



# A qualitative study of the influence of loss aversion and increased social pressure in physician pay-for-performance

John W. Urwin<sup>1,2,3</sup>, Judy A. Shea<sup>1,2</sup>, Shireen E. Matloubieh<sup>1,2</sup>, Kristen L. Caldarella<sup>1,2</sup>, Matt Walters<sup>4</sup>, Akriti Mishra<sup>1</sup>, Ezekiel J. Emanuel<sup>1,2</sup>, Kevin G. Volpp<sup>1,2,5</sup>, Amol S. Navathe<sup>1,2,5</sup>

<sup>1</sup>Department of Medical Ethics and Health Policy, University of Pennsylvania, Philadelphia, PA, USA; <sup>2</sup>Leonard Davis Institute for Health Policy, University of Pennsylvania, Philadelphia, PA, USA; <sup>3</sup>Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA; <sup>4</sup>Advocate Health Care, Chicago, IL, USA; <sup>5</sup>Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA, USA

*Contributions:* (I) Conception and design: JW Urwin, JA Shea, SE Matloubieh, KL Caldarella, JA Shea, EJ Emanuel, KG Volpp, AS Navathe; (II) Administrative support: JA Shea, EJ Emanuel, KG Volpp, AS Navathe; (III) Provision of study materials or patients: SE Matloubieh, KL Caldarella, M Walters, A Mishra, EJ Emanuel, KG Volpp, AS Navathe; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: JW Urwin, JA Shea, SE Matloubieh, KL Caldarella, A Mishra, EJ Emanuel, KG Volpp, AS Navathe; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

*Correspondence to:* Amol S. Navathe, MD, PhD. 1108 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104, USA.

Email: amol@penmedicine.upenn.edu.

**Background:** Medicare and other payers are increasingly relying on behavioral economic principles to design pay-for-performance (P4P) reimbursement programs. However, there remains little evidence regarding how these principles relate to physicians' clinical behaviors. We conducted a qualitative study to assess the efficacy of increased social pressure and loss aversion on influencing physicians' motivations and behaviors.

**Methods:** We interviewed 22 of 33 physicians who completed the Advocate Trial, a recent randomized controlled trial testing increased social pressure and loss aversion in P4P. Seven participants were in the "active control" arm whose incentive scheme was structured the same as the year prior; nine were in the loss aversion arm that had the option to withdraw up to 50% of their expected bonus early (averaging \$19,292); and twelve were in the social pressure arm, whose performance bonus was based 50% on the performance of the group rather than 30%. Transcripts from these interviews were coded by two trained observers using NVivo 11 and analyzed using nodal analysis.

**Results:** Participants in the increased social pressure arm reported increased group cohesiveness and individual empowerment that were associated with changes in clinical behavior. Participants in the loss aversion arm reported little change in clinical behavior. The ineffectiveness of this arm was attributed to low participation in the voluntary incentive.

**Conclusions:** Incorporation of increased social pressure into physician P4P programs can lead to changes in physicians' personal motivations, group dynamics, and clinical behaviors. Process measures should be incorporated into P4P programs, as outcomes measures alone may not accurately capture quality.

**Keywords:** Pay-for-performance (P4P); behavioral economics; loss aversion

Received: 26 November 2019. Accepted: 03 March 2020; Published: 25 June 2020.

doi: 10.21037/jhmhp.2020.03.03

View this article at: <http://dx.doi.org/10.21037/jhmhp.2020.03.03>

## Introduction

An increasing number of healthcare payers are turning away from traditional fee-for-service (FFS) models and are instead looking towards pay-for-performance (P4P) programs to incentivize high-value care (1). In contrast to the traditional FFS programs, these P4P models reimburse physicians not only for *how much* care they provide, but also *how well* they achieve quality metrics, such as blood pressure control and recommended screenings (2).

While P4P programs have generally had mixed success (3-7), recent experimentation with incentive design has shown greater promise (8,9). These newer studies addressed some limitations of the early programs—notably small and poorly designed incentives—and began to incorporate principles from behavioral economics, such as loss aversion and increased social pressure. These behavioral economic principles are attractive to policymakers and payers because they have the potential to achieve higher quality care without driving up costs (10). Further, they are differentiated from traditional economic incentives because the incentive design (e.g., framing the bonus dollars as losses rather than gains or linking to group performance rather than individual performance) is intended to drive differences in performance, not incentive size.

Still, despite the promise of these principles, there is very limited data to support their efficacy in the healthcare setting (11). While there have been several successes in applying these principles to patients' decision-making and in the professional setting outside of healthcare, notably Chicago public school teachers (12-14), there have been few examples of studies examining their effectiveness in changing the clinical behavior of physicians (11).

Given the paucity of data and widespread use of these principles in multiple Medicare programs, including the Oncology Care Model (OCM) (15), Comprehensive Primary Care Plus (CPC+) initiative (16), and Merit-Based Incentive System (MIPS) (17), we performed a randomized controlled study that sought to determine if the behavioral economic principles of loss aversion and social pressure could improve clinical outcomes. While a quantitative analysis did not show a significant difference in quality outcomes (18), it did not analyze process changes or perceptions of how these interventions influenced the motivations of participating physicians. In this study, we sought to determine the perceived influences of these interventions on physicians and how the interventions changed their clinical behaviors and peer interactions.

## Methods

Standardized interviews were conducted with 22 of the 33 physicians who participated in the Advocate Trial.

### *Setting, participants, and trial design*

The Advocate Trial was a randomized controlled trial conducted from January 1, 2016 to December 31, 2016, by the Trinity Physician-Hospital Organization (PHO). Trinity is a clinically integrated physician network that is part of the larger Advocate Physician Partners associated with Advocate Health Care. At the time of the study, Trinity PHO's practice comprised 56 primary care and specialty practices and 300 physicians. It is located in Chicago, Illinois. It was selected for the trial because it had historically lower quality scores compared with other physician networks within Advocate Health Care.

The study population included all physicians affiliated with Trinity who were participating in Advocate's existing P4P program, known as the clinical integration (CI) program, and had been treating at least one patient with one of five chronic conditions (asthma, chronic obstructive pulmonary disease, diabetes, coronary artery disease or ischemic vascular disease, or congestive heart failure) for over a year. Participating physicians (*Table 1*) had a mean age of 57 years old, over 80% were in primary care or internal medicine, and nearly half (45%) of their patient panels were participating in the program. Participating patients (*Table 2*) were mostly black (71%), 85% lived in the 606\*\* zip code (99% lived in one of four zip codes), and they averaged 1.61 chronic conditions per patient.

The trial was divided into three arms. In the "active control" arm (control group), providers received 32% higher bonuses than the year prior (from \$52 per patient to \$68 per patient, or \$3,355 per physician). The decision was made to have an "active control" group with higher wages than the prior year because it was culturally unacceptable to offer higher bonuses to the experimental groups than the control group. Similar to prior years, 30% of their performance was tied to the performance of the group. In the first interventional arm (loss aversion group), physicians received the same larger bonus as the control arm, with the option of withdrawing 50% of their expected annual performance bonus (average \$19,292) as a lump sum at any point during the trial year. If their subsequent performance did not lead to a bonus equal to or greater than the amount they withdrew, they would have to pay back the difference

**Table 1** Characteristics of participating physicians

Characteristics	Active control arm (n=8)	Loss aversion arm (n=7)	Social pressure arm (n=7)
<b>Demographics</b>			
Age (years), range	61 (45–69)	66 (62–72)	54 (41–68)
Tenure (years), mean	18 (3)	18 (3)	10 (9)
Gender	38% female	29% female	43% female
<b>Medical specialty</b>			
Family medicine	3 (38%)	0 (0%)	4 (57%)
Internal medicine	3 (38%)	3 (43%)	1 (14%)
Pediatrics	1 (13%)	1 (14%)	1 (14%)
Other	1 (13%)	3 (43%)	1 (14%)
<b>Number of patients</b>			
Participating in program	857	1,209	1,076
Not participating in program	1,171	1,989	1,386
<b>Performance</b>			
Before intervention	88%	85%	83%
After intervention	92%	89%	89%

at the end of the year. In other words, their motivation during the year would be to avoid losing the upfront bonus they were given at the beginning of the year. In the second interventional arm (social pressure group), physicians also received the same larger bonus as the control arm, but the percent of their individual performance bonus tied to the group's performance was increased from 30% to 50%. In both interventional arms, physicians were given four additional official reviews of their performance data compared to the control arm.

### *Standardized interviews and surveys*

At the end of the trial, all participating physicians were invited to be interviewed using a standardized script (Supplementary file). Of the 33 who participated, 22 agreed to be interviewed. Each interview was led by one researcher from the University of Pennsylvania and was designed to last approximately 20 minutes in length. The list of questions varied slightly based on which arm of the trial they were in, though a set of general experience questions were asked of each physician interviewed. The questions were originally piloted by study authors and revised for clarity. The questions focused on eliciting global feedback and reflecting on the salience of their group's incentives,

changes in group dynamics, difficulties experienced during the trial, and recommendations for the future. All interviews were audio-recorded and subsequently transcribed by an independent third-party contractor.

### **Data analysis**

All interviews transcriptions were uploaded to NVivo 11 for analysis. Three independent observers went through each transcript and generated preliminary “nodes” that reflected recurring themes. The research team then reviewed these preliminary nodes and consolidated them into final nodes. All transcripts were then re-coded by one of the original observers and an additional coder with these final nodes and remaining conflicts were resolved through discussion. During subsequent analysis, these 15 nodes were further categorized into five major domains reflecting broader themes.

Inter-rater reliability for the 15 nodes in the nodal analysis had a mean of 0.84, a median of 1.00 and a range of 0.591–1.000.

### **Results**

During our analysis of the standardized interviews, five major domains were identified (*Table 3*). The most

Table 2 Characteristics of participating patients

Characteristics	Active control arm (n=857)	Loss aversion arm (n=1,209)	Social pressure arm (n=1,076)
<b>Demographics</b>			
Age	63 (15)	65 (13)	61 (13)
Gender	61% female	61% female	64% female
Black or African American	573 (67%)	759 (63%)	873 (81%)
Caucasian or White	81 (9%)	208 (17%)	33 (3%)
Other	59 (7%)	57 (5%)	19 (2%)
Unknown	144 (17%)	185 (15%)	151 (14%)
<b>Area code</b>			
606__	84.7%	83.8%	87.5%
604__	11.1%	10.8%	9.5%
608__	2.1%	2.0%	1.5%
463__	1.9%	2.2%	0.5%
Other	0.2%	1.2%	1.1%
<b>Qualifying illness</b>			
Asthma	54 (6%)	42 (3%)	70 (7%)
Chronic obstructive pulmonary disease	245 (29%)	174 (14%)	193 (18%)
Congestive heart failure	84 (10%)	100 (8%)	92 (9%)
Chronic hypertension	676 (79%)	1043 (86%)	820 (76%)
Diabetes mellitus	261 (30%)	356 (29%)	426 (40%)
Coronary artery disease or ischemic vascular disease	163 (19%)	254 (21%)	189 (18%)
Average no. of chronic diseases, mean (S.D.)	1.73 (0.9)	1.63 (0.8)	1.66 (0.9)

frequently discussed topics were financial incentives (73% participants, 3.6 references per interview), personal motivation to improve performance (68% participants, 2.3 references per interview), and each arm's group dynamic (64% participants, 2.0 references per interview). Other recurrent topics included how the interventions affected clinical care (64% participants, 2.7 references per interview) and barriers to success (68% participants, 2.0 references per interview).

### **Domain 1: financial incentives**

The majority of participants were either optimistic or neutral regarding their assignment. Participants in the control group tended to be pleased because the increased bonuses did not seem to come with any 'strings attached'

and they could continue without changing their own status quo, while those in the loss aversion group were similarly pleased with the *option* of accessing money earlier.

Several participants in the social pressure group, however, expressed concern that their income would be tied to the performance of others. As one physician put it, "I still was kind of upset about it...because I know I have a good score. And I was concerned the people that I saw, when I saw their scores I was like 'Oh, Lord. They're going to close me down. I'm going to lose money.'" Another physician stated their concern even more bluntly, "sharing my money with other physicians' sucks." Interestingly, none of the physicians in the other groups expressed these sentiments despite still having 30% of their incentives tied to others.

The salience of each incentive arm varied on several factors.

**Table 3** Frequency of major domains and subdomains

	Number of sources	Number of references	Number of words referenced
Financial incentives	16	52	5,513
Saliency of incentive	12	25	1,732
Interaction between groups and financial incentive	9	22	1,376
Influence of incentive	10	18	1,250
Other	3	14	1,155
Group dynamics	14	31	2,295
Teamwork	8	25	1,329
Other	11	20	966
Quality of patient care	14	35	4,154
Diligence	10	26	1,848
Patient engagement	5	10	820
Time spent with patients	6	8	620
Other	7	17	866
Motivation	15	50	4,684
Effect of being under observation	13	36	1,867
Group motivation	6	26	1,600
Peer pressure	7	16	777
Other	2	4	440
Barriers to success	15	45	3,807

In the loss aversion group, only two of the eleven physicians actually withdrew money from the virtual account. Of the two, one said that he only did so out of deference to the study, while the other used the funds to hire a physician assistant and remodel her office. There were three major reasons why the nine participants in this group elected not to withdraw any money. First, they did not need the money personally. Second, they did not have any specific business ventures for which they needed money. And third, they didn't want to assume the risk of having to repay the money later. Only one participant said he was unsure of how to access the funds.

There was also broad agreement among those who did not withdraw money that the virtual account had little influence on their decision making. As one participant said, "Let's face it this trial is sort of a once-a-year event. So it's not like you think about it a lot during the year. Once we made this decision it basically gave us some flexibility to take the money at

a sooner time frame. We ultimately decided not to." Another echoed this sentiment, saying, "Once the decision was made we didn't really review it until the end of the year." Interestingly, one physician in a different group expressed frustration that he was not in the loss aversion group because he had specific needs for the money.

In the social pressure group, those who had personal or professional ties to other members were strongly influenced by the incentive, while those lacking those ties did not. As one of the internists said, "more than the incentive structure, the group structure motivated me more. Because I'm working with the chair of medicine. And then you—we talk and we have a nice relation, and these are older people than me. And they are doing all these things. Why am I not supposed to do this?" In contrast, one of the specialists in the group brushed off the group dynamics, saying "most of the time...my performance did not influence the group as much because I'm a specialist...the group was not affected as much, kay, by my performance."

### Domain 2: group dynamics

These different incentive structures within each trial arm translated into starkly different dynamics in the social pressure arm compared to the control and loss aversion arms. While neither the control nor the loss aversion group noted significant changes from the year prior, as group members remained largely independent, the social pressure group organized regular meetings throughout the year and its members actively reached out to their colleagues on an individual level.

The prevailing sense in both the loss aversion group and the control group was one of pacificity. Despite still having 30% of their performance bonus tied to the performance of others, one physician in the loss aversion group said, “our arm was a little bit more hands-off because our success didn’t depend on the next person.” A different participant in the control arm recalled how he “was pleasantly surprised that the incented [arm was] getting together, were trying to encourage each other, and move forward. I was actually impressed.”

By contrast, the physicians in the social pressure group attended meetings every other month, at which they would review their performance and highlight areas of improvement. One physician explained, “At the meetings, we gave out scores and we looked at what percentages we were. We tried to print off peoples’ so we could see that—who is at 30 percent, who is at 50 percent.... so we really broke down the CI score...to see where we were as far as getting those points. So, we tried to analyze that. And I think that was way more information than some guys had ever seen.” Another participant elaborated further, saying “we were discussing...all the obstacles, how to improve the score, and how to improve patient care...because [if] patient care will be improved, the score will improve.”

At these meetings, they would also share best practices and help those whose performance was lagging. One physician described the content of these meetings, saying “we shared techniques. We shared insights. Maybe clarify any misunderstandings or confusion about the various parameters – the various measures.” Another added, “Those of us that did really well...tried to give pointers to people who [didn’t]. And then everybody around the table—it’s like a roundtable discussion or something. And we all got to say well, yeah, this is how I do it.”

This increased cohesiveness also translated to increased communication outside of these meetings, in the form of text messages and, occasionally, phone calls during which they would give each other reminders and tips. High-performing physicians, in particular, felt empowered to reach out directly to lower-performing members. As one

physician put it, being in the arm “motivated me to really help those that I knew were struggling or I knew didn’t historically have good scores...I took it upon myself to really push them and text them and encourage them.” Another echoed this sentiment, saying that he “personally reached out to them to help them be the best they could, especially since I know their score is going to affect me so much.”

These group-based incentives also led to increased team cohesiveness and buy-in. As one participant put it, “I would say everybody motivated each other. And we had a couple of meetings and we discussed what to do and certain kind of things that we were not looking into. We had to talk to other people what they do in their office and kind of do that too.” Another echoed this sentiment, saying “of course, meetings played a role to motivate everyone because, of course, it was a group-based score. So there was motivation...and also, investment, too.”

### Domain 3: quality of patient care

The different incentive structures also translated into notable changes in clinical behavior. The increased salience of the quality metrics—either through increased reimbursement, more frequent proformas, or new group motivations—led members in all groups to become more diligent. One participant, for instance, explained, “We put in more effort. We’re able to provide better care to the patients. And I think we must have influenced or pleased like 20 percent of the patients to do things which they were not thinking about doing yet. So finally, we were able to bring in those changes with patient care.”

Multiple participants—though notably more so in the social pressure group—also described various other ways in which they changed their practice in response to the incentives. Several described spending more time with patients to explain the rationale behind recommended screening tests or vaccinations, while others said that they began offering certain tests themselves. Others described targeting patients with unmet needs, such as missing preventative care measures, by offering earlier follow-up appointments or more proactively reaching out in the form of phone calls. As one physician said, “If it was something with the diabetes or their blood pressure...I might have made a bigger deal of it...I said well, I know we said three months, but now we’re going to two months. Or instead of waiting six months, let me see you back in three months and see if we can get this under control.”

Notably, many of these changes appear to be long-lasting rather than temporary measures. One physician put it, “if you have a set of parameters that you have to meet...on a regular



*basis and then ingrain it into habits of everyone in the office... on a regular basis ...[for] a year or six months, after the study is finished, those habits stay, at least, I'd say, maybe, 80 percent."*

#### **Domain 4: motivation**

The majority of participants said that being under observation for the study did not significantly influence their behaviors, as they were already committed to providing quality care. However, several physicians across all groups acknowledged that being observed by researchers made them want to "shine" even more than usual.

Neither the loss aversion group nor the control group expressed significant changes in their motivations. In stark contrast, the increased social pressure group consistently described increased motivation stemming from largely positive dynamics. As one physician put it, "*Yeah, I would say everybody motivated each other...we discussed what to do and certain kind[s] of things that we were not looking into.*" Another echoed this sentiment saying, "*If I saw people in the hospital... I would mention stuff. 'How are things going? My score's up to this. What's yours?' And after a while I started text[ing] people going 'My score is this. Let's all shoot for this. Let's get it here.'*"

Still, the specter of public visibility and accountability of underperformers at the end of year was apparent. As one participant in the social pressure group noted, "*If I don't score well, then my colleagues will suffer, too. And I didn't want them, at the end of the year, looking back towards me, oh, you didn't score well.*"

Interestingly, there was consensus in both interventional groups that the increased frequency of proformas was helpful in motivating them to improve. As one participant stated, "*The fine tuning and the numbers, it helped. I'm not going to tell you that the last year I am doing something different. I always tried. But then when you see these numbers totally out of the box, you try to put it together.*"

#### **Domain 5: barriers to success**

Despite the positive changes from the new incentives, participants in each arm expressed concerns over barriers that precluded further change.

The most commonly cited barrier was the time and clerical burden that the P4P programs entailed. Several said that coding and mandatory meetings were time-consuming, while others said that they did not have enough time to do everything with patients while still having time to address their immediate concerns. As one participant put it, "*I want*

*to concentrate [on] the patient rather than thinking about all this registry. That is too much for the physician. Sometime...[your] patient has some other complaint, but I was told that [she] doesn't have [a] flu shot...not this, not this...You want to take care of the patient as a patient. You know that patient. This mammogram patient, this colonoscopy patient, this flu shot...is too much work for [the] physician...And it's like paperwork."*

Several others focused on the difficulties inherent in using outcomes measures versus process measures, especially with skeptical patient populations. As one participant explained, "*So many of my patient[s]...in this area, they're not literate—they don't want [a] flu shot. And then I want to force them and even then some of them they get [a] complication [like] this high fever or something like that—they will get mad at me. So you can't force them.*" Another participant would go on to say, "*And then [I] call the patient...[and] I explain...that you need a flu shot, you need [a] colonoscopy. But if they don't go still I don't get credit.*" Given these challenging circumstances, a couple of the physicians expressed their frustration, one even saying, "*The time...dedicated to doing it was not worth the financial remuneration [we] got.*"

## **Discussion**

Physicians who participated in the Advocate Trial expressed significantly different perspectives on the salience of their incentive programs and their subsequent group dynamics, clinical behaviors, and motivations depending on which arm of the study they were enrolled.

While the control and loss aversion group participants generally approved of their incentive program, neither described significant changes in their motivation or group dynamics from the year prior. The control group physicians largely viewed their contracts as a continuation of the status quo, as they have had been working under similar contracts since 2005. The loss aversion group, in turn, conceptualized their incentive as a theoretical benefit they could use if necessary, rather than a tangible sum of money they might lose. As such, only two out of the eleven withdrew any money and none felt 'at risk.' Ironically, it appears to be the salience of loss aversion to physicians that prevented the loss aversion arm from being successful, as the threat of having to repay money was enough to convince nine out of eleven physicians to not withdraw *any* money. In a way, this arm may best have been termed an "opportunity" for loss aversion and programs implementing such incentives should consider whether their program is likely to have a high or low participation rate.

In contrast, the social pressure group widely agreed that the new group incentives caused significant changes in personal motivations, group dynamics, and clinical behavior. Not only did they demonstrate significantly higher levels of teamwork and group cohesion, several members also felt empowered to reach out to others on an individual level. These changes appear to have led to improvements in clinical practice patterns, especially at the margin, including offering additional tests to patients, arranging for closer follow-up, and more proactively reaching out to patients with gaps in care. Still, several members complained of increased uncompensated time in meetings, as well as increased professional and personal pressure from the visibility of their scores.

Regarding barriers, physicians in all arms expressed dissatisfaction with the growing clerical and clinical burdens of these programs and several critiqued reimbursing physicians on outcomes measures rather than process measures, especially with patient populations that are reluctant to accept recommended screenings.

These results are significant for several reasons. First, while a quantitative analysis failed to demonstrate a significant difference in outcomes between the three arms (18), these qualitative results indicate that increased social pressure led to significantly different team dynamics and clinical behaviors compared to the control group. These positive results for social pressure are consistent with prior trials (19,20), while our negative results are at odds with several small-scale interventions involving physicians (21) and non-physicians (11). Second, the underwhelming performance of the loss aversion group appears to stem largely from limitations in incentive design—namely that only two participants took the incentive, and none felt ‘at risk’—rather than physicians being intrinsically less susceptible to loss aversion. Third, the broad criticisms lodged against using outcomes measures as the cornerstone of Advocate Physician Partners’ quality program—especially for patients serving challenging populations—argues for integration of process measures into new P4P programs. The Advocate program has deliberately moved from all process measures to more outcomes measures over time. These criticisms from participants may also reflect the inability to risk adjust quality results for challenging populations (22) and low signal to noise ratio for many outcomes measures (23).

This trial has several key limitations. First, the sample size of the interviews was limited to 22 physicians and followed only one group in the increased social pressure

arm. While this is certainly less than ideal, all physicians were exposed to interventions for one year and this trial represents the largest trial to date examining these principles on physicians’ behaviors. Second, the design of this trial was pragmatic and limited to what was culturally acceptable for Advocate. In turn, there was an “active control” rather than a true control and physicians in the loss aversion arm were not forced to withdraw any money. Third, the baseline performance of each groups was variable, with the active control groups having the highest baseline performance. While different baselines may interfere with objective comparisons, our study focused on behavioral changes and thus should not be affected by different baseline performances. Finally, this trial was limited to a single service area of a large healthcare system in Illinois that served a predominately underserved population.

## Conclusions

While a quantitative study failed to demonstrate significant differences in outcomes between the three arms of the Advocate Trial, this qualitative analysis demonstrates significant process differences in the social pressure group. Moreover, the shortcomings of the loss aversion group appear to stem largely through the design of the incentive rather than a refutation of the principle of loss aversion to physicians.

## Acknowledgments

*Funding:* None.

## Footnote

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/jhmhp.2020.03.03>). ASN reports grants from Pennsylvania Department of Health, grants from Hawaii Medical Services Association, grants from Anthem Public Policy Institute, grants from Commonwealth Fund, grants from Oscar Health, grants from Cigna Corporation, grants from Robert Wood Johnson Foundation, grants from Donaghue Foundation, personal fees from Navvis Healthcare, personal fees from Agathos, Inc., personal fees from University Health System (Singapore), personal fees from Elsevier Press, personal fees from Navahealth, personal fees from Cleveland Clinic, personal fees from the Medicare Payment Advisory Commission, serving as



a board member without compensation for Integrated Services, Inc., and equity from Embedded Healthcare outside the submitted work. Since 2016, EJE reports speaker's fees from Tanner Healthcare System, Mid-Atlantic Permanente Group, American College of Radiology, Marcus Evans, Loyola University, Oncology Society of New Jersey, Good Shepherd Community Care, Remedy Partners, Medzel, Kaiser Permanente Virtual Medicine, Wallace H Coulter Foundation, Lake Nona Institute, Partners Chicago, Pepperdine University, Huron, American Case Management Association, Philadelphia Chamber of Commerce, Blue Cross Blue Shield Minnesota, UnitedHealth Group, Futures Without Violence, CHOP, Washington State Hospital Association, Association of Academic Health Centers, Blue Cross Blue Shield of Massachusetts, American Academy of Ophthalmology, Lumeris, Roivant Sciences Inc, Medical Specialties Distributors LLC, Vizient University Healthcare System, Center for Neuro-Degenerative Research, Colorado State University, Genentech Oncology Inc., Council of Insurance Agents and Brokers, Grifols Foundation, America's Health Insurance Plans, Montefiore Physician Leadership Academy, Greenwall Foundation, Medical Home Network, HFMA, Ecumenical Center – UT Health, American Association of Optometry, Associação Nacional de Hospitais Privados, National Alliance of Healthcare Purchaser Coalitions, Optum Labs, Massachusetts Association of Health Plans, District of Columbia Hospital Association, Washington University, Optum, Brown University, America's Essential Hospitals, National Resident Matching Program, Shore Memorial Health System, Tulane University, and Oregon Health & Science University. Since January 2017, EJE has been a venture partner with Oak HC/FT a firm that invests in health services but not pharmaceuticals or devices. EJE is on the board of two start-ups: Village MD and Oncology Analytics. KGV reported receiving grants from CVS Health, Humana, Merck, Weight Watchers, and Oscar Health Insurance; receiving personal fees from VAL Health; and that he is part owner of VAL Health. The other authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons

Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

1. Khullar D, Chokshi DA, Kocher R, et al. Behavioral economics and physician compensation--promise and challenges. *N Engl J Med* 2015;372:2281-3.
2. Rosenthal MB. Beyond pay for performance--emerging models of provider-payment reform. *N Engl J Med* 2008;359:1197-200.
3. Fendrick AM, Chernew ME. Value-based insurance design: a "clinically sensitive, fiscally responsible" approach to mitigate the adverse clinical effects of high-deductible consumer-directed health plans. *J Gen Intern Med* 2007;22:890-1.
4. Jha AK, Joynt KE, Orav EJ, et al. The long-term effect of premier pay for performance on patient outcomes. *N Engl J Med* 2012;366:1606-15.
5. Van Herck P, De Smedt D, Annemans L, et al. Systematic review: Effects, design choices, and context of pay-for-performance in health care. *BMC Health Serv Res* 2010;10:247.
6. Eijkenaar F, Emmert M, Scheppach M, et al. Effects of pay for performance in health care: a systematic review of systematic reviews. *Health Policy* 2013;110:115-30.
7. de Bruin SR, Baan CA, Struijs JN. Pay-for-performance in disease management: a systematic review of the literature. *BMC Health Serv Res* 2011;11:272.
8. Asch DA, Troxel AB, Stewart WF, et al. Effect of Financial Incentives to Physicians, Patients, or Both on Lipid Levels: A Randomized Clinical Trial. *JAMA* 2015;314:1926-35.
9. Petersen LA, Simpson K, Pietz K, et al. Effects of individual physician-level and practice-level financial incentives on hypertension care: a randomized trial. *JAMA* 2013;310:1042-50.
10. Navathe AS, Sen AP, Rosenthal MB, et al. New strategies for aligning physicians with health system incentives. *Am J Manag Care* 2016;22:610-2.
11. Emanuel EJ, Ubel PA, Kessler JB, et al. Using Behavioral Economics to Design Physician Incentives That Deliver High-Value Care. *Ann Intern Med* 2016;164:114-9.
12. Volpp KG, John LK, Troxel AB, et al. Financial incentive-

- based approaches for weight loss: a randomized trial. *JAMA* 2008;300:2631-7.
13. Volpp KG, Troxel AB, Pauly MV, et al. A randomized, controlled trial of financial incentives for smoking cessation. *N Engl J Med* 2009;360:699-709.
  14. Baker TB, Fraser DL, Kobinsky K, et al. A randomized controlled trial of financial incentives to low income pregnant women to engage in smoking cessation treatment: Effects on post-birth abstinence. *J Consult Clin Psychol* 2018;86:464-73.
  15. CMS. Oncology Care Model 2018 1 November 2018. Available online: <https://innovation.cms.gov/initiatives/oncology-care/>
  16. CMS. Comprehensive Primary Care Initiative 2018 10 August 2018. Available online: <https://innovation.cms.gov/initiatives/comprehensive-primary-care-initiative/>
  17. CMS. The Merit-Based Incentive Payment System: MIPS Scoring Methodology Overview 2015 10 August 2018. Available online: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/MIPS-Scoring-Methodology-slide-deck.pdf>
  18. Navathe AS, Volpp KG, Caldarella KL, et al. Effect of Financial Bonus Size, Loss Aversion, and Increased Social Pressure on Physician Pay-for-Performance: A Randomized Clinical Trial and Cohort Study. *JAMA Netw Open* 2019;2:e187950.
  19. Linder JA, Meeker D, Fox CR, et al. Effects of Behavioral Interventions on Inappropriate Antibiotic Prescribing in Primary Care 12 Months After Stopping Interventions. *JAMA* 2017;318:1391-2.
  20. Meeker D, Linder JA, Fox CR, et al. Effect of Behavioral Interventions on Inappropriate Antibiotic Prescribing Among Primary Care Practices: A Randomized Clinical Trial. *JAMA* 2016;315:562-70.
  21. Torchiana DE, Colton DG, Rao SK, et al. Massachusetts General Physicians Organization's quality incentive program produces encouraging results. *Health Aff (Millwood)* 2013;32:1748-56.
  22. Eijkenaar F. Key issues in the design of pay for performance programs. *Eur J Health Econ* 2013;14:117-31.
  23. Lilford RJ, Brown CA, Nicholl J. Use of process measures to monitor the quality of clinical practice. *BMJ* 2007;335:648-50.

doi: 10.21037/jhmhp.2020.03.03

**Cite this article as:** Urwin JW, Shea JA, Matloubieh SE, Caldarella KL, Walters M, Mishra A, Emanuel EJ, Volpp KG, Navathe AS. A qualitative study of the influence of loss aversion and increased social pressure in physician pay-for-performance. *J Hosp Manag Health Policy* 2020;4:14.

### Arm 1 control group script

*Arm 1 (control arm) will serve as the ‘current usual care’ control in which physician incentives payments are 70% individual/30% group-based and delivered with a performance gap loss frame (i.e., providers are told how much incentive is left on the table). Performance gap loss framing is a way to introduce a component of loss aversion to an otherwise traditional gain frame incentive and was incorporated into the incentive design at Advocate by including a report on unearned dollars and what would have needed to occur to receive the additional incentive.*

The interview should last no more than 20 minutes. Before we get started, do you have any questions for me? As we go through this, please feel free to interrupt me with any questions you may have.

#### Study participation

- ❖ Your participation involves participating in a 20-minute interview.
- ❖ The interview will be audio-recorded so that we do not miss anything important that you say. When the audio is being transcribed (typed out), all identifying information, including any names or institutions you mention in the interview, will be removed from the transcript. Therefore, no one reading the transcript of the interview will know who you are.
- ❖ If at any point during the interviews you feel uncomfortable or you need to stop for any reason, please let the interviewer know.
- ❖ Your participation is voluntary which means you can choose whether or not to participate. You can also choose to skip any question you do not wish to answer or decide you no longer wish to participate at any time.

#### Questions

- ❖ If a member of the research team cannot be reached or you want to talk to someone other than those working on the study, you may contact the Office of Regulatory Affairs with any questions, concerns or complaints at the University of Pennsylvania by calling (215) 898-2614.

#### Physician interview script

- (I) Thinking back to a little over a year ago, we first randomized you to one of three groups.
  - (i) What were your initial thoughts when you

found out what group you were in?

- (ii) What were your feelings about being assigned to your group?
  - (iii) How did your satisfaction with your arm assignment change during the course of the study?
- (II) How did participating in this trial influence you?
    - (i) What was your experience like?
    - (ii) How did being in the study change your interaction with patients?
  - (III) How often did you interact with the online registry to access your performance?
    - (i) How much did you engage in the field opps meetings or with your supplemental ProFormas?
  - (IV) How did participating in the program affect your medical care and/or your practice in other ways you did not expect?
  - (V) If this program were to be rolled out to the rest of APP physicians, what changes would you like to see?
  - (VI) Is there anything else you want to say about the study?

### Arm 2 loss aversion group script

*Arm 2 will test a potentially more potent way of inducing loss aversion by creating an endowment effect for the physician incentives, keeping the 70% individual/30% group-based components constant. This will be done by giving providers the bonus at the beginning of the year in a virtual account and making retention of incentive dollars conditional on performance. This would provide an alternative to the once-yearly payment currently provided by Advocate.*

The interview should last no more than 20 minutes. Before we get started, do you have any questions for me? As we go through this, please feel free to interrupt me with any questions you may have.

#### Study participation

- ❖ Your participation involves participating in a 20-minute interview.
- ❖ The interview will be audio-recorded so that we do not miss anything important that you say. When the audio is being transcribed (typed out), all identifying information, including any names or institutions you mention in the interview, will be removed from the transcript. Therefore, no one reading the transcript of the interview will know who you are.
- ❖ If at any point during the interviews you feel

uncomfortable or you need to stop for any reason, please let the interviewer know.

- ❖ Your participation is voluntary which means you can choose whether or not to participate. You can also choose to skip any question you do not wish to answer or decide you no longer wish to participate at any time.

### Questions

- ❖ If a member of the research team cannot be reached or you want to talk to someone other than those working on the study, you may contact the Office of Regulatory Affairs with any questions, concerns or complaints at the University of Pennsylvania by calling (215) 898-2614.

### Physician interview script

- (I) Thinking back to a little over a year ago, we first randomized you to one of three groups. What were your initial thoughts when you found out what group you were in? What were your feelings about being assigned to your group? How did your satisfaction with your arm assignment change during the course of the study?
- (II) We are particularly interested to learn about how the groups worked and how participation influenced you and your practice.
  - (i) Questions for loss aversion arm
    - ◆ Let's talk about those who were randomized to the pre-funded incentive group. We know most people did not withdraw from the account, which was very interesting to us. Let's talk about those that did not.
    - ◆ Why do you think they didn't access the funds?
    - ◆ What's your impression of how the funds were used for those that did access them?
    - ◆ We heard about excitement over investing the funds back into the practice. What was your experience with that? How were those decisions made at your practice?
  - (ii) Questions for both
    - ◆ What additional activities did you undertake to be successful in this program?
    - ◆ What worked?
    - ◆ What did not work?
- (III) How did participating in this trial influence you? What was your experience like? How did being in the study change your interaction with patients?

- (IV) In general, do you think using the prefunded account is an effective way to improve patient care? What about quality? How so? What changes would you make to the incentive structure?
- (V) How often did you interact with the online registry to access your performance?
  - (i) How much did you engage in the field opps meetings or with your supplemental ProFormas?
- (VI) How did participating in the program affect your medical care and/or your practice in other ways you did not expect?
- (VII) If this program were to be rolled out to the rest of APP physicians, what changes would you like to see?
- (VIII) Is there anything else you want to say about the study?

### Arm 3 social pressure/group-based incentive group script

*Arm 3 will test enhanced social pressure by increasing the group-based performance component of the individual physician's incentive payment to 50% (with other 50% based on individual performance) and will keep the performance gap framing constant.*

The interview should last no more than 20 minutes. Before we get started, do you have any questions for me? As we go through this, please feel free to interrupt me with any questions you may have.

### Study participation

- ❖ Your participation involves participating in a 20-minute interview.
- ❖ The interview will be audio-recorded so that we do not miss anything important that you say. When the audio is being transcribed (typed out), all identifying information, including any names or institutions you mention in the interview, will be removed from the transcript. Therefore, no one reading the transcript of the interview will know who you are.
- ❖ If at any point during the interviews you feel uncomfortable or you need to stop for any reason, please let the interviewer know.
- ❖ Your participation is voluntary which means you can choose whether or not to participate. You can also choose to skip any question you do not wish to answer or decide you no longer wish to participate at any time.

## *Questions*

- ❖ If a member of the research team cannot be reached or you want to talk to someone other than those working on the study, you may contact the Office of Regulatory Affairs with any questions, concerns or complaints at the University of Pennsylvania by calling (215) 898-2614.

## *Physician interview script*

- (I) Thinking back to a little over a year ago, we first randomized you to one of three groups. What were your initial thoughts when you found out what group you were in? What were your feelings about being assigned to your group? How did your satisfaction with your arm assignment change during the course of the study?
- (II) We are particularly interested to learn about how the groups worked and how participation influenced you and your practice.
  - (i) Questions for the group arm
    - ♦ Let's talk about those randomized to the group incentive. We are interested to learn how that played out.
    - ♦ What kind of communication was there between you and others in your group?
    - ♦ How did you work together to achieve the incentive?
  - (ii) Questions for both
    - ♦ What additional activities did you undertake to be successful in this program?
    - ♦ What worked?
    - ♦ What did not work?
- (III) How did participating in this trial influence you? What was your experience like? How did being in the study change your interaction with patients?
- (IV) In general, do you think using a group incentive is an effective way to improve patient care? What about quality? How so? What changes would you make to the incentive structure?
- (V) How often did you interact with the online registry to access your performance?
  - ♦ How much did you engage in the field opps meetings or with your supplemental ProFormas?
  - (iii) Question for group incentive arm
    - ♦ Regarding the physician rankings on the proformas- In proformas, you saw performance of yourself or peers, how did that feel versus your expectations?
    - ♦ What additional information would have been useful to add?
- (VI) How did participating in the program affect your medical care and/or your practice in other ways you did not expect?
- (VII) If this program were to be rolled out to the rest of APP physicians, what changes would you like to see?
- (VIII) Is there anything else you want to say about the study?