REVIEW ARTICLE


S. J. O’BRIEN¹,*, I. A. GILLESPIE², M. A. SIVANESAN², R. ELSON², C. HUGHES² AND G. K. ADAK²

¹ Division of Medicine and Neurosciences, University of Manchester, Manchester, UK
² Environmental and Enteric Diseases Department, Health Protection Agency Centre for Infections, London, UK

(Accepted 8 November 2005, first published online 18 January 2006)

SUMMARY

Systematic national surveillance of outbreaks of infectious intestinal disease (IID) was introduced in England and Wales in 1992 to provide comprehensive information on causative organisms, sources or vehicles of infection and modes of transmission. We compared information from this system with that published in the peer-reviewed literature between 1 January 1992 and 31 January 2003 to assess the potential effect of publication bias on food-safety policy. During the study period 1763 foodborne outbreaks of IID were reported to national surveillance. Fifty-five were published in the peer-reviewed literature. The peer-reviewed literature overestimated the impacts of milk/milk products, miscellaneous foods (e.g. sandwiches) and desserts and underestimated those of poultry, fish and shellfish, red meat/meat products and eggs/egg products. Without systematic surveillance, knowledge of causative organisms, sources or vehicles of infection and modes of transmission, as gleaned from the peer-reviewed literature, would potentially distort food-safety policy.

INTRODUCTION

The United Kingdom (UK) Food Standards Agency has set a target for a 20% reduction in foodborne illness by April 2006 [1]. To achieve this it needs robust information on the vehicles and causes of food poisoning. The majority of foodborne infectious intestinal disease (IID) is sporadic and so the identification of sources of infection in these instances is often unavailable [2]. Thus, epidemiological and microbiological evidence gained in foodborne outbreaks can provide some of the strongest information linking food to illness. The routine surveillance of such outbreaks provides a powerful tool available to policymakers (including the Food Standards Agency), as well as the wider public health community.

The system for the surveillance of IID in England and Wales was introduced in 1992 [3]. The objectives were to:

- identify routes of transmission;
- identify trends in pathogens causing outbreaks;
- identify trends in food vehicles;
- detect new pathogens/vehicles;
- assess the impact of outbreaks in different settings.

This dataset has been interrogated regularly to provide information for individuals from a variety of organizations; from university students on the one hand to policy-makers on the other. Indeed the majority of requests for information are from government agencies – the Food Standards Agency,
the Department of Health, and the Department for the Environment, Food and Rural Affairs. It is imperative, therefore, that the data used to provide this information are as complete as possible.

Alternative sources of intelligence include the peer-reviewed literature. Outbreaks that are published in the literature are likely to be unusual, for example highlighting novel food vehicles. Therefore, we compared foodborne general outbreaks reported to the surveillance system with those appearing in the peer-reviewed literature to assess the potential effect of publication bias on food-safety policy.

**METHODS**

**Surveillance of general outbreaks of IID**

The national system for the surveillance of general outbreaks of IID has been described in detail elsewhere [3, 4]. It is focused on a standard questionnaire which is sent to the lead investigator, who is requested to complete it when the outbreak is over. The questionnaire seeks data on the outbreak setting, the mode of transmission, the causative organism and details of epidemiological and laboratory investigations. Data from returned questionnaires (response rate >70% [5]) are stored in a dynamic database (GSURV) which is derived from Epi-Info [6]. For the purposes of this paper, outbreaks were included initially where the mode of transmission was described as mainly foodborne.

**Literature search strategy**

Relevant literature was obtained by interrogating the online PubMed database [7]. English-language articles published from 1992 to 2003 were considered. Searches of the title word 'outbreak' and the expanded Medical Subject Headings (MeSH) terms ‘disease outbreaks’ and ‘food’ were undertaken. Furthermore, a search of Medline was undertaken by a professional librarian. A search on the MeSH subheadings ‘food poisoning’, ‘epidemiology’, ‘aetiology’, ‘food handling’, ‘food’ and ‘beverages’ was combined with searches of the MeSH term ‘disease outbreaks’ and then limited by the MeSH subheading 'Great Britain’. Reference lists were used to identify additional outbreak papers.

Papers relating to family outbreaks, or those which took place outside England and Wales were excluded, as were those where the mode of transmission was not described as foodborne. Initial reports of outbreaks, published in the *Communicable Disease Report (CDR) Weekly* [8], were excluded, as were outbreak reviews or papers which discuss outbreaks in scant detail (e.g. discussing decreased susceptibility to ciprofloxacin in an outbreak of *Salmonella enterica* serovar Typhimurium definitive phage type (DT) 104 associated with a pasteurization failure at an on-farm dairy [9]).

Two members of the research team (I.A.G. and M.A.S.) reviewed all the output from each search independently of each other in order to identify articles to be included and agreed a common list. Transcription errors when creating the publication dataset were minimized at two stages by highlighting data for inclusion on the manuscripts and a process of rigorous data-checks once these data had been transcribed. Routine validation of the national outbreak dataset takes place as part of ongoing surveillance procedures.

Data collected on general outbreaks (described above) were extracted where available and recorded in a database.

**Calculation of lag periods**

Where the date of publication and the last date of onset of symptoms in an outbreak were available, the publication lag period was calculated. The last date of onset in an outbreak and date of questionnaire completion were used to define the reporting lag for general outbreaks of IID.

**Calculation of Publication Bias Index (PBI)**

We calculated a publication bias index (PBI) to provide a direct measure of the impact of the publication bias for characteristics of interest. The PBI was expressed as the ratio of the percentages of outbreaks by type of report. For example, if 10/50 outbreaks published in the peer-reviewed literature were set on farms, compared with 49/1500 general outbreaks then the PBI would be 20%/3% = 6.1. Where the PBI was >1, outbreaks were over-represented in the literature compared with the outbreak surveillance dataset. Conversely where the PBI was <1, outbreaks were under-represented in the literature compared with the outbreak surveillance dataset.

**Statistical analysis**

The datasets were analysed using Microsoft Excel 2000 (Microsoft Corp., USA), Epi-Info version 6.04b (CDC, Atlanta, GA, USA) and Stata version 8 (StataCorp LP, College Station, TX, USA). Relative
proportions were compared using the \( \chi^2 \) test. For smaller samples Fisher’s exact test was used. Medians and means were compared using the non-parametric \( K \)-sample test on the equality of medians and Student’s \( t \) test respectively.

**RESULTS**

Between 1 January 1992 and 31 December 2003, 1784 potential foodborne general outbreaks of IID in England and Wales were identified in the peer-reviewed literature. Initially, 66 outbreaks were considered eligible for inclusion [9–69]. Upon closer scrutiny, one outbreak occurred outside England and Wales [64], in five outbreaks initial foodborne transmission was followed by person-to-person transmission [17, 38, 52, 57, 67] three described the outbreak in insufficient detail for analysis [9, 69, 70], one was an outbreak review [28] and one was a news report [55]. These outbreaks were excluded from further analysis, leaving 55 foodborne general outbreaks identified in the literature (‘literature outbreaks’). During the same time period, 7658 general outbreaks of IID were reported to the Health Protection Agency (HPA) Centre for Infections (CfI). In 1763 (23%) of these outbreaks the mode of transmission was described as ‘mainly foodborne’.

**Lag period**

The date of publication was available for all the literature outbreaks, and the last onset date of the outbreak was described in 34 outbreaks (65%). The lag-time period from outbreaks taking place to their appearance in the peer-reviewed literature ranged from 7 to 169 months, and the median length was 23 months. The date of entry of data was unavailable for outbreaks before 1996 in GSURV, so these outbreaks were excluded from the calculation of the reporting lag period. It was, therefore, possible to calculate the reporting lag time for 707/939 general outbreaks (75%). The time period between the occurrence of an outbreak and the completion of the national surveillance system questionnaire ranged from 1 to 44 months, and the median length was 13 months. The median publication lag period in literature outbreaks was significantly longer than the median reporting lag in general outbreaks (\( P < 0.001 \)).

**Completeness of data**

There were interesting differences between the sources of data in terms of their completeness. In general, outbreak dynamics and impact (numbers of cases hospitalized or died, month of onset and outbreak duration) were more complete in the outbreak surveillance dataset compared with outbreaks reported in the peer-reviewed literature. More complete in the published reports were data on pathogens, food vehicles, the evidence implicating a food vehicle, as well as the food-handling/hygiene faults that were thought to have contributed to the outbreak.

**Dynamics and impact**

In the 52 literature outbreaks, 2443 people were affected (range 4–361), with 199 hospital admissions (range 0–33) and nine deaths (range 0–3) reported. This corresponds with 39 842 people affected (range 2–530), 1537 hospital admissions (range 0–42) and 68 deaths (range 0–10) in the 1528 general outbreaks. Literature outbreaks were larger than general outbreaks with regard to the mean number of people affected (44 vs. 23, \( P < 0.001 \)), admitted to hospital (6.9 vs. 1.3, \( P < 0.001 \)) or reported to have died (0.6 vs. 0.07, \( P < 0.001 \)). Furthermore, the duration of literature outbreaks (range 2–393 days, mean 31) was significantly longer than general outbreaks (range 1–373 days, mean 8) (\( P < 0.001 \)).

**Outbreak setting**

The various settings for literature and general outbreaks of IID are summarized in Table 1. The majority (55%) of general outbreaks took place in, or were linked to, commercial catering premises (canteens, halls or caterers, hotels, public houses or bars, restaurants, shop caterers), but these premises accounted for less than one fifth of literature outbreaks (18%) (PBI 0.6, \( \chi^2 \) \( P < 0.001 \)). Literature outbreaks were more likely to occur in the community (20% vs. 2%, PBI 10.7, \( P < 0.001 \)) or were linked to shops (20% vs. 7%, PBI 2.8, \( P = 0.002 \)) compared with general outbreaks.

**Pathogens and toxins**

Bacterial pathogens predominated in both literature (88%) and general outbreaks (76%, Table 2). Literature outbreaks reported outbreaks of campylobacteriosis.
(16% vs. 4%, PBI 4.3, P < 0.001) and Shiga toxin-producing *Escherichia coli* O157 (STEC O157) infection (13% vs. 3%, PBI 5.0, P < 0.001) more often than general outbreaks. Outbreaks of *Clostridium perfringens* gastroenteritis were reported relatively commonly in general outbreaks (12%), but were not reported in literature outbreaks.

### Food vehicles

In the majority (45/55, 82%) of literature outbreaks only one food vehicle was reported. In four (7%) and five (9%) outbreaks two and three vehicles were reported respectively and in only one outbreak (2%) the investigators were unable to identify the foodborne vehicle of infection. In just over half (1030/1763, 58%) of the general outbreaks the vehicle of infection was reported. In 162 outbreaks (9%) two vehicles were identified and in 87 outbreaks (5%) three vehicles were reported. It is notable that in over a quarter (484, 27%) of general outbreaks the investigators were unable to identify a vehicle of infection.

### Evidence implicating food vehicles

The evidence implicating foodborne vehicles for literature and general outbreaks is shown in Table 3. Miscellaneous foods (e.g. sandwiches), desserts and milk/milk products appeared to be over-represented in literature outbreaks, whereas poultry, fish and shellfish, red meat/meat products and eggs/egg products appeared to be under-represented.
outbreaks (43/55, 78%) than in general outbreaks (427/1763, 24%) ($P < 0.001$).

**Contributory faults**

The faults, thought to have contributed to an outbreak, were available for almost all (53/55, 96%) of the literature outbreaks but only two thirds (1170/1763, 66%) of the general outbreaks ($P < 0.001$). Furthermore, on average, more faults were reported in literature outbreaks (1.7) than in general outbreaks (1.1, $P < 0.001$). Inadequate heat treatment (PBI 1.7) and cross contamination (PBI 1.8) were over-represented in literature outbreaks (Table 4).

**DISCUSSION**

We have compared foodborne general outbreaks reported to the surveillance system with those appearing in the peer-reviewed literature to assess the potential effect of publication bias on food-safety policy. Few of the foodborne outbreaks reported to CfI led to peer-reviewed publication. Publications in peer-reviewed journals tended to favour the unusual or novel event, which is not necessarily surprising since peer-reviewed journals favour articles providing original findings. However, in order to develop rational policies, policy-makers, enforcers and risk assessors need to know what usually causes people to become ill, as well as what is unusual, because it is in dealing with commonly occurring problems that the greatest health gains are to be made.

It is possible that we might have missed some articles that had been published, underestimating the visibility of the information base. However, two searches were conducted, one by a professional medical librarian. Furthermore, we scanned the references at the end of peer-reviewed articles to identify any that we might have missed.

We applied very stringent criteria for defining papers for inclusion in this study and this might also have underestimated dissemination of the results of outbreak investigations in the public domain. In particular, one potential source of data, not included in our search strategy was the weekly surveillance bulletin produced by the HPA (CDR Weekly) [71]. The reason for this was that although preliminary reports of foodborne outbreak investigations sometimes appear for the purposes of alerting other colleagues to their occurrence and for case-finding, follow-up reports containing the outcome of the investigations are rare. The fact that the same analytical strategy was applied to both datasets means that the results should be fully comparable.

Six papers appeared in the peer-reviewed literature describing outbreaks that were not already reported to the national surveillance scheme [24, 25, 31, 33, 56, 62]. We acknowledge that linking the two datasets is not a simple process, especially as key fields (e.g. details of the outbreak setting) might be omitted from peer-reviewed publications. National outbreak surveillance is a passive system, relying on local investigators (consultants in communicable disease control, environmental health officers or microbiologists) to trigger a report. There is a fail-safe mechanism for the

### Table 3. ‘General outbreaks’ and ‘literature outbreaks’ of infectious intestinal disease, England and Wales, 1992–2002. Reported vehicles of infection

<table>
<thead>
<tr>
<th>Vehicle</th>
<th>Percentage of outbreaks reporting this vehicle</th>
<th>PBI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poultry</td>
<td>15</td>
<td>0.8</td>
</tr>
<tr>
<td>Red meat/meat products</td>
<td>13</td>
<td>0.9</td>
</tr>
<tr>
<td>Fish and shellfish</td>
<td>5</td>
<td>0.5</td>
</tr>
<tr>
<td>Salad/vegetables/fruit</td>
<td>5</td>
<td>1.0</td>
</tr>
<tr>
<td>Sauces</td>
<td>2</td>
<td>0.7</td>
</tr>
<tr>
<td>Desserts</td>
<td>15</td>
<td>1.4</td>
</tr>
<tr>
<td>Milk/milk products</td>
<td>15</td>
<td>5.6</td>
</tr>
<tr>
<td>Water</td>
<td>0</td>
<td>n.c.</td>
</tr>
<tr>
<td>Miscellaneous foods</td>
<td>27</td>
<td>1.9</td>
</tr>
<tr>
<td>Eggs/egg dishes</td>
<td>4</td>
<td>0.7</td>
</tr>
<tr>
<td>Rice</td>
<td>2</td>
<td>0.7</td>
</tr>
</tbody>
</table>

PBI, Publication Bias Index; n.c., not calculable.

### Table 4. ‘General outbreaks’ and ‘literature outbreaks’ of infectious intestinal disease, England and Wales, 1992–2002. Faults thought to have contributed to outbreaks

<table>
<thead>
<tr>
<th>Contributory faults</th>
<th>Number (%) of outbreaks reporting</th>
<th>PBI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infected food handler</td>
<td>6 (11)</td>
<td>0.9</td>
</tr>
<tr>
<td>Inadequate heat treatment</td>
<td>27 (49)</td>
<td>1.7</td>
</tr>
<tr>
<td>Cross contamination</td>
<td>28 (51)</td>
<td>1.8</td>
</tr>
<tr>
<td>Inappropriate storage</td>
<td>16 (29)</td>
<td>1.0</td>
</tr>
<tr>
<td>Other faults</td>
<td>18 (33)</td>
<td>2.9</td>
</tr>
</tbody>
</table>

PBI, Publication Bias Index.
foodborne zoonoses, the majority of which are referred to national reference laboratories. National reference microbiologists also report to the system. The response rate for these known outbreaks is consistently high [5]. However, for non-zoonotic foodborne disease outbreaks, e.g. Norovirus, this failsafe mechanism is not as robust so that the national surveillance dataset might underestimate the true incidence of all foodborne disease outbreaks in England and Wales. The forthcoming European Union Zoonoses Regulations [72], which come into force in 2007, will help to strengthen national foodborne disease outbreak surveillance since there will be a duty to investigate and report centrally all suspected outbreaks of foodborne disease.

Those outbreaks that are published tend to be those where the evidence, microbiological, environmental and epidemiological, is the strongest. The national dataset comprises outbreaks where one or two of those three strands of evidence might be missing. The danger in only considering information for policymaking from the peer-reviewed literature is that the influence of unusual organisms/food vehicles might be over-emphasized. We must acknowledge that the discussion sections of all peer-reviewed papers should place the novel observations in the context of what is already known and how the new findings contribute further to our understanding of the epidemiology of foodborne diseases. This means that while the topics published may be biased, a critical appraisal of the papers need not necessarily lead to a biased assessment of food-safety policy issues. However, the importance of certain pathogens in foodborne disease outbreak causation in England and Wales might be overlooked. For example, during the study period no peer-reviewed papers describing outbreaks of Gram-positive bacteria, Staphylococcus aureus and C. perfringens, appeared in the literature, despite the latter being recorded as the pathogen responsible for some 12% of outbreaks during the surveillance period. Indeed the fact that C. perfringens is one of the Food Standard Agency’s five target pathogens reflects its importance in the outbreak dataset [1]. Relying solely on the literature might lead to an over-emphasis on controlling the milk supply and a failure to pay enough attention to the role of contaminated fish and shellfish in outbreak causation.

The main drawback of only considering information from the outbreak surveillance dataset is that some of the evidence implicating food vehicles and contributory faults is weaker than that which would be accepted for publication. However, the advantage of data collection for the national dataset is that levels of evidence implicating particular food vehicles are available. This means that the quality of evidence linking organisms to food vehicles supplied to policy-makers is transparent.

Our work has shown that, in the absence of systematic national surveillance, knowledge of causative organisms, sources or vehicles of infection and modes of transmission in foodborne disease outbreaks gleaned from the peer-reviewed literature might have the potential to distort food-safety policy. In practice, both types of data are needed so that novel and unusual peer-reviewed findings can be contextualized. Our study emphasizes the need for routine systematic surveillance of outbreaks.

ACKNOWLEDGEMENTS

The authors thank the consultants in communicable disease control, public health physicians, microbiologists, environmental health officers, infection control nurses, and all the staff at the HPA and NHS laboratories, without whose work this surveillance scheme would not function, and also Mrs S. Le Baigue, Mrs C. Hopcroft and Miss C. Penman who maintain the outbreak database at CfI.

DECLARATION OF INTEREST

None.

REFERENCES

Publication bias in foodborne outbreaks of IID 673


