Review

Two Decades of Nursing Home Compare: What Have We Learned?

Medical Care Research and Review
2021, Vol. 78(4) 295–310

DOI: 10.1177/1077558720931652
journals.sagepub.com/home/mcr
SAGE

R. Tamara Konetzka[1], Kevin Yan[1], and Rachel M. Werner[2]

Abstract

Approximately two decades ago, federally mandated public reporting began for U.S. nursing homes through a system now known as Nursing Home Compare. The goals were to provide information to enable consumers to choose higher quality nursing homes and to incent providers to improve the quality of care delivered. We conduct a systematic review of the literature on responses to Nursing Home Compare and its effectiveness in meeting these goals. We find evidence of modest but meaningful response by both consumers and providers. However, we also find evidence that some improvement in scores does not reflect true quality improvement, that disparities by race and income have increased, that risk-adjustment of the measures is likely inadequate, and that several key domains of quality are not represented. Our results support moderate success of Nursing Home Compare in achieving intended goals but also reveal the need for continued refinement.

Keywords

Nursing Home Compare, public reporting, nursing home, quality, information

Quality of care in nursing homes has long presented a policy challenge (Institute of Medicine, 1986). Publicly reporting information about nursing home quality is one of the most prominent policy efforts aimed at improving nursing home quality in the past two decades. The Centers for Medicare and Medicaid Services (CMS) has implemented public reporting through Nursing Home Compare (NHC), a web-based report card on all Medicare- and/or Medicaid-certified nursing homes in the United States. As in the case of other health care report cards, the goal is to provide consumers with information about quality that will inform their choice of nursing homes and increase the probability of going to a high-quality provider. In turn, providers are hypothesized to use the quality information for benchmarking and to face an incentive to score well on the report cards, lest they lose market share. Numerous studies have examined these assertions and the potential for unintended consequences, while media reports have at times cast skepticism on the report card's validity (Thomas, 2014). In this article, we systematically review the empirical evidence on the effectiveness of NHC that has emerged in the two decades since the system was initiated.

History of NHC

In 1998, the first version of the NHC website was launched with information limited to nursing home regulatory deficiencies. In 2000, the available information was expanded to include nurse staffing data. While these quality measures were publicly available, they were not widely disseminated or publicized. Then, in 2002, through the Nursing Home Quality Initiative, CMS released what became widely known as NHC, a web-based guide detailing quality of care at over 17,000 Medicare- or Medicaid-certified nursing homes (CMS, 2002). It included 10 clinical quality measures, six of which measure quality for long-stay residents with chronic care needs and four of which measure quality for patients in postacute care with skilled needs. It also

[1]The University of Chicago, Chicago, IL, USA
[2]University of Pennsylvania, Philadelphia, PA, USA

Corresponding Author:
R. Tamara Konetzka, Department of Public Health Sciences, Department of Medicine, Biological Sciences Division, The University of Chicago, 5841 South Maryland Avenue, Room W255, MC2000, Chicago, IL 60637, USA.
Email: konetzka@uchicago.edu

included information on staffing and rates of regulatory deficiencies. Prior to the 2002 launch, a 7-month pilot was conducted in six states (Colorado, Florida, Maryland, Ohio, Rhode Island, and Washington). The expanded version of NHC was launched nationally in November 2002, allowing consumers to compare quality measures across nursing homes nationwide (CMS, 2002; Harris & Clauser, 2002). The release of this web-based information about nursing home quality was actively promoted to consumers in 2002, with the hope that consumers would use this information to help choose a nursing home.

In June 2008, CMS announced it would make significant changes to the NHC rating system by introducing five-star summary ratings. Starting on December 18, 2008, CMS began publicly rating each nursing home with a star rating ranging from one to five stars, which is still used today. The star ratings give consumers a simplified, composite look at nursing home quality. Nursing home star ratings are based on quality in three domains: health inspections (based on scope and severity of health deficiencies found during the two most recent state inspections and number of repeat visits needed to confirm the correction of deficiencies), staffing (based on case-mix adjusted measures of total nursing hours per resident day and registered nurse hours per resident day), and quality measures (averaged over 4 quarters and originally based on 10 clinical quality outcomes measures). At its implementation, the five-star rating system was one of the largest experiments using a summary rating in a public health care report card. CMS continues to invest resources into this public reporting effort, updating the reported scores on a quarterly basis, amending the list of included measures, improving the interface, and raising awareness among consumers and other stakeholders.

New Contributions

NHC constitutes a major policy initiative designed to improve consumer choice and quality of care for a particularly vulnerable population—individuals who need nursing home care. Findings from evaluations of public reporting in other health care sectors, while related, may not apply to the nursing home sector because of several unique characteristics: a population with high rates of cognitive impairment and other comorbidities; a sector with a long history of quality problems; the long-term nature of the care, which raises the importance of location and proximity to family members; and, for some, the planned nature of nursing home entry allows the time needed to gather information about nursing homes before making a choice. Thus, distributing information about quality may be simultaneously more important and more challenging than in other sectors. Although numerous studies have evaluated the effectiveness of NHC at various points of time and from various perspectives, the evidence about its impact to date has not been systematically reviewed and summarized. Several prior reviews of public reporting included studies of Nursing Home Compare among studies in other health care sectors, but these reviews are generally dated, preceding the five-star system, and many used restrictive inclusion criteria such that few nursing home studies were included (Berger et al., 2013; Mukamel et al., 2014; Totten et al., 2012; Vukovic et al., 2017). This article fills a key gap, providing a comprehensive overview of empirical evidence on the effects of NHC including effects of the five-star system. In addition to providing insights about consumer and provider behavior in response to public reporting of quality, the goal of our review is to inform the evolution and refinement of the NHC system.

Conceptual Framework

We conceptualize the effects of public reporting using tools from the economics of information (Arrow, 1963). Public reporting is intended to work through two main mechanisms: consumer response and provider response. For consumers—whom we define broadly to include prospective nursing home residents, their families, or other individuals involved in the decision to be admitted to a particular nursing home, such as discharge planners and referring providers—demand for nursing home care depends primarily on health status, out-of-pocket price, availability of potential alternatives such as informal caregivers, and perceptions about the quality of care they will receive. However, health care markets are subject to imperfect and asymmetric information, whereby consumers are unable to assess the level of clinical quality of their potential nursing home choices with certainty. Finding valid information may entail high search costs. They may therefore use noisy but lower cost signals for quality such as nonprofit status or "hotel" amenities. Public reporting of health care quality is intended to solve, at least in part, the problem of asymmetric information by making information about quality more readily available to consumers and lowering search costs. Thus, all else equal, consumers should be more likely to choose high-quality providers and less likely to choose low-quality providers under public reporting.

The degree of consumer response to public reporting may differ by the type of nursing home care needed. People searching for short-term, postacute care for rehabilitation after an acute hospitalization may be better able to respond than those searching for long-term, chronic care for functional and/or cognitive impairment. Postacute stays are generally covered by Medicare,

making out-of-pocket price less of an issue, and postacute patients tend to be younger and less cognitively impaired, making the use of quality information more feasible (Werner et al., 2012). On the other hand, decisions about postacute care at the end of an acute hospitalization may be more rushed and subject to bed availability.

This demand response by consumers is expected to prompt a provider response. In the absence of public reporting, imperfect and asymmetric information may lead providers to choose a level of quality where the marginal cost of providing quality is less than the social optimum that would be achieved under perfect information. In other words, providers react to consumers' inability to judge quality by underproviding quality. In the presence of public reporting, which partially solves the information issue, providers face a demand curve that is more responsive to the level of quality. Thus, providers are expected to increase the level of quality in order to maintain or increase market share. Even in the absence of competition and concerns about market share, report cards may provide information to providers against which they can benchmark their quality improvement efforts; nonprofit nursing homes, for example, may include the provision of high-quality care in their objective functions directly.

Whether the hypothesized effects of public reporting are realized depends on a number of assumptions about consumers and providers (Casalino et al., 2007; Mukamel, Weimer, et al., 2007; Werner & Asch, 2005): that consumers know about and can access and understand the reported information; that consumers have a choice of providers willing and able to accommodate them nearby such that they can act on the information; that providers know how to improve quality; and that providers engage in good-faith efforts to broadly improve quality and not just improve scores. The hypothesized effects of public reporting also depend on assumptions about the measures themselves: that they are useful, valid, and reliable representations of quality and cover domains that are salient to consumers. Violations of these assumptions could lead to a dampening of the effects or to unintended consequences, namely: (a) heterogeneity and potential disparities in benefits; (b) selection of healthier residents that could lead to access problems for less healthy residents; (c) targeted improvement that results in no improvement or worsening quality in nontargeted areas, also known as "teaching to the test"; and (d) possible gaming of the system and changes in documentation such that scores improve without true improvements in quality. We review the evidence on each of these unintended consequences as well as direct evidence of consumer and provider response to NHC.

Method

We conducted a systematic review of the literature published between 1998 (when the first version of NHC appeared) and July 2019 to find peer-reviewed empirical articles assessing NHC. A search on PubMed using the search terms "Nursing Home Compare," "Nursing Home Quality Initiative," "Nursing Home Five Star Rating," and "Nursing Home Report Card" resulted in 2622 potential articles. We then conducted a title and abstract review to narrow the pool to articles that included an empirical (quantitative or qualitative) analysis of some aspect of NHC, retaining 55 of the 2,622 articles. We excluded articles that were primarily essays or commentary. We also excluded articles that simply used the NHC measures as outcomes in an analysis but whose intent was not to evaluate or inform NHC. These decisions were independently conducted by two of the authors and any discrepancies were resolved through discussion among all three authors. Finally, we conducted searches using the same search terms in Web of Science and Google Scholar to identify articles not indexed in PubMed, and we checked reference sections of all the articles already identified for citations to articles we might have missed. These efforts resulted in an additional 16 articles that met our inclusion criteria, bringing the total number of articles to 71. Our search strategy is displayed in Figure 1.

We read the 71 included articles to assess methodology and findings. Study designs were evaluated for rigor and the extent to which they enabled causal inference. We did not exclude studies based on weak study designs, as even simple descriptive studies of trends can be informative; however, we placed more weight on studies with more rigorous designs (those using randomized or quasi-experimental designs such as longitudinal panel designs, regression discontinuity, difference-in-differences, or instrumental variables) when drawing conclusions across studies. We classified articles according to whether they mainly assessed consumer response, provider response, or issues related to the presentation and validity of the reported measures. A complete list of included studies, with review categories and study designs, can be found in Appendix Table 1 in the Supplemental Material, available online.

We structure the results of our review according to the broad categories in our conceptual model. First, we review the evidence on consumer response to NHC, followed by provider response to NHC. Next, we review the indirect effects or unintended consequences of NHC, including disparities, selection, teaching to the test, and potential gaming of the data. Finally, we review the evidence on the presentation, validity, and content of the measures themselves.

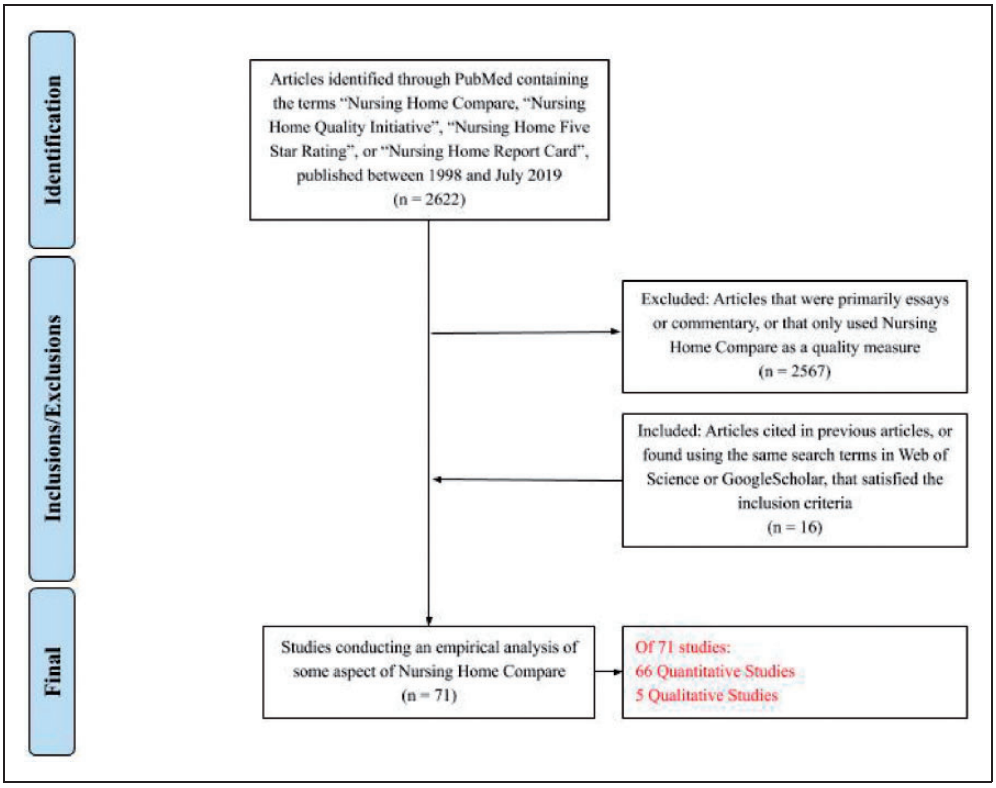

Figure 1. Search strategy for review of nursing home compare.

Results

Consumer Response to NHC

As consumer use of health care report cards is central to their intended purpose, a first-order question is whether consumers use them and respond to their publication. Early evaluations of NHC used surveys and descriptive pre–post trend analysis to assess whether consumers appeared to be shifting toward higher quality nursing homes after NHC was launched. Stevenson (2006) found that occupancy rates of highly rated nursing homes increased slightly after the limited 1998 launch, suggesting a positive but quite small response. Castle (Castle, 2009a, 2009b) conducted a mail survey of family members of newly admitted nursing home residents to probe their use of the 2002 version of NHC, finding that 12% of respondents reported using it.

Three studies employed more rigorous designs to assess whether market share shifted as a result of NHC prior to the five-star system. These studies examined the effects of NHC on market share, defined as the percentage of residents market-wide who go to each nursing home, a common way to measure consumer response. Grabowski and Town (2011) employed a difference-in-differences model based on comparing pilot states, which implemented the 2002 changes 7 months earlier, to nonpilot states, and vice versa when the remaining states implemented. Focusing on long-stay, chronic care residents, they found no significant effect of NHC on market share for four of the five quality measures they studied. In contrast, Feng Lu (2012) used a similar design but found significant effects on market share for most of the five quality measures studied. These two analyses focused on different quality measures, which may explain their divergent results, but in both cases the significant effects were small. Werner et al. (2012) used a discrete choice estimation with an instrumental variable to examine choice of nursing homes for postacute care and found significant market share changes in response to higher scores on some (but not all) postacute quality measures on NHC. Because most postacute care patients are covered by Medicare and therefore can consider quality without considering price, it is plausible that postacute care residents respond more to NHC than the long-stay, chronic care residents. Still the effects for postacute care residents were also small. Overall, quantitative evidence on consumer use of NHC prior to the five-star system is consistent with modest use of the system, with small changes in market share potentially concentrated among postacute care patients.

The evidence on consumer response to NHC after the 2008 launch of the five-star summary score is sparse but more consistent. An analysis of descriptive trends among long-stay residents showed a significant decline in the

number being served in one-star nursing homes and a significant increase in the number being served in five-star nursing homes during the first 2 years of the new system (Konetzka, Grabowski, et al., 2015). A more rigorous analysis of these shifts during the same time period, analyzed using a discrete choice model and controlling for other nursing home characteristics, found that one-star facilities typically lost 8% of their market share and five-star facilities gained 6% of their market share (Werner et al., 2016). Finally, Perraillon et al. (2019) used a regression discontinuity design to assess the causal effect of the switch to star ratings in the first 6 months after they were released, estimating an effect at each star level relative to the star level below. They found that nursing homes with five, four, and three stars received 2% to 5% more admissions than similar nursing homes with one fewer star, but that there was little effect among lower quality nursing homes (two vs. one star). Of note, all three of these studies found practically meaningful effect sizes that were larger than the somewhat inconsistent effects found prior to the five-star system. In terms of encouraging the use of NHC by consumers, the five-star summary ratings appear to be effective.

Some of the same studies that examined consumer response on average also examined heterogeneity of response. One might assume that public reporting would be more effective in moving market share in more competitive markets, that is, when consumers have more choice. Similarly, nursing homes with high occupancy may not be able to increase market share due to capacity constraints, even if very highly rated. Prior to the five-star ratings, the evidence on these factors was mixed. Grabowski and Town (2011), who did not find overall effects on market share among long-stay residents, also did not find that market competition influenced changes in market share after the 2002 launch. In contrast, Werner et al. (2012), who found an overall market share response among postacute patients, found that the effect was driven by nursing homes that were not already at high occupancy. After the five-star ratings were implemented, several studies demonstrated the expected effects: Gains in admissions or market share for highly rated facilities were more pronounced in more competitive markets and among lower occupancy facilities (Perraillon et al., 2019; Werner et al., 2016).

Insights Into Consumer Response. Beyond studies examining whether and to what extent consumers have responded to NHC, research has begun to unpack the mechanisms behind these findings, drawing from both quantitative and qualitative studies. An early national survey showed that use of the Internet was common in searching for information about nursing homes, but that only 12% of respondents could remember using NHC specifically (Castle, 2009a, 2009b). Furthermore, only about a third of those used NHC for quality information, whereas others were looking just for names and addresses of nearby facilities. Awareness of NHC could be enhanced through state health care websites, almost all of which now include links to NHC, but the links are sometimes difficult to find, offer little explanation, and use inconsistent terminology (Liu & Lu, 2015).

Qualitative, interview-based studies conducted in limited geographical areas after the launch of the five-star system produced similar findings to those in the survey, revealing substantial use of the Internet but low awareness and use of NHC itself (Konetzka & Perraillon, 2016; Schapira et al., 2016). Konetzka and Perraillon (2016) found that many consumers, when shown the NHC website, were suspicious of the source of the data and wondered if the nursing homes themselves influenced or created the site, that is, it was not immediately apparent to them that the site was produced by the government. Schapira et al. (2016) also found that consumers wanted more information on the source of the data. Quality was found to be only one factor among many considered by consumers, with recommendations from doctors, the availability of specialized services, distance to family and friends, and the availability of a bed (especially for those on Medicaid) often dominating quality as a constraint or driver of choice (Konetzka & Perraillon, 2016). Limited supply of high-quality nursing homes in low-income neighborhoods remains a difficult challenge (Konetzka, Grabowski, et al., 2015). Overall, these studies provide some explanations for a modest consumer response to NHC—low awareness, mistrust of the data, competing considerations, and constrained choice sets.

Provider Response to NHC

Complementary to the issue of consumer response is the question of whether providers respond to public reporting by improving the quality of care delivered, as this is also a fundamental goal of public reporting policies. Similar to consumer response, the evidence on provider response prior to the five-star system was mixed. Early descriptive studies found small improvements in some of the long-stay quality measures but not in others, with more consistent improvement in the short-stay quality measures (Castle et al., 2007; Castle et al., 2008; Zinn et al., 2005).

Several more rigorous studies used different forms of quasi-experimental design to identify plausibly causal improvement in the clinical quality measures. One study used a survey of nursing home administrators to assess whether and how they were responding to their scores and then interacted an indicator for these

responses with changes over time in five selected quality measures; that study found significant improvements linked to nursing home actions in three of the five measures (Mukamel, Weimer, et al., 2008). Two studies that used the difference in timing of implementation between pilot and nonpilot states to identify effects of NHC found conflicting evidence on whether nursing homes had improved their scores due to NHC (Feng Lu, 2016; Grabowski & Town, 2011). A study that examined changes in vaccination rates after measures of vaccination rates were added to NHC used rates among community-dwelling elderly people as a comparison, and found no effect attributable to NHC (Cai & Temkin-Greener, 2011). Finally, one study that focused on the postacute care measures constructed a comparison group using small facilities, which were exempt from publishing measures on NHC. The study found that nursing homes improved on all three postacute care measures in an absolute sense, and on two of the three relative to small facilities (Werner, Konetzka, Stuart, et al., 2009). A follow-up study examined the extent to which improvements in the average level of postacute care quality were due to within-facility improvement over time versus shifts in market share toward higher quality nursing homes, concluding that both were contributing (Werner et al., 2010). Overall, the evidence points to providers increasing quality in response to the pre-five-star version of NHC, but with the most consistent improvements concentrated among postacute quality measures.

Evidence on quality improvement due to NHC under the five-star system is relatively sparse. Several studies documented substantial upward trends in within-facility reported quality over time (Han et al., 2017; Konetzka, Grabowski, et al., 2015; Werner et al., 2016), consistent with government-sponsored reports showing dramatic improvement in scores (Abt Associates, 2014). In a study of antipsychotic medication use, which was added to NHC and to the five-star calculation in 2013, Bowblis et al. (2015) found large and significant improvements relative to facilities not subject to reporting of the measure. Although the absolute trends in improved scores under the five-star system are substantial, with the exception of the latter study on antipsychotic use, these trends have not been rigorously studied as a causal effect of NHC.

Provider response to NHC may include nonquality responses, such as changes in the price charged to private-pay residents or changes in payer mix and profits. If a nursing home scores well on NHC, the incentive to improve further may be blunted, but the high score is a form of advertising that may simultaneously increase demand and the price that a nursing home is able to charge to private-pay residents. A few studies examined price responses. Clement et al. (2012) examined the immediate aftermath of the 2002 launch of NHC using a pre–post design in one state and found some increases in price, but these were among low-scoring facilities, making it difficult to attribute the response to NHC. Using a more rigorous design, Feng Lu (2012) found no evidence of price changes in response to NHC. However, after the five-star system was implemented, Huang and Hirth (2016) used a difference-in-differences model to reveal increases in private-pay price for highly rated nursing homes relative to low-rated homes, especially in more competitive markets. Finally, two studies examined the relationship between NHC quality scores and profits using difference-in-differences and instrumental variables models. Nursing homes that scored well or improved substantially on NHC were found to have higher subsequent profit margins, largely as a result of increasing occupancy and the share of Medicare days (Park et al., 2011). However, the increases in profit margins were not necessarily large enough to result in further quality improvement (Park & Werner, 2011).

Just as in the consumer studies, some of the studies that examined provider response on average also examined heterogeneity of response. If public reporting is more effective in moving market share in more competitive markets, then providers in more competitive markets should also have greater incentive to improve. Similarly, nursing homes that are already at high occupancy should have less incentive to improve. Empirical evidence on provider response generally supports these hypotheses. Several studies found that the likelihood of a nursing home improving in at least some quality scores and taking actions toward quality improvement was higher in more competitive areas (Castle et al., 2007; Castle et al., 2008; Grabowski & Town, 2011; A. S. Kim, 2016a; Zinn et al., 2008; Zinn et al., 2010) and in areas with low occupancy (Castle et al., 2007; Castle et al., 2008; A. S. Kim, 2016a). In addition, one early survey-based study of the 2002 version of NHC found that administrators of nursing homes in the bottom 20% of scores reported being more likely to take immediate action to improve scores than administrators of homes that scored better (Zinn et al., 2008). Somewhat surprisingly, no difference was found by ownership of nursing homes (for-profit vs. not-for-profit; chain vs. independent) in terms of the quality response to NHC (Feng Lu, 2016; Zinn et al., 2005; Zinn et al., 2010).

Insights Into Provider Response. Beyond studies examining whether and to what extent providers have responded to NHC, several surveys and qualitative studies lend insights into why, how, and when providers respond. In the early years after the 2002 launch, few nursing home administrators believed that NHC played an important role in consumers' decisions, based on lack

of inquiries from consumers about the scores (Mukamel, Spector, et al., 2007; Zinn et al., 2008). Accordingly, although most providers reported checking their scores, only a third said they would use the scores for quality improvement (Castle, 2005).

This began to change quickly over time, with providers increasingly reporting that they not only checked their scores but investigated the cause of low scores (Mukamel, Spector, et al., 2007). After the five-star system was launched, administrators began to closely monitor their scores, use high scores for advertising, and develop strategies to improve scores. Nursing home chains often employed management-level staff to monitor scores and address problems across member facilities (Perraillon et al., 2017). The level of effort and resources devoted to monitoring and improving scores is proportional to the extent to which key constituencies seem to care about the scores (Perraillon et al., 2017; Zinn et al., 2010)—where key constituencies include not only consumers, but increasingly, accountable care organizations, managed care companies, and hospitals looking for postacute partners. Nursing homes with high occupancy rates report that they are less concerned about their scores (Perraillon et al., 2017).

Even if report cards provide incentives to try to improve scores, it is not always clear how to improve quality in each area. Evaluation of a pilot project using the clinical quality measures suggested that improvement on most of the measures involves changing processes of care, but that nursing homes may need technical assistance in order to know how to improve (Kissam et al., 2003). Little research to date has been conducted on which specific quality improvement strategies are effective to improve on the NHC measures.

Unintended Consequences: Disparities

A common concern with public reporting is that the information will disproportionately benefit those who are already advantaged (consumers with higher education, higher income, and/or who do not identify as a racial or ethnic minority), potentially increasing disparities in access to high-quality care. An alternative hypothesis is that public reporting acts as an equalizer, providing information to less advantaged consumers that would otherwise be difficult to find. Studies to date provide some evidence to support the first hypothesis. Werner et al. (2012) found that postacute patients with a higher level of education were more likely than those with lower education to respond to the 2002 launch of NHC. Konetzka, Brauner, et al. (2015) found that, although there was an overall shift toward five-star facilities and away from one-star facilities among long-stay residents after the five-star system was implemented, this shift was substantially smaller for individuals dually enrolled in Medicaid and Medicare (relative to those enrolled in Medicare only). One study used Oaxaca decomposition methods to find that education, health needs, and distance to high-quality nursing homes explained significant portions of the difference in the star rating of nursing home that Medicaid and non-Medicaid residents were admitted to (Sharma et al., 2019).

Surveys and qualitative studies found complementary evidence on heterogeneity of use by income and race. An early survey found that comprehension of the quality information among respondents was high on average but was lower for individuals with lower incomes or whose family member resided in a high-Medicaid facility (Castle, 2009a, 2009b), whereas one of the later qualitative studies found that respondents with lower income or Black race expressed less desire to use NHC when told about it (Schapira et al., 2016). Overall, these studies are consistent in suggesting that NHC may be widening disparities even as consumers shift to higher quality facilities on average.

Mirroring results on the consumer side indicating that consumers with lower income, on Medicaid status, and of minority race and ethnicity are less likely to respond to NHC, examination of provider performance under NHC suggests strong differences in performance by provider characteristics. A common concern is that providers who serve low-income consumers will lack the resources needed for quality improvement and, subsequently, will not reap the financial benefits of scoring well, fueling a cycle of increased disparities. Unroe et al. (2012) found that one-star nursing homes had older patients who were more likely to be Black, have Medicaid, and have more comorbidities relative to five-star nursing homes. Konetzka and Gray (2017) found a clear and monotonic gradient between star rating, percentage of Medicaid, and the income level of the neighborhood in which a nursing home is located, with low income and high Medicaid associated with lower star ratings. Although it is difficult to disentangle true differences in quality from inadequate risk-adjustment, controlling for these factors would not change the star ratings of most facilities (Konetzka & Gray, 2017). One study revealed an effect of NHC on disparities in access as well: After the 2002 launch of NHC, high-scoring nursing homes that were capacity-constrained had significant increases in private-pay admission and decreases in Medicaid-admission compared to low-scoring homes. Thus, even as Medicaid recipients potentially use NHC to choose higher quality nursing homes, those same nursing homes are using their increased market demand to exclude Medicaid recipients and improve their payer mix (He & Konetzka, 2015). Overall, the evidence on provider-level disparities under NHC points to the policy reinforcing resource

differentials and having adverse effects, but the causal relationships have not been fully established.

Unintended Consequences: Selection

A second common concern with public reporting policies is that imperfect risk-adjustment will incentivize providers to engage in selection of patients that will make their scores look better, sometimes referred to as "cherry-picking" or "cream-skimming." This type of selection is difficult to identify in quantitative data, as it is difficult to separate changes in case-mix that are due to provider selection versus consumer-driven changes in nursing home choice, or "sorting." Sicker consumers, for example, may disproportionately avoid low-scoring providers because they have more at stake than healthier consumers.

Using a nursing home fixed effects model, an initial look at the selection issue after the 2002 launch of NHC revealed a significant but small decline in the percentage of patients entering nursing homes with pain or memory limitations that was concentrated in low-scoring nursing homes, but no overall compelling evidence of selection (Mukamel et al., 2009). Werner et al. (2011) examined the same issue using a difference-in-differences model based on differences in timing of implementation for pilot states versus other states, finding that patient sorting was likely to be occurring: High-scoring nursing homes were increasingly admitting higher risk patients while low-scoring nursing homes were increasingly admitting lower risk patients, at least with respect to pain; there was also some evidence that pain was being downcoded on average but little evidence of selection was found. The lack of evidence for substantial selection at admission under NHC is perhaps not surprising, given that both Medicaid and Medicare reimburse on a per diem basis and the payment can be earned long before a resident becomes eligible to be counted in NHC scores. Konetzka et al. (2013) used multiple quasi-experimental designs to find that some selection was occurring on discharge instead, that is, postacute residents were more likely to be hospitalized just before being counted in NHC scores if they were at risk of scoring poorly. In more recent years, this incentive may have been tempered by an increased policy focus on reducing hospital readmissions.

Overall, although there is some evidence of nursing homes selecting healthier patients, the effects are small. This may be due to competing priorities, as found in interviews with nursing home administrators and staff: Nursing homes want to maximize revenues, which requires filling beds, and performing well on NHC is a secondary consideration. In addition, when selection occurs, it is often to avoid litigation, and litigious residents are not necessarily those who would make the nursing home score poorly on NHC (Perraillon et al., 2017).

Unintended Consequences: Teaching to the Test

A third common concern with public reporting is that targeted areas of quality will improve at the expense of areas that are important but not rewarded. This concept of "teaching to the test" is also related to Holmstrom and Milgrom's theory of multitasking in the context of optimal contracts for employee performance when an employee engages in multiple tasks (Holmstrom & Milgrom, 1991). In the context of public reporting and other types of quality improvement incentives, the theory implies that measuring and rewarding quality in some areas can either have positive spillovers to other areas (if there are complementarities in production) or may harm quality in other areas (if resources are shifted to targeted areas). NHC targets clinical quality measures that are arguably specific and may engender specific quality improvement strategies that could "crowd out" quality improvement along other dimensions.

Research on NHC to date lends considerable support to the idea that nursing homes teach to the test. Surveys of administrators after the 2002 launch found that, rather than investing in new quality improvement programs, nursing homes were more likely to refocus existing programs and rearrange priorities to try to improve NHC scores (Mukamel, Spector, et al., 2007; Zinn et al., 2008). Several national studies using quasi-experimental designs followed, showing that untargeted areas of quality stagnated or declined as targeted areas improved (Feng Lu, 2012; Werner, Konetzka, & Kruse, 2009; Werner, Konetzka, Stuart, et al., 2009). One study focused in on the substantial decline in physical restraint use after restraints was included in the 2002 launch of NHC, finding a concurrent increase in antipsychotic medication use (a possible chemical restraint) prior to antipsychotic use being added to NHC (Konetzka et al., 2014). Mukamel et al. (2010) followed a different strategy to capture the idea broadly, examining changes in the ratio of clinical to "hotel" expenditures after NHC. They found that once clinical measures were reported, resources were shifted to clinical expenditures, especially among nursing homes that were lower quality, for-profit, members of a chain, had lower occupancy, or were in more competitive areas. Overall, these studies provide substantial evidence that providers respond to the incentives under NHC by focusing on targeted measures, potentially ignoring other important areas of quality.

Unintended Consequences: Gaming of the Data

Although provider performance on measured outcomes in NHC appears to have improved, policy makers, the media, and the public are increasingly skeptical about whether the improved performance is real, that is, truly

reflects improvements in residents' outcomes. Several media articles drew national attention to the skepticism about the measures, highlighting cases where consumers felt misled by high ratings, choosing nursing homes that fell far short of expectations and where family members experienced adverse events (Thomas, 2014, 2015). The concern stems in part from the fact that two of the three main quality domains under the five-star system, the clinical quality measures and staffing ratings, are based on self-reported data from nursing homes, raising the possibility of improving scores through changes in documentation used for the quality measure without accompanying changes in the underlying construct being measured. Documentation changes could be legitimate increases in attention and accuracy, or could constitute "gaming"—outright falsification of data, or employing questionable strategies to improve scores. In contrast, inspections, the third main quality domain used in the five-star rating, is not based on self-reported data and is therefore given the most weight in the five-star formula (Abt Associates, 2014).

Empirical efforts to identify the degree of gaming in NHC have used a common, if indirect, strategy, implemented in multiple ways. The strategy is based on the assumption that "real" improvement should be at least somewhat correlated either with the mechanisms to improve quality or with related (untargeted) measures of quality where spillovers from true quality improvement would be expected. In terms of mechanisms, one study found that the correlation between outcomes included in the NHC quality measures and associated measures of process decreased after public reporting began, such that very little of the postpublication change in outcomes could be explained by changes in process (Werner et al., 2013). Similarly, very little of the reported improvement in staffing ratios under the five-star system could be corroborated by changes in staffing costs, indicating that either staffing levels were being inflated or jobs were being reclassified as providing direct care (Sharma et al., 2017); this is consistent with earlier survey evidence that poor-scoring facilities were more likely to revise job descriptions (Zinn et al., 2008). In terms of correlation with other measures of quality, one study found that differences in ratings for inspections were correlated with differences in complaints (an untargeted area of quality), but differences in the self-reported domains were not correlated with complaints, suggesting ratings inflation in the self-reported domains (Han et al., 2016); this inflation was also correlated with a financial incentive to score well (Han et al., 2017). Similarly, the correlation between star ratings and potentially preventable rehospitalization rates for post-acute patients became weaker after the five-star ratings were published, indicating potential inflation in the five-star ratings (Ryskina et al., 2018).

Several studies that examined narrower, more direct documentation issues also found evidence of a NHC effect. Werner et al. (2011) found a decrease in average documented pain levels for postacute patients on admission to nursing homes after pain was included as a quality measure in 2002; because other, correlated characteristics of admitted patients were not changing, the decrease was likely due to a coding change. Konetzka, Brauner, et al. (2015) found that nursing home residents with dementia were more likely to be coded as having end-stage disease after NHC was launched, given that many of the quality measures exclude residents with end-stage disease from the denominator.

Finally, qualitative evidence drawn from interviews with nursing home administrators and nursing staff underscores and is consistent with the quantitative evidence that gaming is occurring (Perraillon et al., 2017). Nursing home representatives reported substantial coding-related efforts to improve scores on the clinical quality measures, often led by a centralized director of clinical operations in the case of chain facilities. Strategies included asking residents about pain only after pain medications were given; timing assessments such that adverse events would not be counted in the look-back period for some measures; and counting nondirect-care staff in nursing hours. All of these strategies would potentially improve scores without changing care.

Presentation, Measurement, and Content of NHC

Conceptually, consumer and provider responses to NHC are inherently tied to the presentation of the quality information, the validity of the measures, and whether important areas of quality are excluded. A fairly robust body of evidence exists on each of these topics.

Presentation. Prior to the five-star system, although consumers generally understood the information contained in each quality measure included in NHC (Castle, 2009b), parts of the way NHC presented information were confusing. In particular, several studies found that consumers had a difficult time reconciling that most of the clinical quality measures are adverse events and, thus, lower scores represent higher quality (Gerteis et al., 2007; Schapira et al., 2016). In addition, consumers often had trouble with graphical representations of the data as opposed to text labels such as "high quality" or "low quality" (Castle, 2009b; Gerteis et al., 2007). The addition of the five-star summary rating was more intuitive and familiar, despite difficulty understanding how the rating was calculated (Schapira et al., 2016). The fact that consumer response increased substantially after the five-star system was implemented supports the idea that

a simplified and familiar composite measure constitutes a more effective presentation than the prior system (Perraillon et al., 2019; Werner et al., 2016).

The five-star summary rating was, in part, a solution to presenting numerous individual measures with no systematic way for consumers to aggregate across measures and rank choices. In calculating a summary measure, however, it was necessary for CMS to decide which measures to include how much weight to give each measure. The complex formula for the star ratings reflects these decisions, informed with the help of a technical expert panel to triangulate which measures were both important and reliably measured. An alternative view is that consumers themselves should decide on these weights. Mukamel et al. developed a tool for hospital patients in need of a nursing home placement to create their own ranking of local choices by choosing those measures that they valued most. A randomized trial of the tool showed that, relative to using the standard NHC, consumers using the tool prioritized measures differently than the standard NHC, had higher satisfaction with the process, left the hospital slightly sooner, and were more likely to choose a higher quality facility even in terms of the NHC rating (Mukamel, Amin, Weimer, Ladd, et al., 2016; Mukamel, Amin, Weimer, Sharit, et al., 2016; Weimer et al., 2019). In related work, the same team showed that a contingent valuation approach, asking consumers hypothetically to trade off travel time for nursing home quality, is a valid approach to constructing NHC weights from the consumer perspective (Weimer et al., 2019).

Validity of the Measures. A perhaps inevitable area of concern with the NHC measures is inadequate risk-adjustment. Most providers feel that risk-adjustment could be improved and that lack of adequate risk-adjustment means that their nursing home is not accurately represented on NHC (A. S. Kim, 2016b; Perraillon et al., 2017). The staffing measures and many of the clinical quality measures incorporate some risk-adjustment, while the inspections rating is not risk-adjusted. Research to date on NHC risk-adjustment focuses exclusively on the clinical quality measures. Despite extensive testing of the clinical quality measures in the 1980s and 1990s (Zimmerman et al., 1995), before NHC was launched, vulnerabilities remain in part due to inherent tradeoffs between the desire for validity and the desire to avoid complexity. Studies have examined the effects of including additional factors in risk-adjustment, finding that resident case-mix can explain 50% of the variation in the quality measures that are not risk-adjusted and that unadjusted measures have high false negative rates (Li et al., 2009; Li et al., 2010); that nursing home rankings would change significantly if some risk factors were added using appropriate statistical modeling (Arling et al., 2007; Mukamel, Glance, et al., 2008); and that the inappropriate antipsychotic medication measure in NHC may substantially overstate rates due to inadequate risk-adjustment (Lucas et al., 2014). Overall, although an inherent tradeoff exists between better risk-adjustment and complexity, these studies demonstrate that fairly simple additions to variables used to risk-adjust the clinical quality measures could significantly improve the validity of nursing home rankings in this domain.

Beyond risk-adjustment, several studies criticized the validity of the staffing measures. Some of the criticism focused on the data source used, until recently, to calculate the staffing measures, which drew from approximately annual, point-in-time estimates of staffing levels collected during inspections; those data could be subject to inflation if nursing homes increased their staffing at the time of anticipated inspections and potentially did not reflect long-term performance (Sharma et al., 2017). This problem has been alleviated with the use of ongoing payroll-based staffing data in NHC (Geng et al., 2019). Nonetheless, there are remaining concerns that average staffing ratios are just one perspective on the adequacy of staffing; the reported staffing ratios may not reflect fluctuations in staffing, the experience and skill of staff within each job title, or turnover (Snyder et al., 2019).

A common way to assess the validity of reported quality measures is to examine either the correlation among reported measures or the correlation with broad but unreported measures of quality. Conceptually, one would not expect perfect correlation among reported measures when measures represent different domains of quality, as a provider could legitimately be good in some domains and not in others. At the same time, very low correlations across measures or with broad measures of quality may be interpreted as potential validity problems with the measures themselves, due to the expectation of at least some joint production across domains of quality (e.g., good management should lead to higher quality across domains). Several studies found low correlations among reported measures (Brauner et al., 2018; Saliba et al., 2018). Studies found little correlation between reported measures and quality of life (S. J. Kim et al., 2014) and resident or family satisfaction (Calikoglu et al., 2012; Williams et al., 2016), but it is not apparent that reported measures should be correlated with resident or family satisfaction if measures are valid, given that they are distinct domains of quality.

In terms of broad measures of quality that arguably should be correlated with NHC measures, two measures often studied were the rate of admissions or readmissions to the hospital, measures that were added to NHC in the past several years, and mortality. The evidence connecting NHC to these broader measures is

mixed. Studies examining individual quality measures (not the composite star ratings) generally found weak or no correlations between the quality measures and hospitalization/mortality (Fuller et al., 2019; Neuman et al., 2014; Saliba et al., 2018; Snyder et al., 2019; Unroe et al., 2012; Xu et al., 2019).

Although there is limited direct evidence on the validity of the five-star composite rating, several pieces of evidence support that the star ratings capture meaningful information about nursing home quality, at least at the extremes. First, studies have found the overall star rating to satisfy face validity in terms of nursing home characteristics typically associated with higher or lower quality: five-star nursing homes are more likely to be nonprofit, independent, have more Medicare and private-pay residents, have more educated residents, and a lower percentage of non-White compared with one-star homes, which are more likely to be for-profit, to be part of a chain, have a high Medicaid census, and to serve a less educated, less affluent, and higher percentage of minority population (Konetzka & Gray, 2017; Perraillon et al., 2019; Unroe et al., 2012). Second, several studies that examined the relationship between the overall star rating and hospital admissions/readmissions and mortality found a meaningful relationship whereby nursing homes with higher star ratings had lower rates of these adverse outcomes (Cornell et al., 2019; Unroe et al., 2012). In the most rigorous of these studies, Cornell et al. used variation in postacute patient distances to nursing homes of different star ratings within ZIP codes over time as an instrument to identify effects that account for patient selection into nursing homes of different quality levels; they found sizable effects whereby being admitted to nursing homes of higher star ratings for postacute care led to lower mortality and lower risk of long-term nursing home admission, with somewhat mixed results on hospital admission (Cornell et al., 2019).

A final important issue with respect to validity of the included measures is whether they validly differentiate between levels of quality. Several studies found that only ratings at the extremes (five stars vs. one star) were predictive of rehospitalization for heart failure (Unroe et al., 2012), of resident and family satisfaction (Williams et al., 2016), and patient safety outcomes (Brauner et al., 2018) with no meaningful difference among nursing homes rated two, three, or four stars. However, these studies were descriptive and reflected only associations. The more rigorous study by Cornell et al. found more monotonic relationships between star rating and several adverse outcomes—mortality and long-term nursing home admission—although not for hospital readmissions. Thus, although the five-star composite measure seems to have face validity and predictive validity at the extremes, the evidence is mixed as to whether the star ratings are helpful to consumers choosing among nursing homes that are closer to average.

In summary, after two decades of use and refinement, the validity of NHC and its measures seems to have improved with the addition of the five-star composite measure, claims-based quality measures, and payroll-based staffing data. However, as the studies in this section show, the measures are still in need of refinement in terms of risk-adjustment, lack of correlation among measures that should be correlated, and a composite measure that can better distinguish modest increments in quality.

Content: What Is Missing?. Even if all included measures are valid, response to any public reporting system may be limited if important domains of quality are missing, domains that may be strong drivers of decisions. Two key domains have surfaced in this regard: resident and family satisfaction and end-of-life care, neither of which is directly measured in NHC. Lack of a satisfaction domain is due to the fact that the administrative data sets underlying the NHC measures do not contain information on satisfaction, making additional data collection required if satisfaction were to be reported. One study calling for the addition of a satisfaction domain found that resident and family satisfaction were only weakly correlated with each other, such that one could not substitute for the other (Williams et al., 2016). Several qualitative studies of consumers found that resident and family satisfaction were key items that consumers would like to see added to NHC (Konetzka & Perraillon, 2016; Schapira et al., 2016), consistent with the fact that recommendations from family and friends still weighs heavily in consumer decisions about where to seek nursing home care. Although our review did not find parallel evidence that consumers were looking for information on end-of-life care, nursing homes often serve as the final home for many residents prior to death. Accordingly, researchers have noted the complete lack of measures in this area and have proposed several measures using existing administrative data that could fill this gap (Mukamel et al., 2012; Mukamel, Ladd, et al., 2016). The proposed measures pertain to place of death and use of hospice.

Several other studies pointed to additional measures or areas that could be better represented on NHC. One study proposed a measure of urinary incontinence that could serve as an additional clinical outcome measure (Mukamel et al., 2003). Konetzka and Perraillon (2016) found that consumers would also be interested in information about costs and specialized services offered being posted on NHC, as these were often key drivers of choice that currently need to be ascertained directly from individual nursing homes. Finally, one study found that the NHC measures did not reflect

nursing home performance on patient safety very well, with some key outcomes such as medication errors being buried within the inspections domain (Brauner et al., 2018).

Discussion

We conducted a systematic review of the evidence on the use and effectiveness of NHC that has emerged since NHC's genesis more than two decades ago. Our search revealed a rich body of quantitative and qualitative research that speaks to how consumers have responded, how providers have responded, the unintended consequences, and issues related to the validity and comprehensiveness of the NHC measures. Overall, we found evidence that NHC has been moderately successful in meeting intended goals: NHC affects consumer choices to a modest but meaningful degree, especially under the five-star rating system, and providers pay attention to the system and try to improve their scores. Both effects appear to be stronger for postacute care than for chronic care. Because consumer use of NHC is modest due at least in part to lack of awareness, ongoing campaigns to make more consumers aware of NHC could increase use.

Our review revealed two key unintended consequences of NHC that are of particular concern. First, there are ways for providers to game the system, and evidence that they do, especially with respect to the measures based on self-reported data. Identifying gaming and trying to prevent it is difficult without substantial audits and monitoring. CMS has instead moved toward improving the sources and quality of the data. For example, staffing measures are now based on a payroll-based data collection system; while still self-reported, the required detail and potential for audits is likely to limit the potential for gaming. CMS has also added several claims-based measures to the quality domain to eliminate reliance only on Minimum Data Set data. A related issue to gaming is that providers implement narrow efforts to improve performance on included metrics, that is, they "teach to the test," while not investing broadly in quality. While technically allowed, this does little to achieve the goal of overall nursing home quality improvement. In addition to lessening reliance on self-reported data, future refinements of NHC should consider ways to incent broad-based quality improvement.

The second key unintended consequence of NHC is the exacerbation of disparities by race, ethnicity, and Medicaid enrollment status. Our review supports that traditionally underserved individuals and communities are less likely to use NHC and less likely to end up in higher quality nursing homes as a result of NHC. This is due to both consumer-driven factors (e.g., being less likely to be aware of or to use NHC) and provider-driven factors (e.g., providers in low-income areas being less likely to score well). In part, this is due to entrenched geographic disparities in supply of high-quality nursing homes, as few people want to move far from home for nursing home care, even if higher quality. Countering the exacerbation of disparities under NHC is therefore a difficult challenge, likely requiring substantial additional incentives and resources to motivate quality improvement in nursing homes in low-income neighborhoods.

Our review revealed several important domains of quality that are currently missing from NHC, detracting from the ideal of a more comprehensive picture of quality. These domains are end-of-life care and resident and family satisfaction. In addition, the salient domain of patient safety, while not entirely missing, is not well represented in the current version of NHC. Like other changes to NHC, while the addition of these areas may make NHC more complete, they may also increase complexity and further limit consumer use. The trade-offs between the desire to minimize complexity and the desire to maximize comprehensiveness should be considered carefully for any potential addition to NHC.

Our review has several limitations. First, the designs of included studies varied widely. Although we were unable to discuss the strengths and weaknesses of each study due to the volume of studies, we noted in the text which results were based on stronger quasi-experimental designs. Second, we searched for articles published by July 2019, but NHC remains a popular target of research and relevant studies continue to emerge.

Despite the large body of evidence to date on NHC, several areas are ripe for additional research to inform policy. First, little research has examined the consequences of the fairly recent changes to NHC, the addition of payroll-based staffing data and claims-based quality measures. Second, our review revealed general validity of the five-star composite rating system at the extremes, but unclear ability to make meaningful distinctions among nursing homes between those extremes. Additional research would be helpful to investigate this issue and to assess the need for refinement or development of an alternative rating system to ameliorate this problem. One alternative would be to collapse several categories of the star rating if they do not reflect meaningful differences. Third, newly available data on private-pay prices in nursing homes could lead to additional explorations of the relationship between performance on NHC and prices (Loomer et al., 2019). Finally, a new rule from CMS requires hospital discharge planners to provide patients with information about their potential choices for postacute care, including information about quality and resource use (CMS, 2019). Research will be needed to determine the role of

NHC in meeting these new requirements, and whether awareness and use of NHC increase as a result.

We conclude from our review that overall, NHC is useful and, to a moderate extent, is working as intended, but with some unintended consequences that need to be addressed. Like any public policy tool, the need for ongoing refinement should be expected. Quality measurement and reporting is inherently messy and imperfect, but moving toward the goal of empowering consumers with information about quality seems worth the challenge. Even with ongoing refinement, however, public reporting will never be a panacea. It is just one tool among many and will not erase the need for appropriate payment incentives, regulation, and technical assistance.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Supplemental Material

Supplemental material for this article is available online.

ORCID iDs

R. Tamara Konetzka https://orcid.org/0000-0003-3016-1318
Kevin Yan https://orcid.org/0000-0003-3283-4811

References

Abt Associates. (2014). *Nursing home compare five-star quality rating system: Year five report*. https://www.cms.gov/Medicare/Provider-Enrollment-and-Certification/CertificationandComplianc/Downloads/NHC-Year-Five-Report.pdf

Arling, G., Lewis, T., Kane, R. L., Mueller, C., & Flood, S. (2007). Improving quality assessment through multilevel modeling: The case of nursing home compare. *Health Services Research*, *42*(3 Pt 1), 1177–1199. https://doi.org/10.1111/j.1475-6773.2006.00647.x

Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review*, *53*(5), 941–973.

Berger, Z. D., Joy, S. M., Hutfless, S., & Bridges, J. F. (2013). Can public reporting impact patient outcomes and disparities? A systematic review. *Patient Education and Counseling*, *93*(3), 480–487. https://doi.org/10.1016/j.pec.2013.03.003

Bowblis, J. R., Lucas, J. A., & Brunt, C. S. (2015). The effects of antipsychotic quality reporting on antipsychotic and psychoactive medication use. *Health Services Research*, *50*(4), 1069–1087. https://doi.org/10.1111/1475-6773.12281

Brauner, D., Werner, R. M., Shippee, T. P., Cursio, J., Sharma, H., & Konetzka, R. T. (2018). Does Nursing Home Compare reflect patient safety in nursing homes? *Health Affairs (Millwood)*, *37*(11), 1770–1778. https://doi.org/10.1377/hlthaff.2018.0721

Cai, S., & Temkin-Greener, H. (2011). Influenza vaccination and its impact on hospitalization events in nursing homes. *Journal of Post-Acute and Long-Term Care Medicine*, *12*(7), 493–498. https://doi.org/10.1016/j.jamda.2010.03.005

Calikoglu, S., Christmyer, C. S., & Kozlowski, B. U. (2012). My eyes, your eyes–the relationship between CMS five-star rating of nursing homes and family rating of experience of care in Maryland. *Journal for Healthcare Quality*, *34*(6), 5–12. https://doi.org/10.1111/j.1945-1474.2011.00159.x

Casalino, L. P., Elster, A., Eisenberg, A., Lewis, E., Montgomery, J., & Ramos, D. (2007). Will pay-for-performance and quality reporting affect health care disparities? *Health Affairs (Millwood)*, *26*(3), w405-w414. https://doi.org/10.1377/hlthaff.26.3.w405

Castle, N. G. (2005). Nursing home administrators' opinions of the nursing home compare web site. *The Gerontologist*, *45*(3), 299–308. https://doi.org/10.1093/geront/45.3.299

Castle, N. G. (2009a). Consumers' use of Internet-based nursing home report cards. *Joint Commission Journal on Quality and Patient Safety*, *35*(6), 316–323. https://doi.org/10.1016/S1553-7250(09)35045-X

Castle, N. G. (2009b). The Nursing Home Compare report card: Consumers' use and understanding. *Journal of Aging & Social Policy*, *21*(2), 187–208. https://doi.org/10.1080/08959420902733272

Castle, N. G., Engberg, J., & Liu, D. (2007). Have Nursing Home Compare quality measure scores changed over time in response to competition? *BMJ Quality & Safety*, *16*(3), 185–191. https://doi.org/10.1136/qshc.2005.016923

Castle, N. G., Liu, D., & Engberg, J. (2008). The association of Nursing Home Compare quality measures with market competition and occupancy rates. *Journal for Healthcare Quality*, *30*(2), 4–14. https://doi.org/10.1111/j.1945-1474.2008.tb01129.x

Centers for Medicare and Medicaid Services. (2002). *Nursing Home Compare* http://www.medicare.gov/Nhcompare/Home.asp

Centers for Medicare and Medicaid Services. (2019). *CMS' discharge planning rule supports interoperability and patient preferences*. https://www.cms.gov/newsroom/press-releases/cms-discharge-planning-rule-supports-interoperability-and-patient-preferences

Clement, J. P., Bazzoli, G. J., & Zhao, M. (2012). Nursing home price and quality responses to publicly reported quality information. *Health Services Research*, *47*(1 Pt 1), 86–105. https://doi.org/10.1111/j.1475-6773.2011.01306.x

Cornell, P. Y., Grabowski, D. C., Norton, E. C., & Rahman, M. (2019). Do report cards predict future quality? The case of skilled nursing facilities. *Journal of Health Economics*, *66*(July), 208–221. https://doi.org/10.1016/j.jhealeco.2019.05.008

Feng Lu, S. (2012). Multitasking, information disclosure, and product quality: Evidence from nursing homes. *Journal of*

Economics & Management Strategy, *21*(3), 673–705. https://doi.org/10.1111/j.1530-9134.2012.00341.x

Feng Lu, S. (2016). The role of donations in quality disclosure: Evidence from nonprofit nursing homes. *American Journal of Health Economics*, *2*(4), 431–462. http://dx.doi.org/10.1111/j.1530-9134.2012.00341.x

Fuller, R. L., Goldfield, N. I., Hughes, J. S., & McCullough, E. C. (2019). Nursing Home Compare star rankings and the variation in potentially preventable emergency department visits and hospital admissions. *Population Health Management*, *22*(2), 144–152. https://doi.org/10.1089/pop.2018.0065

Geng, F., Stevenson, D. G., & Grabowski, D. C. (2019). Daily nursing home staffing levels highly variable, often below CMS expectations. *Health Affairs (Millwood)*, *38*(7), 1095–1100. https://doi.org/10.1377/hlthaff.2018.05322

Gerteis, M., Gerteis, J. S., Newman, D., & Koepke, C. (2007). Testing consumers' comprehension of quality measures using alternative reporting formats. *Health Care Financing Review*, *28*(3), 31–45.

Grabowski, D. C., & Town, R. J. (2011). Does information matter? Competition, quality, and the impact of nursing home report cards. *Health Services Research*, *46*(6 Pt 1), 1698–1719. https://doi.org/10.1111/j.1475-6773.2011.01298.x

Han, X., Yaraghi, N., & Gopal, R. (2016). *Five-star ratings for sub-par service: Evidence of inflation in nursing home ratings*. https://www.brookings.edu/research/five-star-ratings-for-sub-par-service-evidence-of-inflation-in-nursing-home-ratings/

Han, X., Yaraghi, N., & Gopal, R. (2017). Winning at all costs: Analysis of inflation in nursing homes' rating system. *Production and Operations Management*, *27*(2), 215–233. https://doi.org/10.1111/poms.12804

Harris, Y., & Clauser, S. B. (2002). Achieving improvement through nursing home quality measurement. *Health Care Financing Review*, *23*(4), 5–18.

He, D., & Konetzka, R. T. (2015). Public reporting and demand rationing: Evidence from the nursing home industry. *Health Economics*, *24*(11), 1437–1451. https://doi.org/10.1002/hec.3097

Holmstrom, B., & Milgrom, P. (1991). Multitask principal agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law Economics & Organization*, *7*(January), 24–52. https://doi.org/10.1093/jleo/7.special_issue.24

Huang, S. S., & Hirth, R. A. (2016). Quality rating and private-prices: Evidence from the nursing home industry. *Journal of Health Economics*, *50*(December), 59–70. https://doi.org/10.1016/j.jhealeco.2016.08.007

Institute of Medicine. (1986). *Improving the quality of care in nursing homes*. National Academy of Sciences.

Kim, A. S. (2016a). Market conditions and performance in the Nursing Home Compare five-star rating. *Journal of Health Politics, Policy and Law*, *41*(5), 939–968. https://doi.org/10.1215/03616878-3632221

Kim, A. S. (2016b). Nursing home report card and performance gap. *Health Care Management Review*, *41*(4), 368–377. https://doi.org/10.1097/hmr.0000000000000080

Kim, S. J., Park, E. C., Kim, S., Nakagawa, S., Lung, J., Choi, J. B., Ryu, W. S., Min, T. J., Shin, H. P., Kim, K., & Yoo, J. W. (2014). The association between quality of care and quality of life in long-stay nursing home residents with preserved cognition. *Journal of Post-Acute and Long-Term Care Medicine*, *15*(3), 220–225. https://doi.org/10.1016/j.jamda.2013.10.012

Kissam, S., Gifford, D., Parks, P., Patry, G., Palmer, L., Wilkes, L., Fitzgerald, M., Petrulis, A. S., & Barnette, L. (2003). Approaches to quality improvement in nursing homes: Lessons learned from the six-state pilot of CMS's Nursing Home Quality Initiative. *BMC Geriatrics*, *3*, Article 2. https://doi.org/10.1186/1471-2318-3-2

Konetzka, R. T., Brauner, D. J., Coca Perraillon, M., & Werner, R. M. (2015). The role of severe dementia in nursing home report cards. *Medical Care Research and Review*, *72*(5), 562–579. https://doi.org/10.1177/1077558715588436

Konetzka, R. T., Brauner, D. J., Shega, J., & Werner, R. M. (2014). The effects of public reporting on physical restraints and antipsychotic use in nursing home residents with severe cognitive impairment. *Journal of the American Geriatric Society*, *62*(3), 454–461. https://doi.org/10.1111/jgs.12711

Konetzka, R. T., Grabowski, D. C., Perraillon, M. C., & Werner, R. M. (2015). Nursing home 5-star rating system exacerbates disparities in quality, by payer source. *Health Affairs (Millwood)*, *34*(5), 819–827. https://doi.org/10.1377/hlthaff.2014.1084

Konetzka, R. T., & Gray, Z. S. (2017). The role of socioeconomic status in nursing home quality ratings. *Seniors Housing and Care Journal*, *25*(1), 3–14.

Konetzka, R. T., & Perraillon, M. C. (2016). Use of Nursing Home Compare website appears limited by lack of awareness and initial mistrust of the data. *Health Affairs (Millwood)*, *35*(4), 706–713. https://doi.org/10.1377/hlthaff.2015.1377

Konetzka, R. T., Polsky, D., & Werner, R. M. (2013). Shipping out instead of shaping up: Rehospitalization from nursing homes as an unintended effect of public reporting. *Journal of Health Economics*, *32*(2), 341–352. https://doi.org/10.1016/j.jhealeco.2012.11.008

Li, Y., Cai, X., Glance, L. G., Spector, W. D., & Mukamel, D. B. (2009). National release of the nursing home quality report cards: Implications of statistical methodology for risk adjustment. *Health Services Research*, *44*(1), 79–102. https://doi.org/10.1111/j.1475-6773.2008.00910.x

Li, Y., Schnelle, J., Spector, W. D., Glance, L. G., & Mukamel, D. B. (2010). The "Nursing Home Compare" measure of urinary/fecal incontinence: Cross-sectional variation, stability over time, and the impact of case mix. *Health Services Research*, *45*(1), 79–97. https://doi.org/10.1111/j.1475-6773.2009.01061.x

Liu, D., & Lu, C. J. (2015). Visibility and findability of the nursing home compare website. *Social Work in Public Health*, *30*(2), 144–156. https://doi.org/10.1080/19371918.2014.969861

Loomer, L., Gandhi, A., Geng, F., & Grabowski, D. C. (2019). Secret shopper data on private prices in the nursing home industry from 2008 to 2010. *Medical Care Research and*

Review. Advance online publication. https://doi.org/10.1177/1077558719879022

Lucas, J. A., Chakravarty, S., Bowblis, J. R., Gerhard, T., Kalay, E., Paek, E. K., & Crystal, S. (2014). Antipsychotic medication use in nursing homes: A proposed measure of quality. *International Journal of Geriatric Psychiatry*, *29*(10), 1049–1061. https://doi.org/10.1002/gps.4098

Mukamel, D. B., Amin, A., Weimer, D. L., Ladd, H., Sharit, J., Schwarzkopf, R., & Sorkin, D. H. (2016). Personalizing Nursing Home Compare and the discharge from hospitals to nursing homes. *Health Services Research*, *51*(6), 2076–2094. https://doi.org/10.1111/1475-6773.12588

Mukamel, D. B., Amin, A., Weimer, D. L., Sharit, J., Ladd, H., & Sorkin, D. H. (2016). When patients customize nursing home ratings, choices and rankings differ from the government's version. *Health Affairs (Millwood)*, *35*(4), 714–719. https://doi.org/10.1377/hlthaff.2015.1340

Mukamel, D. B., Caprio, T., Ahn, R., Zheng, N. T., Norton, S., Quill, T., & Temkin-Greener, H. (2012). End-of-life quality-of-care measures for nursing homes: Place of death and hospice. *Journal of Palliative Medicine*, *15*(4), 438–446. https://doi.org/10.1089/jpm.2011.0345

Mukamel, D. B., Glance, L. G., Li, Y., Weimer, D. L., Spector, W. D., Zinn, J. S., & Mosqueda, L. (2008). Does risk adjustment of the CMS quality measures for nursing homes matter? *Medical Care*, *46*(5), 532–541. https://doi.org/10.1097/MLR.0b013e31816099c5

Mukamel, D. B., Haeder, S. F., & Weimer, D. L. (2014). Top-down and bottom-up approaches to health care quality: The impacts of regulation and report cards. *Annual Review of Public Health*, *35*, 477–497. https://doi.org/10.1146/annurev-publhealth-082313-115826

Mukamel, D. B., Ladd, H., Caprio, T., & Temkin-Greener, H. (2016). Prototype end-of-life quality measures based on MDS 3 data. *Medical Care*, *54*(11), 1024–1032. https://doi.org/10.1097/MLR.0000000000000576

Mukamel, D. B., Ladd, H., Weimer, D. L., Spector, W. D., & Zinn, J. S. (2009). Is there evidence of cream skimming among nursing homes following the publication of the Nursing Home Compare report card? *The Gerontologist*, *49*(6), 793–802. https://doi.org/10.1093/geront/gnp062

Mukamel, D. B., Spector, W. D., Zinn, J., Weimer, D. L., & Ahn, R. (2010). Changes in clinical and hotel expenditures following publication of the nursing home compare report card. *Medical Care*, *48*(10), 869–874. https://doi.org/10.1097/MLR.0b013e3181eaf6e1

Mukamel, D. B., Spector, W. D., Zinn, J. S., Huang, L., Weimer, D. L., & Dozier, A. (2007). Nursing homes' response to the nursing home compare report card. *Journals of Gerontology B: Psychological Sciences and Social Sciences*, *62*(4), S218–S225. https://doi.org/10.1093/geronb/62.4.S218

Mukamel, D. B., Watson, N. M., Meng, H., & Spector, W. D. (2003). Development of a risk-adjusted urinary incontinence outcome measure of quality for nursing homes. *Medical Care*, *41*(4), 467–478. https://doi.org/10.1097/01.MLR.0000053227.95476.02

Mukamel, D. B., Weimer, D. L., & Mushlin, A. I. (2007). Interpreting market share changes as evidence for effectiveness of quality report cards. *Medical Care*, *45*(12), 1227–1232. https://doi.org/10.1097/MLR.0b013e31812f56bb

Mukamel, D. B., Weimer, D. L., Spector, W. D., Ladd, H., & Zinn, J. S. (2008). Publication of quality report cards and trends in reported quality measures in nursing homes. *Health Services Research*, *43*(4), 1244–1262. https://doi.org/10.1111/j.1475-6773.2007.00829.x

Neuman, M. D., Wirtalla, C., & Werner, R. M. (2014). Association between skilled nursing facility quality indicators and hospital readmissions. *Journal of the American Medical Association*, *312*(15), 1542–1551. https://doi.org/10.1001/jama.2014.13513

Park, J., Konetzka, R. T., & Werner, R. M. (2011). Performing well on nursing home report cards: Does it pay off? *Health Services Research*, *46*(2), 531–554. https://doi.org/10.1111/j.1475-6773.2010.01197.x

Park, J., & Werner, R. M. (2011). Changes in the relationship between nursing home financial performance and quality of care under public reporting. *Health Economics*, *20*(7), 783–801. https://doi.org/10.1002/hec.1632

Perraillon, M. C., Brauner, D. J., & Konetzka, R. T. (2017). Nursing home response to Nursing Home Compare: The provider perspective. *Medical Care Research and Review*, *76*(4), 425–443. https://doi.org/10.1177/1077558717725165

Perraillon, M. C., Konetzka, R. T., He, D., & Werner, R. M. (2019). Consumer response to composite ratings of nursing home quality. *American Journal of Health Economics*, *5*(2), 165–190. https://doi.org/10.1162/ajhe_a_00115

Ryskina, K. L., Konetzka, R. T., & Werner, R. M. (2018). Association between 5-star nursing home report card ratings and potentially preventable hospitalizations. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*. Advance online publication. https://doi.org/10.1177/0046958018787323

Saliba, D., Weimer, D. L., Shi, Y., & Mukamel, D. B. (2018). Examination of the New Short-Stay Nursing Home Quality Measures: Rehospitalizations, Emergency Department Visits, and Successful Returns to the Community. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*. Advance online publication. https://doi.org/10.1177/0046958018786816

Schapira, M. M., Shea, J. A., Duey, K. A., Kleiman, C., & Werner, R. M. (2016). The Nursing Home Compare report card: Perceptions of residents and caregivers regarding quality ratings and nursing home choice. *Health Services Research*, *51*(Suppl. 2), 1212–1228. https://doi.org/10.1111/1475-6773.12458

Sharma, H., Konetzka, R. T., & Smieliauskas, F. (2017). The relationship between reported staffing and expenditures in nursing homes. *Medical Care Research and Review*, *76*(6), 758–783. https://doi.org/10.1177/1077558717739214

Sharma, H., Perraillon, M. C., Werner, R. M., Grabowski, D. C., & Konetzka, R. T. (2019). Medicaid and nursing home choice: Why do duals end up in low-quality facilities? *Journal of Applied Gerontology*. Advance online publication. https://doi.org/10.1177/0733464819838447

Snyder, D. J., Kroshus, T. R., Keswani, A., Garden, E. B., Koenig, K. M., Bozic, K. J., Jevsevar, D. S., Poeran, J., & Moucha, C. S. (2019). Are Medicare's nursing home compare ratings accurate predictors of 90-day complications, readmission, and bundle cost for patients undergoing primary total joint arthroplasty? *Journal of Arthroplasty*, *34*(4), 613–618. https://doi.org/10.1016/j.arth.2018.12.002

Stevenson, D. G. (2006). Is a public reporting approach appropriate for nursing home care? *Journal of Health Politics, Policy and Law*, *31*(4), 773–810. https://doi.org/10.1215/03616878-2006-003

Thomas, K. (2014, August 24). Medicare star ratings allow nursing homes to game the system. *The New York Times*. https://www.nytimes.com/2014/08/25/business/medicare-star-ratings-allow-nursing-homes-to-game-the-system.html

Thomas, K. (2015, April 14). In race for Medicare dollars, nursing home care may lag. *The New York Times*. https://www.nytimes.com/2015/04/15/business/as-nursing-homes-chase-lucrative-patients-quality-of-care-is-said-to-lag.html

Totten, A., Wagner, J., Tiwari, A., O'Haire, C., Griffin, J., & Walker, M. (2012). *Closing the quality gap: Revisiting the state of the science* (Vol. 5: Public reporting as a quality improvement strategy). https://pubmed.ncbi.nlm.nih.gov/24422977/

Unroe, K. T., Greiner, M. A., Colon-Emeric, C., Peterson, E. D., & Curtis, L. H. (2012). Associations between published quality ratings of skilled nursing facilities and outcomes of Medicare beneficiaries with heart failure. *Journal of Post-Acute and Long-Term Care Medicine*, *13*(2), 188.e1–188.e6. https://doi.org/10.1016/j.jamda.2011.04.020

Vukovic, V., Parente, P., Campanella, P., Sulejmani, A., Ricciardi, W., & Specchia, M. L. (2017). Does public reporting influence quality, patient and provider's perspective, market share and disparities? A review. *European Journal of Public Health*, *27*(6), 972–978. https://doi.org/10.1093/eurpub/ckx145

Weimer, D. L., Saliba, D., Ladd, H., Shi, Y., & Mukamel, D. B. (2019). Using contingent valuation to develop consumer-based weights for health quality report cards. *Health Services Research*, *54*(4), 947–956. https://doi.org/10.1111/1475-6773.13155

Werner, R., Stuart, E., & Polsky, D. (2010). Public reporting drove quality gains at nursing homes. *Health Affairs (Millwood)*, *29*(9), 1706–1713. https://doi.org/10.1377/hlthaff.2009.0556

Werner, R. M., & Asch, D. A. (2005). The unintended consequences of publicly reporting quality information. *Journal of the American Medical Association*, *293*(10), 1239–1244.

Werner, R. M., Konetzka, R. T., & Kim, M. M. (2013). Quality improvement under nursing home compare: The association between changes in process and outcome measures. *Medical Care*, *51*(7), 582–588. https://doi.org/10.1097/MLR.0b013e31828dbae4

Werner, R. M., Konetzka, R. T., & Kruse, G. B. (2009). Impact of public reporting on unreported quality of care. *Health Services Research*, *44*(2 Pt 1), 379–398. https://doi.org/10.1111/j.1475-6773.2008.00915.x

Werner, R. M., Konetzka, R. T., & Polsky, D. (2016). Changes in consumer demand following public reporting of summary quality ratings: An evaluation in nursing homes. *Health Services Research*, *51*(Suppl. 2), 1291–1309. https://doi.org/10.1111/1475-6773.12459

Werner, R. M., Konetzka, R. T., Stuart, E. A., Norton, E. C., Polsky, D., & Park, J. (2009). Impact of public reporting on quality of postacute care. *Health Services Research*, *44*(4), 1169–1187. https://doi.org/10.1111/j.1475-6773.2009.00967.x

Werner, R. M., Konetzka, R. T., Stuart, E. A., & Polsky, D. (2011). Changes in patient sorting to nursing homes under public reporting: Improved patient matching or provider gaming? *Health Services Research*, *46*(2), 555–571. https://doi.org/10.1111/j.1475-6773.2010.01205.x

Werner, R. M., Norton, E. C., Konetzka, R. T., & Polsky, D. (2012). Do consumers respond to publicly reported quality information? Evidence from nursing homes. *Journal of Health Economics*, *31*(1), 50–61. https://doi.org/10.1016/j.jhealeco.2012.01.001

Williams, A., Straker, J. K., & Applebaum, R. (2016). The nursing home five star rating: How Does it compare to resident and family views of care? *The Gerontologist*, *56*(2), 234–242. https://doi.org/10.1093/geront/gnu043

Xu, D., Kane, R., & Arling, G. (2019). Relationship between nursing home quality indicators and potentially preventable hospitalisation. *BMJ Quality & Safety*, *28*(7), 524–533. https://doi.org/10.1136/bmjqs-2018-008924

Zimmerman, D., Karon, S. L., Arling, G., Clark, B., Collins, T., Ross, R., & Sainfort, F. (1995). Development and Testing of Nursing Home Quality Indicators. *Health Care Financing Review*, *16*(4), 107–127.

Zinn, J., Spector, W., Hsieh, L., & Mukamel, D. B. (2005). Do trends in the reporting of quality measures on the nursing home compare web site differ by nursing home characteristics? *The Gerontologist*, *45*(6), 720–730. https://doi.org/10.1093/geront/45.6.720

Zinn, J. S., Spector, W. D., Weimer, D. L., & Mukamel, D. B. (2008). Strategic orientation and nursing home response to public reporting of quality measures: An application of the miles and snow typology. *Health Services Research*, *43*(2), 598–615. https://doi.org/10.1111/j.1475-6773.2007.00781.x

Zinn, J. S., Weimer, D. L., Spector, W., & Mukamel, D. B. (2010). Factors influencing nursing home response to quality measure publication: A resource dependence perspective. *Health Care Management Review*, *35*(3), 256–265. https://doi.org/10.1097/HMR.0b013e3181e23d64