Quantification of Basic Epidemiological Characteristics: The Example of Human Polyomaviruses

Georg A Funk
University of Basel, Switzerland
Outline

- Learning goals
- Epidemiological characteristics
- Polyomaviruses & their pathogenicity profiles
- Source data, formulae & model selection
- Results (graphs) & summary of characteristics
- Some limitations & context
- Summary / take home message
- Q & A
Learning Goals

- Get an intuition for key epidemiol. characteristics.
- See & understand the pathway (or loop)
  - Initial idea
  - Data collection
  - Model construction & selection
  - Parameter estimation
  - Interpretation
  - Refinement(s)
- Be able to calculate from age-stratified (sero-prevalence) data the force of infection, \( R_0 \) and \( H \).
Epidemiology...

... the ecology of infectious disease(s)

Matthews & Woolhouse, 2005
'Force' of infection ($\lambda$)  
(per capita rate of acquisition of infection)  

Basic reproductive ratio ($R_0$)  
(secondary cases per 'index'-case)  

Herd immunity threshold ($H, p_c$)  
(proportion to be immunised to control infection)  

Further characteristics:  
- Average age of infection ($A$); $A \sim \lambda^{-1}$  
- Transmission parameter ($\beta$)
The 'Force' of Infection ($\lambda$)

- Per capita rate (='velocity') at which susceptible individuals acquire the infection.

$\lambda$ 'large' -> rapid

$\lambda$ 'low' -> slow

Note: $\lambda$ changes as the epidemiol. circumstances change; i.e. $\lambda$ is not necessarily constant over time!
Basic Reproductive Ratio ($R_0$)

- Expected number of secondary cases per primary case in a population where everybody except the 'index'-case is susceptible to infection.
  
  'Index'-case: $I(0) = \text{one individual}$  
  
  (see Luchsinger, p.81)

- Initial multiplication factor when considering events at the population level on a 'per generation' basis.

  Generation time = serial interval = time between catching an infection and passing it on to so. else.
Herd Immunity Threshold (H)

- Proportion of the population to be immunised to reduce $R_0$ below unity, i.e. to control infection. (See Smith, p.21; Luchsinger, p.82)

$R_0 > 1$, 'low' -> bar set low

$R_0 > 1$, 'high' -> bar set high

$R_0 < 1$ -> transient outbreak expected (see Luchsinger, p.81/82)
Basic Model & Assumptions

- 'Classical' Kermack-McKendrick (1927) SIR model

(see Luchsinger, p.80; here with $\beta$ instead of $\lambda$ to avoid confusion)

\[
\begin{align*}
\text{Prop. Susceptibles} & \quad \frac{dS}{dt} = b - \delta S - \beta IS \\
\text{Prop. Infecteds} & \quad \frac{dI}{dt} = \beta IS - \mu I - (\delta + \nu)I \\
\text{Prop. Removeds} & \quad \frac{dR}{dt} = \mu I - \delta R
\end{align*}
\]

- Underlying assumptions:
  - population in demographic equilibrium (i.e. $b=\delta$ and $\delta \approx 0$)
  - random mixing of infecteds with susceptibles
  - infected individuals become immediately infectious
  - negligible pathogen induced host mortality (i.e. $\nu=0$)
  - short infectious period compared with lifespan (i.e. $\mu \gg \delta$)
  - removed ones cannot become infected/infectious any more
Where Are Our Quantities?

- **Simplified epidemic SIR model...**

  \[
  \text{Prop. Susceptibles } \quad \frac{dS}{dt} = -\beta IS \\
  \text{Prop. Infecteds } \quad \frac{dI}{dt} = \beta IS - \mu I = I \mu (R_0 \cdot S - 1) \\
  \text{Prop. Rremoveds } \quad \frac{dR}{dt} = \mu I
  \]

  'Force' of infection: \( \lambda = \beta I \)

  \( 1/\lambda \approx \text{mean time an individual spends in the susceptible class} \)

  Basic reprod. ratio: \( R_0 = \frac{\beta}{\mu} \)

  Herd immun. thresh.: \( H = 1 - \frac{1}{R_0} = 1 - (\mu/\beta), \text{ for } R_0 > 1 \)
Derivation of $H$

- **Goal:** Proportion of infecteds $I$ shall shrink, i.e.

\[
\text{Prop. Infecteds} \quad \frac{dI}{dt} = \beta IS - \mu I \\
= I(\beta S - \mu) \\
= I(\beta(\mu/\mu)S - \mu) \\
= I\mu((\beta/\mu)S - 1) \\
= I\mu(R_0 \cdot S - 1) < 0
\]

$\Rightarrow S_T < 1/R_0$ for $I>0$, $\mu>0$ and $R_0>1$

Herd immun. thresh.: $H = 1 - 1/R_0 = 1 - (\mu/\beta) \approx 1 - S_T$

Suppose $R_0=20$ $\Rightarrow S_T=5\%$ and thus $H=1-0.05=0.95$
Polyomaviruses

- Small (~50 nm), non-enveloped DNA viruses (can infect a variety of vertebrates)

- 8 'human related' polyomaviruses known (5 were discovered in the past 4 years!)
Why Interesting to Study?

- **Ubiquitous virus(es)**
  - No vaccines -> host-pathogen system in endemic equilibrium; i.e. \( I(t) > 0 \) for extended periods of time

- **Disease(s) only when immunity is compromised**
  - HIV / AIDS
  - Transplantation

- **Not much is known...**
  ... in particular reg. epidemiol. characteristics

- **Improve clinico-epidemiological knowledge**
  (help to device approaches for better protecting patients at risk)
Pathogenicity Profiles

- **BKV** -> Nephropathy
- **JCV** -> PML (demyelating brain cells)
- **MCV** -> Merkel Cell Carcinoma
- Other polyomaviruses (SV40, KIV, WUV, TSV) respiratory illnesses? transforming capacity
Types of Data

- **Real-time, ongoing**
  - Raw data
  - Case notification data
    - Number of cases vs. age classes
  - Sero-epidemiological data
    - Proportion seropositive vs. age classes

- **Retrospective**
  - Limitations:
    - Representative sample
    - Fine age stratification
    - No underreporting

- **Limitations**:
  - Representative sample
  - Fine age stratification
  - Sensitive and specific assay
Interrelation of Characteristics

**Raw data**

**Sero-epidemiological data**

Proportion seropositive vs Age classes

**Key quantity**

Force of infection, $\lambda$
- per capita rate of acquisition of infection

**Derivatives**

- **Basic reproductive ratio**, $R_0$
  - secondary cases per (index-) case if $\lambda$ is supposed to be fix over life

- **Herd immunity threshold**, $H$
  - proportion of the population to be immunised to control infection

- **Invasion criterion & cross-immunity**
  - threshold an invading pathogen has to overcome in face of resident pathogen

**Limitations:**
- fine age stratification
- representative sample
- sensitive and specific assay

**Assumption:**
- infection is at its endemic equilibrium

**Life expectancy, $L$**

**100% efficacy of antiviral measures**

**Degree of cross-immunity, $\varepsilon$**
Age-stratified sero-epidemiological surveys
- ideally longitudinal ('follow' an individual; inexistent), here cross-sectional (all age-strata sampled simultaneously)
- sensitive and specific assays (HIA, ELISA)
- at regular time intervals
- a large unbiased sample of the population

PubMed® search: 11 studies providing 22 data sets (7 BKV, 6 JCV, 2 SV40, 1 LPV, 2 KIV, 2 WUV, 2 MCV)
Prevalence in England of Antibody to Human Polyomavirus (B.K.)

SYLVIA D. GARDNER

### Table 11—Haemagglutination-inhibiting Antibody to B.K. Virus in Human Sera

<table>
<thead>
<tr>
<th>Age Group</th>
<th>No. of Sera Tested</th>
<th>No. of Sera with Antibody</th>
<th>Percentage with Antibody</th>
<th>Range of Titers</th>
<th>Percentage Sera with Titres 2,560 or Greater</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-3 months</td>
<td>36</td>
<td>24</td>
<td>67</td>
<td>40-5,120</td>
<td>25</td>
</tr>
<tr>
<td>4-11 months</td>
<td>54</td>
<td>17</td>
<td>37</td>
<td>40-5,120</td>
<td>11</td>
</tr>
<tr>
<td>Primary children (4-6 years)</td>
<td>48</td>
<td>35</td>
<td>73</td>
<td>40-5,120</td>
<td>31</td>
</tr>
<tr>
<td>6-10 years</td>
<td>52</td>
<td>43</td>
<td>83</td>
<td>40-5,120</td>
<td>35</td>
</tr>
<tr>
<td>11-17 years</td>
<td>40</td>
<td>33</td>
<td>83</td>
<td>40-20,480</td>
<td>36</td>
</tr>
<tr>
<td>18-25 years</td>
<td>34</td>
<td>27</td>
<td>79</td>
<td>40-2,560</td>
<td>4</td>
</tr>
<tr>
<td>26-50 years</td>
<td>46</td>
<td>33</td>
<td>72</td>
<td>40-10,240</td>
<td>20</td>
</tr>
<tr>
<td>&gt;50 years</td>
<td>53</td>
<td>60</td>
<td>75</td>
<td>40-20,480</td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>409</td>
<td>254</td>
<td>62</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Source data

<table>
<thead>
<tr>
<th>Polymavirus type, reference, region &amp; assay</th>
<th>Age class</th>
<th>% sero-preval</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>BKV Gardner 1973</td>
<td>0-1</td>
<td>4</td>
<td>54</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>1-5</td>
<td>27</td>
<td>46</td>
</tr>
<tr>
<td>HIA (≥40)</td>
<td>6-10</td>
<td>83</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>11-17</td>
<td>83</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>18-25</td>
<td>79</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>26-50</td>
<td>72</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>&gt;50</td>
<td>75</td>
<td>53</td>
</tr>
</tbody>
</table>

### Estimate

<table>
<thead>
<tr>
<th></th>
<th>FoL (1/y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BKV Gardner 1973</td>
<td>0.004</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>0.262</td>
</tr>
<tr>
<td>HIA (≥40)</td>
<td>0</td>
</tr>
<tr>
<td>18-25 years</td>
<td>-0.026</td>
</tr>
<tr>
<td>&gt;50 years</td>
<td>-0.012</td>
</tr>
<tr>
<td>Total</td>
<td>NA</td>
</tr>
</tbody>
</table>

![Graph of Fraction of susceptibles vs Age (years)](image)

**Q&A**
'Force' of infection ($\lambda$) in interval $i$, $i+1$

$$\lambda_i = -\ln[(1 - p_{i+1}) / (1 - p_i)]$$

$p_i$ -> proportion of those who have experienced infection at age $i$

(Anderson & May, 1983, Appendix 2, Eqs. 2.2 & 2.9; and 1985, Eq. 58)

Basic reproductive ratio ($R_0$)

$$R_0 = (\lambda_c \cdot L) / (1 - \exp(-\lambda_c \cdot L))$$

$L$ -> life expectancy = 80 years, type I mortality;
$\lambda_c$ -> 'childhood' force of infection

Herd immunity threshold ($H$, $p_c$)

$$H = 1 - 1 / R_0$$
Competing candidate models (4-5) ranked by Akaike's information criterion

\[ AIC_c = 2 \cdot p + n \cdot [\ln(2 \cdot \pi \cdot \text{RSS} / n) + 1] + 2 \cdot p \cdot (p + 1) / (n - p - 1) \]

- **penalty for model complexity**
- **goodness of fit**
- **correction for small sample size**

RSS = residual sum of squares; p = number of parameters; n = sample size

**Models with lowest** \( AIC_c \) **score are shown**
(fitted by nonlinear least squares)
Polyomaviruses BKV & JCV

A. BKV, UK 1973
\[ \lambda_c = 0.178 \]
\[ \lambda_A = -0.0050 \]

B. BKV, USA 1973
\[ \lambda_c = 0.304 \]
\[ \lambda_A = -0.0053 \]

C. BKV, Italy 1974
\[ \lambda_c = 0.348 \]
\[ \lambda_A = -0.0078 \]

D. BKV, Japan 1982
\[ \lambda_c = 0.185 \]
\[ \lambda_A = -0.0035 \]

E. BKV, UK 2003
\[ \lambda_c = 0.509 \]
\[ \lambda_A = -0.0050 \]

F. BKV, USA 2009
\[ \lambda_c = 0.138 \]
\[ \lambda_A = -0.0038 \]

G. JCV, USA 1973
\[ \lambda_c = 0.055 \]
\[ \lambda_A = -0.0042 \]

H. JCV, Japan 1982
\[ \lambda_c = 0.104 \]
\[ \lambda_A = -0.0041 \]

I. JCV, UK 1995
\[ \lambda_c = 0.042 \]
\[ \lambda_A = -0.0061 \]

J. JCV, Taiwan 2002
\[ \lambda_c = 0.272 \]
\[ c = 0.276 \]
\[ f = 0.72 \]

K. JCV, UK 2003
\[ \lambda_c = 0.0248 \]
\[ \lambda_A = -0.0053 \]

L. JCV, USA 2009
\[ \lambda_c = 0.791 \]
\[ \lambda_A = 0.0044 \]
\[ c = 0.833 \]
\[ f = 0.17 \]
SV40, KIV, WUV & MCV

Introduction

Methods

Discussion

Summary

Results
Characteristics: Example

Study type: BKV
Author: Gardner, 1973

Model prob.: 0.93

$\lambda_c$ (1/y): 0.178 (95% CI: 0.094-0.263)

$R_0$: 14.3 (95% CI: 7.5-21)

$H_{L=80}$ (in %): 93 (95% CI: 87-95)

Sensitivity
$H_{L=70}$ (in %): 92

delta H: -1.08%
# Summary of Characteristics

<table>
<thead>
<tr>
<th>Virus Type</th>
<th>Study (Author, Year)</th>
<th>Model prob.</th>
<th>F0 (1</th>
<th>y)</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
<th>R0</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
<th>H in % (L=80)</th>
<th>Lower 95% CI</th>
<th>Upper 95% CI</th>
<th>H in % (L=70)</th>
<th>% change H_{L=80} vs H_{L=70}</th>
</tr>
</thead>
<tbody>
<tr>
<td>BKV</td>
<td>Gardner 1973</td>
<td>0.93</td>
<td>0.173</td>
<td>0.0094</td>
<td>0.203</td>
<td>14.3</td>
<td>7.5</td>
<td>21.0</td>
<td>93</td>
<td>87</td>
<td>95</td>
<td>95</td>
<td>-1.06</td>
<td></td>
</tr>
<tr>
<td>SIV</td>
<td>Padgett 1973</td>
<td>0.50</td>
<td>0.055</td>
<td>0.034</td>
<td>0.076</td>
<td>4.4</td>
<td>2.9</td>
<td>6.1</td>
<td>77</td>
<td>66</td>
<td>84</td>
<td>74</td>
<td>-3.65</td>
<td></td>
</tr>
<tr>
<td>SV40</td>
<td>Knowles 2003</td>
<td>0.26</td>
<td>0.003</td>
<td>-0.010</td>
<td>0.015</td>
<td>1.1</td>
<td>0.7</td>
<td>1.7</td>
<td>10</td>
<td>0</td>
<td>42</td>
<td>0</td>
<td>-11.75</td>
<td></td>
</tr>
<tr>
<td>LPV</td>
<td>Knowles 2003</td>
<td>0.77</td>
<td>0.010</td>
<td>0.004</td>
<td>0.017</td>
<td>1.5</td>
<td>1.2</td>
<td>1.9</td>
<td>32</td>
<td>13</td>
<td>45</td>
<td>23</td>
<td>-9.64</td>
<td></td>
</tr>
<tr>
<td>KIV</td>
<td>Kean 2009</td>
<td>0.39</td>
<td>0.64</td>
<td>0.297</td>
<td>0.995</td>
<td>51.3</td>
<td>23.8</td>
<td>78.9</td>
<td>98</td>
<td>96</td>
<td>96</td>
<td>90</td>
<td>-0.28</td>
<td></td>
</tr>
<tr>
<td>WUV</td>
<td>Kean 2009</td>
<td>0.92</td>
<td>0.740</td>
<td>0.342</td>
<td>1.155</td>
<td>59.9</td>
<td>27.3</td>
<td>92.4</td>
<td>98</td>
<td>96</td>
<td>96</td>
<td>99</td>
<td>-0.44</td>
<td></td>
</tr>
<tr>
<td>MCV-339</td>
<td>Kean 2009</td>
<td>0.95</td>
<td>0.154</td>
<td>0.086</td>
<td>0.242</td>
<td>12.3</td>
<td>5.3</td>
<td>19.4</td>
<td>92</td>
<td>81</td>
<td>95</td>
<td>91</td>
<td>-1.26</td>
<td></td>
</tr>
<tr>
<td>MCV-350</td>
<td>Kean 2009</td>
<td>0.94</td>
<td>0.283</td>
<td>0.130</td>
<td>0.445</td>
<td>23.0</td>
<td>10.4</td>
<td>35.6</td>
<td>96</td>
<td>90</td>
<td>97</td>
<td>95</td>
<td>-0.65</td>
<td></td>
</tr>
</tbody>
</table>

**Legend:**
- Model prob.: probability (Akaike weight) of the respective best fitting model.
- F0: force of infection (parameter lamda(c)) reported in each panel of Figure 2.
- R0: estimated basic reproductive ratio obtained by assuming a life expectancy of L=80 years.
- H: herd immunity threshold H for assumed life expectancies L of 80 or 70 years, respectively.
- % change: reduction of H in % if the life expectancy was assumed to be 70 years instead of 90 years.
Cross-Immunity?

- Expectation if cross-immunity would exist between BKV or JCV and SV40
  Invasion criterion: $R_0^{\text{INV}} \left[ 1 - \varepsilon \left( 1 - \frac{1}{R_0^{\text{EST}}} \right) \right] > 1$

- SV40 does not surpass the invasion threshold!
Some Limitations

- Quality of (source) data!
  - Actuality (2/3 of data after year 2000)
  - Accuracy (endemic infection, lasting sero-conversion)
  - Origin (only peer-reviewed studies, critically appraised)

- Sero-reversion of ~5% per decade during adulthood
  (dI/dt<0 due to low S?; any balance between sero-conversion and sero-reversion? ==> longitudinal studies)

- Does cross-immunity hinder SV40 from invasion?
  (neutralizing or binding antibodies? cellular immunity is totally missing)
Broader Context

- Lack of *extended* protection by maternal antibodies
- Rapid acquisition of polyomaviruses at an age when toddlers have increasing numbers of *social contacts*
  - median age: 5-7 years -> familiy (-), compagnons (+)
  - median FoI: ~0.3/y -> compares well with *measles*
- Ten times faster than acquisition of *cytomegalovirus*
- Acquisition during adulthood ~200-fold slower as during childhood (reactivations?)
- Protection of immunocompromised patients must be both *highly efficient* and *well targeted* (infants may be 'vectors')
Summary

- First quantitative study describing polyomavirus circulation, vital to inform virus control strategies.
- Complements recent reports proposing the development of candidate vaccines, e.g. against MCV.
- Conform acquisition profiles of BKV across space (Asia, Europe, USA) and time (1973-2009).
- Sero-conversion during childhood driven by a median force of infection of ~0.3/y ($R_0 \approx 24$).
- Herd immunity thresholds of BKV, KIV, WUV or MCV are comparable with those of measles, i.e. high!
- Antibodies against most polyomaviruses are on the wane any time (albeit slowly, sero-reversion rate ~0.005/y).
Thank You!

"I think it's a Pig virus..."

Questions?
Endemic SIR Model

- Considering demographic turnover, etc., leads to

  Prop. Susceptibles: \[ \frac{dS}{dt} = b - \delta S - \beta IS \]
  Prop. Infecteds: \[ \frac{dI}{dt} = \beta IS - \delta I - \mu I \]
  Prop. Removeds: \[ \frac{dR}{dt} = \mu I - \delta R \]

  'Force' of infection: \[ \lambda = \beta I^*, \text{ with } I^* = \left(\frac{\delta}{\beta}\right)(R_0 - 1) \]

  \( \frac{1}{\lambda} \approx \text{mean time an individual spends in the susceptible class} \)
  \( \approx \text{average age of infection, denoted } A, \text{ with } A \approx \frac{1}{\delta(R_0 - 1)} \)

  Basic reprod. ratio: \[ R_0 = S_0 \cdot \frac{\beta}{\delta + \mu}, \text{ with } S_0 = \frac{b}{\delta} \]

  Herd immun. thresh.: \[ H = 1 - \frac{1}{R_0} \]
Derivation of $\lambda_i$

(see Anderson & May, 1983, Appendix 2, Eqs. 2.2 & 2.9; and 1985, Eq. 58)

Prop. Susceptibles at age $i$
- constant $\lambda$
  $$S(i) = \exp[-\lambda i]$$  
  Eq. 2.1
- age dependent $\lambda$
  $$S(i) = \exp[-\int_{0}^{i} \lambda(x) dx]$$  
  Eq. 2.2

$\lambda_i$ in interval $i+\Delta i$, with $\Delta i$ relatively small
  $$i + \frac{1}{2} \Delta i = \frac{-\ln[S(i+\Delta i)/S(i)]}{\Delta i}$$  
  Eq. 2.9

by means of Eq. 2.2 where $i+\frac{1}{2} \Delta i$ is given as by Eq. 2.9, and observing that $S(i+\Delta i) = (1-p_{i+\Delta i})$ we obtain

  $$\lambda_i = -\ln[(1-p_{i+\Delta i})/(1-p_i)]$$  
  Eq. 58
Why Modelling?

- Modelling allows to...
  - identify gaps in our knowledge
  - take a fresh look at old data or phenomena
  - generate new hypotheses
  - explore them in silico before performing costly & time consuming experiments
  - improve clinical decisions
  - ...

and it makes fun!