Effectiveness and Ethics of Incentives for Research Participation
2 Randomized Clinical Trials

Scott D. Halpern, MD, PhD; Marzana Chowdhury, PhD; Brian Bayes, MS; Elizabeth Cooney, MPH; Brian L. Hitsman, PhD; Robert A. Schnoll, PhD; Su Fen Lubitz, MPH; Celine Reyes, MA; Mitesh S. Patel, MD, MBA, MS; S. Ryan Greysen, MD, MHS, MA; Ashley Mercede, MPH; Jason Karlawish, MD; Alisa J. Stephens-Shields, PhD

IMPORTANCE Incentivizing research participation is controversial and variably regulated because of uncertainty regarding whether financial incentives serve as undue inducements by diminishing peoples' sensitivity to research risks or unjust inducements by preferentially increasing enrollment among underserved individuals.

OBJECTIVE To determine whether incentives improve enrollment in real randomized clinical trials (RCTs) or serve as undue or unjust inducements.

DESIGN, SETTING, AND PARTICIPANTS Two RCTs of incentives that were embedded in 2 parent RCTs, 1 comparing smoking cessation interventions (conducted at smoking cessation clinics in 2 health systems) and 1 evaluating an ambulation intervention (conducted across wards of the Hospital of the University of Pennsylvania) included all persons eligible for the parent trials who did not have prior knowledge of the incentives trials. Recruitment occurred from September 2017 to August 2019 for the smoking trial and January 2018 through May 2019 for the ambulation trial; data were analyzed from January 2020 to July 2020.

INTERVENTIONS Patients were randomly assigned to incentives of $0, $200, or $500 for participating in the smoking cessation trial and $0, $100, or $300 for the ambulation trial.

MAIN OUTCOMES AND MEASURES The primary outcome of each incentive trial was the proportion of people assigned to each recruitment strategy that consented to participate. Each trial was powered to test the hypotheses that incentives served neither as undue inducements (based on the interaction between incentive size and perceived research risk, as measured using a 10-point scale, on the primary outcome), nor unjust inducements (based on the interaction between incentive size and participants' self-reported income). Noninferiority methods were used to test whether the data were compatible with these 2 effects of incentives and superiority methods to compare the primary and other secondary outcomes.

RESULTS There were a total of 654 participants (327 women [50.0%]; mean [SD] age, 50.6 [12.1] years; 394 Black/African American [60.2%], 214 White [32.7%], and 24 multiracial individuals [3.7%]) in the smoking trial, and 642 participants (364 women [56.7%]; mean [SD] age, 46.7 [15.6] years; 224 Black/African American [34.9%], 335 White [52.2%], and 5 multiracial individuals [0.8%]) in the ambulation trial. Incentives significantly increased consent rates among those in the smoking trial in 47 of 216 (21.8%), 78 of 217 (35.9%), and 104 of 221 (47.1%) in the $0, $200, and $500 groups, respectively (adjusted odds ratio [aOR] for each increase in incentive, 1.70; 95% CI, 1.34-2.17; \( P < .001 \)). Incentives did not increase consent among those in the ambulation trial: 98 of 216 (45.4%), 102 of 212 (48.1%), and 92 of 214 (43.0%) in the $0, $100, and $300 groups, respectively (aOR, 0.88; 95% CI, 0.64-1.22; \( P = .45 \)). In neither trial was there evidence of undue or unjust inducement (upper confidence limits of ORs for undue inducement, 1.15 and 0.99; \( P < .001 \) showing noninferiority; upper confidence limits of ORs for unjust inducement, 1.21 and 1.26; \( P = .01 \) and \( P < .001 \), respectively). There were no significant effects of incentive size on the secondary outcomes in either trial, including time spent reviewing the risk sections of consent forms, perceived research risks, trial understanding, perceived coercion, or therapeutic misconceptions.

CONCLUSIONS AND RELEVANCE In these 2 randomized clinical trials, financial incentives increased trial enrollment in 1 of 2 trials and did not produce undue or unjust inducement or other unintended consequences in either trial.

TRIAL REGISTRATION ClinicalTrials.gov Identifier: NCT02697799

© 2021 American Medical Association. All rights reserved.
Difficult enrolling patients in randomized clinical trials (RCTs) is among the greatest barriers to evaluating health interventions. Because methods to improve RCT enrollment would greatly enhance the statistical power, generalizability, and efficiency of trials, there are increasing calls to embed evaluations of novel recruitment strategies within clinical trials. However, to our knowledge, interventions designed to improve RCT enrollment have largely failed.

Many investigators seek to increase enrollment by offering financial incentives for research participation. Paying for participation is sensible because patients and healthy volunteers consider prospects for remuneration when contemplating RCT enrollment. However, the extent to which incentives increase enrollment is uncertain. Further, because of ethical concerns, their use and size are highly variable among otherwise similar studies and are variably regulated by institutional review boards (IRBs). The most commonly cited concern with incentives for research participation is that incentives may represent undue inducements by diminishing peoples’ perceptions of the risks that are associated with research participation, thereby publicly informing consent. Another concern is that incentives represent unjust inducements by encouraging enrollment preferentially among people with lower income levels, thereby creating a system in which the burdens of research participation are borne preferentially by low-income individuals, whereas the knowledge gained from the research would benefit all.

Both of these concerns have been challenged on conceptual grounds, and studies that assess how willing people are to participate in hypothetical RCTs have provided limited empirical support for the view that incentives increase enrollment without yielding either of these unintended ethical consequences. However, evidence of how real incentives affect decisions to participate in real trials is needed to optimally inform research regulations and international practice. We therefore undertook the Randomized Evaluation of Trial Acceptance by Incentive (RETAIN) study to examine, in 2 separate RCTs embedded within different parent trials, the extent to which incentives produced several intended and unintended influences on the decision for patients to enroll. Our primary goals were to evaluate whether incentives increase trial enrollment among eligible patients and whether incentives constitute either undue or unjust inducements.

Methods

Trial Design

RETAIN comprised 2 RCTs. The first trial compared 3 financial incentives for participation in a parent RCT that evaluated smoking cessation interventions among persons with major depressive disorder (NCT02378714), and the second compared 3 incentives for participation in a parent RCT that compared a gamification intervention with usual care to promote ambulation among hospitalized patients (NCT03321279). A third trial of incentives for participation in a parent trial that compared forms of radiation therapy for lung cancer was initiated but halted because logistical barriers in the parent trial (Supplement 1). The protocols for the 2 embedded trials that moved forward are also described in Supplement 1 and were approved by IRBs at the University of Pennsylvania and Northwestern University using a waiver of the requirement for informed consent (eMethods 3 in Supplement 2). The trials were also guided by an ethics advisory board and overseen by a data and safety monitoring board. We followed the guidelines for reporting embedded recruitment trials and several aspects of the Consolidated Standards of Reporting Trials extension for noninferiority trials because portions of the analytic plan entail tests of noninferiority.

Participants

Participants eligible for RETAIN were those approached for informed consent to participate in the smoking or ambulation trials. Embedded RETAIN recruitment in these trials occurred from September 2017 through August 2019 and from January 2018 through May 2019, respectively. For the smoking trial, recruitment first entailed a telephone conversation with a study coordinator, during which initial eligibility was assessed and patient characteristics were measured. Key measures obtained at this stage included income, financial well-being, and score on the Research Attitudes Questionnaire—7 (RAQ-7), a measure with high internal consistency and factorial validity in which higher scores indicate more favorable views of research. Key elements and risks of the trial were then described, and the assigned incentive was revealed (eMethods 1 in Supplement 2). Coordinators used scripted probes to assess whether patients had heard about the use of incentives in the trial. Patients who indicated any prior familiarity were deemed ineligible for RETAIN. The remaining patients were
Table 1. Participant Characteristics by Group in the 2 Embedded Randomized Clinical Trials*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Smoking trial</th>
<th>Ambulation trial</th>
<th>Smoking trial Ambulation trial</th>
<th>Smoking trial Ambulation trial</th>
<th>Smoking trial Ambulation trial</th>
<th>Smoking trial Ambulation trial</th>
<th>Smoking trial Ambulation trial</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.</td>
<td>216</td>
<td>217</td>
<td>221</td>
<td>654</td>
<td>216</td>
<td>212</td>
<td>214</td>
</tr>
<tr>
<td>Age, y (Mean (SD))</td>
<td>49.8 (12.8)</td>
<td>50.1 (12.2)</td>
<td>51.8 (11.2)</td>
<td>50.6 (12.1)</td>
<td>49.0 (15.6)</td>
<td>46.3 (14.9)</td>
<td>45.3 (15.9)</td>
</tr>
<tr>
<td>Missing</td>
<td>0</td>
<td>0</td>
<td>1 (0.5)</td>
<td>1 (0.2)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sex</td>
<td>Male</td>
<td>80 (37.0)</td>
<td>86 (39.6)</td>
<td>85 (38.5)</td>
<td>251 (38.7)</td>
<td>91 (42.1)</td>
<td>88 (41.5)</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>110 (50.9)</td>
<td>110 (50.7)</td>
<td>107 (48.4)</td>
<td>327 (50.0)</td>
<td>125 (57.9)</td>
<td>124 (58.5)</td>
</tr>
<tr>
<td>Missing</td>
<td>26 (12.0)</td>
<td>21 (9.7)</td>
<td>29 (13.1)</td>
<td>76 (11.6)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Race</td>
<td>Black/African American</td>
<td>135 (62.5)</td>
<td>128 (59.0)</td>
<td>131 (59.3)</td>
<td>394 (60.2)</td>
<td>79 (36.6)</td>
<td>70 (33.0)</td>
</tr>
<tr>
<td></td>
<td>White</td>
<td>68 (31.5)</td>
<td>76 (35.0)</td>
<td>70 (31.7)</td>
<td>214 (32.7)</td>
<td>115 (53.2)</td>
<td>110 (51.9)</td>
</tr>
<tr>
<td></td>
<td>More than 1 race</td>
<td>8 (3.7)</td>
<td>6 (2.8)</td>
<td>10 (4.5)</td>
<td>24 (3.7)</td>
<td>1 (0.5)</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td></td>
<td>Missing</td>
<td>4 (1.9)</td>
<td>5 (2.3)</td>
<td>8 (3.6)</td>
<td>17 (2.6)</td>
<td>12 (5.6)</td>
<td>17 (8.0)</td>
</tr>
<tr>
<td>Race</td>
<td>Married or living with a partner</td>
<td>15 (6.9)</td>
<td>23 (10.6)</td>
<td>23 (10.4)</td>
<td>61 (9.3)</td>
<td>98 (45.4)</td>
<td>89 (42.0)</td>
</tr>
<tr>
<td></td>
<td>Never married</td>
<td>37 (17.1)</td>
<td>49 (22.6)</td>
<td>51 (23.1)</td>
<td>137 (20.9)</td>
<td>82 (38.0)</td>
<td>83 (39.2)</td>
</tr>
<tr>
<td></td>
<td>Missing</td>
<td>4 (1.9)</td>
<td>5 (2.3)</td>
<td>8 (3.6)</td>
<td>17 (2.6)</td>
<td>12 (5.6)</td>
<td>17 (8.0)</td>
</tr>
<tr>
<td>Education</td>
<td>Some high school</td>
<td>19 (8.8)</td>
<td>25 (11.5)</td>
<td>15 (6.8)</td>
<td>59 (9.0)</td>
<td>9 (4.2)</td>
<td>10 (4.7)</td>
</tr>
<tr>
<td></td>
<td>High school graduate</td>
<td>56 (25.9)</td>
<td>55 (25.3)</td>
<td>72 (32.6)</td>
<td>183 (28.0)</td>
<td>62 (28.7)</td>
<td>65 (30.7)</td>
</tr>
<tr>
<td></td>
<td>Some college</td>
<td>60 (27.8)</td>
<td>71 (32.8)</td>
<td>87 (39.4)</td>
<td>218 (33.3)</td>
<td>68 (31.5)</td>
<td>56 (26.4)</td>
</tr>
<tr>
<td></td>
<td>College degree</td>
<td>35 (16.2)</td>
<td>46 (21.2)</td>
<td>36 (16.3)</td>
<td>117 (17.9)</td>
<td>65 (30.1)</td>
<td>64 (30.2)</td>
</tr>
<tr>
<td></td>
<td>Missing</td>
<td>46 (21.3)</td>
<td>20 (9.2)</td>
<td>11 (5.0)</td>
<td>77 (11.8)</td>
<td>12 (5.6)</td>
<td>17 (8.0)</td>
</tr>
<tr>
<td>Household income in thousands, $&lt;10</td>
<td>39 (18.1)</td>
<td>36 (16.6)</td>
<td>38 (17.2)</td>
<td>113 (17.3)</td>
<td>8 (3.7)</td>
<td>9 (4.2)</td>
<td>12 (5.6)</td>
</tr>
<tr>
<td>10-19</td>
<td>31 (14.4)</td>
<td>33 (15.2)</td>
<td>39 (17.6)</td>
<td>103 (15.7)</td>
<td>9 (4.2)</td>
<td>14 (6.6)</td>
<td>8 (3.7)</td>
</tr>
<tr>
<td>20-29</td>
<td>23 (10.6)</td>
<td>38 (17.5)</td>
<td>43 (19.5)</td>
<td>104 (15.9)</td>
<td>22 (10.2)</td>
<td>12 (5.7)</td>
<td>16 (7.5)</td>
</tr>
<tr>
<td>30-49</td>
<td>33 (15.3)</td>
<td>38 (17.5)</td>
<td>40 (18.1)</td>
<td>111 (17.0)</td>
<td>37 (17.1)</td>
<td>39 (18.4)</td>
<td>33 (15.4)</td>
</tr>
<tr>
<td>50-99</td>
<td>25 (11.6)</td>
<td>34 (15.7)</td>
<td>37 (16.7)</td>
<td>96 (14.7)</td>
<td>92 (42.6)</td>
<td>80 (37.7)</td>
<td>88 (41.1)</td>
</tr>
<tr>
<td>≥100</td>
<td>13 (6.0)</td>
<td>13 (6.0)</td>
<td>7 (3.2)</td>
<td>33 (5.0)</td>
<td>36 (16.7)</td>
<td>40 (18.9)</td>
<td>44 (20.6)</td>
</tr>
<tr>
<td>Missing</td>
<td>52 (24.1)</td>
<td>25 (11.5)</td>
<td>17 (7.7)</td>
<td>94 (14.4)</td>
<td>12 (5.6)</td>
<td>18 (8.5)</td>
<td>13 (6.1)</td>
</tr>
<tr>
<td>Personal financial well-being Median (Q1-Q3)</td>
<td>3.88 (2.50-5.88)</td>
<td>4.63 (3.13-6.13)</td>
<td>3.88 (2.16-5.75)</td>
<td>4.13 (2.50-5.88)</td>
<td>6.88 (4.19-8.00)</td>
<td>6.94 (3.75-8.47)</td>
<td>7.16 (4.50-8.75)</td>
</tr>
<tr>
<td>Missing</td>
<td>44 (20.4)</td>
<td>20 (9.2)</td>
<td>11 (5.0)</td>
<td>75 (11.5)</td>
<td>13 (6.0)</td>
<td>17 (8.0)</td>
<td>14 (6.5)</td>
</tr>
<tr>
<td>Prior research experience</td>
<td>Favorable prior research experience</td>
<td>79 (36.6)</td>
<td>91 (41.9)</td>
<td>83 (37.6)</td>
<td>253 (38.7)</td>
<td>49 (22.7)</td>
<td>39 (18.4)</td>
</tr>
<tr>
<td></td>
<td>Unfavorable prior research experience</td>
<td>16 (7.4)</td>
<td>22 (10.1)</td>
<td>15 (5.9)</td>
<td>51 (7.8)</td>
<td>16 (7.4)</td>
<td>20 (9.4)</td>
</tr>
<tr>
<td></td>
<td>No prior research experience</td>
<td>79 (36.6)</td>
<td>86 (39.6)</td>
<td>116 (52.5)</td>
<td>281 (43.0)</td>
<td>138 (63.9)</td>
<td>135 (63.7)</td>
</tr>
<tr>
<td></td>
<td>Missing</td>
<td>42 (19.4)</td>
<td>18 (8.3)</td>
<td>9 (4.1)</td>
<td>69 (10.6)</td>
<td>13 (6.0)</td>
<td>18 (8.5)</td>
</tr>
<tr>
<td>Research attitudes questionnaire Mean (SD)</td>
<td>28.1 (4.18)</td>
<td>28.9 (4.26)</td>
<td>28.4 (4.05)</td>
<td>28.5 (4.17)</td>
<td>27.1 (2.92)</td>
<td>27.1 (2.16)</td>
<td>27.2 (2.95)</td>
</tr>
<tr>
<td>Missing</td>
<td>44 (20.4)</td>
<td>21 (9.7)</td>
<td>11 (5.0)</td>
<td>76 (11.6)</td>
<td>13 (6.0)</td>
<td>18 (8.5)</td>
<td>14 (6.5)</td>
</tr>
</tbody>
</table>

Abbreviation: Q, quarter.

* All mean standardized effects were <0.25 in the smoking trial and <0.22 in the ambulation trial.

b Marital status was asked at the initial telephone screen step of the smoking trial starting in July 2018.

c Possible values for financial well-being scale, 1 to 10.

d Possible values for research attitudes questionnaire, 7 to 35.
invited to attend an in-person assessment to read and potentially sign the informed consent form and to confirm full eligibility for the smoking trial.

Recruitment to the ambulation trial followed a similar order but occurred in a single step, with research coordinators conducting in-person recruitment in the hospital rooms of patients (eMethods 1 in Supplement 2). The recruitment strategy for the RETAIN trial that was embedded within the ambulation trial was changed after the first 30 patients were enrolled (Supplement 1). Those 30 patients were excluded from analyses.

Randomization, Interventions, and Masking
Potential participants in both trials were randomized in equal (33.3%) proportions to the 3 incentive groups. Randomization was stratified by site in the embedded smoking trial and by recruiter in the embedded ambulation trial. Within strata, variable block sizes of 3 and 6 were used. Patients approached for the smoking trial were randomized to receive $0, $200, or $500 for participation. Patients approached for the ambulation trial were randomized to receive $0, $100, or $300. These amounts differed based on the time and effort being asked of participants and are consistent with incentive amounts commonly offered in other RCTs. In both trials, the assigned incentive was communicated to potential participants via scripted verbal communication from research coordinators and the informed consent documents of the parent trials. These documents indicated the payment for research participation in the trial summary on the first page and in all subsequent sections that described the costs to participate. In the $0 arms, these sections stated that participants “would not be paid” for enrolling. All other elements of the consent processes for the parent trials were identical across incentive groups. To enable measurement of time spent reviewing the consent documents, we used electronic versions with discrete pages that covered different aspects of the trial.

Research coordinators administering the consent processes in both parent trials were unmasked to the incentive information but adhered to IRB-approved communication scripts. Investigators conducting data analyses (M.C. and A.S.) were unmasked to group assignment once data collection for both trials was complete to enable tests for trend and non-inferiority, but the statistical analysis plan was finalized before unmasking.

Debriefing
After RETAIN outcomes were collected, research coordinators debriefed patients who consented to participate in each parent trial using an IRB-approved debriefing script (eMethods 2 in Supplement 2) that explained the rationales for assigning different incentives and not informing patients about RETAIN. Coordinators revealed that all patients, regardless of the assigned incentive, would receive the maximal incentive ($500 in the smoking trial and $300 in the ambulation trial). We equalized payments to reward people identically for their contributions to the parent RCTs.

Outcomes
The primary outcome in each embedded trial was the proportion of people assigned to each recruitment strategy that consented to participate, defined as whether patients signed the parent trial consent form. Randomized participants who verbalized disinterest in the parent trial before seeing the consent form were considered to have declined consent.

All secondary outcomes were measured after the incentive was verbally revealed. We measured the time in seconds that patients spent reading the electronic consent form on the tablet and the time spent reading the specific pages describing risks. After patients made their choice to consent to participate in the parent trial or not, we administered a questionnaire assessing the perceived risks of patients regarding the parent trial, as measured by a modified version of the compared riskiness scale. We assessed patients’ understanding of the parent trial using 6 questions regarding core trial elements mentioned in the consent form, such as its duration and purpose. The incidence of therapeutic misconception (ie, the belief that the goals of research are similar to the goals of clinical care) was assessed using a 4-item therapeutic misconception tool. Patients’ perceptions of being coerced to participate were assessed using the Perceived Coercion Scale of the MacArthur Admission Experience Survey, with wording modified to suit these trials. Because most patients who declined consent for the parent trials did not complete these postconsent instruments, these outcomes were compared across incentive arms among consenting patients. We had originally planned to measure the effects of incentives on retention until the end of treatment in the parent trials but diverted from this plan because the choice to equalize actual payments, as revealed through debriefing (previously described), precluded any effects of incentives.

Statistical Analysis
Primary analyses used logistic regression, which was adjusted for site and prespecified patient characteristics (Supplement 1), to evaluate the main effects of incentives on consenting to participate in the parent trial and the 3 interaction terms of interest: between incentive and perception of research risk (as the measure of undue inducement) and between incentive and income and incentive and financial well-being (as 2 measures of unjust inducement). In the primary analysis, to maximize power, we entered incentives into logistic models as a continuous variable with values 0, 1, and 2 to promote consistency between the 2 trials. Similarly, in tests for undue and unjust inducement, incentives and the effect modifiers of perceived riskiness, income, and financial well-being were each coded as continuous variables (eTables 1-3 in Supplement 2). We used χ² and Jonckheere tests for trend to compare binary and continuous outcomes, respectively, across ordered incentive arms. We repeated all analyses of secondary outcomes after weighting for the inverse probability of response. We report unweighted results because all weighted results were similar.

We designed the 2 embedded trials such that each was independently powered to determine that incentives did not represent undue inducements, defined by the interaction between incentive size and perceived trial risk, using a
noninferiority margin of 2.0 for the interaction odds ratio (OR). We chose noninferiority analyses rather than traditional superiority tests of these terms because concerns about undue and unjust inducement commonly limit the use of incentives in research, meaning that the burden of proof for practice to change requires establishing that such unintended consequences are unlikely to arise. For example, an incentive-by-risk interaction OR of 2.0 would be found if enrollment dropped from 40% to 25% with increasing risk perception among patients who were not assigned to incentives, but remained constant at 40% across levels of risk perception among patients who received incentives (eTable 4 in Supplement 2).

Assuming that 50% of RETAIN patients would choose to enroll in each parent trial, a true interaction OR of 1, and using a 1-sided significance level of \( P = .04 \) following 2 interim analyses such that the upper 95.73% confidence limit on the observed OR would have to fall below the noninferiority threshold to rule out undue inducement, we calculated that 576 patients in each embedded trial would provide at least 80% power to conclude that undue inducement did not manifest. Similar power would exist to conclude that incentives did not represent unjust inducements, as indicated by an interaction OR of 2.0 or greater between either (1) income and incentive amount or (2) financial well-being and incentive amount.

As a secondary measure of undue inducement, we tested the significance of the interaction between RAQ-7 scores and incentives. We reasoned that if the relationship between patients’ favorable views of research and their odds of enrolling was modified by incentives, then incentives may unduly alter the nonfinancial motivations for people to participate. The embedded smoking trial ended when the parent trial closed, and the embedded ambulation trial ended when the target sample size with complete risk assessment data for RETAIN was achieved. All analyses were conducted using R, version 3.6.3 (R Foundation), or Stata, version 14 (StataCorp).

Results

Participants

All 659 eligible participants in the smoking trial who presented during the RETAIN collaboration were randomized, as were 646 of 652 eligible participants (99.1%) in the ambulation trial. After randomization, 5 patients (0.7%) in the smoking trial and 4 patients (0.6%) in the ambulation trial were excluded for reasons indicated in Figure 1, leaving 654 and 642 patients in these trials, respectively. Table 1 presents the characteristics of patients enrolled in the 3 groups of each trial.

Effects of Incentives on Trial Enrollment

In the smoking trial, consent rates were 21.8%, 35.9%, and 47.1% in the $0, $200, and $300 arms, respectively (\( P < .001 \)). In the ambulation trial, consent rates were 45.4%, 48.1%, and 43.0% in the $0, $100, and $300 arms, respectively (\( P = .62 \)). Analyses adjusted for baseline patient characteristics, including demographic characteristics, financial well-being, and RAQ scores, revealed a significant effect of incentives on consent in the smoking trial (adjusted OR [aOR] for each increase in incentive amount or (2) financial well-being and incentive amount. OR of 2.0 or greater between either (1) income and incentive amount or (2) financial well-being and incentive amount.

Within the Smoking and Ambulation Trials

Table 1 presents the characteristics of patients enrolled in the 3 groups of each trial.
have been reached using noninferiority margins as low as 1.8 and 1.3, respectively, for income, and 1.7 and 1.2 for financial well-being (eTable 9 in Supplement 2). In additional post hoc analyses of undue inducement (eTable 10 in Supplement 2), we found no evidence of undue inducement in the smoking trial (interaction OR, 1.34; 95% CI, 0.91-1.95; P = .03). In the ambulation trial, the significance of the corresponding interaction term (interaction OR, 0.83; 95% CI, 0.38-2.06; P = .048) was slightly greater than the adjusted threshold of P = .043 required to conclude that there was no evidence of undue inducement (eTable 10 in Supplement 2). Incentives did not meet secondary criteria for undue inducement, as they did not modify the relationships between favorable research attitudes and enrollment probabilities in either the smoking (interaction OR, 1.01; 95% CI, 0.95-1.08; P = .65) or ambulation trials (interaction OR, 1.01; 95% CI, 0.88-1.15; P = .94).

Tests of Incentives as Undue or Unjust Inducements
Whereas nearly half the participants in the smoking trial perceived it to have risks, fewer than 10% of participants in the ambulation trial perceived such risks (Table 2). Risk perception did not modify the relationships between incentives and enrollment in either trial (upper confidence limits of interaction ORs, 1.15 and 0.99; both P < .001) (eFigures 1-2 in Supplement 2). In the smoking trial, the upper confidence limit of 1.15 must be cubed to account for the 3-unit difference between the 25th and 75th percentiles of the risk perception distribution (eTable 7 in Supplement 2), yielding an upper limit of 1.52 (eTable 4 in Supplement 2). Because these upper confidence limits were appreciably less than the established noninferiority margin of 2.0, these data were not compatible with incentives operating as undue inducements. Similarly, the data were not compatible with incentives operating as unjust inducements in either trial because neither income (upper confidence limits of ORs, 1.21 and 1.26; P = .01 and P < .001) (Figure 2) nor financial well-being (upper confidence limits of ORs, 1.17 and 1.04; P = .003 and P < .001) (Figure 3) modified the effect of incentives.

Because there was no evidence to guide the choice of the noninferiority margin, we performed post hoc sensitivity analyses to determine the smallest margins at which undue and unjust inducement were ruled out. Similar conclusions that undue inducement was not present would have been reached using noninferiority margins as low as 1.6 and 1.8 in the smoking and ambulation trials, respectively (eTable 8 in Supplement 2). The smallest noninferiority margins at which unjust inducement would be found not to be present were 1.8 and 1.3, respectively, for income, and 1.7 and 1.2 for financial well-being (eTable 9 in Supplement 2). In additional post hoc analyses of undue inducement treating risk perception as binary (0 or >0) because of its distributions (eTable 10 in Supplement 2), we found no evidence of undue inducement in the smoking trial (interaction OR, 1.34; 95% CI, 0.91-1.95; P = .03). In the ambulation trial, the significance of the corresponding interaction term (interaction OR, 0.83; 95% CI, 0.38-2.06; P = .048) was slightly greater than the adjusted threshold of P = .043 required to conclude that there was no evidence of undue inducement (eTable 10 in Supplement 2). Incentives did not meet secondary criteria for undue inducement, as they did not modify the relationships between favorable research attitudes and enrollment probabilities in either the smoking (interaction OR, 1.01; 95% CI, 0.95-1.08; P = .65) or ambulation trials (interaction OR, 1.01; 95% CI, 0.88-1.15; P = .94).

Analyses of Secondary Outcomes
Table 2 presents secondary outcomes for both trials, all of which were evaluated with traditional superiority tests. Increasing incentives were not associated with patients’ perceptions of research risks in the smoking (OR, 0.85; 95% CI, 0.61-1.17) or ambulation trials (OR, 1.34; 95% CI, 0.70-2.69). Patients who received greater incentives spent more time reviewing the consent form risk section in the smoking trial but not the ambulation trial and more total time reviewing consent forms in both trials (Table 2). There were low rates of possible and likely therapeutic misconceptions in both trials, which did not differ by incentive group. The median percentage of correct re-

Table 2. Outcomes by Group in the 2 Embedded Randomized Clinical Trials

<table>
<thead>
<tr>
<th>Outcomea</th>
<th>Smoking trial</th>
<th>Ambulation trial</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$0</td>
<td>$200</td>
</tr>
<tr>
<td>No.</td>
<td>216</td>
<td>217</td>
</tr>
<tr>
<td>Consent rate, No./total No. (%)</td>
<td>47/216 (21.8)</td>
<td>78/217 (35.9)</td>
</tr>
<tr>
<td>Perceived risks of the research, No./total No. (%)</td>
<td>102/216 (47.2)</td>
<td>106/217 (48.8)</td>
</tr>
<tr>
<td>Time spent reviewing consent, s</td>
<td>All, median (IQR)</td>
<td>908 (411.5-1283.8)</td>
</tr>
<tr>
<td>Risk section, median (IQR)</td>
<td>155 (73.5-321)</td>
<td>137 (59.25-257.5)</td>
</tr>
<tr>
<td>Incidence of TM, No./total No. (%)</td>
<td>5/47 (10.6)</td>
<td>8/78 (10.3)</td>
</tr>
<tr>
<td>Likely TM</td>
<td>1/47 (2.1)</td>
<td>0/78 (0)</td>
</tr>
<tr>
<td>Understanding of the trial</td>
<td>Median % (IQR)</td>
<td>66.7 (66.7-88.3)</td>
</tr>
<tr>
<td>Perceptions of influence of coercion, No./ total No. (%)</td>
<td>4/92 (0)</td>
<td>6/92 (6.5)</td>
</tr>
</tbody>
</table>

Abbreviation: TM, therapeutic misconceptions.

a Except perceived risk of research, the rest of the outcomes were analyzed among the consented sample.

b Adjusted P value from the logistic model.
sponses to the 6 trial understanding questions was 67% in all smoking trial groups and 100% in all ambulation trial groups. More than 70% of participants in each smoking trial group perceived no coercion ($P = .91$), as did more than 93% of participants in each ambulation trial group ($P = .48$).

Discussion

These 2 embedded trials of financial incentives for RCT enrollment provide several insights that are important for investigators, research regulatory bodies, and research sponsors. First, we found that incentives substantially improved enrollment in a smoking cessation trial but not in a trial of a behavioral intervention to promote ambulation among hospitalized patients. The higher incentive sizes in the smoking trial are unlikely to explain the different results because the $300 incentive in the ambulation trial was ineffective, whereas the $200 incentive in the smoking trial increased enrollment. A more likely explanation is that the benefits of the ambulation trial relative to the risks were perceived to be high, yielding higher baseline consent rates and fewer patients whose decisions could be altered by incentives. The different results could also be because of unmeasured differences in the characteristics of the samples or differences in the interventions or recruitment settings of the 2 trials.

Second, although many ethicists and research regulators have worried that research incentives may unduly induce participation by masking participants to risk and preventing fully informed consent, in these 2 trials, incentives were not associated with patients’ perceptions of research risks and did not significantly modify patients’ sensitivity to increased risks as a factor that governed their willingness to enroll. For example, in the smoking trial, the data are compatible with an incentive-by-risk interaction OR of up to 1.15, and even that
worst-case scenario would only represent an 8.7% absolute reduction in risk sensitivity among patients who were receiving vs not receiving incentives (Table 4 in Supplement 2). Interpretation of similar analyses in the ambulation trial is limited by the lack of effects of incentives on enrollment, low perceived risks, and because a post hoc analysis with risk modeled as a binary variable was inconclusive regarding undue inducement in this trial. However, the findings that incentives did not modify the relationships between research attitudes and enrollment in either trial supports the conclusion that incentives did not unduly influence trial enrollment decisions.

Third, incentives did not function as unjust inducements, as they were not preferentially motivating across groups with different income levels or financial well-being in either trial. Thus, these trials do not support concerns that incentives could cause undeserved individuals to disproportionately bear the risks and burdens of research.25,27,52

Finally, we found that incentives did not influence important secondary outcomes. In contrast with prior hypotheses,53 incentives did not affect rates of therapeutic misconception, possibly because these rates were so low among patients who consent to participate in these 2 parent trials. Patients were also no more likely to perceive being coerced when offered incentives. Although patients who received incentives spent more time reviewing the consent forms, which was consistent with a prior study that used hypothetical RCTs,54 this increased attention did not translate into improved trial understanding. Ceiling effects, particularly in the ambulation trial in which most people had perfect understanding, may have affected the latter result.

Limitations

This study has limitations. First, we relied on patients’ perceptions of study risks, rather than actual study risks, to determine whether payments were undue inducements. This is a necessary limitation in evaluating incentives within real RCTs because the risks of real trials are not malleable. However, studies using hypothetical RCTs, in which risks were experimentally manipulated, also concluded that payments do not represent undue inducements.30,31

Second, we could not measure certain secondary outcomes, including therapeutic misconception and perceived coercion, among patients who chose not to participate in the parent trials. The absolute rates of these outcomes may have differed among patients who were uninterested in the trials. Nonetheless, the similar rates of these outcomes across incentive arms among patients who did participate in the parent trials provides reassurance that incentives do not affect these end points among those who go on to bear research risks.

Third, although the generalizability of our results is strengthened by conducting 2 embedded trials within different parent trials that differed in their true risks, neither of these parent trials posed particularly high risks. Future tests of incentives of different sizes, and in the context of higher-risk parent trials, including trials that test treatments of serious illnesses, are warranted.

Conclusions

The findings from these 2 RCTs of financial incentives embedded within parent trials with different sample characteristics provide real-world confirmation of hypothetical studies that also had not identified ethical problems with incentives for research participation. Thus, research regulators should relax restrictions on the use of incentives that are designed to improve enrollment in low-risk trials. However, the effectiveness of incentives in improving enrollment rates is likely to vary across trials. Thus, although it will be appropriate in most trials to compensate research participants for their time and burden in engaging with research tasks, further evaluations of incentives that are designed specifically to enhance enrollment are needed to understand the range of trials in which this practice represents money well spent.

ARTICLE INFORMATION

Accepted for Publication: June 25, 2021.
Published Online: September 20, 2021.

Author Affiliations: Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Halpern, Patel, Greysen, Mercede, Reale, Volpp, Karlawish); Palliative and Advanced Illness Research Center, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Halpern, Chowdhury, Bayes, Cooney, Greysen, Volpp, Karlawish, Stephens-Shields); Center for Health Incentives and Behavioral Economics, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Halpern, Chowdhury, Bayes, Cooney, Greysen, Volpp, Karlawish, Stephens-Shields); Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois (Hitsman, Reyes, Schnoll, Lubitz); Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Schnoll, Lubitz); Abramson Cancer Center, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Scholl); Department of Health Care Management, The Wharton School, University of Pennsylvania, Philadelphia (Patel, Volpp); Corporal Michael J. Crescenz VA Medical Center, Philadelphia, Pennsylvania (Patel, Volpp); Family Medicine and Community Health, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Barg); Department of Anthropology, University of Pennsylvania School of Arts and Sciences, Philadelphia (Barg); Penn Memory Center at the Penn Neuroscience Center, Philadelphia, Pennsylvania (Karlawish).

Author Contributions: Dr Halpern takes full responsibility for the integrity of the data and the accuracy of the data analysis. Concept and design: Halpern, Bayes, Reyes, Greysen, Volpp. Acquisition, analysis, or interpretation of data: Halpern, Chowdhury, Bayes, Cooney, Hitsman, Schnoll, Lubitz, Reyes, Patel, Greysen, Mercede, Reale, Barg, Karlawish, Stephens-Shields.

Drafting of the manuscript: Halpern, Chowdhury, Cooney, Reale.
Critical revision of the manuscript for important intellectual content: Halpern, Bayes, Hitsman, Schnoll, Lubitz, Reyes, Patel, Greysen, Mercede, Volpp, Karlawish, Stephens-Shields.
Statistical analysis: Chowdhury, Bayes, Stephens-Shields.
Obtained funding: Schnoll.
Administrative, technical, or material support: Halpern, Chowdhury, Bayes, Cooney, Hitsman, Schnoll, Lubitz, Reyes, Patel, Greysen, Mercede, Reale, Volpp, Karlawish.
Supervision: Cooney, Hitsman, Schnoll, Reale, Barg, Stephens-Shields.

Conflict of Interest Disclosures: Dr Hitsman reported nonfinancial support from Pfizer during the conduct of the study. Dr Schnoll reported nonfinancial support from Pfizer during the conduct of the study and consulting fees from Curaleaf and GSK outside the submitted work.
Effectiveness and Ethics of Incentives for Research Participation

REFERENCES


© 2021 American Medical Association. All rights reserved.
Evidence for the Ethics of Incentivizing Clinical Trial Enrollment?
Sang Ngo, BS; Anthony S. Kim, MD, MS; Winston Chiong, MD, PhD

In clinical research ethics, using monetary incentives for clinical trial enrollment has been controversial. Proponents argue that such incentives can accelerate recruitment in socially beneficial clinical trials, address problems of diversity and external validity in research participation, and appropriately compensate participants for trial-related burdens. However, many institutional review boards have limited the use or magnitude of such incentives out of concern that they could unduly influence prospective participants or unjustly shift the burdens of research participation to people with lower income levels. Yet empirical data are lacking on whether incentives have these purported positive (eg, facilitating enrollment) or negative (undue or unjust influence) effects.

In this issue of JAMA Internal Medicine, Halpern et al2 report results from 2 embedded randomized clinical trials (RCTs) intended to provide empirical evidence to inform these ethical questions. Prospective participants in 2 parent RCTs, an ambulation intervention for hospitalized patients and a smoking cessation intervention for people with major depressive disorder, were first randomized to different enrollment conditions: no monetary incentive, a smaller incentive, and a larger incentive. Incentives increased the rate of consent to enrollment in the smoking cessation trial but not in the ambulation trial. The authors also evaluated whether inducement was undue or unjust, based on the magnitude of the interactions between incentive size and perceived research risk or self-reported income, respectively, on consent to enroll in the parent RCTs. The upper limit of the confidence intervals for the observed interactions excluded a prespecified noninferiority margin (an interaction odds ratio of 2.0), which the authors interpreted as being not compatible with the presence of undue or unjust inducement.

This work is welcome, as it presents experimental data to a bioethical debate that so far has been largely driven by conjecture and competing suppositions. The authors regard their study as having settled the practical and normative debate, concluding, “Thus, research regulators should relax restrictions on the use of incentives designed to improve enrollment in low-risk trials.” However, interpreting the authors’ findings is complex and illustrates some of the challenges inherent to applying empirical data to ethical problems.

An initial challenge is at the level of definitions. Among bioethicists, there is no consensus about what counts as undue inducement or an unjust distribution of research burdens. In this article, the authors have operationalized these constructs based on their own interpretations of undue and unjust inducement, which may not capture all the concerns that scholars have raised about inducement. For example, other experts have argued that undue inducement exists when incentives are great enough to distort participants’ perception of re-