Bayesian Data Analysis Workshop

Orion Pharma, Espoo, 6-7.5.2013

Campylobacter Risk Assessment: Bayesian Inference, Predictive Microbiology and Uncertainty

Jukka Ranta
Risk Assessment Research Unit
Finnish Food Safety Authority Evira

Much of this is based on results from NMDD project led by Maarten Nauta (DTU/Denmark), and current work on a related Finnish Campylobacter project.

Evira

QMRA:
Quantitative Microbial Risk Assessment

”Microbiological risk assessment aims at determining the likelihood and severity of biological agents harmful to the health of the consumer, such as bacteria, viruses or protozoa that spread through food. The risk factors affecting the spread and prevalence of the biological agents and the importance of the transmission routes on the magnitude of the risk, and the effect of different means of risk management are often assessed at the same time. The risk assessment comprises the formation of risk along the whole food production chain from the production of raw materials all the way to the exposure of the consumer.”

Evira

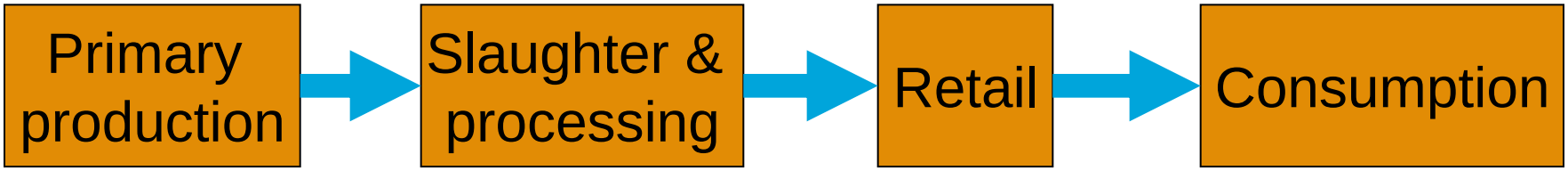
QMRA:
Quantitative Microbial Risk Assessment

”Often however, the assessment is targeted at one specific part of the chain, for example a part that is important for the decision-making.”

Evira

Q(M)RA: modular structures

- Describes causal structures of a production chain in a modular fashion:



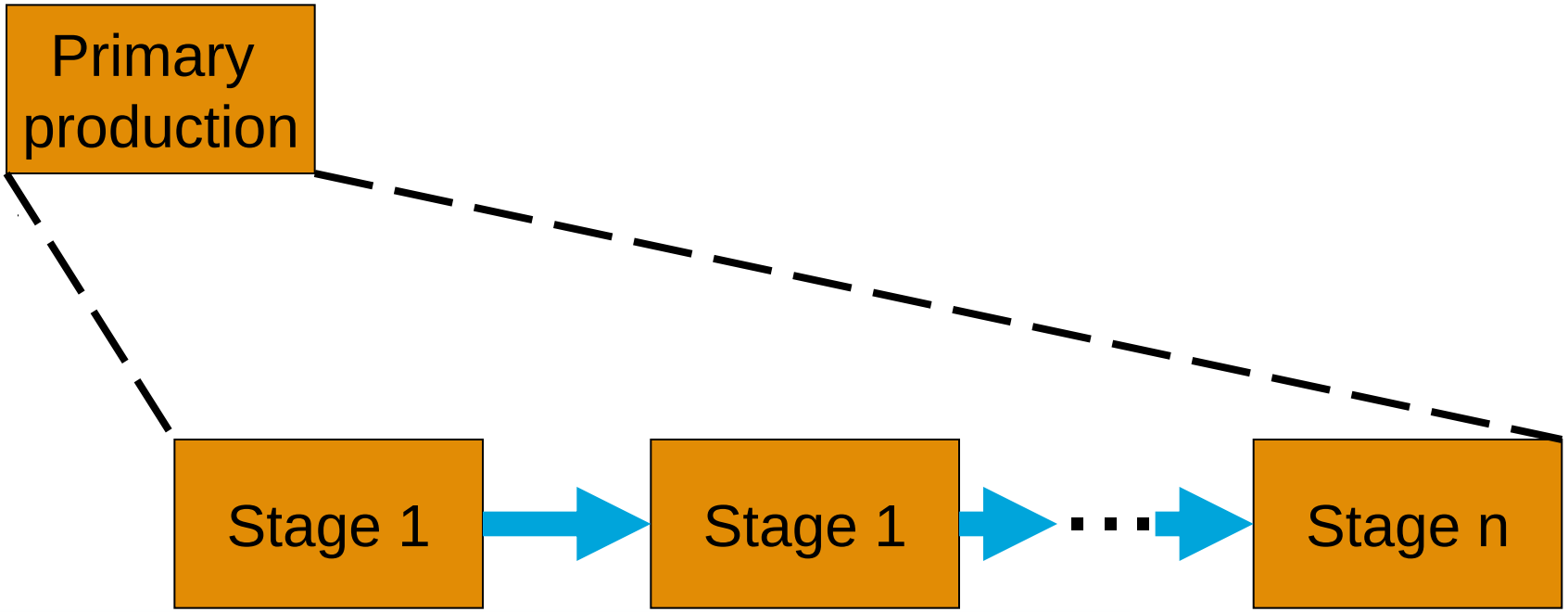
- Prevalence & concentration at given stages?
- Effect of intervention at that stage?

- Risk: the predicted number of human cases?

Evira

QRA: nested modular structure

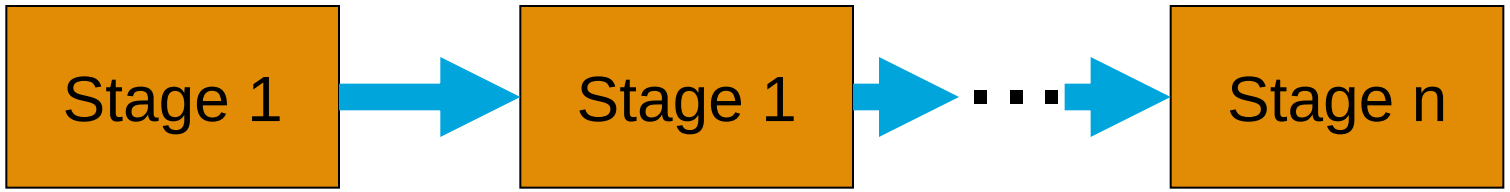
- Each module can further consist of submodules:



Evira

QRA: mechanistic approach

- QRA & mechanistic model approach: description of detailed processes, -- with only partial empirical support due to limited or even nonexisting data.
  - Each variable depends on the previous variable(s) either as a deterministic function, or stochastically.



Evira

Modeling a structured process

- Process structures are often quite known.
  (uncertainty of ’how it works’?)
- Biology of the bacteria is also quite known.
  (uncertainty of new strains?)

- Predictive microbiology: if we know the temperature, and the storage time, and the initial bacteria counts, etc, it seems possible to predict from one stage to the next.

Evira

Knowing structures vs knowing values

- But we only know a few of these conditions.
  - Initial prevalence and concentration can be very uncertain.
  - Bacteria transfer rates between stages not measured.
  - Measurements may not be comparable.
  - Predicting under actual conditions vs experimental conditions.
  - Dose-response for the general population or interesting subgroups not well known.
  - Exact consumer behaviour not known.

Evira

But what needs to be known?

- Intervention may only concern a specific stage, previous stages not of interest?
- Objective: a required safety level at a specific stage (performance objective), not the end point of consumption?
- Ready-to-eat foods don’t need cooking models?
- For a sub-process, inner details not of interest, only its inputs and outputs?
- Relative risk differences, not absolute risks?

- All depends on the questions, and the application.

Evira

Systems with complexity and uncertainty

- José Bernardo: *"Actual scientific research often requires the use of models that are far too complex for conventional statistical methods".*

- No matter how much reduced, the problem still contains 'too many unknowns' so that "n << p".

$(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7 , \ldots , \theta_p)$

Data

Analyze the data or the problem?

Evira

How to deal with the known unknowns?

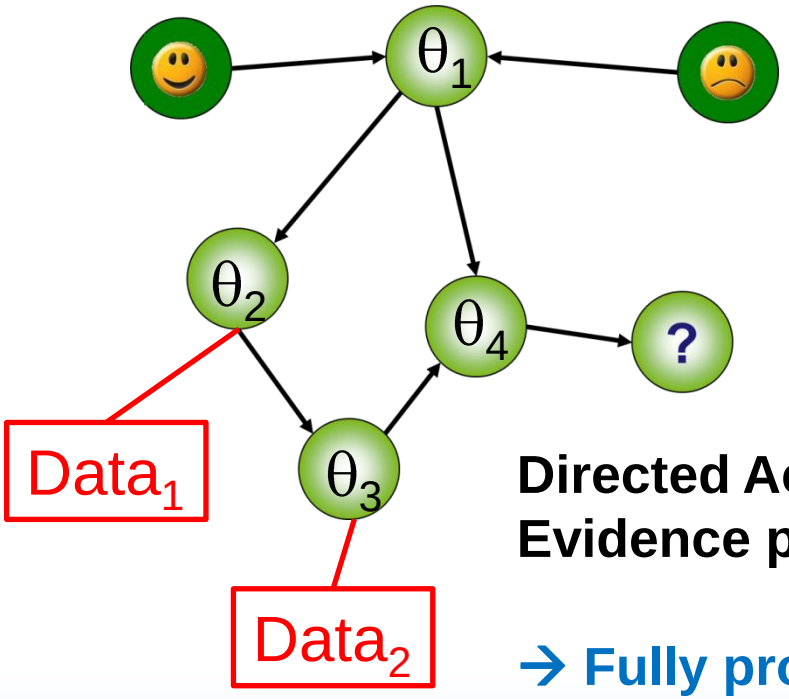
- If there are separate (external) data for a parameter, draw an estimate from that?
- If no data, use expert elicitation?
- If elicitation not feasible, make bold assumptions, and check sensitivity for results?
- Investigate which parameters are most influencial for results?

- *Yet, none of this makes a coherent probabilistic assessment of joint uncertainty, conditional on the stated set of evidence*

Evira

Bayesian evidence synthesis

- Exploit structural information with data & indirect evidence and expert opinions (or literature):



Directed Acyclic Graph (DAG) of the model.
Evidence propagates throughout the model.

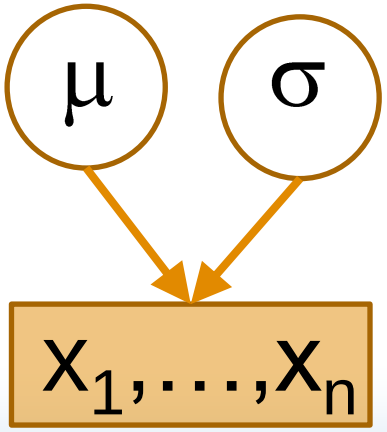
→ Fully probabilistic risk model.

Evira

Directed Acyclic Graph (DAG)

- Perhaps the most useful concept to visualize Bayesian models

- A graph of conditional dependencies
  - e.g. for a normal model for $x_1,\ldots,x_n$
  - goal: posterior distribution of unknowns



Evira

Campylobacter and broilers

- Effect of *microbial criteria* set at slaughter stage on consumer risk?
- Criteria: batch level testing of broiler carcasses: "n=5,c=1,m=1000"

- Issues to be dealt with:
  - between and within batch variability of concentrations.
  - prevalence of contaminated batches.
  - prevalence within batch.
  - resulting prevalence and concentration at consumption? (dose-response).

Evira

Campylobacter and broilers

- Partially informative data about prevalence and concentrations *on carcasses*
  - Consumer risk to be predicted from this.
  - Yet, assumptions needed to bridge this with the resulting dose.

→Hierarchical model !

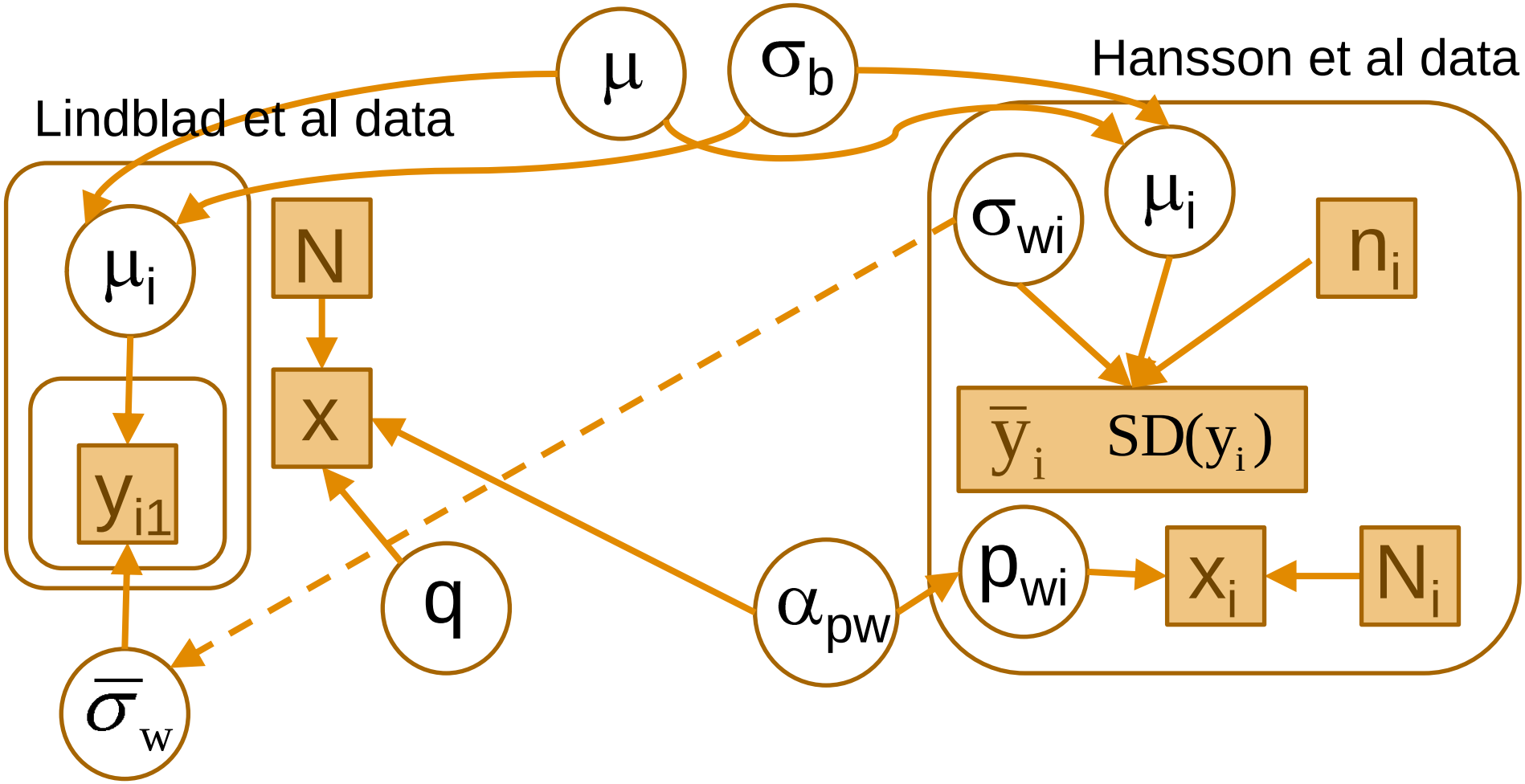
  - For evidence synthesis, this can be made as a hierarchical Bayesian model.

Evira

Beginning with two data sets

- One of the data sets has measurements of concentrations from random batches, but only one per batch.
- One of the data sets has mean and SD of concentrations from a sample, but only from selected positive batches.
- A hierarchical model connects both data sets, this provides the evidence synthesis (for this part).

Evira


Lindblad et al data
Hansson et al data
μ
σb
μi
yi1
N
x
q
σ̄w
σwi
μi
ni
ȳi SD(yi)
αpw
pwi
xi
Ni

Evira

Predicting consumer dose

- Generate prediction for binary status “1/0” & concentration (for contaminated carcass)
  - From posterior predictive distribution: p(y*|data)=

$$\iint\iint \mathrm{N}(\mathrm{y}^*|\,\mu_\mathrm{i},\sigma_\mathrm{w}^2)\mathrm{N}(\mu_\mathrm{i}\,|\,\mu,\sigma_b^2)\mathrm{p}(\mu,\sigma_w^2,\sigma_b^2\,|\,\mathrm{data})\mathrm{d}\mu_\mathrm{i}\mathrm{d}\mu\;\mathrm{d}\sigma_w^2\;\mathrm{d}\sigma_b^2$$

  - Final value is: $y^*\times I_{batch}\times I_{carcass}$

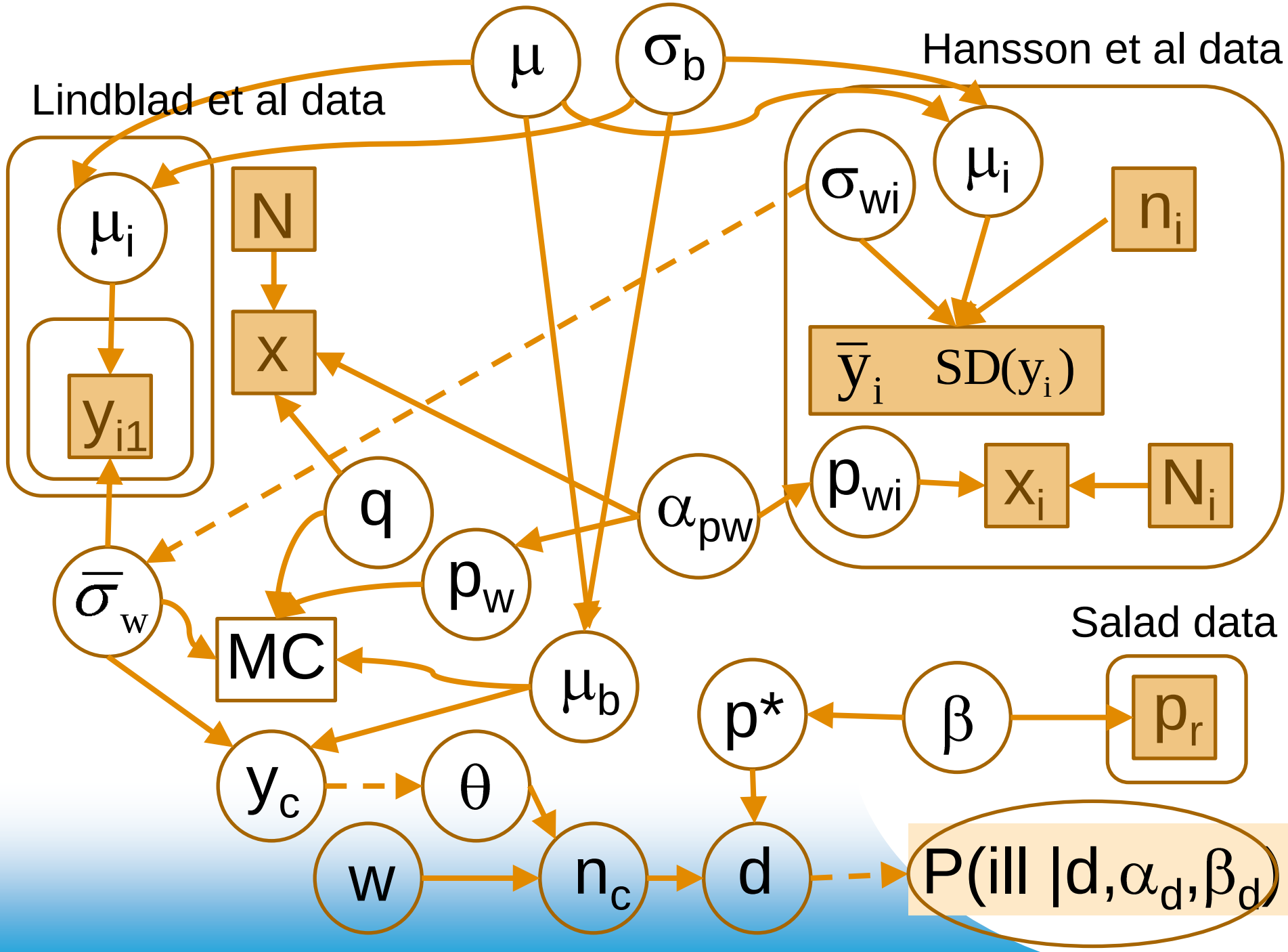
where I_{batch} and $I_{carcass}$ are posterior predictive indicator variables for batch and carcass status.

Evira

Predicting consumer dose

- From this predicted carcass ($\log_{10}$) concentration:
- Assume the resulting meat ($\log_{10}$) concentration is y*-1 (if contamination).
- Distribution of serving size w as log-normal with mean 189g and variance 127 (based on Danish data).
- Resulting bacteria n_c in a serving as Poisson($w10^{y^{*-1}}$).
- Assume meat is cooked well, but cross contamination to salad could occur.
- Resulting bacteria d (dose) in a salad as Bin(n_c,p*).

Evira


Lindblad et al data
Hansson et al data
Salad data
μ
σb
μi
N
x
yi1
σwi
μi
ni
ȳi SD(yi)
pwi
xi
Ni
q
αpw
pw
σ̄w
MC
μb
yc
θ
w
nc
p*
β
pr
d
P(ill |d,αd,βd)

Predicting consumer dose

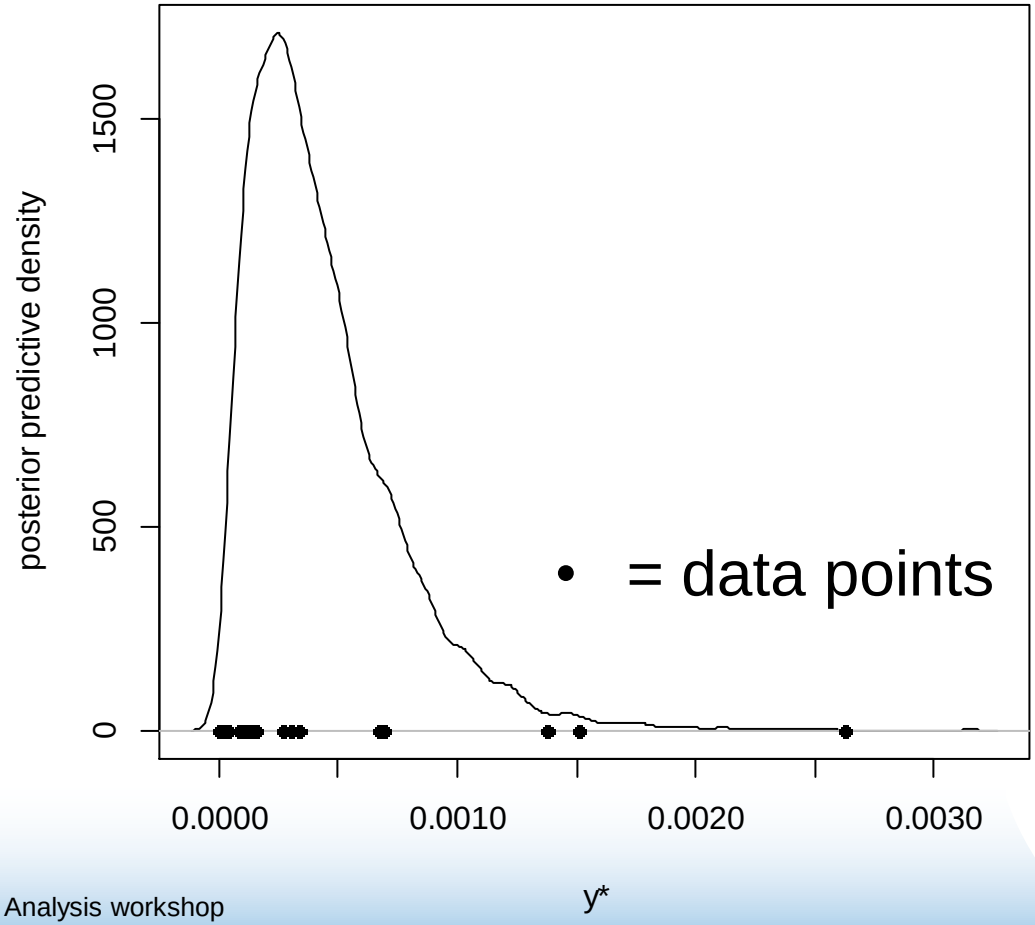
- All these steps are 'forward predictions' with given distributions.
- *No additional Bayesian inference* except for the cross contamination p*
  - 'Salad making experiment': with injected dose of marker bacteria on raw broiler, salad was prepared and resulting concentration in salad was measured → transfer rate p_r in rth experiment.

$$p(\mathrm{p}^* \mid \mathrm{data}) = \int_{-\infty}^{\infty} \mathrm{Beta}(p^* \mid \beta) p(\beta \mid p_1, \ldots, p_r) d\beta$$

Evira

Transfer from broiler to salad

$$p(\mathrm{p}^* \mid \mathrm{data}) = \int_{-\infty}^{\infty} \mathrm{Beta}(p^* \mid \beta)\, p(\beta \mid p_1, \ldots, p_r)\, d\beta$$



Evira

Dose-response model

- With the predicted dose d in salad, apply a dose-response model (α_d,β_d fixed)

$$P(\text{illness} \mid d, \alpha_d, \beta_d) = 0.33\left(1 - \frac{\Gamma(\alpha_d + \beta_d)\Gamma(\beta_d + d)}{\Gamma(\beta_d)\Gamma(\alpha_d + \beta_d + d)}\right)$$

= 0, if d=0.

Conditionally on status: $P(\text{illness} \mid d, I_{carcass}, I_{batch}) = P(\text{illness}|d,\alpha_d,\beta_d)^{x} \; I_{carcass}^{x} \; I_{batch}$

Evira

Effect of microbial criteria

- Every batch is tested
- If MC not met → batch rejected
- If MC met → batch passed

- Predict the consequent prevalence of contaminated carcasses and their concentrations – at the stage where MC is applied

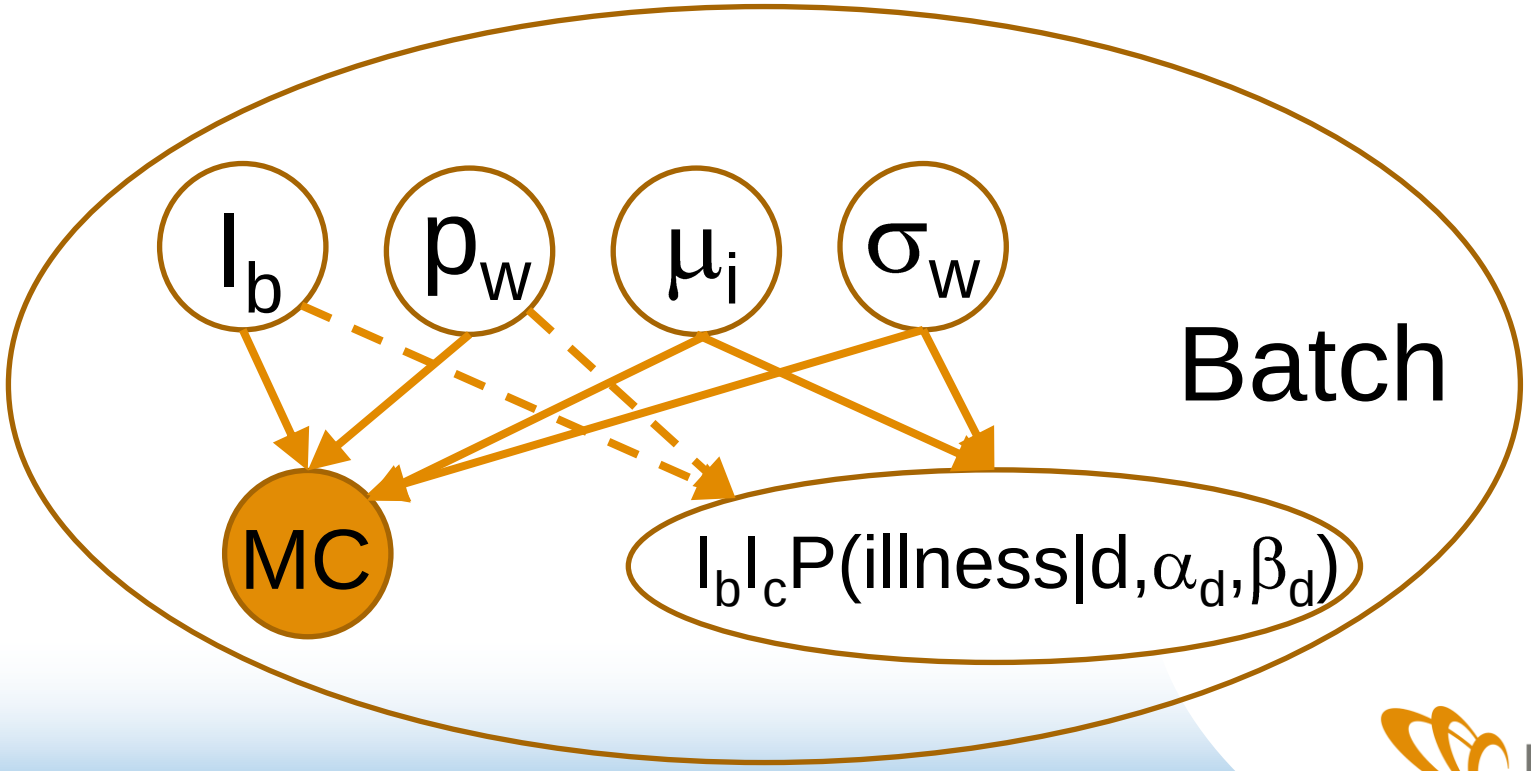
- Predict the consequent consumer risk

- Quantify the effect, accounting for uncertainties

Evira

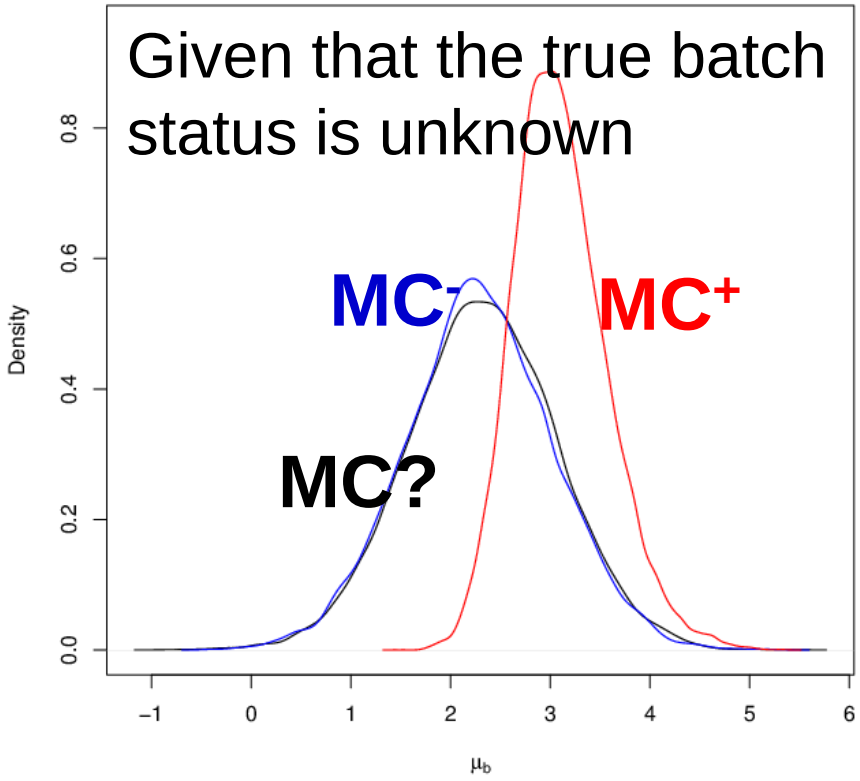
MC result: additional evidence for a batch

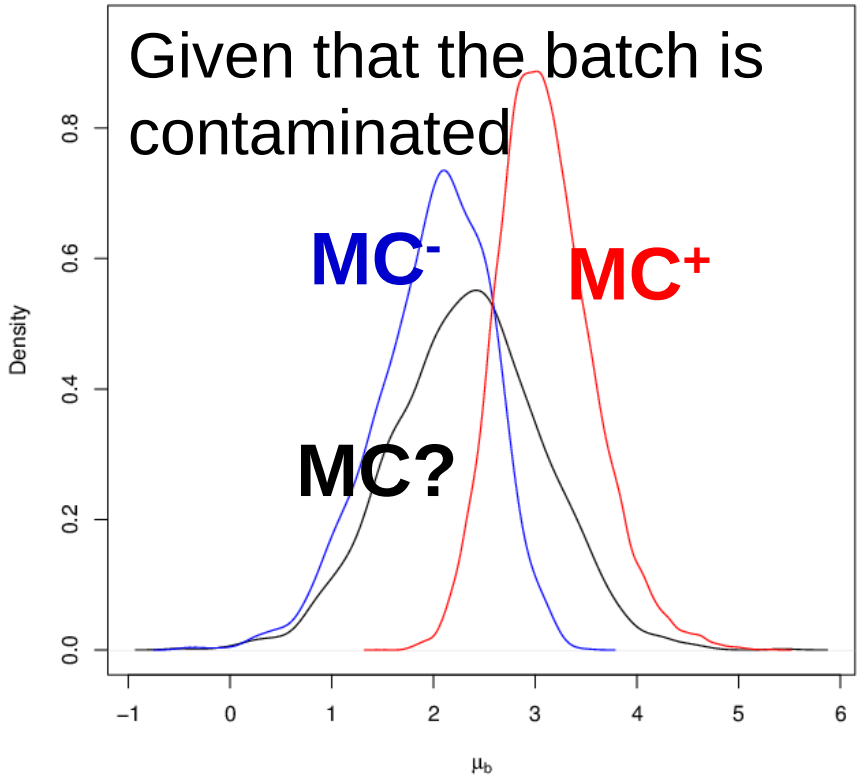
- Compute results conditionally on MC result, to get updated distribution for this batch



Evira

New conditional distributions





This is further reflected on P(illness | MC status)

Evira

These were still preliminary results

- This is currently ongoing work, some structures of the models can still change…
- Main tool is OpenBUGS, under R.

Evira