



DMC Departamento de Métodos
Cuantitativos en Economía y Gestión

Bayesian solutions for handling uncertainty in survival extrapolation

Miguel A. Negrín Hernández

Investigación realizada en colaboración con **Andrew H. Briggs** y **Julian Nam** del Health Economics and Health Technology Assessment de la Universidad de Glasgow.



- **Guión:**

1.- Motivación

2.- Soluciones propuestas en la literatura

3.- Metodología

4.- Resultados

5.- Conclusiones

1.- Motivación

Horizonte temporal observado vs. **Horizonte temporal apropiado**

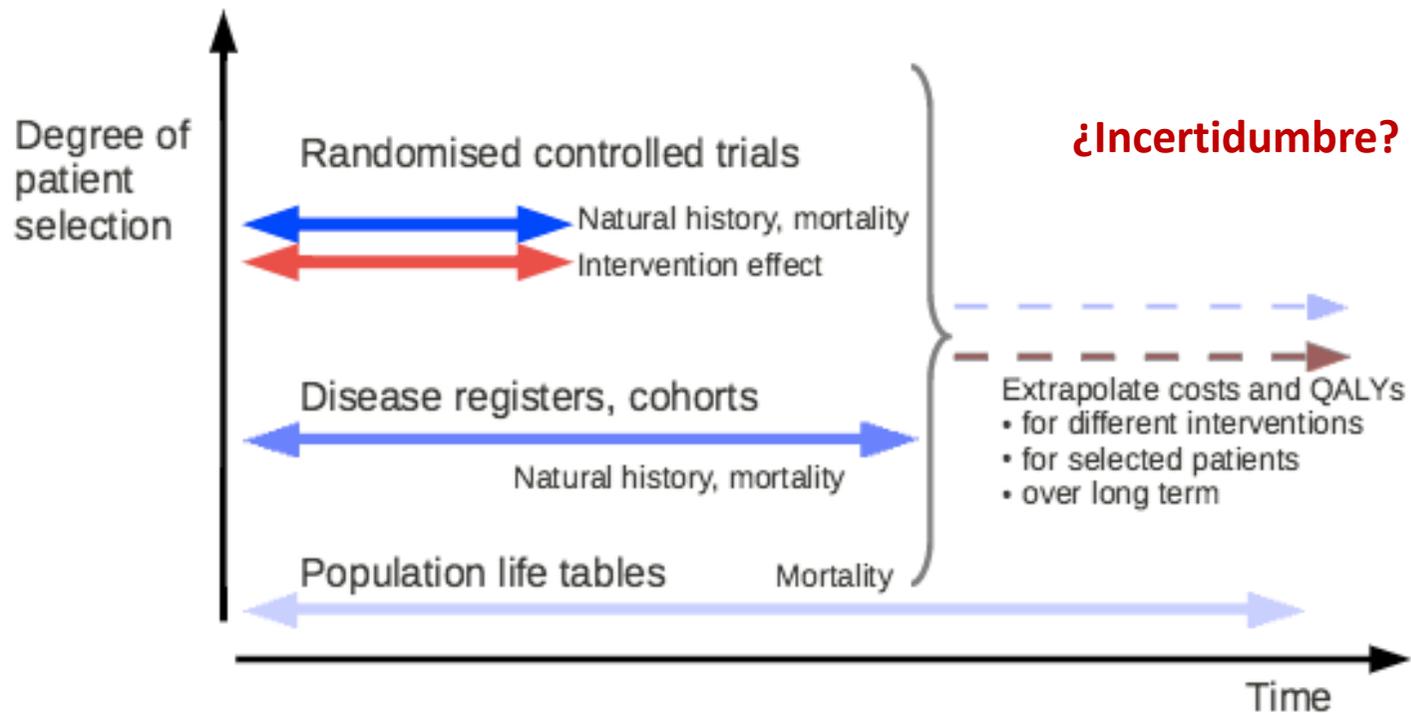
Corto-plazo

Largo plazo

(Ensayo clínicos)

1.- Motivación

En la evaluación económica de tecnologías sanitarias se necesita la extrapolación de la progresión de los resultados (clínicos y de costes) más allá de los datos observados.



1.- Motivación

Supuestos posibles en la extrapolación:

- Estabilidad temporal (Invariabilidad)

Las estimaciones obtenidas en el corto-plazo se mantienen en el largo-plazo. Utilización de recursos, QALYs, costes unitarios, probabilidades de transición entre posibles estados de la enfermedad, etc.

- Modelización

Estimar el comportamiento a largo-plazo en base a la observación a corto-plazo. Datos de mortalidad, recaídas, fallos, etc. (análisis de supervivencia)

1.- Motivación

Briggs et al. (2004) Análisis coste-efectividad de dos prótesis alternativas de reemplazo de cadera: Charnley vs. Spectron

Fuente de datos: Registro Nacional Sueco de todas las artroplastias de cadera (20.145 pacientes) realizadas en hospitales suecos. Periodo 1992-1999 (8 años).

Modelo de decisión:

Modelo de Markov con 5 estados,
horizonte temporal 60 años

Variable relevante:

Tiempo hasta el fallo

Extrapolación:

Distribución Weibull,
de riesgo proporcional

Covariables: Sexo, edad,
diagnóstico inicial

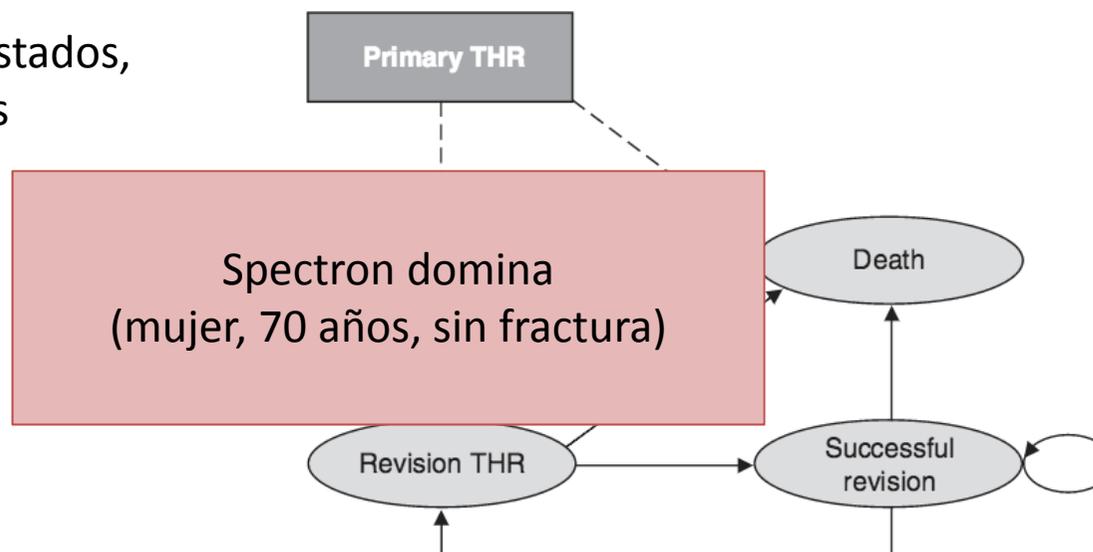


Fig. 1. Diagram showing the structure of the model. **THR** = total hip replacement.

1.- Motivación

Davies et al. (2013) The “Hazards” of Extrapolating Survival Curves. *Medical Decision Making*, 33:369-380

Repiten el análisis duplicando el número de años de seguimiento: Registro Nacional Sueco, Periodo 1992-2007 (16 años).

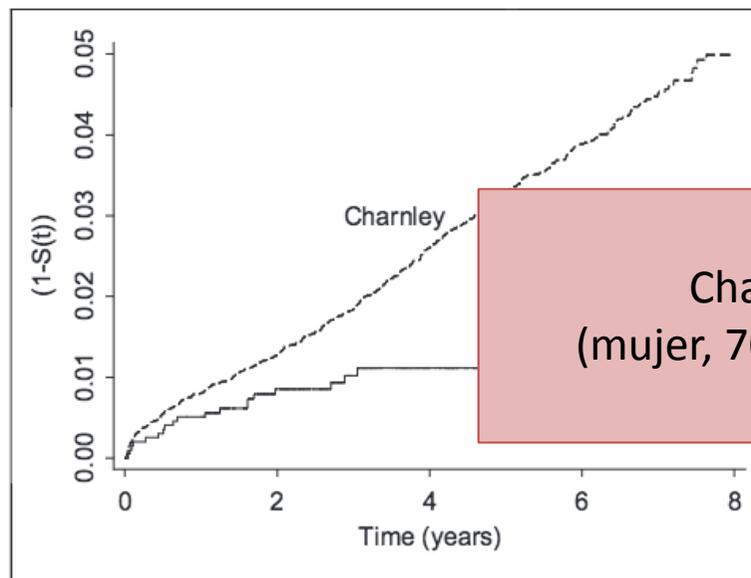


Figure 1 Kaplan-Meier failure curves for replication sample, 1992-1999 (SHORT).

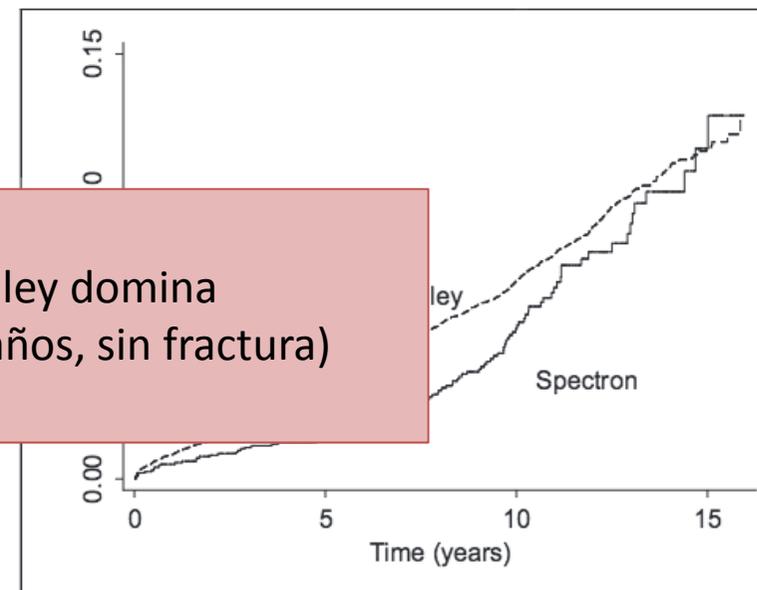


Figure 2 Kaplan-Meier failure curves for replication sample, 1992-2007 (LONG).

2.- Soluciones propuestas en la literatura

Davies et al. (2013):

- Estimar distintos modelos alternativos (exponential, Weibull, Gompertz, lognormal y log-logistic) y seleccionar el del mejor ajuste.
- Descartar el supuesto de riesgo proporcional.
- Añadir flexibilidad a la especificación del modelo (no linealidad de la edad).
- Analizar el impacto del tamaño muestral.

Conclusión: recomendar seguimientos a largo-plazo, sobre todo si el tratamiento analizado es nuevo.

2.- Soluciones propuestas en la literatura

Latimer (2013) Survival analysis for economic evaluations alongside clinical trials – extrapolation with patient-level data inconsistencies, limitations, and practical guide. *Medical Decision Making*, 33:743-754.

Revisión del análisis de supervivencia de 45 evaluaciones económicas del NICE.

- Estimar varios modelos posibles y seleccionar un modelo en base a criterios de ajuste (AIC, BIC, etc.) o criterios de validez externa (opinión de expertos, datos externos). No se detalla el procedimiento.
- Realizar un análisis de sensibilidad usando los modelos “plausibles”.

3.- Metodología

Incorporar en la toma de decisiones la incertidumbre asociado a la extrapolación utilizando los datos disponibles.

- **Combinación de modelos:**

Bayesian Model Averaging

- **Estabilidad de los parámetros:**

Escenarios optimista y pesimista

3.- Metodología. Modelos

Se consideraron 6 modelos alternativos, que incluyen modelos convencionales de 1-2 parámetros: Exponencial, Log-normal, Gamma, Weibull; y modelos más flexibles como la Gamma-Generalizada y Poly-Weibull.

Se estimaron por separado para Charnley y Spectron, por lo que el supuesto de riesgo proporcional no fue asumido.

El ajuste se midió a partir del BIC.

$$\text{BIC} = -2 \log p(D|\hat{\theta}_s, S = s) + d_s \log(n).$$

Distribuciones a priori no-informativas: Distribuciones normales con varianza alta y distribuciones Gamma(1,0)

Las estimaciones se realizaron usando OpenBUGS/WinBUGS.

3.- Metodología. Estabilidad de los parámetros

Davies et al. (2003) mostraron cómo la estimación de los parámetros de la distribución Weibull variaban notablemente al utilizar el seguimiento a 8 o 16 años.

En este apartado pretendemos incorporar la incertidumbre sobre la estabilidad de los parámetros en el modelo, a partir de la observación de la evolución de los parámetros en el corto-plazo y la utilización de escenarios optimistas-pesimista.

El procedimiento es el siguiente:

- A partir de los datos de seguimiento de 8 años creamos submuestras de 4, 5, 6 y 7 años.
- Tomaremos la estimación más “optimista” y más “pesimista” sobre la evolución del riesgo de fallo para definir los escenarios.
- Los modelos para 8 años se reestimarán utilizando como distribución a priori los escenarios optimista y pesimista.

3.- Metodología. Combinación de modelos

Para combinar los **8** modelos se emplea la metodología BMA (Raftery, 1995).

$$E[\Delta|D] = \sum_{s=1}^S \hat{\Delta}_s p(S = s|D),$$

$$\text{Var}[\Delta|D] = \sum_{s=1}^S (\text{Var}[\Delta|S = s, D] + \hat{\Delta}_s^2) p(S = s|D) - E[\Delta|D]^2,$$

Probabilidad a posteriori de los modelos

$$p(S = s|D) = \frac{p(D|S = s)p(S = s)}{\sum_{s=1}^S p(D|S = s)p(S = s)}, s = 1, 2, \dots, S$$

$$p(S = s|D) = \frac{\exp(-0.5BIC_s)p(S = s)}{\sum_{s=1}^S \exp(-0.5BIC_s)p(S = s)}$$

3.- Metodología. Combinación de modelos

BMA 1. Distribuciones a priori desinformativas para los 8 modelos considerados.

En particular se asume una probabilidad a priori de $1/3$ para los modelos optimista y pesimista, y una probabilidad a priori de $1/18$ para los otros 6 modelos.

Con información suficiente, la probabilidad a posteriori del modelo con **mejor ajuste** se aproximará a 1. Sin embargo, mejor ajuste no tiene por qué implicar mayor capacidad predictiva.

BMA 2. Distribuciones a priori informativas que asignen probabilidades a priori mayores a los modelos con peor ajuste con el objetivo de que **todos los modelos tengan la misma probabilidad a posteriori de ser ciertos.**

4.- Resultados. Modelos desinformativos

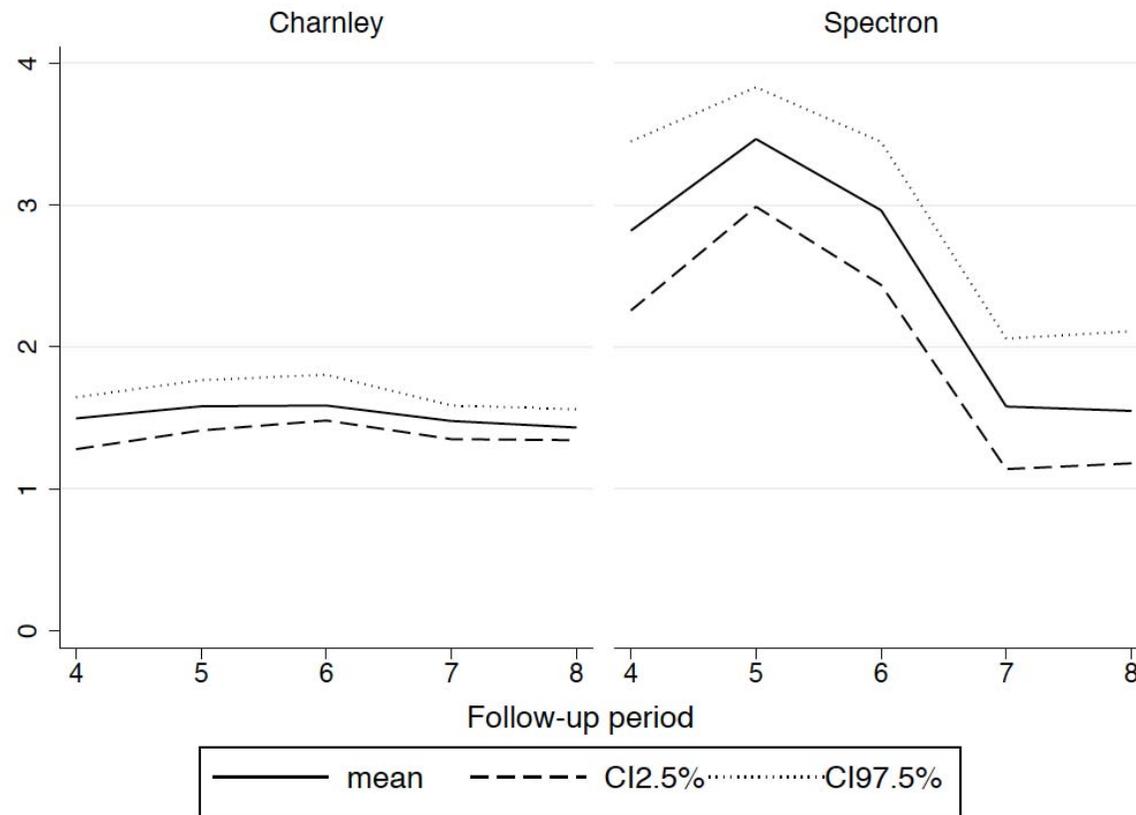
Table 1. Bayesian information criteria (BIC) for each model specification fit to the original 8-year dataset. In addition, the restricted mean survival (RMS) at 8 years and 30 years (revision-free years) with both the 8-year and 16-year datasets (mean and 95% credible interval).

	Charnley				Spectron			
	BIC ^a	RMS(8Y) ^a	RMS(8Y) ^b	Mejor predicción	BIC ^a	RMS(8Y) ^a	RMS(30Y) ^a	RMS(30Y) ^b
Exponential	39836	7.9 (7.9,7.9)	28.5 (28.4,28.8)	28.6 (28.5,28.7)	4050	8.0 (7.9,8.0)	29.5 (29.0,29.7)	28.5 (28.2,29.0)
Lognormal	39800	7.9 (7.9,7.9)	28.3 (28.0,28.6)	28.6 (28.5,28.7)	4062	8.0 (8.0,8.0)	29.6 (29.1,29.9)	27.8 (27.0,28.4)
Gamma	39836	7.9 (7.9,7.9)	27.2 (26.8,27.8)	28.5 (28.4,28.6)	4056	8.0 (8.0,8.0)	29.4 (29.1,29.7)	27.6 (26.5,28.2)
Gen-Gamma	39799	7.9 (7.9,7.9)	28.0 (27.7,28.3)	28.6 (28.4,28.7)	4071	8.0 (8.0,8.0)	29.9 (29.8,30.0)	27.5 (26.5,28.3)
Weibull	39792	7.9 (7.9,7.9)	27.7 (27.4,28.1)	28.5 (28.4,28.7)	4058	8.0 (8.0,8.0)	29.4 (28.3,29.8)	27.1 (26.1,28.0)
Poly Weibull	39842	7.9 (7.9,7.9)	27.8 (27.3,28.3)	28.6 (28.4,28.7)	4097	8.0 (8.0,8.0)	29.2 (27.7,29.9)	27.4 (26.4,28.2)
Range	39792- 39842	7.9- 7.9	27.2- 28.5	28.5- 28.6	4050- 4097	8.0- 8.0	29.2- 29.9	27.1- 28.5

^a Models with 8-year dataset, ^b Models with 16-year dataset

4.- Resultados. Inestabilidad de los parámetros

Figure 3. Estimated shape parameter (mean and 95% credible interval) for Weibull distribution for 4, 5, 6, 7 and 8-year datasets for Charnley (left panel) and Spectron (right panel).



4.- Resultados. Modelos desinformativos

Table 1. Bayesian information criteria (BIC) for each model specification fit to the original 8-year dataset. In addition, the restricted mean survival (RMS) at 8 years and 30 years (revision-free years) with both the 8-year and 16-year datasets (mean and 95% credible interval).

	Charnley				Spectron			
	BIC ^a	RMS(8Y) ^a	RMS(30Y) ^a	RMS(30Y) ^b	BIC ^a	RMS(8Y) ^a	RMS(30Y) ^a	RMS(30Y) ^b
Exponential	39836	7.9 (7.9,7.9)	28.5 (28.4,28.8)	28.6 (28.5,28.7)	4050	8.0 (7.9,8.0)	29.5 (29.0,29.7)	28.5 (28.2,29.0)
Lognormal	39800	7.9 (7.9,7.9)	28.3 (28.0,28.6)	28.6 (28.5,28.7)	4062	8.0 (8.0,8.0)	29.6 (29.1,29.9)	27.8 (27.0,28.4)
Gamma	39836	7.9 (7.9,7.9)	27.2 (26.8,27.8)	28.5 (28.4,28.6)	4056	8.0 (8.0,8.0)	29.4 (29.3,29.5)	27.6 (27.5,27.7)
Gen-Gamma	39799	7.9 (7.9,7.9)	28.0 (27.7,28.3)	28.6 (28.4,28.7)	4071	8.0 (8.0,8.0)	29.9 (29.8,30.0)	27.5 (26.5,28.3)
Weibull	39792	7.9 (7.9,7.9)	27.7 (27.4,28.1)	28.5 (28.4,28.7)	4058	8.0 (8.0,8.0)	29.4 (29.3,29.5)	27.1 (27.0,27.2)
Poly Weibull	39842	7.9 (7.9,7.9)	27.8 (27.3,28.3)	28.6 (28.4,28.7)	4097	8.0 (8.0,8.0)	29.2 (29.1,29.3)	27.4 (26.4,28.2)
Range	39792- 39842	7.9- 7.9	27.2- 28.5	28.5- 28.6	4050- 4097	8.0- 8.0	29.2- 29.9	27.1- 28.5

^a Models with 8-year dataset, ^b Models with 16-year dataset

4.- Resultados. Inestabilidad de los parámetros

Ejemplo ilustrativo: Spectron

Modelo optimista: Gamma generalizada (4 años)

$$E[\gamma|D] = 2, E[\beta|D] = 0.9$$

$$Var[\gamma|D] = 1.1^2, Var[\beta|D] = 2.2^2$$

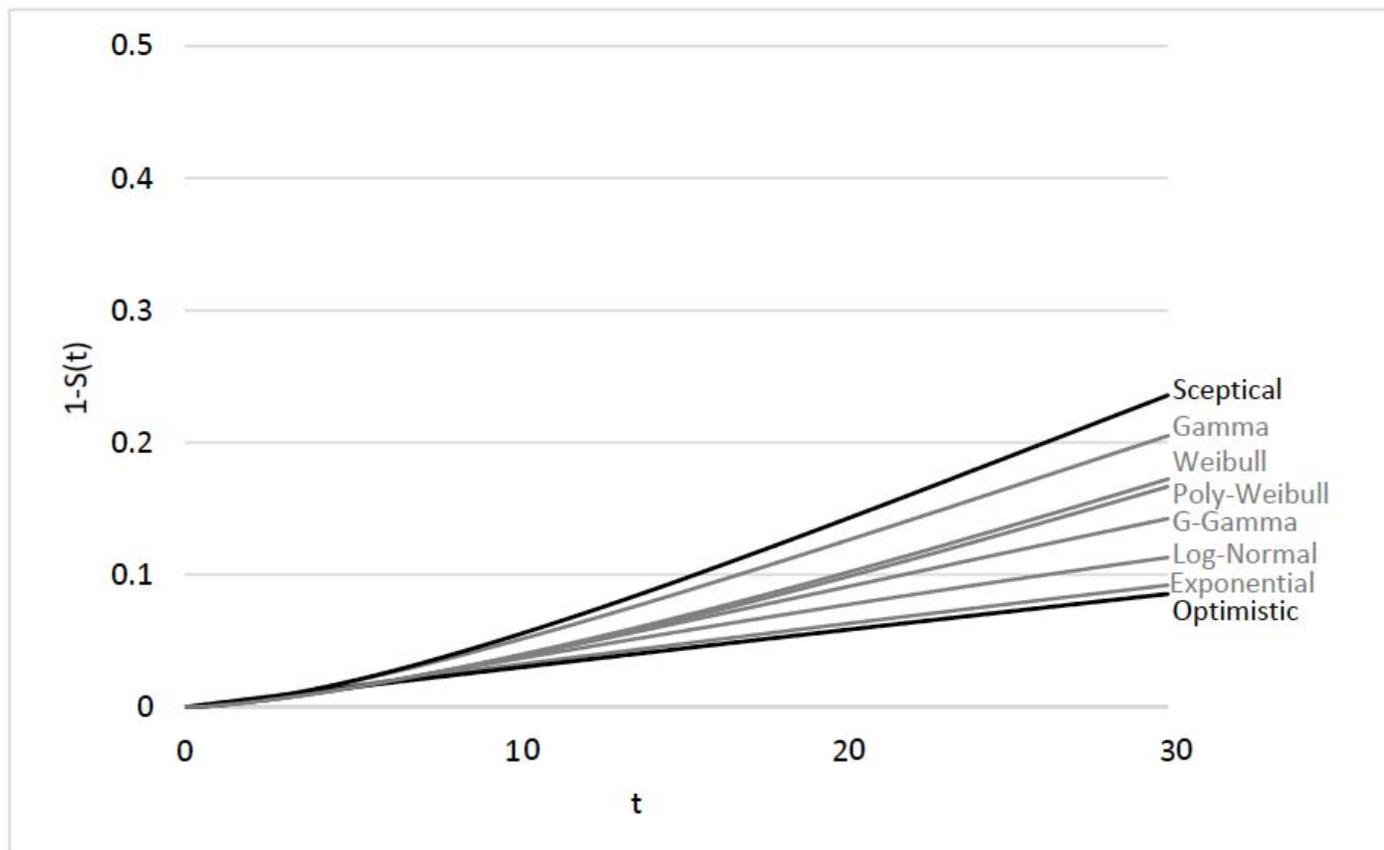
Distribución a priori: $\gamma \sim \text{Gamma}(3.1, 1.6)$ and $\beta \sim \text{Gamma}(0.18, 0.19)$

Modelo pesimista: Poly-Weibull (5 años)

Distribución a priori: $\gamma_1 \sim \text{Gamma}(2.8, 1.6)$ $\gamma_2 \sim \text{Gamma}(97.9, 27.0)$.

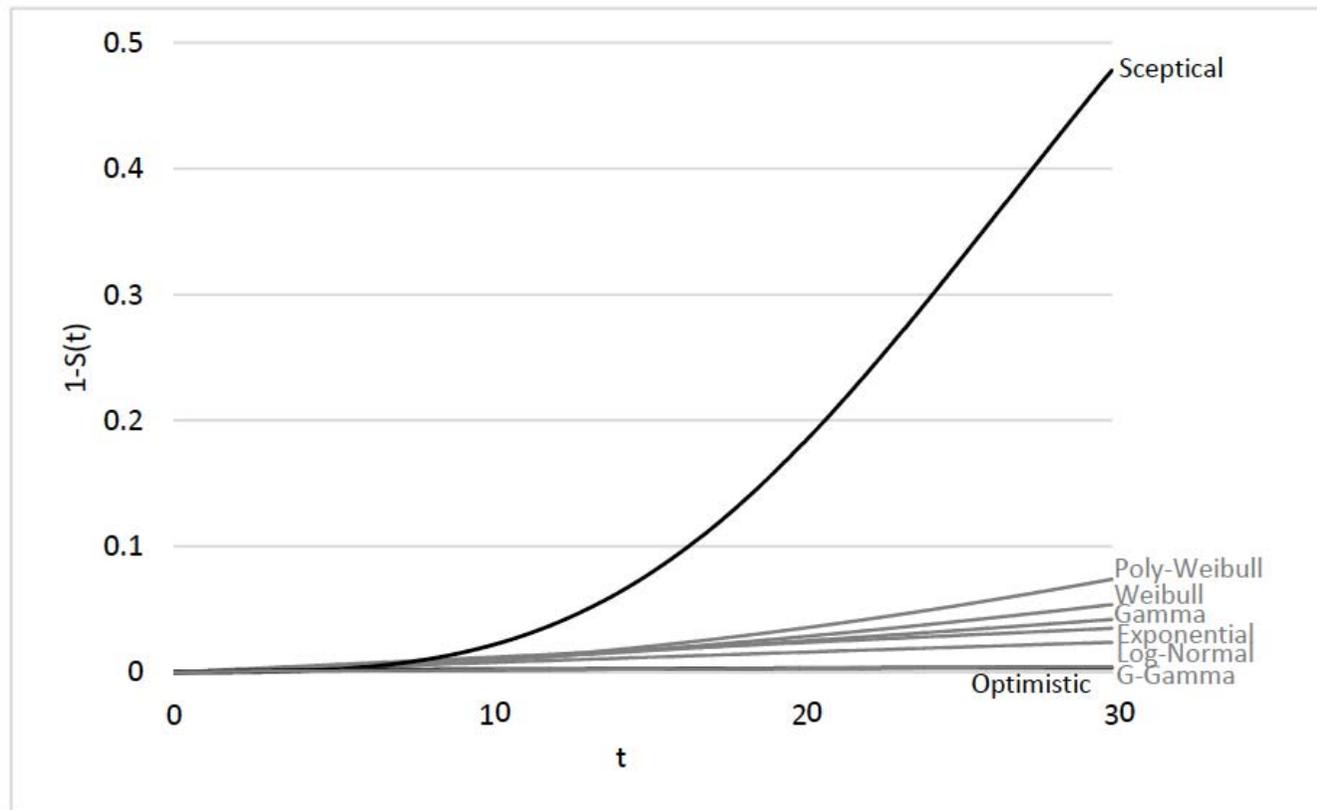
4.- Resultados. Charnley

Figure 1. Extrapolated failure curves for Charnley using different model specifications using the original 8-year dataset.



4.- Resultados. Spectron

Figure 2. Extrapolated failure curves for Spectron using different model specifications using the original 8-year dataset.



4.- Resultados. BMA

BMA1

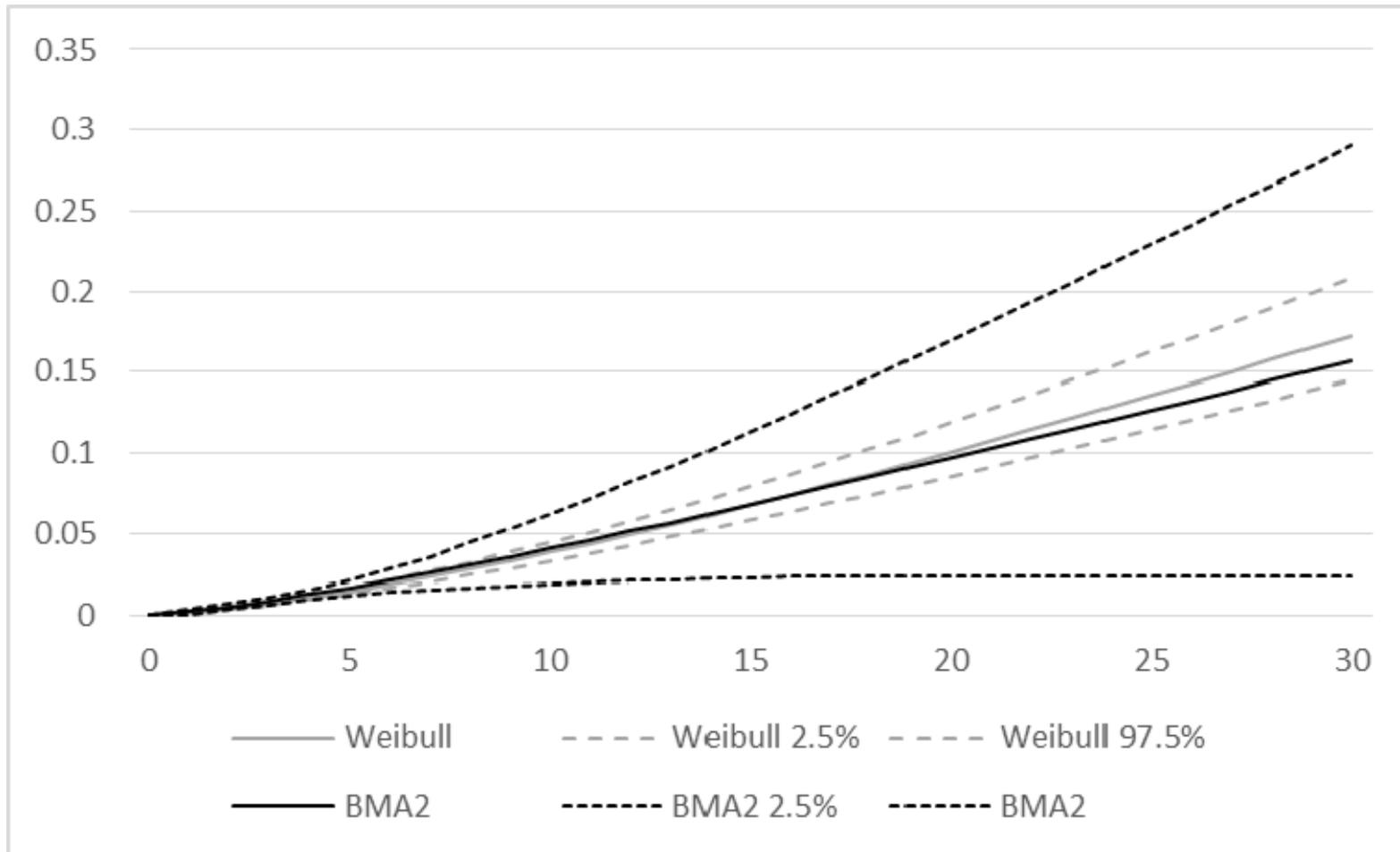
BMA no informativo, asumiendo la misma probabilidad a priori a los modelos optimista, pesimista y no-informativos (1/3).

Concluye que hay 1 modelo claramente mejor, basado en el BIC. En el caso del Charnley sería el modelo Weibull (probabilidad a posteriori del 95.5%), y en el caso del Spectron sería el modelo Exponencial (probabilidad a posteriori del 93.8%)

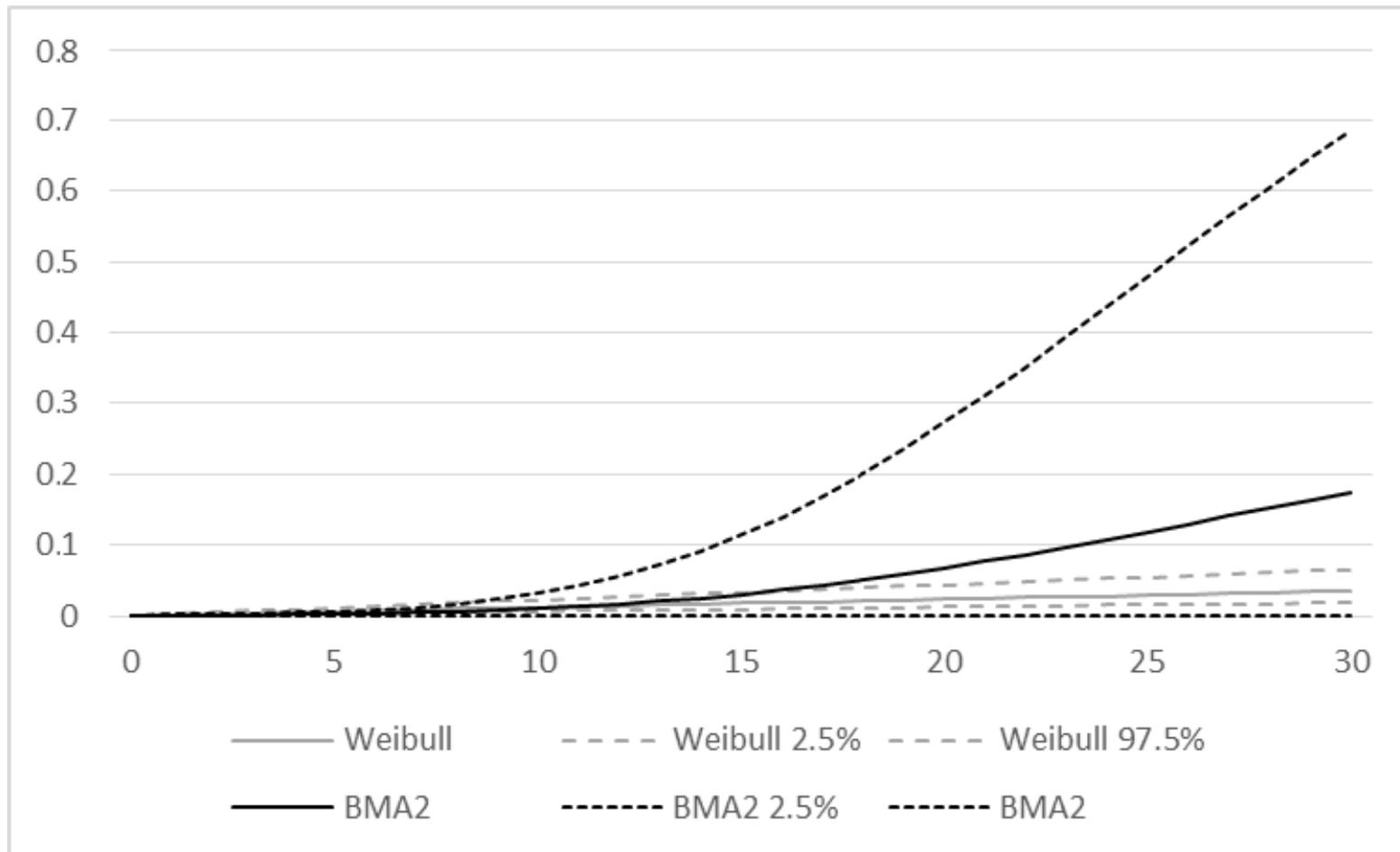
BMA2

El BMA2 asume mayores distribuciones a priori para los modelos con menor ajuste con la intención que la distribución a posteriori sea igual para todos los modelos.

4.- Resultados. BMA2 Charnley



4.- Resultados. BMA2 Spectron



4.- Resultados. Análisis coste-efectividad

Table 2. Mean results of the cost-effectiveness analysis

	Charnley		Spectron		Incremental		ICER	ProbCE*
	Costs (£)	QALYs	Costs (£)	QALYs	Costs (£)	QALYs		
8-year data								
BMA 1	686	8.49	654	8.54	-32	0.05	S dom	1.00
BMA 2	688	8.49	757	8.52	69	0.03	2974	0.93

* Using a WTP of £20,000/QALY

4.- Conclusiones y discusión

- La calidad de la extrapolación realizada dependerá del modelo estructural asumido.
- Considerar un único modelo de supervivencia entre todos los posibles ignora la incertidumbre asociada a la selección de modelos.
- Más aún cuando el criterio para seleccionar el “mejor” modelo se basa en el ajuste y no en la capacidad predictiva.
- La incertidumbre asociada a la estabilidad de los parámetros debe considerarse ya que puede ser elevada con muestras pequeñas, periodos de seguimiento cortos o con pocos eventos.
- La metodología BMA permite combinar la incertidumbre observada con la muestra a corto-plazo.
- Nuestra propuesta concluye que en el corto plazo (8 años) existe cierta incertidumbre sobre la elección de la mejor prótesis (93%).

Recomendación: Proponer revisiones con periodos de seguimientos mayores en aquellos casos en los que exista incertidumbre con los datos observados.

Limitaciones y futuras extensiones

- Únicamente se tiene en cuenta la incertidumbre observada en el corto plazo (cambios de estructura, etc.)
- BMA es sensible a los modelos paramétricos considerados para la supervivencia.
- Predecir el futuro es siempre una tarea difícil.



DMC Departamento de Métodos
Cuantitativos en Economía y Gestión

Bayesian solutions for handling uncertainty in survival extrapolation

Miguel A. Negrín Hernández