Effectiveness and cost-effectiveness of incentives as a tool for prevention of non-communicable diseases: A systematic review

Finkelstein, Eric; Bilger, Marcel; Baid, Drishti

Published in:
Social Science and Medicine

DOI:
10.1016/j.socscimed.2019.05.018

Published: 01/01/2019

Document Version
Publisher's PDF, also known as Version of record

Social Science & Medicine 232 (2019) 340–350

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Review article

Effectiveness and cost-effectiveness of incentives as a tool for prevention of non-communicable diseases: A systematic review

Eric A. Finkelstein[a,*], Marcel Bilger[a,b], Drishti Baid[a]

[a] Programme in Health Services and Systems Research, Duke-NUS Medical School, Singapore
[b] Vienna University of Economics and Business, Austria

ARTICLE INFO

Keywords:
Non-communicable disease
Smoking
Weight loss
Diet
Alcohol
Physical activity
RCT
Cost-effectiveness
Incentive

ABSTRACT

The rising epidemic of non-communicable diseases (NCDs) poses substantial health and economic challenges to both individuals and society. Application of incentive-based strategies based on traditional and behavioural economic theory has emerged as a potential strategy to address rising rates of NCDs. Yet, whether or not incentives truly represent a promising strategy for addressing NCDs has not been systematically addressed nor is it clear whether certain behavioural economic strategies outperform others or simply offering a cash-based incentive for meeting a goal. In this systematic review we aim to determine whether there is an evidence base for any of these strategies. Forty-eight published randomized controlled trials (70 contrasts) evaluating the effectiveness of incentive-based strategies for improvements in NCD risk-factors were reviewed. Our primary conclusion is that there is a lack of compelling evidence that incentives of any form represent a compelling NCD reduction strategy. More evidence for long-term effectiveness and cost-effectiveness is needed to justify third party funding of any incentive based strategy.

1. Introduction

Non-communicable diseases (NCDs) pose substantial health and economic challenges both at the individual and societal levels. NCDs account for an estimated 38 million deaths annually, of which 16 million are premature death (World Health Organization, 2015). Recognizing this as an "invisible epidemic", the World Health Organisation seeks to reduce premature deaths from NCDs by one-third by 2030 (World Health Organization, 2017).

NCD prevention efforts focus primarily on four modifiable behavioural risk factors - physical inactivity, unhealthy diet, and the harmful use of alcohol and tobacco (World Health Organization, 2015). These four risk factors promote four key metabolic/physiological changes (raised blood glucose, blood pressure, cholesterol and obesity) that in turn increase risks for at least four NCDs, including cardiovascular diseases, cancers, respiratory diseases, and diabetes. Given the increasing prevalence of NCDs across all age groups and a disproportionate burden on low and middle-income countries, there is a pressing need for low cost, cost-effective, and scalable NCD prevention strategies.

The use of (dis)incentives to encourage healthy behaviours, either alone or as part of a multi-component intervention, is one strategy under consideration in both lower and higher income settings (Gaarder et al., 2010; Health Promotion Board (Singapore), 2014; Lagarde et al., 2007, 2009; Ranganathan and Lagarde, 2012; USDA, 2018). These could take the form of monetary or non-monetary rewards or penalties for meeting a pre-defined process or outcome measure, such as attending a fitness session or losing weight, or less frequently, a penalty if the goal is not met. Why should we expect incentive-based strategies to work? Classical economic theory assumes that individuals are rational, forward-looking and weigh the costs and benefits of their decisions. If an individual is given a reward (penalty) for engaging in a healthy (unhealthy) behaviour, the theory predicts she will be more (less) likely to do it. Theory also predicts that, ceteris paribus, incentives are more likely to be effective among lower-income populations both because they are likely to have a lower opportunity cost of their time and because the value of the reward represents a larger percentage increase in disposable income available to the participant (Sutherland et al., 2008). However, if an intervention already includes several effective components, then the opportunity cost of even greater changes in behaviour will be that much more difficult, and therefore, we would expect the marginal benefit of incentives to be smaller, all else equal.

Behavioural economists and psychologists recognize that not everyone is rational and many individuals may not perceive a clear cause-and-effect relationship between healthy behaviours and their decreased likelihood of developing a NCD, which may not occur until well into the

* Corresponding author. Programme in Health Services and Systems Research, Duke-NUS Medical School, 8 College Road, Level 4, 169857, Singapore.
E-mail address: eric.finkelstein@duke-nus.edu.sg (E.A. Finkelstein).

https://doi.org/10.1016/j.socscimed.2019.05.018
Received 20 March 2019; Received in revised form 14 May 2019; Accepted 16 May 2019
Available online 17 May 2019

future (Chapman et al., 2001). Incentives address this problem by providing individuals with a short-term reward for health behaviours. These researchers offer additional insights into when incentives may – or may not – work as intended. In doing so, the theory points to several aspects of intervention design that could influence their short and long-term cost-effectiveness. These include design features concerning the size, type, frequency, and duration of incentive payments. The theory also warns us to "pay enough or don't pay at all" (Gneezy and Rustichini, 2000) based on the idea that incentives that are too small may confer a signal of low value and be worse than offering no incentive at all.

Many incentive-based interventions include a lottery component. However, because individuals are risk averse (Kahneman and Tversky, 1979), a guaranteed reward should be more effective at influencing behaviour than a lottery with the same expected value. However, if two lotteries of equal expected value are offered, prospect theory posits that the lottery with the lower probability but a larger pay-out will be more effective, because people overweight small probabilities and underweight moderate and large probabilities.

The type of incentive may also matter. Classical economists argue that cash is best because it is the most fungible. However, non-cash incentives have been posited to be more effective than cash if they are earmarked to pleasurable (hedonic) rewards (Jeffrey, 2009; Jeffrey and Shaffer, 2007; Kelly et al., 2015; R. Thaler, 1985; R. H. Thaler, 1999). Loss aversion predicts that individuals prefer avoiding losses to acquiring equivalent gains, suggesting that deposit contracts should be more effective than rewards of equal value. How often incentives are paid and the duration of rewards programs are two additional considerations. Hyperbolic discounting suggests that individuals discount the future to a larger extent in the short term than in the longer term (Read and Van Leeuwen, 1998; R. Thaler, 1981). Therefore frequent rewards that focus on short-term behaviours or outcomes may be effective at encouraging individuals to overcome the temptation to engage in present-biased behaviour.

NCD prevention ultimately requires forming a habit of healthy behaviours. Classical economic theory suggests that individuals who make a behaviour change due to the prospect of rewards will be less likely to engage in the behaviour once it is no longer incentivised. Contrarily, the psychological theory of operant conditioning suggests that modified behaviours can be sustained by repeatedly associating a reward (or punishment) with that behaviour (Skinner, 1963). This idea is echoed in more recent research suggesting that creating a loop of "cue-routine-reward" can lead to habit formation (Duhigg, 2012). The habit formation literature suggests that interventions with longer durations, with 24 weeks as an oft-cited time frame (Mantzari et al., 2015; Volpp et al., 2009), are more likely to produce lasting effects (Lally et al., 2008; Lally and Gardner, 2013; Neal et al., 2006; Ouellette and Wood, 1998; Wood et al., 2005).

The above-mentioned aspects of incentive design have the potential to influence their effectiveness, costs, and cost-effectiveness. Although many review articles and meta-analyses have focused on the effectiveness of incentives and their ability to influence health behaviours, to date only two studies, both meta-analyses, have attempted to address the relative effectiveness of different types of incentive features targeting multiple risk factors for NCDs (Giles et al., 2014; Mantzari et al., 2015). These studies addressed reward size (Giles et al., 2014; Mantzari et al., 2015), certainty vs. lottery rewards (Mantzari et al., 2015), participant deprivation levels (in terms of their income, employment or socio-economic status) (Mantzari et al., 2015) and length of post-intervention follow-up (Giles et al., 2014).

As with other studies that focus on single risk factors for NCDs, both reviews conclude that incentives work to positively influence health behaviours. However, in contrast to classical economic theory, both studies find that larger rewards only work better in the case of smoking cessation. Mantzari et al. (2015) finds no evidence that lottery payments are inferior to cash but offers some evidence that incentive effects are greater among more deprived participants (only at > 6–12 months from baseline). Concerning habit formation, Mantzari et al. (2015) notes that incentive effects are sustained for up to 3 months post-incentive removal and Giles et al. (2014) find that effect sizes decrease with post-intervention follow-up duration.

Although the prior meta-analyses (Giles et al., 2014; Mantzari et al., 2015) offer some insights, a careful look at the incentive design across the included studies shows significant heterogeneity in many factors, including how the dependent variable is measured, the target population, whether the incentive strategy is used alone or in conjunction with other intervention features, in addition to the size, frequency, duration, and type of reward offered. To account for these differences, the authors made several simplifying assumptions in some cases (e.g., Mantzari et al. (2015) classified rewards as high vs low using a single threshold value that is independent of duration) or completely ignored differences in others, such as frequency of reward payments. As a result, their results may suffer from both a lack of statistical power and systematic bias (e.g., studies with higher reward values also had longer durations). Moreover, the reviews do not answer the primary question posed here, which is whether or not an investment in any of the above strategies would represent a good use of scarce resources as a means for addressing rising rates of NCDs.

The purpose of this systematic review is to revisit and synthesize the evidence of effectiveness and cost-effectiveness of incentives specifically as a tool for NCD prevention. Rather than attempting to meta-analyse the results across very heterogeneous studies, we provide an updated and more comprehensive synthesis of components of intervention design aspects that could potentially influence the effectiveness of incentive strategies based on classical and behavioural economic/psychology theory. In doing so, we also consider a wider evidence base for the effectiveness of economic incentives by (1) placing no restrictions on the type of economic incentive offered (Giles et al., 2014) and (2) including studies that incentivise weight loss given its link to diet, physical activity, and NCDs and the fact that it is a common target for incentives. Unlike prior reviews, we also consider how the frequency of rewards could influence their effectiveness and whether use of incentives may be cost-effective for prevention and treatment of NCDs. This is critical as evidence of cost-effectiveness would strengthen the argument for subvention of incentive-based programs.

Specifically, we provide a detailed synthesis of the evidence for effectiveness and cost-effectiveness of incentives for healthy eating, physical activity, weight loss, smoking cessation and alcohol consumption overall and with respect to the following aspects of intervention design:

i. Target population and setting,
ii. Independent strategy or part of a multicomponent intervention,
iii. Incentive type,
iv. Incentive size,
v. Frequency of opportunity to earn incentives,
vi. Intervention duration, and
vii. Evidence for habit formation among effective studies.

To keep the review tractable and minimize potential sources of bias, we limited it to randomized controlled trials with a measured primary endpoint. We also conduct a quality assessment for all included studies in efforts to identify potential biases and to understand why results may or may not be consistent with the theory. Based on the results of the review we conclude with a summary of the evidence as it pertains to NCD risk reduction and provide recommendations for future research.

2. Methods

This review was conducted in line with PRISMA guidelines for systematic reviews (Moher et al., 2009). The protocol for the review was registered on PROSPERO (CRD42018100638).

2.1. Data sources and search strategy

Two reviewers searched for all English-language published articles from January 1995 until June 2018 across the following four databases: PUBMED (accessed via National Centre for Biotechnology Information), PsycINFO (accessed via OvidSP), EconLit (accessed via Ovid) and EMBASE (accessed via Elsevier). Databases were searched using terms identified from title, abstract, keywords or medical subject headings. The search took the general form of ([incentive type] AND [target behaviour]). For example, the search used for physical activity is described here:

('financial incentive*' OR 'incentive*' OR 'economic incentive*' OR 'monetary incentive*' OR 'cash' OR 'lotter*' OR 'lotto*' OR 'deposit*' OR 'deposit contract*' OR 'voucher*' OR 'ticket*' OR 'bonus program*' OR 'bonus*' OR 'cheque*' OR 'notes*' OR 'monetary benefit*' OR 'cash benefit*' OR 'financial benefit*' OR 'reward*' OR 'financial reward*' OR 'monetary reward*' OR 'cash reward*' OR 'prize*' OR 'cash prize*' OR 'money' OR 'non-cash reward*') AND ('physical activ*' OR 'physical inactiv*' OR 'insufficient physical activ*' OR 'movement*' OR 'bodily movement*' OR 'exercise*' OR 'exercise behavio*r' OR 'physical training*' OR 'training*' OR 'fitness' OR 'sport*' OR 'walk*' OR 'gym' OR 'motor activ*' OR 'physical exercise*' OR 'acute exercise*' OR 'aerobic exercise*' OR 'exercise training*' OR 'run*')

The search terms and MeSH terms used were adapted for the individual databases need. The full search strategy for each target behaviour can be found in Appendix Table 1.

2.2. Study selection

We included all randomized control trials lasting at least four weeks if they comprised economic incentives tied to one of the following: (i) increased physical activity; (ii) weight loss; (iii) smoking cessation; (iv) positive changes in dietary behaviour (in terms of increased fruit and vegetable consumption, reduced consumption/purchase of sugar sweetened beverages, and/or increased consumption/purchase of low-calorie food items) and (iv) reduced alcohol consumption/purchases. Studies were included if the primary outcome was the effectiveness of the incentive-based intervention (in terms of the incentivised behaviour/goal) offered relative to the control group. For studies in which a primary outcome was not reported (n = 11), we first considered the measure of effectiveness that is most commonly reported among studies in the same health domain, then the measure directly tied to the incentive, or, if necessary, the first measure reported when incentives were tied to multiple outcomes. Within included studies, we only included comparisons between arms (referred to as contrasts henceforth) in which the independent effect of incentives could be inferred relative to another intervention arm.

Studies were excluded if the primary outcome was based on self-report as those in the intervention arm have an incentive to misrepresent their behaviour, if the effect of incentives was confounded with other aspects of the intervention, and/or where incentives were tied to group behaviour. We further excluded studies if the incentives were offered to healthcare providers to change patient behaviour, targeted children, or focussed exclusively on those with mental-health and/or substance abuse conditions.

2.3. Data extraction

The initial search identified 58,367 records, and 42,342 were subject to title/abstract screening after duplicates were removed. Screening was conducted by two independent reviewers (YG and DB). 83 articles were identified as eligible for full-text review (Fig. 1 – PRISMA flowchart). Full text screening was performed by two independent reviewers and conflicts were resolved by consensus. Reference lists of included articles were also checked for any articles that were potentially missed from the search, but no unique articles were identified. In total, 48 articles met the inclusion criteria and were included in the study.

Data was extracted using a pre-designed and pre-tested data extraction template to extract the following information from each study: (1) publication details – author, date of publication; (2) study design – inclusion and exclusion criteria, sample size, description of intervention received by treatment and control group; (3) intervention characteristics – type of incentive, incentive size, frequency of incentive payment, duration of intervention, evidence for habit formation, cost and cost-effectiveness of intervention; and (4) effectiveness of intervention – primary outcome of the study. Incentive size was calculated as the maximum incentive the participant could expect to receive if she met all opportunities to earn incentives over the duration of the intervention. For lotteries, maximum pay-out was calculated as the expected value. For non-cash rewards the value of the reward in monetary terms was considered. All monetary figures were converted to Purchasing Power Parity dollars (World Bank, 2017) for 2017 using country specific inflation rates (International Monetary Fund, 2017). Duration of the intervention was defined as the time period between the start of the intervention and the final opportunity to earn an incentive. The data extracted from all studies included in the review are summarised in Appendix Table 2.

2.4. Critical appraisal

Risk of bias assessment was conducted for all articles that met the inclusion criteria using the Cochrane Risk of Bias Assessment tool. This tool, specifically designed for clinical trials, evaluates the methodological quality of studies using the following six criteria: selection bias, performance bias, detection bias, attrition bias, reporting bias, and other biases. Risk of bias assessment was performed by one reviewer (YG) and cross-checked by another (DB). All review team members (EAF and MB) were consulted in case of ambiguities/conflicts. Differences were resolved by discussion. Regardless of risk of bias, all studies reviewed are included in the narrative synthesis, with implications for low quality studies addressed in the Discussion section.

2.5. Outcome measures and data analysis

For all contrasts considered, we tabulated whether the incentive-based intervention was effective or not based on whether the study reported statistically significant differences between the incentive and no-incentive arm on the primary outcome and the direction of the effect, if significant. For studies with multiple incentive arms we considered all contrasts that met our inclusion criteria.

3. Results

A total of 48 full-text articles were reviewed. Among these articles, 70 contrasts allowed for identifying the independent effect of incentives relative to a no-incentive arm. Studies/contrasts reviewed are summarised in Table 1. By health domain: 3 contrasts tied incentives to healthy food purchases; 18 to weight loss; 11 to physical activity; 12 to fitness session attendance; 24 to smoking cessation, and none to alcohol. Two contrasts tied incentives to both weight loss and physical activity goals (Finkelstein et al., 2017; Shin et al., 2017). In 6 studies, incentives were tied to other aspects of a multicomponent intervention (e.g., participation in smoking cessation programs, use of a weight loss website, etc.) in addition to the target behaviour. This review features a small number of studies focussing on diet and a lack of alcohol studies due to exclusion of studies that relied on self-report and studies that focused solely on populations with mental health and/or substance abuse conditions due to concerns about generalizability.

Below we present characteristics of the included studies, including variation in key intervention design aspects likely to influence intervention effectiveness. We then present overall and domain-specific effectiveness results and evidence for how effectiveness varies by

Table 1
Distribution of incentive-based interventions by incentive type and health domain.[a].

	DIET	FITNESS SESSION ATTENDANCE	PHYSICAL ACTIVITY	WEIGHT LOSS	SMOKING CESSATION	Contrasts showing statistically significant effects
CASH	Thorndike et al. (2016)	Acland and Levy (2015) Charness and Gneezy (2009) Jeffrey et al. (1998) C1 Jeffrey et al. (1998) C2 Pope et al. (2013)	Finkelstein et al. (2008) Finkelstein et al. (2016) Harkins et al. (2017) C1 Harkins et al. (2017) C2 Patel et al. (2016a) C1	Finkelstein et al. (2007) C1 Finkelstein et al. (2007) C2 Kullgren et al. (2013) Paloyo et al. (2015) C1 Paloyo et al. (2015) C2 Shin et al. (2017)	Baker et al. (2018) Halpern et al. (2015) C1 Jason et al. (1997) Lamb et al. (2007) Volpp et al. (2006) Volpp et al. (2009)	18 of 23
DEPOSIT CONTRACTS	Nil	Nil	Patel et al. (2016a) C3	John et al. (2011) C1[b] John et al. (2011) C2[b] Kullgren et al. (2016) C1 Kullgren et al. (2016) C2 Kullgren et al. (2016) C3 Volpp et al. (2008) C1	Gine et al. (2010)	6 of 8
LOTTERY	Nil	Nil	Kullgren et al. (2014) C1 Kullgren et al. (2014) C2 Patel et al. (2016a) C2 Patel et al. (2016b)	Patel et al. (2016) C3 Volpp et al. (2008) C2	Crowley et al. (1995) C1 Crowley et al. (1995) C2	1 of 8
NON-CASH REWARDS	French et al. (2017) C1 French et al. (2017) C2	Carrera et al. (2018) C1 Carrera et al. (2018) C2 Carrera et al. (2018) C3 Courneya et al. (1997)	Nil	Patel et al. (2016) C1 Patel et al. (2016) C2	Donatelle et al. (2000) Etter and Schmid (2016) Fraser et al. (2017) Heil et al. (2008) Higgins et al. (2014) C1 Higgins et al. (2014) C2 Kendzor et al. (2015) Ondersma et al. (2012) C1 Ondersma et al. (2012) C2[c] Secades-Villa et al. (2014) Tappin et al. (2015)	12 of 19
COMBINATION OF INCENTIVE TYPES	Nil	Royer et al. (2015) C1 Royer et al. (2015) C2 Wing et al. (1996)	Petry et al. (2013)	Faghri et al. (2014) Finkelstein et al. (2017) Leahey et al. (2015) Petry et al. (2011)	Halpern et al. (2015) C2 Hennrikus et al. (2002) Ledgerwood et al. (2014) C1 Ledgerwood et al. (2014) C2	10 of 12
Contrasts showing statistically significant effects	1 of 3	10 Of 12	5 of 11	13 of 20	19 of 24	

[a] Contrasts reporting statistically significant intervention effects are represented in bold.
[b] Both incentive groups *combined* had a significantly greater effect relative to control and there was no significant difference between the two incentive groups. Hence, we assume that both contrasts had a significant effect.
[c] Ondersma et al. (2012) C2 was the only contrast that reported a statistically significant negative effect. Hence it was not included in the computation of the percentage of contrasts reporting statistically significant effects in the corresponding row and column.

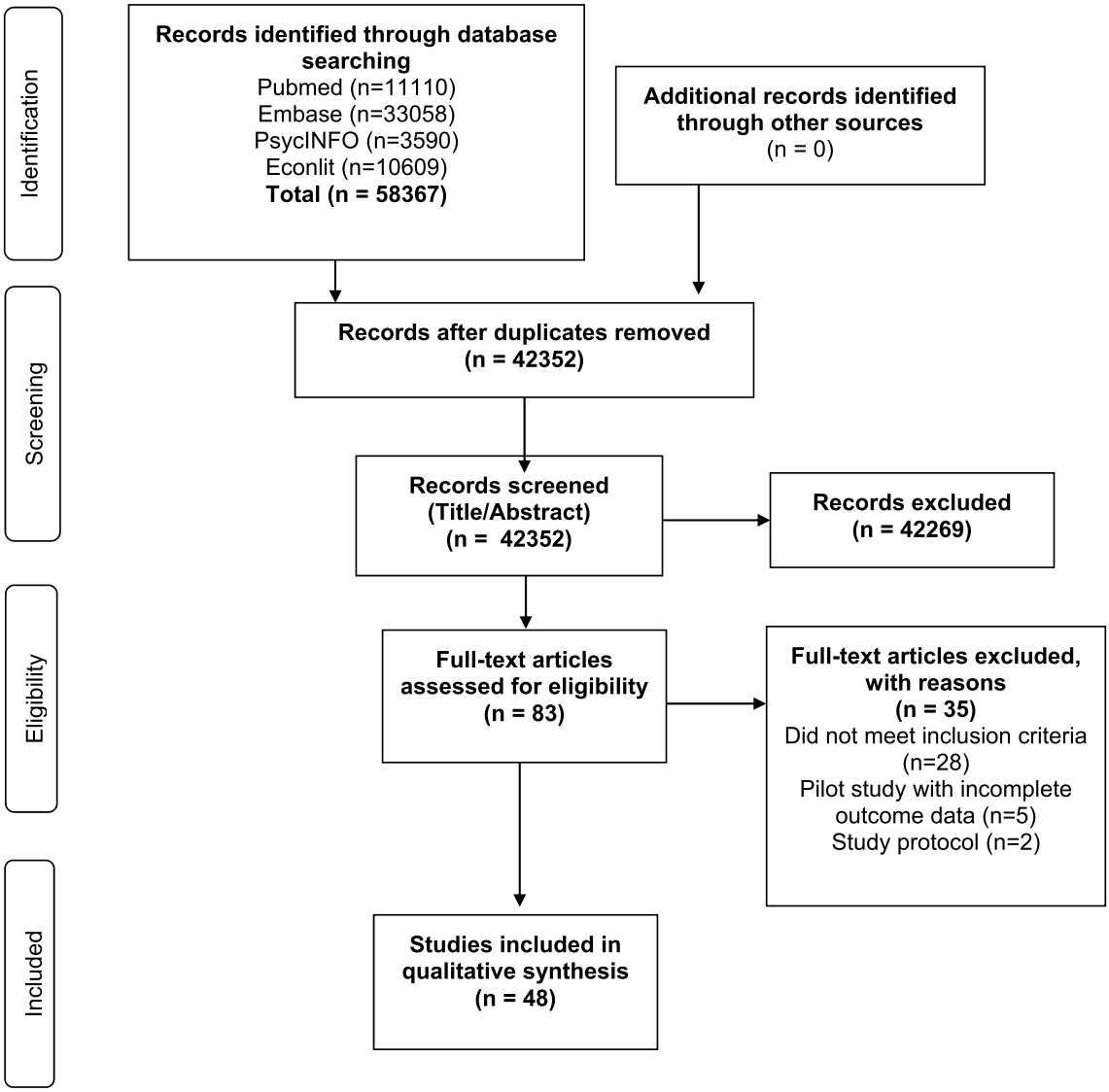

Fig. 1. Prisma flow diagram.

population, comparator group, value, type, frequency, and duration of incentives, followed by the evidence for habit formation and cost-effectiveness. Lastly, we present a risk of bias assessment for the primary outcome of all included studies.

3.1. Characteristics of included studies

3.1.1. Population and setting

The vast majority of studies focused on adult populations in high income countries (68 contrasts), and primarily the US (60 contrasts). Twenty-five of 70 contrasts were evaluated in a community setting (36%), 27 in a workplace environment (39%), 14 focused on outpatients of a medical centre (20%) and 4 in a fitness centre (6%). Within each health domain, the target population differed. All weight loss interventions and 3 interventions incentivising fitness centre attendance or physical activity targeted an overweight and/or obese population. A third of contrasts (8 of 24, (33%)) that tied incentives to smoking cessation specifically targeted pregnant women, a highly motivated subset of smokers. Across all domains, only 6 contrasts specified that their intervention exclusively targeted low-income populations either as part of their inclusion criteria or description of the sample.

3.1.2. Comparator group

For the majority of contrasts (41 contrasts (58%)), the incentive was given as part of a multi-component intervention. In these contrasts, the incremental effect of the incentive was measured relative to a control group that received an intervention without the incentive. For instance, over half of weight loss studies combined incentives with other evidence-based weight loss strategies, including face to face counselling (8 of 20 contrasts, (40%)) or goal-setting (5 of 20 contrasts, (25%)) (Finkelstein et al., 2017; John et al., 2011; Jeffrey T Kullgren et al., 2016; Paloyo et al., 2015; N. M. Petry et al., 2011; Volpp et al., 2008). Two-thirds (16 out of 24 contrasts) of smoking cessation interventions combined incentives with cessation programs including counselling, personalised feedback or cognitive behavioural therapy (11 contrasts) and/or other cessation resources such as nicotine gum, computerized cessation interventions, or social support for cessation (5 contrasts). Two of the 3 diet contrasts combined incentives with a food-assistance program, and the third combined incentives with a social-norm based feedback intervention.

Among the remaining 29 contrasts (41%), 10 contrasts focused solely on incentives compared to a control group that received no intervention and 19 had minimal interventions that were unlikely to more than minimally affect the target behaviour. Minimal interventions included provision of information about fitness programs/smoking

cessation programs (6 contrasts), tips on how to quit smoking or exercise more and/or general fitness/weight loss goals (2 contrasts), pedometers with or without daily/weekly feedback on whether step-goals were met (5 contrasts), access to weighing scales with encouragement to weigh themselves as often as they like (3 contrasts), abstinence monitoring for smokers (1 contrast) and free access to a behavioural smoking cessation program (1 contrast).

3.1.3. Outcome measure

Even within health domains, the primary measure of effectiveness differed, especially among the physical activity domain. In this domain, interventions reported proportion of fitness sessions attended (4 contrasts), proportion of days a step-goal was met (5 contrasts), proportion of participant-days a step-goal was met (4 contrasts), number of visits or average weekly gym visits attended (7 contrasts), whether a target number of fitness centre visits was reached (1 contrasts) or active minutes (2 contrasts). Majority of smoking cessation interventions used a biochemically verified 7 day point-prevalence abstinence rate (verified by either saliva/urine cotinine/breath CO tests) as the primary endpoint (13 contrasts). The primary outcome was sustained abstinence in two studies, wherein participants had to be biochemically confirmed abstinent at more than one time point to be considered abstinent (3 contrasts). Three studies used a biochemical verification of cessation (urine cotinine/breath CO) without a point prevalence abstinence measure. Other measures used were 24-h abstention rates (2 contrasts), percent of negative CO samples (3 contrasts). Among the diet interventions, mean expenditure on fruit in dollars (2 contrasts) or proportion of healthy food purchases (1 contrast) was used. Outcomes were most consistent in weight loss studies, where mean weight loss was used in all but two contrasts that instead focused on the proportion of participants achieving > 5% weight loss (N. M. Petry et al., 2011) or the percentage weight loss relative to baseline (Leahey et al., 2015).

3.1.4. Duration

Length of interventions varied from 4 to 78 weeks. Median duration across all studies was 16 weeks and was largest in the smoking domain (median: 24 weeks). Only a third of contrasts (29 of 70, (41%) lasted 24 weeks or longer, the duration that has been reported necessary for a habit to accrue. Ten of these incentivised weight loss, 1 incentivised physical activity, 1 incentivised both weight loss and physical activity, 3 incentivised fitness session attendance and 14 incentivised smoking cessation. Among contrasts that incentivised smoking cessation, almost half (6 of 14, (43%)) had an intervention period of 24 weeks or greater as they incentivised smoking cessation among pregnant smokers until/post-birth. Few contrasts (7 of 70, (10%)) lasted 52 weeks or greater. Among these, 2 contrasts incentivised fitness session attendance, 2 incentivised smoking cessation and 3 contrasts incentivised weight loss.

3.1.5. Incentive type

The incentives offered can be broadly classified into five types: cash (23 contrasts, (33%)), lotteries (8 contrasts (11%)), deposit contracts (8 contrasts (11%)), non-cash incentives (19 contrasts (27%)) and a combination (12 contrasts (17%)). Two studies (2 contrasts) gave participants an option between a cash incentive and a cash plus deposit contract incentive or a cash incentive and a cash plus lottery incentive. Four studies (5 contrasts) used a lottery tied to non-cash prizes. Four studies (5 contrasts) combined cash incentives along with additional incentives (lottery, non-cash, or a deposit contract incentive).

In most interventions offering deposit contract incentives as the only incentive, the voluntary deposit was combined with a matching incentive (5 of 8 (63%)). This was done in efforts to increase the amount of money participants put at risk to increase intervention effectiveness, and was most common in the weight loss domain. This enabled participants to earn a 100% return (3 contrasts), a 100% return and a fixed return per day (1 contrast) or a 200% return (1 contrast) on the amount deposited by the participant.

Various types of non-cash incentives were offered either as an independent incentive strategy or as part of a bundle of incentives across studies. Thirteen contrasts used vouchers for retail items or gift cards, 3 contrasts used fitness centre memberships, 2 contrasts used rebates that could be used for groceries and 2 contrasts used adjustments in health insurance premiums and 4 contrasts used a range of “small” (eg. Toiletries, water bottles, food, etc) and “large” (fitness equipment, DVD players, etc.) non-cash incentives in a lottery. In one study (1 contrast), participants selected an item from an online retail store at the start of the intervention that they would later receive as an incentive if they met their fitness goals. In one study, a voucher incentive was used but the rewards that could be claimed for the voucher was not reported. Among interventions including non-cash incentives, 17 of 26 contrasts (65%) were more likely to be perceived as hedonic (gift cards, retail items, fitness centre memberships etc.) as compared to the 9 other contrasts which were either tied to utilitarian items (6 contrasts) (such as grocery vouchers, adjustments in health insurance premiums, etc.) or were cash-like in that they could be exchanged for a large range of rewards (3 contrasts) that may or may not be perceived as hedonic.

3.1.6. Frequency of opportunities to earn incentives

Most contrasts (56 of 70, (80%)) offered multiple opportunities to earn incentives. Many offered daily incentives (17 contrasts), on alternate days (1 contrast), thrice weekly (4 contrasts), weekly (11 contrasts), while others offered monthly incentives (3 contrasts). Frequent chances to earn incentives were most common among physical activity interventions (11 of 11 contrasts) where incentives were tied to daily/weekly step-goals. Among other domains, frequent chances to earn incentives were based on interim weight loss goals on a daily/weekly basis (13 contrasts), attendance at daily/weekly fitness sessions (4 contrasts), meeting weekly physical activity minute goals (1 contrast), weekly/monthly food purchases (3 contrasts), participating in smoking cessation program sessions (4 contrasts) or verified abstinence at multiple points during the intervention (15 contrasts). The longest period in which an incentive could be earned was 6 months, for smoking abstinence.

Among 20 contrasts, variable incentive schedules were used instead of offering incentives at a regular frequency. The vast majority of these were smoking cessation interventions (16 of 20 contrasts (80%)). Almost two-thirds (13 contrasts) of these offered more frequent chances to earn incentives in the initial stages of the intervention to encourage initial behaviour change. Among the remaining 7 contrasts, a variable payment resulted from the fact that incentives were tied to participation/completion of other intervention components in addition to a targeted risk-factor reduction (4 contrasts), depended on participants to self-report meeting a target and claim their reward (2 contrasts). In one contrast incentives were only offered toward the end of the intervention.

3.1.7. Incentive size

Maximum potential incentive pay-outs ranged from $29–1736 across all studies, with a median value of $360. The largest potential pay-outs were in the weight loss and smoking domains, with maximum potential pay-outs ranging from $47 to 1544 (median: $556) and $155–1736 (median: $402) respectively. Maximum potential pay-outs for the other domains are as follows: diet (range: $31–136, median: $136), fitness session attendance (range: $29–738, median: $108) and physical activity (range: $114–853; median: $320). These pay-outs are highly influenced by the intervention duration. The median maximum potential pay-out was $293 for interventions that lasted 24 weeks or less, and $738 for longer interventions.

3.2. Evidence for the effectiveness of interventions

In total, over two-thirds of contrasts (48 of 70, (69%)) reported statistically significant evidence that the use of incentives improved the

targeted risk factors. Thirteen contrasts (13 of 20, (65%)) were effective for weight loss, 5 contrasts (5 of 11, (45%)) for physical activity, 10 (10 of 12, (83%)) contrasts for fitness session attendance, 1 contrast (1 of 3, (33%)) for healthy food purchases, and 19 contrasts (19 of 24, (79%)) for smoking cessation. Considering the subset of studies (8 contrasts) in which incentives were tied to multiple targets, either tied to multiple risk-factors or tied to program participation/completion in addition to risk-factor reductions, all but one were effective (7 of 8 contrasts (88%)).

3.2.1. Effectiveness according to comparators

Among interventions in which incentives were used as part of a multi-component strategy, 29 out of 41 (71%) were effective. Among interventions where incentives were used as a standalone strategy or coupled with minimal interventions 17 out of 29 (58%) were effective.

Three multi-arm RCT's directly tested the relative effectiveness of incentive strategies when used independently or when combined with other interventions. One study (French et al., 2017) found that while giving incentives along with a food assistance program was not effective relative to a control group that received only food assistance, incentives were found to increase weekly fruit purchases when the intervention combined incentives with dietary restrictions on unhealthy food (sweet baked goods, sugar sweetened beverages, etc.) and food assistance. Another study found that while an incremental standalone incentive was ineffective relative to a control group that received usual care, the incremental effect of incentives when added to a computer-delivered smoking cessation program had an adverse effect on smoking cessation outcomes (Ondersma et al., 2012). J. T. Kullgren et al. (2014) tested the relative effectiveness of using incentives alone or combined with a peer-network based intervention, but neither strategy was found to be effective.

The limited number of studies and mixed results offers little support for the hypothesis that the incremental effectiveness of incentives will be smaller when used as part of a multi-component intervention.

3.2.2. Effectiveness according to population and setting

Five of 6 contrasts (83%) that targeted low-income groups were found to be effective. This proportion of effective interventions was slightly lower among interventions targeting the general population in community settings (14 of 18 contrasts, (78%)) and even lower among workplace interventions that generally targeted participants of higher socio-economic status (15 of 27 contrasts, (55%)).

Although there are no RCTs with direct tests of this hypothesis, the larger proportion of effective interventions targeting low-income groups compared to interventions in workplace or community settings is consistent with the hypotheses that incentives may be more likely to be effective among this population subset.

3.2.3. Effectiveness according to incentive type

Across all studies, the majority (18 of 23 contrasts, (78%)) of cash based interventions were effective. With one exception, this was also true among other incentive types: deposit contracts (6 of 8, (75%)), non-cash incentives (12 of 19, (63%)), and combined incentives (potentially including lotteries) (10 of 12, (83%)). The one exception is cash-based lotteries in the absence of another form of incentives. For these, only 1 contrast (1 of 8, (13%) was effective.

Ten contrasts offered non-cash incentives that were more likely to be perceived as hedonic, and the vast majority were effective (80%). All interventions that combined incentives with hedonic rewards were effective (7 contrasts). Among the remaining 9 contrasts that comprised non-cash incentives that were less likely to be perceived as hedonic, 4 (44%) were found to be effective.

The review identified three studies with head-to-head comparisons of different incentive types. One study (Patel et al., 2016b) offered a direct test of the relative effectiveness of a lottery incentive versus a cash incentive of equivalent value. Neither incentive strategy resulted in a statistically significant effect. No studies offered a direct test of whether lotteries comprising a low-probability large reward are more effective than a high-probability small reward, as prospect theory would predict. Two studies examined the effectiveness of deposit contracts versus other gain-based incentives in a multi-arm RCT (Halpern et al., 2015; Patel et al., 2016b), but only one finds results consistent with theory (Patel et al., 2016b). Limited uptake of deposit contracts could potentially explain why deposit contracts did not outperform equivalent cash-based incentives in one study (Halpern et al., 2015), but did so in the other when deposit contracts were not self-funded but comprised upfront allocation of funds from the researchers (Patel et al., 2016b).

No studies directly tested the effectiveness of hedonic rewards compared to cash. However, one study (Carrera et al., 2018) compared two non-cash incentives of equivalent value of which one incentive was cash-like (an Amazon gift-card) while the other could have been perceived as hedonic (pre-selected item chosen from the Amazon website by the participant before the intervention began). While the item-based incentive lead to slightly more fitness centre visits than the equivalent cash-like incentive, the difference was small and not statistically significant. Thus, we have no direct evidence that a hedonic reward would be more effective than cash.

Although three other studies involved more than one incentive type, we could not infer the relative effectiveness of one over the other as: (i) maximum incentive value was not equal across the two incentive arms (lottery and deposit contracts in Volpp et al. (2008)) or (ii) individuals had a choice of which arm to join to the independent effect could not be independently identified (cash and deposit contracts in Faghri and Li (2014); cash and lottery in Finkelstein et al. (2017)).

Overall, the current evidence base examining the relative effectiveness of different incentive types is limited. Existing literature does not inform on whether hedonic rewards are more effective than cash equivalents and provides both limited and mixed evidence on whether deposit contracts work better than cash equivalents, as theory would predict. However, the evidence base does suggest that pure cash-based lotteries are unlikely to be effective, and that among non-cash incentives, hedonic rewards appear to be more promising than those less likely to be perceived as hedonic.

3.2.4. Effectiveness according to frequency of opportunity to earn incentives

Among interventions that provided participants multiple opportunities to earn incentives, more than two-thirds (40 of 56 contrasts, (71%)) were found to be effective. In particular, the majority of interventions that offered multiple chances to earn incentives in a week (16 of 22 contrasts, (73%)) were effective. The proportion of interventions that were found to be effective was lower among contrasts which offered less frequent chances to earn incentives: 7 of 11 contrasts (64%) that offered chances to earn incentives on a weekly basis, 1 out of 3 (33%) contrasts that do so on a monthly basis, and 6 of 14 (43%) contrasts less frequently than on a monthly basis or only tied an incentive to one goal achievable at the end of the intervention.

Only one weight loss study (Patel et al., 2016a) compared the effectiveness of two equivalent incentive designs that varied in frequency of incentive pay-outs: a daily lottery incentive, and a one-time incentive pay-out in the form of health insurance premium adjustments that began immediately once a weight loss goal was reached. Neither intervention was found to be effective.

Although no RCTs are available that directly support this claim, the review suggests that, consistent with theory, interventions with more frequent opportunities to earn incentives are more likely to be effective.

3.2.5. Effectiveness according to incentive value

Among the quartile of contrasts which offered the smallest incentives in terms of maximum incentive pay-out per month, 8 out of 18 (44%) contrasts were found to be effective. In the quartile with the largest pay-outs 13 out of 18 (72%) were found to be effective.

Four studies directly tested if larger incentives work better than smaller ones. Among these, one weight loss study and one fitness centre attendance study found this hypothesis to be true (Carrera et al., 2018; Finkelstein et al., 2007), while one other relying on deposit contracts for weight loss (Jeffrey T Kullgren et al., 2016) did not. In one study using cash rewards for weight loss (Paloyo et al., 2015), larger incentives were found to have a greater effect for women and migrants, but not for the treatment group as a whole.

Our review suggests that larger incentives are more likely to be effective than smaller ones. Although limited, we also find some direct evidence from head-to-head trials to support the hypothesis that increasing the value of incentives can increase the magnitude of the intervention effect.

3.2.6. Effectiveness according to intervention duration

Contrary to expectations, the proportion of effective interventions was lower among interventions lasting less than 24 weeks (28 of 42 (67%)) as compared with interventions with a duration of 24 up to 52 weeks (18 of 21 (86%)). However, the proportion of effective interventions was much lower among the few contrasts that lasted 52 weeks or longer (3 out of 7 (43%)).

Whereas these results suggest that incentives are likely to be effective at reducing NCD risk-factors in between 1 and 12 months, evidence for greater than one year is not compelling.

3.3. Evidence for habit formation

Over half of the contrasts (40 out of 70, (57%)) tested for sustained intervention effects at a time-point after the last incentive could be earned. Reporting evidence for habit formation was most common among physical activity interventions (10 of 11 contrasts (91%)) relative to interventions in the domains of fitness session attendance (7 of 12 contrasts (58%)), smoking (13 of 24 contrasts (54%)) and weight loss (10 of 20 contrasts (50%)). No diet studies tested for sustained effects. The longest follow-up assessment post-incentive removal varied significantly: ranging from 4 to 100 weeks (median: 13 weeks).

Less than a third of contrasts (11 of 40 (28%)) found evidence for habit formation at the longest follow-up time point. A slightly greater proportion of interventions lasting 24 weeks or longer found evidence for sustained effects (5 of 14 contrasts, (36%)) as compared with shorter ones (6 of 26 contrasts, (23%)). For interventions incentivising smoking cessation, weight loss and fitness centre attendance, less than half of contrasts found evidence for habit formation (smoking cessation: 5 of 13 (38%), weight loss: 3 of 10 (30%), fitness centre attendance (2 of 7 (29%)). The evidence for sustained effects of physical activity interventions is especially weak; 1 of 10 (10%) contrasts found evidence for habit formation). Evidence for habit formation is summarised in Appendix Table 3.

The current evidence base offers no direct evidence for testing the hypothesis that longer interventions (lasting 24 weeks or greater) are more likely to result in habit formation than shorter ones. Overall, the evidence for sustained effects post-incentive removal is weak for all health domains, and especially so for physical activity.

3.4. Cost and cost-effectiveness of interventions

The average incentive pay-outs per capita were reported or could be estimated from given information for half of the contrasts (36 of 70, (51%)). This was most common in the domain of smoking (14 of 24 contrasts, (58%)), followed by weight loss (11 of 20 contrasts, (55%)). None of the diet studies reported intervention costs. Per capita monthly incentive pay-outs ranged from - $4 to $245 (median: $44). For two contrasts, the pay-outs were negative ($2 and $4) due to the use of deposit contracts which resulted in a net loss of participants' money.

Only one study reported a cost-effectiveness finding that could be used to assess whether the intervention represents good value for money. Tappin et al. (2015) reported a cost-effectiveness ratio of $698 per QALY gained for smokers in the intervention group who received shopping vouchers in addition to usual care relative to those who received only usual care, and notes that this is cost-effective relative to established thresholds.

Average incentive pay-outs and efforts to measure cost-effectiveness are reported in Appendix Table 4. Note that we only report incentive pay-outs and do not consider labour and other costs associated with administering the interventions, as most studies did not provide those values.

3.5. Risk of bias assessment

The heterogeneity in all of the above contrasts clearly demonstrates the potential for bias when trying to meta-analyse the results. However, the potential for bias remains even within studies. Although our inclusion criteria that limited the review to randomized control trials minimizes risks for selection bias, many studies reported insufficient information to assess the integrity of the random sequence generation (n = 28) and/or concealment of treatment allocation prior to baseline measurements (n = 34). Despite these concerns, only two studies were considered to be at greater than low risk for selection bias. One study described a non-random component in the sequence generation process (Giné et al., 2010) and in the other study allocation to treatment arms was not concealed prior to enrolment (Carrera et al., 2018).

Additional bias may result if evaluators are aware of the allocation prior to assessments. Blinding of assessors was reported in only 12 studies and three studies reported that outcome assessors were not blinded. Regardless, the likelihood of evaluator bias was judged to be minimal in all studies.

Non-random missing outcome data is another potential source of bias. When data was missing, reasons for missing outcomes and method of handling missing data were clearly stated in only 21 studies. Attrition bias is unclear in the remaining studies as they provided insufficient information about the reasons for missing outcome data, reasons for exclusion from analyses or a comparison of the rate of missing responses by intervention arm.

All but one study (n = 47) pre-specified the primary outcome and the method of analysis, and were judged to have a low risk of bias due to selective outcome reporting. This was unclear for one study (Donatelle et al., 2000) as the methods of statistical analysis was not pre-specified.

Our inclusion criteria also required an objectively measured primary endpoint. This minimizes self-report bias by participants but does not rule out other sources of bias. Six studies that use step-counters/trackers to measure their primary outcome were considered to have high risk of bias because one cannot distinguish between a true intervention effect and a 'wear-time' effect that arises due to the fact that incentivised participants may simply be wearing their pedometers longer to earn incentives. Despite being objectively measured, all of the pedometer based studies are considered to have a risk of bias.

Finally, ten studies used a primary outcome that was a proxy measure of effectiveness: food purchases or fitness session attendance. As individuals can meet these objectives without improving their health, say by attending fitness sessions but not necessarily increasing their activity levels, these studies provide weak evidence in support of incentives for improving diet quality or fitness or truly reducing risk factors for NCDs. A full description of the risk of bias for each study can be found in Appendix Table 5.

4. Discussion

The aim of this review was to synthesize the evidence of effectiveness and cost-effectiveness of incentive-based strategies as a tool for NCD prevention. A systematic review of this nature is important because prior reviews fall short of answering our primary question, which

is whether there is truly an evidence base to support the use of incentives, large or small, traditional or behavioural, alone or as part of a multi-component strategy, as a tool for addressing rising rates of NCDs.

4.1. Is there a role for incentives as a tool for NCD prevention?

Our primary finding is that although the overwhelming majority of contrasts reviewed (69%) find a statistically significant intervention effect, the existing evidence base does not provide a strong case for incentives as a tool for NCD prevention in any of the included domains. This results because, upon consideration of potential biases and limitations identified in the risk of bias assessment, we find that 15 of the 70 (21%) contrasts, including all in the food consumption domain, focused on a proxy measure of effectiveness: food purchases or fitness session attendance. As individuals can meet these objectives without improving their health, these studies provide only weak evidence in support of incentives for NCD prevention. Ten of the remaining 11 studies in the physical activity domain are likely biased because they are based on pedometers and do not differentiate an increase in wear time from an increase in activity. Therefore, these studies also provide weak evidence. Once these 25 studies are removed, there remains only one study in the physical activity domain (using sealed accelerometers) and no studies in the domain of diet and alcohol consumption. Therefore, the evidence for effectiveness, let alone cost-effectiveness, is clearly lacking in these domains.

For weight loss, 13 of 20 contrasts (65%) reported a statistically significant intervention effect, whereas for smoking cessation 19 of 24 (79%) contrasts did so. For these domains, the evidence of effectiveness is more compelling, but the question remains as to whether financing the interventions is a good use of scarce resources. In other words, is there any evidence of cost-effectiveness? For 7 of 13 contrasts that report statistically significant effects in the weight loss domain, average incentive pay-outs and the difference in mean weight loss (kg) was reported. This allowed for the computation of the average cost paid per kg of weight lost. This figure ranged from a negative $3/kg lost (i.e., cost saving) to $73/kg, with a median of $47/kg lost. For comparison, average cost per kg lost for WW (formerly Weight Watchers) is $134 (Finkelstein and Kruger, 2014) and the average cost per kg lost for the weight loss drug Qsymia is $327 (Finkelstein et al., 2015) and both of these have been shown to be cost-effective based on established benchmarks for cost-effectiveness. However, in those cases weight loss was based on 12 month results and on larger weight losses; 3.2 and 6.7 kg compared to a median of 2.9 kg over 12–35 weeks for the incentive-based studies. It is unclear whether the weight loss results in the incentive studies are large enough and/or would be sustained long enough to be clinically relevant. On the latter point, consistent with prior reviews (Giles et al., 2014; Mantzari et al., 2015), evidence for sustained effectiveness post-incentive removal is weak. Only 3 of 10 wt loss contrasts (33%) which report evidence for habit formation find statistically significant evidence of sustained effects. Therefore, we believe current evidence for the use of incentives in the weight loss domain is also not compelling as a means to reduce NCDs. This leaves smoking cessation as the remaining target where incentives may present a compelling case.

The existing literature on cost-effectiveness of behavioural interventions for smoking cessation, such as cognitive behaviour therapy or telephone counselling, report cost per quit estimates ranging from $1846/quit to $3397/quit (Ali et al., 2018). For 10 of 19 contrasts in the smoking cessation domain that allow for generated comparable estimates, incremental costs per quitter ranged from $274/quit - $20,503/quit (median: $2212/quit), suggesting that some of the strategies may be cost-effective. However, less than half of the smoking cessation interventions which test for sustained effects post-incentive removal find evidence for habit formation (5 of 13 (38%)) which casts doubt on the cost-effectiveness claim. Therefore, although one study reports their intervention to be cost-effective (Tappin et al., 2015) on the whole the evidence base remains weak even for this domain.

4.2. Behavioural economics: hype or help?

In addition to assessing the evidence base for incentives as a tool for NCD prevention, the review aimed to identify whether one type of incentive strategy may work better than others or that incentives work better for any population subset. The overall heterogeneity in various aspects of intervention design across studies and the very few studies (n = 8) that test specific behavioural economic hypotheses does not allow for making strong statements about any of the predictions based on size, type, or frequency of opportunities to earn incentives. However, the review does offer some suggestions for those looking to design incentive-based studies.

The review suggests that multi-pronged incentive strategies may be most likely to be effective. This includes strategies that target both participation and outcomes or those that tie incentives to both shorter and longer term behaviours and outcomes. Effectiveness may be more likely by offering more frequent opportunities to earn incentives and by targeting lower income populations. We also found some evidence that non-cash rewards which are likely to be perceived as hedonic represent a promising strategy and, consistent with prior reviews, that larger incentives appear more likely to generate statistically significant improvements, at least in the weight loss and fitness domains. With one exception, Ondersma et al. (2012), we also found no evidence of an undermining effect of incentives, as prior research has suggested (Gneezy et al., 2011). Yet, although interventions using these incentive strategies may be most likely to be effective, whether they are cost-effective for NCD prevention remains unknown.

The review further suggest strategies that appear least compelling. This includes pure-cash based lotteries, possibly because many participants are risk averse and dislike lotteries (Finkelstein et al., 2017), and deposit contracts. Concerning the latter, whereas interventions involving deposit-contracts are likely to be effective, potential for scale-up is limited due to low uptake, even when third party matching is offered.

Beyond these conjectures, whereas behavioural economic theory suggests that intervention design aspects have the potential to influence intervention effectiveness, results of the review reveal that evidence to support this claim is both limited and mixed in the context of incentives for NCD risk-factor reduction.

4.3. Future research and policy implications

Contrary to the implications from many prior studies, literature reviews, and meta-analyses on the use of traditional and behavioural economic incentives as a tool to influence risk factors for NCDs, based on our review we conclude that the evidence base is not compelling at this time (Giles et al., 2014; Mantzari et al., 2015; Sutherland et al., 2008). If one wants to make a stronger case for incentives in NCD prevention, we would recommend conducting studies that include a clinical endpoint, are at least 12 months in duration, and that monitor habit formation beyond the incentive period. Although this has not always been the case, all future studies focussing on incentives should conform to best practices for conducting and reporting results of randomized controlled trials, as outlined in the CONSORT statement (Schulz et al., 2010). Researchers should also systematically report on the cost-effectiveness of incentives as they would for other health interventions using best practices (Husereau et al., 2013) as a lack of evidence for cost-effectiveness will make even effective interventions difficult to justify to policymakers.

In terms of testing behavioural economic theories, future trials should also adhere to these recommendations and be powered such that differences in effect sizes are clinically meaningful. At this time the evidence from randomized trials is weak that any behavioural strategy is cost-effective when compared to a simple cash payment for meeting a

pre-determined goal, such as weight loss or quitting smoking. To make a case for these strategies, researchers should design trials that allow for directly testing this hypothesis.

4.4. Strengths and limitations of the review

This is the first systematic review to our knowledge that comprehensively examines intervention design aspects such as incentive type, frequency of opportunities to earn incentives, and cost-effectiveness of incentive-based interventions targeting four key risk-factors for NCDs: diet, alcohol consumption, physical activity and weight loss. Consideration of the four key modifiable behaviours linked to NCDs (diet, alcohol consumption, physical inactivity, and smoking) and a commonly incentivised risk-factor (weight loss) enables us to consider a broad evidence base that is necessary for understanding the role of incentives as a tool for NCD prevention. Furthermore, this review identifies and critically examines the direct evidence for several hypotheses motivated by behavioural economic theory which are often used to design incentive-based interventions, but have not been systematically assessed for effectiveness in the context of NCD prevention. These findings are relevant for policymakers considering the incorporation of incentives in large-scale NCD prevention programs.

Limitations of this review include the omission of non-English language articles, evidence published before 1995 and evidence from unpublished or grey literature. This may be a potential source of publication bias. Another limitation was that we did not consider differences in methods of estimating intervention effects across studies or heterogeneity in how the dependent variable was measured both across and within health domains. It is possible that studies that did not show effectiveness were underpowered due to small sample sizes, measurement error, or conservative methods of estimation. Finally, the review outcomes need to be carefully interpreted for generalizability as we did not identify any studies in the domain of alcohol consumption and very few studies in the domain of diet and because results from trial participants may not represent what would happen in the real world (Deaton and Cartwright, 2018).

Despite these limitations we believe that our primary conclusion on the lack of evidence of incentives as a cost-effective strategy for addressing rising rates of NCDs is justified. To put this conclusion into context, if incentives for NCD prevention were being considered for subvention by NICE or another HTA agency and evaluated as are medicines, we believe these agencies would not recommend subvention due to a lack of compelling evidence of long term effectiveness and cost-effectiveness. For those wishing to make the case, this evidence is sorely needed.

5. Contributions

All authors have contributed to all stages of the conceptualizing, drafting, editing and finalizing of this Review.

Declaration of interest

None.

Acknowledgements

The authors and research assistants are supported by internal funding from the Duke-NUS Medical School, Singapore. We would like to thank research assistants, Ms. Shweta Mital, Ms. Isha Chaudhry and Ms. Yingyu Guo, for their help in finalizing this Review.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2019.05.018.

References

Ali, A., Kaplan, C.M., Derefinko, K.J., Klesges, R.C., 2018. Smoking cessation for smokers not ready to quit: meta-analysis and cost-effectiveness analysis. Am. J. Prev. Med. 55 (2), 253–262 ISSN: 0749-3797.
Baker, T.B., Fraser, D.L., Kobinsky, K., Adsit, R., Smith, S.S., Khalil, L., et al., 2018. A randomized controlled trial of financial incentives to low income pregnant women to engage in smoking cessation treatment: effects on post-birth abstinence. J. Consult. Clin. Psychol. 86, 464–473.
Carrera, M., Royer, H., Stehr, M., Sydnor, J., 2018. Can financial incentives help people trying to establish new habits? Experimental evidence with new gym members. J. Health Econ. 58, 202–214.
Chapman, G.B., Brewer, N.T., Coups, E.J., Brownlee, S., Leventhal, H., Levanthal, E.A., 2001. Value for the future and preventive health behavior. J. Exp. Psychol. Appl. 7, 235–250.
Charness, G., Gneezy, U., 2009. Incentives to exercise. Econometrica 77, 909–931.
Courneya, K.S., Estabrooks, P.A., Nigg, C.R., 1997. A simple reinforcement strategy for increasing attendance at a fitness facility. Health Educ. Behav. 24, 708–715.
Crowley, T.J., Macdonald, M.J., Walter, M.I., 1995. Behavioral anti-smoking trial in chronic obstructive pulmonary disease patients. Psychopharmacology 119, 193–204.
Deaton, A., Cartwright, N., 2018. Understanding and misunderstanding randomized controlled trials. Soc. Sci. Med. 210, 2–21.
Donatelle, R.J., Prows, S.L., Champeau, D., Hudson, D., 2000. Randomised controlled trial using social support and financial incentives for high risk pregnant smokers: significant Other Supporter (SOS) program. Tobac. Contr. 9, 67–69.
Duhigg, C., 2012. The Power of Habit: Why We Do what We Do in Life and Business. Random House.
Etter, J.-F., Schmid, F., 2016. Effects of large financial incentives for long-term smoking cessation: a randomized trial. J. Am. Coll. Cardiol. 68, 777–785.
Faghri, P.D., Li, R., 2014. Effectiveness of financial incentives in a worksite diabetes prevention program. Open Obes. J. 6, 1.
Finkelstein, E.A., Brown, D.S., Brown, D.R., Buchner, D.M., 2008. A randomized study of financial incentives to increase physical activity among sedentary older adults. Prev. Med. 47, 182–187.
Finkelstein, E.A., Haaland, B.A., Bilger, M., Sahasranaman, A., Sloan, R.A., Nang, E.E.K., et al., 2016. Effectiveness of activity trackers with and without incentives to increase physical activity (TRIPPA): a randomised controlled trial. The Lancet Diabetes & Endocrinology 4 (12), 983–995 ISSN: 2213-8587.
Finkelstein, E.A., Kruger, E., 2014. Meta-and cost-effectiveness analysis of commercial weight loss strategies. Obesity 22, 1942–1951.
Finkelstein, E.A., Kruger, E., Karnawat, S., 2015. Cost-effectiveness analysis of Qsymia for weight loss. Pharmacoeconomics 33, 699–706.
Finkelstein, E.A., Linnan, L.A., Tate, D.F., Birken, B.E., 2007. A pilot study testing the effect of different levels of financial incentives on weight loss among overweight employees. J. Occup. Environ. Med. 49, 981–989.
Finkelstein, E.A., Tham, K.-W., Haaland, B.A., Sahasranaman, A., 2017. Applying economic incentives to increase effectiveness of an outpatient weight loss program (TRIO)–a randomized controlled trial. Soc. Sci. Med. 185, 63–70.
Fraser, D.L., Fiore, M.C., Kobinsky, K., Adsit, R., Smith, S.S., Johnson, M.L., et al., 2017. A randomized trial of incentives for smoking treatment in Medicaid members. Am. J. Prev. Med. 53, 754–763.
French, S.A., Rydell, S.A., Mitchell, N.R., Oakes, J.M., Elbel, B., Harnack, L., 2017. Financial incentives and purchase restrictions in a food benefit program affect the types of foods and beverages purchased: results from a randomized trial. Int. J. Behav. Nutr. Phys. Act. 14, 127.
Gaarder, M.M., Glassman, A., Todd, J.E., 2010. Conditional cash transfers and health: unpacking the causal chain. J. Dev. Eff. 2, 6–50.
Giles, E.L., Robalino, S., McColl, E., Sniehotta, F.F., Adams, J., 2014. The effectiveness of financial incentives for health behaviour change: systematic review and meta-analysis. PLoS One 9, e90347.
Giné, X., Karlan, D., Zinman, J., 2010. Put your money where your butt is: a commitment contract for smoking cessation. Am. Econ. J. Appl. Econ. 213–235.
Gneezy, U., Meier, S., Rey-Biel, P., 2011. When and why incentives (don't) work to modify behavior. J. Econ. Perspect. 25, 191–210.
Gneezy, U., Rustichini, A., 2000. Pay enough or don't pay at all. Q. J. Econ. 115, 791–810.
Halpern, S.D., French, B., Small, D.S., Saulsgiver, K., Harhay, M.O., Audrain-McGovern, J., et al., 2015. Randomized trial of four financial-incentive programs for smoking cessation. N. Engl. J. Med. 372, 2108–2117.
Harkins, K.A., Kullgren, J.T., Bellamy, S.L., Karlawish, J., Glanz, K., 2017. A trial of financial and social incentives to increase older adults' walking. Am. J. Prev. Med. 52, e123–e130.
Health, 2014. Promotion Board (Singapore).
Heil, S.H., Higgins, S.T., Bernstein, I.M., Solomon, L.J., Rogers, R.E., Thomas, C.S., et al., 2008. Effects of voucher-based incentives on abstinence from cigarette smoking and fetal growth among pregnant women. Addiction 103, 1009–1018.
Hennrikus, D.J., Jeffery, R.W., Lando, H.A., Murray, D.M., Brelje, K., Davidann, B., et al., 2002. The SUCCESS project: the effect of program format and incentives on participation and cessation in worksite smoking cessation programs. Am. J. Public Health 92, 274–279.
Higgins, S.T., Washio, Y., Lopez, A.A., Heil, S.H., Solomon, L.J., Lynch, M.E., et al., 2014. Examining two different schedules of financial incentives for smoking cessation among pregnant women. Prev. Med. 68, 51–57.
Husereau, D., Drummond, M., Petrou, S., Carswell, C., Moher, D., Greenberg, D., et al., 2013. Consolidated health economic evaluation reporting standards (CHEERS) statement. Cost Eff. Resour. Allocation 11, 6.

International Monetary Fund, 2017. World Economic Outlook Database.
Jason, L.A., Salina, D., McMahon, S.D., Hedeker, D., Stockton, M., 1997. A worksite smoking intervention: a 2 year assessment of groups, incentives and self-help. Health Educ. Res. 12, 129–138.
Jeffery, R.W., Wing, R.R., Thorson, C., Burton, L.R., 1998. Use of personal trainers and financial incentives to increase exercise in a behavioral weight-loss program. J. Consult. Clin. Psychol. 66, 777.
Jeffrey, S.A., 2009. Justifiability and the motivational power of tangible noncash incentives. Hum. Perform. 22, 143–155.
Jeffrey, S.A., Shaffer, V., 2007. The motivational properties of tangible incentives. Compens. Benefit Rev. 39, 44–50.
John, L.K., Loewenstein, G., Troxel, A.B., Norton, L., Fassbender, J.E., Volpp, K.G., 2011. Financial incentives for extended weight loss: a randomized, controlled trial. J. Gen. Intern. Med. 26, 621–626.
Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica: Journal of the Econometric Society 263–291.
Kelly, K., Presslee, A., Webb, A., 2015. The Effects of Tangible Rewards versus Cash Rewards in a Sales Tournament: A Field Experiment. Incentive Research Foundation.
Kendzor, D.E., Businelle, M.S., Poonawalla, I.B., Cuate, E.L., Kesh, A., Rios, D.M., et al., 2015. Financial incentives for abstinence among socioeconomically disadvantaged individuals in smoking cessation treatment. Am. J. Public Health 105, 1198–1205.
Kullgren, J.T., Harkins, K.A., Bellamy, S.L., Gonzales, A., Tao, Y.Y., Zhu, J.S., et al., 2014. A mixed-methods randomized controlled trial of financial incentives and peer networks to promote walking among older adults. Health Educ. Behav. 41, 43S–50S.
Kullgren, J.T., Troxel, A.B., Loewenstein, G., Asch, D.A., Norton, L.A., Wesby, L., et al., 2013. Individual- versus group-based financial incentives for weight loss a randomized, controlled trial. Ann. Intern. Med. 158, 505.
Kullgren, J.T., Troxel, A.B., Loewenstein, G., Norton, L.A., Gatto, D., Tao, Y., et al., 2016. A randomized controlled trial of employer matching of employees' monetary contributions to deposit contracts to promote weight loss. Am. J. Health Promot. 30, 441–452.
Lagarde, M., Haines, A., Palmer, N., 2007. Conditional cash transfers for improving uptake of health interventions in low-and middle-income countries: a systematic review. Jama 298, 1900–1910.
Lagarde, M., Haines, A., Palmer, N., 2009. The impact of conditional cash transfers on health outcomes and use of health services in low and middle income countries. Cochrane Database Syst. Rev. (4), CD008137 ISSN: 1361-6137.
Lally, P., Chipperfield, A., Wardle, J., 2008. Healthy habits: efficacy of simple advice on weight control based on a habit-formation model. Int. J. Obes. 32, 700–707.
Lally, P., Gardner, B., 2013. Promoting habit formation. Health Psychol. Rev. 7, S137–S158.
Lamb, R.J., Morral, A.R., Kirby, K.C., Javors, M.A., Galbicka, G., Iguchi, M., 2007. Contingencies for change in complacent smokers. Exp. Clin. Psychopharmacol 15, 245–255.
Leahey, T.M., Subak, L.L., Fava, J., Schembri, M., Thomas, G., Xu, X., et al., 2015. Benefits of adding small financial incentives or optional group meetings to a web-based statewide obesity initiative. Obesity 23, 70–76.
Ledgerwood, D.M., Arfken, C.L., Petry, N.M., Alessi, S.M., 2014. Prize contingency management for smoking cessation: a randomized trial. Drug Alcohol Depend. 140, 208–212.
Mantzari, E., Vogt, F., Shemilt, I., Wei, Y., Higgins, J.P., Marteau, T.M., 2015. Personal financial incentives for changing habitual health-related behaviors: a systematic review and meta-analysis. Prev. Med. 75, 75–85.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., Group, P., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 6, e1000097.
Neal, D.T., Wood, W., Quinn, J.M., 2006. Habits—a repeat performance. Curr. Dir. Psychol. Sci. 15, 198–202.
Ondersma, S.J., Svikis, D.S., Lam, P.K., Connors-Burge, V.S., Ledgerwood, D.M., Hopper, J.A., 2012. A randomized trial of computer-delivered brief intervention and low-intensity contingency management for smoking during pregnancy. Nicotine Tob. Res. 14, 351–360.
Ouellette, J.A., Wood, W., 1998. Habit and intention in everyday life: the multiple processes by which past behavior predicts future behavior. Psychol. Bull. 124, 54.
Paloyo, A.R., Reichert, A.R., Reuss-Borst, M., Tauchmann, H., 2015. Who responds to financial incentives for weight loss? Evidence from a randomized controlled trial. Soc. Sci. Med. 145, 44–52.
Patel, M.S., Asch, D.A., Rosin, R., Small, D.S., Bellamy, S.L., Eberbach, K., et al., 2016a. Individual versus team-based financial incentives to increase physical activity: a randomized, controlled trial. J. Gen. Intern. Med. 1–9.
Patel, M.S., Asch, D.A., Roy Rosin, M., Small, D.S., Bellamy, S.L., Heuer, J., et al., 2016b. Framing financial incentives to increase physical activity among overweight and obese adults. Ann. Intern. Med. 164, 385–394.
Patel, M.S., Asch, D.A., Troxel, A.B., Fletcher, M., Osman-Koss, R., Brady, J., et al., 2016c. Premium-based financial incentives did not promote workplace weight loss in a 2013–15 study. Health Aff. 35, 71–79.
Petry, N.M., Andrade, L.F., Barry, D., Byrne, S., 2013. A randomized study of reinforcing ambulatory exercise in older adults. Psychol. Aging 28, 1164.
Petry, N.M., Barry, D., Pescatello, L., White, W.B., 2011. A low-cost reinforcement procedure improves short-term weight loss outcomes. Am. J. Med. 124, 1082–1085.
Pope, L., Harvey-Berino, J., 2013. Burn and earn: a randomized controlled trial incentivizing exercise during fall semester for college first-year students. Prev. Med. 56, 197–201.
Ranganathan, M., Lagarde, M., 2012. Promoting healthy behaviours and improving health outcomes in low and middle income countries: a review of the impact of conditional cash transfer programmes. Prev. Med. 55, S95–S105.
Read, D., Van Leeuwen, B., 1998. Predicting hunger: the effects of appetite and delay on choice. Organ. Behav. Hum. Decis. Process. 76, 189–205.
Royer, H., Stehr, M., Sydnor, J., 2015. Incentives, commitments, and habit formation in exercise: evidence from a field experiment with workers at a fortune-500 company. Am. Econ. J. Appl. Econ. 7, 51.
Schulz, K.F., Altman, D.G., Moher, D., 2010. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMC Med. 8, 18.
Secades-Villa, R., Garcia-Rodriguez, O., Lopez-Nunez, C., Alonso-Perez, F., Fernandez-Hermida, J.R., 2014. Contingency management for smoking cessation among treatment-seeking patients in a community setting. Drug Alcohol Depend. 140, 63–68.
Shin, D.W., Yun, J.M., Shin, J.H., Kwon, H., Min, H.Y., Joh, H.K., et al., 2017. Enhancing physical activity and reducing obesity through smartcare and financial incentives: a pilot randomized trial. Obesity 25, 302–310.
Skinner, B.F., 1963. Operant behavior. Am. Psychol. 18, 503.
Sutherland, K., Christianson, J.B., Leatherman, S., 2008. Impact of targeted financial incentives on personal health behavior A review of the literature. Med. Care Res. Rev. 65, 36S–78S.
Tappin, D., Bauld, L., Purves, D., Boyd, K., Sinclair, L., MacAskill, S., et al., 2015. Financial incentives for smoking cessation in pregnancy: randomised controlled trial. BMJ 350, h134.
Thaler, R., 1981. Some empirical evidence on dynamic inconsistency. Econ. Lett. 8, 201–207.
Thaler, R., 1985. Mental accounting and consumer choice. Mark. Sci. 4, 199–214.
Thaler, R.H., 1999. Mental accounting matters. J. Behav. Decis. Mak. 12, 183–206.
Thorndike, A.N., Riis, J., Levy, D.E., 2016. Social norms and financial incentives to promote employees' healthy food choices: a randomized controlled trial. Prev. Med. 86, 12–18.
USDA, 2018. Healthy Incentives Pilot Final Evaluation Report.
Volpp, K.G., John, L.K., Troxel, A.B., Norton, L., Fassbender, J., Loewenstein, G., 2008. Financial incentive-based approaches for weight loss a randomized trial. Jama-Journal of the American Medical Association 300, 2631–2637.
Volpp, K.G., Levy, A.G., Asch, D.A., Berlin, J.A., Murphy, J.J., Gomez, A., et al., 2006. A randomized controlled trial of financial incentives for smoking cessation. Cancer Epidemiol. Biomark. Prev. 15, 12–18.
Volpp, K.G., Troxel, A.B., Pauly, M.V., Glick, H.A., Puig, A., Asch, D.A., et al., 2009. A randomized, controlled trial of financial incentives for smoking cessation. N. Engl. J. Med. 360, 699–709.
Wing, R.R., Jeffery, R.W., Pronk, N., Hellerstedt, W.L., 1996. Effects of a personal trainer and financial incentives on exercise adherence in overweight women in a behavioral weight loss program. Obes. Res. 4, 457–462.
Wood, W., Tam, L., Witt, M.G., 2005. Changing circumstances, disrupting habits. J. Personal. Soc. Psychol. 88, 918.
World Bank, 2017. World Development Indicators.
World Health Organization, 2015. Noncommunicable Diseases Prematurely Take 16 Million Lives Annually, WHO Urges More Action.
World Health Organization, 2017. Sustainable Development Goals (SDGs).