As clean as they look? Food hygiene inspection scores, microbiological contamination, and foodborne illness

Janet Fleetwood\textsuperscript{a,∗}, Shamim Rahman\textsuperscript{b,1}, Darren Holland\textsuperscript{b}, David Millson\textsuperscript{b,2}, Laura Thomson\textsuperscript{b,3}, Guy Poppy\textsuperscript{c}

\textsuperscript{a} Drexel University, Dornsife School of Public Health, 3215 Market Street, Philadelphia, PA, 19104, USA
\textsuperscript{b} Food Standards Agency, Aviation House, 125 Kingsway, London, WC2B 6NH, UK
\textsuperscript{c} University of Southampton, Biological Sciences, Life Sciences Building 85, Highfield Campus, Southampton, SO17 1BJ, UK

\textbf{ARTICLE INFO}

\textbf{Keywords:}
Restaurant inspection
Microbiological contamination
Foodborne illness

\textbf{ABSTRACT}

This work describes the relationship between compliance with food hygiene law as reflected in food hygiene scores; measures of microbiological contamination of food samples taken from consumer-facing food businesses in England, Northern Ireland and Wales; and outbreaks of foodborne illness. This paper demonstrates an association between the results of food hygiene inspections done by trained inspectors, using a rigorous and consistent procedure, with microbiological contamination of actual food samples from those premises. A proposed theoretical model further demonstrates the reduction in foodborne illness that would result if there were increased compliance with food hygiene law.

\section{1. Introduction}

The primary aim of this work was to determine whether there is a relationship between compliance with food hygiene law as reflected in premises’ food hygiene scores; objective measures of microbiological contamination of food samples taken from food business premises including restaurants, caterers and retailers; and outbreaks of foodborne illness. Specifically, Local Authorities applied the Food Law Codes of Practice to consumer facing food premises in England, Northern Ireland and Wales. Our analysis indicates that visual food hygiene inspections done by trained inspectors measuring compliance with food law are useful at detecting illness-inducing microbiological contamination and can be an important tool in reducing foodborne illness outbreaks.

\subsection{1.1. Impact of foodborne illness}

In the United Kingdom, there are an estimated 500,000 cases of foodborne illness per year from known pathogens, and over a million cases if unknown pathogens were included in the count (Tam, Larose, & O’Brien, 2014). In 2008, the annual cost of foodborne illness was estimated at approximately £1 billion (Food Standards Agency, 2016a, p. 8). This cost and the incidence, along with the illnesses’ very unpleasant symptoms, could be reduced by improved food hygiene (USDA, 2016).

\subsection{1.2. Food hygiene inspections and the Food Standards Agency (FSA)}

The Food Standards Agency (FSA) promotes microbiological safety of food in England, Northern Ireland and Wales by developing strategies for reducing foodborne illness, dealing with microbiological food hazards and foodborne disease outbreaks, promoting food safety management, and providing education and guidance to producers, retailers, food premises, and the public. The FSA emphasizes a risk based approach. England, Wales and Northern Ireland publish their own Code of Practice; i.e. Food Law Code of Practice (England), Food Law Code of Practice (Wales), and Food Law Code of Practice (Northern Ireland).

In 2010, the Food Standards Agency introduced the Food Hygiene Rating Scheme (FHRS) which uses the results from inspections to produce a rating for each of the premises. This rating is based on combining three compliance scores ranging from 25 to 0 or 30 to 0 on
factors including compliance with food hygiene and safety procedures, including food handling practices and temperature control; level of compliance with structural requirements, such as structural cleanliness, ventilation, etc.; and confidence in management procedures. Each of the three elements is scored from 25 to 0 or 30 to 0 and then numerical scores from the intervention-rating scheme in the Food Law Code of Practice are mapped to the six food hygiene ratings to produce an overall food hygiene rating from ‘0’ (Urgent Improvement Necessary) to ‘5’ (Very Good), based on the hygiene standards found at the time (Table 1). The higher the score, the more compliant the premises. These scores are conveyed to proprietors and then made public on the FSA website.

In addition, door or window stickers are provided to help consumers make informed choices about where to eat based on the hygiene rating of the premises. In England, display of the stickers is voluntary, while mandatory display was introduced in Wales in November 2013 and in Northern Ireland in October 2016. The requirements of compliance with food hygiene law for food businesses in England, Northern Ireland and Wales during this study are outlined in the Food Law Code of Practice 2012 (Food Standards Agency, 2012a, 2012b, p. 18). These requirements have been consistent from 2012 through the present and are summarized in Food Law Code of Practice 2017 (Food Law Code of Practice, 2017).

Our work was structured to test the relationship between indicators of food safety, i.e. microbiological sample results and foodborne outbreaks, and several consistent measures of compliance with food hygiene law. The measures generally used include Poor Compliance, Broad Compliance, and Full Compliance. These measures of compliance are used throughout the food regulation sector in the United Kingdom. For simplicity, we have removed Poor Compliance, preserving the other two which are important and align closely to the system used by risk hygiene law. The measures generally used include Poor Compliance, breaks, and several consistent measures of compliance with food hygiene law for food businesses in England, Northern Ireland and Wales during this study are outlined in the Food Law Code of Practice 2012 (Food Standards Agency, 2012a, 2012b, p. 18). These requirements have been consistent from 2012 through the present and are summarized in Food Law Code of Practice 2017 (Food Law Code of Practice, 2017).

1.3. Study aims and objectives

The objectives of our analysis were threefold:

1. To analyze the relationship between compliance with food hygiene law in England, Northern Ireland and Wales, as documented by food hygiene scores and objective measures of microbiological contamination of food samples;
2. To analyze the relationship between the food hygiene rating system and foodborne illness outbreaks;
3. To create a model to show the importance of maintaining food hygiene standards by estimating how the number of foodborne illnesses might change if all premises were Broadly Compliant and, alternatively, how the number of foodborne illness might increase if all premises were Not Broadly Compliant according to the rating scheme described in section 2.4.

1.4. Strategies to reduce foodborne illness: posted hygiene ratings

As the consumption of food outside the home increases, efforts to identify effective strategies to reduce foodborne illness often focus on restaurants and other food premises. Inspections monitor hygiene while educating proprietors about remediation of deficiencies and, if made public, can inform consumers. One strategy for educating consumers, and potentially reducing foodborne illness, is generating food hygiene ratings or scores through food hygiene inspections. However, the strength of the relationship between food hygiene ratings and foodborne illness requires analysis of two key questions. First, do ratings that indicate poor food hygiene affect consumer behavior; i.e., are consumers less inclined to eat in premises with poor posted scores or grades? Second, and most crucial, is there an association between poor hygiene ratings and foodborne illness outbreaks?

Consumers use a variety of observable cues when making decisions about where to eat, including the premise’s image, atmosphere, popularity, and reputation. Accurate assessment of restaurant hygiene can be challenging for consumers, although websites may provide consumers with access to hygiene scores or ratings (Worsfold, 2006; Worsfold, 2007a, 2007b). Readily visible posted hygiene inspection ratings, such as stickers on doors or windows, can be one important source of information about food hygiene for consumers (Henson, Majowicz, Masakure, & et.al, 2006). In a 2017 survey, most respondents in England (54%), Northern Ireland (54%) and Wales (75%) were aware of the Food Hygiene Rating System; the highest levels to date. 44% of respondents in England, Northern Ireland and Wales said they would definitely eat or purchase food from somewhere based on the FHRs rating, showing an increase from previous surveys. Many consumers report often or sometimes checking the FHS rating before deciding to purchase food from an establishment (England 46%, Northern Ireland 47%, Wales 53%), and there has been a significant increase over time in those who avoid a food business due to poor hygiene. Most telling, the majority of respondents (80%) reported that they would not consider buying food from a business with a worse rating than their chosen minimum acceptable rating, which was often 3 or 4 out of 5, although some reported that they would buy food from a very poorly-rated establishment if there were few other alternatives, if it was a place the respondent had frequented before, or if they needed get food quickly. (FHS Tracker Wave 6, October 2017). Similarly, in a survey in New York City that occurred 12 and 18 months after restaurant hygiene posting began, 88% of respondents said they considered hygiene scores when making dining decisions (Wong, McKelvey, Ito, et al., 2015). In short, many consumers report that hygiene ratings, especially when posted rather than accessible only online, affect their choices about where to eat.

Evidence also shows that there are economic effects of posted restaurant hygiene scores. Food premises posting poor hygiene scores, or poor letter grades, suffer negative economic effects as consumers choose to eat elsewhere. This loss of revenue following a poor hygiene rating may motivate proprietors to improve their hygiene practices (Jin & Leslie, 2003) (Jin & Leslie, 2009) (Henson et al., 2006). In fact, since the implementation of posted letter grading in New York City, restaurant hygiene ratings have consistently improved, demonstrating the responsiveness of proprietors to the postings. In short, posted hygiene scores can be a useful tool for informed consumer choices, and can also provide an economic incentive for increasing compliance with food regulations.

However, linking specific cases of foodborne illness to individual food establishments is very challenging. Even knowing consumer
preferences, incubation periods for foodborne illnesses vary, consumers frequently do not report gastrointestinal symptoms to their doctors, and consumers’ memories are unreliable about what, when, and where they ate. In addition, samples of possibly contaminated food are rarely preserved so it is difficult to definitively identify a source. Few studies correlate restaurant hygiene ratings with foodborne illness outbreaks as it is extremely difficult to show a direct causal relationship.

A widely cited study in Los Angeles, California showed that foodborne illness rates decreased after posted hygiene scores were implemented in 1998 (Simon, Leslie, Run, et al., 2003). Another analysis of the impact in Los Angeles showed that posting hygiene grades caused health inspection scores to rise, consumer demand to respond to changes in posted hygiene grades, and the number of foodborne illness hospitalizations to decrease (Jin & Leslie, 2003). Similarly, a study in New York City showed that foodborne illness was down and restaurant revenue was up since the start of letter grading, with salmonella infections falling by 14% during the first full year of letter grading – hitting the lowest level in two decades. Meanwhile, in the neighboring states of Connecticut and New Jersey, salmonella ratings were unchanged. This study shows that food hygiene ratings, when available to consumers, can change consumer behavior and the incidence of foodborne illness (Bloomberg, 2012).

Studies connecting hygiene ratings to foodborne illness are conflicting. A study in Seattle-King County, Washington in 1986–1987 concluded that routine restaurant inspections can predict outbreaks of foodborne illness. In that matched case-control study, restaurants with poor inspection scores and violations of temperature controls, especially for hazardous foods, were five and ten times more likely to have outbreaks than restaurants with better compliance (Irwin, Ballard, Grendon, & Kobayashi, 1989). These appeared to be small toxin-mediated events caused by improper heating, cooling, holding, or food storage. A study done in Tennessee from 1993 to 2000 concluded that mean hygiene scores of restaurants that have foodborne illness outbreaks did not differ from restaurants with no reported outbreaks (Jones, Pavlin, LaFleur, Ingram, & Schaffner, 2004a) (Jones et al., 2004b). A later Tennessee study, from June 2002 to June 2003, showed norovirus was the most common foodborne disease agent in outbreaks (Hedberg, Smith, Kirkland, et al., 2006). Toxin-mediated events, such as those in Washington, are more likely to be responsive to food hygiene efforts than norovirus, possibly explaining the difference in results between the Washington study and the 1993–2000 Tennessee study. Causal relationships are complex, and it is essential to holistically consider the agent, the food item, and the route of transmission in any analysis of inspection results and risk of foodborne illness (Lee & Hedberg, 2016). A study of restaurant inspections and predictive success in 1995 in Miami-Dade County, Florida did not predict outbreaks based on poor inspection scores (Cruz, Katz, & Suarez, 2001). However, a more recent retrospective study in Minnesota showed that more hygiene violations were recorded at restaurants that had foodborne outbreaks of norovirus, Clostridium perfringens infection or toxin-type illness, and Salmonella infection (Petran, White, & Hedberg, 2012a) (Petran, White, & Hedberg, 2012b).

More general concerns about hygiene ratings focus on consistency between inspectors. A 2012 study of over 700,000 health inspections, primarily in San Diego and New York, argued that food hygiene regulatory design, implementation, and practice were frequently flawed (Ho, 2012). Similarly, the 1993–2000 Tennessee study found that the inspection scores varied over time, by region, and by the person who performed the inspection, raising questions about training and consistency of Tennessee inspectors (Jones et al., 2004a, 2004b). To address potential consistency issues in the U.K., the FSA undertakes national consistency exercises to keep variations to a minimum, and commissioned research looking at inspectors and whether their experiences or sequence of observations influenced scores. The research shows inspections are reasonably consistent, accurate, and robust (FSA, 2015) (FSA, 2018). Further work will continue and focus on data-driven initiatives that enhance consistency.

Overall, while some analyses show a connection between hygiene ratings, consumer behavior, proprietor response, and foodborne illness outbreaks, some studies have negative results and some note problems with validity and reliability of the inspections - something which rigorous training and consistent, clear protocols for hygiene inspections can address. Taken together, these studies provide prima facie evidence that many consumers respond to posted scores and use them to make dining decisions, and proprietors can improve their hygiene scores through better hygiene practices and better compliance with food hygiene regulations (Wong et al., 2015).

2. Methods

2.1. Food sampling procedures in England, Northern Ireland and Wales

To examine the relationship between food business compliance with hygiene standards and microbiological contamination of samples, we used food sampling to compare the mean percentage of unsatisfactory samples at premises with different compliance levels. Routine samples were taken by food safety officers, enabling the FSA to have access to many samples and to use the results in our analysis.

Samples results can be recorded in two systems, the United Kingdom Food Surveillance System (UKFSS) and the Local Authority Enforcement Management System (LAEMS). The United Kingdom Food Surveillance System samples for this study were gathered during routine inspection and enforcement activities between January 2013–June 2015 in accordance with the Food Law Codes of Practice issued in April 2012 through April 2015 when it was revised (Food Standards Agency, 2012a, 2012b, p. 18) (Food Standards Agency, 2015). LAEMS, or Local Authority Enforcement Monitoring System, samples were gathered 2009–2010 to 2013–2014.

During the period that samples analyzed in this project were taken, as well as currently, samples taken for microbiological examination were gathered by trained and competent officers according to specific protocols and requirements. Chapter 6 of Food Law Code of Practice (England, April 2012), stipulates that all samples for microbiological examination must be taken by authorized officers who are properly trained, experienced and competent in the appropriate methods and techniques for the specific type of sampling and analysis. Similar Codes of Practice are also published by Northern Ireland and Wales. Local authorities take samples for microbiological examination to be taken by authorized officers who are properly trained, experienced and competent in the appropriate methods and techniques for the specific type of sampling and analysis. Local Authorities “make returns,” or send the FSA their LAEMS data annually. All staff at those laboratories are fully trained and qualified in accordance with national legislation, and all samples are sent to laboratories accredited by the United Kingdom Accreditation Service, according to approved transport methods. (Food Standards Agency, 2013).

While sampling can occur during an inspection, it is often done separately as part of an overall sampling strategy. Therefore, there is not a complete overlap between sampling and inspection data. Either way, the sample result does not form part of the rating, which is made at the inspection’s conclusion. Should any sample prove unsatisfactory, whether or not it is taken at the same time as an inspection, the Local Authority must determine whether a further inspection is needed. In other words, an unsatisfactory sample can trigger an inspection but it is not part of the inspection rating itself, hence the importance of ensuring that FHRS rating used in our analysis was prior to any sampling result (Food Law Code of Practice, England, 2017; Section 5.3.5 Revision of the Frequency).

2.2. Food sampling data sources

The FSA has two different systems which hold sample results. The first source of sample data is the LAEMS (Food Standards Agency, 2017a, 2017b, 2017c, 2017d, 2017e). Each year all Local Authorities
upload data that has been generated from their local system(s) on which they record information on food law enforcement activities. Once uploaded onto the LAEMS, the local authority data are aggregated to the pre-defined categories required by the FSA. The identifier is the same one used throughout LAEMS for that year so samples can be linked to inspection results. Although a sample may be tested for several different microbes, only one result is given which is based on whether it is satisfactory for all microbes tested for or not. One unsatisfactory result for any microbe tested for will be sufficient for the sample to be regarded as unsatisfactory. No detail is given for the microbes tested for nor for which specific test(s), if any, were unsatisfactory.

The second source of sample data available centrally to the FSA is the UK Food Surveillance System (UKFSS). This is a national database for central storage of analytical results from feed and food samples taken by enforcement authorities (Local Authorities and port health authorities) as part of their official controls. These databases hold details on test results for specific microbes and as such is more detailed than LAEMS. However, entering results in this system is not compulsory and was used by mainly Northern Irish and Scottish local authorities (the latter are out of the scope of this work) in terms of microbiological samples during the period of this study.

The data held centrally by the Food Standards Agency was less complete for the period of analysis than the full results held by the Local Authorities. UKFSS data is entered directly by labs, so for that data the FSA received the results directly – this data is complete. However, for most samples, the local authorities get the results and then, at the end of the year, upload sample data into LAEMS. The FSA then gets these data via them but not the lab, so this data does not have all the detail that the local authorities receive.

Our aim was to determine if food hygiene inspection scores are associated with microbiological contamination of food samples. Therefore, this analysis was done in two parts, the first using the LAEMS database which covers more Local Authorities and holds many more sample results although the details of the tests taken is less specific, and the second using UKFSS which has more detailed results but smaller coverage and fewer samples. These two parts can be regarded as separate studies. Although all local authorities responsible for food hygiene or food standards should make a return, some are of low quality so not all are included in the analysis.

2.3. Inspection data

Local authority samples were recorded in the Local Authority Enforcement and Monitoring System, the UK Food Surveillance System, or both. Table 2 summarizes the FHRS rating and the number of premises from which samples were taken for LAEMS and UKFSS. See Table 2.

The analysis using matched FHRS and UKFSS data used 7115 samples from 2774 premises and 26 Local Authorities from Northern Ireland were gathered from January 2013–June 2015. A subsequent reorganization has resulted in just 11 local authorities in Northern Ireland. To ensure that the compliance score given to the premises was not influenced by the result sample in LAEMS, consecutive years were linked with compliance scores taken from the first year and sample results from the second year.

The 7115 samples in UKFSS had all been tested for the same 9 organisms or indicators. These were B. cereus, Bacillus Species, C. perfringens, E. coli, Enterobacteriaceae, Listeria species, Salmonella, Staphylococcus aures and other species, and Total Variable Count 30 (TVC 30). Samples which were not tested for these 9 organisms were excluded from the analysis.

LAEMS holds information on interventions undertaken during the year for each premises, sample results, and the latest inspection scores at the end of year. From these scores, it is possible to calculate either the FHSR rating for premises types covered by the FHRS scheme or the FHSR rating the premises would have had if it was covered by the scheme. In this paper, this is referred to as the “equivalent FHRS score.” This allows us to link the equivalent FHRS results directly with any sample results for that premises, as both are held in LAEMS and use the same premises ID for a specific year. By using the equivalent FHRS score we can also include data from LAs before they joined the FHRS scheme. To ensure that the compliance scores given to premises were not influenced by the microbiological sample results, data from consecutive years was combined and the compliance scores were taken from the first year, and the microbiological sample results from the second year.

For the analysis using the UKFSS database it was not possible to use the equivalent FHRS score from LAEMS as there was no way to link the two datasets. Instead premises in UKFSS and the FHRS database were linked using fuzzy matching using addresses. Fuzzy matching is a technique used to link records from different datasets where they have a common identifier but where the match between the two is not 100% precise. In this case, where we are using addresses, reasons for an imprecise match included differences in spelling, completeness, punctuation or spacing. Fuzzy matching finds correspondences between segments of a text and gives an overall probability of a true match. For each sample at each premises in UKFSS, the most recent FHRS rating for that premises before the sample data was matched from the FHRS database. This ensured the FHRS ratings given to premises were not influenced by the microbiological sample results. The samples from all years of data in UKFSS were considered together as separating into individual years would mean the data set sizes were too small to draw statistically significant conclusions.

It is also important to note the caveat that parameters such as Total Variable Count 30 (TVC30) and Bacillus Species are measures of sub-optimal preparation and food spoilage rather than of contamination with pathogens. Indeed, in some circumstances, spoilage of food may reduce the levels of pathogen contamination. Therefore, analysis was also conducted with TVC30 removed as an example to test whether this changed the results. The results were not significant for the UKFSS work with TVC30 removed. Other parameters such as Enterobacteriaceae and total E. Coli are indicators of faecal contamination, which is in itself an indirect measure of the risk of contamination with pathogens. Also, some of the pathogens most prevalent as the cause of foodborne disease in humans (e.g. Campylobacter and Norovirus) are not routinely tested for so could not be included in this analysis as is discussed below. Only those samples that had been tested for all main 9 specific pathogens or indicators were considered (Table 3). The criteria used to determine whether a sample is unsatisfactory follow the Commission Regulations of the European Communities (Commission, 2005).

2.4. Inspection data analysis

T-tests were used to determine if there was a statistically significant relationship between equivalent FHRS scores (0, 1, 2, 3, 4, 5) and sample test results (satisfactory/unsatisfactory). These FHRS equivalent scores were grouped by the authors into the following categories:
These were: additional undertakings on samples of different food products within the United Kingdom’s food regulation community as measures of compliance with food hygiene law.

For LAEMS data where there were more samples, statistical tests were additionally undertaken on samples of different product types. These were:

- Materials in contact (e.g. cutting board)
- Protein foods (e.g. meat; game such as grouse, pheasant, or rabbit; poultry, dairy; fish; eggs)
- Complex foods (e.g. ices and desserts, prepared dishes, broths and soups)

And different business types:

- Restaurants and caterers
- Retailers

2.5. Outbreak data and FHRS historical data

An outbreak of foodborne illness is defined as two or more linked cases of the same disease. However, data on outbreaks are limited as it is difficult to confirm a direct relationship between microbiological samples from restaurants and outbreaks at those restaurants. Most cases of foodborne disease cannot be linked to a specific source and are therefore classified as sporadic.

The aim of this analysis was to determine if there is a relationship between compliance with food hygiene regulations as demonstrated by a food premises' FHRS rating and the association with an outbreak at that premises. The outbreak data were provided by Public Health England (PHE) and covered the period 2010–2014. There were 298 outbreaks over those 4 years, and 260 occurred at “restaurants,” i.e. catering premises, hotel/guest houses, pubs/clubs, restaurants/cafés/canteens, restaurant/caterers, other, schools/colleges, or take away facilities.

The outbreak data were matched to FHRS historical data using fuzzy matching based on the premises names and addresses in the two datasets. The most recent FHRS rating prior to the outbreak was taken as a measure of compliance at time of outbreak to ensure the FHRS rating was not influenced by outbreak occurrence. Some restaurants could not be matched by name and address, while others could be matched but did not have an FHRS rating prior to the outbreak. Within the period of study (2010–2014), four premises saw repeated outbreaks.

3. Results

3.1. Results from analysis based on LAEMS sample data

Overall, Not Broadly Compliant premises are likely to have a larger proportion of unsatisfactory samples than Broadly Compliant (Fig. 1). Not Fully Compliant premises are likely to have a larger proportion of unsatisfactory samples than Fully Compliant premises (Fig. 2). All the differences are significant at the 99% except the BC and PC figures for 2009 which are not significant.

When the individual elements of the FHRS score were considered separately there was a statistically significant increase in the proportion of unsatisfactory samples for each increase in a premises’ hygiene score, apart from scores between 15 and 20 (Fig. 3). There was a statistically significant increase in the proportion of unsatisfactory samples for each increase in a premises’ confidence in management score (Fig. 4). The mean percentage of unsatisfactory samples in a premises based on structural compliance was smaller for scores 0 and 5 than for all other scores except 25 (Fig. 5).

In addition, Fig. 6 shows the relationship between the proportions of unsatisfactory samples at a premises by product group from which the sample is taken, and food business compliance, as measured by FHRS rating. The difference in unsatisfactory samples between premises rated 5 (Fully Compliant) and the other two groups is statistically significant at the 99% level for all three product groups. The difference between

### Table 3

<table>
<thead>
<tr>
<th>Test</th>
<th>Unsatisfactory (proportion)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B.cereus</td>
<td>0.4%</td>
</tr>
<tr>
<td>Bacillus Species</td>
<td>0.4%</td>
</tr>
<tr>
<td>Clostridium perfringens (C.perfringens)</td>
<td>0.0%</td>
</tr>
<tr>
<td>E.coli</td>
<td>0.7%</td>
</tr>
<tr>
<td>Enterobacteriaceae</td>
<td>5.0%</td>
</tr>
<tr>
<td>Listeria species</td>
<td>0.0%</td>
</tr>
<tr>
<td>Salmonella</td>
<td>0.0%</td>
</tr>
<tr>
<td>Staphylococcus aureus and other species</td>
<td>0.1%</td>
</tr>
<tr>
<td>Total Variable Count 30 (TVC30)</td>
<td>5.1%</td>
</tr>
</tbody>
</table>

* Fully Compliant (FC) - Premises for which compliance levels have been assessed as equivalent to an FHRS rating of 5 at their most recent food hygiene inspection. Fully Compliant premises are a subset of the Broadly Compliant category i.e. all Fully Compliant premises are also Broadly Compliant.
* Broadly Compliant (BC) - Premises for which compliance levels have been assessed as equivalent to an FHRS rating of 3, 4 or 5 at their most recent food hygiene inspection.

These two categories were selected as they are commonly used within the United Kingdom’s food regulation community as measures of compliance with food hygiene law.

For LAEMS data where there were more samples, statistical tests were additionally undertaken on samples of different product types. These were:

- Materials in contact (e.g. cutting board)
- Protein foods (e.g. meat; game such as grouse, pheasant, or rabbit; poultry, dairy; fish; eggs)
- Complex foods (e.g. ices and desserts, prepared dishes, broths and soups)

And different business types:

- Restaurants and caterers
- Retailers

![Fig. 1 LAEMS Microbiological Sample Analysis – By Year (Not Broadly Compliant vs. Broadly Compliant) (All significant at 99% except 2009 which is not significant).](image-url)
premises rated 3, 4 and premises rated 0, 1, 2 (Not Broadly Compliant) was not statistically significant for any of the product groups. Although the different product groups have different proportions of unsatisfactory samples due to different levels of risk associated with the product groups, the relationship with FHRS rating is in the same direction i.e. the proportion of unsatisfactory samples is lower for more compliant premises for all three product groups (Fig. 6).

Fig. 7 shows the relationship between the proportion of unsatisfactory samples at a premises by business type and FHRS rating. The results illustrated are for 2013–2014 but the pattern is similar in other years. The differences in unsatisfactory samples between Not Broadly Complaint and Broadly Compliant premises, and Not Broadly Compliant and Fully Compliant premises, are statistically significant at the 99% level for both restaurants/caterers and for retailers. The proportion of unsatisfactory samples in a premises is generally smaller for retailers than for restaurants and caterers at each compliance level (Fig. 7).

3.2. Analysis based on UKFSS sample data

The mean percentage of unsatisfactory samples in premises rated 5 (6.4%) is lower than in premises rated 0–4 (8.2%), and this is statistically significant at the 95% level. The mean percentage of unsatisfactory samples in premise rated 3, 4, 5 (7.2%) is lower than in premises rated 0, 1, 2 (9.1%), but this is not statistically significant. Thus, the mean proportion of Unsatisfactory Samples in a Premises was higher in those premises that were Not Fully Compliant when compared with those premises that were Fully Compliant (Fig. 8). Similarly, the mean proportion of Unsatisfactory Samples in a Premises that was Not Broadly Compliant was higher than those which were rated Broadly Compliant, but this was not a statistically significant result (Fig. 9).

3.3. Results of foodborne illness outbreak analysis

The outbreak analysis showed that, while there was limited data about outbreaks and few outbreaks could be definitively matched to specific premises, there was a statistically significant relationship between a premises’ compliance category and the chance of an outbreak, with Broadly Compliant premises having fewer outbreaks per restaurant years than Not Broadly Compliant premises (Fig. 10).

4. Discussion & impact of results

4.1. Summary of findings

The relationship between compliance with food law, documented via hygiene ratings, and both microbiological contamination of food samples and foodborne outbreaks indicates an association between better scoring establishments and reduced food safety risks. The strength of this analysis is the separate sources of data that all provide results pointing in the same direction – namely higher levels of compliance with food hygiene law is associated with lower risk for the two
measures of food safety. While the outbreaks result is only at 95% statistically significant, when used side by side with the LAEMS microbiological results - which are often at 99% and consistent across various splits such as year, product type, restaurant/retailer - we have compelling evidence. This evidence is further supported by the microbiological results using UKFSS data which trends the same way.

4.2. Model estimating impact of increased compliance with food hygiene regulations

To assess the likely impact of our findings, we created a model that estimates how the number of foodborne illnesses may change if all premises were Broadly Compliant and, alternatively, what it might increase to if all premises were Not Broadly Compliant. This was to show the importance of maintaining food hygiene standards. This model has two main assumptions. First, that the associations found in this paper between broad compliance and measures of food safety are a cause and effect, and second, that these measures of food safety are directly proportional to foodborne illness.

From the results of the three studies; the LAEMS sample study, the UKFSS sample study, and the Outbreak analysis; it was possible to estimate the difference in risk between Broadly Compliant and Not Broadly Compliant premises. From the LAEMS analysis the mean proportion of unsatisfactory samples in a premises in 2014–15 was 21.5% for Not Broadly Complaint premises and 15.3% for Broadly Compliant Premises. This gives a relative ratio of 1.41 (21.5% divided by 15.3%). Not Broadly Compliant premises were 41% more likely to have an unsatisfactory result than Broadly Compliant premises. From the outbreaks analysis, Not Broadly Complaint premises were 110% more likely to have an outbreak than Broadly Compliant Premises. For UKFSS the difference between Broadly Complaint and Not Broadly Compliant premises was not statistically significant - similar relative risk figures could be produced for the different between Fully Complaint and Not Fully Compliant premises instead.

Currently about 94% of premises are Broadly Compliant. If this were to increase to 100% then, based on the relative risk rates above, we might expect foodborne illness acquired eating out to decrease by between 2.5% and 6.2%. This may not seem that large because most business are already compliant. However more strikingly, if no (0%) premises were Broadly Compliant then we might expect foodborne illness acquired eating out to increase by between 37% and 97%.

The absolute numbers of cases from these estimates will depend on how many foodborne illness cases are acquired eating out. Although the FSA estimates there are 1 million cases of foodborne illness in the UK per year, due to the difficulty of tracing foodborne illness to source it is less clear what proportion of these 1 million cases are acquired by people eating out as opposed to being acquired from consuming food prepared in the home. The FSA has recently commissioned a systematic review to better understand what research has already been done to estimate these proportions, but results from this work were not available at the time of writing this paper. In the meantime, we created scenarios that assume that the proportion of foodborne illness acquired in food premises is either 25% or 50% of the total, i.e. 250,000 and 500,000 cases. For the first scenario, a simple calculation suggests that if all premises were Broadly Compliant foodborne the number of cases would drop from 250,000 to between 234,500 and 244,000 cases a year.
based on the relative risk ratios produced in the LAEMS and Outbreak analysis respectively. If all premises were Not Broadly Compliant the numbers would increase from 250,000 to between 343,500 and 492,500 cases per years. For the second scenario, the equivalent changes would be a drop from 500,000 to between 469,000 and 488,000 cases if all premises were to become Broadly Compliant and an increase from 500,000 to between 687,000 and 985,000 cases if all premises were to become Not Broadly Compliant.

Our model assumes changes in foodborne disease are similar to rates of unsatisfactory samples and that all other factors are equal. While not all pathogens are tested for regularly (e.g. *Campylobacter* or norovirus are rarely tested for), the majority of unsatisfactory samples are due to unsatisfactory results for indicator organisms, such Enterobacteriaceae. These are used as general indicators of food hygiene. The assumption is that premises which have unsatisfactory results for any of the pathogens and indicators organisms tested for are more likely to be at higher risk from all pathogens and foodborne disease more generally. While this assumption is clearly not perfect, it allows us to make some estimation as to the likelihood of risk reduction for compliant premise, enabling us to develop a model that is indicative, although not definitive, in its quantification of the possible positive impact of improving food hygiene on reducing foodborne illness.

5. Conclusions, limitations and future work

This analysis demonstrates an association between compliance with food hygiene law and levels of microbiological contamination in food samples drawn from food premises. Further, the model quantifies the possible positive impact on reducing foodborne illness by improving food hygiene and compliance with food hygiene law. Detecting public health benefits of this type is extremely complex. Our analysis strongly suggests the important link and the real potential benefits of vigilant surveillance and stronger policies to raise hygiene standards.

Regarding limitations, first, we recognize that food hygiene inspection scores represent only a snapshot in time. Inspections are unannounced, and it possible to visit premises on a particularly bad or unusually good day and generate a rating that does not reflect the usual conditions. Sewers sporadically overflow; rodent infestations erupt suddenly and may disappear only briefly after pest control efforts. Yet visual inspections, when done by well-trained inspectors adhering to a systematic process and using a comprehensive list, are still one of the best means currently at our disposal for assessing hygiene. Knowing that the ratings align with microbiological data further bolsters our reliance on hygiene inspections as a first line of consumer protection.

Second, the microbiological samples in this analysis were not taken as part of a structured survey design, but during the day-to-day work of hundreds of food safety officers. However, this drawback is more than compensated by the large numbers of samples available, namely 37,304 for the LAEMS and 7115 from UKFSS. Also, the analysis using the LAEMS data was further handicapped by not knowing the microbiological tests applied and relying on an overall unsatisfactory result. However, this was not the case with the UKFSS dataset where tests undertaken were known and this data gave similar results.

Third, although this study analyzed a wide variety of illness-causing organisms, it did not assess for norovirus or *Campylobacter* – organisms causing illnesses with symptoms similar to those caused by other pathogens. *Campylobacter* is the most common foodborne pathogen in the UK, and the FSA has recently led a successful reduction campaign (Food Standards Agency, 2018; Food Standards Agency, 2016). Norovirus is the most common intestinal illness pathogen in the US and the UK. The
US Centers for Disease Control estimates that about 50% of foodborne illness outbreaks are caused by norovirus (CDC, 2016) (Food Standards Agency, ND) (Scallan, Hoekstra, Angulo, & et.al, 2011). Although most cases of norovirus are due to person-to-person contact, many are due to contaminated food and drink. Food can become contaminated with norovirus when being grown, shipped, or prepared; surfaces such as door handles or faucets can also be contaminated with norovirus; and workers can spread norovirus to customers through food handling. Moreover, a recent study of human noroviruses found no association between hygiene levels and norovirus contamination on surfaces and suggests that swabbing is necessary to reveal actual surface contamination of norovirus (Maunula, Ronnvist, Aberg, Lunden, & Nevas, 2017). Given all these complex factors, further studies and expanded reduction strategies for Campylobacter and norovirus are warranted but are beyond the scope of this paper.

Fourth, we cannot underestimate the human element in food safety and the need for additional studies, guidelines, and regulations regarding food workers. The food safety culture – how people behave with respect to food safety, the behavior of workers, and the way the organization operates - directly affect the risk of harm to consumers (Nayak & Waterson, 2017) (Gould, Rosenblum, Nicholas, & et. al, 2013). Those behaviors include willingness to abide by food safety regulations, and may reflect lack of knowledge about transmission of microorganisms (Maunula et al., 2017). In particular, improper hand washing or improper use of gloves, dirty uniforms, and contaminated handles or knobs, spread microorganisms (Garayoa, Abundancia, Diez-Leturia, & Vitas, 2017). Further, food service workers often work when they are sick, contributing to the spread of illness. A recent CDC study in the U.S. showed that 12% of food workers had vomiting or diarrhea during two or more work shifts in the previous year (Centers for Disease Contro, 2017). This was less likely in premises that had a policy for workers to inform their manager when they were ill, had on-call workers who could come in, and had a more experienced manager. Similarly, a recent study of UK restaurants showed 32% of chefs and catering students had worked within 48 h of suffering from diarrhea or vomiting (Jones et al., 2017). Knowledge of, and willingness to abide by, food safety regulations, and the need to work when suffering from gastrointestinal symptoms – an endemic issue in food worker culture – must be addressed. While beyond the scope of this paper, more research into effective education, practices, and policies pertaining to sick food workers and food safety culture is clearly warranted.

Fifth, critics of posted hygiene ratings may point to incidents of bribery that may affect the credibility of ratings. Occasionally bribes have been initiated by proprietors or solicited by inspectors. Media reports from the 1970s discuss a single incident of attempted payoff in Washington DC and several more in New York City (Barker & Pala, 1972), (Raab, 1977) (Purdum, 1988). Notably, these occurred before hygiene ratings were visibly posted outside establishments, therefore prior to the economic pressures imposed by posting. In 1998, one inspector was arrested in Los Angeles, prompting creation of a hotline (Los Angeles Times, 1998). Bribery allegations resurfaced in New York in 2011 and 2016, and in Jacksonville, Florida in 2012. (Reddy, 2011)
Like other public agencies, health departments in U.S. cities have policies prohibiting bribes, and we found no evidence that bribery increased after visible posting or that bribery in restaurant hygiene ratings was more common than bribery in other areas, such as building inspections (Celona & Eustachewich, 2017) (Sewell & Zahniser, 2011). Moreover, we found no evidence of bribery in the U.K., where this study was conducted and where initiatives evaluating reliability and consistency are in place.

Finally, data sharing could be improved within the broader food control system which spans farms to fork. Relatively new food safety management systems and accredited assurance schemes provide important data which currently may not be optimally integrated into the food control system in the U.K. Visual inspections provide important evidence, but integrating industry assurance data, as well as making better use of social media to inform and educate consumers - as was recently done in the successful UK Campylobacter initiative - could help strengthen food safety as a system (Food Standards Agency, 2017a, 2017b, 2017c, 2017d, 2017e) (Barnes, 2017).

While these limitations pose challenges to connecting compliance with food hygiene law, microbiological contamination, and foodborne illness outbreaks, they do not change the clear trend substantiated by separate data sources. Better compliance with food law results in better hygiene ratings for food premises. When these ratings are generated by trained food inspectors following rigorous guidelines, these ratings provide important data which currently may not be optimally integrated into management systems and accredited assurance schemes provide important data which currently may not be optimally integrated into the food control system which spans farms to fork. Relatively new food safety management systems and accredited assurance schemes provide important data which currently may not be optimally integrated into the food control system in the U.K. Visual inspections provide important evidence, but integrating industry assurance data, as well as making better use of social media to inform and educate consumers - as was recently done in the successful UK Campylobacter initiative - could help strengthen food safety as a system (Food Standards Agency, 2017a, 2017b, 2017c, 2017d, 2017e) (Barnes, 2017).

While these limitations pose challenges to connecting compliance with food hygiene law, microbiological contamination, and foodborne illness outbreaks, they do not change the clear trend substantiated by separate data sources. Better compliance with food law results in better hygiene ratings for food premises. When these ratings are generated by trained food inspectors following rigorous guidelines, these ratings have a strong relationship to a lower likelihood of foodborne illness outbreaks. Strategies to encourage or nudge compliance are clearly warranted. The evidence-based recommendations for food hygiene can help ensure that food service establishments are, indeed, as clean as they look.

Conflicts of interest

The authors have no conflicts of interest in conducting this work.

Acknowledgements

The authors wish to acknowledge Public Health England for providing data on outbreaks. We would like to thank Robin Clifford at the Food Standards Agency who provided statistical help and advice. Finally, we greatly appreciate the many food safety officers whose diligent work enhances food safety while making this analysis possible.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.foodcont.2018.08.034.

References


