

# Development of the Health Incentive Program Questionnaire (HIP-Q) in a cardiac rehabilitation population

Marc S. Mitchell, M.Sc.,<sup>1,2</sup> Jack M. Goodman, Ph.D.,<sup>1</sup> David A. Alter, M.D./Ph.D.,<sup>2</sup> Paul I. Oh, M.D.,<sup>3</sup> Guy E. J. Faulkner, Ph.D.<sup>1</sup>

<sup>1</sup>Faculty of Kinesiology and Physical Education,  
University of Toronto, 55 Harbord St.,  
Toronto, ON M5S 2W6, Canada

<sup>2</sup>Institute for Clinical Evaluative Sciences, Toronto, ON, Canada

<sup>3</sup>Toronto Rehabilitation Institute,  
University Health Network, Toronto,  
ON, Canada

Correspondence to: M Mitchell  
marc.mitchell@uhn.ca

Cite this as: *TBM* 2015;5:443–459  
doi: 10.1007/s13142-015-0330-3

## Abstract

The purpose of this study was to develop a questionnaire to facilitate the design of acceptable financial health incentive programs. A multiphase psychometric questionnaire development method was used.

Theoretical and literature reviews and three focus groups generated a pool of content areas and items. New items were developed to ensure adequate content coverage. Field testing was conducted with a convenience sample of cardiac rehabilitation (CR) patients ( $n=59$ ) to establish face and construct validity ( $p=0.021$ ) and reliability (intraclass coefficients=0.42–0.87). The final questionnaire is comprised of 23 items. This questionnaire builds on previous attempts to explore acceptability by sampling a wider range of instrumental and affective attitudes and by measuring the effect of program features on the likelihood of incentive program participation. Future research is now needed to examine whether tailoring incentives to preferences assessed by the questionnaire improves uptake and effectiveness.

## Keywords

Exercise, Incentives, Cardiac rehabilitation, Prevention, Motivation

## INTRODUCTION

The societal costs of chronic disease are enormous. Employers bear their share of this burden as they pay more for unhealthy employees in health costs, disability, and absenteeism expenses. In 2013, for example, US employers paid \$9157 (US) per active employee in health costs—up from \$7486 in 2009 [1]. This number is expected to increase by 4.4 % (twice the rate of inflation) in 2014. In response, many employers in the USA (and elsewhere) have added wellness programs to their package of benefits. Low levels of employee engagement have unfortunately been a hallmark of these programs. Web-based wellness programs are particularly susceptible to attrition [2]. To boost engagement, two thirds of large US employers now offer financial incentives for wellness program participation [1]. Companies are forging ahead with less than optimal incentive schemes, however, limiting returns-on-investment. For example, by offering incentives in the delayed, less salient form of health insurance premium

## Implications

• **Practice:** To drive clinically and economically significant health behavior change, wellness incentives should be specially designed to maximize acceptability and uptake.

• **Policy:** Tools such as the Health Incentive Program Questionnaire are needed to optimize the development of incentive-based workplace wellness and public health policies.

• **Research:** Future research is needed to examine whether tailoring incentives to preferences assessed by the questionnaire improves uptake and effectiveness.

reimbursements (61 % of US employers do so) [1] for the attainment of hard-to-achieve biometric standards (58 % of US employers do so) [1], companies risk squandering scarce resource on weak behavioral stimuli [3, 4].

To optimize incentive program design, the range of incentive program features should be considered in the design process (see Table 1 for a list of incentive design features and attributes). To date, not enough attention has been paid to these features even though they appear to moderate effectiveness [5]. Even subtle variations in incentive design, for example, can have a profound impact on target group “acceptability” [6, 7], a critical precondition to successful incentive program implementation [4]. Since incentives for health remain a contentious topic (about half of survey respondents think they are unfair, coercive, a breach of privacy, or a waste of limited resource) [8–14], a tool is needed to assess target group acceptability in advance of implementation. This tool could be used to identify acceptability moderators and preferred incentive features. Learning more about preferred incentive program features, and how these vary for individuals and groups with shared characteristics (e.g., older employees, lower income earners), should inform the design of more refined, effective, cost-effective, and marketable (e.g., “custom incentives”) incentive programs.

**Table 1** | Financial health incentive design features and the range of attributes for each (examples in parentheses)

Features	Attributes
1. Form	(a) Cash (\$10 cash, cheque) (b) Voucher (iTunes, grocery, transit, Amazon) (c) Specific good/service (gym shoes, dietician consultation) (d) <i>Reimbursement</i> (existing expense reimbursed, like gym membership fee or health insurance premium) (e) <i>Donation</i> (value of incentive earned donated to charity of choice)
2. Magnitude	Continuous variable (often expressed as dollars (US) per week or month) <sup>a</sup>
3. Target	(a) <i>Self-regulatory behavior</i> (self-monitoring, scheduling, seeking social support) (b) Behavior (exercise, medication adherence) (c) Outcome (BMI < 25 kg/m <sup>2</sup> , BP < 140/90)
4. Timing of assessment	(a) Completion of incentive program (6 months) (b) Set intervals (daily/weekly assessments) (c) Random intervals (10 assessments over 6 months) (d) Dependent intervals (varying intervals based on previous performance)
5. <i>Type of assessment</i>	(a) <i>Self-report</i> (exercise diary submission) (b) <i>Objective, direct assessment</i> (face-to-face) (c) <i>Objective, indirect assessment</i> (pedometer, photo of weight on scale)
6. Reward immediacy <sup>b</sup>	Continuous variable (often expressed as days or weeks between assessment and reward)
7. Certainty	(a) Certain (\$50 for meeting A1C target) (b) Certain chance (1 in 4 chance of \$25) (c) Uncertain chance (1 in 100 chance of \$500) (d) Mix (\$50 and a 1 in 100 chance of \$500)
8. Schedule	(a) Uniform (\$50 lump sum for meeting goal) (b) Indexed (\$1 for each gym visit) (c) Escalating (\$1 for first 10 gym visits, \$2 for next 10, etc.) (d) Random (\$1 to \$50 for gym visits)
9. Dispensing type	(a) Resetting (discreet reward at time of each achievement) (b) Aggregative (“passbook saving”—information on running tally given) (c) <i>Mix</i> (accumulated incentives lost if discreet goal not met, “Go back to zero if missed gym visit.”)
10. Participant investment	(a) Opportunity cost only (time) (b) Deposit contract (own money lost if fail to achieve goal) (c) Matching (“double or nothing”) (\$50 of own money lost if fail, \$50 extra gained if successful)
11. Information disclosure	(a) Factual (information given about meeting or failing to meet goal) (b) Counterfactual (information given about reward lost by failing to meet goal, i.e., regret)
12. <i>Duration</i>	Continuous variable (often expressed in weeks or months incentive available; maybe indefinitely)
13. <i>Source</i>	(a) <i>Self or significant others</i> (spouse, friend) (b) <i>Group members</i> (incentive plan members) (c) <i>Government</i> (d) <i>Employer</i> (e) <i>Insurance company</i> (f) <i>Other (noninsurance) companies</i>
14. Recipient	(a) Individual (cash for weight lost) (b) Group (reward for >50 % group attendance) (c) Significant other(s) (spouse, parent) (d) <i>Charitable organization</i>

This table represents a combination of works published by Klein and Karlawish [24], Adams et al. [25], and Mitchell and Faulkner [2]. Italicized items have not been previously published

<sup>a</sup> Magnitude is ideally considered in relation to individual/group socioeconomic circumstance

<sup>b</sup> Consider when (1) behavior/outcome is assessed, (2) when it is rewarded, and (3) the time between assessment and reward

Therefore, the purpose of this study was to develop a new survey, called the Health Incentive Program Questionnaire (HIP-Q), to measure target group acceptability and identify preferred incentive program structures.

Despite conducting this study with a convenience sample of cardiac rehabilitation (CR) patients, lessons learned may be applicable to employee-directed incentives as well. Incentives that are more readily accepted

by target groups may lead to greater program uptake, increase the potential for sustained behavior change, and ultimately yield cost-effective outcomes.

## METHODS

A multiphase psychometric questionnaire development method was used to develop the HIP-Q [15]. The HIP-Q is designed specifically to optimize incentives *for exercise*, since the authors of this study planned to deploy the questionnaire for the first time in a CR context, where exercise is the cornerstone therapy. Of all the chronic disease risk factors, physical inactivity is arguably the most important [16] and has been recently shown to increase health care costs in US employees with metabolic syndrome by about 30 % [17]. Incentives can target any number of health behaviors though (e.g., medication adherence, weight loss) and the HIP-Q was designed to easily incorporate them (e.g., by replacing the word “exercise” with “take my medications” or “lose weight” in HIP-Q item #1 stem: “For me, getting paid cash or healthy vouchers to exercise/take my medications/lose weight *would be...*” A convenience sample of English-speaking CR participants was recruited from the Toronto Rehabilitation Institute’s (Toronto Rehab) to participate in the development of the HIP-Q. Participants were recruited between February and October 2013 by a researcher at the beginning of their weekly CR sessions. The sample broadly reflected the sociodemographic profile (e.g., gender, marital status, income) of cardiac patients in Ontario (see Table 2 for population and sample characteristics). Drawing from the Toronto Rehab CR population was deemed appropriate since it is suggested that high-risk, high-cost groups, such as older adults living with cardiovascular disease, should be the initial targets of incentive interventions [18, 19]. The research ethics boards of the University Health Network and the Univer-

sity of Toronto approved this study. The development of the questionnaire consisted of five phases.

### Phase 1: identifying content areas and items

#### *Step 1—literature review*

A review of relevant behavior change theories served a “heuristic purpose” [15] suggesting content areas that the authors could use to begin to shape the HIP-Q as well as to identify and/or phrase items. A systematic review of the literature examining incentives for exercise adherence in adults [20] and an overview of related papers and reviews also added to the inventory of content areas and items.

#### *Step 2—focus groups*

Three focus groups of five to six CR participants were conducted to explore opinions of incentives, and elucidate content areas and items that may not have emerged from the theoretic/literature reviews. The focus group methodologies have been previously reported [21].

### Phase 2: new item generation

#### *Step 3—drafting new items*

New items were developed to ensure adequate coverage in the HIP-Q. The 14 features of incentive programs in Table 1 were used to guide this step, ensuring all features (and attributes) were considered in the incentive design process.

### Phase 3: content validity

#### *Step 4—expert consultation*

Once new items were written, a draft of the HIP-Q and an accompanying review guide were sent to five international experts with experience conducting incentive research, writing about incentives, implementing incentives, developing surveys, and/or working with CR patients. Four out of five experts held Ph.D. degrees in health behavior change or related fields. In the review, guide experts were asked about the appropriateness and clarity of items using a 4-point Likert scale. Items with mean scores below 3 were discarded or reconsidered. Experts were also asked if the HIP-Q sampled all relevant content given its stated purpose, and to recommend additional content areas/items, if needed. Lastly, experts were asked to recommend different approaches to scaling, and to reword items, as required.

### Phase 4: face validity

#### *Step 5—pretesting*

The HIP-Q was pretested in one-on-one interviews to ensure that it was comprehensible for the target population before pilot testing it with a larger group. To explore whether the questions were clear, individuals were asked

**Table 2** | Population and sample sociodemographic characteristics

	Population <sup>a</sup> , n=1807	Sample, n=59
Age (years, mean±SD)	65.4±10.4	66.0±10.9
Female	450 (25)	13 (22)
Caucasian	1446 (83)	47 (80)
Married	1392 (78)	37 (63)
Retired	905 (52)	36 (61)
Postsecondary education	1312 (75)	43 (73)
Household income (Canadian dollars)		
<\$35,000	N/A	12 (20)
<\$65,000	N/A	27 (46)
>\$65,000	730 (50) <sup>b</sup>	32 (54)
>\$95,000	N/A	22 (37)

Numbers in parentheses represent the percent of participants within the condition (column) possessing the given attribute

<sup>a</sup> Sociodemographic characteristics of cardiac inpatients from 11 Ontario hospitals enrolled in the Cardiac Rehabilitation Care Continuity through Automatic Referral Evaluation study (1807 out of 2635 recruited patients were enrolled) [36]

<sup>b</sup> Annual family income >\$50,000 (Canadian)

to “think aloud” through their responses to identify problem items. Problem items were rewritten or eliminated. Frequency of endorsement was tested and discarding item alternatives was considered if endorsed by very few or very many (endorsement rate of 0.20 and 0.80, respectively). Pretesting continued until no new concerns arose. Time for HIP-Q completion was recorded.

#### Phase 5: construct validity and reliability

##### *Step 6—pilot testing*

The final draft of the HIP-Q (23 total items) was piloted through paper-and-pencil self-administration in a convenience sample of 59 CR patients to test construct validity and reliability. In line with self-determination theory, the authors hypothesized that individuals scoring lower on the Relative Autonomy Index (RAI; summary measure of intrinsic motivation to exercise calculated using the Behavioral Regulation to Exercise Questionnaire (BREQ-3)) [22] would favor incentive program participation, as indicated by a “likely” or “very likely” response to HIP-Q item #2: “In general, how likely would you be to participate in an incentive program that paid you \$40 a month for exercising 15 min a day, 3 days a week?”

##### Statistical analyses

The magnitude and statistical significance of the relationship between the RAI and the “likelihood of participation” response was evaluated using Pearson’s correlation coefficient. To test reliability, a 7-day test-retest was conducted and intraclass coefficients (ICC) were computed. Identifying the number of patients leaving more than 10 % of items unanswered also tested completeness of item responses, or who incorrectly answered items. All data analyses were conducted using the Statistical Package for the Social Sciences 22.0 (SPSS).

## RESULTS

### Phase 1: identifying content areas and items

#### *Step 1—literature review*

The theoretical review conducted during phase 1 highlighted the value of grounding incentives in health behavior change theory. A full outline of the theoretical considerations informing the development of the HIP-Q is reported elsewhere [23]. In keeping with self-determination theory in particular, for incentives to drive sustained health behavior change, the authors suggest incentives be designed in a way that fulfills the basic psychological needs of competence (experiencing mastery), autonomy (a sense of ownership over behavior), and/or social relatedness (feeling socially connected to others). The HIP-Q therefore included items that aimed to identify health behaviors or outcomes that prospective participants could realistically achieve (to increase confidence). The HIP-Q’s purpose was to aid in the delivery of custom incentives to increase feelings of ownership and autonomy. Last, incentives related to social outcomes (e.g., charitable donations) or that promote social

interaction (e.g., incentives for group success) were included in the HIP-Q as plausible program options. The three psychological needs described in self-determination theory were carefully considered, therefore, in the development of the HIP-Q.

The literature review undertaken by the authors [20] uncovered two papers that outlined and defined the set of incentive design features (and their associated attributes) [24, 25]. So not to neglect important features in the design process, the HIP-Q was formatted according to these published features (there are 11 in total), as well as three additional features emerging from the authors’ review [20] (i.e., type of assessment, duration of incentive program, source of incentive; see Table 1), with the aim of using the data to customize incentive programs. According to a 2013 Consensus Statement, *Guidance for a Reasonably Designed, Employer-Sponsored Wellness Program Using Outcomes Based Incentives*, to build employee acceptance all “reasonably designed” incentive programs should consider the full range of incentive approaches when looking to increase wellness program uptake and engagement [26]. In addition, a questionnaire developed and used by Long et al. to examine opinions of incentives was discovered during the literature review phase of this study [27]. This questionnaire was not validated but provided a foundation from which to build an updated, more comprehensive, and psychometrically sound incentive-focused questionnaire. The questions developed by Long et al. were used as the initial basis for the HIP-Q items.

#### *Step 2—focus groups*

The focus group results have been previously reported [21]. Briefly, a thematic analysis of the focus group data revealed that participants’ ethical concerns with incentives were prominent, but were mitigated in considering a range of program features, including source (e.g., government vs. private company) and type (e.g., cash vs. voucher) of incentive, as well as incentive target (e.g., behavior vs. outcome) (see Appendix 1 for an overview of focus group themes and acceptability moderators). Identifying the features most likely to elicit strong (negative) reactions in this sample focused the authors’ attention on key content areas (i.e., design features), ensuring that these areas/features were adequately addressed in the questionnaire.

### Phase 2: new item generation

#### *Step 3—drafting new items*

Since ethical concerns were prominent in the focus groups (consistent with the literature) [6–13], the Long et al. questionnaire was expanded using Spector’s and Ajzen’s lists of categories to include seven total pairs of instrumental (e.g., Necessary/Unnecessary) and affective (e.g., Fun/Not Fun) attitudes. A paired comparison technique was used here, where respondents were asked to indicate which attitudinal opposite they agreed with most on a 7-point Likert scale



(uneven to offer a “neutral position,” and with most points labeled to ease cognitive requirements) [15].

A paired comparison technique using 7-point Likert scales was also used to identify features that may increase the “likelihood” of incentive program participation as well as to identify preferred incentive design features. The “likelihood of participation” and incentive design preference items were deemed to be more directly relevant for employers and others interested in investing in incentives for health than broader attitudinal items. Notably, it was not suitable for every incentive design feature from Table 1 to be represented in the HIP-Q. In particular, new items exploring features #4, #8, and #9 (timing of assessment, schedule, and dispensing type) were not drafted given the overlap with feature #6 (reward immediacy). To further limit redundancy, feature #5 (type of assessment, e.g., self-report) was not explicitly represented in the HIP-Q either given similarities with feature #3 (incentive target, e.g., self-monitoring).

One categorical item was drafted to identify specific voucher preferences, since vouchers may be perceived as more acceptable and meaningful than cash alone [21]. In total, 28 new items were drafted (replacing the Long et al. items) to accommodate a more comprehensive assessment of attitudes around incentives and to determine whether acceptability varies with design features/attributes. Several steps were taken to ensure that the newly devised items were psychometrically sound including using words that do not require greater than a 6th grade reading level.

### Phase 3: content validity

#### *Step 4—expert consultation*

Mean appropriateness and clarity scores, as given by content experts, ranged from 3.2 to 4.0, and thus, no items were discarded due to low scores. Seven items were edited, as per reviewer suggestions, to increase clarity. To ease cognitive requirements, the number of response alternatives for the paired comparisons was reduced to five (from seven). The instructions and stems in this section were also edited for clarity. The depth to which certain items explored the role of design feature attributes in moderating acceptability (e.g., certain vs. uncertain rewards) was deemed to be unnecessarily complex, and potentially confusing, by three experts, and so these items were rewritten. Once recommendations from experts were incorporated, a “readability score” (7.6 Fleisch-Kincaid Grade Level) was generated using Microsoft Word. This score was interpreted with caution given some of the limitations outlined by Streiner and Norman [15].

### Phase 4: face validity

#### *Step 5—pretesting*

Eight participants completed the HIP-Q and participated in one-on-one interviews. No new concerns arose during the final three interviews and so sampling

ceased at this point. Missing values on one or more items occurred in three participants (37.5 %). Paired comparison items (for instrumental and affective attitudes) were reformatted to include *labeled check boxes* (see Fig. 1, or the full HIP-Q in Appendix 2), rather than numbers (1–5) to be circled, as the inherent values of numbers confused some of the participants (e.g., “So ‘1’ is the highest?”).

Using paired comparisons to examine the impact of *subtle* feature attribute variations on opinions confused some participants (see Fig. 2). For this reason, questions about feature attribute preferences were reformatted to simpler categorical judgments (see Fig. 1, item 4), with fewer “variations” presented, bringing the total number of HIP-Q items to 23, from 28. Endorsement rates of item alternatives did not fall outside a priori parameters and so no item alternatives were eliminated. The average time to completion was 13 min 21 s.

### Phase 5: construct validity and reliability

#### *Step 6—pilot testing*

The HIP-Q was pilot tested with CR patients through self-administration to test construct validity ( $n=59$ ) and reliability ( $n=32$ ). Seventy-one percent (17/24) of the respondents with RAIs below the group mean (i.e., more externally controlled—“I exercise because my doctor told me to.”) indicated that they would be likely/very likely to participate in an incentive program compared to 51 % (18/35) of those above the mean (e.g., “I exercise because I enjoy it.”). As well, RAI and likelihood of participation were correlated ( $p=0.021$ ) supporting the authors’ a priori hypothesis that less self-determined respondents would self-report being more likely to participate in an incentive intervention. An examination of BREQ-3 subscales yielded similar results, with 71 % of more “externally regulated” respondents (15/21, vs. 52 % of those less “externally regulated”) indicating they would be likely/very likely to participate in an incentive program. Ten (16.9 %) respondents either did not answer, or incorrectly answered, 10 % or more of the items. For instrumental and affective attitude items, the ICCs were 0.76 and 0.60, respectively. For categorical items, the ICCs ranged from 0.42 to 0.87 (see Fig. 1 for a sample of HIP-Q items and Appendix 2 for the full questionnaire).

## DISCUSSION

The aim of this study was to develop a valid and reliable questionnaire for the purpose of customizing health incentives. This is the latest attempt to develop a novel incentive design tool, the first study to consider the broad range of incentive design features in the development of such a tool, and the first to psychometrically evaluate a health incentives program questionnaire. Although this study was conducted in a CR context, there is no obvious reason that the HIP-Q cannot be applied in other contexts and for other health behaviors given its focus on core design features of incentives. Preferences may certainly vary across

1. For each **pair of words** below, check the box  that best represents how you feel about being paid to exercise.

“For me, getting paid cash or healthy vouchers to exercise **would be...**”

<b><u>Effective</u></b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<b><u>Not effective</u></b>
	Strongly Agree	Agree	Neutral	Agree	Strongly Agree	

2. In general, how likely would you be to participate in an incentive program that paid you **\$40 a month for exercising 15 minutes a day, 3 days a week**? Please circle one (1) option.

Very unlikely	Unlikely	Neutral	Likely	Very likely
1	2	3	4	5

3. “I would be **more likely** to participate in an incentive program if I was...” Please check  one (1) of the boxes.

- Paid cash
- Paid with vouchers, like grocery store or gym membership vouchers
- Able to donate my incentive to my favourite charity
- The ‘type’ of incentive doesn’t matter to me
- I would **NOT** participate in an incentive program

*If you were being paid to exercise for 15 minutes a day, 3 days a week, for a month, which incentive program ‘feature’ below would you prefer?*

4. The ‘**guaranteed**’ or the ‘**lottery**’ incentive feature? Please check  one (1) of the boxes.

- Get paid \$40 for sure – ‘guaranteed’ incentive
- Have a 1 in 10 chance of winning \$300 – ‘lottery’ incentive
- I don’t have a preference
- I would **NOT** participate in an incentive program

Fig 1 | Sample Health Incentive Program Questionnaire (HIP-Q) items

populations and contexts, and further validation work will be needed to demonstrate this.

**Psychometric properties**

The HIP-Q demonstrated content, face, and construct validity. Informed by the extant literature and expert review, the HIP-Q adequately covers the relevant information. As well, items were interpretable by the target population during pretesting, and pilot testing demonstrated that the HIP-Q (item #2) is significantly related to RAI (calculated using the BREQ-3) [22], consistent with

self-determination theory, increasing confidence in responses. HIP-Q test-retest reliability was partly supported as well, with 12 out of 23 items demonstrating “Good” reliability (ICC≥0.7). Since the purpose of the HIP-Q was to assess as many design features as possible, items with less than satisfactory reliability (n=11; ICC<0.7) were not discarded. Those interested in implementing incentives should interpret item responses with caution until further validation is conducted. Notably, affective attitude items (e.g., Good vs. Bad) yielded different responses over time (ICC=0.60). Affective attitudes around incentives may be nebulous, changing over time, perhaps with the presentation of new information, or in

1. For each **pair of words** use the scale to indicate which one best represents how you would complete the sentences below.

If I were participating in a **weekly** incentive program (reward paid out each week), I would prefer...

	Strongly prefer		Don't care		Strongly prefer	
\$10 for sure	1	2	3	4	5	1 in 4 chance (25%) for \$40
\$10 for sure	1	2	3	4	5	1 in 4 chance (25%) for <u>\$35</u>
\$10 for sure	1	2	3	4	5	1 in 10 chance (10%) for \$100
\$10 for sure	1	2	3	4	5	1 in 10 chance (10%) for <u>\$90</u>

**Fig 2 |** Item from the “pretesting” draft of the Health Incentive Program Questionnaire using paired comparisons to examine the impact of subtle feature attribute variations on incentive program preferences

different settings, or with more time for personal reflection on a contentious topic. Before drawing firm conclusions regarding the reliability of these items, further study is warranted with a larger sample.

#### Application

The HIP-Q is a comprehensive incentive design tool that has several potential applications. Since low program uptake is a barrier to successful implementation, the HIP-Q may be used to identify the overall acceptability of interventions. Although effectiveness was not tested in this article, the authors presume that higher acceptance of incentive designs may lead to greater effectiveness, as has been suggested [4]. Not only may the HIP-Q be used to identify perceived levels of effectiveness and acceptability (instrumental and affective attitudes), but it may also be used to establish how likely individuals would be to sign-up for an incentive program. The HIP-Q allows for the identification of feature attributes that may boost likelihood of participation as well, providing incentive program sponsors with information to customize incentive packages so they are more readily accepted by target groups. Building a repository of incentive program design preferences over time and linking these to sociodemographic, health status, and health behavior characteristics may help segment incentive interventions in the future.

The HIP-Q may also help shed light on the question of “incentive direction” (i.e., Should companies

implement financial health incentives, or penalties?). For instance, HIP-Q items #6 and #7 ask respondents how likely they would be to “wager” their own money in an incentive program (called a “deposit contract”). The answers to these questions may give companies a sense of how willing employees would be to pay an enrollment fee, with the chance to earn their money back in the program. This incentive structure is becoming increasingly common [1] and is the one championed in the Patient Protection and Affordable Care Act—where employees can earn back up to 50 % of their health insurance premium (their “deposit,” so to speak) [28]. Regarding the implementation of a financial penalty over and above the cost of insurance, our position is that penalties are more likely to generate resistance [29], limit enrolment [29, 30], discriminate disadvantaged groups [31], and undermine intrinsic motivation [23]—damaging the potential for sustained health behavior change [32].

Regarding tailored incentive programs, a growing body of research is examining how individual characteristics (e.g., age, income, confidence to exercise, weight status, consumer habits/preferences) moderate incentive effectiveness, and how these characteristics interact with incentive design features/attributes to produce health behavior change. For instance, John et al. determined in a sample of low and high income adults that the higher income individuals were more sensitive to lottery-based (vs. certain) and voucher-type (vs. cash) incentives compared to their lower

income counterparts [33]. As this body of research develops, interventionists will be in a better position to match individual characteristics to preferred and/or more effective incentive design features/attributes.

For instance, in the future, older adults may be offered “certain chance” incentives (e.g., 1 in 5 chance of winning \$25) given their suspected inclination toward lottery-based interventions [24]; higher income individuals may receive larger incentives (1.2 % or more of disposable income, as has been suggested) [3]; individuals identified as less confident in their ability to exercise may be offered incentives for more achievable, “self-regulatory” behaviors (e.g., wearing a pedometer) as opposed to the attainment of difficult to achieve biometric outcomes (e.g., lose 10 lb) [23]; overweight adults could receive escalating incentives to drive regular exercise over longer periods [34]; and individuals preferring grocery store over iTunes vouchers may receive the credit they prefer and value the most [35]. Collecting relevant sociodemographic and health-related information in future studies will assist with the matching of personal characteristics with more promising incentive approaches.

#### Study limitations

The HIP-Q builds on previous attempts to explore incentive acceptability by sampling a wider range of instrumental and affective attitudes and by measuring the effect of program features, including type, source, timing, and certainty of incentive, on the likelihood of incentive program participation. It was not suitable to have all 14 features from Table 1 represented in the HIP-Q, however. Asking prospective participants about all possible incentive design subtleties proved challenging, in part because it was difficult to fully explore feature nuances in a succinct questionnaire and phrase items in a way that was comprehensible to the target group. Rather, the HIP-Q ended up focusing on those features most likely to moderate opinions in a Canadian CR population [21] and increase probability of incentive program participation.

Owing to the complexity of some of the items, the final iterations of the HIP-Q were simplified (using more general categorical judgements) to maintain the psychometric qualities of the questionnaire. Several questions exploring the subtleties of incentive program design were omitted on account of their perceived complexity, leaving several features only superficially explored (#1–3, 6–7, 10, 12–13). Nonetheless, this questionnaire is the first to the authors’ knowledge to examine the impact of multiple incentive program features on acceptability and, in this sense, makes a novel contribution to wellness incentive programming and research/evaluation.

The following limitations should also be noted. Regarding low test-retest reliability during the pilot testing phase of this study, HIP-Q items were completed *following* the completion of several related questionnaires (demographic and health-related surveys), and thus, responder fatigue may partially account for this.

For the affective attitude items, the authors suggest that reliability was low because opinions actually shifted over time, rather than CR patients not fully comprehending the questions, since similarly phrased instrumental attitude items demonstrated “Good” reliability. More research examining affective attitudes is needed. The study sample of CR patients was a convenience sample, and thus, the generalizability of the psychometric properties of the HIP-Q is limited. Certainly, as incentives grow in popularity, it will be worth testing the questionnaire among people in different settings (e.g., younger employees in different sectors), especially within the context of workplace wellness programs targeting multiple health behaviors.

Though several steps were taken to ensure that HIP-Q items were psychometrically sound, they are not without limitation. For example, the stem for item #1 includes both “cash” and “healthy vouchers” which may be problematic for two reasons. First, cash and vouchers may not be equally acceptable due to (a) dead weight loss of the voucher if it is for an item the recipient does not value as much as the giver and (b) time discounting associated with future use of the voucher vs. immediate value of cash. Second, the use of the positive word “healthy” before voucher is not balanced by a similarly positive word before cash. Although these issues cannot be fixed post hoc, they should be acknowledged as limitations. Future validation studies will aim to optimize the psychometrics of the HIP-Q and maximize its generalizability.

#### CONCLUSIONS

Financial health incentive programs should be carefully designed, considering the range of available features and attributes in the design process, as well as the impact of contextual factors on incentive acceptability and effectiveness. Even subtle variations in incentive program design can have profound effects. The newly developed HIP-Q has the potential to be a useful tool for assessing attitudes of incentives and examining the role of design features in moderating acceptability and uptake. Taken together, the HIP-Q may serve as a practical incentive design tool, used to increase financial health incentive program enrolment and participation. Further research is now needed to examine whether tailoring incentives to preferences assessed by the HIP-Q improves uptake and effectiveness.

**Acknowledgments:** This work was supported by the Canadian Institutes of Health Research (CIHR) [grant number 305843], the Ontario Centres of Excellence (member of the Ontario Network of Excellence), and Cookson James Loyalty Inc. Guy Faulkner is supported by a CIHR-Public Health Agency of Canada (CIHR-PHAC) Chair in Applied Public Health. The authors acknowledge the contributions of Karen Dobson and the rest of the cardiac rehabilitation program staff at Toronto Rehab for supporting this study. We also thank the study participants for taking the time to participate and the reviewers for their helpful insight.

**Conflict of interest:** Cookson James Loyalty Inc., the private sector funding partner, did not influence the design, collection, or interpretation of the data presented. In addition, the authors have no fiduciary interest in the company.



**Compliance with ethical standards:** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix 1

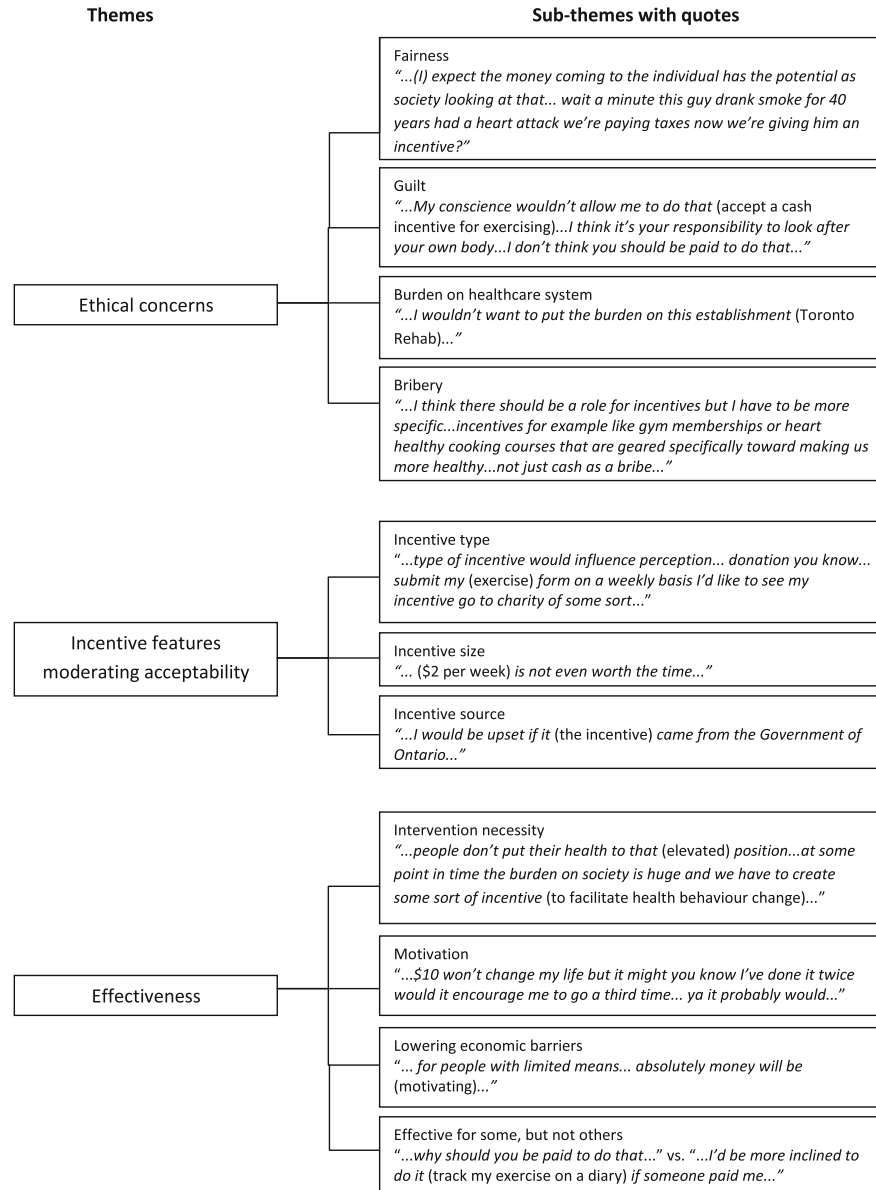


Fig 3 | Themes, sub-themes and illustrating quotes from Mitchell et al. [2] focus group study

**Appendix 2. The complete Health Incentive Program Questionnaire**

**The Health Incentive Program Questionnaire (HIP-Q)**

It may be that a financial health incentive, like getting paid to exercise, could help you start and/or maintain an exercise program. You could be paid in cash, or

with healthy vouchers, to do your exercise. You could earn grocery or drug store vouchers, gym discounts, magazine subscriptions or even make charitable donations for exercising regularly! The answers you give in this questionnaire are very important and will help us design a health incentive program *just for you!*

**Your opinion**

1. For each **pair of words** below, check the box  that best represents how you feel about being paid to exercise.

“For me, getting paid cash or health-related vouchers to exercise **would be...**”

a) 

<b><u>Good</u></b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<b><u>Bad</u></b>
	Strongly Agree	Agree	Neutral	Agree	Strongly Agree	

b) 

<b><u>Not fun</u></b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<b><u>Fun</u></b>
	Strongly Agree	Agree	Neutral	Agree	Strongly Agree	

c) 

<b><u>Effective</u></b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<b><u>Not effective</u></b>
	Strongly Agree	Agree	Neutral	Agree	Strongly Agree	

d) 

<b><u>Silly</u></b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<b><u>Sensible</u></b>
	Strongly Agree	Agree	Neutral	Agree	Strongly Agree	

e) 

<b><u>Useful</u></b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<b><u>Useless</u></b>
	Strongly Agree	Agree	Neutral	Agree	Strongly Agree	

f) 

<b><u>Necessary</u></b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<b><u>Unnecessary</u></b>
	Strongly Agree	Agree	Neutral	Agree	Strongly Agree	

## Question 1 continued...

“For me, getting paid cash or health-related vouchers to exercise **would make me feel...**”

g)

<b><u>Guilty</u></b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<b><u>Not guilty</u></b>
	Strongly Agree	Agree	Neutral	Agree	Strongly Agree	

2. In general, how likely would you be to participate in an incentive program that paid you **\$40 a month for exercising 15 minutes a day, 3 days a week?**

Please circle one (1) option.

<b>Very unlikely</b>	<b>Unlikely</b>	<b>Neutral</b>	<b>Likely</b>	<b>Very likely</b>
1	2	3	4	5

*Incentive programs can be designed in different ways. The next questions asks which design ‘features’ would make you more likely to participate in an incentive program that paid you **\$40 a month for exercising 15 minutes a day, 3 days a week.***

3. “I would be more likely to participate in an incentive program if incentives were...” Please check  one (1) of the boxes.

- Paid for by the government
- Paid for by a company I trust
- Who pays for the incentive doesn’t matter to me
- I would **NOT** participate in an incentive program

4. “I would be more likely to participate in an incentive program if I was...” Please check  one (1) of the boxes.

- Paid cash
- Paid with vouchers, like grocery store or gym membership vouchers
- Able to donate my incentive to my favourite charity
- The ‘type’ of incentive doesn’t matter to me
- I would **NOT** participate in an incentive program

5. “I would be more likely to participate in an incentive program if I was paid to...” Please check  one (1) of the boxes.

- Do certain behaviours, like exercise regularly (15 minutes a day, 3 days a week)
- Achieve certain outcomes, like improve my fitness (the amount of time I can exercise)
- The incentive ‘target’ doesn’t matter to me
- I would NOT participate in an incentive program

6. **How likely would you be to wager your own money in an incentive program?** For example, you get \$40 of your own money back if you meet the goal of exercising 15 minutes a day, 3 days a week, for a month. If you do not meet the goal, your \$40 goes to other program participants.

Please circle one (1) option below.

<b>Very unlikely</b>	<b>Unlikely</b>	<b>Neutral</b>	<b>Likely</b>	<b>Very likely</b>
1	2	3	4	5

7. **How likely would you be to wager your own money if it meant that you could double your money?** For example, if you wagered \$40 of your own money then you would receive \$80 back if you exercised 15 minutes a day, 3 days a week, for a month - called ‘double or nothing’. If you do not meet the goal, your \$40 goes to other program participants.

Please circle one (1) option below.

<b>Very unlikely</b>	<b>Unlikely</b>	<b>Neutral</b>	<b>Likely</b>	<b>Very likely</b>
1	2	3	4	5

8. What is the **MINIMUM** amount you would need to be paid in cash or vouchers, if any, **each month** to motivate you to exercise for 15 minutes a day, 3 days a week? Please circle one (1) monthly dollar amount.

\$0	\$10	\$20	\$30	\$40	\$50	\$60	\$70	\$80 or more
-----	------	------	------	------	------	------	------	--------------

If you were being paid to exercise for 15 minutes a day, 3 days a week, for a month, which incentive program *'feature'* below would you prefer?

9. a) The **'daily'** or the **'lump sum'** incentive? Please check  one (1) of the boxes.

- Get paid \$3.33 right away on the days that I do my exercise (can earn up to \$40 in a month)
- Wait until the end of the month to get paid a 'lump sum' of up to \$40 (get paid at the end of the month for the days you exercised)
- I don't have a preference
- I would NOT participate in an incentive program

b) The **'daily'** or the **'lump sum'** incentive? Please check  one (1) of the boxes.

- Get paid \$3.33 right away on the days that I do my exercise (can earn up to \$40 in the month)
- Wait until the end of the month to get paid a 'lump sum' of up to **\$30 (a little less)**
- I don't have a preference
- I would NOT participate in an incentive program

c) The **'daily'** or the **'lump sum'** incentive? Please check  one (1) of the boxes.

- Get paid \$3.33 right away on the days that I do my exercise (can earn up to \$40 in the month)
- Wait until the end of the month to get paid a 'lump sum' of up to **\$50 (a little more)**
- I don't have a preference
- I would NOT participate in an incentive program



*If you were being paid to exercise for 15 minutes a day, 3 days a week, for a month, which incentive program 'feature' below would you prefer?*

10. The **'guaranteed'** or the **'lottery'** incentive? Please check  one (1) of the boxes.

- Get paid \$40 for sure – 'guaranteed' incentive
- Have a 1 in 10 chance of winning \$300 – 'lottery' incentive
- I don't have a preference
- I would NOT participate in an incentive program

11. a) The **'individual'** or the **'group'** incentive? Please check  one (1) of the boxes.

- Get paid \$40 if I exercise 15 minutes a day, 3 days a week, for a month
- Get paid \$40 if at least half the people in my 'group' (group from work or in your exercise class, for example) exercise 15 minutes a day, 3 days a week, for a month
- I don't have a preference
- I would NOT participate in an incentive program

b) The **'individual'** or the **'group'** incentive? Please check  one (1) of the boxes.

- Get paid \$40 if I exercise 15 minutes a day, 3 days a week, for a month
- Get paid **\$50 (a little more)** if at least half the people in my 'group' exercise 15 minutes a day, 3 days a week, for a month
- I don't have a preference
- I would NOT participate in an incentive program

**\*Go to question 12 on the next page.**

12. Please indicate how you would prefer to receive your incentive, if at all. Please check  one (1) of the following boxes.

- In person pick-up (you go to a specific location to pick up your payment)
- By mail (received within 5 to 7 days)
- By courier (received within 1 to 2 days)
- By email transfer through your internet banking (received immediately)
- Using a special website, like the AIRMILES website (received immediately)
- Using your 'smart phone' or 'tablet' (received immediately)
- I don't have a preference
- I would NOT participate in an incentive program.

13. How long, if at all, would you want to participate in an incentive program? Please check  one (1) of the following boxes.

- 1-3 months
- 4-6 months
- 7-12 months
- More than 1 year
- I don't have a preference
- I would not participate in an incentive program

**\*Go to question 14 on the next page.**



- change health related behaviour. *Soc Sci Med*. 2012; 75(12): 2509-2514. doi:10.1016/j.socscimed.2012.09.033.
8. Bonevski B, Bryant J, Lynagh M, et al. Money as motivation to quit: a survey of a non-random Australian sample of socially disadvantaged smokers' views of the acceptability of cash incentives. *Prev Med*. 2012; 55(2): 122-126. doi:10.1016/j.ypmed.2012.06.001.
  9. Bonevski B, Bryant J, Paul C. Encouraging smoking cessation among disadvantaged groups: a qualitative study of the financial aspects of cessation. *Drug Alcohol Rev*. 2011; 30(4): 411-418. doi:10.1111/j.1465-3362.2010.00248.x.
  10. Lynagh M, Bonevski B, Symonds I, et al. Paying women to quit smoking during pregnancy? Acceptability among pregnant women. *Nicotine Tob Res*. 2011; 13(11): 1029-1036. doi:10.1093/ntr/ntr108.
  11. Park JD, Metlay J, Asch JM, et al. The New York Times readers' opinions about paying people to take their medicine. *Health Educ Behav*. 2012; 39(6): 725-731. doi:10.1177/1090198111428645.
  12. Park JD, Mitra N, Asch DA. Public opinion about financial incentives for smoking cessation. *Prev Med*. 2012; 55(Suppl): S41-S45. doi:10.1016/j.ypmed.2012.06.013.
  13. Promberger M, Brown RCH, Ashcroft RE, et al. Acceptability of financial incentives to improve health outcomes in UK and US samples. *J Med Ethics*. 2011; 37(11): 682-687. doi:10.1136/jme.2010.039347.
  14. Blondon K, Klasnja P, Coleman K, et al. An exploration of attitudes toward the use of patient incentives to support diabetes self-management. *Psychol Health*. 2014; 29(5): 552-563. doi:10.1080/08870446.2013.867346.
  15. Streiner D, Norman G. *Health Measurement Scales: A Practical Guide to Their Development and Use*. 4th ed. New York: Oxford University Press; 2008.
  16. Lee IM, Shiroma EJ, Lobelo F, et al. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet*. 2012; 380(9838): 219-229. doi:10.1016/S0140-6736(12)61031-9.
  17. Burton WN, Chen CY, Li X, et al. The association of self-reported employee physical activity with metabolic syndrome, health care costs, absenteeism, and presenteeism. *J Occup Environ Med*. 2014; 56(9): 919-926.
  18. Ades PA, Gaalema DE. Coronary heart disease as a case study in prevention: potential role of incentives. *Prev Med*. 2012; 55(Suppl): S75-S79.
  19. Loewenstein G, Asch DA, Volpp KG. Behavioral economics holds potential to deliver better results for patients, insurers, and employers. *Health Aff*. 2013; 32(7): 1244-1250. doi:10.1377/hlthaff.2012.1163.
  20. Mitchell M, Goodman JM, Alter DA, et al. Financial incentives for exercise adherence in adults: systematic review and meta-analysis. *Am J Prev Med*. 2013; 45(5): 658-667.
  21. Mitchell MS, Goodman JM, Alter DA, et al. 'Will walk for groceries': acceptability of financial health incentives among Canadian cardiac rehabilitation patients. *Psychol Health*. 2014; 29(9): 1032-1043. doi:10.1080/08870446.2014.904863.
  22. Wilson PM, Rodgers WM, Loitz CC, et al. "It's who i am...really!" The importance of integrated regulation in exercise contexts. *J Appl Biobehav Res*. 2006; 11(2): 79-104.
  23. Mitchell M, Faulkner G. A "nudge" at all? The jury is still out on financial health incentives. *Healthc Pap*. 2012; 12(4): 31-36.
  24. Klein E, Karlawish J. Challenges and opportunities for developing and implementing incentives to improve health-related behaviors in older adults. *J Am Geriatr Soc*. 2010; 58(9): 1758-1763. doi:10.1111/j.1532-5415.2010.03030.x.
  25. Adams J, Gilles EL, McColl E, et al. Carrots, sticks and health behaviours: a framework for documenting the complexity of financial incentive interventions to change health behaviours. *Health Psychol*. 2013; 8(3): 100-110.
  26. Consensus Statement of the Health Enhancement Research Organization, American College of Occupational and Environmental Medicine, American Cancer Society and American Cancer Society Cancer Action Network, et al. Guidance for a reasonably designed, employer-sponsored wellness program using outcomes-based incentives. *Occup Environ Med*. 2012; 54(7): 889-896. doi:10.1097/JOM.0b013e3182620214.
  27. Long JA, Helweg-Larsen M, Volpp KG. Patient opinions regarding 'pay for performance for patients'. *J Gen Intern Med*. 2008; 23(10): 1647-1652. doi:10.1007/s11606-008-0739-1.
  28. Patient Protection and Affordable Care Act of 2010 (PPACA), Pub. L. no. 111-148 § Sec. 3143; 2010.
  29. Volpp KG, Galvin R. Reward-based incentives for smoking cessation: how a carrot became a stick. *JAMA*. 2014; 311(9): 909-910. doi:10.1001/jama.2014.418.
  30. Farooqui MA, Tan YT, Bilger M, et al. Effects of financial incentives on motivating physical activity among older adults: results from a discrete choice experiment. *BMC Public Health*. 2014; 14: 141.
  31. Madison K, Schmidt H, Volpp KG. Smoking, obesity, health insurance, and health incentives in the Affordable Care Act. *JAMA*. 2013; 310(2): 143-144. doi:10.1001/jama.2013.7617.
  32. Promberger M, Marteau TM. When do financial incentives reduce intrinsic motivation? Comparing behaviors studied in psychological and economic literatures. *Health Psychol*. 2013; 32(9): 950-957. doi:10.1037/a0032727.
  33. John LK, Loewenstein G, Volpp KG. Empirical observations on longer-term use of incentives for weight loss. *Prev Med*. 2012; 55(Suppl): S68-S74. doi:10.1016/j.ypmed.2012.01.022.
  34. Jeffery RW, Wing RR, Thorson C, et al. Use of personal trainers and financial incentives to increase exercise in a behavioral weight-loss program. *J Consult Clin Psychol*. 1998; 66(5): 777-783. doi:10.1037//0022-006x.66.5.777.
  35. Hunter RF, Tully MA, Davis M, et al. Physical activity loyalty cards for behavior change: a quasi-experimental study. *Am J Prev Med*. 2013; 45(1): 56-63. doi:10.1016/j.amepre.2013.02.022.
  36. Shanmugasagaram S, Oh P, Reid RD, et al. Cardiac rehabilitation barriers by rurality and socioeconomic status: a cross-sectional study. *Health Int*. 2013; 12: 72.