ERG commentary on Responses to the Appraisal Consultation Document (ACD): Total hip replacement and resurfacing arthroplasty for end-stage arthritis of the hip (review of technology appraisal guidance 2 and 44)

Consultee comments on the ACD, together with brief responses from the ERG where necessary, are summarised in Table 1.

Table 1. Summary table of consultee comments

	Comment	*ERG Response*
DePuy		
	The role of ODEP in the assessment of projected revision rate is important and should be referenced in section 1.1 of the guidance	No comment required. We referred to ODEP and their benchmarks in our report.
	There are significant unintended consequences when section 1.2 of the guidance explicitly recommends acquisition cost as the main driver of prosthesis selection provided the benchmark revision rate has been achieved	No comment required.
	We request that the British Orthopaedic Association and/or British Hip Society are represented at the second TAC meeting on 20th November 2013.	No comment required.
	Clarification that a hip prosthesis is not a commodity.	No comment required.
NICE THA Assessment ~ HW Comments (JRI)		
	According to the revision rates shown in table 6 and 7 in the overview and tables 67 and 68 (pages 213 and 214) in the assessment report a 10 year 5% revision rate benchmark means uncemented MoP and CoC are OK in men and women over 65 years, but not in men under 65 years and only uncemented MoP, not CoC, is OK in women under 65 years. This goes against conventional wisdom that uncemented hip implants do better in younger, more active patients and that ceramic on ceramic bearings should be used in this patient demographic because of superior wear rates.	Whether different benchmarks should be according to gender and age is an issue beyond the remit of the ERG.
	The price data provided by NHSSC significantly overstates the prices actually paid for all the THR implant categories.	The price of prosthesis was accrued from NHS supply chain and there is no evidence available of what is **the actual / accurate cost of different types of prosthesis.** The price of prosthesis obtained from NHS supply chain approximately reflected the price of prosthesis across the different manufactures. We recognise that the price of prosthesis submitted by JRI was less than from other manufactures, however the JRI submission to NICE failed to reach the tome deadline and

		ERG were unable to test them in tested as a sensitivity analysis.
	The assessment concludes that cemented CoP dominates the other THR categories in terms of clinical and cost effectiveness. However, there is no uncemented CoP category included. The "ceramic bearing" comparison is therefore between cemented CoP and uncemented CoC. In table 11, the price difference between an uncemented CoC and a cemented CoP is £1,900, whereas in reality it is in the region of £500 - £600.	No comment required.
Smith and Nephew		
	Total hip replacement and resurfacing arthroplasty prostheses are recommended as treatment options for people with end-stage arthritis of the hip only if the prosthesis has a rate (or projected rate) of less than 5% at 10 years; If more than one type of prosthesis meeting the above criteria is suitable for a patient, the prosthesis with the lowest acquisition costs should be chosen" Submitted paper Heintzberg 2013 (paper). Cost-ultity of Metal on Metal hip resurfacing compared to conventional total hip replacement in young active patients with OA	See additional documentation below this table.
Stryker		
	How will ODEP define the different categories in light of the reduction of the 10 year revision rate to 5%? Will this impact on the failure rates at other time points eg 3, 5, 7 years?	See additional documentation below this table for further information.
	The Committee agreed that there should be a 'stopping rule' – how will this be managed? There will be a number of Contracts/Tenders with lengthy timelines that do not meet this criteria.	No comment required.
	The recommendations do not refer to the causes for revision which may influence the revision rate. Also, the combination of different manufacturers product and the variety therin may influence revision rates, and affect the rate of different components. It is also unclear how products with revision rates lower than 5% will be preferenced.	No comment required.
British Orthopeadic Association		
	It is not stated what is acceptable as a source for this evidence. If it is NJR data then please see Comment on 4.1.16 below concerning data quality.	There is a lack of RCTs with UK patients with follow up beyond 1 year; the NJR is the best available data source for UK patients and is widely accepted worldwide as an important resource. The assessment group commented on the quality of the NJR.
	This recommendation is too naïve to be workable, would not necessarily lead to the lowest overall cost of arthroplasty for an institution, may reduce quality of service and ultimately patient safety.	No comment required.
	A document called "*Commissioning Guidance for Adult Hip Pain*" is being prepared using NICE approved methodology	No comment required.

	by a collaboration between the BOA, BHS, RightCare, RCS England, CSP, BSR and RCGP. The document has been through public consultation and peer review and is to be signed off imminently by NHS England and the Royal College of Surgeons. This document defines best practice in investigation and treatment of adults with hip pain and should be referred to in this section.	
	Use of hip resurfacing implants has fallen to 1% of patients undergoing arthroplasty in 2012. Most orthopaedic surgeons carrying out arthroplasty have already abandoned use of resurfacing implants. Excellent results are reported from specialist centres in young and truly active patients, particularly males with large femoral heads.	No comment required.
	There are publications refuting this finding on mortality (Hunt et al. 90-day mortality after 409 096 total hip replacements for osteoarthritis, from the National Joint Registry for England and Wales: a retrospective analysis. The Lancet, Volume 382, Issue 9898, Pages 1097 -1104, 28 September 2013).	Interesting paper Published 2013 - hence not within our inclusion criteria.
	Although data has been collected on the NJR since 2003 there were significant problems with compliance and incomplete data entry in the early years.	NJR is the best available. The assessment group commented on the quality of the NJR.
	Extrapolating revision rates beyond 9 years using NJR data is extremely problematic not only because the quality of the NJR data is unknown, but also because the mechanisms for failure of devices change with increasing follow-up; some mechanisms are unexpected and they cannot be predicted using early data (Exempli Grati: re-operation for problems with first generation polyethylene uncemented cups).	These issues were appreciated and discussed in the report. No other UK data with appreciable follow up is available.
	Existing NJR data are inadequate to predict revision rates and modes of failure into the second decade.	NJR is the best available. The assessment group commented on the quality of the NJR.
	There are very many assumptions in these models particularly on costs of follow-up/treatment of complications over the lifetime of the patient.	All published economic models of hip arthroplasty have made similar types and numbers of assumptions.
	Data quality for revision rates reported by the NJR has never been validated.	NJR is the best available. The assessment group commented on the quality of the NJR.
	The clinical outcome in terms of PROM, complication rate and revision rate and the cost of revision surgery will be different depending on what type of hip replacement is revised and the indication for revision.	Unfortunately relevant data for each of the THR categories does not exist.
	Empirical use of the list price of implants will be of almost no value in calculating the real cost to an institution of providing implants to fulfill all clinical needs.	No comment required.
	New implants to the market should now be introduced according to the Beyond Compliance initiative driven by BOA/MHRA.	No comment required.
Nick Welch - lay representative		
	Total hip replacement and resurfacing arthroplasty prostheses arerecommended as treatment options for people	No comment required.

	with end-stage arthritis of the hip only if the prosthesis has a rate (or projected rate) of revision of less than 5% at 10 years. This would appear to be contrary to NICE guidance of 'less than 10% at 10 years'. I have difficulty with 'or projected rate' of revision. Early experience with a particular device usually provides a greater failure rate in the first year or two – it would not be reasonable to extrapolate such to the 10 years expectation. Both Charnley and Exeter hips had many failures at the start of their useage. Also, this may well preclude many of the combinations (including from differing producers) preferred by the operating surgeon. Not all hybrid combinations have track records. If more than one type of prosthesis meeting the above criteria is suitable for a patient, the prosthesis with the lowest acquisition costs should be chosen. Should the surgeon, and support team's, experience with a particular device not be also taken into account. Further, it is common practice for there to be discounts available according to the number of devices ordered.	
Department of Health		
	Has no substantive comments to make	No comment required.
MHRA		
	Metal stems are available in various degrees of modularity and some text could be added to reflect this: The head of the femur (thigh bone) is replaced with either a single piece metal stem and head or a modular component consisting of a metal stem **(which may consist of more than one pieces)**, with a metal, ceramic or ceramicised metal head.	No comment required.
	The 36mm relates to the diameter of the femoral head. Therefore suggesting: THRs may also vary by femoral head size, with a large head defined as being greater than, or equal to 36 mm **in diameter.**	No comment required.
	Current conventional description of the type of the THRs is head material/cup, not the other way around. Therefore suggesting changes as follows: The average list prices for THRs across the manufacturers were: £1557 for cemented **metal-on-polyethylene**; £3016 for cementless **metal-on-polyethylene**; £3869 for cementless ceramic-on-ceramic; £2650 for hybrid **metal-on-polyethylene** and £1996 for cemented ceramic-on-polyethylene.	No comment required.
ABPI		
	We were disappointed to note the absence of any practicing surgeon members of the British Orthopaedic Association (BOA) or British Hip Society (BHS) at the TAC meeting on 18th September 2013 when the Committee made its preliminary recommendations. It is critical that recommendations be appropriately informed through professional input across the range of clinical experience across different types of fixation and technology. To that end we feel strongly that the BOA and BHS should provide the TAC with independent expert opinion and very much hope the second committee meeting will take account of this	No comment required.

	point.	
	The ABHI believe NICE's preliminary recommendations do not account for the importance of surgeon training and experience when choosing types of prostheses. The provision of training, education and support by the manufacturer is an important component of the decision to select a particular prosthesis to ensure long-term clinical outcomes for the patient and we do not consider that sufficient emphasis has been placed on this. We request that the Committee carefully consider this important aspect as context against which section 1.2 must be balanced when making its final recommendations to ensure the outcome is pragmatic and relevant to the NHS.	No comment required.
	Section 1.2 of the draft guidance currently recommends use of the prosthesis with the lowest acquisition cost. We believe that whilst this is one important factor for the NHS, patient characteristics, surgeon training and experience with a particular prosthesis and non-implant procedure costs (e.g. consumables and theatre time) should be taken into account in the guidance.	No comment required.

In the following sections we respond to more substantive issues raised by Smith & Nephew and others.

Smith & Nephew

Smith & Nephew discuss the following aspects of the ERG analyses which they consider to be flaws:

[1] Cost effectiveness of the Birmingham Hip Resurfacing (BH) prosthesis versus THR: Smith & Nephew contend that the Canadian cost utility analysis published by Heinzbergen et al 2013 demonstrates that the BH is cost effective relative to THR;
[2] Brands; Smith & Nephew contend that prosthesis brands, in particular the BH prosthesis, should have been taken into account.

Below the ERG consider these points and present some analyses of the rates of revision experienced by NJR recipients of the BH, comparing these with analysis of BH revisions recently published by the designer of the BH (McMinn et al. BMJ 2012;344:e3319) .

Cost effectiveness

Our search of relevant cost effectiveness literature had to be terminated at a fixed date in order to meet timelines of the project. The paper by Heintzbergen et al 2013 fell outside our search time line. This study is not comparable with the ERG analysis because it compares a single brand of RS with generic THR which was not stratified by bearing surface class or by brand. In short there is an inconsistency of approach in the Heintzbergen study: thus we note that Heintzbergen et al 2013

compares the brand "Birmingham Hip" (because it was the most commonly used RS) with conventional generic THR, rather than with different THR brands, or with. In this respect the title of the paper (*Cost-Utility of Metal-on-Metal Hip Resurfacing Compared to Conventional Total Hip Replacement in Young Active Patients with Osteoarthritis*) is misleading.

For revision rates the authors used Australian Registry data but weighted this according to frequency of usage in Alberta. The authors only report year 1 revision (as a transition probability); nowhere in the paper is revision data beyond year one available; there is no comparison indicating if Alberta revision experience is concordant with that in the Australian registry.

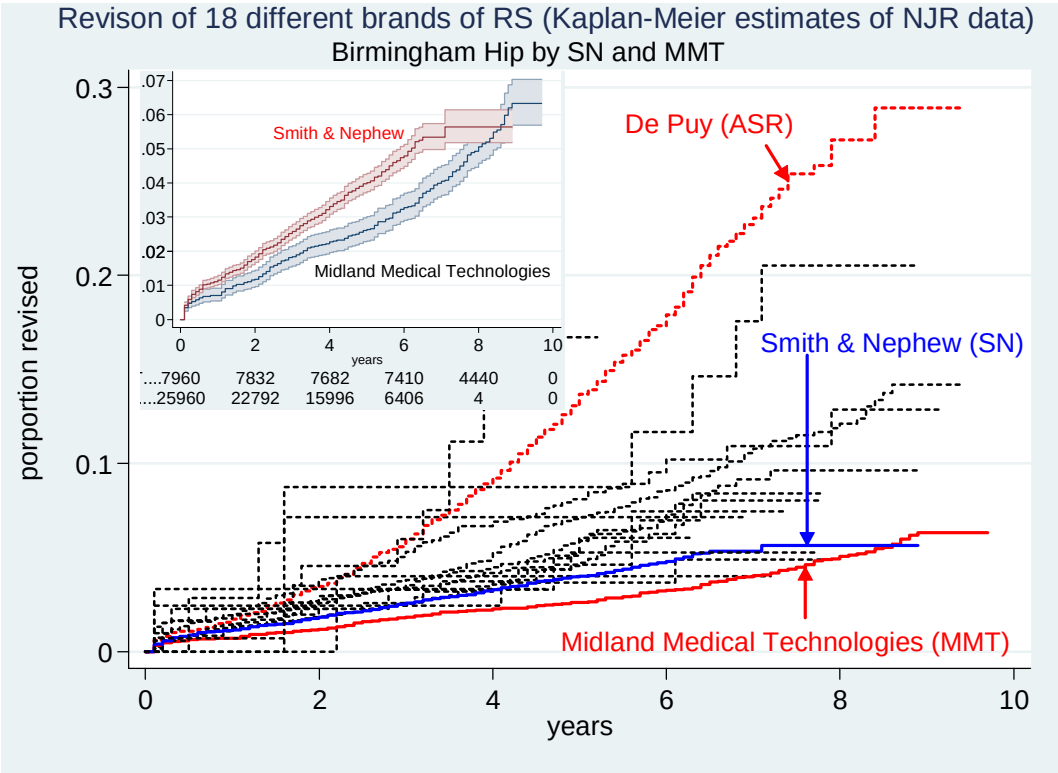
The authors found that there was a 58% probability that RS with BH was cost effective relative to THR at a willingness to pay of Can $50,000/QALY. For women of various ages THR was dominant to RS and the absolute differences between compared groups for men were very small; this led the authors to conclude that: "*THA is still a reasonable option if the patient or clinician prefers, given the small absolute differences between the options and the confidence ellipses around the cost-effectiveness estimates.*"

Brands

Comparing different brands was beyond the remit for this MTA; if different brands of RS and THR prostheses were to be compared the number of different brands would exceed 40; no investigators have attempted this, including the cost utility study of Heinzbergen et al. itemised by S & N and the recent analysis of the BH published in BMJ by McMinn et al. (BMJ 2012;344:e3319).

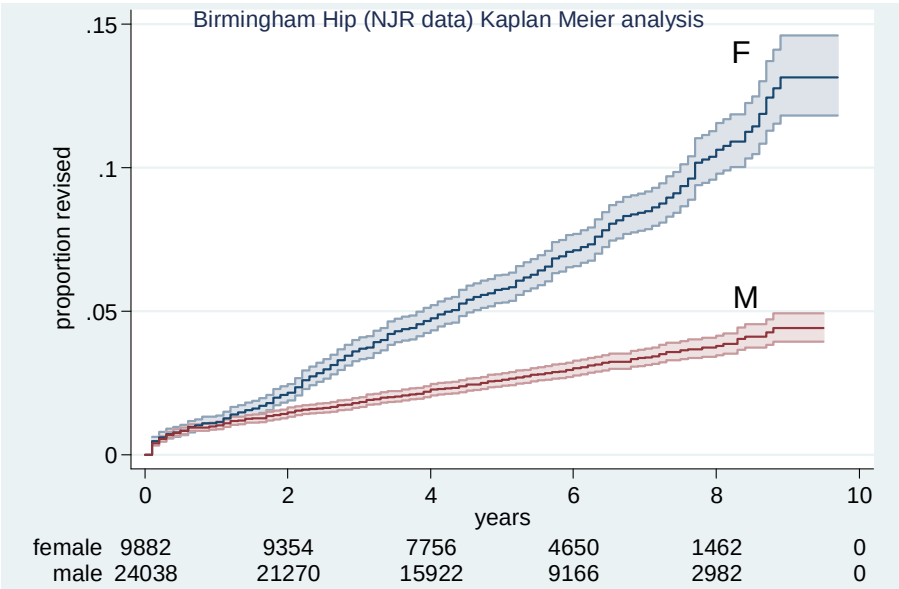
Smith & Nephew point to the superior revision performance of the BH brand relative to that of other brands of RS prosthesis. ERG can confirm that NJR recipients of the BH do indeed experience lower revision rates than recipients of nearly all other RS brands (Figure 1). The NJR contains two data sets for the BH, namely Smith & Nephew and MMT sets. Revision for the two BH sets differed somewhat (see Figure 1).

Figure 1. Revision for different brands of RS prosthesis



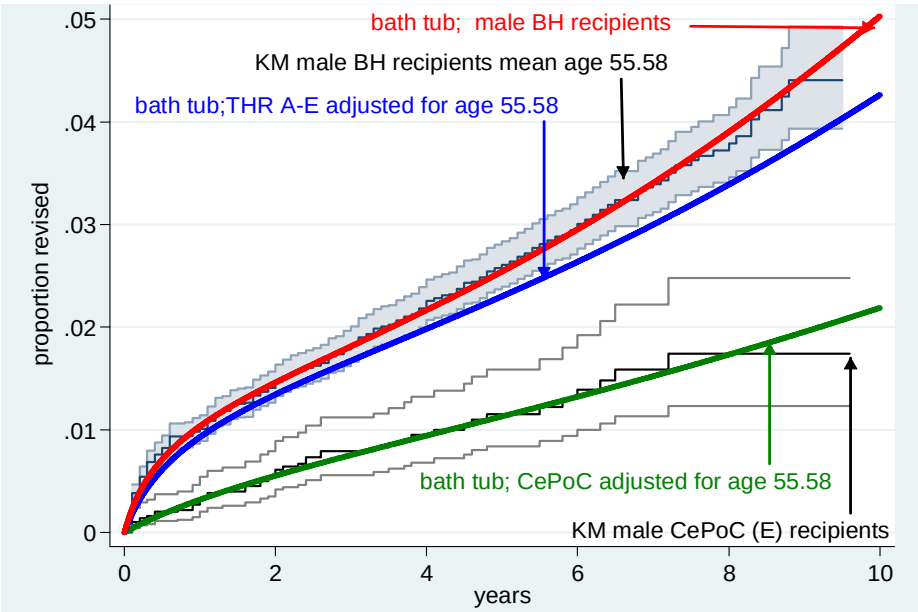
When revision experienced by NJR recipients of the BH is compared between men and women very different rates are observed (Figure 2). By about 9 years the revision rate for women approaches about 14%. This gender difference is similar to that for RS as a generic class. McMinn et al. 2012 did not report the revision rate for female recipients of the BH. It is apparent that for females the BH RS prosthesis is probably unlikely to represent a good option relative to THR (for which revision rates for women are generally well below 5% at 10 years).

Figure 2. Revision for NJR BH recipients according to gender



The mean age for male NJR BH recipients was 55.58 years. The bath tub model of the KM analysis for this group is shown in Figure 3 and predicts a revision rate at 10 years of 5%. Bath tub models for male NJR THR categories A- E and category CePoC (E) adjusted for age of 55.6 years are also shown in Figure 3.

Figure 3. Bath tub modelled revision for the BH and recipients of THR categories A-E and category CePoC (E)



The bath tub model predicts a 10 year revision rate of 4.3% category A-E recipients and 2.2% for the best performing THR category (CePoC (E)) compared with the higher rate of 5% for the BH, the best performing RS brand.

It is pertinent to point out that in the recently published analysis of revision rates for the Birmingham Hip (BMJ 2012;344:e3319) this brand was compared not with THR brands but with two generic forms of THRs, that is with "cemented" and "uncemented", **NOT** with different brands amongst these generic forms. These investigators concluded that "*Both uncemented total hip replacements and Birmingham hip resurfacings had a higher revision rate than cemented total hip replacements*" This is wholly in line with the current ERG analyses. Due to lack of data none of these analyses were able to adjust for patient activity levels of the prosthesis recipients. A potential advantage of RS has been the assertion that for active patients the post recovery activity levels are superior after RS relative to THR. However the recent RCT of RS versus THR for young active UK patients found no between group difference in Paffenbarger activity (mJ/week) levels at 12 months post-surgery (effect size 1.17 (95% CI: −4.99 to 7.32).

In summary

It is evident that Smith & Nephew find it appropriate to compare the best performing brand of RS with generic forms of THR, both in the analysis of revision rates and in estimates of cost effectiveness. Such comparisons do not seem equitable; if "best brands" of RS are to be compared with THR then best performing brands of THR are the appropriate comparator(s).

CONSULTEE COMMENTS REGARDING BENCHMARK

The focus of the remaining consultees' comments is on implementation of a benchmark revision rate. Two main areas of concern can be identified: how should revision rate be analysed and how valid is a linear interpolation from a 10-year benchmark?

In the following section the ERG provide a brief overview regarding these concerns.

Revision rate analysis

The NJR and manufacturers in general do not have available follow up revision data to 10 years. Therefore some modelling and extrapolation is required to establish a 10 year revision rate. There are two main analytical procedures that may be adopted namely Kaplan Meier (KM) and competing risk (CR) analysis. Each has been variously used recently in studies analysing revision rates. Competing risk analysis of revision was undertaken in the 9th NJR annual report (2012), while in 10th annual report (2013) the analyses reverted to the Kaplan-Meier approach.

Using NJR data for patients who received THR categories A to E the ERG found that the best parametric fits for extrapolation to 10 years for KM analyses were bath tub, lognormal and Weibull distributions while for CR analyses best fits were provided by lognormal and Weibull models. The modelled 10 year revision rates using these methods for each of the five THR categories are summarised in Table 2.

Table 2. Estimated cumulative incidence of revision at 10 years according to method of analysis

Category of THR	**Models based on Kaplan Meier analyses**			**Models based on competing risks analyses**	
	Bath tub	*Weibull*	*Lognormal*	*Weibull*	*Lognormal*
A-E	3.23	2.72	2.56	2.47	2.42
CePoM (A)	2.84	2.42	2.27	2.13	2.09
CeLPoM (B)	3.91	3.48	3.29	3.16	3.10
CeLCoC (C)	4.65	3.96	3.67	3.72	3.62
HyPoM (D)	3.03	2.51	2.37	2.26	2.21
CePoC (E)	2.10	1.97	1.96	1.83	1.79

Only one estimate exceeded 4% revision at 10 years. The differences in cumulative incidences of revision for any individual THR category, according to different models based on KM analyses, was small, reaching a maximum of 1%. Similarly Weibull and lognormal models for any particular THR type based on CR analyses differed only slightly (by <0.5%).

The NJR populations which received the five categories of prosthesis differed in terms of mean age and mix of male and female patients. The mean age across the five categories was close to 70 years and 35% were male. When the models were adjusted for age 70 years and gender mix 35% male the 10 year predicted revision rates were as shown in Table 3 and very similar to the unadjusted predicted rates.

Table 3. Modelled 10 year revision rates for a population 70 years old and 35% male

Category of THR	**Models based on Kaplan Meier analyses**			**Models based on competing risks analyses**	
	Bathtub	*Weibull*	*Lognormal*	*Weibull*	*Lognormal*
A-E	3.22	2.73	2.58	2.55	2.50
CePoM (A)	3.03	2.61	2.44	2.43	2.38
CeLPoM (B)	3.97	3.42	3.24	3.14	3.08
CeLCoC (C)	4.35	3.65	3.45	3.44	3.35
HyPoM (D)	3.97	2.57	2.42	2.41	2.36
CePoC (E)	1.96	1.83	1.74	1.71	1.67

Linear interpolation from 10 year revision rate

Because revision accumulate more rapidly in the first year or so post-surgery a linear interpolation of the 10 modelled revision rate may provide a more demanding benchmark than the 10 year benchmark. The shape of the modelled revision rates shown in Figure 4 illustrate that for CR based models the modelled 5 year rate is greater than a 5 year rate predicted by linear interpolation from the modelled 10 year rate; for the bathtub models based on KM analysis this is not always the case.

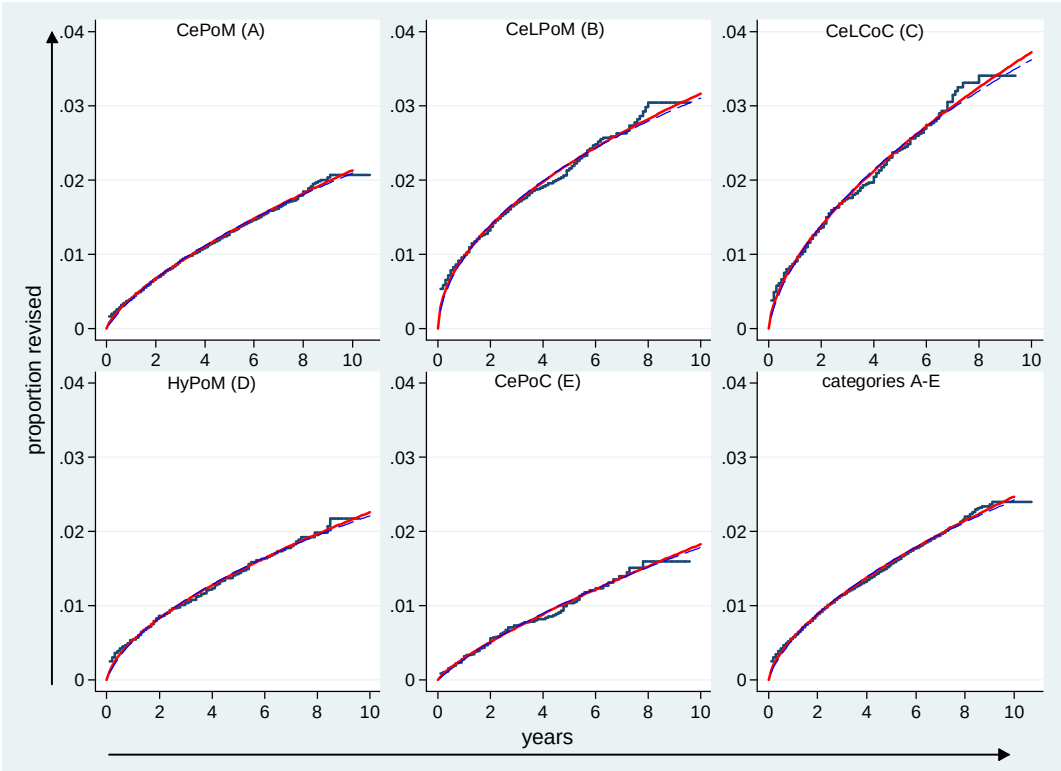
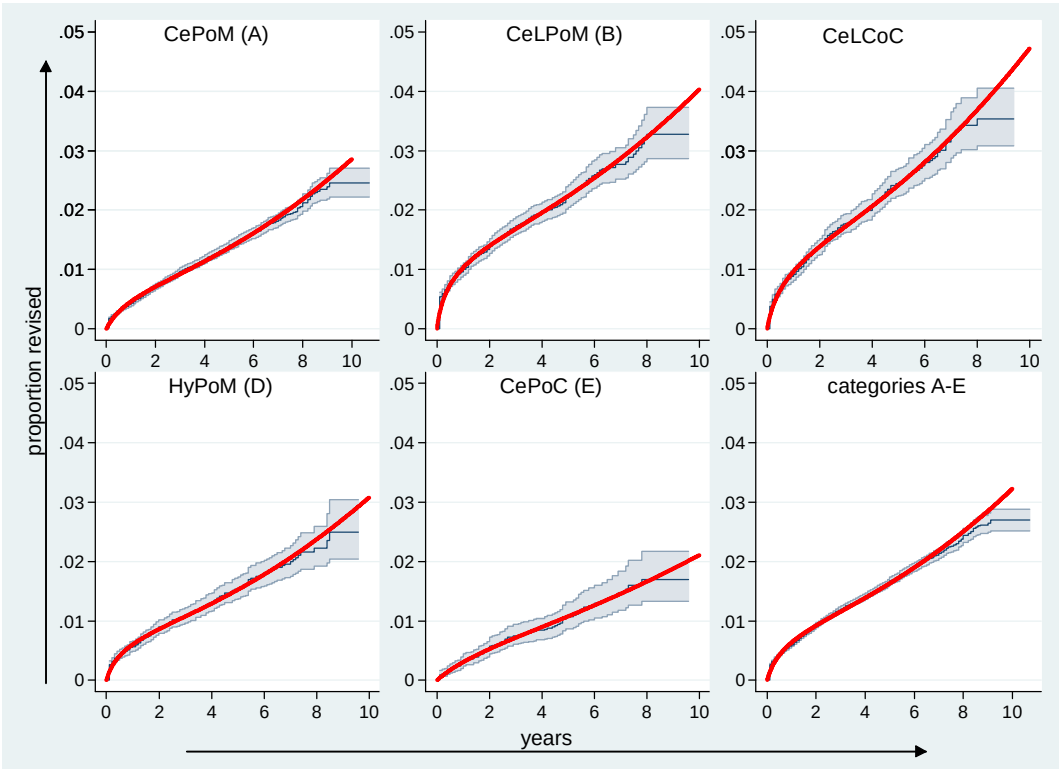



Figure 4. Upper panel: Weibull (solid line) & lognormal (dashed line) fits to CR analysis (stepped line) of time to revision; lower panel: unadjusted bath tub parametric fit (smooth line) to KM estimates of time to revision