Epidemics 10 (2015) 31–34

Contents lists available at ScienceDirect

Epidemics

journal homepage: www.elsevier.com/locate/epidemics

Five challenges in modelling interacting strain dynamics

Paul S. Wikramaratna[a,b], Adam Kucharski[c], Sunetra Gupta[b], Viggo Andreasen[d], Angela R. McLean[b], Julia R. Gog[e,f,*]

[a] Institute of Evolutionary Biology, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom
[b] Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, United Kingdom
[c] Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London WC1E 7HT, United Kingdom
[d] Department of Mathematics and Physics, Roskilde University, DK-4000 Roskilde, Denmark
[e] Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, United Kingdom
[f] Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA

ARTICLE INFO

Article history:
Received 22 March 2014
Received in revised form 26 July 2014
Accepted 28 July 2014
Available online 7 October 2014

Keywords:
Multiple strains
Cross-immunity
Adaptive immunity
Pathogen evolution
Mathematical modelling

ABSTRACT

Population epidemiological models where hosts can be infected sequentially by different strains have the potential to help us understand many important diseases. Researchers have in recent years started to develop and use such models, but the extra layer of complexity from multiple strains brings with it many technical challenges. It is therefore hard to build models which have realistic assumptions yet are tractable. Here we outline some of the main challenges in this area. First we begin with the fundamental question of how to translate from complex small-scale dynamics within a host to useful population models. Next we consider the nature of so-called "strain space". We describe two key types of host heterogeneities, and explain how models could help generate a better understanding of their effects. Finally, for diseases with many strains, we consider the challenge of modelling how immunity accumulates over multiple exposures.

Introduction

Mathematical models of "fully immunising" infections like measles have become useful tools for understanding the dynamics of some infections and, in several instances, planning their control. However, there are many common and important infections that simply do not follow the assumptions of "measles models". For examples measles models fail to illuminate the behaviour of pathogens that exist as inter-related families of strains or variants. Influenza, norovirus, malaria and dengue virus are just a few of the infectious agents for which the interacting dynamics of different strains form a crucial part of their biology.

Recent decades have seen substantial progress in building a stronger understanding of how partial host immunity acts to structure the population dynamics of such infections. However, much remains to be done. In this review we lay out five challenge areas that we believe remain open and important for advancing our understanding of strain dynamics.

1. Translating from single host to population strain models

To make tractable population models of multiple strains, modellers typically make relatively simple assumptions about the nature of "partial cross-immunity". For example, a partially immune host can be treated as having a lower rate of becoming infected (reduced susceptibility) or a lower rate of transmitting the infection to others (reduced transmissibility), but reality for some systems may be a mix of the two, and also other effects such as reduced duration of infection (Park et al., 2009). A major challenge for the area, and indeed one that underpins all the other challenges presented here, is how to choose an appropriate abstraction: for a given system, how should we move from the full immunodynamics at the individual host level to reasonable assumptions that can be incorporated into population models? We explore some aspects of these choices and their consequences here.

If we define 'reduction in host susceptibility' as a reduction in the probability that an infection can take place within an individual host, then if an infection does indeed occur does it proceed as normal, or not? If the infection does not proceed, is that individual then also protected in subsequent challenges? On the other hand, defining 'reduction in host infectiousness' as a reduction in the probability of onward transmission, does this occur because the

* Corresponding author at: Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, United Kingdom.
Tel.: +44 1223 760429.
E-mail address: jrg20@cam.ac.uk (J.R. Gog).

http://dx.doi.org/10.1016/j.epidem.2014.07.005

peak in pathogen load is generally diminished, and/or is there simply faster clearance of pathogen following that peak? Further, is it the overall pathogen load that matters, or only the pathogen load in particular tissues, and how do these different factors combine to favour certain pathogen strains?

The rate of pathogen clearance, or the duration of infectiousness, is affected by partial immunity, and this will consequently shape both the pathogen population that is available for onward transmission. In particular, the immune system is designed to rapidly recall pre-existing cross-reactive immune responses from memory and any such responses therefore have a selective advantage over novel immune responses within infected hosts. The repertoire of partial immunity within an individual host might not therefore change much as a result of an exposure in which little if any pathogen replication takes place, but if the host, in effect, ignores some of the antigenic novelties of this particular pathogen then this may be to the pathogen's long-term advantage. Equally, recalled responses may result in a poorer quality response than that seen in naïve individuals in some circumstances – see original antigenic sin (Kucharski and Gog, 2012a; Haaheim, 2002) – and/or result in immunopathogenesis (Ubol and Halstead, 2010).

A careful consideration of what immune escape means at the level of a single host may also help us develop more informative models of pathogen evolution at the population level. If each mutation makes only a subtle change to the antigenic profile of the pathogen, how do such novel variants overcome their inevitable disadvantage in frequency compared to the wildtype within a host? On the other hand, if each mutation makes a sufficiently large change to the antigenic profile of the pathogen so that it totally escapes from pre-existing responses, does this not then drastically affect the subsequent profile of pathogen load within these individuals? In attempting to answer these sorts of questions, recent work on the influenza virus neatly illustrates alternative mechanisms of escape: viruses that bind to target cells faster – giving the immune system less time to eliminate them – may be favoured in the presence of a strong immune response (Hensley et al., 2009; Yuan and Koelle, 2013). Simultaneous infection by different variants of the same pathogen may also affect the possibilities and propensities for antigenic change (see [reference] paper in this volume.)

We must consider how the various arms of the immune response act both in concert and independently to moderate the extent of infection and the types of selection that they impose on a within host level and how this then manifests at the level of host populations. In the case of viruses for example, at the most basic level, antibody responses can act to eliminate cell-free virus, whereas CD8+ T cells destroy infected cells. Thus, the latter response only comes into play once infection is underway but an antiviral antibody response has the chance to terminate an infection before any replication takes place. Meanwhile, the role of innate immunity in regulating pathogen load has often been ignored since it would be expected to target strains indiscriminately. Nevertheless, this is not necessarily the case: elements of the innate immune system have now been shown to behave in an adaptive manner (Vivier et al., 2011). Each type of response may therefore act in different ways to suppress pathogen replication (i.e. exert different selection pressures), not to mention that each tends to target fundamentally different antigens, and therefore ultimately differentially affect pathogen evolution.

Without a better understanding of the above, it will be difficult to build models that explicitly link within to between host dynamics. Nevertheless, provided comprehensive within-host infectious datasets are available, a recent approach sidesteps this problem by estimating genetic covariance functions for the pathogen traits of interest (Day et al., 2011; Mideo et al., 2011). Not only does this yield interesting insights, for example into the interplay between the age distribution of infections and the genetic constraints imposed on the pathogen, but it should also ultimately complement and help validate future developments in our understanding of mechanistic within-host processes.

2. The nature of strain space

A critical issue in modelling multi-strain systems is how to best represent the relationships between the biological entities involved. A common approach is to consider strain space as a continuous volume, surface or even, most simply, a single dimension in which the Euclidean distance between points serves as a measure of their antigenic and/or genetic similarity (Lin et al., 1999, 2003; Gog and Swinton, 2002; Bedford et al., 2012). This approach is appropriate when the targets of immunity (or epitopes) can experience continuous variation, and also substitute for each other in the sense that strong immunity against one epitope can compensate for a lack of immunity against another. Alternatively, one can argue that antigenic distance should be regarded as the minimum distance across all dimensions (Kryazhimskiy et al., 2007). In both cases, an implicit assumption is that antigenic space is homogenous.

Variation in epitopes can also be considered as discrete, with strains occupying nodes $\{i,j,k,\ldots\}$ where i designates a particular variant of epitope 1, j designates a particular variant of epitope 2, etc. As with continuous systems, a function has to be chosen to express the risk of infection with (or transmission of) a strain in relation to the fraction of epitopes previously encountered. This could simply be proportional to the Hamming distance between two strains or epitope strings, or take some more complex form (e.g. Zinder et al., 2013). In Gupta et al. (1998), for example, exposure to any fraction of epitopes is sufficient to reduce transmission by $(1-\gamma)$ where γ represents the strength of immune selection (Gupta et al., 1998). An alternative formulation was explored in Recker et al., where risk of re-infection was zero when the fraction of epitopes already encountered was 1 even though the exact strain may not have been encountered previously; however it was only possible to write down a set of equations for a 2-locus, 2-allele system (Recker et al., 2007). Much further work needs to be done to link particular behaviours to different functional forms in both discrete and continuous systems.

An important challenge is to identify which of these structures adequately represents the "strain space" of a particular host–pathogen system as this affects the dynamics in important ways. Many of these systems exhibit polarisation (i.e. discrete strain structure) but not all will exhibit cyclical and/or chaotic behaviour (Gomes et al., 2002). The mode and tempo of mutation (i.e. the rate of generation of novel antigenic types) can also be important (Zinder et al., 2013; Bedford et al., 2012; Minayev and Ferguson, 2009). One would hope that a combination of molecular and serological studies should enable us to assign the most appropriate structure to each system, but this may not be easy. For example, antibodies typically target conformational epitopes on the surface of pathogens. As such, mutations within the epitope itself may completely, subtly or not at all change its biochemical properties, or mutations outside the epitope may nevertheless affect the presentation of the epitope and the strength of antibody binding. In other words, there is not necessarily a clear mapping from genotype to antigenic phenotype and multiple substantially different genotypes may correspond to identical phenotypes, whilst other closely related genotypes may quite far apart in phenotypic space. These concepts form the basis of 'neutral network' theories of strain space (Koelle et al., 2006) but many problems remain when trying to match the increasingly available genetic data to model driven hypotheses of antigenic evolution.

3. Host heterogeneity: population structure

When population heterogeneity is incorporated into transmission models, it has the potential to substantially change the dynamics of an infection. Certain types of heterogeneity are relatively straightforward to add to strain models. It has been observed that strain diversity can increase if contact network (Buckee et al., 2004) or community structure (Buckee et al., 2007) is included in a simple multi-strain model. Further, variation in social contacts between different age groups can influence the interaction between epidemic dynamics and population-level immunity after multiple infections (Kucharski and Gog, 2012a).

Other kinds of heterogeneity are more challenging to incorporate into models. To ensure computational tractability, many strain models do not track all possible combinations of past infections; instead they gather these combinations into summary variables. These summary variables are used to capture the immune state of the population (Gog and Swinton, 2002; Gupta et al., 1998; Kucharski and Gog, 2012b) with the potential for substantial generalizations (Ferguson and Andreasen, 2002). To make this simplification possible, models assume a population without immigration: other than via births or deaths, nobody enters or leaves the pool of potential hosts. However, the simplified model structure breaks down when models include movement of people, rather than just transmission of infection between locations or social groups. If new individuals arrive or leave, the values of the summary variables – which were tailored to the original population – no longer reflect the true level of population immunity.

A key challenge is to understand how population structure and movement influences strain dynamics over long timescales. In particular, it is important to find ways to extend population models of disease strains to account for changes in host demographics, such as individual movements between different locations or fluctuations in birth rate. These changes could have a major impact on the immune composition of the host population, and hence the evolutionary dynamics of an infection. A surplus of partially immune hosts could increase selection pressure exerted on a pathogen, and influence the proliferation of new antigenic variants.

Individual-based models could be used to examine how heterogeneities interact with strain dynamics in this manner. Such models would make it possible to explore the effects of population heterogeneity on pathogen competition and the appearance of new strains (Ferguson et al., 2003). For example, there is the potential to examine how changes in population structure have influenced the emergence and re-emergence of influenza and dengue fever subtypes in the past. Multi-strain models could also show how spatial movements might combine with strain interactions to affect disease incidence.

4. Host heterogeneity: variation in immune response

Until now our discussion has focussed on the heterogeneity that is associated with the hosts' position in the transmission network, and their previous exposures to related strains. This approach to cross-immunity implies that hosts are identical in their immune response in the sense that if two hosts were exposed to exactly the same sequence of challenges during their lives, then they would mount the same immune response – at least in a probabilistic sense. There is, however, ample evidence that variation in immune type, nutritional status, and age may affect the specificity of the immune response. For example while children produce a monoclonal response to influenza infection older individuals produce polyclonal responses (Cobey and Pascual, 2011). Furthermore, certain hosts may respond to conserved epitopes while others recognise variable epitopes (Gupta and Galvani, 1999). Host immune responses also suffer restriction by the Major Histocompatibility Complex, and modelling this interface between host and pathogen (Penman et al., 2013; Fryer et al., 2010) remains a key challenge.

Host specificity of the cross-immunity response may be hard to identify even in challenge experiments. For example imagine that we have a population that had all been previously infected with strain A. When challenged with a related strain B, we observe that 40% of the hosts become infected. We may then conclude either that each individual host has a 40% chance of being reinfected when challenged (and that this is a random effect at point of challenge) or that 40% of the host population will be reinfected when challenged (and that these individuals are predetermined by some inherent variability). Most multi-strain studies assume the former of these interpretations (but see (Cobey and Pascual, 2011; Simpson and Roberts, 2012)) and it is currently unclear how inherent heterogeneities in immune response affect disease transmission.

5. Accumulation of immunity from multiple strains

Suppose we have a defined strain space, and we know how exposure to one strain confers partial immunity to another strain. With three or more strains we would still be left with an ambiguity about how immunity from multiple exposures accumulates. For example, suppose strain A reduces susceptibility to strain C by a factor 0.5, and strain B reduces susceptibility to strain C by a factor 0.6. How immune to strain C is someone who has had both strains A and B? To put the problem in mathematical terms, we require a function that depends on the set of strains previously seen (Andreasen et al., 1997), which somehow we must specify or extend from the function evaluated for each strain singly. Generally modellers have selected their model for convenience, or indeed without comment about how immunity accumulates: this aspect of model choice often goes uncommented on and therefore perhaps unnoticed, but has potential to alter the dynamics of a strain model.

One extreme adopted by many modellers is so-called "product" cross-immunity (Gog and Grenfell, 2002), where the factors reducing susceptibility or transmissibility accumulate geometrically: the reduction is therefore 0.3 in the above example. Another extreme is "minimum" cross-immunity, where essentially the strain (or one of the strains) that confers the strongest immunity is the one that is counted. In the above example, it means that the 0.5 reduction from strain B counts, and strain C "sees" no difference between those previously infected by B, and those infected by both B and A. The minimum and product each can offer mathematical convenience, but how valuable that is depends on the model framework. Minimum cross-immunity means that potentially only a small amount of information is needed per host, for example with a linear strain space for annual influenza: for each host only their most recent strain infection shapes their immune response to newer strains (Andreasen, 2003). The product assumption means that all past exposures contribute to the current immunity, but sometimes the product gives some mathematical independence that can be used to simplify the model variables (Gog and Swinton, 2002).

Epitope models (see 'nature of strain space') are neither of the above: their structure is more complex, but based on consideration of how immunity may depend on contributions from multiple "parts" of the pathogen. Epitope models can sidestep the issue of how immunity accumulates, as cross immunity is no longer a function of a set of strains seen, but rather a set of epitopes seen. However, if our epitope space is fine enough that there are similar epitopes, we will be back with the same questions.

There are three aspects to this challenge that would enhance our ability to model infectious diseases. Firstly there is the mathematical challenge to characterise exactly what the effects of different assumptions about how immunity accumulates are. When does it not matter which we choose, and when would the dynamics depend on choice? Some progress has been made in comparing the different modelling assumptions (Dawes and Gog, 2002; Gomes et al., 2002) but so far these have been specific to certain simple strain spaces. The second aspect is to develop frameworks that allow relatively simple models to be developed with broader assumptions possible than now, so that we are not constrained to one choice for tractability (Kucharski and Gog, 2012a,b). The third and perhaps most urgent part is that we need a better empirical understanding of how partial immunity accumulates for different pathogens and hosts. This comes back to the challenge of translating from within-host dynamics to tractable population models.

Summary

All models – be they mathematical, animal or experimental – are a necessary abstraction of reality. We use models precisely because we cannot hope to recreate the system in full either for practical or ethical reasons. The challenges we present here involve understanding how such abstractions affect disease dynamics, with the aim of using models to identify the critical determinants driving the system's behaviour.

Model validation is central to such work. This includes an assessment of the individual building blocks of the model and an analysis the overall model behaviour: are the underlying biological assumptions plausible, and does the model capture the system it is trying to emulate? As different models can lead to the same behaviour, the generation of testable inferences that allow discrimination between models is crucial. Validation may take the form of qualitative comparisons, with the model behaving in a dynamically similar way to the real system, or quantitative inference, whereby specific model outputs are compared with empirical measurements in statistical framework.

It may be the case that a model can reproduce observed dynamics despite omitting a known biological phenomenon. This does not mean that the model is invalid; rather, it may indicate that this particular phenomenon is not as important as was initially supposed (i.e. it does not radically affect the behaviour of the model). However, we cannot know this until we have made comparisons between models that both include and exclude the phenomenon of interest.

The intention of this review is to highlight phenomena that have been little studied to date, and outline potential comparisons that could be made between different biological assumptions. When incorporated into appropriate models, these features may provide useful insight into the biology of both pathogens that exist as multiple strains and their target hosts. Some of these challenges may prove more difficult than others, but each represents a new and exciting frontier for future research.

Acknowledgment

The authors wish to thank Katia Koelle and an anonymous reviewer for helpful comments.

References

Andreasen, V., Lin, J., Levin, S.A., 1997. The dynamics of cocirculating influenza strains conferring partial cross-immunity. J. Math. Biol. 35 (7), 825–842.
Andreasen, V., 2003. Dynamics of annual influenza A epidemics with immunoselection. J. Math. Biol. 46 (6), 504–536.
Bedford, T., Rambaut, A., Pascual, M., 2012. Canalization of the evolutionary trajectory of the human influenza virus. BMC biology 10 (1), 38.
Buckee, C.O.F., Koelle, K., Mustard, M.J., Gupta, S., 2004. The effects of host contact network structure on pathogen diversity and strain structure. Proc. Natl. Acad. Sci. U. S. A. 101 (29), 10839–10844.
Buckee, C., Danon, L., Gupta, S., 2007. Host community structure and the maintenance of pathogen diversity. Proc. R. Soc. B: Biol. Sci. 274 (1619), 1715–1721.
Cobey, S., Pascual, M., 2011. Consequences of host heterogeneity, epitope immunodominance, and immune breadth for strain competition. J. Theor. Biol. 270 (1), 80–87.
Day, T., Alizon, S., Mideo, N., 2011. Bridging scales in the evolution of infectious disease life histories: theory. Evolution 65 (12), 3448–3461.
Dawes, J.H.P., Gog, J.R., 2002. The onset of oscillatory dynamics in models of multiple disease strains. J. Math. Biol. 45 (6), 471–510.
Ferguson, N.M., Andreasen, V., 2002. The influence of different forms of cross-protective immunity on the population dynamics of antigenically diverse pathogens. The IMA Volumes in Mathematics and Its Applications, vol. 126. Springer, New York, NY, pp. 157–170.
Ferguson, N.M., Galvani, A.P., Bush, R.M., 2003. Ecological and immunological determinants of influenza evolution. Nature 422 (6930), 428–433.
Fryer, H.R., Frater, J., Duda, A., Roberts, M.G., Phillips, R.E., McLean, A.R., SPARTAC Trial Investigators, 2010. Modelling the evolution and spread of HIV immune escape mutants. PLoS pathogens 6 (11), e1001196.
Gog, J.R., Swinton, J., 2002. A status-based approach to multiple strain dynamics. J. Math. Biol. 44 (2), 169–184.
Gog, J.R., Grenfell, B.T., 2002. Dynamics and selection of many-strain pathogens. Proc. Natl. Acad. Sci. U. S. A. 99 (26), 17209–17214.
Gomes, M.G.M., Medley, G.F., Nokes, D.J., 2002. On the determinants of population structure in antigenically diverse pathogens. Proc. R. Soc. Lond. B: Biol. Sci. 269 (1488), 227–233.
Gupta, S., Ferguson, N., Anderson, R., 1998. Chaos, persistence, and evolution of strain structure in antigenically diverse infectious agents. Science 280 (5365), 912–915.
Gupta, S., Galvani, A., 1999. The effects of host heterogeneity on pathogen population structure. Philos. Trans. R. Soc. Lond. B: Biol. Sci. 354 (1384), 711–719.
Haaheim, L.R., 2002. Original antigenic sin. A confounding issue? Dev. Biol. 115, 49–53.
Hensley, S.E., Das, S.R., Bailey, A.L., Schmidt, L.M., Hickman, H.D., Jayaraman, A., Yewdell, J.W., 2009. Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift. Science 326 (5953), 734–736.
Koelle, K., Cobey, S., Grenfell, B., Pascual, M., 2006. Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans. Science 314 (5807), 1898–1903.
Kryazhimskiy, S., Dieckmann, U., Levin, S.A., Dushoff, J., 2007. On state-space reduction in multi-strain pathogen models, with an application to antigenic drift in influenza A. PLoS Comput. Biol. 3 (8), e159.
Kucharski, A.J., Gog, J.R., 2012a. The role of social contacts and original antigenic sin in shaping the age pattern of immunity to seasonal influenza. PLoS Comput. Biol. 8 (10), e1002741.
Kucharski, A.J., Gog, J.R., 2012b. Age profile of immunity to influenza: effect of original antigenic sin. Theor. Popul. Biol. 81 (2), 102–112.
Lin, J., Andreasen, V., Levin, S.A., 1999. Dynamics of influenza A drift: the linear three-strain model. Math. Biosci. 162 (1), 33–51.
Lin, J., Andreasen, V., Casagrandi, R., Levin, S.A., 2003. Traveling waves in a model of influenza drift. J. Theor. Biol. 222 (4), 437–445.
Mideo, N., Nelson, W.A., Reece, S.E., Bell, A.S., Read, A.F., Day, T., 2011. Bridging scales in the evolution of infectious disease life histories: application. Evolution 65 (11), 3298–3310.
Minayev, P., Ferguson, N., 2009. Improving the realism of deterministic multi-strain models: implications for modelling influenza A. J. R. Soc. Interface 6 (35), 509–518.
Park, A.W., Daly, J.M., Lewis, N.S., Smith, D.J., Wood, J.L., Grenfell, B.T., 2009. Quantifying the impact of immune escape on transmission dynamics of influenza. Science 326 (5953), 726–728.
Penman, B.S., Ashby, B., Buckee, C.O., Gupta, S., 2013. Pathogen selection drives nonoverlapping associations between HLA loci. Proc. Natl. Acad. Sci. U. S. A. 110 (48), 19645–19650.
Recker, M., Pybus, O.G., Nee, S., Gupta, S., 2007. The generation of influenza outbreaks by a network of host immune responses against a limited set of antigenic types. Proc. Natl. Acad. Sci. 104 (18), 7711–7716.
Simpson, J., Roberts, M., 2012. Modelling heterogeneous host immune response in a multi-strain system. J. Theor. Biol. 304, 60–65.
Ubol, S., Halstead, S.B., 2010. How innate immune mechanisms contribute to antibody-enhanced viral infections. Clin. Vaccine Immunol. 17 (12), 1829–1835.
Vivier, E., Raulet, D.H., Moretta, A., Caligiuri, M.A., Zitvogel, L., Lanier, L.L., Yokoyama, W.M., Ugolini, S., 2011. Science 331 (6013), 44–49.
Yuan, H.Y., Koelle, K., 2013. The evolutionary dynamics of receptor binding avidity in influenza A: a mathematical model for a new antigenic drift hypothesis. Phil. Trans. R. Soc. B: Biol. Sci. 368 (1614), 20120204.
Zinder, D., Bedford, T., Gupta, S., Pascual, M., 2013. The roles of competition and mutation in shaping antigenic and genetic diversity in influenza. PLoS Pathogens 9 (1), e1003104.