Bayesian modelling of *Escherichia coli* O157:H7 dose response incorporating age as a covariable

M. L. Delignette-Muller - CNRS UMR 5558
S. Jaloustre - CNRS UMR 5558 - ANSES
H. Bergis - ANSES

NCSC 2010, 09/28/2010

Escherichia coli O157 :H7

- a Shiga toxin-producing *Escherichia coli*
- linked to a large number of human infections due to consumption of contaminated food products (especially ground beef)
- potential severe clinical manifestations : hemolytic uremic syndrome (HUS)
- most common cause of acute renal failure in children, preferentially among children under 10 years

Dose-response characterization

From which data ?

- experimental data
  - on animals ? Difficult to extrapolate
  - on human volunteers ? Low susceptibility and ethically not reasonable
  - on children (higher susceptibility) ? Ethically not reasonable
- outbreak data and surveillance data
  disparate data with much uncertainty
  ⇒ appropriate methods need to be developed
  bayesian methods may be of interesting in such a case

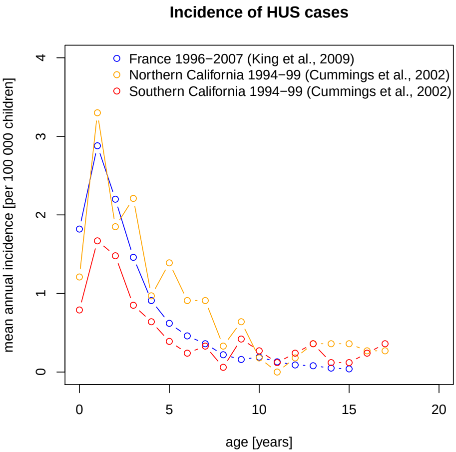
Recent works

- *Teunis et al., 2008*
  A meta-analysis on 8 outbreaks :
  hierarchical model linking the risk of illness (diarrhea cases) to the ingested dose
  (estimation of parameters by bayesian inference)
- *Delignette-Muller and Cornu, 2008*
  An analysis of one well-documented french outbreak : two models linking the risk of severe complication (HUS cases) to the ingested dose for two age classes ("$<$ 5 years" and "5-10 years")
  (estimation of parameters by bayesian inference).
  Greater susceptibility in the class "$<$ 5 years".

Objective

To build a unique *E. coli* O157 :H7 dose-response model for children under 16 years
incorporating age as a covariable,
from outbreak data used in our previous work
and surveillance data.

Surveillance data



Decrease of susceptibility with age with a lower value for babies certainly due to a lower exposure before food diversification.

Outbreak data

Largest community-wide outbreak of *E. coli* O157 :H7 in France, associated with consumption of contaminated frozen beef burgers in households.

- Estimated total number of contaminated consumed patties : $N = 2155$
  (information from the distributor)
- Estimation of ingested dose : depends on the initial contamination level ($C[cfu.g^{-1}]$) and of the consumption preference (CP) and the serving size (S)

A challenge : to incorporate age as a covariable in the estimation of all parameters of the exposure model (N_{age}, CP_{age}, S_{age}).

Exposure model

Ingested dose D for one consumer

$$D_{age} \sim Poisson(C \times S_{age} \times 10^{-R_{age}})$$

R_{age} depending on CP_{age}

- C : initial mean contamination level on frozen ground beef $(cfu.g^{-1})$
- S_{age} : serving size (g)
- R_{age} : number of decimal reduction due to cooking
- CP_{age} : consumption preference
  (raw, rare, medium or well-done)

Initial mean contamination level C

Estimated from microbial detection and counts performed on 22 frozen patties sampled from the contaminated batch

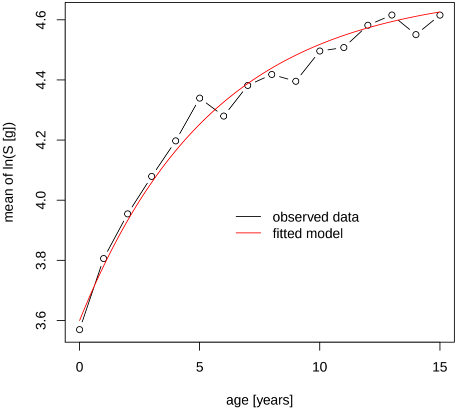
Bayesian estimation :
$5.8 \; cfu.g^{-1}$
95% credibility interval :
$[3.2; 9.4]$

Modelling of the initial mean contamination level

$\mathrm{C} = 5.8 \; cfu.g^{-1}$

Serving size S_{age} (1)

Consumption data (890 ground beef intakes by children under 16 years old) : strong effect of age on mean of $ln(S)$ distribution and no significant effect on standard deviation



Serving size S_{age} (2)

Modelling of the serving size distribution by a lognormal distribution,
with the mean described as a function of age.
Estimation of parameters from consumption data, using bayesian inference with non-informative priors.

Modelling of the serving size

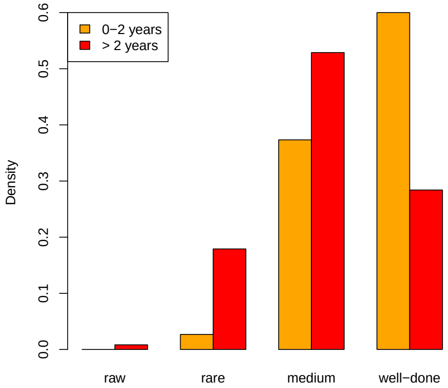
$ln(S_{age}) \sim N(\mu_{age}, \sigma)$
with $\mu_{age} = \mu_0 + \Delta\mu \times (1 - e^{-\beta \times age})$
with $\mu_0 = 3.6$, $\Delta\mu = 1.1$, $\beta = 0.18$ and $\sigma = 0.40$.

Consumption preference CP_{age} (1)

Estimation of the *CP* distribution from a consumer survey (on 589 children under 16 years old) based on 4 photos of ground beef patties illustrating different consumption preferences.

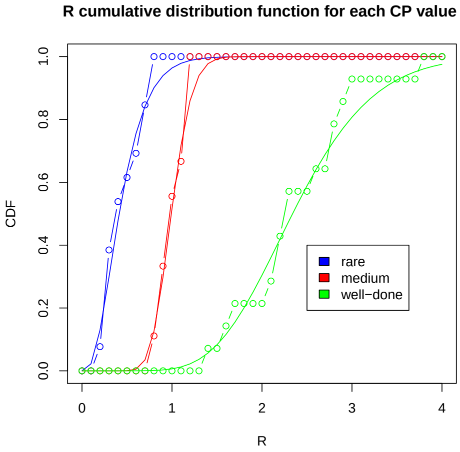
Consumption preference CP_{age} (2)

Two distributions : one for children between 0 and 2 years old and one for children over 2 years old.



Number of decimal reduction due to cooking $R(CP_{age})$

Characterization of R distribution for each cooking preference from experimental data by gamma distributions (36 ground beef patties experimentally contaminated and cooked)



Number of patties consumed in each age group : N_{age}

Proportions p_{age} of total ground beef servings consumed in each age group (length 1 year) estimated from **mean numbers of patties consumed per year** and **demographic data**
From consumer surveys (890 ground beef intakes by children under 16 years old and 1288 on adults and children over 16 years),
Mean ground beef consumption :

- under one year old : 13.6 patties per year,
- from one year old to 15 years old : 35.8 patties per year.
- over 15 years old : 23.3 patties per year.

Estimated numbers of intakes per age class during the outbreak

$N_{age} = p_{age} \times N$ with $N = 2155$
$N_{age} = [7; 38; 38; 38; 39; 39; 37; 37; 36; 37; 36; 36; 36; 37; 38; 39]$
for $age = 0$ to 15 years old
and $N_{>15 years} = 1587$

Dose response model

Simple choice of a single-hit model,
characterized by the parameter r (probability of HUS from a single ingested cell), here defined as an exponential function of age (from the trend observed on surveillance data)

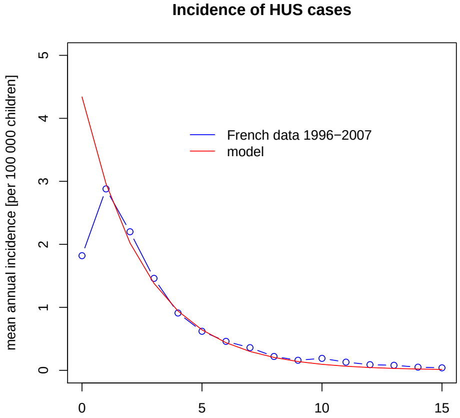
Modelling of the HUS outcome (HUS or not) for chidren

$HUS_{age} \sim Bernouilli(HUSrisk_{age})$
$HUSrisk_{age} = 1 - (1 - r_{age})^{D_{age}}$
D_{age} defined in exposure model and
$r_{age} = r_0 \times e^{-k \times age}$ for age from 0 to 15 years old

Two parameters to estimate : r_0 and k.

Preliminar estimation of k from surveillance data

Estimation of the parameter of the exponential trend from french surveillance data excluding babies (< 1 year), and assuming a common exposition on other age groups : $k = 0.38$ (95% credibility interval $[0.33; 0.43]$)



Estimation of k and r_0 from outbreak data

Bayesian estimation from outbreak data using the previously developped exposure model to describe the dose D as a function of age, with

- informative prior for k : the posterior distribution from the previous bayesian estimation
- non-informative prior for r_0

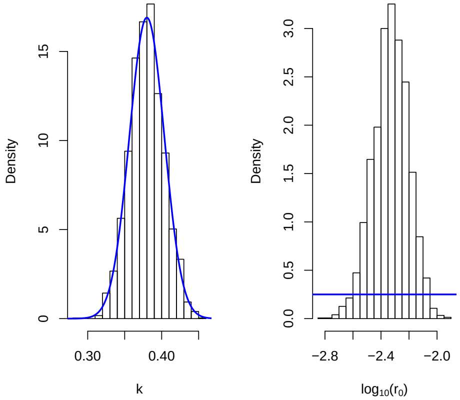
Estimated parameters

$k = 0.38$ (95% credibility interval $[0.33; 0.43]$)
$log_{10}(r_0) = -2.33$ (95% credibility interval $[-2.58; -2.09]$)

Posterior distributions of parameters

Comparison of posterior distributions (histograms) to prior ones

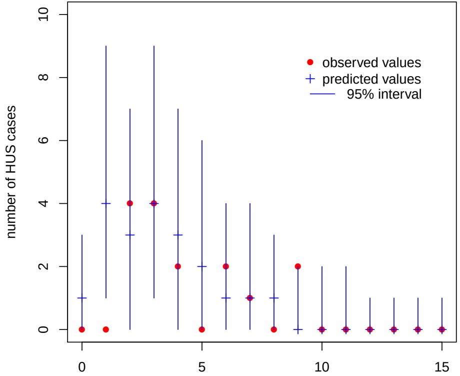
No more information on k, but its ponctual estimation is robust to prior : unchanged using a uniform non-informative prior for k



Goodness-of-fit of the model

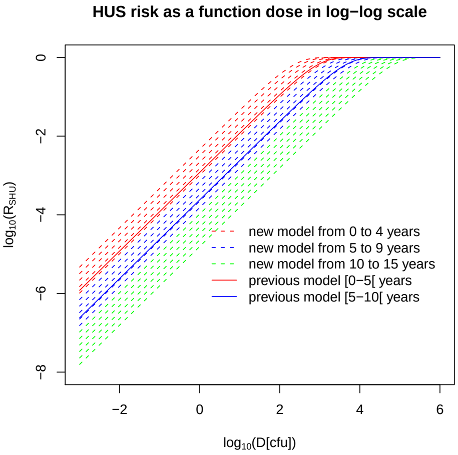
Comparison between observed numbers of HUS cases and predicted ones with credibility intervals :

One observed value outside the 95% credibility interval (the consumption of ground beef may be overestimated in our model for children of one year old).



Comparison with previously proposed models for HUS risk

Agreement between the HUS new model and previous ones. The new model describes a larger variability due to age and a greater risk for younger children in each age group.



Conclusion

- The proposed model was built on only one dataset and with some oversimplified hypotheses and should thus be validated under other data (but HUS data are scarce).
- It is the first attempt to incorporate age as a covariable in a dose-response model for microbial risk assessment in food.
- Bayesian inference was performed using JAGS (developed by Martyn Plummer).
  The great flexibility offered by bayesian tools such as JAGS and WinBUGS is very interesting in dose-response modelling using outbreak data, especially to incorporate non trivial models for the dose.