Epidemiol. Infect. (2012), **140**, 643–654. © Cambridge University Press 2011
doi:10.1017/S0950268811001257

Eliciting expert opinion on the effectiveness and practicality of interventions in the farm and rural environment to reduce human exposure to *Escherichia coli* O157

P. CROSS[1]*, D. RIGBY[2] AND G. EDWARDS-JONES[1]

[1] *School of the Environment, Natural Resources and Geography, College of Natural Sciences, Bangor University, Gwynedd, UK*
[2] *Economics, School of Social Sciences, University of Manchester, UK*

(*Accepted 8 June 2011; first published online 7 July 2011*)

SUMMARY

Few hard data are available on emergent diseases. However, the need to mitigate and manage emergent diseases has prompted the use of various expert consultation and opinion elicitation methods. We adapted best-worst scaling (BWS) to elicit experts' assessment of the relative practicality and effectiveness of measures to reduce human exposure to *E. coli* O157. Cattle vaccination was considered the most effective and hand-washing was considered the most practical measure. BWS proved a powerful tool for expert elicitation as it breaks down a cognitively burdensome process into simple, repeated, tasks. In addition, statistical analysis of the resulting data provides a scaled set of scores for the measures, rather than just a ranking. The use of two criteria (practicality and effectiveness) within the BWS process allows the identification of subsets of measures judged as potentially performing well on both criteria, and conversely those judged to be neither effective nor practical.

Key words: Best-worst scaling, conjoint analysis, epidemiology, *E. coli* O157, expert elicitation, interventions, pathogens.

INTRODUCTION

Newly emergent pathogens in the environment pose important challenges to public health policy. Since 1980, 87 new human pathogen species have been discovered, many of which are associated with animal reservoirs [1, 2] such as *Campylobacter*, *Escherichia coli* O157 and *Yersinia* [3]. Emerging pathogen risks are, by definition, associated with incomplete evidence bases and there is a recurring demand for the scientific community to inform policy-makers about risk management even though systematic evidence is often limited [4].

In the absence of a systematic evidence base, alternative approaches for the management of livestock diseases have been sought. These approaches seek to assist the management of uncertainty [4–6]. Often this involves making recommendations on the best available data, while acknowledging the uncertainty involved in the evidence and hence in the resulting recommendations. Expert opinion is a source of information often sought in this context, and identifying expert consensus regarding risks and the appropriate means of their management represents an attractive option in many situations. However, the appropriate means by which experts are identified and the process by which their opinions are elicited remains contested. In addition, even if the views of the correct experts for consultation are identified, there may be no, or only partial, consensus among them. This is not

---

* Author for correspondence: Dr P. Cross, School of the Environment, Natural Resources and Geography, College of Natural Sciences, Bangor University, Gwynedd LL57 2UW, UK.
(Email: afs202@bangor.ac.uk)

surprising as the extent of consensus within a scientific community may be related to the level of uncertainty regarding the issue at hand and such uncertainty is endemic to the management of emergent pathogen risks.

A number of methods have been used to elicit expert opinion regarding pathogenic risk. Some studies have employed standard questionnaire approaches, for example, asking expert panels to score a range of measures using Likert scales or similar response formats [7, 8]. Other studies have asked an expert panel to weight the relative importance of risk factors [5, 8]. A number of studies have also employed market research techniques such as conjoint analysis [6, 9] in which respondents complete structured survey tasks (making choices, ranking, etc.), generated via an experimental design, in order to explore trade-offs between candidate measures as a means of prioritizing them (see [6, 10, 11]). Conjoint analysis originates in market research where it is used to derive estimates of individuals' preferences for a product or service, and/or its component characteristics [12]. There are various forms of conjoint analysis [adaptive conjoint, choice-based conjoint (CBC), etc.] but they have in common the understanding that a product comprises a number of attributes the importance of which differs across consumers [13].

The application of these techniques to disease management has involved, *inter alia*, identifying the relative effectiveness of potential management measures and how they might be bundled into effective interventions. Cross *et al.* [10], for example, use adaptive conjoint analysis (ACA) to explore the effectiveness of trade-offs between interventions proposed for the control of bluetongue. The ACA approach used by Cross and colleagues is well suited to situations with a small number of control measures, often differing in their levels, which would typically be bundled together by practitioners. It is less suitable when there is a very large set of stand-alone control measures under consideration. In this situation a form of conjoint analysis, termed best-worst scaling (BWS) may offer more potential [14].

We apply the technique to investigate the management of *E. coli* O157 in agriculture and the wider rural environment. *E. coli* O157 was reported as a human pathogen in 1982 [15] and a number of outbreaks have subsequently occurred, typically associated with foodborne infection [16, 17]. More recent outbreaks have shown environmental contamination to be an increasingly common pathogen transmission pathway [18–20]. These outbreaks, alongside evidence on shedding levels, persistence in the environment [21, 22] and re-colonization [23] have highlighted the need to focus on environmental aspects of mitigation [24].

Given increased calls in the UK for action on the management of *E. coli* O157 risk, there is a need to evaluate potential management measures in farm and rural settings. This need for action has become more urgent in light of recent outbreaks such as that at Godstone farm in Surrey in 2009 when 93 people became infected, with 17 developing complications and eight requiring kidney dialysis [25]. Hence there is an imperative to understand, and potentially reconcile, expert opinion on the potential of a range of interventions to reduce the risk of *E. coli* O157 exposure to humans in an environmental setting.

This study investigates control measures applicable in the farm and rural environment. The scope of the farm control measures was limited to pre-lairage for livestock and the farm gate for vegetables. The orientation on the farm and rural environment is one of the novel aspects of the research since many other studies [26, 27] have focused on risks of exposure and associated control measures occurring later in the food chain, often concerning food processing, storage and handling rather than direct environmental exposure. The considered measures affect both direct environmental human exposure (e.g. via camping, petting farms, etc.) and exposure via food (reducing shedding levels of livestock entering the food chain). The focus on the farm and rural environment served to both address a knowledge gap and to keep the range of control measures considered manageable.

This paper describes and critically reflects upon the use of a novel technique (BWS) to elicit expert opinion regarding the 'effectiveness' and 'practicality' of measures to manage *E. coli* O157 risk in the farm and rural environment. It has a methodological focus in terms of outlining and testing a novel approach to expert elicitation. It has a substantive focus also, concerning the relative practicality and effectiveness of O157 control measures, and the potential complementarities and conflicts regarding their performance using these two criteria.

METHODOLOGY

BWS

BWS is a choice-based technique that requires respondents to make repeated choices between sets of options [28]. We first explain the method in general before indicating how it is applied to pathogen management in this study.

In a BWS study respondents are presented with repeated, varying, sets of (typically four or five) items and asked to identify at the extremes of their preferences. For example in a set of four control measures (A–D), the respondent indicates which they consider to be the 'most effective' and 'least effective' measures. If A is more effective than D then information is obtained for five of the six possible pairs of measures within the set. It is known that A is viewed as more effective than B, C and D (three pairs) and that B and C are more effective than D (two pairs). The only pair on which no information is available is B–C. Similarly, for a 5-item set we gain information on 18/25 paired combinations. The process is repeated with new sets (generated via an experimental design) and in each case choices are made at the extreme using whatever criteria is specified by the researcher. The data are analysed by counting procedures or by estimating choice models to derive importance weights across the items featured in the sets.

The BWS method is typically used when information is sought over a large set of items [29]. It holds a number of advantages over other rating and ranking techniques [30]. First, approaches that use some form of scaling (e.g. scoring an item on a scale of 1–5) assume that the distance properties on the scale are equal. Where scaling properties are unknown the possibility of transforming these into parametric data is not reliable [31]. Second, scales can be interpreted differently between individuals. For instance, one respondent may score an item as a '4' while another scores it a '5' when both respondents rate the item equally but have different conceptions of the value of the scores on the scale. Third, respondents are not necessarily compelled to choose one option over another, and frequently they score all of the items with equal importance [32]. This relates to another problem common to rating/ranking approaches, namely the lack of discriminatory power between items [31]. Finally, interpretation of the scores is comparatively simple as the number of times an item is selected as best, minus the number of times an item is selected as worst, approximates the true scale value [28, 29].

In this study we use the approach in a novel manner. We define the sets in terms of *E. coli* O157 control measures and use two choice criteria with those sets: effectiveness (most/least effective) and, in a second stage, practicality (most/least practical to implement). The resulting data are then analysed and the perceptions of the interventions' performance in terms of both practicality and effectiveness assessed. We now turn to the research process before setting out the analytical models employed and the results derived from them.

Identification of O157 control measures and expert sample

Intervention inclusion

Measures to reduce human exposure to *E. coli* O157 relating to the farm and rural environment were identified from peer-reviewed papers and grey literature. The scope of the farming system considered was pre-lairage for livestock and the farm gate for vegetables. The initial list comprised 99 interventions which was deemed excessive for a cognitively bearable BWS exercise and hence an initial expert consultation was undertaken to reduce the set to a more manageable size.

Expert group recruitment

The composition of the expert panel was contingent upon the scope of the interventions selected for evaluation. The panel reflected the study's focus on interventions applicable in the rural and farm environment.

Experts were identified via authorship of relevant publications and from peer networks. A snowball technique was adopted whereby invited experts were also encouraged to suggest other members in their field of expertise who might be willing to participate in the study. Such an approach taps into pre-existing professional expert networks and allows the researchers to distance themselves from the expert selection process [8]. Follow-up contact was made by phone and e-mail to encourage participation. Experts were invited from several different disciplines and academic sectors including public health; veterinary science; food microbiology; agricultural/environmental microbiology; clinical microbiology; land management and pathogen management; risk assessment; communicable disease epidemiology; molecular ecology (details of the expert sample can be seen in Table 1).

Expert elicitation round 1: shortlisting

The first round of expert consultation required experts to decide which of the 99 interventions should be

Table 1. *Categories of respondent expertise invited to participate and their response rates*

	Round 1		Round 2	
Category of expertise	Invited	Responded	Invited	Responded
Public health	19	8	21	10
Veterinary	9	6	15	9
Microbiology (food)	8	4	11	6
Microbiology (agricultural/environmental)	3	2	8	6
Microbiology (clinical)	2	1	7	1
Risk assessment	1	1	0	0
Business	2	1	2	1
Land management	9	8	4	1
Molecular ecology			2	1
Anonymous responses				6
Total	**53**	**31**	**70**	**41**

Table 2. *Example of a 'practicality' best-worst scaling choice set*

Most practical		Least practical
☐	Only treated or batch-stored solid manures and slurries should be applied to land before drilling/planting.	☐
☐	Vaccinate cattle to control pathogen colonization and faecal excretion of *E. coli* O157.	☐
☐	No application of manure to ready-to-eat crops within 12 months of harvest and 6 months of drilling/planting.	☐
☐	Prevent children under the age of 11, and other vulnerable groups, coming into contact with animals at petting, or public visitor farms	☐
☐	Use probiotics to reduce *E. coli* O157 shedding rates (e.g. *E. coli* and *Lactobacillus* strains)	☐

retained for further evaluation based upon their effectiveness in reducing human exposure to *E. coli* O157. Experts classified each intervention as 'priority retain', 'retain', 'don't retain' and 'don't know'. These votes were then scored with values of '2', '1', '–1' and '0' respectively (the full set of interventions is available upon request).

Expert elicitation round 2: BWS

In the second round respondents were asked to re-evaluate the shortlist of measures from round 1 using the BWS methodology. This involved respondents assessing 12 sets, each containing five measures, with the combinations of measures within sets determined by an experimental design. An example of a BWS set is shown in Table 2. In the first stage respondents indicated the 'most' and 'least' effective measures in each set they faced. The process was then repeated with the 'most' and 'least' practical measure chosen in each set.

Analysis

The BWS data on perceptions of the measures' practicality and effectiveness can be analysed in a number of ways. The first form of analysis involves counting rather than estimation. The analyst calculates on what proportion of occasions it was shown each measure was selected as 'most' and 'least'. A more sophisticated analysis involves estimation of practicality and effectiveness scores via a choice model based on random utility (RU) theory [33] which dominates the empirical analysis of choice in many fields. We briefly explain the approach below.

Faced with a BWS set the respondent is asked to indicate the best and worst performing measure, hence the respondent is choosing the two measures with the maximum difference in performance between them. In a set of K measures there are $K(K-1)$ best-worst combinations. The objective is to retrieve estimates of the sample's performance scores that best explain the observed pattern of best-worst choices, and it is the RU choice model which permits this. While such choice models typically involve people choosing a single, most preferred option in this study we consider scales defined in terms of practicality, and then effectiveness. We consider there to be scale of practicality on which measures can be located and refer to ϕ_{A} as the position of measure A on that scale. Respondent n's unobserved practicality score for measure A is given by

$$P_{n\mathrm{A}} = \phi_{\mathrm{A}} + {}_{n\mathrm{A}}, \tag{1}$$

where $\varepsilon_{n\mathrm{A}}$ is an error term, the inclusion of which creates a probabilistic rather than deterministic choice model. The probability of person n choosing any pair of best-worst choices, for example measures A and D, is given by the probability that $(P_{n\mathrm{A}} - P_{n\mathrm{D}})$ exceeds all other $K(K-1)$ performance differences within the BWS set.

The model is statistically implementable via the assumption that the error term, $\varepsilon_{n\mathrm{A}}$, has an extreme value type I (Gumbel) distribution. This means that the probability that A and D are chosen as most and least practical, respectively, is given by the standard conditional logit formulation:

$$\frac{\mathrm{e}^{\phi_{\mathrm{A}} - \phi_{\mathrm{D}}}}{\sum_{b=1}^{K} \sum_{\mathrm{W}=1}^{K} \exp^{\phi_b - \phi_{\mathrm{W}}} - K}. \tag{2}$$

Maximum-likelihood estimation of formula (2) involves retrieval of estimates of the ϕ performance scores which maximize the likelihood of the observed pattern of best-worst choices being observed. The approach is relative, i.e. the estimated practicality and effectiveness scores are relative to each other on an arbitrary scale, with one measure's performance score normalized at zero for identification purposes.

The conditional logit model in formula (2) does not, as specified, include personal characteristics and as such does not allow for the investigation of heterogeneity among the sample. One could extend the model to include characteristics, for example we could allow, and test whether, specific performance scores differ across respondents of differing expertise or demographic profiles. An alternative approach to the accommodation of heterogeneity is the infinite mixture, or mixed logit, model [34, 35]. In this model the importance scores are assumed to be drawn from a distribution the mean and standard deviation of which are estimated. An attraction of this model is that as well as identifying a point estimate of the mean importance score, estimates can be derived for each survey respondent, conditional on the best-worst choice data and the estimated population parameters. More formally, person n's performance score for measure A ($\phi_{n\mathrm{A}}$) is drawn from a distribution with mean (ϕ^*_{A}) and standard deviation σ_{A}. Person n's performance score deviates from the sample mean (ϕ^*_{A}) via a disturbance term, ∇, where $\nabla \sim N(0, 1)$:

$$\phi_{n\mathrm{A}} = \phi^*_{\mathrm{A}} + \phi_{\mathrm{A}} \nabla_{n\mathrm{A}}, \tag{3}$$

The derivation of respondent-specific performance scores is particularly attractive in this study given the desire to examine the degree of consensus or disagreement among the expert sample on the performance of the O157 control measures. We refer readers to [34, 35] for more on the estimation of the mixed logit model, noting only that we estimate the model using Bayesian methods [36] with the sampler run for 20 000 iterations (the 'burn in') before the parameters were recorded, followed by another 20 000 iterations with which to summarize the posterior distribution of the measures' performance scores.

RESULTS

Expert elicitation round 1

For the first round, 53 experts were contacted and 31 completed the survey giving a response rate of 58·5 %. In round 2 the snowball sampling process meant 70 experts were invited to complete the survey, 41 of whom did so, giving a response rate of 60 %. Thirty-five of the respondents in round 2 gave details of their expertise and these are described in Table 1. Response rates were higher for agricultural/environmental microbiologists (75 %) compared to public health experts (48 %) which possibly reflects a greater level of confidence for the latter group ranking interventions primarily located in the rural and agricultural environment.

The resultant rankings and the associated suggestion that the top 30 measures would be taken forward to round 2 for more detailed assessment of their practicality and effectiveness were circulated (full details of the results from this round are available upon request). Respondents were invited to appeal against the inclusion/exclusion of any measures but no such appeals were forthcoming. As a result the top 30 measures (Table 3) were included in round 2 which used BWS.

BWS

Within both the practicality and effectiveness stages of the BWS survey 41 respondents completed a total

Table 3. *List of top 30 interventions selected in round 1 for further appraisal in round 2*

No.	Intervention
1	Encourage farmers and farm visitors to wash hands following contact with farm animals
2	Remove farm animals from proximity of private water supplies (e.g. at least 50 m from well, borehole or other private water supply by fencing-off)
3	Remove high shedding animals prior to slaughter (possibly using some form of cow-side test)
4	Prohibit recreational activities (such as walking and camping) on land where manure, slurry or abattoir waste have been applied, or animals and faeces present, in the previous 4 weeks
5	Monitoring of private water supplies to identify those with either high indicator counts, or those in areas of high risk. These supplies would need to be treated (e.g. by ozonation, chlorination or ultra-violet treatment)
6	Stop run-off from adjacent manured fields using vegetative buffer strips of between 2 and 6 m to control contamination of ready-to-eat crops
7	Keep livestock and pets out of ready-to-eat crop areas, using fencing for example
8	No application of manure to land at high risk of direct flow to watercourses (e.g. adjacent to a watercourse, borehole or road culvert, or areas with a dense network of open drains
9	Locate solid manure heaps and slurry pits at least 50 m away from watercourses, field drains and ready-to-eat crops
10	Require 4 weeks between spreading of waste manure and animals grazing in a field
11	Prevent children under the age of 11, and other vulnerable groups, coming into contact with animals at petting, or public visitor farms
12	Reduce leakage from septic tanks in rural areas (e.g. an annual inspection with owner required to pay for any necessary works/repairs)
13	No slurry or livestock manure to be applied to high-risk fields (i.e. high risk of transport into adjacent areas watercourses (e.g. when soils saturated or frozen, or heavy rain expected)
14	Livestock bedding must be kept dry (e.g. ensuring a 'squelch-score' of 1 or 2, where: 1 = very dry, 2 = dry, 3 = squelchy, 4 = very wet, 5 = soaked and slippery)
15	Keep livestock away from packing and storage areas of ready-to-eat crops
16	Require manure handling to be included in a food safety hazard analysis, or HACCP plan, and a COSHH assessment, if growing ready-to-eat crops and spreading manure on same site
17	Fence-off streams from livestock
18	Fallen fruit not to be used for human consumption (fresh fruit, unpasteurized juice) if livestock have grazed there within the preceding 12 months
19	Vaccinate cattle to control pathogen colonization and faecal excretion of *E. coli* O157
20	Store batches of slurry for at least 90 days before spreading on fields
21	Require in-house water troughs to be cleaned every day
22	No application of manure to ready-to-eat crops within 12 months of harvest and 6 months of drilling/planting
23	Only treated or batch-stored solid manures and slurries should be applied to land before drilling/planting
24	Eliminate contamination of ready-to-eat crops from aerosol and windborne drift during manure spreading by prohibiting spreading within ~500 m of ready-to-eat-crops
25	Use probiotics to reduce *E. coli* O157 shedding rates (e.g. *E. coli* and *Lactobacillus* strains)
26	Groups of young stock should not be mixed once established
27	Reduce cattle stocking densities by 50%
28	Separate clean roof water from farmyard areas contaminated with faeces to reduce storage capacity
29	Prevent contact with neighbouring cows via double fencing
30	Ban the disposal of untreated abattoir waste to land

HACCP, Hazard Analysis Critical Control Point; COSHH, control of substances hazardous to health.

of 492 BWS sets, making two choices within each one. We first considered the choice data descriptively rather than via model estimates. These estimates provide a powerful descriptive of the typical assessment of a measure and also the degree of consensus. For example measure 27 (reducing stocking densities by 50%) was ranked least practical in almost 50% of the sets in which it appeared and most practical in less than 5% of the sets. In contrast measure 19 (cattle vaccination) was ranked least practical in 35% of the sets but also most practical in 30% of the sets. Cattle vaccination was voted the most effective measure in 72% of the sets in whch it appeared and was rarely chosen as 'least effective'.

Assessment of the measures' relative practicality and effectiveness was also investigated via estimation of the choice model described in the Analysis section. Table 4 reports results from this model estimated

Table 4. *Sample mean point estimates of the effectiveness and practicality scores for all 30 measures*

Measure	Mean effectiveness	95% CI	Mean practicality	95% CI
1	6·066	4·772–7·360	8·990	8·313–9·666
2	7·805	7·054–8·555	4·281	3·494–5·069
3	7·662	6·824–8·499	1·218	0·592–1·844
4	4·596	3·557–5·634	1·909	1·054–2·763
5	6·342	5·206–7·478	3·777	2·696–4·857
6	2·139	1·407–2·871	1·150	0·752–1·548
7	3·514	2·785–4·243	4·190	3·463–4·917
8	4·109	3·343–4·875	3·798	3·056–4·540
9	4·118	3·460–4·775	4·577	3·876–5·279
10	1·349	1·118–1·581	5·931	5·091–6·770
11	5·801	4·713–6·889	4·166	2·974–5·358
12	0·509	0·223–0·796	1·756	1·222–2·291
13	6·334	5·659–7·010	3·654	2·928–4·380
14	1·689	0·947–2·432	1·682	1·059–2·305
15	1·808	1·370–2·246	6·612	5·588–7·636
16	3·566	2·689–4·443	5·256	4·229–6·282
17	1·061	0·574–1·548	2·316	1·604–3·027
18	0·659	0·443–0·876	4·823	3·929–5·718
19	9·081	8·263–9·899	3·845	2·380–5·309
20	3·403	2·578–4·228	3·941	3·163–4·719
21	0·252	0·119–0·386	1·647	0·987–2·307
22	3·936	3·172–4·699	3·932	3·160–4·704
23	3·890	3·169–4·612	2·510	1·904–3·115
24	1·135	0·749–1·521	0·773	0·509–1·037
25	3·573	2·440–4·706	2·644	1·845–3·443
26	1·834	1·006–2·661	1·679	1·267–2·091
27	1·910	1·053–2·768	0·622	0·294–0·949
28	0·551	0·159–0·942	2·219	1·756–2·682
29	0·219	0·119–0·318	1·791	1·228–2·354
30	1·090	0·713–1·467	4·313	3·619–5·007

(separately) on both the practicality and effectiveness BWS data. The results in Table 4 show the sample mean point estimates of the practicality and effectiveness scores for all 30 measures. These estimates have been rescaled to sum to 100 to aid interpretation and comparison across the practicality and effectiveness models. In addition to the point estimates, 95% confidence intervals are provided for each parameter estimate.

Cattle vaccination, the removal of animals from the proximity of private water supplies (PWS); the removal of high shedding animals prior to slaughter and the monitoring (and where appropriate treatment) of PWS; were estimated to be the most effective measures. The results indicate that the annual inspections of septic tanks, the daily cleaning of in-house water troughs and the use of double fencing to prevent contact between (sub)herds were, on average, regarded as least effective.

In terms of practicality, encouraging farmers and farm visitors to wash hands following contact with farm animals, keeping livestock away from stores of ready-to-eat crops and a 4-week grazing ban following the spreading of animal waste were seen as the most practical measures. Halving stocking densities, banning manure spreading within 500 m of ready-to-eat crops and using vegetative buffer strips to control contamination of ready-to-eat crops from run-off were seen as the least practical to implement.

The mean practicality and effectiveness scores of the measures are combined in Figure 1. The estimated scores are zero-centred and plotted in practicality and effectiveness space in which the axes represent the mean (zero) practicality and effectiveness scores within the sample. Hence measures plotted above 0 on the y-axis have higher than average effectiveness scores, those with scores over 0 on the x-axis have above average practicality scores. Those measures

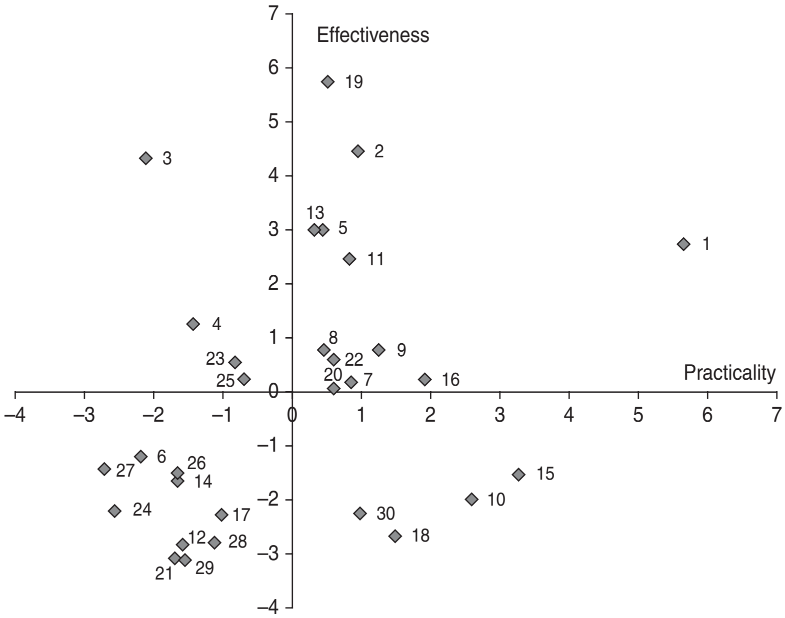

Fig. 1. Zero-centred scatterplot of mean effectiveness and practicality scores for the 30 control measures.

located in the upper right quadrant are regarded as performing relatively well in terms of practicality and effectiveness, those in the bottom left quadrant perform relatively poorly on both criteria. Measures located in the remaining two quadrants highlight the trade-off between the practicality and effectiveness of some measures. Those interventions located above in the upper right-hand segment of the 2 × 2 plot might be considered candidate interventions for consideration by policy-makers as they score positively for effectiveness and/or practicality.

As indicated in the Analysis subsection of the Methodology section, the choice model allows the estimation of respondent-specific scores as well as the estimate of the mean score for the sample as a whole. Analysis of these scores for a single measure allows a consideration of the degree of consistency or disagreement among the sample regarding the measure's performance. This is highlighted in Figure 2 which shows the spread of practicality and effectiveness scores for two example measures. There was a high degree of agreement regarding the effectiveness of vaccination to reduce pathogen loads, but opinion appeared to be divided regarding the practicality of its implementation. For some measures there was little consistency regarding either criteria, leading to a more uniform distribution of scores in practicality and effectiveness space as shown for intervention 16 [Hazard Analysis Critical Control Point (HACCP) requirement for manure spreading].

DISCUSSION

This study has proposed and implemented BWS as an expert opinion elicitation tool. BWS has been used to evaluate the practicality and effectiveness of measures to reduce human exposure to *E. coli* O157 in rural and farm environments. Following an initial short-listing process 30 measures were included in the BWS survey.

The BWS process, completed online, allowed the involvement of many experts from a diverse range of expertises (public health, environmental microbiology, epidemiology, etc.) in varied, often geographically remote, locations. Asking those experts to simply rank the full set of 30 measures would have been extremely cognitively demanding, hence the use of BWS with the repeated ranking over smaller subsets. The study achieved a high response rate with low levels of dropout mid-BWS process, which we regard as evidence that the process was cognitively bearable and intuitive.

The *ex-post* analysis employed counts and the estimation of a choice model on the best-worst data which allowed a full, scaled ranking to be derived. The model results include both sample-mean, and respondent-specific, estimates of the measures' practicality and effectiveness. Analysis of the individuals' scores for a single measure allows the degree of consensus or disagreement to be investigated.

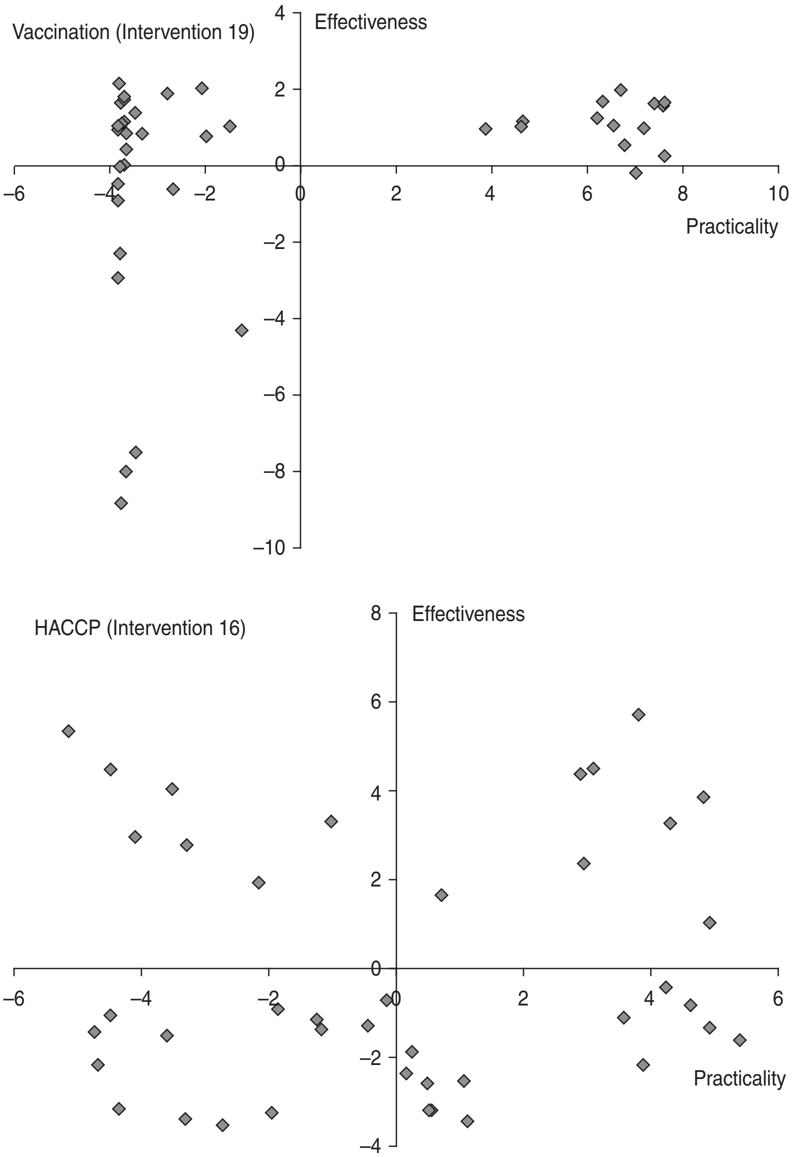

Fig. 2. Respondent-specific practicality and effectiveness scores for two control measures (each symbol denotes the practicality and effectiveness scores of a single expert).

We regard the BWS as highly flexible and suitable when evaluation needs to be multi-dimensional. In this case two facets of performance were considered (practicality and effectiveness). This allows the location of each measure in a multi-dimensional performance space (Fig. 1). This 2 × 2 plot is easily interpretable and provides a powerful stimulus for further discussion, as has been evident from our use of it at, *inter alia*, disease management workshops and meetings of the Advisory Committee on the Microbiological Safety of Food and meat industry bodies.

As to the substantive results, hand washing was regarded as by far the most practical measure. The separation, in time or space, of livestock waste from water supplies, water courses, ready-to-eat crops and the general public were common to many of the measures seen as most practical and effective. In addition some more technical 'fixes' were identified as most effective: concerning livestock and PWS. The three measures viewed as most effective were vaccination, removal of high shedding animals, and monitoring and treatment of PWS.

Some measures which recent research [37] has suggested as promisingly effective in this context were regarded as less effective (requiring in-house water troughs to be cleaned daily, maintaining dry livestock bedding). This raises the issue of managing the tension between a common expert opinion and fragments of research evidence which only some may be party to, or accept. The issue of bedding management was highlighted by the Griffin Report into the Godstone outbreak. Another notable feature of our results is that the prevention of children under the age of 11, and other vulnerable groups, from coming into contact with animals at petting, or public visitor farms was regarded as both highly effective and practical.

We note that the Godstone outbreak took place between the first and second rounds of our elicitation process which may have affected the perceived performance of both this measure and hand washing. It is evident that an empirical microbiological evaluation of the exposure hazard would remain unaffected by such an outbreak and underlines the importance of basing policy decisions on experimental evidence where possible. Any such sensitivity of the experts' effectiveness assessments to a recent or current high profile outbreak (such as Godstone or the 2011 outbreak in Germany) may be regarded as a weakness in any expert elicitation approach. It could, however, be regarded as simply reflecting an updating of knowledge and beliefs among the panel in the light of new evidence provided by such an outbreak.

The identification of hand washing as a practical and effective means of reducing exposure concurs with the findings of the Griffin Report [25]. However, the implementation of the measure raises many issues for further thought and analysis. Questions include how best to design hand-washing facilities into the farm layout and whose responsibility it is to ensure that children visiting open farms wash their hands (Griffin argues it is the parents' responsibility). Further, there is the danger that a sole measure approach could lead to the neglect of other risk factors, for example, a focus on hand washing may be at the expense of other critical control measures, for example the initial faecal contact.

BWS is best suited to the evaluation of large sets of stand-alone measures, as was the case here with 99 candidate measures. If one was taking a much smaller set of measures, with multiple levels, and trying to identify the 'best' bundles of those measures then alternative methods (ACA, CBC) would be more suitable. Having identified the best performing measures in the BWS study one could then consider bundling measures. Within such bundles, measures may have purely additive effects; however, in some cases they may enhance or reduce the impact of other measures. This could be assessed by modelling the cross-impact balances of intervention bundles [38].

This relates to a further notable point that, in order to keep the process focused and manageable for respondents, we focused on a single pathogen. There may, however, be spillover effects to other pathogens and policy targets. A next phase could model the potential of measures and bundles to add value through impacts on other target organisms, such as *Cryptosporidium* and norovirus, or contribute to the meeting of other requirements (e.g. the Water Framework Directive). This would entail re-presenting the finalized intervention list to a different panel of experts whose expertise focused on the secondary target pathogens and viruses. The same BWS methodologies would be employed but the initial research question would ask experts to evaluate the 30 interventions based upon their ability to reduce exposure to *Cryptosporidium* and/or norovirus.

An additional extension of the research would be to extend the sample beyond research or regulatory experts to, for example, farmers. An assessment of the consistency (or otherwise) of the assessments of the two groups would be revealing. Ultimately it will be farmers' perceptions of the practicality and effectiveness of the proposed control measures that will determine the levels of adoption *in situ*.

ACKNOWLEDGEMENTS

This research was undertaken as part of the project 'Reducing *E. coli* O157 Risk in Rural Communities' (RES-229-31-0003) funded under the UK Research Councils' Rural Economy and Land Use Programme. The authors thank all those who participated in the study. Advice from Dr Natasha Valeeva of Wageningen University, research assistance from Dr Seda Erdem and comments from colleagues on the project are all gratefully acknowledged.

DECLARATION OF INTEREST

None.

REFERENCES

1. **Woolhouse M, Gaunt E.** Ecological origins of novel human pathogens. *Critical Reviews in Microbiology* 2007; **33**: 231–242.
2. **Morgan D,** ***et al.*** Assessing the risk from emerging infections. *Epidemiology and Infection* 2009; **137**: 1521–1530.

3. **Tauxe RV.** Emerging foodborne pathogens. *International Journal of Food Microbiology* 2002; **78**: 31–41.
4. **Henson S.** Estimating the incidence of food-borne Salmonella and the effectiveness of alternative control measures using the Delphi method. *International Journal of Food Microbiology* 1997; **35**: 195–204.
5. **Fish R, *et al.*** Unruly pathogens: eliciting values for environmental risk in the context of heterogeneous expert knowledge. *Environmental Science & Policy* 2009; **12**: 281–296.
6. **van Schaik G, *et al.*** Adaptive conjoint analysis to determine perceived risk factors of farmers, veterinarians and AI technicians for introduction of BHV1 to dairy farms. *Preventive Veterinary Medicine* 1998; **37**: 101–112.
7. **Sørensen JT, *et al.*** Expert opinions of strategies for milk fever control. *Preventive Veterinary Medicine* 2002; **55**: 69–78.
8. **Garabed RB, *et al.*** Use of expert opinion for animal disease decisions: an example of foot-and-mouth disease status designation. *Preventive Veterinary Medicine* 2009; **92**: 20–30.
9. **Valeeva NI, *et al.*** Improving food safety at the dairy farm level: farmers' and experts' perceptions. *Review of Agricultural Economics* 2005; **27**: 574–592.
10. **Cross P, Williams P, Edwards-Jones G.** Differences in the perceptions of farmers and veterinary surgeons about the efficacy of mitigation strategies for controlling bluetongue. *Veterinary Record* 2009; **165**: 397–403.
11. **Valeeva NI, *et al.*** Improving food safety within the dairy chain: an application of conjoint analysis. *Journal of Dairy Science* 2005; **88**: 1601–1612.
12. **Orme BK.** *Getting Started with Conjoint Analysis: Strategies for Product Design and Pricing Research*. Madison, USA: Research Publishers, 2006.
13. **Mennecke BE, *et al.*** A study of the factors that influence consumer attitudes toward beef products using the conjoint market analysis tool. *Journal of Animal Science* 2007; **85**: 2639–2659.
14. **Coltman T, Devinney T, Keating B.** Best-worst scaling approach to predict customer choice for 3PL services. *Journal of Business Logistics* 2010; **32**: 139–152.
15. **Riley LW, *et al.*** Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *New England Journal of Medicine* 1983; **308**: 681–685.
16. **Bell BP, *et al.*** A multistate uutbreak of *Escherichia coli* O157:H7-associated bloody diarrhea and hemolytic uremic syndrome from hamburgers: the Washington experience. *Journal of the American Medical Association* 1994; **272**: 1349–1353.
17. **Kassenborg HD, *et al.*** Farm visits and undercooked hamburgers as major risk factors for sporadic *Escherichia coli* O157:H7 infection: data from a case-control study in 5 FoodNet Sites. *Clinical Infectious Diseases* 2004; **38** (s3): S271–S278.
18. **Locking ME, *et al.*** Risk factors for sporadic cases of *Escherichia coli* O157 infection: the importance of contact with animal excreta. *Epidemiology and Infection* 2001; **127**: 215–220.
19. **O'Brien S, Adak G, Gilham C.** Contact with the farming environment as a major risk factor for sporadic cases of Shiga toxin-producing *Escherichia coli* O157 infection in humans. *Emerging Infectious Diseases* 2001; **7**: 1049–1051.
20. **Parry SM, *et al.*** Risk factors for and prevention of sporadic infections with vero cytotoxin (shiga toxin) producing *Escherichia coli* O157. *Lancet* 1998; **351**: 1019–1022.
21. **Ogden ID, *et al.*** Long-term survival of *Escherichia coli* O157 on pasture following an outbreak associated with sheep at a scout camp. *Letters in Applied Microbiology* 2002; **34**: 100–104.
22. **Ogden ID, MacRae M, Strachan N.** Is the prevalence and shedding concentrations of *E. coli* O157 in beef cattle in Scotland seasonal? *FEMS Microbiology Letters* 2004; **233**: 297–300.
23. **Khanna R, *et al.*** Environmental prevention of human disease from verocytotoxin-producing *Escherichia coli*. *Nephrology Dialysis Transplantation* 2008; **23**: 1819–1822.
24. **Solecki O, *et al.*** Can the high levels of human verocytotoxigenic *Escherichia coli* O157 infection in rural areas of NE Scotland be explained by consumption of contaminated meat? *Journal of Applied Microbiology* 2007; **103**: 2616–2621.
25. **Griffin G.** Review of the major outbreak of *E. coli* O157 in Surrey, 2009. Report of the Independent Investigation Committee, June 2010 (www.griffininvestigation.org.uk/).
26. **Lynch MF, Tauxe RV, Hedberg CW.** The growing burden of foodborne outbreaks due to contaminated fresh produce: risks and opportunities. *Epidemiology and Infection* 2009; **137** (Special Issue 3): 307–315.
27. **Newell DG, *et al.*** Food-borne diseases – the challenges of 20 years ago still persist while new ones continue to emerge. *International Journal of Food Microbiology* 2010; **139** (Suppl. 1): S3–S15.
28. **Auger P, Devinney T, Louviere J.** Using best–worst scaling methodology to investigate consumer ethical beliefs across countries. *Journal of Business Ethics* 2007; **70**: 299–326.
29. **Marley AAJ, Louviere JJ.** Some probabilistic models of best, worst, and best-worst choices. *Journal of Mathematical Psychology* 2005; **49**: 464–480.
30. **Lusk JL, Briggeman BC.** Food values. *American Journal of Agricultural Economics* 2009; **91**: 184–196.
31. **Jaeger SR, *et al.*** Best-worst scaling: an introduction and initial comparison with monadic rating for preference elicitation with food products. *Food Quality and Preference* 2008; **19**: 579–588.
32. **Lee JA, Soutar GN, Louviere J.** Measuring values using best-worst scaling: the LOV example. *Psychology and Marketing* 2007; **24**: 1043–1058.
33. **McFadden D (ed.).** *Conditional Logit Analysis of Qualitative Choice Behavior*. New York: Academic Press, 1974.
34. **McFadden D, Train K.** Mixed MNL models for discrete response. *Journal of Applied Econometrics* 2000; **15**: 447–470.

35. **Train K.** *Discrete Choice Methods and Simulation*. Cambridge: Cambridge University Press, 2003.
36. **Rigby D, Balcombe K, Burton M.** Distributional assumptions and mixed logit model performance: preferences and GM foods. *Environmental and Resource Economics* 2009; **42**: 211–226.
37. **Ellis-Iversen J, *et al*.** Farm practices to control *E. coli* O157 in young cattle – a randomised controlled trial. *Veterinary Research* 2008; **39**: 3.
38. **Weimer-Jehle W.** Cross-impact balances: a system-theoretical approach to cross-impact analysis. *Technological Forecasting and Social Change* 2006; **73**: 334–361.