Vector-borne disease models: R0 and beyond

Hartemink, N.

Publication date
2015

Document Version
Final published version

Published in
The impact of a decade (2004-2015) of research on vector-borne diseases

Citation for published version (APA):
The impact of a decade (2004-2015) of research on vector-borne diseases
Contributors

Authors:

Co-Editors: Renaud Lancelot, Sylvie Laurens and Andrew Lewer
With the collaboration of: Marieta Braks, Simon Carpenter, Guy Hendrickx and Heikki Henttonen

Graphic design and layout: Laurence Laffont

EDEN (GOCE-CT-2003-010284) is an Integrated Project (2004-2009) of the 6th Framework Programme (FP6) funded by the European Commission under the Research DG, Environment Directorate, priority 6.3: Global Change and Ecosystems
EDENext (261504) is a research project (2011-2014) of the 7th Framework Programme (FP7) funded by the European Commission under the Health component.

This book was funded by EU grant FP7-261504 EDENext and is catalogued by the EDENext Steering Committee as EDENext420 (http://www.edenext.eu). The contents of this publication are the sole responsibility of the authors and don’t necessarily reflect the views of the European Commission.
Vector-borne disease models: \( R_0 \) and beyond

**INTRODUCTION**

The basic reproduction number, \( R_0 \), has a long history, especially in the field of vector-borne diseases. This paper provides a brief overview of the use of this concept in vector-borne disease epidemiology as well as new developments and recent applications, notably within the EDEN and EDENext framework. Furthermore, it describes a novel framework for the inclusion of landscape factors and animal movements in vector-borne disease risk assessments, which can be merged with more quantitative approaches in order to achieve a mechanistic and realistic description of the spatial and temporal variation in vector-borne disease risk.

**\( R_0 \) FOR VECTOR-BORNE DISEASES**

The basic reproduction number, \( R_0 \), is a key concept in the epidemiology of infectious diseases. It is defined as the expected number of secondary cases caused by one infectious individual introduced into a naïve population. It can be used as a measure of the potential success of invasion of a pathogen into a population; if the value of \( R_0 \) is higher than 1, an outbreak of the infectious agent is possible, whereas if \( R_0 \) is less than 1, the infection will die out (Anderson and May 1991, Diekmann and Heesterbeek 2000).

The theory underlying \( R_0 \) models for vector-borne diseases (hereafter VBDs) originated in the early 20th century, when Ronald Ross first developed a theoretical framework to describe the transmission of malaria. His work was extended by several others, including George MacDonald in the 1950s and 1960s and, by 1970, the foundations of the Ross-MacDonald theory were established (MacDonald 1957, Smith et al. 2012). The by now well-known Ross-MacDonald formula for the transmission of malaria is usually written as:

\[
R_0 = \frac{ma^2bc p^n}{(-\ln(p))r}
\]

where \( m \) is the ratio of mosquitoes to humans, \( a \) the mosquito biting rate (on humans), \( b \) and \( c \) the pathogen transmission efficiencies, \( p \) the daily survival rate of mosquitoes, \( r \) the recovery rate in humans (i.e. the reciprocal of the infective period of the host) and \( n \) the duration of the extrinsic incubation period (EIP). Since the development of this Ross-Macdonald model for malaria, hundreds of models for mosquito-borne diseases have been derived and published, and most of them resemble the original models quite closely (Reiner et al. 2013). The Ross-MacDonald formula applies to situations where one vector species and one host species are involved in the disease transmission.

For those situations where multiple types of infected individuals are involved or where multiple transmission routes are possible, the \( R_0 \) value can be derived using the next-generation matrix approach (Diekmann and Heesterbeek 2000). The first step in this approach is to identify the different types-at-birth in the system, in other words to categorise individuals by their state at the moment they become infected. These types-at-birth differ in traits that affect their ability to produce secondary cases, such as infectivity, contacts, life history traits, possible transmission routes etc. For VBDs, there are at least two types-at-birth (host and vector), but often, there are more. For a system with \( m \) types-at-birth, the next-generation matrix (hereafter NGM) will be...
an $m \times m$ matrix, where each element $k_{ij}$ equals the expected number of new cases of type-at-birth $i$ caused by one individual of type-at-birth $j$.

The basic reproduction number $R_0$ is calculated as the dominant or largest eigenvalue of the NGM (Diekmann, Heesterbeek and Roberts 2010, Diekmann, Heesterbeek and Britton 2013). For a 2 by 2 matrix, the matrix looks like this:

$$K = \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix}$$

and the corresponding expression for $R_0$ is:

$$R_0 = \frac{1}{2} \left( (k_{11} + k_{22}) \pm \sqrt{4k_{12}k_{21} + (k_{11} - k_{22})^2} \right)$$

For a VBD system with only one vector, one host and no direct transmission (i.e. $k_{11}$ and $k_{22}$ are equal to zero), this approach gives the same result as the Ross-MacDonald formula. In fact, the Ross-MacDonald $R_0$ is the square of the $R_0$ obtained by the NGM method. That is, the Ross-MacDonald formula calculates a person-to-person basic reproduction number (via a mosquito, so there are actually two transmission steps), whereas the NGM method calculates a reproduction number from individual to individual (i.e. the average of the vector-to-host and host-to-vector reproduction), so for one transmission step.

The main advantage of the NGM approach is that it can deal with situations with multiple hosts and/or multiple vectors, as well as with different transmission routes. Direct transmission (for example, via transovarial transmission in mosquitoes, or transplacental or direct contact transmission in hosts) can be easily incorporated, as was shown in a study on the impact of direct transmission between crows on WNV transmission (Hartemink et al. 2007), part of the EDEN project. Another EDEN modelling study illustrated how the NGM method can be applied to a complex disease system, namely tick-borne diseases transmitted by *Ixodes ricinus* ticks, which involves five different types of infected individuals and three different transmission routes (Hartemink et al. 2008). Efficient numerical methods for the calculation of the eigenvalues exist, also for a 3 by 3 (or higher-dimensional) matrix (Stoer and Bulirsch 1983) and most common statistical software packages can routinely perform these calculations.

It is important to note that some modelling studies which calculate $R_0$ assume exponential rates. It means that authors use a fixed rate for mortality and for becoming infectious. Since the extrinsic incubation period (EIP), that is, the period required for pathogen development in the vector, is more likely to last a fixed number of days (during which no vector becomes infectious), this assumption is often not met. This is particularly so when the EIP is long compared to the life span of the vector, which is often the case in vector-borne disease systems, and it is better to use the assumption used in the Ross-MacDonald formula, which is more realistic from a biological point of view (Hartemink, Cianci and Reiter 2015).

$R_0$ maps are maps indicating values of $R_0$ for different locations, calculated for a given infectious agent when taking the local conditions into account with respect to hosts (and possible vectors) and their environment. They can be used to identify areas with a higher probability of a major outbreak after this agent is introduced. This concept has been used to develop risk maps for directly transmitted diseases such as foot-and-mouth (Ferguson, Donnelly and Anderson 2001, Keeling et al. 2001), avian influenza (Boender et al. 2007) and classical swine fever (Boender et al. 2008). Within the EDEN project, $R_0$ maps were constructed for two vector-borne disease systems: bluetongue (Hartemink et al. 2009) and canine leishmaniasis (Hartemink et al. 2011).
Several $R_0$ maps have now been developed and published for VBDs, including maps for bluetongue virus (Racloz et al. 2008, Brugger and Rubel 2013), chagas disease (Cordova et al. 2014) and tick-borne diseases such as Lyme disease (Wu et al. 2013) and Crimean-Congo haemorrhagic fever (CCHF) (Estrada-Peña et al. 2013).

**IMPACT OF TEMPERATURE**

$R_0$ maps can also reflect the temperature dependency of disease transmission. Several processes influencing transmission dynamics are known to vary with temperature, and this can be taken into account by making the corresponding parameters in the expression for $R_0$ temperature-dependent. Temperature can have opposing effects on different ingredients of $R_0$, a phenomenon that has been discussed by, amongst others, Rogers and Randolph (2006). Therefore, caution should be taken when using models to predict the effect of climate change; for instance, models taking only the effect of temperature on the pathogen development rate into account are just too simplistic (Reiter 2001, Reiter et al. 2004). Parameters that are known to be temperature-dependent for many VBD systems are i) the biting rate, ii) the rate of development of the pathogen in the vector (which determines the length of the extrinsic incubation period or EIP) and iii) the mortality of the vector. Even parameters often assumed to remain constant (such as the transmission efficiencies $b$ and $c$ in the expressions for $R_0$) are affected by temperature in one or more vector species. For example, *Culicoides nubeculosus*, a midge that is normally refractory to the strains of bluetongue virus circulating in Europe at the present time (i.e. $b=0$), is capable of transmitting bluetongue virus when its larvae or pupae are kept at higher than normal temperatures, as is discussed in Rogers and Randolph (2006).
While $R_0$ and the NGM approach offer a promising tool to assess the risk of establishment, by means of the per-generation growth rate of a starting epidemic, as well as a way to quantify the contributions of different ingredients or transmission routes to this growth rate, there are also some limitations. One of them is that it is not straightforward to include animal movements and landscape factors into the $R_0$ calculation. One of the assumptions underlying the $R_0$ concept is that there is local homogeneous mixing, that is, we consider a population where all individuals can get in contact with each other. An $R_0$ map thus gives local values of the expected number of infected individuals resulting from one infected individual in a naïve population present at that specific location. An $R_0$ map at 1km resolution assumes that there is homogeneous mixing within the 1km grid cell or pixel, but no substantial exchange of individuals between cells. Also the configuration of certain landscape factors, and its effect on how animals move through the landscape, cannot easily be incorporated in $R_0$ formulas, although some studies on metapopulations have looked at the effect of migration rates (for example, Jesse and Heesterbeek 2011).

**ROLE OF LANDSCAPE FACTORS, CONNECTIVITY AND FRAGMENTATION**

Landscape, and elements characterising it, such as land use, land cover, composition and structure, have been suggested to have a considerable impact on the dynamics and therefore the occurrence of VBDs (Ostfeld, Glass and Keesing 2005, Lambin et al. 2010). Landscape, defined as all the visible features of an area of land (Oxford Dictionary), includes the physical elements of landforms (such as mountains, hills, water bodies such as rivers, lakes, ponds and the sea), living elements of land cover (including indigenous vegetation), human elements (including different forms of land use, buildings and structures) and transitory elements such as lighting and weather conditions (Human Geography Dictionary).

These landscape elements will affect the presence, spread and persistence of vector-borne and zoonotic diseases as they may influence the presence of vectors, reservoir hosts, susceptible hosts (including humans) and the pathogen, as well as their overlap in time and space. Land use will, among other things, affect the number of people present in an area and the way they use the landscape, which in turn affects their chances of getting in contact with the pathogen, via the vector or, for example, in the case of hantavirus, via contaminated soil. Land cover will, at least partly, determine the availability and quality of habitat of the different species which interact with the pathogen, i.e. the vectors and hosts. Many of these aspects had been covered in EDEN, as summarised in the paper by Lambin et al. (2010). In EDENext, the focus shifted, based on the notion that it is not just the landscape composition, but also the landscape configuration and connectivity that will affect the suitability of an area for a pathogen to establish, spread and/or persist. Habitat connectivity and habitat fragmentation are likely to have an effect on the dispersal of animals and thereby of the pathogens they carry.

Characterising and quantifying the level of connectivity of a habitat or landscape is not straightforward. Notwithstanding the importance of the concept of connectivity, there is no generally accepted and employed formal definition (Crooks and Sanjayan 2006) and many different measures for connectivity have been applied in various research fields. In the field of landscape ecology, it has long been recognised that habitat connectivity is determined by two different components (Bennett 1990, 2003). The structural component of connectivity is determined by the spatial arrangement of different types of habitats in the landscape. It refers to the mappable, spatial arrangement of habitats. The behavioural component of connectivity relates to the behavioural response of individuals and species to the physical structure of the landscape.

Consequently, even though living in the same landscape, species with contrasting behavioural responses (to habitat disturbance, for example) will experience different levels of connectivity (Bennett 2003). In the field of vector-borne disease epidemiology, this distinction between the structural and behavioural connectivity is often not made and landscapes are classified as well connected or fragmented mostly based on human perception of the landscape. Given the fact that the effect of landscape configuration on connectivity, population density and contact rates is species- and disease-system specific, and the fact that the way to measure connectivity is not well-defined (and in many cases probably overlooks the way that animals, including humans, actually use the landscape), it is not surprising that different studies on the effect of habitat connectivity and fragmentation on disease risk come up with different answers.
In EDENext, a modelling study looked into the effect of fragmentation on animal movement and tick-borne disease transmission by means of an agent-based model, which takes landscape-specific dispersal rates into account. This study predicts a positive effect of fragmentation level on disease risk (Li et al. 2012). Another study, using a multilevel statistical analysis to find landscape factors associated with the occurrence of hantavirus cases in humans, found a positive association with well-connected forest. It is clear that there is no straightforward relationship between landscape configuration and pathogen transmission; the relationship will be different for each disease system and will depend on the way the vector and host species use the landscape, their dispersal capacity, their habitat requirements and many other factors.

**RESOURCE-BASED HABITAT CONCEPT (RBHC)**

In order to overcome the problem of not being able to include landscape factors and animal movements in our risk assessments, several members of the EDENext modelling team worked together to explore approaches used in other research fields to see whether these approaches would be applicable in the VBD context. It became clear that the research field of vector-borne disease risk modelling shares many goals and challenges with the field of conservation biology. In both fields, the aims include the mapping of other research fields to see whether these approaches would be applicable in the VBD context. It became clear that the research field of vector-borne disease risk modelling shares many goals and challenges with the field of conservation biology. In both fields, the aims include the mapping of organisms of interest and the study of environmental factors which help the prediction of presence, abundance or dynamics of focal species.

One particular concept, the resource-based habitat concept, has been explored in great detail and seems to be a promising framework to identify systematically the different ecological resources that are necessary for the completion of the transmission cycle and to relate these resources to (combinations of) landscape features and other environmental factors (Hartemink et al. 2014). The framework has been developed in the field of conservational biology, with special application to butterfly conservation, to help identify functional habitats from a mechanistic, biological viewpoint (Dennis, Shreeve and Van Dyck 2003, Vanreusel and Van Dyck 2007). The concept facilitates the study of functional habitats of the species involved in pathogen transmission (i.e. pathogen, vectors and hosts) by regarding the resources required by each of the species and linking them to specific environmental features at the landscape level.

Using the movement capacities of the animals involved, the potential functional habitat range (i.e. the delineated zones that correspond to the combined and integrated information on resources, utilities and scale of movement) can be determined for each species. Once the functional habitat of the vector and host species is determined, the functional habitat of the pathogen can be delineated, by looking at the areas when transmission can occur (i.e. where functional habitats of vector and host overlap and other (thermal) requirements of the pathogen are met).

**REFERENCES**


Diekmann O, Heesterbeek JAP and Roberts MG (2010). The construction of next-generation ma-
In this way, the concept provides new insight into spatial and temporal variation in transmission opportunities and exposure that ultimately determine disease risks. Since application of the concept requires a systematic inventory of the resource required by each of the interacting species, it may also be a tool to identify knowledge gaps. An interesting advantage of the method is that it forces us to think about the spatial configuration of resources and how these influence disease risk, which may lead to novel control options; changing the spatial configuration of key resources across the landscape may decrease the overlap of functional habitat of vector and host, or may decrease the vector population, if essential resources for the vector are not placed within flight range distance.

An interesting example is a study in South America, where transmission of oropouche virus was decreased by removing rotten banana stumps or cacao husks, which are breeding sites for the main vector, C. paraensis, in the vicinity of houses (Hoch, Roberts and Pinheiro 1986). Removal of bluetongue-virus vector (Culicoides spp.) breeding sites has also been shown to be useful in reducing vector abundance in North America for C. sonorensis (Linley, Evans and Evans 1970; Mullens and Rodriguez 1989; Carpenter, Mellor and Torr 2008a), even though this has not always translated into reduction of bluetongue infection as
measured by sero-conversion rate (Mayo et al. 2012). So, in summary, RBHC allows for the integration of landscape factors and animal movements in epidemiological spatial risk assessments and hence bridges the gap between existing mechanistic modelling approaches that ignore landscape factors and dispersal and satellite image-based approaches that are based on statistical inference only.

**MERGING RBHC AND $R_0$: INCORPORATING LANDSCAPE FACTORS AND MOVEMENT RANGE IN A QUANTITATIVE WAY**

While RBHC may therefore offer promising scope to deal with spatial modelling of VBD risks, it is currently only a conceptual model that gives a qualitative risk assessment, but lacks the advantage of an integrating quantitative approach, such as $R_0$. The next step would be to take the RBHC framework to the next level, and make it a quantitative approach, by parameterizing it for a case study. This would allow us to take advantage of the properly weighted way in which the $R_0$ formula combines all the ingredients of the transmission cycle and the spatially explicit and biologically realistic way in which RBHC incorporates the functional use of the landscape.

**ACKNOWLEDGEMENT**

This work would not have been possible without the dedication and commitment of all the researchers involved in the EDEN and EDENext modelling teams. We thank the following institutions:

- Université Catholique de Louvain, Belgium
- Université Libre de Bruxelles, Belgium
- University of Oxford, United Kingdom
- University of Utrecht, The Netherlands


