Introduction	Modelling	Results	Discussion
000000	0000000	000000000	00000
			1

Co-infections by non-interacting pathogens are not independent and require new tests of interaction

Frédéric Hamelin

Joint work with Nik Cunniffe, Linda Allen et al.

Webinaire de Modélisation et Statistique en Santé des Animaux et des Plantes

24 novembre 2020

Introduction	Modelling	Results	Discussion
●000000	000000	000000000	00000

CO-INFECTIONS AND INTERACTIONS

HPV AS AN EXAMPLE FOR SIMILAR PATHOGENS

Co-infection is the simultaneous infection of one host by multiple pathogen species or strains or clones or types, from now on **pathogens**.

Human papillomavirus

- many HPV types
- high-risk (oncogenic) types
- co-infections are common
- within-host interactions between types is debated
- vaccinating against one type may impact other types



Introduction	Modelling	Results	Discussion
000000	000000	000000000	00000

CO-INFECTIONS AND INTERACTIONS

LYME DISEASE AS AN EXAMPLE FOR DISTINCT PATHOGENS

Ticks are "toxic soup" and coinfections are the rule, including babesiosis, anaplasmosis, Rocky Mountain spotted fever, *etc.*



©lymedisease.org

Introduction	Modelling	Results	Discussion
000000	000000	000000000	00000

CO-INFECTIONS AND INTERACTIONS

MALARIA AS AN EXAMPLE OF TESTING FOR INTERACTIONS

Forbes (1907) and Cohen (1973) introduced

ratio of observed coinfections to

S = the number expected if pathogens were statistically independent

lf

- S > 1 pathogens positively associated
- S = 1 no significant association
- S < 1 pathogens negatively associated

Statistical associations are taken as a signal of biological interactions





FIG. 1. MALARIAL INFECTIONS IN 6 COUNTIES IN TRINIDAD AND TOBAGO

Estimates (x) and 99% confidence intervals (vertical bars) of s as a function of parasite rate (fraction of the population with malarial infection). See Table 1, based on Downs, Gilletter, and Shannon (1945). In this and all of the following figures, y is the ratio of the number of mixed infections obpexted II direct were no interactions among the parasite species.

Introduction	Modelling	Results	Discussion
000000	000000	000000000	00000

HOW DID I CAME TO THIS TOPIC

MAIZE LETHAL NECROSIS AS A MOTIVATING EXAMPLE

MLN is a disease caused by co-infections by two viruses

MLN Distribution and Spread 2011- 2016



Data Sources: KALRO, KEPHIS, FAO, EIAR, NARO, RAB, MARI, UCG, IITA, CIMMYT

NIMBioS Working Group "Multiscale Vectored Plant Viruses" focused on MLN

Introduction	Modelling	Results	Discussion
0000000	0000000	000000000	00000

CO-INFECTION MODEL TO INFORM CONTROL OF MLN

HILKER ET AL (2017) PHYTOPATHOLOGY

- Analyzed a (very) simple model of MLN
- Assumed the two viruses do not interact (within the season)
- Multiplicative prevalence in field data support this (strong) assumption
- Multiplicative prevalence means statistical independence:



Percentage with SCMV x Percentage with MCMV (S x M/100)

Fig. 3. Field data support the assumption of independent transmission of Maize chlorotic mottle virus (MCMV) and Sugarcane mosaic virus (SCMV). MLN = maize lethal necrosis, Field data as reported by Mahuku et al. (2015) are replotted, showing the results of survey D in Kenva in 2013 and 2014

 $P(\text{co-infection}) \approx P(\text{infected with SCMV}) \times P(\text{infected with MCMV})$

Introduction	Modelling	Results	Discussion
0000000	000000	000000000	00000

INTERACTION TESTS ARE BASED ON INDEPENDENCE

- Methods for cross-sectional data based on independence tests
- ► Simple example:
 - observations (N = 1000)

	Pathogen A			
		Infected	Not infected	Sum
Pathogen B	Infected	25	175	200
	Not infected	75	725	800
	Sum	100	900	1000

Expectation under independence hyp.

		Pathogen A		
		Infected	Not infected	Sum
Pathogen B	Infected	20	180	200
	Not infected	80	720	800
	Sum	100	900	1000

Chi-square test:

 $\chi^2(1) = 1.74 \Rightarrow p > 0.05 \Rightarrow$ no evidence from data the pathogens interact

May be dressed up in more complex statistics: log-linear or other regressions accounting for confounding factors (e.g. risk group)

Introduction	Modelling	Results	Discussion
000000	0000000	000000000	00000

QUESTION

Is it right to assume that independence means non-interaction?

Introduction	Modelling	Results	Discussion
000000	000000	000000000	00000

SIMPLEST 2 NON-INTERACTING PATHOGENS MODEL

- β_1, β_2 : infection rates
- μ : natural death rate
- *I_i*: prevalence of pathogen *i* = 1,2

$$\dot{I}_i = \beta_i I_i (1 - I_i) - \mu I_i,$$

$$\bullet \ F_i = \beta_i I_i = \beta_i (J_i + J_{1,2})$$

$$\begin{array}{rcl} \dot{J}_1 &=& F_1 J_{\emptyset} - (F_2 + \mu) J_1 \; , \\ \dot{J}_2 &=& F_2 J_{\emptyset} - (F_1 + \mu) J_2 \; , \\ \dot{J}_{1,2} &=& F_2 J_1 + F_1 J_2 - \mu J_{1,2} \; . \end{array}$$

Sinaly-infected by pathogen 1 Infection Death Infection $(\beta_1 I_1)$ $(\beta_2 I_2)$ (u) Death (µ) Uninfected Co-infected Death (µ) Infection Infection $(\beta_2 l_2)$ $(\beta_1 I_1)$ Sinaly-infected by pathogen 2

► $J_{\emptyset} = 1 - J_1 - J_2 - J_{1,2}$

Introduction	Modelling	Results	Discussion
000000	000000	000000000	00000

PREVALENCES ARE NOT MUTIPLICATIVE



For $R_{0,i} = \beta_i / \mu > 1$, the equilibrium prevalence of co-infection is

$$\bar{J}_{1,2} = \left(\frac{\beta_1 + \beta_2}{\beta_1 + \beta_2 - \mu}\right) \bar{I}_1 \bar{I}_2 \ge \bar{I}_1 \bar{I}_2 \,.$$

Co-infections are more likely than expected by chance.

Introduction	Modelling	Results	Discussion
000000	000000	000000000	00000

PREVALENCES ARE POSITIVELY CORRELATED

Stochastic version: pathogen prevalences correlated



 I_1 and I_2 go down simultaneously whenever co-infected dies.

$$\operatorname{cov}\left(\frac{l_1}{N},\frac{l_2}{N}\right) = \frac{(\beta_1 + \beta_2)(\beta_1 - \mu)(\beta_2 - \mu)\mu}{N\beta_1\beta_2(\beta_1 + \beta_2 - \mu)(\beta_1 - \mu + \beta_2 - \mu)} \ge 0.$$

APPARENTLY NOT A WELL KNOWN RESULT

May & Nowak (1995) Proc. Roy. Soc. B Cited 351 times

Assumes prevalences are multiplicative

Would have been useful to reduce the complexity of co-infection models

Unfortunately this independence assumption is not correct

9 · Super- and Coinfection: The Two Extremes

133

Box 9.3 SI models accounting for coinfection

With i_j denoting the fraction of individuals harboring strain j (possibly in addition to various other strains), a simple model for coinfection is

$$\frac{di_j}{dt} = i_j [\beta_j (1 - i_j) - d - \overline{\alpha}_j], \qquad j = 1, \dots, n.$$
(a)

The total population size of hosts is assumed to be held constant, and is normalized to one. The infectivity (transmission rate) of strain j is denoted by β_j . Strain j can invade any host that is not already infected by strain j. Thus $\beta_j i_j (1 - i_j)$ is the rate at which new infections with strain j occur.

There is a natural death rate d and a disease induced death rate $\overline{\alpha}_j$ which denotes the average death rates of hosts infected by strain j, and is assumed to be given by the strain with the highest virulence in the host. We define p_j as the probability that a host is not infected with a strain *more* virulent than j. That is,

$$p_j = \prod_{k=j+1}^{n} (1-i_k)$$
 (b)

Note that $p_n = 1$ and $p_i = (1-i_{j+1})p_{j+1}$. The fraction of hosts that are uninfected is given by $p_0 = \prod_{k=1}^n (1-i_k)$. The probability that k is the most virulent strain found in a host is $i_k p_k$, and

$$\overline{\alpha}_j = \alpha_j p_j + \sum_{k=j+1}^{n} \alpha_k i_k p_k$$
. (c)

This coinfection model is completely defined by Equations (a) to (c). We note that infection and death rules are devised such that if the strains are randomly assorted relative to each other, this continues to be the case, so that Equation (a) remains correct.

Nowak & Sigmund (2002)

Introduction	Modelling	Results	Discussion
000000	0000000	000000000	00000

KUCHARSKI AND GOG (2012)'S RESULT

- Model reduction in multi-strain influenza models
- Continuous age-structured model
- ► Prevalences are multiplicative within infinitesimal age-classes only: ∀a > 0,

$$j_{1,2}(a) = i_1(a)i_2(a)$$

• We recover previous result with p(a) the age distribution:

$$J_{1,2} = \int_{\mathbb{R}} j_{1,2}(a)p(a)da = \int_{\mathbb{R}} i_1(a)i_2(a)p(a)da$$
$$\geq \left(\int_{\mathbb{R}} i_1(a)p(a)da\right)\left(\int_{\mathbb{R}} i_2(a)p(a)da\right) = I_1I_2$$

if $i_1(a)$ and $i_2(a)$ are increasing with *a* (Harris inequality).

Introduction	Modelling	Results	Discussion
000000	0000000	000000000	00000

KUCHARSKI AND GOG (2012)'S INTERPRETATION

Toxoplasma gondii AND Helicobacter pylori AS EXAMPLES

- One way of understanding the result is in terms of aging
- Individuals acquire more infections as they age
- ► As age increases, so does the probability of being infected
- ► Therefore, the prevalences of pathogens are correlated



Introduction	Modelling	Results	Discussion
000000	000000	000000000	00000

QUESTION

Mathematics say non-interacting pathogens are not independent, but does that change anything in practice?

Introduction	Modelling	Results	Discussion
000000	000000	●00000000	00000

DATA-DRIVEN TESTS OF PATHOGEN INTERACTIONS

MALARIA AS AN EXAMPLE

Cross-sectional data on co-infection in one of two forms:

1. Full infection profiles

TABLE 2									
Distribution	of	single	and	mixed	infections	in	the	study	group'

	Observed (O)			(E = O) ²	
	n = 590	96	Expected (E)	E	
P.f. only	323	54.7	331.6	0.2	
P.m. only	1	0.2	20.4	10.0	
P.o. only	1	0.2	11.2	10.8	
P.f. + P.m.	41	6.9	43.9	0.2	
P.f. + P.o.	12	2.0	24.2	6.2	
P.m. + P.o.	-	-	1.5	1.5	
P.f. + P.m. + P.o.	27	4.6	3.2	177.0	
Negative	185	31.4	153.9	6.3	

* Pf = Plasmodium falciparum; Pm = Plasmodium malariae; Po. = Plasmodium ovale; - = not detected; χ² = 219.1, degrees of freedom = 7, P < 0.00001.</p>

$2^3 = 8$ possible combinations of 3 pathogens

2. Multiplicity of infection



Figure 2. Distribution of multiple clone infections. Distribution of multiplicity of infection as detected by the markers msp1F3 and MS16 as well as both markers combined. Only samples with positive results for both markers were included (n = 1050). doi:10.371/journal.pntd.0001424.g002

distribution of the number of pathogens hosted per host

Introduction	Modelling	Results	Discussion
000000	000000	00000000	00000

1. NON-INTERACTING DISTINCT PATHOGENS (NIDP)

- Analogous n pathogen model
- Each pathogen has a distinct

$$R_{0,i} = \frac{\beta_i}{\mu}$$



Find equilibrium values

e.g.
$$I_1 = J_1 + J_{1,2} + J_{1,3} + J_{1,2,3}$$

Simple enough for recursive solution

$$\bar{J}_{\Gamma} = \frac{\sum_{i \in \Gamma} \left(R_{0,i} - 1 \right) \bar{J}_{\Omega_i}}{1 + \sum_{i \notin \Gamma} \left(R_{0,i} - 1 \right)}.$$

- Γ: Some combination of pathogens
- $\Omega_i = \Gamma \setminus \{i\}$: One fewer pathogen

Introduction	Modelling	Results	Discussion

2. NON-INTERACTING SIMILAR PATHOGENS (NISP)

- Multiplicity of Infection (Mol) data: as many parameters in NiDP as data points, so NiDP is overparameterized
- NiSP makes the strong assumption that all pathogens are interchangeable: R_{0,i} = R₀ (1 parameter, more parsimonious).



Probability individual carries *k* distinct infections given *n* pathogens in total Leads to even simpler recursive solution:

$$ar{M}_k = rac{(n-k+1)(R_0-1)}{(n-k)(R_0-1)+1}ar{M}_{k-1}\,.$$

Introduction	Modelling	Results	Discussion
000000	000000	00000000	00000

USING THE MODELS IN PRACTICE ON REAL DATA

Form of model – NiDP or NiSP – **driven by data** (and whether or not NiSP assumption of equal R_0 acceptable)

Compare fit of model against bi/multi-nomial distributions

- i.e. compare against assuming statistical independence
- fit models via maximum likelihood
- compare using AIC

Test whether NiSP or NiDP is sufficient to explain the data, i.e. test whether data shows evidence of interaction

- Monte Carlo goodness of fit test
- Repeatedly simulate fitted model and check if likelihood of the data is too far in the tail of distribution over all sims



n = 25 (strains), N = 5412 (hosts) n = 102 (strains), N = 475 (hosts)

Introduction 0000000	Modelling	Results ○○○○○●○○○○	Discussion
	<u> </u>		

USING THE NISP CT'D



NiSP outperforms binomial

- ► in all cases
- particularly in the tails (co-infection much more frequent than expected)

For HPV (Panel A)

- goodness of fit
- i.e. data set consistent with idea types of HPV do not interact

Lack of fit in other cases meaning either:

- interaction
- epidemiological differences between pathogens (recall NiSP assumes equal R0)
- confounding factor (other than age)

USING NIDP TO UNDERSTAND INFECTION PROFILES MALARIA AS AN EXAMPLE

3 species: Plasmodium falciparum, P. malariae, P. ovale

Am. J. Trop. Med. Hyg., 1963), 1980, pp. 121-337 Copyright © 1980 by The American Society of Tropical Medicine and Hygiets

> A LONGITUDINAL STUDY OF HUMAN MALARIA IN THE WEST AFRICAN SAVANNA IN THE ABSENCE OF CONTROL MEASURES: RELATIONSHIPS BETWEEN DIFFERENT PLASMODIUM SPECIES, IN PARTICULAR P. FALCIPARUM AND P. MALARIAE*

LOUIS MOLINEAUX, JOHN STOREY, JOEL E. COHEN,† AND ANTHONY THOMAS World Health Organization, 1211 Geneva 27, Suitterland, and Rockefler University, New York, New York 10021

Abtract. The research project on the epidemiology and control of multitat conducted in the Garki Dirict, Anno State, joinfly by the Government of Neigeria and the World Heikh Organization included among its objectives the study of the baseline epidemiology prior to respect to the relationships name the three species of Ramonim present, P_i deforms, P_i making and P_i multiple study of the priority of the

 $\begin{array}{l} R_{0,F} = 2.47, \, R_{0,M} = 1.15, \\ R_{0,O} = 1.03, \, \Delta \, \text{AIC} = 362.2, \\ \rho(\text{GoF}) = 0.40 \end{array}$



No evidence from these data that the pathogens interact

REVISITING THE ANALYSIS OF HOWARD ET AL (2001)

Methods for estimation of associations between multiple species parasite infections

S. C. HOWARD*, C. A. DONNELLY and M.-S. CHAN

Wellcome Trust Centre for the Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3FY

No.	Reference	Country		Ages	Species	Sig. associations
68	Campbell et al. (1987)	Kenya	147	Children	FMO	FM', MO'
69	Campbell et al. (1987)	Kenya	142	Children	FMO	FM ⁺ , MO ⁺
70	Campbell et al. (1987)	Kenya	196	Children	FMO	FM [*] , MO [*]
71	May et al. (1999)	Nigeria	230	Children	FMO	FM ⁺ , MO ⁺
72	May et al. (1999)	Nigeria	59	Children	FMO	
73	May et al. (1999)	Nigeria	142	Children	FM	
74	Legger et al. (1923)	Nigeria	250	Children	FMV	FM ⁻
75	Bedier et al. (1924)	Nigeria	135	Children	FMV	FM ⁻
76	Knowles & White (1930)	Sierra Leone	809	Children	FMV	FV', MV'
77	Alifrangis et al. (1999)	Tanzania	126	Children	FMO	
78	Hellgren et al. (1994)	Tanzania	163	Children	FMO	
79	Thomson et al. (1994)	The Gambia	1465	Children	FMO	FM
80	Gbary et al. (1988)	Togo	707	Children	FMO	FM ⁴ , FO ⁴
81	Brown et al. (1970)	Uganda	2899	Children	FM	FM
82	Dorolle (1927)	China	643	Children	FMV	FV-, MV'
83	Perry (1911)	India	351	Children	FMV	
84	Phillips (1923)*	India	645	Children	FMV	FV ⁻
85	Aiver (1924)*	India	57	Children	FMV	
86	Roy (1926)*	India	309	Children	FV	
87	Christophers & Shortt (1921)	Iraq	76	Children	FMV	
88	Lalor (1913)*	Myanmar	207	Children	FMV	FV-
89	Strickland et al. (1988)	Pakistan	2891	Children	FV	
90	Maitland et al. (1996)	Vanuatu	292	Children	FV	
91	Maitland et al. (1996)	Vanuatu	211	Children	FV	FV-
92	Maitland et al. (1996)	Vanuatu	302	Children	FV	FV ⁻
93	Wilding et al. (1995)	Gabon	2192	Mixed	FMO	
94	Landgraf et al. (1994)	Ghana	1048	Mixed	FMO	FM-
95	Deloron et al. (1989)	Kenya	222	Mixed	FMO	
96	Deloron et al. (1989)	Kenya	245	Mixed	FMO	
97	Deloron et al. (1989)	Kenya	253	Mixed	FMO	FM
98	Deloron et al. (1989)	Kenya	225	Mixed	FMO	
99	May et al. (1999)	Nigeria	159	Mixed	FMO	
100	Trape et al. (1992)	Seneral	2465	Mixed	FMO	FM [*]
101	Trape et al. (1994)	Senegal	8539	Mixed	FMO	FM ¹ , FO ¹ , MO ¹
102	Molineaux et al. (1980)	Sudan	7026	Mixed	FMO	FM [*] , FO [*] , MO [*]
103	Molineaux et al. (1980)	Sudan	6526	Mixed	FMO	FM ⁺ , FO ⁺ , MO ⁺
104	Nhonoli et al. (1974)	Tanzania	360	Mixed	FM	
105	Nhonoli et al. (1974)	Tanzania	450	Mixed	FMV	FM
106	Wilson (1936)	Tanzania	3393	Mixed	FMV	FM*, FV*, MV*
107	Rosenberg & Maheswary (1982)	Bangladesh	1093	Mixed	FV	FV-

One century of data:

73 data sets on malaria (multiple species), covering a period ranging from 1911 to 1999

Perry, R. (1911). Malaria in the jeypore Hill Tract and adjoining coastland. Paludism, 5, 32.



Log-linear modelling to find interactions in malaria co-infection data

MY ROLE AS A FRENCH-SPEAKING CO-AUTHOR



SOMBOLINE DO HOMENO I	
Seance du 10 Janvier 1923	
Prophere de E. Columna, Président	
adjustment die Prevalent - ward is in die	
0000903003088	
a Instation, Ma PLANTING, P. Ross, G. Kardacrowsky La Théories house en	1.0
Algerts	10
F. GORTHY SILE. BRATHART - IN NOTION ROOT IN SUPPORT AN INCOME.	
Filter - Annual - Contribution & Claude dd in man an ferningi -	1
E. Barnaris et il. Passimero, Collans des parseires brokesanifamero alterno cher le	1
in the second se	12
har soon figure de mangersten deugenten. B. Van facturere - Einster de monter an Barnin (Altrase Orientale)	- FF
adaronata	
to. Factorizeta, or Action its large d'Explorizet our difference orygonassement, Calvary de	100
Regelite date dat beer divers	14 C
Ed. Dinicipy, 6. De l'accacician des differences plenataonese etter ches, à propos de pretonite Proprioane accasilitat.	14
Société de l'Ouest-Africain	
Source da vý Denadve spre	100
Practic carlled de l'assessible exemple	6
Nave tonin par le faction	
the second secon	3r
A. BUCLAT OF A LINE	14
H. LEUTS IN A. BALKET HERETAR protocor ballin das portaurs value humanian at on viru-	14
A. Laure at & Dirate Prophetone & Grade-is mainer that he berides its timber.	6
ALIGNETTAL - Propasses 2-9, Subers, discusses,	- R
mark, or pressent construction	_
er Etablissements POULENC FRI	ERES
CS LINDISSERTER PROPERTY IN COMPANY	
100 Boolevant Satat.Germain, PARIS	
PRODUTS CIENTIQUES	FERS
E B INF PROTOTO - COLO	DINTE
REALITYS IL COLO	ALCOLOGICAL D
all Designed and the second se	
Chiefavis - Librens	TITREES
Contraction of the second seco	

ATTAICON L

Introduction	Modelling	Results	Discussion
000000	000000	00000000	00000

COMPARISON WITH HOWARD ET AL. (2001)

Results using NiDP differ in $(12 + 4)/41 \approx 39\%$ of cases



Testing for interaction differs from independence tests

Introduction	Modelling	Results	Discussion
000000	000000	000000000	0000

SUMMARY

Non-interacting pathogen are not statistically independent

- the prevalence of co-infection is always greater than the product of the prevalences (positive correlation)
- statistically independent pathogens may well be interacting (not presented)
- confirms that statistical independence is far from equivalent to the absence of biological interaction between pathogens

Novel interaction tests based on simple epidemic models

- simple models challenge previous methods based on statistical independence
- simple models (with no explicit age structure) intrinsically correct for age as a confounding factor
- epidemic models make it unnecessary to keep track of age

Introduction	Modelling	Results	Discussion
000000	0000000	000000000	0000
			1

LIMITS AND PROSPECTS

Very simple two parameter SIS model (transmission, mortality)

- the positive correlation is due to host mortality (or ageing)
- results are valid for chronic (long-lasting) infections
- which represent a large fraction of co-infections

Clearance

- ► the model can accomodate specific pathogen clearance
- doubles the number of parameters, but may still be fitted
- qualitative results are unchanged in all cases tested

Virulence

- disease-induced mortality was assumed to be zero
- otherwise pathogens interact at the host population scale
- but virulence could be included in the model as well

Introduction	Modelling	Results	Discussion
0000000	0000000	000000000	00000

ECOLOGICAL PERSPECTIVES

Macro-parasites (e.g. worms)

- Macroparasite data show positive associations between parasites known to interact negatively (Fenton et al 2014)
- ► SIS models are inedaquate for macro-parasites
- Our approach could be extended to macroparasites

Meta-communities

- Pathogens are species which form meta-populations occupying discrete patches (hosts)
- In meta-community ecology, interactions between species are often inferred from co-occurrence data
- Most methods are based on statistical associations
- Our approach could be extended to metacommunities

Introduction	Modelling	Results	Discussion
000000	0000000	000000000	00000
			1

TAKE-HOME MESSAGE

To detect interactions, methods based on

statistical independence and random distributions

should be replaced with methods based on

model-based distributions assuming no biological interactions

as a null expectation.

Introduction	
0000000	

Modelling

Results

Discussion

00000

THANKS FOR LISTENING!



- ► The NIMBioS Working Group, particularly
 - Nik Cunniffe (Cambridge)
 - Linda Allen (Texas Tech)
- Hamelin, F. M., Allen, L. J., Bokil, V. A., Gross, L. J., Hilker, F. M., Jeger, M. J., ... & Cunniffe, N. J. (2019). Coinfections by noninteracting pathogens are not independent and require new tests of interaction. PLoS Biology, 17(12).