Abstract—Biological processes occur at distinct but interlinked scales of organization. Yet, mathematical models are often focused on a single scale. Recently, there has been a significant interest in creating and using models that link the within-host dynamics and population level dynamics of infectious diseases. These types of multi-scale models, called immuno-epidemiological models, fall in four categories, dependent on the type of the epidemiological component of the model: network or individual based models (IBM), “nested” age-since-infection structured models, ordinary differential equation (ODE) models, and “size-structured” models. Immuno-epidemiological multi-scale models have been used to address a variety of questions, including what is the impact of within-host dynamics on population-level quantities such as reproduction number and prevalence, as well as questions related to evolution of the pathogen or co-evolution of the pathogen and the host. Here we review the literature on immuno-epidemiological modeling as well as the main insights these models have created.

Index Terms—within-host models, between-host models, immuno-epidemiology, mathematical models, differential equations, reproduction number, evolution.

AMS Subject Classification: 92D30, 92D40

I. Introduction

Biological processes occur at nested scales of organization. In infectious diseases, the dynamical interplay between the microparasite and the host immune system has a strong impact on the epidemiological characteristics of the disease, such as pathogen shedding, population level transmission, disease-induced host mortality and recovery. Yet, traditionally, differential equation modeling of infectious diseases has been strictly separated by biological scale of organization. Within-host modeling of infectious diseases has been drawing significant attention in the last century. Simple differential equation models, developed to describe a number of diseases such as HIV, HCV, malaria, flu and others, have lead to dramatically improving our understanding of how microscopic processes develop and affect the host health. Several books are devoted to within-host (immunological) modeling of infectious diseases [63], [50], [40] and multiple articles develop and use such models to answer an array of biological questions regarding the pathogen and its interplay with the immune system and target cells. At the same time numerous differential equation models have been developed to model the dynamics of specific diseases or in general the distribution of pathogens on population level. Between-host (epidemiological) models have addressed a variety of questions related to public health. Multiple books focus on
the contributions of mathematics to epidemiology ([10], [20], [18], [33] just to mention a few). Some important public health questions that can be addressed with epidemic models include the fraction of the population that needs to be vaccinated to eradicate a disease, the reproduction number of various disease outbreaks and what efficacy do control measures have.

Pathogen reproduction, transmission and evolution are processes that span several scales of biological organization, i.e. intracellular, within-host, and population scales. Answering effectively public health questions at the population level often requires the understanding and the “lifting” of processes from within-host levels to the population level. Unfortunately, very rarely do mathematical models encompass multiple scales of biological organization. Here we will review the relatively recent models linking the within-host scale with the between-host scale. The emergent area of linked data, models and knowledge is called immuno-epidemiology. Hellriegell defines immuno-epidemiology as the area that “combines individual- and population-oriented approaches to create new perspectives” [30]. We define the mathematical immuno-epidemiology as the area in which mathematical dynamical models of within-host disease processes are interlinked to population-level dynamical models of disease spread to allow for novel results. The concept of immuno-epidemiology is not new but in the last half of the 20th century it was primarily linked to the interaction of immunology and epidemiology of macroparasitic diseases and malaria [27], [54], [28]. Until the 21st century little has been done linking immunological and epidemiological ideas in microparasitic diseases, such as viral pathogens.

Why is it important to develop and study multi-scale infectious disease models? (1) Because immunological considerations predict important epidemiological determinants, such as disease prevalence and reproduction number [19]. Such important epidemiological quantities can be explicitly related (for arbitrary parameter values) to host pathogen load and immune responses. (2) Data exist on both scales; linking them is essential for quantitatively coupling processes across scales. Furthermore, the biological understanding gained from more data will be more comprehensive and accurate. (3) Incorporating explicit immune responses is important in diseases, such as dengue, where disease severity depends on the strength of this response [14]. (4) These models are essential for elucidating the role of within-host disease dynamics for pathogen evolution [47], [15].

In this paper we review models linking within-host and between-host processes and some of the insights that have resulted from them. First, we detail some immunological models used in the coupled frameworks. The variability of these modeling techniques stems from the various types of epidemiological models from which they’ve arisen. There are four basic types of epidemiological models being employed: network or individual based epidemiological models, reviewed in subsection 3.A, ODE epidemiological models, reviewed in subsection 3.B, size-structured epidemiological models, reviewed in section 3.C, and age-since-infection structured epidemiological models, reviewed in subsection 3.D. Finally, in section 4, we provide further discussion of the immuno-epidemiological modeling and its implications to biology.

II. WITHIN-HOST MODELS

Within-host models are dynamical models that represent, in caricature, the interaction of the pathogen with the host replication machinery or immune defenses within a single host individual. Roughly speaking these models can be classified into three groups: models that depict the
reproduction processes of the pathogen within the host; models that depict the pathogen with the immune responses; and models that include both the replication of the pathogen and the immune responses. The simplest within-host models are of the first two types. We introduce here an example of each of the first two types as these types are the ones typically used in linked models.

A. Within-host model of the pathogen replication cycle

Within-host models that represent the pathogen replication cycle assume a typical viral pathogen that replicates using the machinery of host cells, called target cells. To introduce such a model, let \( x(\tau) \) be the number of pathogen-free (healthy) target cells (in the blood) and \( y(\tau) \) be the number of infected target cells. The amount of pathogen is denoted by \( P(\tau) \). The within-host replication model has been used for many viral diseases before [50].

\[
\begin{align*}
x' &= r - bPx - \mu x, \\
y' &= bPx - dy, \\
P' &= \nu dy - (\delta + s)P,
\end{align*}
\]

where \( r \) is the replication rate of target cells, \( b \) is the infection rate, \( \mu \) is the clearance rate of healthy cells, \( d \) is the clearance rate of infected cells, \( \nu \) is the number of pathogen particles released from lysis of an infected cell, \( \delta \) is the clearance rate of pathogen, and \( s \) is the shedding rate. The dynamics of the pathogen and target cells is shown in Figure 1.

Model (II.1) has been completely analyzed [17]. The reproduction number of the pathogen is given by

\[
\mathcal{R}_0 = \frac{rvb}{\mu(\delta + s)}.
\]

The model has two equilibria, an infection-free equilibrium \( E_0 = (r/\mu, 0, 0) \) and an infection equilibrium \( E^* = (x^*, y^*, P^*) \) where

\[
x^* = \frac{\delta + s}{\nu b}, \quad y^* = \frac{\mu(\delta + s)}{\nu bd (\mathcal{R}_0 - 1)}, \\
P^* = \frac{\mu}{b} (\mathcal{R}_0 - 1)^{\frac{1}{\mathcal{R}_0}}.
\]

B. Models with immune response

The adaptive immune response includes a cellular component which includes various types of T cells and humoral response which consists of B cells and antibodies. Simple pathogen-immune response models typically include only the pathogen \( P(\tau) \) and one type immune response cells, B-Cells, \( B(\tau) \). The pathogen replicates according to the Malthus model or logistic model. B cells kill the pathogen and B cell production is stimulated by the pathogen. B cells are cleared at rate \( d \):

\[
\begin{align*}
P' &= rP \left(1 - \frac{P}{K}\right) - \epsilon PB, \\
B' &= aP - dB,
\end{align*}
\]

where \( r \) is the parasite growth rate, \( K \) is its carrying capacity, \( \epsilon \) is the killing rate of the immune response, \( a \) is the activation rate of the immune response and \( d \) is the clearance of the immune response. The dynamics of model (II.4) is illustrated in Figure 2.

The model has two equilibria: \((0,0)\) which is always unstable and a coexistence equilibrium...
Fig. 2. The dynamics of the pathogen and the immune response. Parameter values are $r = 1$, $K = 1000$, $\epsilon = 0.1$, $a = 0.2$.

$(P^*, B^*)$ where

$$P^* = \frac{rdK}{\epsilon a K + rd}, \quad B^* = \frac{raK}{\epsilon a K + rd}. \quad (II.5)$$

Immunological models vary immensely in complexity and detail. Differences also stem from the particular processes in the disease being modeled. For our purposes here these two simple within-host models are sufficient to illustrate the concepts. In the next section, we introduce various epidemiological models.

III. IMMUNO-EPIDEMIOL O GICAL MODELS

The epidemiological component of the immuno-epidemiological models can take several different forms based on existing modeling frameworks, or just epidemiologically relevant quantities, such as the reproduction number, expressed in terms of the immunological variables. In this section we consider four types of immuno-epidemiological models structured by the type of epidemic model.

A. Network epidemic models

In network epidemic models, individuals are nodes in a network. Each individual or node can exhibit its own within-host dynamics (see Figure 3). The models can show the impact of the individual immune dynamics on population-level transmission of disease. Tucknell [57] and Kostova [34] introduced some of the first immuno-epidemiological models where the epidemiological component is a network. Kostova linked a network of $n$ within-host models and showed that even if the immune response clears the infection in each individual when isolated, while these individuals are in a network, the pathogen persists in each one of them and on a “population level”.

Vickers and Osgood [58] suggest that increased variance among people’s ability to respond to an infection, while maintaining the average immune responsiveness, may worsen the overall impact of an outbreak within a population. Furthermore, high values for the network connectivity reduced the timing between peak viral levels in neighboring individuals. A network based immuno-epidemiological model was applied by Vickers and Osgood [59] to study treatment of chlamydia which suggested that treatment applied up to the third day post...
infection has significant chance of preventing transmission of the disease to the nearest neighbor. Delivering treatment past the 3rd day post infection allows for infection of nearest neighbor as well as reinfection.

Lukens et al. [37] use a simple ODE model of influenza A and links that model, through infectivity, to large-scale agent-based population-level model to study influenza A epidemics. The authors obtain a map of the parameters of the immune model that characterizes clinical phenotypes of influenza infection and immune response variability across the population. At the population-level, effectively the authors simulate epidemics in Allegheny County, Pennsylvania and consider both age-specific and age-independent severity assumptions.

One of the serious drawbacks of network and agent-based immuno-epidemiological models is that very few population level quantities that describe the disease distribution can be computed analytically in closed form. In particular, in network models there are difficulties for computing the reproduction number and the prevalence of the disease. Consequently, little can be learned of the effect of the immune response on these quantities outside of extensive simulations [52].

The next three classes of immuno-epidemiological models remedy this shortcoming but some of them assume that all infected individuals undergo the same within-host dynamics, an assumption that is largely unrealistic. Still these models have contributed immensely to our further understanding of the mutual impact of within-host and between-host processes.

B. ODE immuno-epidemiological models

One way to obtain a simpler ODE immuno-epidemiological model for chronic diseases is to consider the immune model at infectious equilibrium. In this case one can make the parameters of a simple ODE epidemic model dependent on the equilibria values of the pathogen and/or the immune response. For instance, a simple SI model becomes:

$$ S' = \Lambda - \beta(P^*)SI - m_0S, $$
$$ I' = \beta(P^*)SI - (m_0 + m_1(P^*, B^*))I, $$

(III.1)

where $P^*$ and $B^*$ are given by (II.5) and $m_0$ is the natural death rate, $m_1$ is the disease-induced death rate, $\beta$ is the transmission rate, and $\Lambda$ is the recruitment rate. One can then investigate how within-host parameters, pathogen load and immune response affect the epidemiological quantities, such as disease-induced mortality, prevalence and reproduction number. These conclusions are not necessarily equivalent to conclusions obtained from other type immuno-epidemiological models such as the nested models considered below.

Another opportunity to connect the within-host and between host dynamical systems in an ODE model emerges in environmentally transmitted disease. In this case a chronological-time-structured within-host ODE system is linked to a chronological-time-structured ODE epidemiological system through the pathogen load in the environment (see Figure 4).

Using this novel modeling scenario Feng at al [12], [21], [22] investigate the transmission of Toxoplasma gondii. Feng at al [21] link a dynamic within-host model of the type (II.1), where the infection of target cells depends on population-level prevalence $I$, with an SI epidemic model much like (III.1), where transmission depends on viral load. Within-host and between-host reproduction numbers are computed but for the linked model analysis suggests that long-term behavior of the infection may depend on the initial conditions. Articles [12], [22] contain the environment as an explicit variable and find that infection may persist on population level even if the isolated between-host
reproduction number is less than one; a result that is facilitated by the within-host dynamics.

C. Size-structured PDE immuno-epidemiological models

Size-structured PDE immuno-epidemiological models are perhaps the most complex type of immuno-epidemiological models. In this case the epidemiological model consists of physiologically structured PDEs in which the structural independent variables are the dynamical variables of the ODE immune model. The first such model was proposed by [45] (see also [44]) where the “size structure” variable is the immune response. Analysis of this model revealed similarities to age-since-infection structured models particularly because the structural variable is strictly increasing in time. This model was further extended by [23] to model where the population-level density of infected individuals is structured by both the viral load and the immune response. [23] also transforms the size-structured model to an age-since-infection model where the independent variables are age-since-infection and initial pathogen inoculum. Considering a specific within-host model that allows for both pathogen extinction and unbounded growth, the authors investigate the population level impact of the initial inoculum and of the isolation threshold.

A somewhat different modeling approach to the same modeling scenario is given by [6], [7], [48]. The authors suggest coupling a classical HIV/HCV within-host model, given by equations (II.1) with a size-structured SIR epidemic model. The authors establish well-posedness in the special case when the density of infected individuals is structured only by the viral load. Furthermore, they develop numerical methods and discuss the impact of the (fixed) number of target cells and the burst size on the epidemic [6], [7]. The results allow them to determine the distribution of the density of the infected individuals by their viral load.

In general, the size-structured approach presents interesting mathematical challenges, such as the potential for measure-valued solutions, modeling issues related to creating and linking the models, computational issues with the large number of independent variables [6], [7], [48]. Because of these issues, creating the within-host model is an important step. Furthermore, the computational and analytical problems with the large number of independent variables somewhat restrict the incorporation of significant realism into the immune system; a problem that can be addressed by incorporating fewer within-host model dependent variables as independent variables in the epidemiological model. There are still a lot of interesting and open questions, related to this approach.

D. Nested models

Nested models are a relatively recent class of models suggested for the first time by Gilchrist and Sasaki [25]. The main advantages of nested models relative to size-structured models is that
these models allow for the use of very realistic and specific to a given disease models. Establishing well-posedness for these models is generally not very problematic [51]. Furthermore, since the number of independent variables in the PDE part of the model is restricted to two, there are few difficulties with computation, independent of the complexity of the immunological or epidemiological components.

Because of their advantages, since the Gilchrist and Sasaki paper, nested models have acquired significant popularity. The nested models “nest” a time-post-infection structured immune dynamics model into a time-post-infection and chronological time structured epidemiological model. Nested models also link the within-host model with the epidemiological model through the parameters of the epidemiological models that are expressed in terms of within-host dependent variables. To introduce a simple nested immuno-epidemiological model, let $S(t)$ denote the number of susceptible humans and $i(\tau, t)$ be the density of infected humans. In the simplest case, we can use an SI epidemiological model. The model takes the form:

$$S' = \Lambda - S \int_{0}^{\infty} \beta(\tau) i(\tau, t) d\tau - m_0 S,$$

$$i_{\tau} + i_t = -\left(m_0 + m_1(\tau)\right)i(\tau, t),$$

$$i(0, t) = S \int_{0}^{\infty} \beta(\tau) i(\tau, t) d\tau,$$

(III.2)

where $m_0$ is the natural death rate, $m_1$ is the disease-induced death rate, $\Lambda$ is the recruitment rate and $\beta$ is the transmission rate. Model (III.2) can be linked with either of the within-host models. The transmission rate $\beta(\tau)$ depends on the pathogen load. Experimental evidence suggests that the transmission probability does not increase linearly with the pathogen load but in a saturating fashion [35]. We will use the following simple function of saturating growth:

$$\beta(\tau) = \frac{c P(\tau)}{Q + P(\tau)}$$

where $c$ is the contact rate and $Q$ is the half-saturating constant. The disease-induced mortality can be linked in multiple ways to the immune system. It is thought that two distinct processes lead to disease-induced mortality in the host. On one side is the pathogen itself, and on the other is the immune response. We take the disease-induced mortality generated by the pathogen proportional to the pathogen load. The disease-induced mortality generated by the immune response has been taken to be proportional to the growth of the immune response $\alpha BP$ [25], [53]. We take here the disease-induced mortality proportional to $B^2$. The square guarantees that at low values of $B$, the immune response almost has no impact on the disease-induced mortality while at high levels of $B$, it has significant impact. This way a trade-off exists between the necessity of the immune response to be vigorous enough to clear the virus, but not too vigorous to kill the host. The disease-induced mortality is then given by

$$m_1(\tau) = \nu r P(\tau) + \xi B^2$$

where $\nu$ and $\xi$ are constants of proportionality. $\xi = 0$ if we are working with immune model (II.1).

One of the main disadvantages of the nested immuno-epidemiological models relative to the network and size-structured models is that they assume that all individuals exhibit the same immune dynamics. To remedy this disadvantage one may consider a multi-group immuno-epidemiological model where the different groups exhibit different immune dynamics. The multi-group model is somewhat complicated and obtaining analytical results on it is not easy. This weakens one of the great advantages of the nested models, namely that basic epidemiological quantities, such as the reproduction number and the prevalence, can be computed in analytical form. A schematic diagram of nested models is
The reproduction number of the immuno-epidemiological model (III.2) depends on the within-host variables and parameters:

\[
\mathcal{R}_0 = \frac{\Lambda}{m_0} \int_0^\infty \beta(\tau) e^{-m_0 \tau} e^{-\int_0^\tau m_1(\sigma) d\sigma} d\tau.
\]  

(III.3)

There are two types of questions being addressed with nested models: How does the within-host pathogen dynamics affect the population-level reproduction number and prevalence? What are the evolutionary and co-evolutionary consequences of the pathogen and host within-host evolution?

Several articles have suggested that the dependence of the reproduction number \(\mathcal{R}_0\) on the pathogen reproduction rate \(r\) may be non-monotone \([31], [53], [43]\). As the within-host pathogen reproduction rate \(r\) grows, it should be increasing the population-level reproduction number \(\mathcal{R}_0\); however the increased pathogen load is increasing host mortality, which in turn leads to decrease in the population-level reproduction number. This creates a hump-shaped form of \(\mathcal{R}_0\) as a function of \(r\), which is well-described in the literature. The dependence of \(\mathcal{R}_0\) on immunological parameters has been further discussed in \([32], [41], [34], [16], [8]\). The dependence of prevalence on the some of the within-host parameters may also be counter-intuitive. For instance, prevalence may decrease with increase of \(b\) in model (II.1) \([42]\). This in turn implies that within-host medication that lower the infection of target cells, that is decrease \(b\), de facto increase the population-level prevalence of the disease. This paradox has been observed in practice in HIV, where medications lower within-host virus load and increase survivability of infected individuals which leads to increasing prevalence. Furthermore, amplification of the HIV epidemic has been observed through nested models \([55]\).

More mathematical questions related to immuno-epidemiological models, such as well-posedness and optimal control are addressed in \([51]\).

The importance of multi-scale immuno-epidemiological modeling is best highlighted by its role in studying evolution \([47]\). Gilchrist and Sasaki were among the first to address co-evolution \([25]\) using multi-scale approaches but since then evolution of virulence has been attracting significant attention. Because evolution involves multiple interacting strains, a number of approaches have been developed to handle the emergent complexity \([3], [4], [5], [13], [15], [24], [26], [36], [38]\). One possible way is to model the strains explicitly on within-host and/or between-host scales \([42], [13]\). In the absence of trade-off mechanisms, the strain that maximizes its between-host epidemiological reproduction number dominates; a result first established rigorously mathematically in \([11]\). In the presence of trade-offs there is coexistence between the strains and invasibility is governed by population-level invasion numbers. Nested models further reveal \([2]\) that the optimal virulence in a co-infection model increases with multiple infections and that in a linked within-host and between-host co-infection model, an evolutionary stable strategy (ESS) can turn into a branching point \([1]\). An ESS is a strategy which, if adopted in the population, cannot be invaded by any other strategy. To see this from a pairwise invasibility plot (PIP)
(see Figure 6), we draw a vertical line through the singular strategy and confirm that the vertical line lies entirely in the non-invasibility region.

Nested models have the potential to link in a natural way the virus reproduction rate, population-level fitness, and population level disease-induced mortality. For HIV, the reproduction rates of the virus increase at a moderate rate and the virulence is slightly higher than the level that maximizes the population-level transmissibility of the virus [39]. For Hepatitis C, slowly replicating strains have a higher fitness and produce more population-level secondary infections while strains with higher replication rates dominate within a host [38]. For Influenza A, the relative importance of virulence and viral clearance by the immune system on the viral fitness and persistence was found to depend on the temperature [29].

Day et al [15] develop the mathematical theory that bridges the nested immuno-epidemiological models to quantitative genetics and evolution of traits. Using this framework, the authors show that the trade-off between transmission and virulence, studied early on in multi-scale models in [24], is an interplay of the genetic variation of the pathogen and the population-level dynamics of the disease. This framework is the backbone for future research at the interface of dynamic population modeling and quantitative genetic modeling.

Conclusions from multi-scale nested models for chronic diseases can be derived by writing the epidemiological quantities in term of the infected equilibrium of the within-host model [9], [26], [46]. This approach bridges to the ODE immuno-epidemiological models, discussed in section III.B. Following this approach article [26] found that within-host selection favors viral production rates $\nu$ that maximize virulence but between-host pathogen fitness is maximized at some intermediate virulence and viral production rate. Article [9] extends the results in [26] by incorporating superinfection.

IV. DISCUSSION

These new, more integrative modeling approaches, each have strengths and weaknesses mathematically and in terms of their levels of biological realism, but they have on the whole led to new and biologically counter intuitive insights. This is exemplified by even early examples of these models, such as Kostova [34], which demonstrated population-level disease persistence even when individual immune responses are able to clear infection and result in immunity. A second example is the population level persistence of virus, even when the between host reproduction number is less than 1, as shown by Feng et al.[22]. This result leads to the surprising interpretation that immune responses do not necessarily tip the balance of interactions in favor of the host and decrease disease prevalence. Similarly, drug administration may mimic the host immune system and increase disease prevalence (e.g., references [42] and [54]).

At another level, when we focus on the evolution of virulence, immuno-epidemiological models again lead to counter-intuitive results (e.g., [26]). The results of these analyses, that selection
within an individual can favor different pathogen traits than selection among individuals, highlight that the within-host/among-host model structure characteristic of these models meets the requirements for trait-group selection to play a role in evolutionary dynamics [60], where virus in individuals constitutes trait groups from which virus emerge, intermix and infect susceptible individuals. Wilson’s model has been criticized for unrealistic, or at least uncommon, assumptions of population structure (reviewed in [61]), however, infectious disease may represent a common context in which individual- and group-level selection both act strongly and at times in conflict.

Ultimately, understanding the evolution of virulence might be informed by models designed to understand trait evolution in the context of multi-level selection such as those discussed in [56]. In general, because there are sometimes conflicts between within-host and among-host virulence optima, new insights will likely come from relaxing the biologically-unrealistic, albeit simplifying, assumption that virulence evolves to its within-host optimum prior to the epidemiological dynamics of the model. Again, modeling approaches appropriate for relaxing this assumption might be informed by the body of literature focused in the modeling of group and individual level selection, especially those approaches that consider the continuous nature of the dynamics of within-individual and among-individual processes [56].

ACKNOWLEDGMENTS

The authors acknowledge support from NSF grants DMS-1515661/DMS-1515442.

REFERENCES

an environmentally-driven infectious disease, within-host and between-host dynamics in Nonlinear Dyn. 68 (2012), p. 401-411.


