Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Incorporating Expert Opinion in Statistical Analysis

Philip Cooney
Ph.D. Candidate Computer of Statistics and Computer Science

11/04/22

Motivations

Including Expert Opinion – (Kaldane et al 1998)

- It's a good idea
- Consensus in literature is that expert opinion should be elicited on the observable space and not on regression coefficients
- Experts should not be asked to estimate moments of a distribution (except possibly the first moment); they should be asked to assess quantiles or probabilities of the predictive distribution
- Frequent feed-back should be given to the expert during the elicitation process;
- Experts should be asked to give assessments both unconditionally and conditionally on hypothetical observed data – Think that this is included because Kaldane's approach uses this.
- Most up to date review of Elicitation provided by (Mikkola et al 2021)

Difficulty eliciting opinions

Garthwaite et al 2005

Overconfidence: *Difficulty assessing tails of the distribution*

Judgement by representativeness: This is applicable for questions relating to conditional probability P(B|A) without thinking about the unconditional probability.

Judgement by availability: Recall is affected by factors such as familiarity and newsworthy events also impact disproportionately on our memory, so you might overestimate the probability of a plane crash with fatalities

Anchoring: *Example Question:* % of African countries in UN, each participant was given a random percentage

Conservatism: Anchored to the prior probability, can't do Bayes Theorem in their head!

Law of small numbers: Assuming law of large numbers applies to small numbers

Hindsight bias: Arise when people are asked to assess their a priori probability of an event that has actually occurred

Previous Approaches

Kaldane et al 1980

Was the first and considered this for linear regression:

This approach requires the expert to consider responses conditional not only on covariates but also realisations of the response at different design points.

Table 2. Elicited conditional assessments

Point	*CA*	*PCCA*	*Initial percentiles*		*New observation*	*Conditional assessments*					*Percentile*
			50th	*75th*		*1*	*2*	*3*	*4*	*5*	
1	0.0	0.0	0.0	2.1	−0.88	**−0.75**	**−0.50**	**0.25**	**4.50**		50th
							3.00				75th
2	1.0	0.0	0.0	3.5	1.37		**2.5**	**1.5**	**6.0**		50th
								3.6			75th
3	0.0	0.5	1.0	3.1	2.62			**3.5**	**5.5**		50th
									9.0		75th
4	1.0	0.5	5.0	8.5	7.37				**6.5**	**3.5**	50th
5	0.0	1.0	2.0	4.1							

Previous Approaches

Bedrick et al 1996

Conditional mean priors (CMP) & Data Augmentation Priors (DAP)

The expert in this case contributes a distribution, perhaps based on elicited moments or modes, conditional on the values of the covariates at a design point.

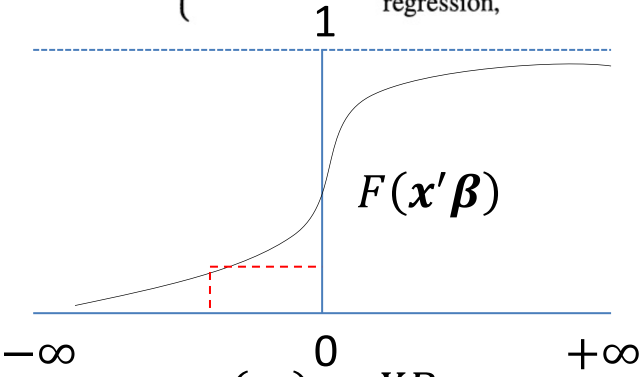
The form of the prior may depend on the chosen link function and likelihood.

Previous Approaches

Bedrick et al 1996

Consider a dichotomous outcome:

$$F(\mathbf{x}'\boldsymbol{\beta}) = \begin{cases} e^{\mathbf{x}'\boldsymbol{\beta}}/[1+e^{\mathbf{x}'\boldsymbol{\beta}}] & \text{logistic regression} \\ \Phi(\mathbf{x}'\boldsymbol{\beta}) & \text{probit regression} \\ 1-\exp[-e^{\mathbf{x}'\boldsymbol{\beta}}] & \text{complementary log-log regression,} \end{cases}$$



$$\mathrm{g}(\mathbf{m}) = XB$$

$$m_i = E(y_i|\boldsymbol{x_i})$$

Likelihood of the data

$$L(\boldsymbol{\beta}) \propto \prod_{i=1}^{n} F(\mathbf{x}_i'\boldsymbol{\beta})^{N_i y_i}[1-F(\mathbf{x}_i'\boldsymbol{\beta})]^{N_i(1-y_i)}.$$

We have a beta prior for the m_i's at each elicited design point

$$\pi(\tilde{\mathbf{m}}) \propto \prod_{i=1}^{p} \tilde{m}_i^{a_{1i}-1}(1-\tilde{m}_i)^{a_{2i}-1}. \quad (1)$$

We have an (inverse) function which maps beta to m.

$$g^{-1}(\mathbf{x}'\boldsymbol{\beta}) = F(\mathbf{x}'\boldsymbol{\beta})$$

The induced prior on $\boldsymbol{\beta}$ for the model $g^{-1}(\mathbf{x}'\boldsymbol{\beta}) = F(\mathbf{x}'\boldsymbol{\beta})$ is

$$\pi(\boldsymbol{\beta}) \propto \prod_{i=1}^{p} F(\tilde{\mathbf{x}}_i'\boldsymbol{\beta})^{a_{1i}-1}[1-F(\tilde{\mathbf{x}}_i'\boldsymbol{\beta})]^{a_{2i}-1} f(\tilde{\mathbf{x}}_i'\boldsymbol{\beta}). \quad (2)$$

For the logistic model, $f(\cdot) = F(\cdot)(1-F(\cdot))$. With the probit model, $f(\tilde{\mathbf{x}}_i'\boldsymbol{\beta}) \propto \exp(-\frac{1}{2}(\tilde{\mathbf{x}}_i'\boldsymbol{\beta})^2)$, and with the complementary log-log model, $f(\tilde{\mathbf{x}}_i'\boldsymbol{\beta}) \propto \exp(\tilde{\mathbf{x}}_i'\boldsymbol{\beta} - \exp(\tilde{\mathbf{x}}_i'\boldsymbol{\beta}))$.

Previous Approaches

Bedrick et al 1996

Previous example is the CMP but a special case of this is the DAP

reparameterizing beta(a_{1i}, a_{2i}) as beta$(\tilde{w}_i\tilde{y}_i, \tilde{w}_i(1-\tilde{y}_i))$ in (1). For the logistic model $f(\cdot) = F(\cdot)(1-F(\cdot))$, the induced prior (2) is in the same form as the likelihood; namely, it is proportional to a likelihood based on the "prior observations" $(\tilde{y}_i, \tilde{\mathbf{x}}_i, \tilde{w}_i\colon i = 1, \ldots, p)$.

- Beta parameters have a #successes & #failures interpretation
- Beta distribution CMP induces a DAP only for logistic regression

Can see whether individual cases in the data or in the prior are particularly influential

Provides the CMP and DAP for generalized linear model (GLM)

Not straightforward to find the hyper-parameters for DAGs in some examples

Previous Approaches

Hosack 2017

Relies on the assumption that a reasonable prior for the <u>linear</u> part (β's) of the GLM is Multivariate Normal (MVN)

At various design points:

1. What value do you believe gives a 50% chance that the true answer is lower? (This is $f_{1/2}$).
2. Assume that the true value is really **below** $f_{1/2}$. Given this information, what value do you believe gives a 50% chance of being above or below the true value? (This is $f_{1/4}$).
3. Assume that the true value is really **above** $f_{1/2}$. Given this information, what value do you believe gives a 50% chance of being above or below the true value? (This is $f_{3/4}$).

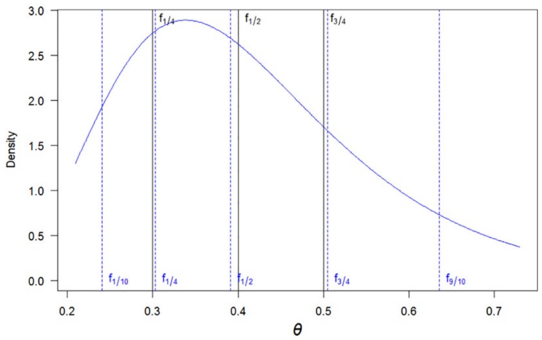


Fig. 1. Example of graphical feedback to experts for an elicited distribution. The elicited target is θ. Probability density function (blue curve) with x-axis bounded by the 90% central CI of the fitted distribution. The elicited fractiles are displayed as the black solid vertical lines, and the fitted fractiles as the blue dashed vertical lines. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

After elicitation find the parameters of the MVN distribution which minimizes the discrepancy between the predicted quantiles and the elicited quantiles (as measured by Kullback–Leibler divergence although other measures possible)

Proposed approach

- **Originally developed for survival analysis, other work in this area include (Singpurwella 1988), (Johnson 1996), (Ouwens 2018), (Cope et al 2019)**
- **Ouwens and Cope are probably most general although both have limitations**
- **Proposed approach in theory will work for any statistical model (including GLMs)**
- **Can be thought of a Penalization of the likelihood function (Frequentist) or a Generalized Bayesian update (Bissiri et al 2016)**

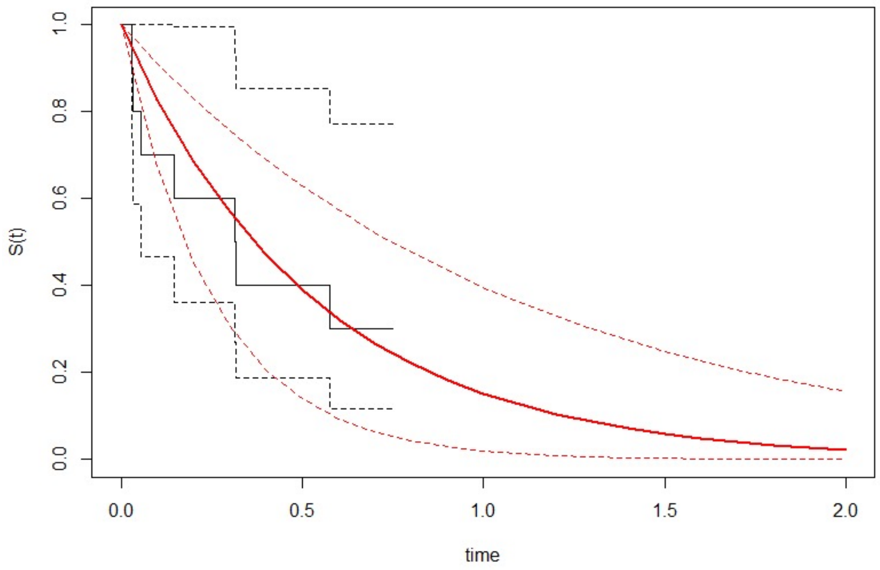
Exponential Survival Model

Simplest of all parametric survival models:

$$h(t) = \theta$$
$$S(t) = 1 - F(t) = exp(-\theta t)$$
$$LL(\theta) = D\theta - \theta T$$

Where D is the number of events and T is the cumulative time for under observation for all data.

h(t) is hazard function; S(t) is survivor function; F(t) is cumulative distribution function.



Expert Opinion

A term that penalizes the likelihood based on the discrepancy between the experts prior and the quantity predicted by the parameters

$$\pi(\theta|D,\mu_{\text{expert}},\sigma^2_{\text{expert}}) \propto L(\theta|D)\pi_{t^*}(\theta|\mu_{\text{expert}},\sigma^2_{\text{expert}})\pi(\theta),$$

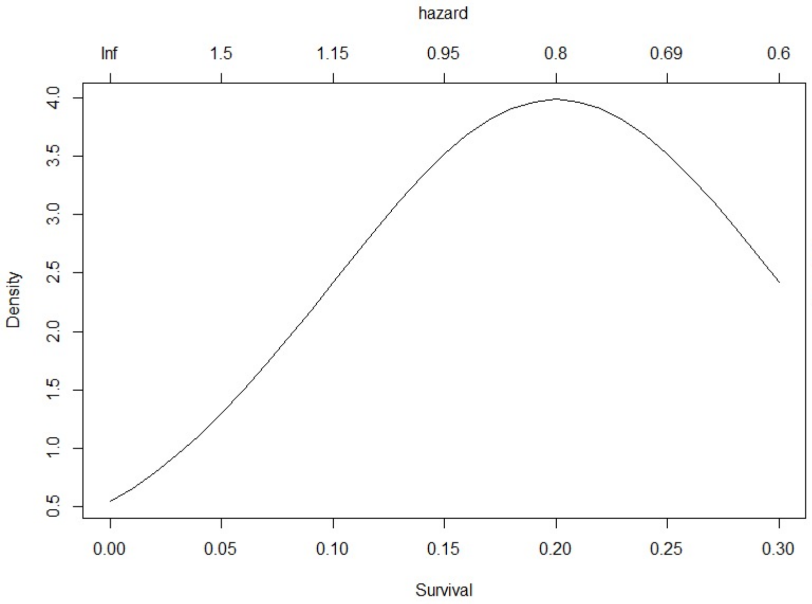
Vague prior

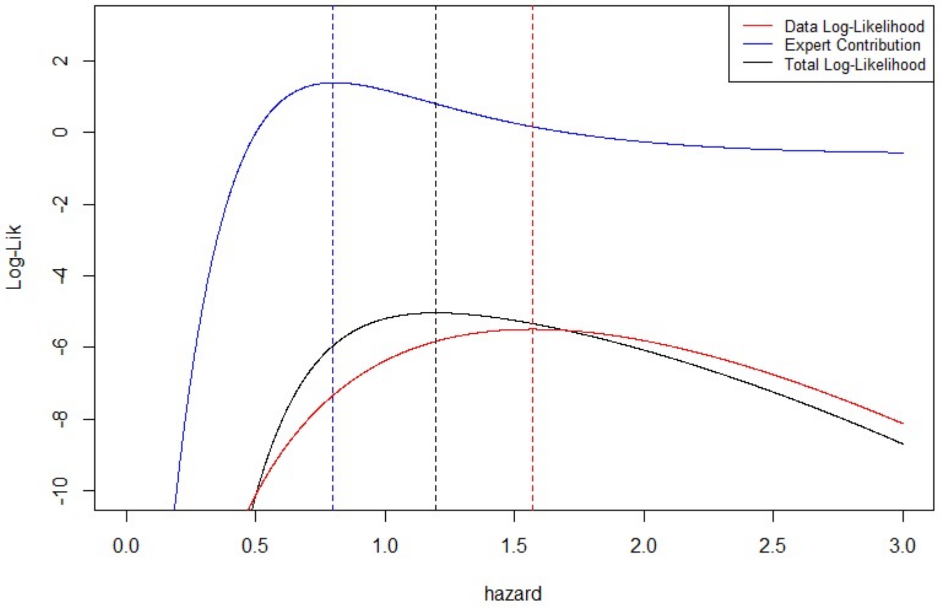
Assume an exponential model; where the expert gives the survival to be a normally with mean μ_{expert} and sd σ_{expert}; we have the following penalization term

$$\pi_{t^*}(\theta|\mu_{\text{expert}},\sigma^2_{\text{expert}}) \propto \exp\left\{-\frac{1}{2}\left(\frac{\exp(-\theta t^*)-\mu_{\text{expert}}}{\sigma_{\text{expert}}}\right)^2\right\}$$

In this example we set μ_{expert} = 0.2 and $\sigma_{expert} = 0.1$

Exponential Survival Model (2)

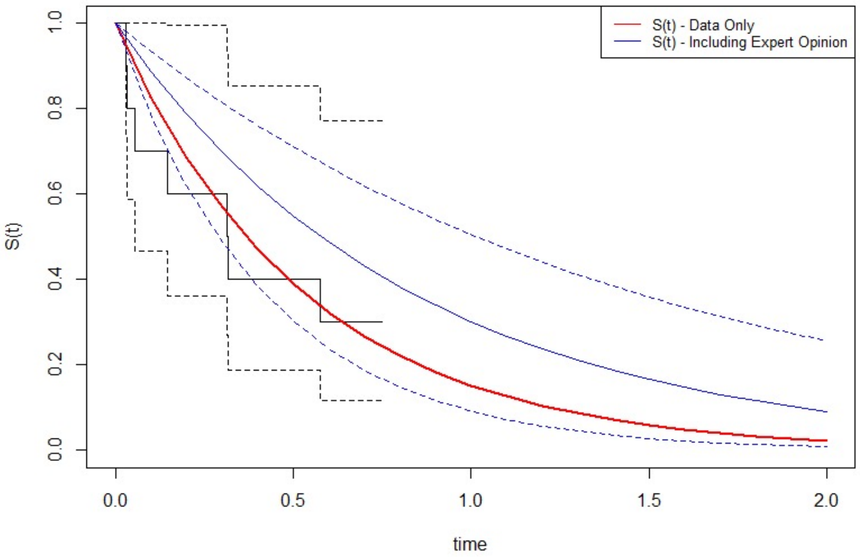




The value of the hazard least penalized by the Expert is 0.8 (blue line). The MLE (maximum likelihood estimate) for the data alone is D/T = 1.57; while the MLE considering both the data and expert is 1.2

Exponential Survival Model (3)

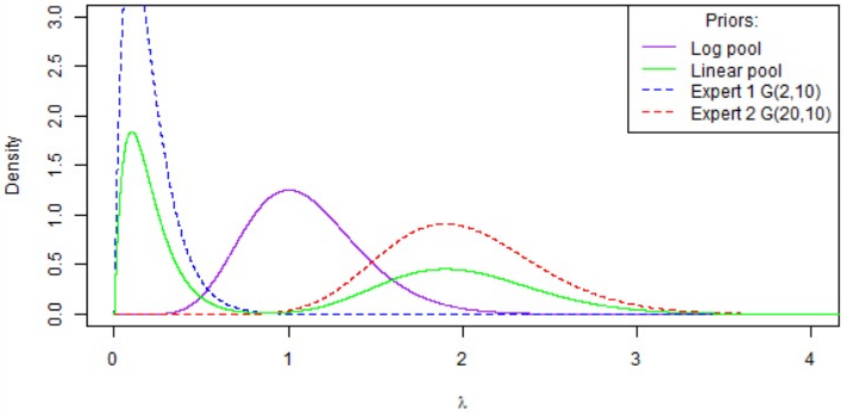
Survival Function with Expert Opinion



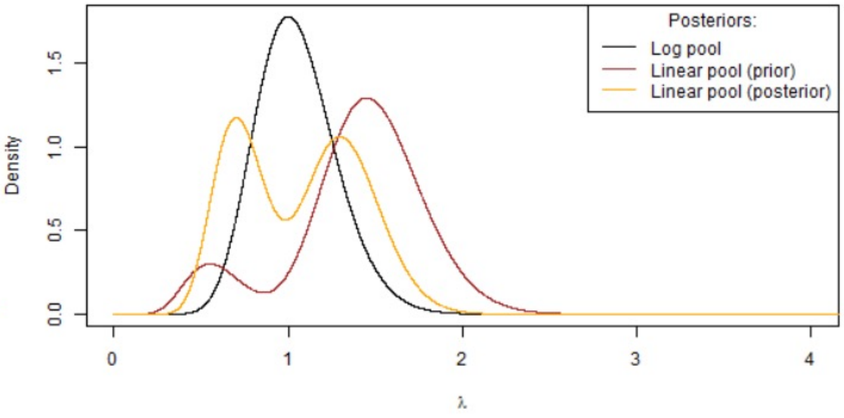
- Other more complex parametric models can be fit e.g. Weibull etc
- Covariates can also be included
- Straightforward to consider expert opinion on medians, means, mean differences, multiple time-points (although may need to space them appropriately)
- Easy to incorporate complex/aggregated prior beliefs, histogram priors etc.

Expert Opinion

Also considered expert opinion pooling and how this can be implemented



(a) Linear and logarithm pooling of opinions



(b) Posterior distributions for linear and logarithm pooling

Linear Regression

Comparison with Indirect Package

```
> summmary_carseats
# A tibble: 3 x 7
  ShelveLoc     n mean_Sales sd_Sales se_Sales CI_lower CI_upper
  <fct>     <int>      <dbl>    <dbl>    <dbl>    <dbl>    <dbl>
1 Bad          96       5.52     2.36    0.240     5.05     5.99
2 Medium      219       7.31     2.27    0.153     7.01     7.61
3 Good         85      10.2      2.50    0.271     9.68    10.7
>
```

Data

```
> as_tibble(expert_opinion)
# A tibble: 3 x 7
  ShelveLoc     n mean_Sales sd_Sales se_Sales CI_lower CI_upper
  <fct>     <dbl>      <dbl>    <dbl>    <dbl>    <dbl>    <dbl>
1 Bad          10       5.52     2.36    0.745     4.06     6.98
2 Medium       10       7.31     2.27    0.717     5.90     8.71
3 Good         10      12.2      2.50    0.791    10.7     13.8
```

We imagine Expert is wrong about the impact of good Shelve Location

Comparison with Indirect Package

Model Results

```
> lm.jags_penal
Inference for Bugs model at "4", fit using jags,
 2 chains, each with 20000 iterations (first 10000 discarded), n.thin = 2
 n.sims = 10000 iterations saved
                   mu.vect sd.vect  2.5%    25%    50%    75%  97.5%  Rhat n.eff
beta[1]              5.523   0.231 5.070  5.367  5.525  5.679  5.971 1.001  3000
beta[2]              1.783   0.279 1.235  1.593  1.784  1.976  2.324 1.002  2500
beta[3]              4.874   0.337 4.220  4.650  4.871  5.100  5.540 1.001  8700
linpred_expert[1]    5.523   0.231 5.070  5.367  5.525  5.679  5.971 1.001  3000
linpred_expert[2]    7.306   0.155 6.996  7.201  7.306  7.409  7.611 1.001 10000
linpred_expert[3]   10.397   0.242 9.923 10.236 10.398 10.559 10.878 1.001 10000
sigma                2.349   0.084 2.190  2.291  2.346  2.404  2.518 1.001  8800
```

```
> lm.jags
Inference for Bugs model at "3", fit using jags,
 2 chains, each with 10000 iterations (first 5000 discarded), n.thin = 5
 n.sims = 2000 iterations saved
        mu.vect sd.vect  2.5%   25%   50%   75% 97.5%  Rhat n.eff
beta[1]   5.522   0.233 5.052 5.366 5.521 5.685 5.962 1.001  2000
beta[2]   1.784   0.280 1.249 1.587 1.779 1.973 2.345 1.001  2000
beta[3]   4.878   0.334 4.234 4.653 4.870 5.101 5.555 1.002  2000
sigma     2.348   0.084 2.195 2.289 2.345 2.405 2.520 1.002   760
```

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Thank You

References

Bedrick, Edward J., et al. “A New Perspective on Priors for Generalized Linear Models.” Journal of the American Statistical Association, vol. 91, no. 436, 1996, pp. 1450–60, https://doi.org/10.2307/2291571. Accessed 11 Apr. 2022.

Bissiri, P. G., C. C. Holmes, and S. G. Walker. “A General Framework for Updating Belief Distributions.” Journal of the Royal Statistical Society. Series B (Statistical Methodology) 78, no. 5 (2016): 1103–30. http://www.jstor.org/stable/44682909.

Cope, S., Ayers, D., Zhang, J., Batt, K. & Jansen, J. P. (2019), ‘Integrating expert opinion with clinical trial data to extrapolate long-term survival: a case study of CAR-T therapy for children and young adults with relapsed or refractory acute lymphoblastic leukemia’, BMC Med Res Methodol 19(1), 182.

Hosack, G. R., Hayes, K. R., and Barry, S. C. (2017). “Prior elicitation for Bayesian generalised linear models with application to risk control option assessment.” Reliability Engineering & System Safety, 167: 351 – 361. Special Section: Applications of Probabilistic Graphical Models in Dependability, Diagnosis and Prognosis.

Hosack, Geoffrey R. "Indirect prior elicitation for generalised linear models with R package indirect."

References

Johnson, W. (1996), "Predictive Influence in the Lognormal Survival Model," in Prediction and Modeling in Statistics and Econometrics: Essays in Honor of Seymour Geisser, eds. J. Lee, A. Zellner, and W. Johnson, Amsterdam: Elsevier.

Kadane, J. B., & Wolfson, L. J. (1998). Experiences in Elicitation. Journal of the Royal Statistical Society. Series D (The Statistician), 47(1), 3–19. http://www.jstor.org/stable/2988424

Kadane, J. B., Dickey, J. M., Winkler, R. L., Smith, W. S., & Peters, S. C. (1980). Interactive Elicitation of Opinion for a Normal Linear Model. Journal of the American Statistical Association, 75(372), 845–854. https://doi.org/10.2307/2287171

Mikkola, Petrus, Osvaldo A. Martin, Suyog Chandramouli, Marcelo Hartmann, Oriol Abril Pla, Owen Thomas, Henri Pesonen et al. "Prior knowledge elicitation: The past, present, and future." arXiv preprint arXiv:2112.01380 (2021).

Ouwens, M. (2018), 'Use of clinical opinion in the estimation of survival extrapolation distributions'. URL: https: // www. ispor. org/ docs/ default-source/ presentations/ 91714pdf. pdf? sfvrsn= a5b2756f_ 0

Paul H Garthwaite, Joseph B Kadane & Anthony O'Hagan (2005) Statistical Methods for Eliciting Probability Distributions, Journal of the American Statistical Association, 100:470, 680-701, DOI: 10.1198/016214505000000105

Singpurwalla, Nozer D. "An Interactive PC-Based Procedure for Reliability Assessment Incorporating Expert Opinion and Survival Data." Journal