Mark Chadwick proposal:

“Would the following calculations be of valuable to further understand the issues associated with having a calibrated evaluated file (ie tweaked, as we do now in endf7 and endf8 and in Jeff, Jendl), versus keeping such files uncalibrated to the extent possible?

What I would like to understand is the extent to which it matters, based on how it impacts an adjusted file’s subsequent performance.

So consider two starting evaluates files, say
ENDF-1. (calibrated), and
ENDF-2 (uncalibrated)
(And note these 2 files would be largely the same except say for different nubar values ; their covariance data would be the same)

Run each of these through your adjustment Bayesian approach, based on a suite of integral criticals. These make output files, let’s call them:
Adjusted-1
adjusted-2

Then, assess the extent to which each of these adjusted 1,2 files differ regarding their different predictions for other applications. (I know the definition of "other applications" will perhaps matter, depending on how similar the application is to the criticals that were used in the adjustment process).

If they perform differently in terms of predicted criticality and reaction rates for application problems, then this would highlight the importance of the issue;

if instead they perform in essentially the same way, then it would suggest that calibration is not of practical negative importance for those who make and use adjusted libraries. (Whilst calibration has obvious benefits for those who don't!).

Is this a calculation worth trying out?

Some answers/comments already, e.g. P.Talou:

The so-called "calibration" or "lucky draw" or "library optimization" or whatever it's called does not account for those post-evaluation correlations. While the mean values are adjusted, there is no "record" of this adjustment in the evaluation process.

I would again point to Kawano et al's paper (NSE 153, 1, 2006) where we played with a toy calculation, adjusting the Pu239 fission cross section and covariance to reproduce Jezebel k_eff within the reported 0.2% experimental uncertainty on the integral data. The result was that the cross section almost did not change, since k_eff was already reproduced by the prior cross section, but the posterior covariance now includes negative correlations, as expected.

This is the main reason why this "calibration" process should not have much of an impact on your post-evaluation adjustment process, *as long as* the evaluated covariance is not modified at the evaluation stage.

Otherwise, it would mean double-counting.

This is the approach proposed recently by Rochman and Bauge. It is somewhat appealing but also confusing, since now part of our knowledge coming from integral data is used *explicitly* in the evaluation process. To avoid double-counting, you would now have to be very careful about which integral data have been used and which have not.