

**Response to questions sent on behalf of
the Environment, Food and Rural Affairs Select Committee, [8 November 2007]**

*F.J. Bourne, C.A. Donnelly, D.R. Cox, G. Gettinby, J.P. McInerney,
W.I. Morrison & R. Woodroffe*
Former members, Independent Scientific Group on Cattle TB

1. *Sir David King makes specific criticism of the ISG's interpretation of confidence intervals for the detrimental effect of proactive badger culling outside the specified area. Specifically, King's paper notes that several of these confidence intervals extend below zero, and are fairly wide (para 43) - with the implication that they should be treated with caution. Can you explain the ISG's thoughts on King's analysis in greater detail?*

We discuss this in detail in our response to Sir David King's report. Here is the relevant section of this response:

(i) Interpretation of statistical confidence intervals in subgroup analysis

ISG13 King *et al.* (2007) place great weight on interpretation of results presented in Figure 5.2B of our final report (Bourne *et al.*, 2007). This figure shows the results of a subgroup analysis, stratifying overall RBCT results into beneficial and detrimental effects recorded at different distances from the boundaries of proactive trial areas. As is often the case in scientific studies, each effect was presented as a point estimate, associated with 95% confidence limits. The latter give a measure of the uncertainty associated with the point estimate; there will be substantial uncertainty when small sample size prevents a precise estimate from being obtained from a subset of the data.

ISG14 In paragraph 43 of King *et al.* (2007), the detrimental effect of badger culling on cattle TB incidence outside culling area boundaries is dismissed since “*Three out of four [confidence intervals] go through zero (i.e. one cannot be confident that the overall effect is detrimental)*”. This interpretation is incorrect, as it omits one very crucial proviso in the interpretation of our subgroup analysis. The limits attached to any subgroup of data concern what can be learned from that subgroup of data on its own. While it is useful to know this, it is rarely the primary focus of analysis, and was not so in this case. Indeed if the data are broken into a large number of small subgroups, each on its own will have substantial uncertainty and be indecisive on any issue of concern, even if the overall picture is entirely clear on the point under study, as was the case here. The overall picture must be studied, not fragments of it. This is what we have done in all our analyses, subject to tests of the uniformity of the effect under study.

ISG15 In the interests of consistency, we note that the same analysis showed that all five of the subgroup estimates of beneficial effects of culling inside proactive areas had confidence intervals which included zero. Thus, were the same (erroneous) interpretation to be placed on those findings, evidence of the beneficial effects of culling should likewise have been dismissed.

2. *Table 5.10 of the ISG's Final Report shows that the point estimates from the initial cull and the first follow-up culls are (29.3% and 27.3%) are both significant at the 5% level (0.015 and 0.044 respectively). What were the null and alternative hypotheses in these cases?*

The null (H_0) and alternative (H_A) hypotheses were in each case:

H_0 : There was **no** effect of proactive badger culling on the incidence of all (confirmed and unconfirmed) TB breakdowns among cattle herds up to 2km outside RBCT trial areas.

H_A : There was an effect of proactive badger culling on the incidence of all (confirmed and unconfirmed) TB breakdowns among cattle herds up to 2km outside RBCT trial areas.

[Note that the alternative hypothesis is two-sided. In other words, the null hypothesis would be rejected in light of sufficient evidence of either an effect which increased incidence or an effect which decreased incidence.]

So the conclusions are that there was evidence of a detrimental effect of proactive culling on the incidence of all TB breakdowns among cattle herds up to 2km outside RBCT trial areas both from the initial cull ($p=0.015$) and from the first follow-up cull ($p=0.044$), based on RBCT location data (Table 5.10, page 101 of the Final Report). Similarly, there was evidence of a detrimental effect of proactive culling on the incidence of confirmed TB breakdowns among cattle herds up to 2km outside RBCT trial areas both from the initial cull ($p=0.016$) and between the initial and first follow-up culls ($p=0.021$), based on RBCT location data (Table 5.8, page 98 of the Final Report).

It is worth noting that the estimates based on all (confirmed and unconfirmed) breakdowns were (in 5 of the 6 analyses) smaller in magnitude than those obtained in analyses of confirmed TB breakdowns only (compare Table 5.8, page 98 and Table 5.10, page 101 of the ISG Final Report). This was similarly the case within RBCT trial areas (compare Table 5.3, page 92 and Table 5.5, page 95 of the ISG Final Report). While our analyses revealed consistent and statistically significant effects of badger culling on the incidence of confirmed cattle herd breakdowns, there were no similar effects on unconfirmed breakdowns. For this reason we considered it appropriate, from a biological perspective, to focus primarily on effects on confirmed breakdowns, rather than all breakdowns. However, disruptions and costs result from both confirmed and unconfirmed breakdowns, although unconfirmed breakdowns are typically shorter in duration. Both confirmed and unconfirmed breakdowns result in the compulsory slaughter of reactor cattle, movement restrictions on the herd, and additional testing of cattle. Hence, from an economic point of view preventing unconfirmed breakdowns would be desirable, whether or not they indicate the presence of disease.

We would also wish to draw the Committee's attention to further analyses of the incidence of confirmed breakdowns, presented in paragraph A20 of our response to Sir David King's report.

3. *Appendix D of the ISG's report (p205) provides summary data on triplets recruited to the RBCT. Some of the sample sizes recorded (for total badgers culled, culled badgers with TB) look to be fairly small once the total number of culls is taken into account. Does this affect the robustness of the RBCT statistical results and the conclusions that can be drawn from them?*

The numbers of badgers culled per year under the proactive and reactive strategies are detailed in Tables 2.4 and 2.6 (pages 49 and 50), respectively. Regarding the numbers of badgers culled in reactive areas, fewer badgers were taken in reactive areas due to the more limited geographic scale of the reactive culling strategy, compared with the repeated

widespread culling in proactive trial areas. Furthermore, the suspension of reactive culling in November 2003 meant that reactive culling was undertaken for fewer triplet years than proactive culling.

The statistical precision of the RBCT results concerning TB incidence in cattle herds was a function of:

- a. the per-annum incidence levels in trial areas,
- b. the number of years of observation,
- c. the number of triplets (10) and
- d. the consistency of culling effects between the triplets.

The precision was not a function of the numbers of badgers removed.

The number of badgers taken from a trial area on a particular cull is a function of the underlying density of badgers within that area and the proportion of badgers removed. Underlying density will be influenced by the suitability of the environment for badgers, suppression of the population by previous culls (both before and during the RBCT), immigration into the trial area from outside, and the natural growth rate of the population due to reproduction and survival of badgers. The proportion of badgers removed is influenced by the trapping method used, as well as by season, weather conditions, land access, disruption and occupation of traps by non-target species; however our published analyses (Donnelly *et al.* 2007; Woodroffe *et al.* 2007) detected no evidence of an effect of the last three factors on overall capture rates.

We discussed the impact of culling efficiency in paragraphs 10.10-10.14 (pages 164-5 of the ISG Final Report), and concluded that “improvements in culling efficiency – if implemented in isolation from other changes – are unlikely to generate benefits [in terms of reducing TB incidence in cattle herds] greater than those recorded in the RBCT”. This is primarily because, in the absence of geographical barriers to badger movement, improvements in trapping success would be largely offset by badger immigration into culled areas. We went on to discuss, in paragraphs 10.15-10.24 (pages 165-167 of the ISG Final Report), alternative approaches based on proactive culling and concluded that there was, at present, no practicable culling method likely to generate benefits substantially greater than those achieved by RBCT proactive culling.

While the ‘*sample size*’ of badgers did not influence the precision of analyses concerning the incidence of TB in cattle herds, these sample sizes did directly affect the precision of estimates obtained in analyses concerning the prevalence of *M. bovis* infection in badgers. While the sample sizes were large enough for us to detect various important effects (for example, both proactive and reactive culling were associated with elevated *M. bovis* infection prevalence in badgers), it is of course very possible that additional effects would have been found to be statistically significant had sample sizes (i.e. numbers of culled badgers) been substantially larger. This is, of course, always the case in such a study.

4. *Medical research trials sometimes use significance tests at 1%. Can the ISG's reasons for testing at the 5% level of significance be further explained?*

In assessing the evidence in our data we did not use any specific significance level but reported the precise p values achieved (rather than “ $p < 0.05$ ”, for example). One quite widely used interpretation is that if p is roughly 0.05 there is quite strong evidence of an effect beyond the play of chance and if p is roughly 0.01 very strong evidence and so on. But we

did not set hard and fast rigid rules for the interpretation of isolated effects. (This is a widely used approach for interpretation, to be contrasted with rules of hard and fast decision making.) A brief reference discussing these issues is given in Cox and Snell (1981) with a more extended account written for non-statisticians presented in Cox (1982). Furthermore, while p values are a way of reporting the results of statistical tests, they do not define the practical importance of the results.

Although initial power calculations were based on the likelihood of detecting a significant effect (this in turn depending on the assumed significance threshold), later discussions focussed on the precision with which any effect could be estimated (paragraph 2.18, page 37 of the ISG Final Report). We typically reported both the confidence interval and the p-value associated with estimates, where the confidence interval quantifies the precision of that estimate.

5. *If economic costs were excluded from your analysis or calculated in a way that might lead to reduced expense, would the Committee be correct to conclude that the ISG might have reached a different recommendation concerning the desirability of culling?*

Our analyses indicate that modest overall benefits could be obtained by culling badgers in a simultaneous and coordinated fashion across large areas of the countryside, and repeating such culls over several years. However, our experience of overseeing such culling in the course of the RBCT – though over only ten 100km² areas – leads us to very seriously doubt that culling could be performed in the manner that would be required, particularly following dissolution of the Defra Wildlife Unit. Our results provide consistent and robust evidence that failing to meet any of the necessary criteria – that is, if culls were conducted asynchronously, or in a patchy manner, or not repeated – their effect would be to increase, rather than reduce, the incidence of cattle TB, and to contribute to its geographical spread. This was our primary reason for recommending against a badger culling policy.

These conclusions were not dependent on the economic assessments included in Chapter 9, and would not have been changed by any plausible adjustments to the cost figures used there. The economic appraisal simply reinforced our view that culling could not contribute meaningfully to control policy, but introduced the added perspective that, to be sustainable and defensible, measures also needed to have some economic justification. The evidence showed that culling along the lines of the RBCT could not meet this requirement.

In considering other culling methods it was clear their impact on the incidence of cattle TB would need to be incomparably greater than that achieved by cage trapping if they were to approach economic justification. For example, licensed culling by farmers using snares (by far the least expensive means of culling) would need to be 10 times more effective at reducing cattle TB incidence than was RBCT culling (para 9.24). Our discussion of culling options in Chapter 10 of our final report led us to conclude that such enhanced performance is highly improbable. While we did not formally discuss shooting or gassing setts using tractor exhaust, which would undoubtedly be less costly, we consider it unlikely that these could be conducted in the widespread, simultaneous, coordinated and repeated manner that would be needed to achieve benefits. Hence the use of such methods would be likely to increase, rather than reduce, the incidence of cattle TB. Moreover, we are concerned that such practices would not meet necessary welfare concerns. Individual farmers, or small groups of farmers, might feel inclined to try them in pursuit of their own personal benefit, but analysis suggests even so that their financial gains may be illusory while the effect of badger social group

perturbation implies potentially severe financial costs to their neighbours.

References

- Bourne, J., Donnelly, C.A., Cox, D.R., Gettinby, G., McInerney, J.P., Morrison, W.I., & Woodroffe, R. (2007) Bovine TB: the scientific evidence Defra www.defra.gov.uk/animalh/tb/isg/pdf/final_report.pdf, London.
- Cox, D.R. (1982). Statistical significance tests. *British Journal of Clinical Pharmacology*, 14, 325-331.
- Cox, D.R. and Snell, E.J. (1981). *Applied statistics*. London: Chapman and Hall, pp37-39.
- Donnelly, C.A., Wei, G., Johnston, W.T., Cox, D.R., Woodroffe, R., Bourne, F.J., Cheeseman, C.L., Clifton-Hadley, R.S., Gettinby, G., Gilks, P., Jenkins, H.E., Le Fevre, A.M., McInerney, J.P., & Morrison, W.I. (2007) Impacts of widespread badger culling on cattle tuberculosis: concluding analyses from a large-scale field trial. *International Journal of Infectious Disease*, 11, 300-308.
- King, D., Roper, T.J., Young, D., Woolhouse, M.E.J., Collins, D.A., & Wood, P. (2007) Tuberculosis in cattle and badgers: a report by the Chief Scientific Adviser, Sir David King <http://www.dius.gov.uk/publications/Expert%20advice%20on%20badgers%20and%20cattle%20TB.pdf>, London.
- Woodroffe, R., Gilks, P., Johnston, W.T., Le Fevre, A.M., Cox, D.R., Donnelly, C.A., Bourne, F.J., Cheeseman, C.L., Gettinby, G., McInerney, J.P., & Morrison, W.I. (2007) Effects of culling on badger abundance: implications for tuberculosis control. *Journal of Zoology*, doi:10.1111/j.1469-7998.2007.00353.x.