

tPRiors: Bayesian prevalence estimation with elicited priors

Konstantinos Pateras

1. University of Thessaly, Department of Public Health
2. University of Utrecht, Department of Biostatistics

kostas.pateras@gmail.com

September 22, 2021

Plans within and after training

Chapter 0 -
Prologue

Chapter 1 -
Warming up
(tPRiors -
overview)

Chapter 2 -
tPRiors
specifics -
Single
population

Chapter 3 -
tPRiors
specifics -
Multiple
populations

Epilogue

1 www.menti.com, code: 7651 6327

Chapter 0 -
Prologue

Chapter 1 -
Warming up
(tPRiors -
overview)

Chapter 2 -
tPRiors
specifics -
Single
population

Chapter 3 -
tPRiors
specifics -
Multiple
populations

Epilogue

- 1 Accessible to non-statisticians / field experts
- 2 Spread true prevalence estimation concept
- 3 Collection of methods + additional analyses

Chapter 0 -
PrologueChapter 1 -
Warming up
(tPRiors -
overview)Chapter 2 -
tPRiors
specifics -
Single
populationChapter 3 -
tPRiors
specifics -
Multiple
populations

Epilogue

- 1 Accessible to non-statisticians / field experts
- 2 Spread true prevalence estimation concept
- 3 Collection of methods + additional analyses
- 4 Personal: Learn GUI, Shiny, Rshiny, tcl-tk
- 5 Rstudio shiny contest 2021 [Check if interested]

Chapter 0 -
Prologue

Chapter 1 -
Warming up
(tPRiors -
overview)

Chapter 2 -
tPRiors
specifics -
Single
population

Chapter 3 -
tPRiors
specifics -
Multiple
populations

Epilogue

- 1 Developed with R and Shiny
- 2 No heavy IT skills required
- 3 1st draft in 2-3 months

Chapter 0 -
Prologue

Chapter 1 -
Warming up
(tPRiors -
overview)

Chapter 2 -
tPRiors
specifics -
Single
population

Chapter 3 -
tPRiors
specifics -
Multiple
populations

Epilogue

- 1 Developed with R and Shiny
- 2 No heavy IT skills required
- 3 1st draft in 2-3 months
- 4 Increased learning curve on topic

What's included?

Chapter 0 - Prologue

Chapter 1 - Warming up (tPRiors - overview)

Chapter 2 - tPRiors specifics - Single population

Chapter 3 - tPRiors specifics - Multiple populations

Epilogue

- 8 distinct prevalence models included
- 5 ways to elicit priors (powered by priorGen [Kostoulas 2019])
- 20 modelling set-ups
- ∞ modelling strategies
- Single/Multiple populations - Apparent/True prevalence - No zero/Zero prior prevalence - Informative/ Non Informative

How to work with tPRiors?

Chapter 0 -
Prologue

Chapter 1 -
Warming up
(tPRiors -
overview)

Chapter 2 -
tPRiors
specifics -
Single
population

Chapter 3 -
tPRiors
specifics -
Multiple
populations

Epilogue


- Install with devtools
- Download repository and click Run App through the global.R script [Fastest]
- Use <https://publicintegratedhealth.shinyapps.io/tPRiors>

Starting page

[Start](#)
[Set up](#)
[Priors](#)
[Input & Output](#)
[Report](#)
[Acks](#)

Bayesian true prevalence inference via elicited priors

| tPRiors |



Below the user can find a brief description of the shiny application functions and options.

- (a) In tab (Set up) the user following questions can fix the parameters of the analysis (Choose model, priors, special characteristics)
- (b) In tab (Priors) the user can elicitate the prior distribution(s) with the aid of sliders and visual confirmation
- (c) In tab (Model) the user inputs the observed data and Jags sampling characteristics. A basic inference plot is presented. For multiple population the model may take some time to run.
- (d) In tab (Report) the program returns a dynamic output that changes based on (a), (b) and (c).
- (e) In tab (Acks) acknowledgments and useful links can be found.

Settings may still be changed even after a tab has been fixed by the user. Though, we advise users to perform a 'Reset' of tPRiors when they want to change a previously fixed setting.

The development of tPRiors was funded by H2020 project unCovEr:Unravelling Data for Rapid Evidence-Based Responses. More details can be found in the manuscript, K Pateras and P Kostoulas, tPRiors: An R Shiny tool for generating prior and producing posterior distributions for disease prevalence

4
prior elicitation approaches

8
prevalence model variations

3
preloaded datasets for demonstration

Set-up page

Start **Set up** Prior(s) Export & Output Report Help

Status: 'Not set'

Please first select options and press 'Fix'

Do you have single or multiple populations/clusters

☐ Single ☒ Multiple

Do you want to model the true or the apparent prevalence?

☒ True prevalence ☐ Apparent prevalence

Do you want to account for zero true prevalence?

☒ No ☐ Yes

Would you specify informative priors?

☒ Yes ☐ No

If No, then the apparent prevalence will be modelled

Which measure of central tendency or dispersion would you like to use for the true prevalence prior?

☒ Mean

Reset tPRiors Fix setup!

Prior page

Start
Set up
Prior(s)
Report & Output
Report
About

Prevalence
Sensitivity
Specificity

Set a prevalence prior

Specify your prior belief about the Mean True prevalence :

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

Is the percentile the upper limit of the Mean ?

☐ Yes ☒ No

Specify the upper or lower limit for the Mean at the specified level of confidence:

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

Specify the level of confidence that the true value of the Mean is greater or lower than the percentile value:

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

Caution! Press 'set priors' only when all values are set.
In case 1 slider gets stuck, select 'Patch', wait and select 'Fixed' to continue

☒ Fixed ☐ Patch

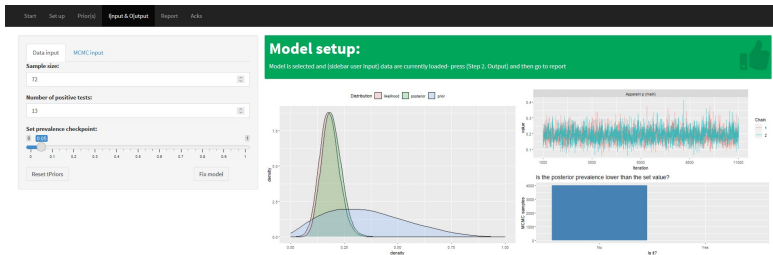
Reset tPriors Example Set prior(s)?

Prior setup:

Setup done - Please click a prior and press select



Input & Output page



Report page


[Start](#) [Set up](#) [Prior\(s\)](#) [Input & Output](#) [Report](#) [Ack](#)

Status report:
The report will become available below

Input

Model

Output

 Download data

Calculation in progress. This may take a while...

Report page

[Start](#)
[Set up](#)
[Prior\(s\)](#)
[Input & Output](#)
[Report](#)
[Index](#)

Status report:

The report will become available below

tPRiors-dynamic-report

15 September, 2021

General information

This report has automatically been generated by the shiny web-application tPRiors as an R Markdown document based on your data input and prior selection. The web-application can be found at <https://konstantinos.shinyapps.io/tPRiors/>. We advise users that after observing the results of this report to avoid re-updating their prior beliefs to avoid hampering the credibility of these results.

The following section describes your input. During set-up the user assumed that:

1. Multiple modelled.
2. No zero prevalence was modeled and.
3. between the Apparent and True prevalence the True prevalence was modeled and
4. (the) Mean was used to elicitate prior knowledge.

If the true prevalence (inline equation test: π_x) is modelled the following relation is utilized to acquire its posterior distribution, inline equation test: $\pi_x = \pi_y - S_p(1 - \pi_y) - (1 - S_p)$, where inline equation test: S_p, S_e denotes the specificity and sensitivity of the diagnostic test and inline equation test: π_y the apparent prevalence.

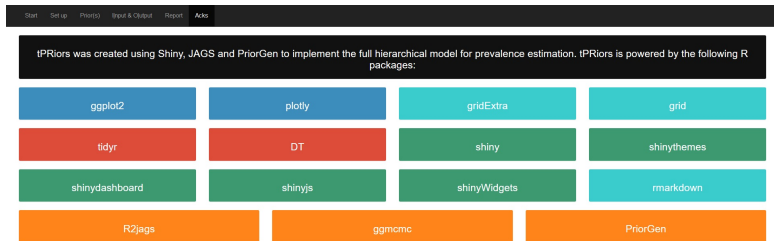
The elicited prevalence prior

The selected prior distribution of the True prevalence has the following descriptive characteristics and density plot.

```
## [1] "Summary of True prevalence Beta(0.1350,0.5477) prior"
```

| ## | Min. | 1st Q. | Median | Mean | 2nd Q. | Max. |
|----|-----------|-----------|-----------|-----------|-----------|-----------|
| ## | 0.0000000 | 0.0001284 | 0.1003370 | 0.2042702 | 0.3105904 | 0.9999992 |

Acks page



Single population toy-example

Chapter 0 -
Prologue

Chapter 1 -
Warming up
(tPRiors -
overview)

Chapter 2 -
tPRiors
specifics -
Single
population

Chapter 3 -
tPRiors
specifics -
Multiple
populations

Epilogue

- Prevalence of a moderate adverse event (fatigue) of a vaccine.
- 2 Observed 50 events in a sample of 1000.
- Prior information:
 - 1 A publication stated that average risk for fatigue is low.
 - 2 Experts report that fatigue is observed with very high sensitivity but moderate specificity.

Set-up page

Start Set up Outputs About tPRiors Support Help

Do you have single or multiple population(s) clusters?
☒ Single ☐ Multiple

Do you want to model the true or the apparent prevalence?
☒ True prevalence ☐ Apparent prevalence

Do you want to account for zero true prevalence?
☒ No ☐ Yes

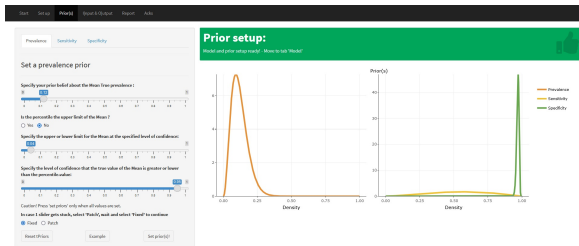
Would you specify information priors?
☒ Yes ☐ No
 If No, then the apparent prevalence will be modelled.

Which measure of central tendency or dispersion would you like to use for the true prevalence prior?
☒ Mean ☐ Median ☐ Mode ☐ Percentiles

Reset defaults Go to setup

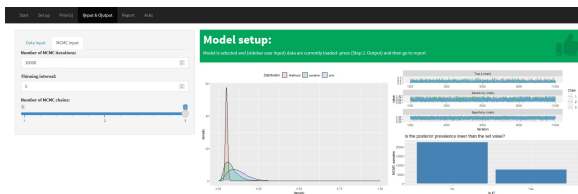
Status: 'Set'
 Your input assumes that: 1. Single will be modelled, 2. No zero prevalence will be modelled, 3. the True prevalence will be modelled and 4. No informative prior(s) will be modelled.

- 1 Single - True - No zero - Informative - Mean.
- 2 If Apparent ($Se=Sp=1$), if Non-informative, then 3 prior prevalence choices available.
- 3 Currently, selected measure applies to all priors.



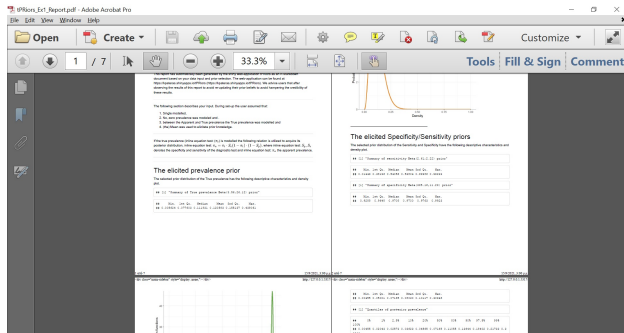
- 1 Does not currently support raw hyper-parameter values
- 2 Some prior specifications can be non-appropriate
- 3 Caution! If "Set" is selected, change values slowly to avoid technical errors.

Input & Output page



- 1 Meshing with MCMC input should not change inference.
- 2 Change input data to check the dynamics of posterior, prior, likelihood.

Report page



1 Open "tPRiors_Ex1_Report.pdf", discuss

Further analyses

Chapter 0 -
Prologue

Chapter 1 -
Warming up
(tPRiors -
overview)

Chapter 2 -
tPRiors
specifics -
Single
population

Chapter 3 -
tPRiors
specifics -
Multiple
populations

Epilogue

- 1 Check ESS, diagnostics, extra plots,
- 2 Check models, input data, output data
- 3 Use ggmcmc for clearer diagnostic plots

Exercise - Analysis and replication of results! (15-20')

Chapter 0 -
Prologue

Chapter 1 -
Warming up
(tPRiors -
overview)

Chapter 2 -
tPRiors
specifics -
Single
population

Chapter 3 -
tPRiors
specifics -
Multiple
populations

Epilogue

- Check manuscript Perquin et al. 2008 GitHub/day3
- Calculate the posterior prevalence of dementia for 65+ participants
- Experts discuss that:
 - 1 the average risk for dementia is relatively low but depends on age range.
 - 2 Conducted test known to show high sensitivity and very high specificity.
 - 3 Work individually. Report posterior median and IQR of the true prevalence.

Exercise - Analysis and replication of results! (15-20')

Chapter 0 -
Prologue

Chapter 1 -
Warming up
(tPRiors -
overview)

Chapter 2 -
tPRiors
specifics -
Single
population

Chapter 3 -
tPRiors
specifics -
Multiple
populations

Epilogue

- Check manuscript Perquin et al. 2008 GitHub/day3
- Calculate the posterior prevalence of dementia for 65+ participants
- Experts discuss that:
 - 1 the average risk for dementia is relatively low but depends on age range.
 - 2 Conducted test known to show high sensitivity and very high specificity.
 - 3 Work individually. Report posterior median and IQR of the true prevalence.
 - 4 www.menti.com (1588 8807)

Set-up page

Start
Set up
Prior(s)
Input & Output
Report
Acks

Do you have single or multiple populations/clusters

☐ Single ☒ Multiple

Do you want to model the true or the apparent prevalence?

☒ True prevalence ☐ Apparent prevalence

Do you want to account for zero true prevalence?

☒ No ☐ Yes

Would you specify informative priors?

☒ Yes ☐ No

If No, then the apparent prevalence will be modelled

Which measure of central tendency or dispersion would you like to use for the true prevalence prior?

☒ Mean

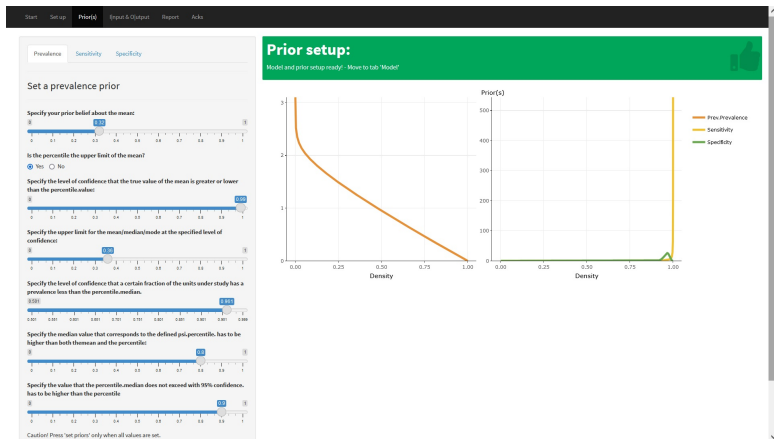
Reset tPRiors Fix setup!

Status: 'Set'

Your input assumes that: 1. Multiple will be modelled, 2. No, zero prevalence will be modelled, 3. the True prevalence will be modelled and 4. Yes informative prior(s) will be modelled

- 1 Multiple - True - NonZero - Informative
- 2 One elicitation measure, if non-informative is selected,
 $\psi \sim IG(0.1, 0.1)$.

Prior page



- 1 Conditions applied, caution how to set values
- 2 If plotted, change one at a time, observe change (not suggested)

Input & Output page

Start Set up Prior(s) **Input & Output** Report Ack

Data input **MCMC input**

Download example file

Load dataset:
☐ Preload ☒ Upload

Choose .xls/.csv file

Browse... Dementia_Updated.xls
 Upload complete

Set prevalence checkpoint:
 0.05

Reset tPriors Fix model

Model setup:

Please set up the model, priors and input data in Tabs 'Set up', 'Priors' and 'Model' before you move forward

Step 1. Data Step 2. Output

Show 10 entries Search:

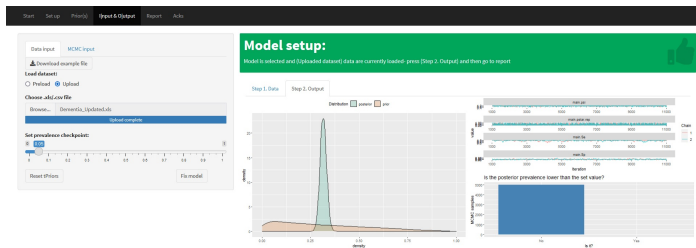
| Study | Country | Years. | positive | n | Appa.Prev | Cis_95. |
|-------------------------|------------|----------------------|----------|------|-----------|---------------|
| 1 Ravaglia et al. | Italy | 1999 | 60 | 961 | 0.062 | (0.047,0.078) |
| 2 Tognoni et al. | Italy | 2000 | 100 | 1662 | 0.06 | (0.049,0.072) |
| 3 Gascon-Bayarri et al. | Spain | 2002 | 165 | 1754 | 0.094 | (0.08,0.108) |
| 4 Fish et al. | UK | 2003 | 88 | 1664 | 0.053 | (0.042,0.064) |
| 5 Bermejo-Pareja et al. | Spain | 1994-1995, 1997-1998 | 306 | 5278 | 0.058 | (0.052,0.064) |
| 6 Mathillas et al. | Sweden | 2000-2002, 2005-2007 | 287 | 895 | 0.321 | (0.29,0.351) |
| 7 Tola-Arribas et al. | Spain | 2009-2010 | 184 | 2170 | 0.085 | (0.073,0.097) |
| 8 Lucca et al. | Italy | 2002-2010 | 894 | 2501 | 0.357 | (0.339,0.376) |
| 9 Perquin et al. | Luxembourg | 2008 | 53 | 1377 | 0.038 | (0.028,0.049) |

Showing 1 to 9 of 9 entries

Previous 1 Next

- 1 Load preselected data or upload .csv or .xls file. Must contain two columns named 1. positive and 2. n.
- 2 Download example file for a template.

Input & Output page



- 1 Necessary step to check Step 2. Output before moving to report!
- 2 Diagnostics look cramped...

Exercise - Analysis and replication of results! (15-20')

Chapter 0 -
Prologue

Chapter 1 -
Warming up
(tPRiors -
overview)

Chapter 2 -
tPRiors
specifics -
Single
population

Chapter 3 -
tPRiors
specifics -
Multiple
populations

Epilogue

Download article Bacigalupo et al. 2018. Produce posterior box-plots for the study-specific prevalences of dementia using tPRiors and the multiple population models. Use similar SE and SP priors to previous example. 3 goals! Work in pairs but individually.

- Goal 1 is to produce the posterior study-level boxplot.
- Goal 2 is to save all information for reproducibility.
- Goal 3 is to re-produce the posterior study-level boxplot of your peer.

Bibliography [1 more slide]

Chapter 0 -
PrologueChapter 1 -
Warming up
(tPRiors -
overview)Chapter 2 -
tPRiors
specifics -
Single
populationChapter 3 -
tPRiors
specifics -
Multiple
populations

Epilogue

- Konstantinos Pateras, Polychronis Kostoulas. tPRiors: Bayesian prevalence estimation with elicited priors, Submitted
- Bagipulo et al. A Systematic Review and Meta-Analysis on the Prevalence of Dementia in Europe 2018
- Perquin et al and cognitive complaints in the context of high cognitive reserve: A population-based study. 2015
- Polychronis Kostoulas priorGen, R packages, 2019

Final thoughts

www.menti.com
Code: 2118 8043