deeper justification of the modelling choices. I agree with all these comments.

The central aspect of the study appears to be that more virulent infections are treated more. However, I miss any consistent part of the manuscript, such a subsection, that would discuss this issue at a reasonable depth. Information is provided here and there, but I cannot see any formula that would model this.

=> I now discuss in additional details the rationale for assuming the greater virulence leads to more treatment.

In terms of modelling, for the short-term evolution (i.e. Price equation approach), the fact that more virulent infections are treated more is captured by the covariance between virulence and treatment rate, as indicated in the initial version of the manuscript:

> *The third term in equation 6a deserves some explanations because it involves the rate at which infections are treated (γ_i). In most epidemiological models, this rate is assumed to be the same for all the strains. Here, this term allows us to assume that treatments are triggered by infection symptoms, which are themselves correlated to virulence. If the covariance is positive, the term acts as a selective force against virulence in the drug sensitive compartment.*

For the adaptive dynamics, this relationship has limited effect since the treatment rate γ does not appear in the fitness of the drug-resistant strain. It does however appear in that of the drug-sensitive strain, and, in the appendix, we discuss potential trade-offs between γ (treatment rate) and α (virulence).

=> I added sentences in the methods in the short-term and long-term evolution paragraphs to further explain how we model the link between virulence/symptoms and treatment rate.

=> I expanded the adaptive dynamics section in the methods, to introduce the trade-

offs in the equations as well as the evolutionary stability conditions. I also present more results about adaptive dynamics in the main text and in the abstract.

Moreover, Figure S3A that appears to describe this is given only in the appendix, and its legend is quite poor; How the points are obtained? Is the dashed line a fit? I think a deeper discussion of this has to precede model formulation.

Figure S3 illustrates the numerical implementation of the Price equation, i.e. the simulation of a finite number of strains. Practically, we assume a multivariate normal distribution for α and γ such that the variance of α is 1/6, the variance of γ is 1/12, and the covariance between α and γ is the same as that between γ and α and is 1/10. We draw 10 random pairs of (α,γ) in this distribution. Transmission rate β is either constant or determined directly using the values of α of the 20 strains and the formula in equation 5. Dashed lines show the results of a linear model fit between α and γ for panel A, and the theoretical trade-off relationship from equation 5 for panel B.

=> I now give details about the way the data is generated and improved the caption for Figure S3.

I admit I do not fully agree with the general idea that more treatment of more virulent individuals generally means faster recovery. I can imagine that severely symptomatic individuals, despite receiving more treatment, may recover much later than individuals with mild symptoms and less treatment.

This is partly the case because in the case of treatment failure, i.e. evolution of resistance, infections will not be cleared at all. Furthermore, in the case where there is no resistance, even if complete recovery can take more time, what matters from an epidemiological standpoint is the transmission. Here, we do assume that for non-resistant infections transmission stops upon treatment initiation.

=> I now explicitly introduce a rate of natural recovery (\nu), which is shown not to affect the results.

Furthermore, it seems the parameter rho is constant in simulations and does not depend on treatment level and symptoms severity. Is this realistic? Is not more likely that when more treatment is provided then there is more likely that treatment fails (or vice versa), so that rho cannot be considered constant?

As indicated in the original version, there is an identifiability issue between ρ and γ. To disentangle between the two would require adding additional details, for instance an explicit recovered class (R) with a finite duration of immune memory. However, as for other model precisions, this would also greatly increase the number of compartments to follow. Importantly, assuming that more virulent infections lead more often to treatment failure would yield the same results, which does reinforce the applicability of the model for the Price equation part.

=> This identifiability issue is now mentioned again in the Discussion.

Is seems the only result that is reflected in the abstract comes from Figure 3B. Honestly, I do not see the shown difference worth the merit given in the abstract and am curious how small changes in model parameters may change the curves. On the other hand, why results from adaptive dynamics are not reflected in the abstract?

I indeed put more emphasis on the short-term evolution results rather than on the long-term ones. For the latter, I only wrote that “the exact shape of genetic trade-offs govern long-term evolutionary dynamics”. More details can indeed be added because the nature of the strain present at equilibrium will vary (in addition to the equilibrium virulence value). Furthermore, quite embarrassingly, I realised I forgot to refer to the study by Porco et al (2005) on how treatments may affect virulence evolution depending on the underlying trade-offs. They do not mention the evolution of resistance but their scenarios cover part of the adaptive dynamics results presented here.

=> I now give additional details about long-term evolutionary dynamics.

Regarding adaptive dynamics, I am curious why calculations conserving evolutionary and

convergence stability are not provided and a typical pairwise invasibility plot is not shown.

The long-term evolutionary dynamics performed using the adaptive dynamics framework lead to invasion fitness expressions that are classical in the field and have been analysed in details since van Baalen & Sabelis (1995, Am Nat), the case where the drug-sensitive strain invades being the most interesting. Furthermore, additional studies such as that by Kisdi (2006) have shown that the exact shape of the trade-offs can determine evolutionary stability and that even slight changes in the trade-off function can have dramatic differences in terms of evolutionary outcomes.

=> I now refer to earlier studies on the evolutionary dynamics associated with such fitness expressions and the limits of adaptive dynamics approach due to uncertainty about trade-off shapes.

My final remark concerns the use of Price equation technique. As Figure 2 suggests, even at quite short temporal scale the approximation power of the Price equation system is quite poor, as the author himself admits, so I really do not see utility of it. Why the author keeps it here? I can imagine moving everything of it to an appendix and say it does not work well, but I regret not providing it in any form, but as it is now I doubt it is useful.

In a way, the Price equation approximation is as poor as the adaptive dynamics'. Indeed, the Price equation investigates short-term dynamics so the further away we move from the initial conditions, the less useful it will be (unless the variance-covariance matrix can be updated). For the adaptive dynamics it is the other way around and unless we wait for what some might say is an unrealistic amount of time, its prediction will be completely off, even from a qualitative point of view as illustrated for instance with Figure 3B and the inversion between the more and less virulent strains.

=> I now further underline the limitations of both approaches as we move away from the initial conditions (for the Price equation) or the long-term equilibrium (for the adaptative dynamics) in the Discussion.

In summary, I also find the study interesting and useful, but at the same time think the author should revise the manuscript in order to present motivation, model formulation and results presentation in a more balanced way.

Thanks you for the appreciation and the suggestions.

Response to Reviewer #1

Summary:

This article aims to mathematically describe a co-evolutionary dynamics of virulence and drug resistance of a pathogen in a population infected by drug-sensitive and drug-resistant strains. The proposed model assumes that there is a correlation between virulence and the rate at which infecteds receive treatment. Co-evolution of the pathogen virulence and drug resistance under the virulence-treatment relationship is investigated. My comments and questions are as follows.

1. I think the introduction section could benefit from adding specific examples of viral, bacterial or fungal infections to which the proposed hypothesis/model, that is, infected individuals with increased virulence receive treatment more quickly and their recovery is thus faster, could be applied.

As pointed out in the answer to Dr Berec, the initial version of the manuscript was potentielly misleading because the result of the treatment is that individuals stop transmitting the parasite, which can be different from full recovery. Note also that in the case of the Price equation, similar results are achieved if we assume that the proportion of treatment failure increases with resistance, which is consistent with genetic associations between the two traits. Furthermore, following Porco et al (2005 JTP), we now cite two textbooks to underline the (intuitive) idea that treatment rate may increase with symptom severity:

- Tuckett, D. (Ed.), 1976. An Introduction to Medical Sociology. Tavistock, London.

- Twaddle, A.C., 1979. Sickness behavior and the sick role. G. K. Hall, Boston.

I also realised these ideas of the link between symptoms and treatments are discussed by Ewald (1980, JTB). Finally, the question is also raised by Canguilhem in his book on the normal and the pathological.

=> I added a sentence in the introduction to further discuss this issue in the Introduction.

2. The model 1a-1c assumes that all infecteds receive treatment though at different rates depending on the virulence of the strain type 'i'. A fraction (1 - rho_i) recovers and a fraction rho_i transitions to the resistant state. Here, if I understand correctly, recovery occurs immediately upon treatment administration, is positively correlated with virulence (more virulent implies more rapid treatment and thus recovery) and if there is no treatment (gamma = 0), there is no recovery. I wonder if having the natural recovery of drug-resistant strain (traded-off with virulence) in the model might be reasonable since people can spontaneously recover from a large number of viral and bacterial infections. Also, the model assumes that all infecteds receive treatment; perhaps having treated and untreated classes would be interesting to incorporate and investigate. For the case of self-medication/prophylaxis, a paper here https://doi.org/10.1111/j.0014-3820.2005.tb01008.x might be worth looking at.

Natural recovery was indeed not mentioned, which is an issue. However, from a mathematical standpoint, as long as this recovery is independent from the other infection traits and can be assumed to be constant, then it can be merged with the \mu term (presented initially as baseline mortality).Exploring potential trade-offs with this recovery rate with treatment rate would lead to similar dynamics as the ones described since the natural recovery term would act as the virulence. To fully explore the implications of natural recovery, it would be necessary to include a former "Recovered" host class (ideally with a short-lived immune memory).

I did not know about the model by Boni and Feldman, which is a nice population genetics analysis of self-medication and antibiotic resistance evolution. However, unless I am mistaken, the strains only differ by their resistance status, which affects their recovery rate (i.e. there are no differences in transmission and virulence). Furthermore, in terms of model structure, hosts either belong to antibiotic users or non-antibiotic users. These make it quite different from the current setting, where there is strain variation in other traits than recovery and where the treatment is a dynamical process. However, the reviewer is entirely correct in that references to earlier models on the evolution of antibiotic resistance were too scarce!

=> I now refer to the review by Spicknall et al (2013, Am J Epidemiol) on modelling the evolutionary epidemiology and drug resistance (and as it turns out, this review does cite the article by Boni & Feldman).

3. The model assumes that drug treatment directly affects recovery. I could imagine that there might be indirect effect of the drug through e.g. decreasing the infectious pathogen load/virulence and thus potentially accelerating recovery. Is that something that could be investigated via the model and affect the co-evolutionary dynamics?

The treatment rate could indeed be correlated with other traits and the Price equation nicely allows to capture this on short-term dynamics. As mentioned above, dynamics for natural recovery, if it was assumed to be negatively correlated with treatment rate (i.e. infections that clear early are less treated) would lead to the opposite trend of virulence, since, as explained above, virulence and natural recovery amount to the same thing from a parasite point of view. For transmission rate, studying such a link is possible mathematically but, from a biological standpoint, it is more difficult to envisage what potential correlations might be... Would infections that transmit more be treated more rapidly, e.g. if we assume that symptoms correlate with transmission? Or would it be the opposite, e.g. if we assume that asymptomatic transmission is more important?

Regarding co-evolutionary dynamics, the one mentioned by the reviewer is actually already present in the study. Indeed, as can be seen on panels 2C and 2E, the decrease in virulence comes along with a decrease in treatment rate. It would be interesting to extend this trait coevolution by allowing for correlations between more than two traits (e.g. virulence, treatment failure and treatment rate) but this would require more data to parameterise the model because currently then number of scenarios to explore would be daunting.

=> I now mention that other traits (co)evolve with virulence.

4. In the relationships for virulence and transmission of the resistant strains, alphai^R = a*alpha and betai^R = a*beta, the assumption is that a >1 or 0 <= b < 1, so that drug-resistance results in higher virulence and decreased transmission. It might be useful to elaborate more on these assumptions and the choice of their values.

As indicated in the main text, parameters a and b are associated to fitness cost related to the drug resistance. Since bacterial fitness depends on infection duration and transmission to new hosts, the only two means to achieve this cost are to increase virulence ($a>1$) or to decrease transmission rate ($b<1$). These values have limited effects on the results and, in fact, the figures in the main text assume that $a=b=1$ (i.e. no fitness cost). In theory, as long as the R from the drug resistant strain is lower than that of the (untreated) drug-sensitive strain, we have a fitness cost. This means that we could have a variety of pairs (a,b) that satisfy the conditions (e.g. strains that are less virulent but barely transmit or strains that transmit well but are very virulent).

=> I now clarify that in the main text the fitness cost is neglected in the numerical simulations. I also added references to the biological bases of this fitness cost.

5. The trade-off between virulence and treatment rate (p. 19, Fig. S3 A) assumes a linear positive relationship between the two. I imagine that this relationship could be sigmoid so that those infected with low to mid-virulent pathogens would receive treatment at slower rates. I would be curious how this would affect the outcome.

Here, in absence of clear data, we used the most parsimonious hypothesis (i.e. a linear trade-off). In the price equation approach, assuming a saturating trade-off would have limited effects since the equations rely on variances and covariances. However, the non-equilibrium dynamics would be affected, and so could the adaptative dynamics results. For the latter, this was discussed in Appendix with the two possible equilibria: a trade-off between virulence and treatment rate would only matter if the drug-resistant strain cannot invade.

=> I added details about the Adaptive dynamics results in the main text and now mention the potential effect of this saturating trade-off on the results.

6. On Fig. S3 A, do these points represent the values used in the simulations with multiple strains? Perhaps it would be good to have a more detailed description of the used methods also in the text so that one does not need to look into the R code.

I apologize for the lack of clarity.

=> As explained to Dr Berec, I now explain how the figure was generated.

7. In the equation S9c, rho is assumed to be a function of virulence. Is this assumption used anywhere else in the manuscript or simulations?

As indicated above, in this model, there is an identifiability issue between the treatment rate and the treatment failure ratio: both lead to similar results in this model. To generate different dynamics, we would need a more detailed model, e.g. with an explicit short-lived immunity, but this would also mean many more host classes and would likely require a purely numerical exploration.

=> The identifiability issue is now mentioned again in the Discussion.

Response to Reviewer #2

The present manuscript offers a mathematical view on why the resistant and more virulent strain can be selected in epidemics. I have found interesting that the author highlighted the important aspect that this can be in particular attributed to public health policies.

Thanks!

Although I like and appreciate mathematics, I think it would be beneficial to give some ground for the performed simulations: for example, to adjust some values and time scale for a particular pathogen in mind. Otherwise, it is difficult to understand, e.g., what the time units for Fig.1 (is the x-axis in ages like for HIV or TB or in weeks like for influenza?).

Host demography has a very limited role in this model so the time unit is largely arbitrary. More precisely, we will obtain the exact same plot if infections are short and transmission rate is high (e.g. influenza) or if infections are long and transmission rate is low (e.g. TB).

=> I now specify that the default time unit in the simulation is weeks, but that similar results would be observed for short infections with high transmission rate or long infections with low transmission rate.

It would be also great to see more connection of the framework with real examples/diseases outlined in the introduction. Even though the discussion is quite detailed, I wish to see the same for Introduction section. In this respect, there is an imbalance between Introduction and Discussion sections.

=> I now discuss in more details the link between symptoms and treatments (with two references), as well as the bibliography on modelling the evolutionary epidemiology of drug resistance.

Main remarks: L32: I have encountered a small contradiction in the framework. As the author writes about incubation period vs illness onset, but his main model does not include the "E" compartment. How the results would be changed or not changed if one would consider SEIR model instead of SIR model and how important the inclusion of the "E" compartment in view of the evolutionary dynamics?

Actually, the initial version of the model explicitly distinguished between asymptomatic and symptomatic hosts (so more of an SAIR model than an SEIR). The idea was then that symptomatic hosts (I) were treated more rapidly than asymptomatic hosts (A). In the end, the results were similar with the dynamics of a trait x in the asymptomatic hosts being affected by a term -Cov(\sigma,x), where \sigma is the rate of becoming symptomatic. The main reason for removing the A class was that the SAIR model requires 4 differential equations for each trait, since its value can be expressed in A, I, A_r, and I_r hosts, where the r subscript indicates drug resistant infections.

=> I now mention in the discussion that an alternative approach could consist in explicitly modelling symptomatic vs. asymptomatic hosts but that this would lead to twice as many differential equations without affecting the results qualitatively.

L63: I think that the examples make sense, but, e.g., TB is not modeled by a simple SIR model. In this view, the paragraph is more suited for introduction part. Then the author writes about parasites in L70, but all examples above except of Plasmodium falciparum are not parasites. Hence there is some mixture of terms, and it can be confusing.

Here the purpose was more to discuss the genetic architecture of the model but I agree that envoking it in the model section is potentially misleading.

=> I moved the discussion of the genetic structure of the model to the introduction and yo the discussion.

Minor: L28: Would it be possible to clarify the term "pleiotropic action" for a general reader?

=> I now make the word pleiotropic more explicit.

L30: "the latter trait is usually binary rather than continuous." – the author does not explain much on why it is important. At the moment that sentence seems out of place.

=> I rewrote the paragraph on modelling the evolution of drug resistance.

L32: I would advice to change the word order: the incubation period and the illness onset, because the former comes first.

=> This sentence was removed in the revised version.

Introduction: I have noticed the gap in the literature dates, and it looks suspicious, because there should be some scientific works in the last few years. For example, Sylvain Gandon probably tackled some relevant problems, but it is just my guess. For example, the work Todd L. Parsons et al 2018 RSIF (doi:10.1098/rsif.2018.0135) looks relevant at least for the notion.

About Sylvain Gandon's work, I am not aware or more recent work than the Day & Gandon (2012) article, which was cited in the introduction. I had a look at his (large) publication list

but all I could find since this article was the following reference on the role of refugia zones on the evolution of resistance, which seems marginally relevant: *Park A. W., Haven J., Kaplan R. & Gandon S. (2015) Refugia and the evolutionary epidemiology of drug resistance. Biology letters. 11(11):20150783* http://dx.doi.org/10.1098/rsbl.2015.0783

Perhaps it would help if the reviewer was more explicit about what he/she has in mind...

About the study by Parsons et al, I am not sure I see the direct relationship because it is a (very neat) exploration of the role of stochasticity in virulence evolution but, again, I fail to see the link with drug resistance.

L85: I would write "a disease-free equilibrium" instead of "a trivial state".

Done.

Response to reviewer #3

This paper is an interesting addition to the literature on the interaction between epidemiological processes. The author states and then illustrates than differential treatment of individuals based on the virulence of their infecting pathogen strains can cause changes in the absolute and relative levels of virulence in the resistant and sensitive strains.

Thanks!

The analysis appears robust and the conclusions are interesting. The analysis of the impact of noise in transmission rates is appreciated, but a more general consideration of the robustness of the results to changes in parameter values would add to the manuscript, as if the effect occurs only in a small region of the parameter space (in this, admittedly, highly abstracted model), it is less likely to have large biological relevance in the real world.

I can of course generate more figures with additional parameter values but the results are analytical: for the short-term evolution, they originate from equation 6, and for the adaptative dynamics, the details were in the Appendix and have been moved to the main text. The figures are only intended to provide a graphical illustration for a wider audience.

=> I now further specify that the results are obtained analytically and that the figures are intended to provide a graphical illustration.

On the topic of modelling drug resistance as a continuous trait, the author admits that that the inclusion would render the model unable to show the effect stated (due to modelling the modelling framework). While this is not an inditement of the work in this paper, it might be worth the author's time to add to this section to describe ways that continuous modelling of virulence could be performed. Just off hand, it appears that this could be done using (for instance) something analogous to the integro-partial differential equation framework of Delgado-Eckert et al in The evolution of virulence in RNA viruses under a competition-colonization trade-off (https://arxiv.org/pdf/1001.3101.pdf), albeit at the cost of the simple mathematical formulation, that is a strength of this paper. Something to consider.

Partial differential equations are indeed a way to study the evolution of resistance as a continuous trait and we are actually about to submit a manuscript using this technique in the team. There are also several models that treat resistance as a quantitative trait.

The difficulty is more to treat *both* virulence and resistance as a quantitative trait. Indeed, at the level of an infection trait, drug resistance affects the value of the life-history traits (virulence, transmission, or recovery rate) expressed in a treated host, which we denote in the model (α, β, γ). If resistance is quantitative, as in this study, we have two sets of parameters in untreated hosts, (α^S, β^S, γ^S) in drug-sensitive infections and (α^R, β^R, γ^R) in drug-resistant infections. Still in our model, we assume a variety of strains with different virulence, with potential correlations between virulence and traits such as transmission or recoveryrate. Notation-wise, for each strain *i*, we can define (α_i^S, β_i^S, γ_i^S) and (α_i^R, β_i^R, γ_i^R). But if resistance is quantitative, it means there is continuum for the resistance traits. Notationwise, this would appear like (α_{ij}, β_{ij}, γ_{ij}), where *i* is the strain number (related to the virulence) and j the strain

number (related to the resistance). As can be seen from the notation, the two (continuous) trait spaces of virulence and resistance are colliding. Unless we have detailed data on the underlying mutation kernel, this model will be identical to an evolutionary model with a notation (α_k, β_k, γ_k), i.e. with a single continuous trait evolving.

=> The difficulty to model both virulence and drug resistance as continuous traits is now explained in the Discussion.

However, overall I enjoyed reading the paper and it seems like a valuable addition to the literature!

Thanks!

General notes:

Line 53 & 54: “by a drug resistant host” - it is the pathogen that is drug resistant, not the host

Indeed!

The section after equations 2a and 2b is confusing. It eventually becomes clear the author is using x-bar as a generic stand in for the weighted mean of any parameter in the equations, but this would be best stated explicitly, so the reader is not attempting to find and work out what trait x and y actually correspond to.

=> The notations are now introduced before the equations. I also added a Table summarising the notations used.

The equilibria are defined have no strain subscripts at line 84, but equilibria 2 and 3 are given with subscripts. Is this meant to imply that each strain persists with its own equilibrium level given by their own strain specific parameters? If so, this would be better explicitly stated.

Thank you for spotting this. The equilibria conditions should indeed be for a strain *i*. On the long run, some strains will go extinct.

Equation 4 - as far as I can see, the subscript m is not defined either in the text or the appendix. It first appears after the calculation of the eigenvalues in appendix A, but is not commented on. Its meaning should be explicitly stated.

The adaptive dynamics section was indeed too short.

=> I now explain the resident (r) / mutant (m) notation in further details. I also give more details about long-term evolutionary dynamics.

Line below equation 4 - apparition is presumably mean to be appearance?

Apologies for the use of a French word...

Line 89 - I assume the dichotomy being referred to is which of the eigenvalues dominates and therefore determines the fate of the mutant? But this should be clarified.

Yes, it does refer to the two possible states. We now discuss in more details the adaptive dynamics results.

Line 148 - misspelling of average

Corrected!