TECHNICAL COMMENTARY

Monetary Incentives

NeuRA Discover. Conquer. Cure.

SCHIZOPHRENIA LIBRARY

Introduction

Monetary incentives have been proposed as a form of positive behavioural reinforcement, given as a reward for the implementation of target behaviour. In the real world this can include employment or living allowances, but they can be applied experimentally as lower values which have greater symbolic significance. At the lowest end these incentives become valueless tokens (see Token Economies table).

The notion of behavioural reinforcement may have applicability in schizophrenia as patients can be unresponsive to everyday social rewards, such as praise[1]. Monetary rewards may help to increase motivation, for example in trying to quit smoking[2].

Method

We have included only systematic reviews (systematic literature search, detailed methodology with inclusion/exclusion criteria) published in full text, in English, from the year 2000 that report results separately for people with a diagnosis of schizophrenia, schizoaffective disorder, schizophreniform disorder or first episode schizophrenia. Reviews were identified by searching the databases MEDLINE, EMBASE, CINAHL, Current Contents, PsycINFO and the Cochrane library. Hand searching reference lists of identified reviews was also conducted. When multiple copies of reviews were found, only the most recent version was included. Reviews with pooled data are prioritised for inclusion.

Review reporting assessment was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist that describes a preferred way to present a meta-analysis[3]. Reviews rated as having less than 50% of items checked have been excluded from the library. The PRISMA flow diagram is a suggested way of providing information about studies included and excluded with reasons for exclusion. Where no flow diagram has been presented by individual reviews, but identified studies have been described in the text, reviews have been checked for this item. Note that early reviews may have been guided by less stringent reporting checklists than the PRISMA, and that some reviews may have been limited by journal guidelines.

Evidence was graded using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group approach where high quality evidence such as that gained from randomised controlled trials (RCTs) may be downgraded to moderate or low if review and study quality is limited, if there is inconsistency in results, indirect comparisons, imprecise or sparse data and high probability of reporting bias. It may also be downgraded if risks associated with the intervention or other matter under review are high. Conversely, low quality evidence such as that gained from observational studies may be upgraded if effect sizes are large, there is a dose dependent response or if results are reasonably consistent, precise and direct with low associated risks (see end of table for an explanation of these terms)[4]. The resulting table represents an objective summary of the available evidence, although the conclusions are solely the opinion of staff of NeuRA (Neuroscience Research Australia).

Results

We found three systematic reviews that met inclusion criteria[1, 2, 5].

- Low quality evidence is unclear as to the benefits of monetary incentive programs.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Hjorthoj C, Fohlmann A, Nordentoft M

Treatment of cannabis use disorders in people with schizophrenia spectrum disorders - A systematic review

Addictive Behaviors 2009; 34(6-7): 520-525

View review abstract online

Comparison	**Contingency management (cash or vouchers in return for clean urine sample) vs. treatment as usual for reducing cannabis use.**
Summary of evidence	**Low quality evidence (unable to assess consistency or precision) is unclear as to the benefit of monetary incentives for reducing cannabis use in the long term.**
Cannabis use	
2 studies (N = 25) reported short-term benefits of monetary incentives in response to clean urine samples (indicating no cannabis use), which were not maintained in the longer term.	
Consistency in results‡	Unable to assess, no measure of consistence is reported.
Precision in results§	Unable to assess, no measure of consistence is reported.
Directness of results‖	Direct

Michalczuk R, Mitchell A

Monetary incentives for schizophrenia

Cochrane Database of Systematic Reviews 2009; (4): CD007626. doi: 10.1002/14651858.CD007626.pub2

View review abstract online

Comparison	**Monetary reward in exchange for target behaviour plus standard care vs. standard care.**
Summary of evidence	**Low quality evidence (small sample, imprecise, direct) is unclear as to the benefit of monetary incentive.**

Target behaviour: average number of dolls assembled per day in return for reward	
A significant, large effect favouring the 'no reward' group for achieving target behaviour; 1 RCT, N = 25, *d* = -0.96, 95%CI -1.81 to -0.11, *p* = 0.03	
Consistency in results	Not applicable (1 RCT)
Precision in results	Imprecise
Directness of results	Direct

Tsoi DT, Porwal M, Webster AC

Interventions for smoking cessation and reduction in individuals with schizophrenia

Cochrane Database of Systematic Reviews 2010; (6): Art. No.: CD007253 doi: 10.1002/14651858.CD007253.pub3

View review abstract online

Comparison	**Contingent Reinforcement (CR) using money with Transdermal Nicotine Patch (TNP) vs. CR alone vs. no active intervention (self-quit) for aiding smoking cessation.**
Summary of evidence	**Low quality evidence (direct, small sample, unable to assess consistency) is unable to determine the benefits of contingency reinforcement for smoking cessation.**
Smoking abstinence	
1 RCT, N = 80, showed 32.5% of participants expressed an interest in quitting smoking. Abstinence rates were significantly higher in the CR with TNP (50%) compared to the CR group (27.8%) or the self-quit group (10%) at the end of the 36 week trial. The CR with TNP group had significantly lower scores on the Fagerstrom Test for Nicotine Dependence (FTND) at 20 and 36 weeks compared to both comparison groups. The CR with TNP group had significantly lower expired carbon monoxide level at the end of the trial compared to self-quit, but not CR alone.	
Cigarettes smoked per day (CPD)	

Monetary Incentives

CPD was lower at week 36 in the CR with TNP group compared to the self-quit group. CR alone showed no difference to either comparison group by the end of the trial.	
Consistency in results	Not applicable (1 RCT). No measures of consistency is reported
Precision in results	No measures of precision is reported.
Directness of results	Direct

Explanation of acronyms

CI = Confidence Interval, CO = Carbon monoxide, CPD = Cigarettes Per Day, CR = Contingent Reinforcement, *d* = Cohen's *d*, FTND = Fagerstrom Test for Nicotine Dependence, *g* = Hedges' *g* = standardized mean differences (see below for interpretation of effect size), N = number of participants, *p* = statistical probability of obtaining that result ($p < 0.05$ generally regarded as significant), TNP = Transdermal Nicotine Patch, vs = versus

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Explanation of technical terms

* Bias has the potential to affect reviews of both RCT and observational studies. Forms of bias include; publication bias - trials that are not formally published tend to show less effect than published trials, further if there are statistically significant differences between groups in a trial, these trial results tend to get published before those of trials without significant differences; language bias – only including English language reports; funding bias - source of funding for the primary research with selective reporting of results within primary studies; outcome variable selection bias; database bias - including reports from some databases and not others; citation bias - preferential citation of authors. Trials can also be subject to bias when evaluators are not blind to treatment condition and selection bias of participants if trial samples are small[6].

† Different effect measures are reported by different reviews.

Weighted mean difference scores refer to mean differences between treatment and comparison groups after treatment (or occasionally pre to post treatment) and in a randomised trial there is an assumption that both groups are comparable on this measure prior to treatment. Standardised mean differences are divided by the pooled standard deviation (or the standard deviation of one group when groups are homogenous) which allows results from different scales to be combined and compared. Each study's mean difference is then given a weighting depending on the size of the sample and the variability in the data. 0.2 represents a small effect, 0.5 a medium effect, and 0.8 and over represents a large teffect[6].

Prevalence refers to how many existing cases there are at a particular point in time. Incidence refers to how many new cases there are per population in a specified time period. Incidence is usually reported as the number of new cases per 100,000 people per year. Alternatively some studies present the number of new cases that have accumulated over several years against a person-years denominator. This denominator is the sum of individual units of time that the persons in the population are at risk of becoming a case. It takes into account the size of the underlying population sample and its age structure over the duration of observation.

Reliability and validity refers to how accurate the instrument is. Sensitivity is the proportion of actual positives that are correctly identified (100% sensitivity = correct identification of all actual positives) and specificity is the proportion of negatives that are correctly identified (100% specificity = not identifying anyone as positive if they are truly not).

Odds ratio (OR) or relative risk (RR) refers to the probability of a reduction (< 1) or an increase (> 1) in a particular outcome in a treatment group, or a group exposed to a risk factor, relative to the comparison group. For example, a RR of 0.75 translates to a reduction in risk of an outcome of 25% relative to those not receiving the treatment or not exposed to the risk factor. Conversely, an RR of 1.25 translates to an increased risk of 25% relative to those not receiving treatment or not having been exposed to a risk factor. An RR or OR of 1.00 means there is no difference between groups. A medium effect is considered if RR > 2 or < 0.5 and a large effect if RR > 5 or < 0.2[7]. lnOR stands for logarithmic OR where an lnOR of 0 shows no difference between groups. Hazard ratios measure the effect of an explanatory variable on the hazard or risk of an event.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Correlation coefficients (eg, r) indicate the strength of association or relationship between variables. They are an indication of prediction, but do not confirm causality due to possible and often unforseen confounding variables. An r of 0.10 represents a weak association, 0.25 a medium association and 0.40 and over represents a strong association. Unstandardised (*b*) regression coefficients indicate the average change in the dependent variable associated with a 1 unit change in the dependent variable, statistically controlling for the other independent variables. Standardised regression coefficients represent the change being in units of standard deviations to allow comparison across different scales.

‡ Inconsistency refers to differing estimates of treatment effect across studies (i.e. heterogeneity or variability in results) which is not explained by subgroup analyses and therefore reduces confidence in the effect estimate. I^2 is the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance) - 0% to 40%: heterogeneity might not be important, 30% to 60%: may represent moderate heterogeneity, 50% to 90%: may represent substantial heterogeneity and 75% to 100%: considerable heterogeneity. I^2 can be calculated from Q (chi-square) for the test of heterogeneity with the following formula[6];

$$I^2 = \left(\frac{Q - df}{Q}\right) \times 100\%$$

§ Imprecision refers to wide confidence intervals indicating a lack of confidence in the effect estimate. Based on GRADE recommendations, a result for continuous data (standardised mean differences, not weighted mean differences) is considered imprecise if the upper or lower confidence limit crosses an effect size of 0.5 in either direction, and for binary and correlation data, an effect size of 0.25. GRADE also recommends downgrading the evidence when sample size is smaller than 300 (for binary data) and 400 (for continuous data), although for some topics, this criteria should be relaxed[8].

‖ Indirectness of comparison occurs when a comparison of intervention A versus B is not available but A was compared with C and B was compared with C that allows indirect comparisons of the magnitude of effect of A versus B. Indirectness of population, comparator and or outcome can also occur when the available evidence regarding a particular population, intervention, comparator, or outcome is not available so is inferred from available evidence. These inferred treatment effect sizes are of lower quality than those gained from head-to-head comparisons of A and B.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

References

1. Michalczuk R, Mitchell A. Monetary incentives for schizophrenia. *Cochrane Database of Systematic Reviews*. 2009; (4).
2. Tsoi D, Porwal M, Webster A. Interventions for smoking cessation and reduction in individuals with schizophrenia. *Cochrane Database of Systematic Reviews*. 2010; **6**(pp CD007253).
3. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMAGroup. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *British Medical Journal*. 2009; **151**(4): 264-9.
4. GRADEWorkingGroup. Grading quality of evidence and strength of recommendations. *British Medical Journal*. 2004; **328**: 1490.
5. Hjorthoj C, Fohlmann A, Nordentoft M. Treatment of cannabis use disorders in people with schizophrenia spectrum disorders - A systematic review. *Addictive Behaviors*. 2009; **34**(6-7): 520-5.
6. CochraneCollaboration. Cochrane Handbook for Systematic Reviews of Interventions. 2008: Accessed 24/06/2011.
7. Rosenthal JA. Qualitative Descriptors of Strength of Association and Effect Size. *Journal of Social Service Research*. 1996; **21**(4): 37-59.
8. GRADEpro. [Computer program]. Jan Brozek, Andrew Oxman, Holger Schünemann. *Version 32 for Windows*. 2008.