Diagnostic Accuracy, part 2 & Practice-Based Evidence

Evidence-Based Practice in Speech-Language Therapy (SHSC 2033)

Session 9

Thomas Klee & Elizabeth Barrett

香 港 大 學
THE UNIVERSITY OF HONG KONG

Outline

1. Diagnostic accuracy: post-test probability
2. Practice-based evidence
3. Evidence from client preferences
4. Group discussion

Post-test Probability

Screening outcomes [1]

	Language delay	Language normal	Total
Screen +	10	2	12
Screen −	1	51	52
Total	11	53	64

[1] Klee, Pearce, and Carson (2000)

Classification accuracy measures [2]

```
# Example 1: calculate classification measures from Klee et al. (2000)
dat <- as.table(matrix(c(10, 2, 1, 51), nrow = 2, byrow = TRUE))
rval <- epi.tests(dat, conf.level = 0.95)
print(rval) # for point estimates and CIs rounded to 2 DPs
```

```
##           Outcome +    Outcome -      Total
## Test +           10            2         12
## Test -            1           51         52
## Total            11           53         64
##
## Point estimates and 95 % CIs:
## -------------------------------------------------------
## Apparent prevalence                   0.19 (0.10, 0.30)
## True prevalence                       0.17 (0.09, 0.29)
## Sensitivity                           0.91 (0.59, 1.00)
## Specificity                           0.96 (0.87, 1.00)
## Positive predictive value             0.83 (0.52, 0.98)
## Negative predictive value             0.98 (0.90, 1.00)
## Positive likelihood ratio            24.09 (6.11, 95.02)
## Negative likelihood ratio             0.09 (0.01, 0.61)
## -------------------------------------------------------
```

[2] Calculated using *epiR* package of R (R Core Team, 2019; Stevenson, 2017)

What does **prevalence** refer to here?

- **Apparent** prevalence: $12/64 = .19$ [3]
- **True** prevalence: $11/64 = .17$ [4]
- Do these accurately reflect population values?

[3] Based on number of **test positives** in the sample.

[4] Based on number of **cases** in the sample.

Be careful...

- If the prevalence (base rate) of the disorder in the research sample is different from the prevalence in the population, the estimates reported for PPV and NPV will not be accurate.
- Notice the results of our sample.

Klee et al sample

```
# Example 1: calculate classification measures from Klee et al. (2000)
dat <- as.table(matrix(c(10, 2, 1, 51), nrow = 2, byrow = TRUE))
rval <- epi.tests(dat, conf.level = 0.95)
print(rval) # for point estimates and CIs rounded to 2 DPs
```

```
##           Outcome +    Outcome -      Total
## Test +           10            2         12
## Test -            1           51         52
## Total            11           53         64
##
## Point estimates and 95 % CIs:
## -------------------------------------------------------
## Apparent prevalence                   0.19 (0.10, 0.30)
## True prevalence                       0.17 (0.09, 0.29)
## Sensitivity                           0.91 (0.59, 1.00)
## Specificity                           0.96 (0.87, 1.00)
## Positive predictive value             0.83 (0.52, 0.98)
## Negative predictive value             0.98 (0.90, 1.00)
## Positive likelihood ratio            24.09 (6.11, 95.02)
## Negative likelihood ratio             0.09 (0.01, 0.61)
## -------------------------------------------------------
```

Summary

- Sensitivity and specificity aren't affected by prevalence.
- Neither are the likelihood ratios, since they are based on sensitivity and specificity.
  - LR+ $= \mathit{Sensitivity}/(1 - \mathit{Specificity})$
  - LR– $= (1 - \mathit{Sensitivity})/\mathit{Specificity}$
- This is why LRs are preferable to PPV and NPV.

LR+

	Condition present	Condition absent
Index test +	True positive	False positive
Index test −	False negative	True negative

- LR+ = Sensitivity / (1 - Specificity)
- LR+ = proportion of TPs / proportion of FPs
- LR+ = proportion of (+) test results in those with the condition / proportion of (+) test results in those without the condition

Interpreting LRs [5]

- LRs of > 10 or < 0.1 indicate large and often conclusive changes from **pre- to post-test probability**.
- LRs of 5 to 10 and 0.1 to 0.2 indicate moderate shifts in **probability**.
- LRs of 2 to 5 and 0.2 to 0.5 indicate small (but sometimes important) shifts in **probability**.
- LRs of 1 to 2 and 0.5 to 1 alter **probability** to a small (and rarely important) degree.

[5] Guyatt, Rennie, Meade, and Cook (2008, p. 208)

Post-test probability can be calculated if we know. . .

1. the LR associated with the test result
2. the pre-test probability of the condition [6]

[6] taking the clinical setting and purpose into account (e.g., screening vs clinical assessment)

Basis of Bayes' theorem

We possess knowledge at any given point in time.

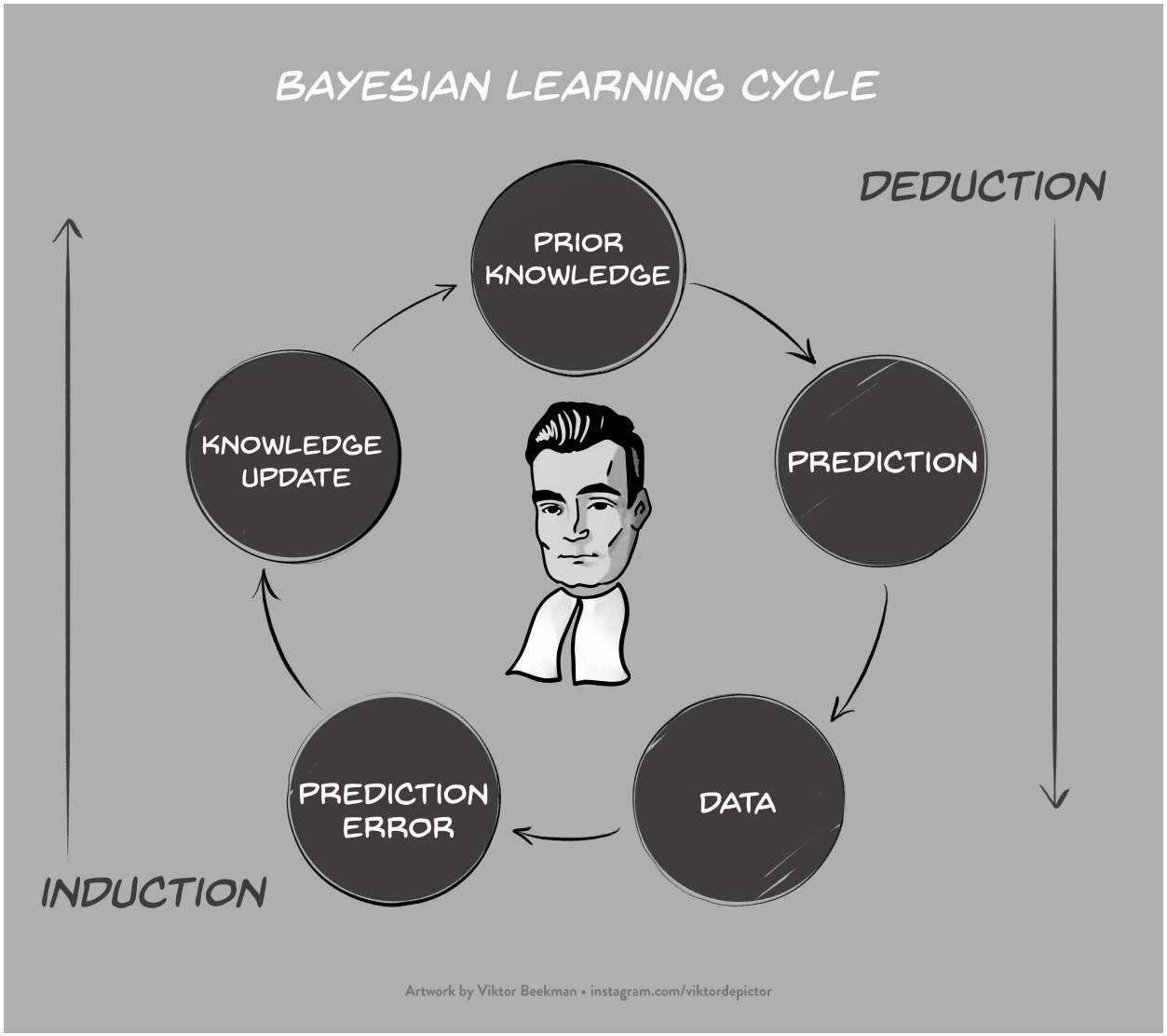
↓

We encounter new information (data).[7]

↓

We revise (update) our knowledge.

[7] Based on our own experience, high quality research evidence, etc.


BAYESIAN LEARNING CYCLE
DEDUCTION
PRIOR KNOWLEDGE
PREDICTION
DATA
PREDICTION ERROR
KNOWLEDGE UPDATE
INDUCTION
Artwork by Viktor Beekman • instagram.com/viktordepictor

Bayes' theorem

"Bayes' theorem is concerned with conditional probability. That is, it tells us the probability that a theory or hypothesis is true if some event has happened." [8]

"...by updating our initial belief about something with objective new information, we get a new and improved belief." [9]

[8] Silver (2012, p. 243)

[9] McGrayne (2011, p. xi)

Applying Bayesian thinking to clinical measures

We possess knowledge at any given point in time.

Prior (pre-test) probability

↓

We encounter new information (data).

LR

↓

We revise (update) our knowledge.

Posterior (post-test) probability

LR Nomogram[10]

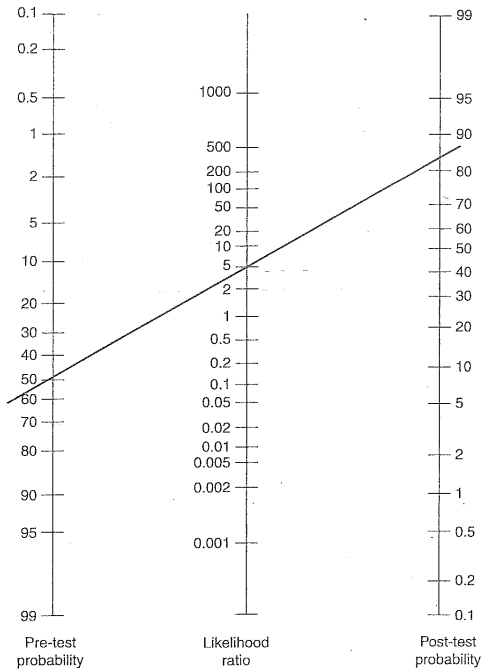


FIGURE 8–1 Nomogram for converting pre-test likelihoods (*left column*) to post-test likelihoods (*right column*) by drawing a straight line from the pre-test likelihood through the likelihood ratio for the test result (5). (From Fagan TJ. Nomogram for Bayes's theorem. *N Engl J Med* 1975;293:257, with permission.)

[10] Estimate post-test probability with a ruler (Haynes, Sackett, Guyatt, & Tugwell, 2006, p. 285)

Estimate post-test probability with R [13]

- Example: With a pre-test probability of 13%,[11] what's the post-test probability of having the condition, given the test result?

```
library(epiR)
epi.nomogram(se = 0.90909091, sp = 0.96226415,
lr = NA, pre.pos = 0.13, verbose = FALSE)
```

  * Given a positive test result, the post-test probability of being disease positive is 0.78
  * Given a negative test result, the post-test probability of being disease negative[12] is 0.01

[11] Based on prevalence data reported in Beitchman et al (1986) and Tomblin et al (1996)

[12] I think this should read "disease positive".

[13] with *epiR* package (R Core Team, 2019; Stevenson, 2017).

Interpretation

- Given a positive screening result, there is a 78% chance the child has the condition being screened for (language disorder).
- Given a negative screening result, there is a 1% chance the child has the condition.
- However, before concluding this, **notice which pre-test probability** these were based on.
  - ▷ Adjust the pre-test probability if the sample prevalence differs from the population.
  - ▷ Post-test probabilities are identical to PPV and NPV when the proportion of cases in the sample is the same as in the population.

What if this was a standardized test given to a child referred for assessment? [15]

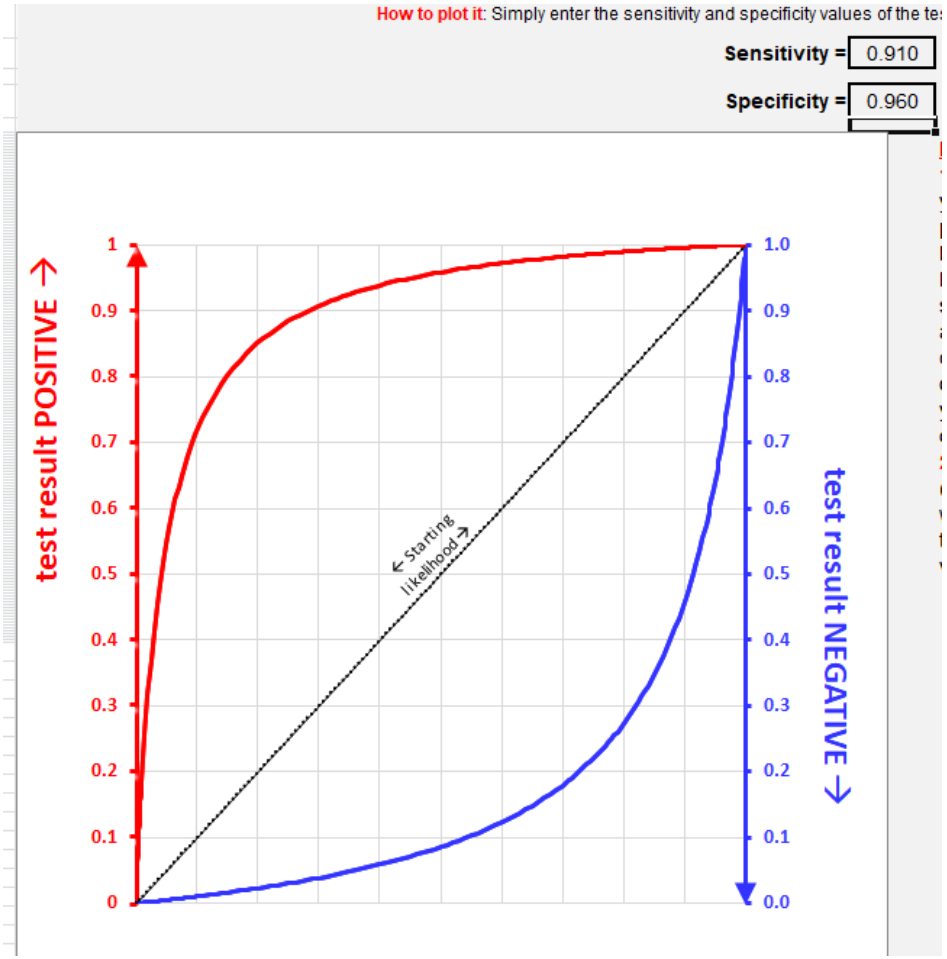
```
library(epiR)
epi.nomogram(se = 0.90909091, sp = 0.96226415,
lr = NA, pre.pos = 0.50, verbose = FALSE)
```

* Given a positive test result, the post-test probability of being disease positive is 0.96
* Given a negative test result, the post-test probability of being disease negative[14] is 0.09

[14] I think this should read "disease positive".

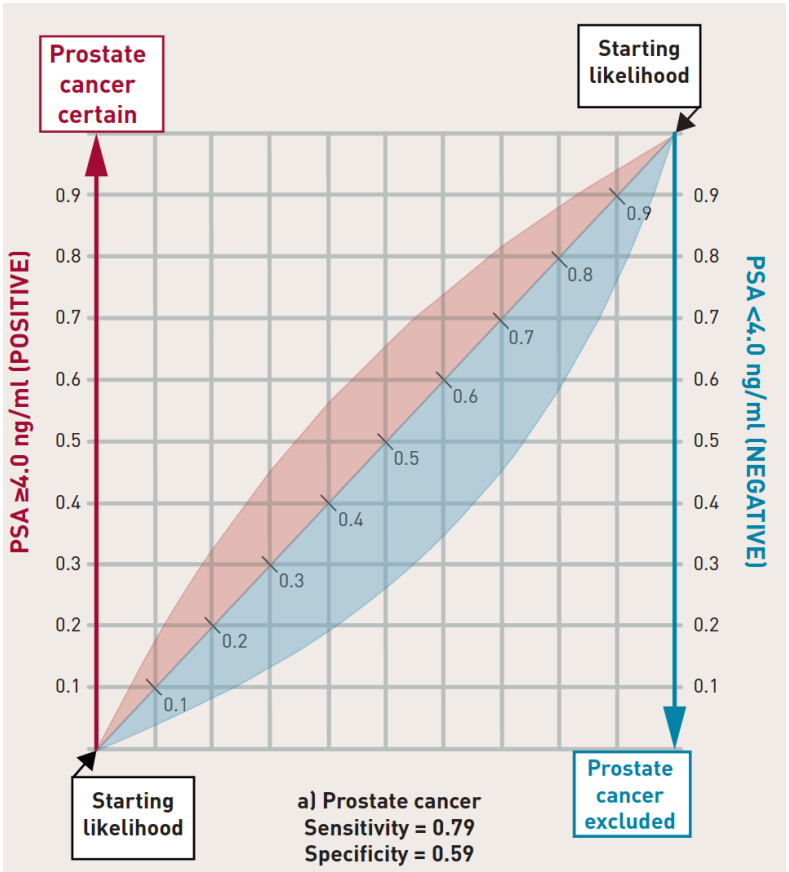
[15] Hypothetical clinical population where half those referred for assessment are diagnosed with language disorder.

Estimate post-test probability with a leaf plot[16]



[16] Coulthard and Coulthard (2019); data based on Klee et al. (2000).

Leaf plot for PSA results[17]

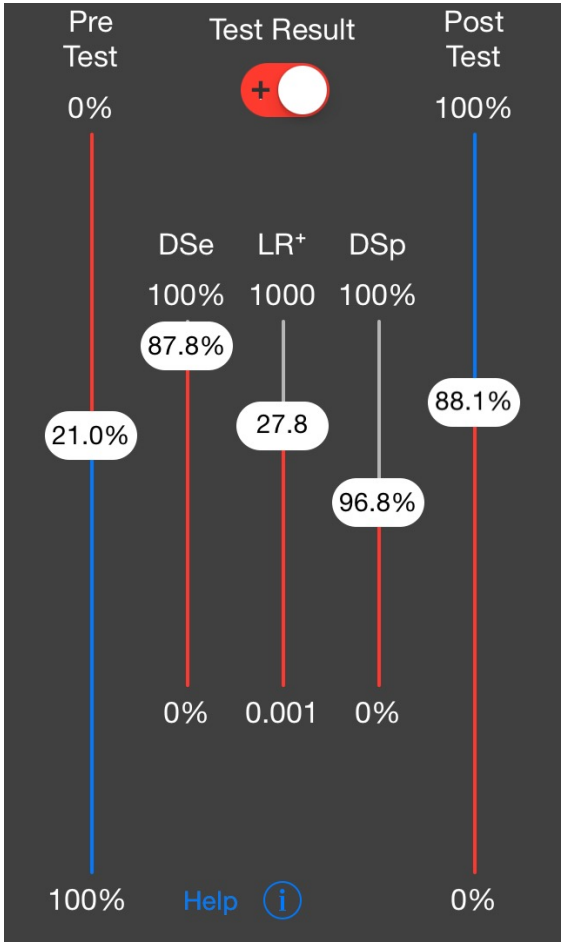


[17] Coulthard and Coulthard (2019)

Try it yourself with the person next to you

- Download leaf plot calculator from `https://childhealthafrica.org/downloads`.
- Open it in MS Excel.
- Enter the sensitivity and specificity values from Klee et al. (2000) study.
- Explore how different pre-test probabilities affect post-test probability.
- Discuss how varying these numbers affects how you interpret test results.

Estimate post-test probability with DocNomo



(Free for iOS devices. Values displayed are from another study.)

Practice-Based Evidence

E^3BP

"...the conscientious, explicit, and judicious integration of best available

1. *external evidence from systematic research,*
2. *evidence internal to clinical practice, and*
3. *evidence concerning the preferences of a fully informed patient."* [18]

[18] Dollaghan (2007, p. 2)

Practice/patient-based evidence

"The applicability of the external evidence from even a very strong study to an individual patient will need to be tested rather than assumed." [19]

[19] Dollaghan (2007, p. 115)

Practice/patient-based evidence

- How effective is this intervention with this client?
- Involves evaluating effectiveness of intervention with your own clients
- CAPE: Checklist for Appraising Patient/Practice Evidence
- "Not. . . a critical appraisal form but rather a checklist of issues to consider in seeking valid, important evidence from clinical practice." [20]

[20] Dollaghan (2007, p. 115)

CAPE: foreground question

P For your client or patient

I is the intervention

O associated with the outcome you've identified

C compared with contrasting intervention condition?

CAPE: validity of the evidence [21]

1. Observational (A-B) or Experimental (e.g., A-B-A, alternating treatments, multiple baseline)?
2. Randomization (e.g., targets to conditions, order of conditions)?
3. Stable baseline(s)?
4. Adequate length of treatment phase(s)?
5. Treatment consistency, fidelity?
6. Potential nuisance factor(s)?
7. Valid and reliable measures?
8. Measures administered with blinding?

[21] Dollaghan (2007, p. 116)

CAPE: importance of the evidence [22]

1. Magnitude of treatment effect?
2. Evidence of maintenance, generalization, and social validity of treatment effect?
3. Substantial cost-benefit advantage?

[22] Dollaghan (2007, p. 116)

Client Preferences

Client preferences

- Concerns "the preferences of fully informed patients relating to the clinical options they face" [23]
- Client-centered care
- Qualitative research can inform this area of clinical practice.
- Establish a common ground with your client.
- Prepare information on clinical options.
- CAPP: Checklist for Appraising Patient Preferences

[23] Dollaghan (2007, p. 123)

CAPP: foreground question [24]

- **P** For this client, once fully informed of costs, risks, benefits
- **I** is one clinical option
- **O** a preferred outcome
- **C** compared with other clinical options?

[24] Dollaghan (2007, p. 128)

Useful resources[25]

Checklists for clinicians (Dollaghan, 2007)

CAPE to appraise patient/practice evidence (p. 159)
CAPP to appraise evidence on patient preferences (p. 161)

Description of diagnostic accuracy measures

`http://www.students4bestevidence.net/ebm-for-diagnostic-tests/`

Estimating post-test probability

- `https://ebm-tools.knowledgetranslation.net/calculator/diagnostic/`
- `http://www.cebm.net/catmaker-ebm-calculators/`
- `https://itunes.apple.com/us/app/docnomo/id901279945?mt=8`
- *R* software's *epiR* package (Stevenson, 2017)

[25] URLs checked on 2019-03-25

Group discussion

- Break up into your assigned groups and discuss:
  1. how you could begin to collect practice-based evidence with your own clients;
  2. how you could use CAPP with your own clients.

References I

Coulthard, M. G., & Coulthard, T. (2019). The leaf plot: a novel way of presenting the value of tests. *British Journal of General Practice*, 1–2.

Dollaghan, C. A. (2007). *The handbook for evidence-based practice in communication disorders*. Baltimore, MD: Paul H. Brookes Publishing Co.

Guyatt, G., Rennie, D., Meade, M. O., & Cook, D. J. (2008). *Users' guides to the medical literature: essentials of evidence-based clinical practice* (2nd ed.). New York: McGraw Hill.

Haynes, R. B., Sackett, D. L., Guyatt, G. H., & Tugwell, P. (2006). *Clinical epidemiology: how to do clinical practice research* (3rd ed.). Philadelphia: Lippincott Williams & Wilkins.

References II

Klee, T., Pearce, K., & Carson, D. K. (2000). Improving the positive predictive value of screening for developmental language disorder. *Journal of Speech, Language, and Hearing Research*, *43*, 821–833.

McGrayne, S. B. (2011). *The theory that would not die: how Bayes' rule cracked the enigma code, hunted down Russian submaries & emerged triumphant from two centuries of controversy*. New Haven, CT: Yale University Press.

R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from `https://www.r-project.org/`

Silver, N. (2012). *The signal and the noise: why so many predictions fail but some don't*. New York: Penguin Press.

References III

Stevenson, M. (2017). *epiR: tools for the analysis of epidemiological data (R package version 0.9-93).* Retrieved from `https://cran.r-project.org/package=epiR`