

# Impact of a Letter-Grade Program on Restaurant Sanitary Conditions and Diner Behavior in New York City

Melissa R. Wong, MPH, Wendy McKelvey, PhD, Kazuhiko Ito, PhD, Corinne Schiff, JD, J. Bryan Jacobson, MPH, and Daniel Kass, MSPH

Restaurant food safety is increasingly important, with almost half of the US food dollar spent on restaurant food<sup>1</sup> and about one third of caloric intake from foods prepared outside the home.<sup>2</sup> In New York City (NYC), residents eat out nearly 1 billion times each year.<sup>3</sup> Although most diners do not get sick, foodborne pathogens cause millions of preventable illnesses in the United States annually.<sup>4</sup> The exact proportion of restaurant-attributable foodborne illness is unknown, but national surveillance in the United States found that two thirds of reported foodborne outbreaks from 1998 through 2008 occurred in the restaurant or deli setting,<sup>5</sup> and consumption of food prepared outside the home has been linked to an increased risk of sporadic foodborne diseases.<sup>6</sup>

Regular inspection of restaurants for food safety is a core function of local health authorities, guided by the US Food and Drug Administration (FDA) Food Code.<sup>7</sup> Although all states have sanitation codes modeled after the FDA Food Code,<sup>8</sup> implementation methods vary by jurisdiction. The NYC Department of Health and Mental Hygiene (hereafter, Health Department) is charged with inspecting restaurants, coffee shops, bars, nightclubs, employee or university cafeterias, bakeries, and fixed-site food stands (hereafter, restaurants). Its inspection program uses a scoring system to measure compliance with the NYC Health Code, which is updated regularly to maintain consistency with the FDA Food Code and the New York State Sanitary Code. Restaurants are entitled to an impartial review of inspection results by an administrative tribunal, which can improve an assigned score and reduce associated monetary fines.

Before letter grading, the Health Department aimed to inspect restaurants at least once per year and imposed monetary fines for violations cited at inspections. Inspection results were available on the Health Department Web site. However, financial disincentives and the

**Objectives.** We evaluated the impact of the New York City restaurant letter-grading program on restaurant hygiene, food safety practices, and public awareness.

**Methods.** We analyzed data from 43 448 restaurants inspected between 2007 and 2013 to measure changes in inspection score and violation citations since program launch in July 2010. We used binomial regression to assess probability of scoring 0 to 13 points (A-range score). Two population-based random-digit-dial telephone surveys assessed public perceptions of the program.

**Results.** After we controlled for repeated restaurant observations, season of inspection, and chain restaurant status, the probability of scoring 0 to 13 points on an unannounced inspection increased 35% (95% confidence interval [CI] = 31%, 40%) 3 years after compared with 3 years before grading. There were notable improvements in compliance with some specific requirements, including having a certified kitchen manager on site and being pest-free. More than 91% (95% CI = 88%, 94%) of New Yorkers approved of the program and 88% (95% CI = 85%, 92%) considered grades in dining decisions in 2012.

**Conclusions.** Restaurant letter grading in New York City has resulted in improved sanitary conditions on unannounced inspection, suggesting that the program is an effective regulatory tool. (*Am J Public Health*. 2015;105:e81–e87. doi:10.2105/AJPH.2014.302404)

Web site posting were insufficient to drive improvements across the industry, with most restaurants cited for multiple public health hazards. Mean inspection scores and restaurant sanitary conditions were stagnant (D. Kass, email communication, February 2009).

In an effort to improve restaurant food safety and increase transparency of inspection information, the Health Department launched its letter-grade program on July 27, 2010. The program uses public disclosure of inspection scores in the form of letter grades at point of decision-making; a more finely tuned, risk-based inspection schedule; and financial incentives to encourage high food-safety standards. It began after an 18-month planning process that included a public announcement of the intent to begin letter grading; meetings with restaurant industry representatives, food safety experts, and regulators from a jurisdiction with a restaurant sanitary grade program; promulgation of 2 regulations subject to notice and comment; and training and education for restaurateurs. The process was covered by the

media, and by July 2010, restaurateurs were aware of the program and anticipating the launch.<sup>9,10</sup>

We evaluated the impact of the restaurant letter-grade program by assessing (1) hygiene and food-safety practices as characterized by inspection outcomes before and after program implementation and (2) public response to the program measured by 2 population-based telephone surveys.

## METHODS

The NYC restaurant inspection program has been using a point system to score inspections since 2005.<sup>11</sup> Presence and severity of violations contribute to an inspection score. Under the grading program, an inspection score of 0 to 13 points is in the A-range; 14 to 27 points is in the B-range; and 28 or more points is in the C-range. Restaurants scoring 0 to 13 points on the first inspection of their inspection cycle (initial inspection) are issued an A grade. Restaurants not earning an A grade on initial



inspection receive a full reinspection no less than 7 days later. The grade card is issued based on the reinspection score. The initial inspection and any reinspection together are an “inspection cycle.” Upon completion of an inspection cycle, there is an interval before the next cycle. Restaurants earning an A grade on initial inspection of a cycle are inspected in 11 to 13 months. Restaurants scoring 28 or more points on either initial or reinspection of a cycle have a 3- to 5-month interval. The remaining restaurants scoring 14 to 27 points on either initial inspection or reinspection of a cycle have a 5- to 7-month interval.

Before the grading program was launched, the Health Department aimed to conduct at least 1 inspection in all restaurants annually. Restaurants scoring 28 or more points received a follow-up compliance inspection about 1 month later. A score of 28 points or higher could result in a restaurant being placed on a twice-yearly inspection schedule. Administrative violations (e.g., expired permit) were included in the scoring system before implementation of letter grading, but they are not included under the grading program.

Health Department inspectors cite violations with standardized forms on handheld computers. They also collect data on restaurant descriptors such as cuisine, service method to customer (e.g., wait service, counter service), venue description (e.g., restaurant, bar), and chain status (15 or more national outlets). Inspectors are trained in the classroom and under an experienced inspector in the field before they are allowed to work independently.

### Data Analysis

We analyzed preadjudicated inspection scores and points for violations cited on initial or reinspections conducted between July 27, 2007, and July 26, 2013. We subtracted administrative violation points from pregrading inspection scores to make pregrading scores more comparable with postgrading.

We calculated measures that used “most recent initial inspection” among restaurants in business as of July 27 in each year. “Most recent initial inspection” is used in crude analyses to depict a restaurant’s usual sanitary conditions closest to the specified period end date. We consider initial inspections of a cycle the best indicator of usual sanitary conditions

because they occur at the longest interval after the previous inspection and they are unannounced to operators. Crude metrics were percentage of restaurants scoring in the A-, B-, or C-range; percentage scoring 40 points or higher (85th percentile score on initial inspection in the program’s first year); median inspection score; and average points for specific violations or violation groups. Average violation points characterize both presence and severity of violations over time.

We assessed performance on reinspection of a cycle by calculating percentage of restaurants scoring in the A-range on reinspection among those with B-range or C-range initial inspection scores. We tracked the percentage of restaurants with A, B, or C grades on a cycle that went on to earn an A grade on their next cycle.

We modeled the probability of scoring 0 to 13 points (A-range score) across all initial inspections in all 43 448 restaurants by fitting a binomial regression model that included 5 indicators of time: 13 to 36 months before grading (reference), 0 to 12 months before grading, 0 to 12 months after grading, 13 to 24 months after grading, and 25 to 36 months after grading. We fit restaurant random intercepts to account for repeated observations and variation across individual restaurants. We used indicator variables to adjust for potential confounding by season of inspection (January–March, April–June, July–September, October–December), because pest and holding temperature–related violations increase during the warmest season and the distribution of inspection date varied over time.<sup>12</sup> We did not think chain restaurant status was a potential confounder because the distribution before and after grading remained constant, but we included it to estimate the probability that a chain restaurant scored 0 to 13 points relative to a nonchain. We also ran the fully adjusted model for the subset of restaurants with inspections in the first and last year of the study ( $n = 7059$ ) to evaluate whether improvement differed among the most stable restaurants.

To assess whether an excess or deficit in the frequency of inspection scores around grade cut-offs could have biased our results, we estimated the underlying (unbiased) smooth frequency distribution of scores by fitting a generalized additive model with penalized splines<sup>13</sup> and used the smoothed distribution

to estimate the “bias-corrected” percentage of A-range scores in the postgrading period. The percentage of A-range scores across initial inspections in the postgrading period dropped only slightly from 30.7% to 27.4% upon correction. We therefore deemed it unnecessary to correct for potential bias resulting from an excess or deficit of scores around grade cut-offs.

We conducted analyses in SQL Management Studio 2008 R2 (Microsoft, Redmond, WA), SAS version 9.2 (SAS Institute, Cary, NC), and R version 3.0.1 (R Project, Vienna, Austria).

### Public Perception Surveys

The Health Department worked with Baruch College Survey Research (BCSR) to conduct 2 English/Spanish bilingual telephone surveys in July 2011 and February 2012 to assess public perceptions of the grading program. Landline samples on a random-digit-dial design and respondents were selected randomly within the household; cell phones were randomly selected from a mobile number database for NYC county telephone numbers. Respondents were screened for NYC residency and age of 18 years or older.

In July 2011 and January 2012, 502 and 511 adults completed surveys, respectively. Based on the American Association for Public Opinion Research (AAPOR) standard definitions,<sup>14</sup> response rates were 26% and 22%, and cooperation rates were 60% and 51%, respectively. AAPOR response rates incorporate estimates of the proportion of respondents of unknown eligibility that might have been eligible. Data were weighted to the US Census 2009 American Community Survey to ensure the samples represented the age, gender, race, Hispanic origin, and borough distribution of NYC adults. Confidence intervals (CIs) for proportions were calculated with SAS version 9.3 (SAS Institute, Cary, NC).

### RESULTS

Approximately 24 000 restaurants operate in NYC on any given day. A total of 43 892 restaurants were in business at some point between July 2007 and July 2013, and 46% (20 005) of those were in operation at some point both before and after grading. During the 3 years before grading, 31 226 restaurants



operated. Of those, 41% were newly opened for business and 36% went out of business. In the 3 years since grading began, 32 700 restaurants operated. Of those, 39% were newly opened for business and 27% went out of business (Table 1). The distribution of restaurant types was nearly identical before and since grading was instituted.

**Inspections**

The percentages of A-range scores on recent unannounced initial inspection were similar during the 3 years before grading and have improved since grading. The proportion of restaurants with A-range scores went from 28% in July 2008 to 31% in July 2010, with an additional increase to 46% by July 2013 (Figure 1). With more restaurants achieving A-range scores after grading, the median initial inspection score went from 21 points as of July 2008 and 20 points as of July 2010 to 17 points as of July 2013.

After we controlled for chain status, season of inspection, and correlation within restaurants, the probability of attaining an A-range score on an unannounced initial inspection among all restaurants increased 26% (success ratio [SR] = 1.26; 95% CI = 1.22, 1.31) by the 2-year mark (Table 2). The SR increased at the 3-year mark to 1.35 (95% CI = 1.31, 1.40). Compared with the warmest season (July–September), the other seasons exhibited higher SRs, with the highest (SR = 1.30; 95% CI = 1.26, 1.35) in the coldest season (January–March). The SRs for the subset of restaurants in business during the whole period (data not shown in Table 2) were slightly higher—1.32 (95% CI = 1.25, 1.40) and 1.41 (95% CI = 1.33, 1.49) for the 2- and 3-year mark, respectively. Chain restaurants showed a high SR for both all restaurants (SR = 3.46; 95% CI = 3.31, 3.61) and the subset of restaurants operating during the whole period (SR = 3.79; 95% CI = 3.54, 4.07).

Certain critical food safety violations contributed fewer average points in July 2013 compared with the 2 years before grading (Table 3). In July 2012, the average points given to all restaurants declined substantially for evidence of any type of vermin (rats, mice, flies, or roaches), inadequate hand-washing facilities, and no food safety–certified supervisor on-site. Points given for improper storage or use of equipment or utensil and inadequate food worker hygiene also declined to a lesser extent. These overall point reductions were maintained in July 2013. Meanwhile, average points increased for improperly maintained food contact surfaces, and the points given for inadequate protection of food from contamination, cross contamination, and holding food at improper temperatures increased slightly (Table 3). Although average points for temperature and cross-contamination violations increased slightly, average severity of cited violations decreased (data not shown).

We observed inverse trends for C-range scores on recent initial inspection over time. The proportion of C-range scores decreased from 29% as of July 2008 and 27% as of July 2010 to 22% as of July 2013 (Figure 1). The percentage of extreme C-range (≥ 40 points)—scoring restaurants dropped from 14% in the year before grading to 13% in July 2011, dropping to 7% in July 2012, and increasing to 9% in July 2013, while the 80th percentile decreased from 36 points in July 2008 to 30 points in July 2013.

Three years after grading, more restaurants corrected unsanitary conditions observed on initial inspection of most recent inspection cycle. In July 2013, 45% of restaurants requiring reinspection earned A grades upon reinspection, up from 34% in July 2011. Likewise, there was a decrease in the proportion of poorly performing restaurants that did not improve on reinspection (28+ point scores on both initial and reinspection). The proportion of restaurants that scored poorly on both initial and reinspection dropped from 28% as of July 2009 to 22% as of July 2013.

When we tracked performance from inspection cycle to inspection cycle, we found that 80% and 79% of A-grade restaurants maintained their A grade on their next cycle at 2 and 3 years after grading, respectively. Among B-grade restaurants, 53% and 54%

**TABLE 1—Restaurant Characteristics Before and After Grading: New York City, NY, 2007–2013**

Characteristic	Operating Between July 27, 2007, and July 26, 2010 (n = 31 226), No. (%)	Operating Between July 27, 2010, and July 26, 2013 (n = 32 700), No. (%)
<b>Borough</b>		
Bronx	3 267 (10)	3 222 (10)
Brooklyn	7 538 (24)	8 047 (25)
Manhattan	11 828 (38)	12 584 (38)
Queens	7 307 (23)	7 552 (23)
Staten Island	1 286 (4)	1 295 (4)
<b>Chain restaurants</b>		
Yes	3 393 (11)	3 627 (11)
No	27 833 (89)	29 073 (89)
<b>Restaurant type<sup>a</sup></b>		
Wait service restaurant or diner	8 589 (31)	10 564 (33)
Quick-service establishment with take-out or limited seating	12 706 (46)	14 443 (45)
Baked goods, ice cream, or cafe only	3 539 (13)	4 050 (13)
Bar or wine bar	1 206 (4)	1 289 (4)
Cafeteria and banquet-style service or deli buffet	992 (4)	1 236 (4)
Food service at attraction	534 (2)	662 (2)
Missing	3 660	456

Note. The city restaurant letter-grading program began on July 27, 2010. All gradable restaurants (or pregrading equivalent) in operation between July 27, 2007, and July 26, 2013, included.

<sup>a</sup>Percentage excludes missing values.