![Cochrane Library logo]

Cochrane Database of Systematic Reviews

Virtual reality for treatment compliance for people with serious mental illness (Review)

Välimäki M, Hätönen HM, Lahti ME, Kurki M, Hottinen A, Metsäranta K, Riihimäki T, Adams CE

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Virtual reality for treatment compliance for people with serious mental illness

Maritta Välimäki[1,2], Heli M Hätönen[1], Mari E Lahti[1,3], Marjo Kurki[1], Anja Hottinen[1], Kiki Metsäranta[1], Tanja Riihimäki[4], Clive E Adams[5]

[1]Department of Nursing Science, University of Turku, Turku, Finland. [2]Turku University Hospital, Turku, Finland. [3]Health and Social Science, Turku University of Applied Sciences, Turku, Finland. [4]Psychiatric Unit, Outpatient Clinic for Adolescents, The Helsinki University Central Hospital, Espoo, Finland. [5]Cochrane Schizophrenia Group, The University of Nottingham, Nottingham, UK

Contact address: Maritta Välimäki, Department of Nursing Science, University of Turku, Turku, Finland. mava@utu.fi.

ABSTRACT

Background

Virtual reality (VR) is computerised real-time technology, which can be used an alternative assessment and treatment tool in the mental health field. Virtual reality may take different forms to simulate real-life activities and support treatment.

Objectives

To investigate the effects of virtual reality to support treatment compliance in people with serious mental illness.

Search methods

We searched the Cochrane Schizophrenia Group Trials Register (most recent, 17th September 2013) and relevant reference lists.

Selection criteria

All relevant randomised studies comparing virtual reality with standard care for those with serious mental illnesses. We defined virtual reality as a computerised real-time technology using graphics, sound and other sensory input, which creates the interactive computer-mediated world as a therapeutic tool.

Data collection and analysis

All review authors independently selected studies and extracted data. For homogeneous dichotomous data the risk difference (RD) and the 95% confidence intervals (CI) were calculated on an intention-to-treat basis. For continuous data, we calculated mean differences (MD). We assessed risk of bias and created a 'Summary of findings' table using the GRADE approach.

Main results

We identified three short-term trials (total of 156 participants, duration five to 12 weeks). Outcomes were prone to at least a moderate risk of overestimating positive effects. We found that virtual reality had little effects regarding compliance (3 RCTs, n = 156, RD loss to follow-up 0.02 CI -0.08 to 0.12, *low quality evidence*), cognitive functioning (1 RCT, n = 27, MD average score on Cognistat 4.67 CI -1.76 to 11.10, *low quality evidence*), social skills (1 RCT, n = 64, MD average score on social problem solving SPSI-R (Social Problem

Solving Inventory - Revised) -2.30 CI -8.13 to 3.53, *low quality evidence*), or acceptability of intervention (2 RCTs, n = 92, RD 0.05 CI -0.09 to 0.19, *low quality evidence*). There were no data reported on mental state, insight, behaviour, quality of life, costs, service utilisation, or adverse effects. Satisfaction with treatment - measured using an un-referenced scale - and reported as "interest in training" was better for the virtual reality group (1 RCT, n = 64, MD 6.00 CI 1.39 to 10.61,*low quality evidence*).

Authors' conclusions

There is no clear good quality evidence for or against using virtual reality for treatment compliance among people with serious mental illness. If virtual reality is used, the experimental nature of the intervention should be clearly explained. High-quality studies should be undertaken in this area to explore any effects of this novel intervention and variations of approach.

PLAIN LANGUAGE SUMMARY

Virtual reality needs to be explored to show its effects for treatment of schizophrenia

People with schizophrenia often have problems in their processes of thinking and understanding, resulting in poor insight into their illness and poor organisational skills. These factors along with experiencing unpleasant side effects of medication can contribute to people with mental health problems often not taking their medication, unwilling to follow treatment and non-attendance at appointments. This can sometimes lead to a loss of contact with the mental health team and relapse. Virtual reality (VR) is a modern, experimental, computerised and real-time technology that uses visual graphics, sounds and other sensory input which creates an interactive computer world. It includes, for example, the use of sensors attached to the hands and fingers allowing virtual reality users to track their position and movement. Virtual reality creates a computerised environment that simulates real life and everyday activities. This could help people learn in a safe and friendly environment to improve their decisions and attitudes about treatment, so encouraging people to take or comply with their medication.

So far, virtual reality has been used in the assessment and treatment of a range of psychiatric disorders and social anxieties, some of which include, fear of flying, public speaking anxiety, spider phobia, and post-traumatic stress disorder. There are also a few studies that examine the emotional responses of people with schizophrenia during a computer simulation with characters displaying happy, neutral, and angry emotions. Virtual reality has also been used for people with schizophrenia in social skills training and to improve processes of thinking and understanding. This review investigates the effects of virtual reality in helping support the treatment and taking of medication for people with serious mental illness.

The most recent search for randomised trials was run in September 2013, only three short studies with a total of 156 people could be included. People with schizophrenia were randomised to a) skills training sessions that used virtual reality to deliver the training or b) sessions of skills training using other methods to deliver the training or c) standard care. All evidence from the trials was low quality and no real effects were found. At present, there is no clear evidence for or against using virtual reality for encouraging people with mental illness to take their medication. If virtual reality is used for people with serious mental illness, it will be of an experimental nature. There is a need to gather more good quality information on the effects of virtual reality for people with mental illness and high quality studies need to be undertaken. At this stage, the effects of virtual reality are experimental, novel and innovative but largely untested.

This summary has been written by a consumer, Ben Gray of RETHINK.

VIRTUAL REALITY AS A SOLE METHOD + STANDARD PROFESSIONAL CARE compared with STANDARD PROFESSIONAL CARE for treatment compliance for people with serious mental illness

Patient or population: patients with treatment compliance for people with serious mental illness
Settings: specialist centres
Intervention: VIRTUAL REALITY AS A SOLE METHOD + STANDARD PROFESSIONAL CARE
Comparison: STANDARD PROFESSIONAL CARE

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	**Corresponding risk**				
	STANDARD PROFESSIONAL CARER	**VIRTUAL REALITY AS A SOLE METHOD + STANDARD PROFESSIONAL CARER**				
Compliance Measured using: loss to follow-up Follow-up: 5-12 weeks	**Low[1]**		See comment	156 (3 studies)	⊕⊕◯◯ **low**[2,3]	Risks were calculated from pooled risk differences
	20 per 1000	0 per 1000 (-2 to 2)				
	Moderate[1]					
	70 per 1000	1 per 1000 (-6 to 8)				
	High[1]					
	120 per 1000	2 per 1000 (-10 to 14)				
Mental state: any outcome	See comment	See comment	Not estimable	0 (0)	See comment	No study reported this outcome.

Functioning: 1. Cognitive - average endpoint total Measured using: Cognistat Follow-up: 12 weeks		The mean functioning: 1. cognitive - average endpoint total in the intervention groups was **4.67 higher** (1.76 lower to 11.1 higher)		27 (1 study)	⊕⊕◯◯ **low**[2,4]	
Functioning: 2. Social - average change in social problem solving Measured using: SPSI-R Follow-up: 5 weeks[5]		The mean functioning: 2. social - average change in social problem solving in the intervention groups was **2.3 lower** (8.13 lower to 3.53 higher)		64 (1 study)	⊕⊕◯◯ **low**[2,4]	
Quality of life: any outcome	See comment	See comment	Not estimable	0 (0)	See comment	No study reported this outcome.
Satisfaction with treatment: Average change score - interest in social skills training Measured using: un-referenced measure Follow-up: 5 weeks[5]		The mean satisfaction with treatment: average change score - interest in social skills training in the intervention groups was **6 higher** (1.39 to 10.61 higher)		64 (1 study)	⊕⊕◯◯ **low**[2,6]	
Acceptability of intervention Measured using: leaving the study early for any reason Follow-up: 5 to 12 weeks	**Moderate**		See comment	92 (2 studies)	⊕⊕◯◯ **low**[2,7]	Risks were calculated from pooled risk differences
	70 per 1000	**4 per 1000** (-6 to 13)				

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; **RR:** Risk ratio;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Moderate risk roughly - that of control group.
[2] Risk of bias: rated 'moderate' - randomisation described in one study out of three studies, concealment not described, outcome assessment blinded in one out of three studies.
[3] Indirectness: rated 'serious' - no direct measure of compliance with treatment.
[4] Indirectness: rated as 'serious' - unclear clinical meaning of scores.
[5] Indirectness: rated as 'moderate' - we have assumed this follow-up from duration of intervention.
[6] Indirectness: rated 'very serious' - unclear meaning of un-referenced scale.
[7] Indirectness: rated as 'serious' - unclear how leaving the study early was conducted.

B A C K G R O U N D

Description of the condition

Schizophrenia is a chronic disabling mental illness affecting approximately 1% of the population. It is associated with a broad array of cognitive impairments (Buchanan 2005), while language and communication deficits are also common (McClellan 2005). Impairments in attention and executive function, visual learning and working memory occur (Kirkpatrick 2005), which are critically important and account for much of the variance in poor social and occupational functional outcomes (Buchanan 2005) and treatment compliance (Johansen 2011). In more detail, hallucinations and delusions are often called positive symptoms of schizophrenia. Negative symptoms, e.g. loss of pleasure, loss of initiative, poverty of speech and affective blunting, are associated with poor functional capacity and further difficulties in social relationships. (APA 1994). People with schizophrenia and other non-affective psychotic disorders have a significantly increased risk of having mobility limitations as well as weak handgrip strength and reduced visual acuity. People with schizophrenia and other non-affective psychotic disorders have more limitations in everyday functioning, deficits in verbal fluency and in memory than general population. More severe negative symptoms, depression, older age, verbal memory deficits, worse expressive speech and impaired distance vision are associated with limitations in everyday functioning (Erickson 2011; Viertiö 2011).

Description of the intervention

Virtual reality (VR) is a modern computerised real-time technology using graphics, sounds and other sensory input, which creates an interactive computer-mediated world (Riva 2005; Kim 2009; Reger 2011). It includes, for example, a primarily visual VR environment and use of sensors attached to the hands and fingers allowing VR users to track their position and orientation of the hands as additional responses to the VR environment (Beuter 2004; Gregg 2007). VR applications have been developed for the assessment and treatment of psychiatric disorders, such as social phobia, fear of flying, fear of heights, public speaking anxiety, spider phobia, body image disturbance (Gregg 2007), agoraphobia (Malbos 2011) and post-traumatic stress disorder (Gerardi 2010; McLay 2011; Reger 2011). There are few studies examining the perception of emotion and the emotional responses of people with schizophrenia during a simulated social encounter with virtual avatars displaying happy, neutral, and angry emotions (Park 2009). VR has been used in schizophrenia patients social skills training (Ku 2007; Dyck 2010; Park 2011). It has been used as psychosocial interventions focusing on improved cognitive functions (Da Costa 2004). Further, studies concerning VR and people with serious mental illness have

shown that it is a potential means in delivering exposure therapy for a range of phobias and anxiety disorders (Gerardi 2010; Reger 2011). VR solutions have also been developed for use in the mental health field for post-traumatic stress disorders, obsessive-compulsive disorder (Kim 2009), male sexual dysfunction, or attention deficit disorder (Gregg 2007). Moreover, VR applications have been used to assist in the cognitive assessment and rehabilitation of patients with brain injury and schizophrenia (Esteves Moreira da Costa 2004; Gregg 2007).

How the intervention might work

Many factors contribute to patients' poor compliance including poor illness insight, a negative attitude toward medication, substance abuse, and disorganisation (Goff 2011). People with schizophrenia often have cognitive deficiencies, which are related to poor compliance (Goodman 2005). Interventions improving patients' compliance with treatment have shown to be long term and complex (Haynes 2008) consisting of multiple components such as advising acceptance of illness, drawing analogies with treatment for chronic medical disease, involving the patient in decision making and improving cognitive functioning. Moreover, trusting and encouraging relationships can support patients' adherence with treatment. (Haynes 2008; Goff 2011)
Technology-supported approaches such as VR have the potential to promote compliance by improving patients' cognitive functioning and integrating the key components of complex treatments (Coons 2011). VR can be an alternative, patient-friendly assessment and treatment tool in the mental health field (Kim 2009). While people with schizophrenia experience difficulties with activities in everyday life, VR can be used to test and support their performance using an environment that simulates real-life activities (Josman 2009) and compliance with treatment. For example, by using VR, patients can learn in a virtually simulated world how to cope with their medication, or how to cope with their fears hindering compliance. The system can be 'stand-alone', automatised, semi-automatised or manualised and managed by healthcare staff.

Why it is important to do this review

VR has the potential to provide complex and long-term interventions for supporting patients' compliance. Improved cognitive function may lead to other effects in daily life and self-esteem among people with schizophrenia. A VR cognitive training programme in rehabilitation may offer the potential for significant gains in the cognitive functioning of people with chronic schizophrenia (Chan 2010). Using VR applications for patients could enhance training outcomes by boosting their motivation for social skills training (Ku 2007). Further, VR holds great promise as a useful method to enhance patients' and their relatives' knowledge and understanding about mental illness.

VR has been growing rapidly within the last decade in the field of psychology and costs have diminished (Riva 2005; Gerardi 2010). However, there are barriers to using VR solutions because of a lack of standardisation in VR devices and software. In addition, PC-based systems lack flexibility and capabilities to individualize environments for patients (Riva 2005). Currently there is limited evidence available to determine the benefits of VR for treatment compliance for people with serious mental illness. Enhancing the recovery may lead to better self-management and quality of life for this vulnerable patient group. This review seeks to investigate evidence for the value of VR possibilities to increase the potential methods for the treatment of people with serious mental illness.

OBJECTIVES

To investigate the effectiveness of virtual reality to support treatment compliance in people with serious mental illness.

METHODS

Criteria for considering studies for this review

Types of studies

We included all relevant randomised controlled trials. If a trial had been described as 'double blind' but implied randomisation, we would have included such trials in a sensitivity analysis (see Sensitivity analysis). If there was no substantive difference within primary outcomes, they would have remained in the analyses. If their inclusion resulted in statistically significant differences, we would not have added the data from these lower quality studies to the results of the better trials, but we would have presented such data within a subcategory. We excluded quasi-randomised studies, such as those allocating by alternate days of the week. Where people were given additional treatments within virtual reality, we only included data if the adjunct treatment was evenly distributed between groups and it was only the virtual reality intervention that was randomised.

Types of participants

Adults, however defined, with schizophrenia or related disorders, including schizophreniform disorder, schizoaffective disorder and delusional disorder, again, by any means of diagnosis. We included only trials where the majority of participants had a diagnosis of schizophrenia.

We were interested in making sure that information is as relevant to the current care of people with schizophrenia as possible so

proposed, if possible, to clearly highlight the current clinical state (acute, early post-acute, partial remission, remission) as well as the stage (prodromal, first episode, early illness, persistent) and as to whether the studies primarily focused on people with particular problems (for example, negative symptoms, treatment-resistant illnesses).

Types of interventions

1. Virtual reality as a sole method

For the purpose of this review we defined virtual reality (VR) as a computerised real-time technology using graphics, sounds and other sensory input, which creates the interactive computer-mediated world as a therapeutic tool.

2. Virtual reality and standard professional care

Virtual reality used as a therapeutic tool additional to standard professional care.

3. Standard professional care

Standard professional care.

Types of outcome measures

We divided all outcomes into short term (less than six months), medium term (seven to 12 months) and long term (over one year).

Primary outcomes

1. Compliance

1.1 Loss to follow-up - loss of contact with the psychiatric care team (including loss to follow-up in outpatients and failure of psychiatric team to re-establish contact)
1.2 Compliance with medication
1.3 Attendance at appointments
1.4 Relapse (both incidence of and time to relapse)

Secondary outcomes

1. Mental state

1.1 Clinically important change in general mental state
1.2 Average change in general mental state scores
1.3 Clinically important change in specific symptoms (positive symptoms of schizophrenia, negative symptoms of schizophrenia, depression, mania)
1.4 Average change in specific symptom scores

---

2. Functioning

2.1 Clinically important change in functioning
2.2 Average change in social functioning scores
2.3 Clinically important change in specific aspects of functioning
2.4 Average change in specific aspects of functioning

3. Insight

3.1 Clinically important change in average level of insight
3.2 Average change in level of insight scores
3.3 Clinically important change in specific aspects of insight
3.4 Average change in specific aspects of insight

4. Skills

4.1 Clinically important change in general skills
4.2 Average change in general skill scores
4.3 Clinically important change in specific aspects of skills
4.4 Average change in specific aspects of skills

5. Behaviour

5.1 Clinically important change in general behaviour
5.2 Average change in general behaviour scores
5.3 Clinically important change in specific aspects of behaviour
5.4 Average change in specific aspects of behaviour

6. Quality of life

6.1 Clinically important change in quality of life
6.2 Average change in quality of life scores
6.3 Clinically important change in specific aspects of quality of life
6.4 Average change in specific aspects of quality of life

7. Satisfaction with treatment

7.1 Clinically important change in satisfaction with treatment
7.2 Average change in satisfaction with treatment scores
7.3 Clinically important change in specific aspects of satisfaction with treatment
7.4 Average change in specific aspects of satisfaction with treatment

8. Acceptability of treatment

8.1 Leaving the studies early - any reason
8.2 Leaving the studies early - specific technological reason

9. Costs

9.1 Direct costs
9.2 Indirect costs

10. Service utilisation

10.1 Admitted to psychiatric hospital
10.2 Mean days spent in psychiatric hospital per month
10.3 Number of contacts with own doctor for mental health problems
10.4 Number of contacts with psychiatric out-patient services
10.5 Crisis attendance due to mental health problems

11. Adverse events

11.1 Suicide attempts
11.2 Death (all causes)
11.3 Adverse effects - any

12. 'Summary of findings' table

We used the GRADE approach to interpret findings (Schünemann 2008) and used GRADE profiler (GRADEPRO) to import data from RevMan 5 (Review Manager) to create 'Summary of findings' tables. These tables provided outcome-specific information concerning the overall quality of evidence from each included study in the comparison, the magnitude of effect of the interventions examined, and the sum of available data on all outcomes we rated as important to patient-care and decision making. We selected the following main outcomes for inclusion in the 'Summary of findings' table.
1. Compliance
2. Mental state
3. Functioning (Cognitive and Social)
4. Quality of life
5. Satisfaction
6. Acceptability of intervention

Search methods for identification of studies

Electronic searches

1. Cochrane Schizophrenia Group Trials Register (May 2011, November 2011, July 2012, September 2013)

We searched the register using the phrase:
[(*virtual* OR *VR* OR *second life* OR *facebook* OR *Twitter* OR *3rd generation* OR *third generation* OR *video* OR *hypermedia* OR *Computer* in title, abstract, index terms of REFERENCE)]

This register is compiled by systematic searches of major databases, handsearches and conference proceedings (see group module).

2. PubMED database search (November 2011)

We searched 2nd of November 2011, PubMED database using the phrase: (virtual reality and schizophrenia).

Searching other resources

1. Reference searching

We inspected references of all identified studies for further relevant studies.

2. Personal contact

We contacted the first author of each included study for information regarding unpublished trials.

Data collection and analysis

Selection of studies

Review authors MV, HH, ML, MK, AH, KM and TR independently inspected citations from the searches and identified relevant abstracts. Where disputes arose, the full report was acquired for more detailed scrutiny. Full reports of the abstracts meeting the review criteria were obtained and inspected by MV, HH, ML, MK, AH, KM and TR. If it had not been possible to resolve disagreement by discussion, we would have attempted to contact the authors of the study for clarification. We also aimed to independently re-inspect a random sample of 20%. This was not done due to the low number of included studies.

Data extraction and management

1. Extraction

Review authors MV, HH, ML, MK, AH, KM and TR extracted data from all included studies. Any disagreements were discussed, decisions documented and, if necessary, we contacted the authors of studies for clarification. With remaining problems CEA helped clarify issues and these final decisions were documented. Data presented only in graphs and figures would have been extracted, but included only if two review authors independently had the same result. We contacted the authors of the studies through an open-ended request in order to obtain missing information and clarification. If studies were been multi-centre, where possible, we would have extracted data relevant to each component centre separately.

2. Management

2.1 Forms

We extracted data onto standard, simple forms.

2.2 Scale-derived data

We included continuous data from rating scales only if:
a. the psychometric properties of the measuring instrument have been described in a peer-reviewed journal (Marshall 2000); however, if unreferred measures were used due to limited numbers of validated measures, this has been identified in analysis table; and
b. the measuring instrument has not been written or modified by one of the trialists for that particular trial.

Ideally, the measuring instrument should either be i. a self-report or ii. completed by an independent rater or relative (not the therapist). We realised that this is not often reported clearly, therefore we noted this in 'Description of studies'.

2.3 Endpoint versus change data

There are advantages of both endpoint and change data. Change data can remove a component of between-person variability from the analysis. On the other hand, calculation of change needs two assessments (baseline and endpoint), which can be difficult in unstable and difficult to measure conditions such as schizophrenia. We decided primarily to use endpoint data, and only used change data if the former were not available. Endpoint and change data were combined in the analysis as we used mean differences (MD) rather than standardised mean differences throughout (Higgins 2011).

2.4 Skewed data

Continuous data on clinical and social outcomes are often not normally distributed. To avoid the pitfall of applying parametric tests to non-parametric data, we applied the following standards to all data before inclusion:
a) standard deviations (SDs) and means are reported in the paper or obtainable from the authors;
b) when a scale starts from the finite number zero, the SD, when multiplied by two, is less than the mean (as otherwise the mean is unlikely to be an appropriate measure of the centre of the distribution, (Altman 1996);
c) if a scale starts from a positive value (such as the Positive and Negative Syndrome Scale (PANSS) (Kay 1986), which can have values from 30 to 210), the calculation described above should be modified to take the scale starting point into account. In these cases skew is present if 2 SD > (S-S min), where S is the mean score and S min is the minimum score.
Endpoint scores on scales often have a finite start and end point and these rules can be applied. If we had found skewed endpoint data from studies of less than 200 participants, we would have entered these data as 'other' data within the data and analyses section rather than into a statistical analysis. Skewed data pose less of a problem when looking at means if the sample size is large and we would have entered such data into statistical syntheses.

When continuous data are presented on a scale that includes a possibility of negative values (such as change data), it is difficult to tell whether data are skewed or not. Change data from large and small trials were entered into analyses.

2.5 Common measure

To facilitate comparison between trials, we would have converted variables that can be reported in different metrics, such as days in hospital (mean days per year, per week or per month) to a common metric (e.g. mean days per month).

2.6 Conversion of continuous to binary

Where possible, efforts were made to convert outcome measures to dichotomous data. This can be done by identifying cut-off points on rating scales and dividing participants accordingly into 'clinically improved' or 'not clinically improved'. It is generally assumed that if there is a 50% reduction in a scale-derived score such as the Brief Psychiatric Rating Scale (BPRS, Overall 1962) or the PANSS (Kay 1986), this could be considered as a clinically significant response (Leucht 2005; Leucht 2005a). If data based on these thresholds were not available, we would have used the primary cut-off presented by the original authors.

2.7 Direction of graphs

Where possible, we entered data in such a way that the area to the left of the line of no effect indicates a favourable outcome for virtual reality. Where keeping to this makes it impossible to avoid outcome titles with clumsy double-negatives (e.g. 'Not un-improved') we reported data where the left of the line indicates an unfavourable outcome. This was noted in the relevant graphs.

Assessment of risk of bias in included studies

REview authors MV, HH, ML, AH, MK, KM and TR worked independently to assess risk of bias by using criteria described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011) to assess trial quality. This set of criteria is based on evidence of associations between overestimate of effect and high risk of bias of the article such as sequence generation, allocation concealment, blinding, incomplete outcome data and selective reporting.

If the raters disagreed, the final rating was made by consensus, with the involvement of another member of the review group. Where inadequate details of randomisation and other characteristics of trials were provided, we contacted the authors of the studies in order to obtain further information. Non-concurrence in quality assessment was reported. When disputes arose as to which category a trial was to be allocated, resolution was made by discussion. The level of risk of bias was noted in both the text of the review and in the Summary of findings for the main comparison.

Measures of treatment effect

1. Binary data

For binary outcomes, instead of calculating risk ratio (RR) as stated in our protocol, we calculated a standard estimation of the risk difference (RD) and its 95% confidence interval (CI) (see Differences between protocol and review). This was because the RD can be calculated for any study, even when there are no events in either group. It has been shown that RD is more intuitive (Boissel 1999) than odds ratios and that odds ratios tend to be interpreted as RD by clinicians (Deeks 2000). For statistically significant results we initially intended to calculate the number needed to treat to provide benefit /to induce harm statistic (NNTB/H), and its 95% CI using Visual Rx (http://www.nntonline.net/) taking account of the event rate in the control group. This, however, has been superseded by Summary of findings for the main comparison and calculations therein.

2. Continuous data

For continuous outcomes, mean difference (MD) between groups were estimated. We preferred not to calculate effect size measures (standardised mean difference SMD). However, if scales of very considerable similarity had been used, we would have presumed there was a small difference in measurement, and we therefore would have calculated effect size and transformed the effect back to the units of one or more of the specific instruments.

Unit of analysis issues

1. Cluster trials

Studies increasingly employ 'cluster randomisation' (such as randomisation by clinician or practice) but analysis and pooling of clustered data poses problems. Firstly, authors often fail to account for intra-class correlation in clustered studies, leading to a 'unit of analysis' error (Divine 1992) whereby P values are spuriously low, confidence intervals unduly narrow and statistical significance overestimated. This causes type I errors (Bland 1997; Gulliford 1999).

We did not include any cluster trials. If cluster trials had been included, we would have presented their data in a table, with a (*) symbol to indicate the presence of a probable unit of analysis error, if clustering had not been accounted for in primary studies. In subsequent versions of this review, if cluster trials are included, if necessary, we will contact the first authors of studies to obtain intra-class correlation coefficients (ICCs) for their clustered data and adjust for this by using accepted methods (Gulliford 1999). In cases where clustering is incorporated into the analysis of primary studies, we will present these data as if from a non-cluster randomised study, but adjust for the clustering effect.

We have sought statistical advice and have been advised that the binary data as presented in a report should be divided by a 'design effect'. This is calculated using the mean number of participants per cluster (m) and the ICC [Design effect = 1+(m-1)*ICC] (Donner 2002). If the ICC is not reported it will be assumed to be 0.1 (Ukoumunne 1999). If cluster studies have been appropriately analysed taking into account ICCs and relevant data documented in the report, synthesis with other studies will be possible using the generic inverse variance technique.

2. Cross-over trials

We did not include cross-over trials in this version of our review. A major concern of cross-over trials is the carry-over effect. It occurs if an effect (e.g. pharmacological, physiological or psychological) of the treatment in the first phase is carried over to the second phase. As a consequence on entry to the second phase the participants can differ systematically from their initial state despite a wash-out phase. For the same reason cross-over trials are not appropriate if the condition of interest is unstable (Elbourne 2002). As both effects are very likely in severe mental illness, we would have only used data from the first phase of cross-over studies.

3. Studies with multiple treatment group

Where a study involved more than two treatment arms, if relevant, we would have presented the additional treatment arms in comparisons. If data were binary we would have simply added and combined within the two-by-two table. If data were continuous we would have combined data following the formula in section 7.7.3.8 (Combining groups) of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011). Where the additional treatment arms were not relevant, these data would not have reproduced. Our included studies did not have more than two treatment arms.

Dealing with missing data

1. Overall loss of credibility

At some degree of loss of follow-up, data must lose credibility (Xia 2009). We chose that, for any particular outcome, should more than 50% of data be unaccounted for, we did not reproduce these data or use them within analyses, (except for the outcome 'leaving the study early'). If, however, more than 50% of those in one arm of a study were lost, but the total loss was less than 50%, we would have marked such data with (*) to indicate that such a result may well be prone to bias.

2. Binary

In the case where attrition for a binary outcome was between 0% and 50% and where these data were not clearly described, we presented data on a 'once-randomised-always-analyse' basis (an intention-to-treat (ITT) analysis). Those leaving the study early were all assumed to have the same rates of negative outcome as those who completed, with the exception of the outcome of death and adverse effects. For these outcomes the rate of those who stayed in the study - in that particular arm of the trial - were used for those who did not. A sensitivity analysis was undertaken to test how prone the primary outcomes were to change when 'completer' data only were compared to the ITT analysis using the above assumptions.

3. Continuous

3.1 Attrition

In the case where attrition for a continuous outcome was between 0% and 50% and completer-only data were reported, we presented and used these data.

3.2 Standard deviations

If standard deviations (SD) had not been reported, we first would have tried to obtain the missing values from the authors. If not available, where there were missing measures of variance for continuous data, but an exact standard error (SE) and confidence intervals were available for group means, and either 'P' value or 't' value available for differences in mean, we would have calculated them according to the rules described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011). When only the SE is reported, SDs are calculated by the formula $SD = SE * \sqrt{n}$. Chapters 7.7.3 and 16.1.3 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011) present detailed formulae for estimating SDs from P values, t or F values, confidence intervals, ranges or other statistics. If these formulae did not apply, we would have calculated the SDs according to a validated imputation method, which is based on the SDs of the other included studies (Furukawa 2006). Although some of these imputation strategies can introduce error, the alternative would have been to exclude a given study's outcome and thus to lose information. We would have examined the validity of the imputations in a sensitivity analysis excluding imputed values.

3.3 Last observation carried forward

We anticipated that in some studies the method of last observation carried forward (LOCF) would be employed within the study report. As with all methods of imputation to deal with missing data, LOCF introduces uncertainty about the reliability of the results (Leucht 2007). Therefore, if LOCF data had been used in

the trial, and less than 50% of the data had been assumed, we would have reproduced these data and indicated that they were the product of LOCF assumptions.

Assessment of heterogeneity

1. Clinical heterogeneity

We considered all included studies initially, without seeing comparison data, to judge clinical heterogeneity. We simply inspected all studies for clearly outlying people or situations, which we had not predicted would arise. When such situations or participant groups arose, these were fully discussed.

2. Methodological heterogeneity

We considered all included studies initially, without seeing comparison data, to judge methodological heterogeneity. We simply inspected all studies for clearly outlying methods, which we had not predicted would arise. When such methodological outliers arose these were fully discussed.

3. Statistical heterogeneity

3.1 Visual inspection

We visually inspected graphs to investigate the possibility of statistical heterogeneity.

3.2 Employing the I^2 statistic

Heterogeneity between studies was investigated by considering the I^2 method alongside the Chi^2 'P' value. The I^2 provides an estimate of the percentage of inconsistency thought to be due to chance (Higgins 2003). The importance of the observed value of I^2 depends on i. magnitude and direction of effects and ii. strength of evidence for heterogeneity (e.g. 'P' value from Chi^2 test, or a confidence interval for I^2). An I^2 estimate greater than or equal to around 50% accompanied by a statistically significant Chi^2 statistic, would have been interpreted as evidence of substantial levels of heterogeneity (Section 9.5.2 - Higgins 2011). When substantial levels of heterogeneity were found in the primary outcome, we explored reasons for the heterogeneity (Subgroup analysis and investigation of heterogeneity).

Assessment of reporting biases

Reporting biases arise when the dissemination of research findings is influenced by the nature and direction of results (Egger 1997). These are described in Section 10 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011). We are aware that funnel plots may be useful in investigating reporting biases

but are of limited power to detect small-study effects. We did not use funnel plots for outcomes because there were less than 10 included studies.

Data synthesis

We understand that there is no closed argument for preference for use of fixed-effect or random-effects models. The random-effects method incorporates an assumption that the different studies are estimating different, yet related, intervention effects. This often seems to be true to us and the random-effects model takes into account differences between studies even if there is no statistically significant heterogeneity. There is, however, a disadvantage to the random-effects model. It puts added weight onto small studies which often are the most biased ones. Depending on the direction of effect, these studies can either inflate or deflate the effect size. We chose the fixed-effect model for all analyses. The reader is, however, able to choose to inspect the data using the random-effects model.

Subgroup analysis and investigation of heterogeneity

1. Subgroup analyses - only primary outcomes

1.1 Clinical state, stage or problem

We proposed to undertake this review to provide an overview of the effects of virtual reality for people with schizophrenia in general. We also tried to report data on subgroups of people in the same clinical state, stage, with similar problems and settings. However, this was not possible due to heterogeneity of the studies.

2. Investigation of heterogeneity

We reported high inconsistency in studies. First, we investigated whether data had been entered correctly. Second, if this was the case, the graph was visually inspected and outlying studies were successively removed to see if homogeneity was restored. For this review we decided that should this occur with data contributing to the summary finding of no more than around 10% of the total weighting, data would be presented. However, if not, the data were not pooled and issues were discussed. We know of no supporting research for this 10% cut off but are investigating use of prediction intervals as an alternative to this unsatisfactory state.

When unanticipated clinical or methodological heterogeneity were obvious we stated hypotheses regarding these for future reviews or versions of this review. We did not anticipate undertaking analyses relating to these.

Sensitivity analysis

1. Implication of randomisation

Our aim was to include trials in a sensitivity analysis if they were described in some way as to imply randomisation. For the primary outcomes we included these studies and if there was no substantive difference when the implied randomised studies were added to those with better description of randomisation, then all data were employed from these studies.

2. Assumptions for lost binary data

Where assumptions had to be made regarding people lost to follow-up (see Dealing with missing data), we compared the findings of the primary outcomes when we used our assumption compared with completer data only. If there was a substantial difference, we reported results and discussed them but continued to employ our assumption.

Where assumptions had to be made regarding missing SDs data (see Dealing with missing data), we compared the findings on primary outcomes when we used our assumption compared with complete data only. A sensitivity analysis was undertaken to test how prone results were to change when 'completer' data only were compared to the imputed data using the above assumption. If there was a substantial difference, we reported results and discussed them but continued to employ our assumption.

3. Risk of bias

We analysed the effects of excluding trials that were judged to be at unclear risk of bias across one or more of the domains of randomisation (implied as randomised with no further details available) allocation concealment, blinding and outcome reporting for the meta-analysis of the primary outcome. If the exclusion of trials at unclear risk of bias did not substantially altered the direction of effect or the precision of the effect estimates, then data from these trials were included in the analysis

4. Imputed values

If cluster trials had been included, we would also have undertaken a sensitivity analysis to assess the effects of including data from trials where we had used imputed values for ICC in calculating the design effect in cluster randomised trials.

If substantial differences were present, we would have noted the direction or precision of effect estimates in any of the sensitivity analyses listed above. We would not have pooled data from the excluded trials with the other trials contributing to the outcome, but would have presented them separately.

5. Fixed and random effects

All data was synthesised using a fixed-effect model, however, we also synthesised data for the primary outcome using a random-effects model to evaluate whether this altered the significance of the results. No significant differences between these two models were found.

R E S U L T S

Description of studies

Results of the search

Despite extensive initial searches in May 2011, we found only 33 references, from which only one was considered relevant (Figure 1). An additional search in November 2011 identified one more study (Figure 2). A further search in July 2012 did not identify any new studies suitable for inclusion in the analysis (Figure 3). Searches undertaken in September 2013 identified two further relevant studies, one to be included, and one an ongoing study (Figure 4). In total, three studies were included in the analysis.

Figure 1. Study flow diagram - Search of May 2011.

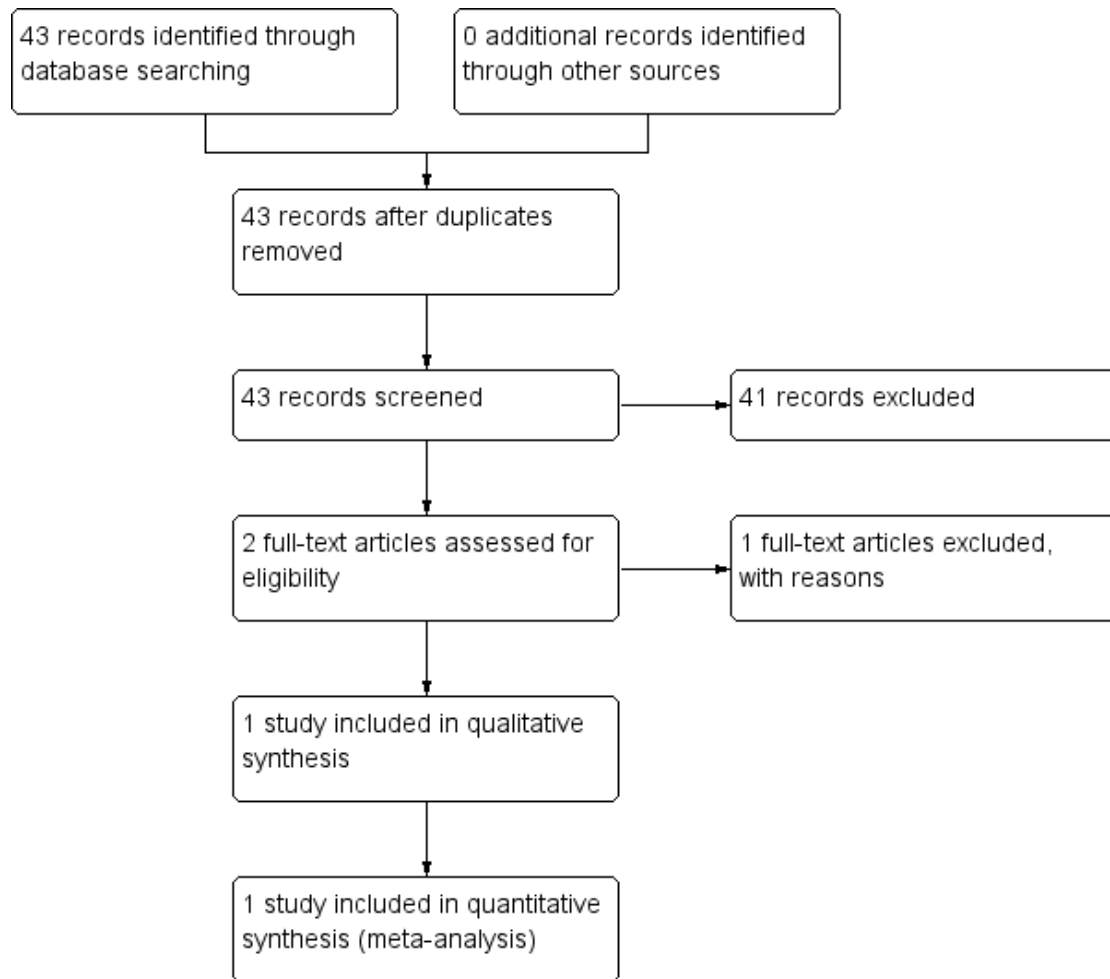**Figure 2. Study flow diagram - search of November 2011.**

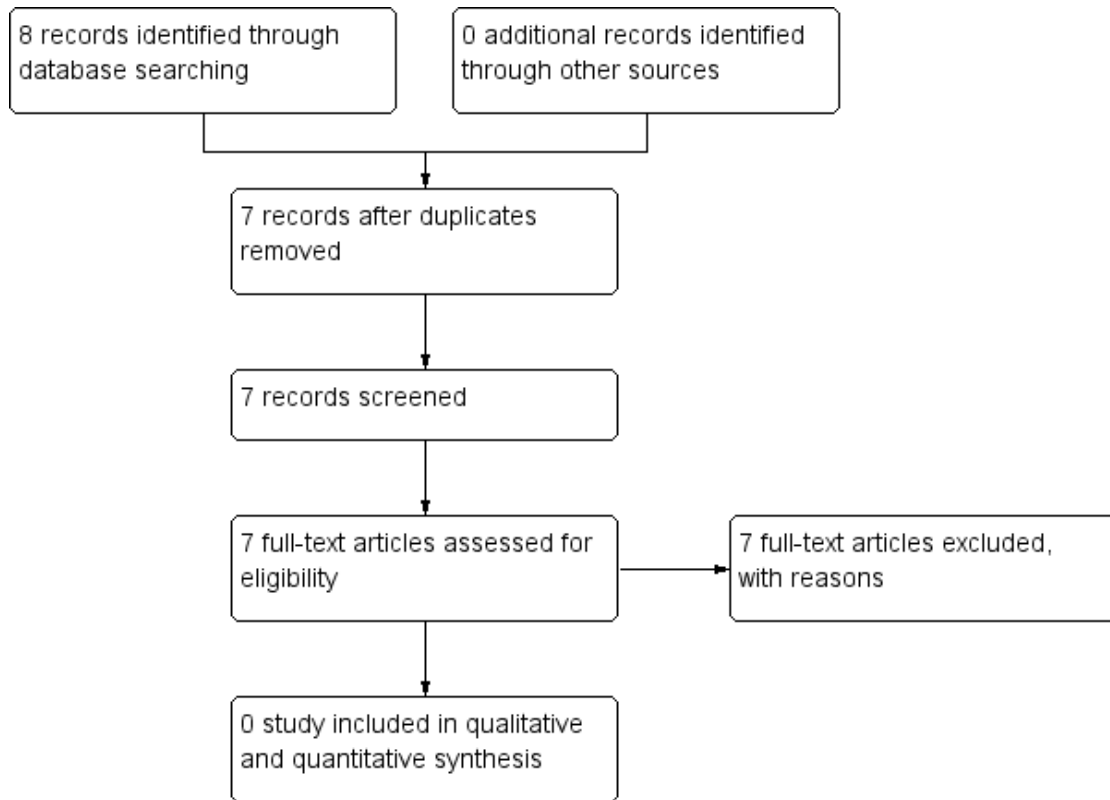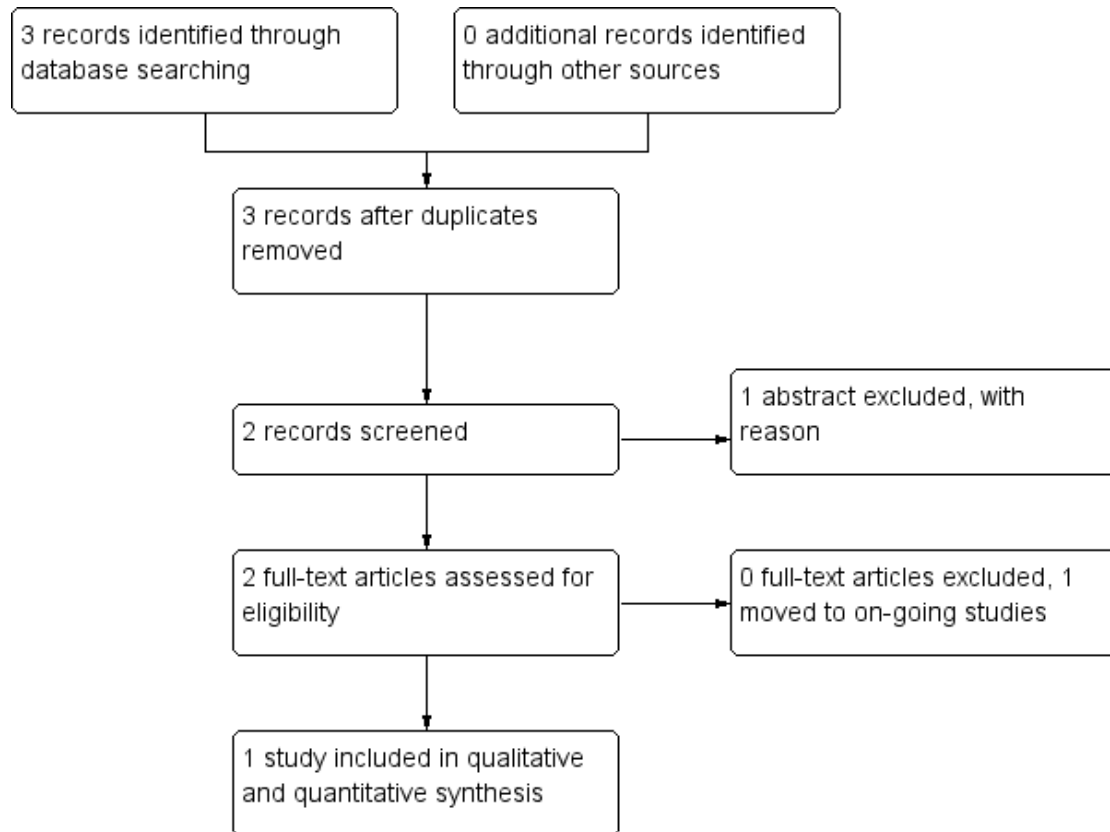**Figure 3.   Study flow diagram - Search of July 2012**

Figure 4. Study flow diagram - Search of December 2013

Included studies

For an overview of included studies please see Characteristics of included studies.

1. Methods

All the included studies (Chan 2010; Park 2011; Tsang 2013) were stated to be randomised, although, the description of allocation varied. For further details please see Allocation (selection bias) below. In one study (Tsang 2013) the assessors were blinded. The other two studies (Chan 2010; Park 2011) did not use blinding nor did they provide information related to attempts to blind. Overall, the length of the trials varied from five weeks (Park 2011; Tsang 2013) to three months (Chan 2010). However, these follow-up times are our assumptions as the reports did not provide clear information about length of follow-up. In Park Park 2011 we assumed that follow-up was five weeks as the intervention was this length of time. The same assumption applied to Chan 2010, but in this case it was three months. All included studies used a parallel study design and were located in one research centre.

2. Participants

All studies included people with a diagnosis of schizophrenia or related disorder. Diagnoses were undertaken using the Structural Interview for DSM-IV Axis I (APA 2000). In total, there were 156 participants who were either adults or older adults. In Chan 2010, the population was a little older (mean age ~ 66 years). In Park 2011 and Tsang 2013 people were younger adults aged between 18 and 45. Altogether there were 86 men and 81 women. About two thirds of the participants in Chan 2010 had only primary education, while in Park 2011 the average length of education was about 13 years.

The setting of the included studies varied from the inpatient hospital setting to long-stay care unit. In Chan 2010 participants were recruited from a long-stay care setting that provided long-term residential care and support services to people who were functionally impaired and unable to live independently in the commu-

nity. In Tsang 2013 participants were inpatients attending a vocational rehabilitation programme at a psychiatric hospital. Conversely, participants in Park 2011 were all inpatients in one mental health hospital who were treated with intensive psychiatric care for two to four weeks. In Park 2011, the average age at illness onset was 24 years, the average duration of illness six years, and participants had less than three previous psychiatric admissions. Chan 2010 evaluated participants' mental status using the Mini Mental State Examination (MMSE) (mean score ~ 18), while Park 2011 measured participants' symptoms and their Positive and Negative Symptoms Scale (PANNS) score (mean score ~ 72); Tsang 2013 used the Brief Psychiatric Rating Scale (BPRS) (mean sore ~ 21 SD 6).

Park 2011 excluded people who abused substances, those with head trauma, neurological illness or physical illness that could affect brain functioning as well as other Axis I diagnosis (APA 2000). Tsang 2013 excluded people with physical handicaps, ECT treatment during the previous 12 months, drug abuse during previous 30 days, or a history of mental retardation. The Chan 2010 exclusion criteria were reported indirectly. They included participants only if they had no functional restriction in upper extremities, no specific concomitant disease that would have compromised cognitive functions (e.g. mental retardation, dementia, or brain injury).

3. Intervention

All three included studies used virtual reality (VR) programmes to deliver skills training sessions.

The VR program used in Chan 2010 was a two dimensional VR programme that enables a person to engage in a series of simulated tasks within a VR environment through video contact. The system comprises more than 20 programmed VR activities, which can be adjusted to the level of difficulty in speed, directionality, and/or number of distracters. Users wearing red gloves stand or sit in a demarcated area, viewing a large screen that displays one of the series of simulated tasks, such as catching a virtual ball or playing goalie in a virtual soccer game. A single-camera, vision-based tracking system is placed in front of the demarcated area and captures the real-time image and movement of the person for processing. The person's video-captured image is processed by the system on the same plane as the screen graphical animations that react in real time in response to his or her movement. The person can control his or her movements within the virtual environments and can interact with graphic objects as depicted in this environment. The VR programme consists of two VR activities: ball and bird, and shark bait using IREX. These activities were selected because they tapped into fluid intelligence.

Park 2011 included a personal computer for rendering and providing the virtual environment, a head mounted display for displaying the virtual environment in a more immersive manner, and a position tracker for following the head direction in real time. The participants were able to move their heads to direct their gaze

in a natural manner, and the display of the virtual environment depended on the orientation data obtained from the participants' head direction. VR role-plays were displayed through two different panels: an HMD and a 120 inch screen. Social skill training using virtual reality role-playing (SST-VR) included core features of role-playing games. For example, the participant was provided with a joystick and buttons to operate his/her avatar, which produced the first-person perspective view. By using the joystick and buttons, he/she freely moved and interacted with avatars in the virtual space.

In the Tsang 2013 study, participants attended a VR-based vocational skills training in a virtual boutique scenario (VRVTS). Before training, the participants were briefed on the training procedures. They were required to attend 10 sessions of training (30 minutes each) with a specific topics (e.g. identifying clothes, checking stock, sorting clothes, checking clothes). To ensure better adaption and to observe any cases of cyber sickness, the participants were allowed to browse the VR scenario for five to 10 minutes during the first session.

Standard care used as a control group varied in the included studies. First, in Chan 2010 study, participants in the control group attended the usual program in the facility and were to receive the VR programme three months later. Park 2011 used verbal, writing, picture, and video supplies as simulators of the scenes and social skills training (SST) therapists as the actors were used in TR role-plays. Finally, Tsang 2013 used work-simulated workshops in the occupational therapy department including packaging tasks, typing, and cleansing tasks. Participants also attended at least three hours of prevocational skills training in every working day during hospitalisation.

4. Outcomes

The duration of follow-up in the included trials varied between five weeks to three months. Details of the scales used in this review to quantify different outcomes are provided below. Data from these scales were reported either in continuous form or as binary figures.

4.1 Binary data

Compliance was reported as loss to follow-up and being withdrawn by the trialist. Acceptability of the intervention was reported as leaving the studies early due to any reason.

4.2 Scale-derived outcomes

4.2.1 Functioning: Cognistat

Cognistat - The Neurobehavioral Cognitive Status Examination (Eisenstein 2002; Johansson 2012) is a brief neuropsychological screening test designed to assess seven major areas of cognitive

functioning: orientation, attention, calculations, constructions, memory, language, reasoning (Chan 2010). Cognistat has been developed for psychiatric and psychogeriatric patients (Eisenstein 2002; Johansson 2012).

4.2.2 Skills

a. Assertiveness (RAS) is a six-point Likert scale. Scale is including 30 items assessing assertiveness (Park 2011; Jenerette 2010). The RAS is widely used as both an index of general assertiveness and as a means for evaluating the effectiveness of assertiveness training (Jenerette 2010).

b. Relationship Change Scale (RCS) was initially developed as a 27-item questionnaire (Park 2011). Questionnaire measures a person's perception of changes in a relationship with a significant other with regard to satisfaction, communication, trust, sensitivity, openness, and understanding (Coleman 2005).

c. Social Behaviour Scale (SBS) was including 29-items related to voice, non verbal and conversational skills (Park 2011). The scale consisted six items on voice quality, nine items on nonverbal skills and 14 on conversational properties (Park 2011).

d. Social Problem Solving Inventory - Revised (SPSI-R) measures individuals cognitive, affective, or behavioural responses to real life problem solving situations (Park 2011). The scale consists of a 25-item self-report measure and it was five-point Likert scale (Park 2011, Wakeling 2007).

4.2.3 Satisfaction with treatment

a. Interest-in-Participation assesses interest in social skill training. This scale includes two items assessing participants' interest in the current session and their expectation for the next session on a scale of one to five. Average score of the two items was used as a proxy measure of motivation (Park 2011).

b. Generalisation of the skill was assessed by using initially a developed scale contributing four STT constituents (material, content, therapist and structure) to overall improvement by using five-point Likert scale. Higher score reflects greater contributions, meaning that the participant more efficiently applied the learned skills into specific social knowledge. The questionnaire also asked participants to identify which session was most helpful (Park 2011).

4.3 Missing outcomes

One trial reported outcomes related to functioning by using Brief Neuropsychological Cognitive Examination (BNCE). These data were not able to be used because of missing means and SD values (Tsang 2013). Two trials reported outcomes related to behaviour. One used Volitional Questionnaire (VQ) and adverse effects by using the Simulator Sickness Questionnaire (SSQ) (Chan 2010). Data on VQ, however, were reported only for the intervention group and data on SSQ were impossible to use. Another trial reported work perspective by using Vocational Cognitive Rating Scale (VCRS), but these data were not able to be used because of missing means and SD values (Tsang 2013). There were no data at all on mental state, insight, behaviour, quality of life, costs, service utilisation, or adverse effects.

Excluded studies

We excluded 37 studies (full-text). Twenty-seven were excluded because the intervention was not virtual reality. Park 2009 did use virtual reality (VR) in their study, but it was excluded because the purpose of the trial was to examine the effects of two drugs, not the effects of the VR itself; both groups in Park 2009 used VR.

Awaiting classification

No studies await classification.

Ongoing

There is one ongoing study (UKCRNID12951 2012).

Risk of bias in included studies

Our overall impression of risk of bias in the included studies is represented in Figure 5. There is, at the very least, a moderate risk of bias in all outcomes and therefore a risk of overestimating any positive effects of VR for people with serious mental illness.

Figure 5. 'Risk of bias' summary: review authors' judgements about each risk of bias item for each included study.

Allocation

All included studies were stated to be randomised. In two studies (Chan 2010; Park 2011) there were no more details available. We therefore classified both trials as of 'unclear' quality with a moderate risk of selection bias and of overestimate of positive effect.

Blinding

Chan 2010 stated that blinding was not used and risk of observer bias was rated as high. Park 2011 did not report whether blinding had been used. We therefore rated the risk of observer bias as 'unclear'. This gathers further potential for overestimate of positive effects and underestimate of negative ones. Tsang 2013 reported that assessors were blinded to group assignment.

Incomplete outcome data

Chan 2010 excluded the information of two participants to try and eliminate the effect of alteration in psychotropic medication. We felt we had to rate this as high risk of overestimating positive results. Park 2011 described participant flow but there were no data on those who left early. We therefore classed this aspect of the study as of unclear risk.

Selective reporting

We did not have any protocols for the included studies. All included in this version of the review seemed to report all outcomes measured.

Other potential sources of bias

All studies were small and publication bias was very possible. Chan 2010 rated Simulator Sickness Questionnaire (SSQ) and Volitional Questionnaire (VQ) only from those in the intervention group. We are not quite sure why. There were no other obvious potential sources of bias.

Effects of interventions

See: **Summary of findings for the main comparison** VIRTUAL REALITY AS A SOLE METHOD + STANDARD PROFESSIONAL CARE compared with STANDARD PROFESSIONAL CARE for treatment compliance for people with serious mental illness
For dichotomous data we calculated the risk difference (RD) and 95% confidence intervals (CI). For continuous data we calculated mean differences (MD), again with 95% CIs.

1. VIRTUAL REALITY AS A SOLE METHOD + STANDARD CARE vs. STANDARD CARE

1.1 Compliance

All trials (n = 156) reported no difference between groups regarding compliance in terms of loss to follow-up (RD 0.02 CI -0.08 to 0.12). The one relevant study (n = 29) found that no one left of their own volition but two people were removed by the trialists (RD 0.14 CI -0.06 to 0.35, Analysis 1.1) because they had swapped medication and it was felt that this could distort the results of the virtual reality.

1.2 Functioning

1.2.1 Cognitive

One study (Chan 2010) (n = 27), using Cognistat, found no difference between interventions for an overall endpoint score (MD 4.67 CI -1.76 to 11.10, Analysis 1.2).

1.2.2 Social

One study (Park 2011) (n = 64) reported the data on average endpoint change scores in specific skills (assertiveness, relationship change, social behaviour, social problem solving). No data were statistically significant, or suggested an effect (1 RCT, n = 64, MD average score on social problem solving SPSI-R -2.30 CI -8.13 to 3.53) (Analysis 1.3). This group also undertook a series of sub-analyses of this scale but it was not clear if these were *post hoc* and therefore prone to spurious significant findings. We are also unsure whether these sub-scores are validated in themselves and have therefore reported them only in Table 1.

1.3 Satisfaction with treatment

One study (Park 2011) (n = 64) found a significant difference between groups regarding the interest in social skills training (MD 6.00 CI 1.39 to 10.61), and generalisation of the skill in terms of applying the learned skills into specific social knowledge (MD 5.10 CI 1.03 to 9.17) Analysis 1.4).

1.4 Acceptability of intervention

Two studies (Chan 2010; Tsang 2013) (n = 92) reported the acceptability of the intervention in terms of leaving the study early for any reason (RD 0.05 CI -0.09 to 0.19, Analysis 1.5).

DISCUSSION

For an overall summary of our findings - please see Summary of findings for the main comparison.

Summary of main results

Comparison 1. VIRTUAL REALITY AS A SOLE METHOD + STANDARD CARE versus STANDARD PROFESSIONAL CARE

About 12% of people were lost to follow-up by around five to 12 weeks. This was a primary outcome of this review and we cannot be quite sure even of this result as in one study two people were removed by the trialists because they had swapped medication and it was felt that this could distort the results of the virtual reality (VR). This illustrates how all data are based on small pioneering studies that are methodologically problematic and confronting novel problems. Larger trials overcoming these problems are needed. There is really no discernable effect of the intervention for cognitive testing or measures of social function, or acceptability of treatment. Neither was there really any suggestion of an effect. The finding for satisfaction is neither direct nor strong. It is possible that the VR method really does have no effect.

We have no data on mental state, insight, behaviour, quality of life, costs, service utilisation, or adverse effect. The results of this review provide both very limited information, but also limited hope regarding the effects of VR to support treatment compliance for people with severe mental illness.

Overall completeness and applicability of evidence

Virtual reality is new approach in mental health care and the studies included in this review can be considered pioneering. These data are limited, patchy and difficult to apply anywhere but most specialised centres. We assume, however, that interventions using new technologies such as VR will be developed and we expect studies to be broader ranging, and using technology that becomes ever-more accessible.

Quality of the evidence

Overall, the quality of reporting varied (Figure 5). Out of three studies, two did not clearly describe how the randomisation was conducted. None of the studies described allocation adequately. Blinding of participants and personnel was not described clearly in any study, while blinding of outcome assessment was also inadequately described in two studies. Further, one study was considered poor quality regarding incomplete outcome data and it

reported results for intervention group only. There is a moderate risk of overestimating the estimate of effect and therefore the results of this review need to be considered as high risk of overestimating positive results. In addition, there were concerns related to indirectness in terms of no direct measure used, unclear clinical meaning of scores or unreferred scales and how leaving the study early was conducted. Therefore, even if the trials had been well-powered, we would have had problems interpreting the findings because of this limited quality.

Potential biases in the review process

Virtual reality is a rather new approach. Therefore, the search might not have identified all relevant studies. Such studies are possible to undertake for work only published as dissertations and these can be difficult to identify. Also, we do have an interest in this area and could, feasibly, have been biased in our selection and data extraction. We, however, have tried to be as transparent as possible about this and leave the reader to decide.

Agreements and disagreements with other studies or reviews

We know of no other systematic reviews in the area of VR for treatment compliance for people with serious mental illness. There is, however, one Cochrane review conducted among people with stroke, which found limited evidence that use of VR and interactive video gaming may be beneficial in improving arm function and activities of daily living when compared with conventional therapy (Laver 2011).

AUTHORS' CONCLUSIONS

Implications for practice

1. For people with schizophrenia

There is no clear evidence for or against using virtual reality for treatment compliance among people with serious mental illness. The intervention is purely experimental for people with schizophrenia. Virtual reality has been used with other patient groups, such as people with stroke, and has been shown to be promising. There is a need to gather more information on its effects for this particular client group.

2. For clinicians

This review was unable to provide sufficient evidence to inform clinicians about the value of virtual reality to support treatment compliance for people with serious mental illness. If virtual reality is used, the experimental nature of the intervention should be clearly explained to the patient.

3. For policymakers

There is no evidence for or against using virtual reality for treatment compliance among people with serious mental illness. Moreover, no data exist about the economic consequences of using virtual reality in clinical practice. More high-quality studies should be undertaken in this area to explore the costs of this novel intervention and variations of approach.

Implications for research

1. General

Given that there is insufficient research to say whether virtual reality is effective or not, there is a need for further research to establish its value - or lack of it - in this population. In this fast moving area it is likely that additional studies are currently being undertaken. All should report to the standards required by CONSORT.

2. Specific

2.1 Other reviews

Excluded studies suggest that there are potentially several more reviews in this broad area covering a range of treatment options (Table 2).

2.2 Other trials

We do feel that the story for virtual reality on compliance is not complete and that larger studies are needed. We do realise that designing such studies needs great care and attention to detail and that simply reviewing past studies is only part of that process. However, we have given considerable thought to the existing studies and do suggest an outline for a relevant trial (Table 3).

REFERENCES

References to studies included in this review

Chan 2010 *{published data only}*
> Chan CLF, Ngai EKY, Leung PKH, Wong S. Effect of the adapted virtual reality cognitive training program among Chinese older adults with chronic schizophrenia: A pilot study. *International Journal of Geriatric Psychiatry* 2010;**25** (6):643–9.

Park 2011 *{published data only}*
> Park KM, Ku J, Choi SH, Jang HJ, Park JY, Kim SI, et al. A virtual reality application in role-plays of social skills training for schizophrenia: a randomized, controlled trial. *Psychiatry Research* 2011;**189**(2):166–72. [PUBMED: 21529970]

Tsang 2013 *{published data only}*
> Tsang MM, Man DW. A virtual reality-based vocational training system (VRVTS) for people with schizophrenia in vocational rehabilitation. *Schizophrenia Research* 2013;**144** (1-3):51–62.

References to studies excluded from this review

Benedict 1994 *{published data only}*
> Benedict RH, Harris AE, Markow T, McCormick JA, Nuechterlein KH, Asarnow RF. Effects of attention training on information processing in schizophrenia. *Schizophrenia Bulletin* 1994;**20**(3):537–46. [MEDLINE: 7973469]

Bryson 2003 *{published data only}*
> Bryson G, Bell MD, Wexler B. Cognitive rehabilitation in schizophrenia: a comparison of two methods. *Schizophrenia Research* 2003;**60**:320.

Chinman 2004 *{published data only}*
> Chinman M, Young AS, Schell T, Hassell J, Mintz J. Computer-assisted self-assessment in persons with severe mental illness. *Journal of Clinical Psychiatry* 2004;**65**(10): 1343–51. [MEDLINE: 15491237]

David 2012 *{published data only}*
> David AS, Ster IC, Zavarei H. Effect of video self-observations vs. observations of others on insight in psychotic disorders. *Journal of Nervous and Mental*

Disease 2012;**200**(4):358–61. [MEDLINE: BIOSIS: PREV201200322149]

Dunn 2001 *{published data only}*
Dunn LB, Lindamer LA, Schneiderman LJ, Jeste DV. A novel, computer-based enhancement of informed consent in older patients with psychotic disorders. Proceedings of the 14th Annual Meeting of the American Association for Geriatric Psychiatry; 2001 Feb 23-26; San Francisco, California, USA. 2001.

Field 1997 *{published data only}*
Field CD, Galletly C, Anderson D, Walker P. Computer aided cognitive rehabilitation: possible application to the attentional deficit of schizophrenia, a report of negative results. *Perceptual and Motor Skills* 1997;**85**:995–1002. [MEDLINE: 9399309]

Fiszdon 2004 *{published data only}*
Fiszdon JM, Bryson GJ, Wexler BE, Bell MD. Durability of cognitive remediation training in schizophrenia: performance on two memory tasks at 6-month and 12-month follow-up. *Psychiatry Research* 2004;**125**(1):1–7. [MEDLINE: 14967547]

Gelkopf 1994 *{published data only}*
Gelkopf M, Sigal M, Kramer R. Therapeutic use of humor to improve social support in an institutionalized schizophrenic inpatient community. *Journal of Social Psychology* 1994;**134**(2):175–82. [MEDLINE: 8201815]

Genevsky 2010 *{published data only}*
Genevsky A, Fisher M, Garrett CT, Cook C, Vinograd S. Increased serum bdnf induced by cognitive training in schizophrenia is negatively associated with baseline SAA. *Schizophrenia Research* 2010;**117**(2):368.

Han 2008 *{published data only}*
Han DH, Renshaw PF, Sim ME, Kim JI, Arenella LS, Lyoo IK. The effect of internet video game play on clinical and extrapyramidal symptoms in patients with schizophrenia. *Schizophrenia Research* 2008;**103**(1-3):338–40.

Jaugey 2012 *{published data only}*
Jaugey L, Urben S, Pihet S, Halfon O, Holzer L. Short-and long-term outcomes of a randomized controlled trial of a computer-assisted cognitive remediation (cacr) program in adolescents with psychosis or at high risk of psychosis. *Biological Psychiatry* 2012;**8**(Suppl 1):84S.

Kim 2007 *{published data only}*
Kim K, Kim JJ, Kim J, Park DE, Jang HJ, Ku J, et al. Characteristics of social perception assessed in schizophrenia using virtual reality. *Cyberpsychology & behavior: the impact of the Internet, multimedia and virtual reality on behavior and society* 2007;**10**(2):215–9. [PUBMED: 17474838]

Marsh 2010 *{published data only}*
Marsh PJ, Langdon R, McGuire J, Harris A, Luckett G, Polito V, et al. The development of the mental-state reasoning training (msr) program: Phase i and ii. *Australian and New Zealand Journal of Psychiatry* 2010;**44**(1 Suppl): A36.

Medalia 1998 *{published data only}*
Medalia A, Aluma M, Tryon W, Merriam AE. Effectiveness of attention training in schizophrenia. *Schizophrenia Bulletin* 1998;**24**(1):147–52. [MEDLINE: 9502552]

NCT00507988 2007 *{published data only}*
NCT00507988. Complex problem solving training in schizophrenic patients. http://www.clinicaltrials.gov 2007.

NCT00655239 2008 *{published data only}*
NCT00655239. Effectiveness of neuroadaptive cognitive training in adolescents at risk for psychosis. http://www.clinicaltrials.gov 2008.

NCT00694889 2008 *{published data only}*
NCT00694889. Randomized clinical trial of intensive computer-based cognitive remediation in recent-onset schizophrenia. http://www.clinicaltrials.gov 2008.

NCT00712075 2008 *{published data only}*
NCT00712075. Functional rehabilitation for older patients with schizophrenia. http://www.clinicaltrials.gov 2008.

NCT00995553 2009 *{published data only}*
NCT00995553. Skills training with computers. http://www.clinicaltrials.gov 2009.

NCT01027962 2009 *{published data only}*
NCT01027962. Intensive Computerized Brain Training (ICBT) in youth with Early Onset Psychosis (EOP). http://www.clinicaltrials.gov 2009.

NCT01036282 2009 *{published data only}*
NCT01036282. Cognitive skills training using computer for patients with severe mental illness. http://www.clinicaltrials.gov 2009.

NCT01422902 2011 *{published data only}*
NCT01422902. Evaluation of a cognitive adaptive e-treatment in schizophrenia-diagnosed adults. http://ClinicalTrials.gov/show/NCT01422902 2011.

Park 2009a *{published data only}*
Park KM, Ku J, Park IH, Park JY, Kim SI, Kim JJ. Improvement in social competence in patients with schizophrenia: a pilot study using a performance-based measure using virtual reality. *Human Psychopharmacology* 2009;**24**(8):619–27. [MEDLINE: 19946937]

Rotondi 2010 *{published data only}*
Rotondi AJ, Anderson CM, Haas GL, Eack SM, Spring MB, Ganguli R, et al. Web-based psychoeducational intervention for persons with schizophrenia and their supporters: One-year outcomes. *Psychiatric Services* 2010; **11**:1099–105. [MEDLINE: CN-00772198]

Steinwachs 2011 *{published data only}*
Steinwachs DM, Roter DL, Skinner EA, Lehman AF, Fahey M, Cullen B, et al. A web-based program to empower patients who have schizophrenia to discuss quality of care with mental health providers. *Psychiatric Services* 2011;**62** (11):1296–302. [MEDLINE: 22211208]

Subramaniam 2011 *{published data only}*
Subramaniam K, Woolley JD, Luks TL, Garrett C, Nagarajan S, Vinograd S. Cortical plasticity in

schizophrenia patients after computerised cognitive training: Behavioral and fmri assessments of working memory performance. *Biological Psychiatry* 2011;**9**(Suppl 1):239.

Vauth 2001 *{published data only}*
Vauth R, Barth A, Stieglitz RD. Evaluation of a cognitive strategy training in vocational rehabilitation of schizophrenic outpatients [Evaluation eines kognitiven Strategietrainings in der ambulanten beruflichen Rehabilitation Schizophrener]. *Zeitschrift für Klinische Psychologie und Psychotherapie* 2001;**30**(4):251–8.

Vinogradov 2011 *{published data only}*
Vinogradov S, Subramaniam K, Herman A, Woolley J, Luks T, Fisher M, et al. Neuroscience-informed cognitive training in schizophrenia "normalizes" brain-behavior associations in auditory and verbal working memory processes: Evidence from meg and fmri. Proceedings of the 50th Annual Meeting of the American College of Neuropsychopharmacology; 2011 Dec 4-8; Waikoloa, Hawaii. 2011.

Westermann 2012 *{published data only}*
Westermann S, Kesting ML, Lincoln TM. Being deluded after being excluded? How emotion regulation deficits in paranoia-prone individuals affect state paranoia during experimentally induced social stress. *Behavior Therapy* 2012;**43**(2):329–40.

Wirshing 2003 *{published data only}*
Wirshing DA, Mintz J, Kern R, Boyd J. A videotape intervention to enhance the informed consent process for patients participating in clinical research trials. *Schizophrenia Research* 2003;**60**:306.

References to ongoing studies

UKCRNID12951 2012 *{published data only}*
UKCRNID12951. Examining whether virtual reality can help make people feel safer. http://public.ukcrn.org.uk/Search/StudyDetail.aspx?StudyID=12951 2012.

Additional references

Altman 1996
Altman DG, Bland JM. Detecting skewness from summary information. *BMJ* 1996;**313**(7066):1200.

APA 1994
American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)*. 4th Edition. Washington DC: American Psychiatric Association, 1994.

APA 2000
American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR*. 4. American Psychiatric Association, 2000.

Beuter 2004
Beuter L, Harwood T. Virtual reality in psychotherapy training. *Journal of Clinical Psychology* 2004;**60**(3):317–30.

Bland 1997
Bland JM. Statistics notes. Trials randomised in clusters. *BMJ* 1997;**315**:600.

Boissel 1999
Boissel JP, Cucherat M, Li W, Chatellier G, Gueyffier F, Buyse M, et al. The problem of therapeutic efficacy indices. 3. Comparison of the indices and their use [Apercu sur la problematique des indices d'efficacite therapeutique, 3: comparaison des indices et utilisation. Groupe d'etude des Indices d'efficacite]. *Therapie* 1999;**54**(4):405–11. [PUBMED: 10667106]

Buchanan 2005
Buchanan RW, Carpenter WT. Concept of schizophrenia. In: Sadock BJ, Sadock VA editor(s). *Kaplan & Sadock´s Comprehensive Textbook of Psychiatry*. 8th Edition. Vol. **II**, Philadelphia, USA: Lippincott Williams & Wilins, 2005.

Coleman 2005
Coleman EA, Tulman L, Samarel N, Wilmoth MC, Rickel L, Rickel M, et al. The effect of telephone social support and education on adaptation to breast cancer during the year following diagnosis. *Oncology Nursing Forum* 2005;**32**(4):822–9. [PUBMED: 15990911]

Coons 2011
Coons MJ, Roehrig M, Spring B. The potential of virtual reality technologies to improve adherence to weight loss behaviors. *Journal of Diabetes Science and Technology* 2011;**5**(2):340–4.

Da Costa 2004
Da Costa R, De Carvalho L. The acceptance of virtual reality devices for cognitive rehabilitation: a report of positive results with schizophrenia. *Computer Methods and Programs in Biomedicine* 2004;**73**:173–82.

Deeks 2000
Deeks J. Issues in the selection for meta-analyses of binary data. Proceedings of the 8th International Cochrane Colloquium; 2000 Oct 25-28; Cape Town. Cape Town: The Cochrane Collaboration, 2000.

Divine 1992
Divine GW, Brown JT, Frazier LM. The unit of analysis error in studies about physicians' patient care behavior. *Journal of General Internal Medicine* 1992;**7**(6):623–9.

Donner 2002
Donner A, Klar N. Issues in the meta-analysis of cluster randomized trials. *Statistics in Medicine* 2002;**21**:2971–80.

Dyck 2010
Dyck M, Winbeck M, Leiberg S, ChenY, Mathiak K. Virtual faces as a tool to study emotion recognition deficits in schizophrenia. *Psychiatry Research* 2010;**179**:247-52.

Egger 1997
Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;**315**:629–34.

Eisenstein 2002
Eisenstein N, Engelhart CI, Johnson V, Wolf J, Williamson J, Losonczy MB. Normative data for healthy elderly persons with the neurobehavioral cognitive status exam (Cognistat). *Applied Neuropsychology* 2002;**9**(2):110–3. [PUBMED: 12214821]

Elbourne 2002

Elbourne D, Altman DG, Higgins JPT, Curtina F, Worthingtond HV, Vaile A. Meta-analyses involving cross-over trials: methodological issues. *International Journal of Epidemiology* 2002;**31**(1):140–9.

Erickson 2011

Erickson M, Jaafari N, Lysaker P. Insight and negative symptoms as predictors of functioning in a work setting in patients with schizophrenia. Psychiatry Research 2011; Vol. 189, issue 2:161–5.

Esteves Moreira da Costa 2004

Esteves Moreira Da Costa RM, Vidal De Carvalho LA. The acceptance of virtual reality devices for cognitive rehabilitation: a report of positive results with schizophrenia. *Computer Methods and Programs in Biomedicine* 2004;**73**: 173–82.

Furukawa 2006

Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. *Journal of Clinical Epidemiology* 2006;**59**(7):7–10.

Gerardi 2010

Gerardi M, Cukor J, Difede J, Rizzo A, Rothbaum B. Virtual reality exposure therapy for post-traumatic stress disorder and other anxiety disorder. *Current Psychiatry Report* 2010;**12**:298–305.

Goff 2011

Goff DC, Hill M, Freudenreich O. Treatment adherence in schizophrenia and schizoaffective disorder. *Journal of clinical Psychiatry* 2011;**72**(4):e13.

Goodman 2005

Goodman C, Knoll G, Isakov V, Silver H. Negative attitude towards medication is associated with working memory impairment in schizophrenia patients. *International Clinical Psychopharmacology* 2005;**20**(2):93–6.

Gregg 2007

Gregg L, Tarrier N. Virtual reality in mental health: a review of the literature. *Social Psychiatry and Psychiatric Epidemiology* 2007;**42**:343–534.

Gulliford 1999

Gulliford MC. Components of variance and intraclass correlations for the design of community-based surveys and intervention studies: data from the Health Survey for England 1994. *American Journal of Epidemiology* 1999;**149**: 876–83.

Haynes 2008

Haynes RB, Ackloo E, Sahota N, McDonald HP, Yao X. Interventions for enhancing medication adherence. *Cochrane Database of Systematic Reviews* 2008, Issue 2. [DOI: 10.1002/14651858.CD000011.pub3]

Higgins 2003

Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**: 557–60.

Higgins 2011

Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Jenerette 2010

Jenerette C, Dixon J. Developing a short form of the simple Rathus assertiveness schedule using a sample of adults with sickle cell disease. *Journal of Transcultural Nursing : official journal of the Transcultural Nursing Society / Transcultural Nursing Society* 2010;**21**(4):314–24. [PUBMED: 20592057]

Johansen 2011

Johansen R, Hestad K, Iversen VC, Agartz I, Sundet K, Andreassen OA, et al. Cognitive and clinical factors are associated with service engagement in early-phase schizophrenia spectrum disorders. *Journal of Nervous and Mental Disease* 2011;**199**(3):176–82.

Johansson 2012

Johansson M, Wressle E. Validation of the Neurobehavioral Cognitive Status Examination and the Rivermead Behavioural Memory Test in investigations of dementia. *Scandinavian Journal of Occupational Therapy* 2012;**19**(3): 282–7. [PUBMED: 21105840]

Josman 2009

Josman N, Schenirderman A, Klinger E, Shevil E. Using virtual reality to evaluate executive functioning among persons with schizophrenia: A validity study. *Schizophrenia Research* 2009;**115**:270-7.

Kay 1986

Kay SR, Opler LA, Fiszbein A. *Positive and Negative Syndrome Scale (PANSS) Manual*. North Tonawanda, NY: Multi-Health Systems, 1986.

Kim 2009

Kim K, Kim CH, Kim SY, Roh D, Kim SI. Virtual reality for obsessive-compulsive disorder: past and the future. *Psychiatry Investigation* 2009;**6**:115–21.

Kirkpatrick 2005

Kirkpatrick B, Tek C. Schizophrenia: Clinical features and psychopathology concepts. In: Sadock BJ, Sadock VA editor(s). *Kaplan & Sadock´s Comprehensive Textbook of Psychiatry*. 8th Edition. Vol. **II**, Philadelphia, USA: Lippincott Williams & Wilins, 2005.

Ku 2007

Ku J, Han K, Lee HR, Jang HJ, Kim KU, Park SH, et al. VR-based conversation training program for schizophrenia: a preliminary clinical trial. *Cyberpsychology & Behaviour* 2007;**10**:567–74.

Laver 2011

Laver KE, George S, Thomas S, Deutsch JE, Crotty M. Virtual reality for stroke rehabilitation. *Cochrane Database of Systematic Reviews* 2011, Issue 9. [DOI: 10.1002/14651858.CD008349.pub2]

Leucht 2005

Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel RR. What does the PANSS mean?. *Schizophrenia Research* 2005;**79**(2-3):231–8. [PUBMED: 15982856]

Leucht 2005a

Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel R. Clinical implications of brief psychiatric rating scale scores. *British Journal of Psychiatry* 2005;**187**:366–71. [PUBMED: 16199797]

Leucht 2007

Leucht S, Engel RR, Bauml J, Davis JM. Is the superior efficacy of new generation antipsychotics an artifact of LOCF?. *Schizophrenia Bulletin* 2007;**33**(1):183–91. [PUBMED: 16905632]

Malbos 2011

Malbos E, Rapee R, Kavakli M. Isolating the effect of virtual reality based exposure therapy for agoraphobia: a comparative trial. *Studies in Health Technology and Informatics* 2011;**167**:44–50.

Marshall 2000

Marshall M, Lockwood A, Bradley C, Adams C, Joy C, Fenton M. Unpublished rating scales: a major source of bias in randomised controlled trials of treatments for schizophrenia. *British Journal of Psychiatry* 2000;**176**: 249–52.

McClellan 2005

McClellan JM. Early-onset schizophrenia. In: Sadock BJ, Sadock V editor(s). *Kaplan & Sadock´s Comprehensive Textbook of Psychiatry*. 8th Edition. Vol. **II**, Philadelphia, USA: Lippincott Williams & Wilins, 2005:3307-14.

McLay 2011

McLay R, Webb-Murphy J, Spira J, Wiederhold M, Pyne J, Wiederhold B. A randomized, controlled trial of virtual reality-graded exposure therapy for post-traumatic stress disorder in active duty service members with combat-related post-traumatic stress disorder. *Cyberpsychology, Behavior, and Social Networking* 2011;**14**(4):223–9.

Overall 1962

Overall JE, Gorham DR. The brief psychiatric rating scale. *Psychological Reports* 1962;**10**:799–812.

Park 2009

Park I, Kim J, Ku J, Jan H, Park H, Kim C, et al. Characteristics of social anxiety from virtual interpersonal interactions in patients with schizophrenia. *Psychiatry* 2009; **72**(1):79.

Reger 2011

Reger GM, Holloway KM, Candy C, Rothbaum BO, Difede J, Rizzo AA, et al. Effectiveness of virtual reality exposure therapy for active duty soldiers in a military mental health clinic. *Journal of Traumatic Stress* 2011;**24**(1):93-6.

Riva 2005

Riva G. Virtual reality in psychotherapy: Review. *Cyber Psychology & Behaviour* 2005;**8**(3):221–30.

Schünemann 2008

Schünemann HJ, Oxman AD, Vist GE, Higgins JPT, Deeks JJ, Glasziou P, et al. Chapter 12: Interpreting results and drawing conclusions. In: Higgins JPT, Green S editor(s). *Cochrane Handbook for Systematic Reviews of Interventions*. The Cochrane Collaboration, 2008:359–83.

Ukoumunne 1999

Ukoumunne OC, Gulliford MC, Chinn S, Sterne JAC, Burney PGJ. Methods for evaluating area-wide and organisation-based intervention in health and health care: a systematic review. *Health Technology Assessment* 1999;**3**(5): 1–75.

Viertiö 2011

Viertiö S, Tuulio-Henriksson A, Perälä J, Saarni SI, Koskinen S, Sihvonen M, et al. Activities of daily living, social functioning and their determinants in persons with psychotic disorder. European Psychiatry 2012; Vol. 27, issue 6:409–15.

Wakeling 2007

Wakeling HC. The psychometric validation of the Social Problem-Solving Inventory--Revised with UK incarcerated sexual offenders. *Sexual Abuse : a journal of research and treatment* 2007;**19**(3):217–36. [PUBMED: 17333399]

Xia 2009

Xia J, Adams CE, Bhagat N, Bhagat V, Bhoopathi P, El-Sayeh H, et al. Loss to outcomes stakeholder survey: the LOSS study. *Psychiatric Bulletin* 2009;**33**(7):254–7.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Chan 2010

Methods	Allocation: randomised. Blindness: open. Duration of the follow-up: 3 months.* Design: parallel. Location: single located.
Participants	Diagnosis: schizophrenia. N = 29. Age: 60 and over, mean ~66 years (SD ~6). Sex: 18 males, 9 females. History: functionally impaired and are unable to live independently in the community. Having schizophrenia more than 20 years. MMSE score mean in both group ~18 (SD ~2) Excluded: not reported. Setting: Long-stay care, long-term residential care.
Interventions	1. Virtual reality program**: IREX*** system involving red gloves and movement within a demarcated area, viewing large screen: 10 session 15 minutes twice a week. N = 14 2. Usual care: usual program in the facility. N = 15.
Outcomes	Compliance: lost to follow-up. Functioning: Cognistat. Acceptability of the intervention: leaving the study early - any reason Not able to use: Adverse effects: SSQ (only reported for intervention group). Behaviour: VQ (only reported for intervention group).
Notes	* Assumed as participants in control group were given virtual programme "3 months later". ** We assume that virtual reality group also got usual care *** The Interactive Rehabilitation Exercise System

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	"Randomly assigned" - no further details.
Allocation concealment (selection bias)	Unclear risk	No details.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	No description of any attempt.

Chan 2010 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Undertaken by "another occupational therapist" - unclear of blinded
Incomplete outcome data (attrition bias) All outcomes	High risk	"To eliminate the effect of alteration in psychotropic medication, information of two participant was excluded from analysis"
Selective reporting (reporting bias)	Unclear risk	No protocol available.
Other bias	High risk	SSQ and VQ only asked of intervention group. Small-study publication bias possibility

Park 2011

Methods	Allocation: randomised. Blindness: open. Duration of the follow-up: 5 weeks.* Design: parallel. Location: single located.
Participants	Diagnosis: schizophrenia. N = 91, analyses 63. Age: mean ~28 years (SD ~8) in experimental, mean ~31 years (SD ~8) in control Sex: 34 males, 30 females. History: first onset mean ~22 years (SD ~6) in experimental, mean ~25 years (SD ~7) in control, duration of illness mean ~6 years (SD ~6) in experimental, mean ~6 years (SD ~6) in control, previous psychiatric admissions mean ~2 (SD ~2) in experimental, mean ~3 (SD ~3) in control. PANSS total score mean ~73 (SD ~13) in experimental, mean ~71 (SD ~13) in control Excluded: substance abuse, head trauma, neurological illness or physical illness that could affect brain functioning Setting: Severance Mental Health Hospital, Yonsei University College Medicine
Interventions	1. Social skill training using virtual reality role-playing (SST-VR), VR system included personal computer for virtual environment and head mounted display for displaying the virtual environment in a more immersive manner, and a position tracker : 10 session, in 5 weeks. N = 32 2. Social skill training using traditional role-playing (SST-TR):10 session, in 5 weeks. N = 31
Outcomes	Compliance: loss to follow-up. Satisfaction with treatment: motivation and generalization for each session. Acceptability of the intervention: leaving the study early - any reason Data not usable: Social skills (RAS, RCS, SBS, SPSI-R). No numeric data available
Notes	* Assumed as intervention was 5 weeks.

Park 2011 (*Continued*)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	"Randomly assigned" - no further details.
Allocation concealment (selection bias)	Unclear risk	No details.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	No description of any attempt.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No description of any attempt.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	They have presented participant flow in "fig. a describes the participants progress in this study"
Selective reporting (reporting bias)	Unclear risk	No protocol available.
Other bias	Unclear risk	Small-study publication bias possibility.

Tsang 2013

Methods	Allocation: randomised with a computational random number generation Blindness: the assessors were blinded to the group assignment Duration: duration of the intervention 5 weeks. Design: parallel. Location: inpatients attended a vocational rehabilitation program at the psychiatric hospital
Participants	Diagnosis: schizophrenia (DSM-IV). N = 95. Age: mean 41 years. Sex: M 33, F 42. History: hospitalised patients with a variety of reasons (relapse, unmanageable behaviour problem in the community). Duration of hospitalisation for acute patients about 6-8 weeks, about 15 weeks for extended patients. Patients had about 13 years of education, and their duration of illness was about 15 years. Patients (with Chinese ethnicity) had normal or mild psychotic symptoms (an average score of BPRS 21, SD 6) Exclusion: physical handicaps (e.g. blindness), undergone ECT therapy during the past 12 months, an episode of drug abuse during the past 30 days, a history of mental retardation or other neurological disease and development disabilities

Interventions	1. Virtual reality based vocational skills training in a virtual boutique scenario (VRG); and prevocational skills training in work-simulated workshops. N = 33
	2. Prevocational skills training in work-simulated workshops. (CG) N = 30
Outcomes	Compliance: Loss to follow-up
	Acceptability of the intervention: Leaving the study early - any reason
	Data not able to use - mean and SD not available:
	Functioning: General cognitive ability (Brief Neuropsychological Cognitive Examination, BNCE)
	Behaviour: Work perspective (Vocational Cognitive Rating Scale, VCRS)
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomly assigned to the three groups (VRG, TAG and CG) using a computational random number generator
Allocation concealment (selection bias)	Unclear risk	No further details.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	No further details.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The assessors were blinded to the group assignment. Independent assessors, who did not know the expected results of the training programs, were responsible for the pre-test and post-test outcome assessments at baseline and post-intervention
Incomplete outcome data (attrition bias) All outcomes	Low risk	Attrition rate 21%.
Selective reporting (reporting bias)	Unclear risk	No protocol available.
Other bias	Unclear risk	-

BNCE - Brief Neuropsychological Cognitive Examination
BPRS - Brief Psychiatric Rating Scale
DSM IV - Diagnostic and Statistical Manual fourth edition
ECT - electroconvulsive therapy
IREX - two-dimensional VR programme that enables a person to engage in a simulated task within a VR environment through video contact

MMSE - Mini Mental State Examination
N - number of participants
PANSS - Positive and Negative Syndrome Scale
RAS - Rathus Assertiveness Schedule
RCS - Relationship Change Scale
SD - standard deviation
SSQ - Simulator Sickness Questionnaire
SST-VR - Social skill training using virtual reality role-playing
SST-TR - Social skill training using traditional role-playing
SBS - Social Behavior Scale
SPSI-R - Social Problem Solving Inventory
VCRS - Vocational Cognitive Rating Scale
VQ - Volitional Questionnaire

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Benedict 1994	Allocation: randomised. Participants: chronic schizophrenia. Intervention: attention training with computers + day treatment program vs. day treatment program alone, attention training involved interaction with computers but not creating or being involved with an interactive world
Bryson 2003	Allocation: randomised. Participants: schizophrenia. Intervention: individualised computer-based cognitive rehabilitation (ICBCR) remediation vs. cognitive remediation therapy (CRT), individualised computer-based cognitive rehabilitation involved interaction with computers but not creating or being involved with an interactive world
Chinman 2004	Allocation: randomised. Participants: schizophrenia, schizoaffective disorder or bipolar disorder Intervention: 2 standardised symptom survey via audio computer-assisted self-interviewing survey self-administered through Internet browser vs. 2 standardised symptom survey via an in-person interview vs. distraction task between the 2 standardised symptom survey administrations patients were shown 20 minutes of video, brief interview conducted after surveys. These intervention involved interaction with computers but not creating or being involved with an interactive world and intervention was not treatment focus
David 2012	Allocation: randomised. Participants: schizophrenia. Intervention: viewing videos of themselves vs. video of a same-sex actor displaying psychotic symptoms does not involve interactive world
Dunn 2001	Allocation: randomised. Participants: schizophrenia, schizoaffective disorder, mood disorder with psychotic features, or psychosis not otherwise specified Intervention: computer-based enhanced consent (EC) vs. routine consent (RC), computer-based enhanced consent involved interaction with computers but not creating or being involved with an interactive world and the computer was not the focus of treatment

Field 1997	Allocation: matched control design, no further details. Participants: schizophrenia. Intervention: computer-aided cognitive rehabilitation vs. graphics-based computer games, both involved interaction with computers but not creating or being involved with an interactive world and intervention was not treatment focus
Fiszdon 2004	Allocation: randomised, stratified according to intake cognitive functioning Participants: schizophrenia or schizoaffective disorder. Intervention: neurocognitive enhancement therapy/remediation + work therapy vs work therapy alone, cognitive enhancement involved interaction with computers but not creating or being involved with an interactive world
Gelkopf 1994	Allocation: randomised, cluster trial. Participants: chronic schizophrenia. Intervention: video projection of humor movies vs. video projections of other movies, no interaction with computers nor creating or being involved with an interactive world
Genevsky 2010	Allocation: randomised. Participants: schizophrenia. Intervention: cognitive training using software vs. commercially available video games vs. healthy comparison receiving cognitive training using software. Cognitive training program involved interaction with computers but not creating or being involved with an interactive world
Han 2008	Allocation: no information provided. Participants: schizophrenia. Interventions: Internet video games vs. watching movies, no interaction with computers nor creating or being involved with an interactive world
Jaugey 2012	Allocation: randomised. Participants: adolescents.
Kim 2007	Allocation: not randomised, case control.
Marsh 2010	Allocation: randomised. Participants: schizophrenia. Intervention: manual-based program developed around videos, computer games and group games vs. social activities control group, manual-based program involved interaction with computers but not creating or being involved with an interactive world
Medalia 1998	Allocation: randomised. Participants: schizophrenia. Intervention: individual computerised attention remediation vs. individual sessions during which they viewed video documentaries, individual computerised attention remediation involved interaction with computers but not creating or being involved with an interactive world
NCT00507988 2007	Allocation: randomised. Participants: schizophrenia. Intervention: computer-based cognitive remediation vs. commercially available computer games vs. healthy

	comparator, cognitive remediation training involved interaction with computers but not creating or being involved with an interactive world
NCT00655239 2008	Allocation: randomised. Participants: schizophrenia or schizoaffective disorder. Intervention: computer-based cognitive remediation vs. commercially available computer games, cognitive remediation program involved interaction with computers but not creating or being involved with an interactive world
NCT00694889 2008	Allocation: randomised. Participants: schizophrenia or schizoaffective disorder. Intervention: computer-based cognitive remediation vs. commercially available computer games, cognitive remediation program involved interaction with computers but not creating or being involved with an interactive world
NCT00712075 2008	Allocation: randomised. Participants: older patients with schizophrenia. Intervention: computer-assisted (PDA) cognitive behavioral social skills training (CBSST) vs. cognitive behavioural social skills training (CBSST) vs. PDA only. All these interventions involved interaction with computers but not creating or being involved with an interactive world
NCT00995553 2009	Allocation: randomised. Participants:schizophrenia or schizoaffective disorder. Intervention:neurocognitive enhancement therapy/remediation + work therapy vs work therapy alone, cognitive enhancement involved interaction with computers but not creating or being involved with an interactive world
NCT01027962 2009	Allocation: randomised. Participants: early onset psychosis. Intervention: intensive computerised brain training using software vs. commercially available computer games vs. healthy comparison with no computer activity, intensive computerised brain training using software involved interaction with computers, and computer skills not creating or being involved with an interactive world
NCT01036282 2009	Allocation: randomised. Participants: schizophrenia or schizoaffective disorder. Intervention: computerised cognitive remediation, CRT vs. social cognition training, computerised cognitive remediation involved interaction with computers but not creating or being involved with an interactive world
NCT01422902 2011	Allocation: randomised. Participants: schizophrenia, 18 Years and older. Intervention: non-plasticity-based Computer Treatment: Active Comparator vs. Plasticity-based Computer Treatment: Active Comparator involves interaction with computers but not creating or being involved with interactive world
Park 2009a	Allocation: randomised. Participants: schizophrenia. Intervention: aripiprazole + virtual reality vs. risperidone + virtual reality, evaluating effects of different drugs with outcomes including those relevant to virtual reality, not evaluating virtual reality alone

Rotondi 2010	Allocation: randomised. Participants: schizophrenia or schizoaffective disorder. Intervention: online tele-health intervention vs. treatment as usual, online tele-health program involved interaction with computers but not creating or being involved with an interactive world
Steinwachs 2011	Allocation: randomised. Participants: schizophrenia. Intervention: An interactive web-based intervention featuring actors simulating a patient discussing treatment concerns, no virtual reality
Subramaniam 2011	Allocation: randomised. Participants: schizophrenia. Intervention: Cognitive training program vs. computer-games involves interaction with computers but not creating or being involved with interactive world
Vauth 2001	Allocation: randomised. Participants: schizophrenia or schizoaffective disorder. Intervention: involved interaction with computers but not creating or being involved with an interactive world
Vinogradov 2011	Allocation: randomised. Participants: schizophrenia. Intervention: Cognitive training program vs. computer-games involves interaction with computers but not creating or being involved with interactive world
Westermann 2012	Allocation: randomised. Participants: schizophrenia. Intervention: both groups used Cyberball ball-tossing game.
Wirshing 2003	Allocation: randomised. Participants: schizophrenia. Intervention: intervention using video to provide information to patients, not using virtual reality

CBSST - cognitive behavioral social skills training
CRT - cognitive remediation therapy
EC - enhanced consent
ICBRC - individualised computer based cognitive rehabilitation
PDA - computer-assisted
RC - routine consent
vs - versus

Characteristics of ongoing studies *[ordered by study ID]*

UKCRNID12951 2012

Trial name or title	Examining whether virtual reality can help make people feel safer
Methods	No further details.
Participants	The participants will be 30 patients with persecutory delusions attending treatment services, typically with a case note diagnosis of schizophrenia, schizoaffective disorder or delusional disorder. They will report feeling paranoia when around other people and using within-situation safety behaviours. The delusions will meet the criteria of Freeman & Garety (2000)
Interventions	Virtual reality (computer environments).
Outcomes	No further details.
Starting date	No further details.
Contact information	Dr Katherine Pugh Warneford Hospital University Department of Psychiatry Warneford Lane Headington Oxford Oxfordshire OX3 7JX UNITED KINGDOM Tel: 01865226468 katherine.pugh@psych.ox.ac.uk
Notes	

DATA AND ANALYSES

Comparison 1. VIRTUAL REALITY AS A SOLE METHOD + STANDARD PROFESSIONAL CARE vs. STAN-DARD PROFESSIONAL CARE

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Compliance	3		Risk Difference (M-H, Fixed, 95% CI)	Subtotals only
1.1 loss to follow-up	3	156	Risk Difference (M-H, Fixed, 95% CI)	0.02 [-0.08, 0.12]
1.2 being withdrawn by trialist	1	29	Risk Difference (M-H, Fixed, 95% CI)	0.14 [-0.06, 0.35]
2 Functioning: 1. Cognitive - average endpoint total score (Cognistat, high is good)	1	27	Mean Difference (IV, Fixed, 95% CI)	4.67 [-1.76, 11.10]
3 Functioning: 2. Social - average change in specific aspects of skills (various scales)	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
3.1 assertiveness (RAS, high is good)	1	64	Mean Difference (IV, Fixed, 95% CI)	0.60 [-2.59, 3.79]
3.2 relationship change (RCS, high is good)	1	64	Mean Difference (IV, Fixed, 95% CI)	-2.00 [-7.83, 3.83]
3.3 social behaviour (SBS, high is good)	1	64	Mean Difference (IV, Fixed, 95% CI)	-1.70 [-5.03, 1.63]
3.4 social problem solving (SPSI-R, high is good)	1	64	Mean Difference (IV, Fixed, 95% CI)	-2.30 [-8.13, 3.53]
4 Satisfaction with treatment: Average change scores (unreferenced measure)	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
4.1 interest in social skills training (high is good)	1	64	Mean Difference (IV, Fixed, 95% CI)	6.00 [1.39, 10.61]
4.2 generalization of the skill (high is good)	1	64	Mean Difference (IV, Fixed, 95% CI)	5.10 [1.03, 9.17]
5 Acceptability of intervention: leaving the study early for any reason	2	92	Risk Difference (M-H, Fixed, 95% CI)	0.05 [-0.09, 0.19]

ADDITIONAL TABLES

Table 1. Additional table from Cognistat - sub analyses

Cognitive domain	VR M (SD)	SC M (SD)
Orientation	4.00 (2.66)	3.87 (3.02)
Attention	7.83 (0.58)	7.60 (0.91)

Table 1.   Additional table from Cognistat - sub analyses   *(Continued)*

Comprehension	4.25 (1.48)	4.27 (1.39)
Repetition	8.08 (2.54)	6.60 (2.75)
Naming	4.67 (2.50)	4.27 (1.22)
Constructions	1.17 (1.34)	6.60 (2.75)
Memory	2.92 (2.23)	2.60 (2.03)
Calculation	2.67 (1.15)	1.87 (1.19)
Similarities	0.17 (0.58)	0.07 (0.26)
Judgement	1.91 (1.73)	0.87 (1.85)

SC - standard care

SD - standard deviation

M - mean

VR - virtual reality

Table 2.   Suggested future reviews

Review title	Studies
Computer games for schizophrenia	Genevsky 2010
Informed consent	Wirshing 2003
Combining virtual reality with antipsychotic medication for schizophrenia	Park 2009a
Using humour for schizophrenia	Gelkopf 1994

Table 3.   Suggested design of trial

Methods	Allocation: randomised, clearly described and concealed. Blinding: open. Duration: 3 years.
Participants	Diagnosis: people with schizophrenia. N = 500. Age: over 18 years. Sex: both. History: not too clinically acute.

Table 3. Suggested design of trial (*Continued*)

	Excluded: injuries or condition that may compromise cognitive function
Interventions	1. Virtual reality program related medication use in home environment, using virtual reality social skills program. 5 times, 30 minutes 2. Standard care.
Outcomes	Compliance Relapse Satisfaction with treatment
Notes	* For 20% difference in binary outcome to be apparent power calculation of 0.05, 80% power needs 150 per group

N - number of participants

CONTRIBUTIONS OF AUTHORS

Maritta Välimäki - initiation of the review, protocol production, data extraction, analysis, data interpretation, writing the final report.

Heli Hätönen - protocol production, data extraction, analysis, data interpretation, writing the final report.

Mari Lahti - protocol production, data extraction, analysis, data interpretation, writing the final report.

Anja Hottinen - protocol production, data extraction.

Marjo Kurki - protocol production, data extraction.

Kiki Metsäranta - protocol production, data extraction, writing the final report.

Tanja Riihimäki - protocol production, data extraction.

Clive Adams - protocol production, analysis, data interpretation, writing the final report.

DECLARATIONS OF INTEREST

All authors: We do have an interest in information technology use for people with schizophrenia, such as mobile technology, Internet, social media or virtual reality games. Grants and research funding for the topics are listed in 'External sources'.

SOURCES OF SUPPORT

Internal sources

- The University of Turku, Finland.

The University of Turku offer research facilities to conduct the review.

External sources

- The Academy of Finland, Finland.

Grant number: 132581

- The Hospital District of Southwest Finland, Finland.
- Turku University Hospital, Finland.

EVO 13893

- Finnish Cultural Foundation, Finland.
- Foundations' Professor Pool, Finland.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We searched 2nd of November 2011 PubMED database by using the phrase: (virtual reality and schizophrenia). This search found Park 2011 study included in this review.

For binary outcomes we calculated a standard estimation of the risk difference (RD), not the risk ratio (RR). We presented the RD to show all data, as one trial did not have any loss to follow-up. Using Relative Risk makes no discernable difference.

Taken from the most recent methodology of the Cochrane Schizophrenia Group, it acknowledges changes in Cochrane methodology since the protocol of this review was published. We do not think the use of these new methods affects the results or integrity of the review.

Due to a limited number of validated instruments used in included studies, two unreferenced measures reported by Park 2011 were included into our analysis regarding patient satisfaction with treatment.

INDEX TERMS

Medical Subject Headings (MeSH)

*Patient Compliance; *User-Computer Interface; Mental Disorders [therapy]; Randomized Controlled Trials as Topic; Schizophrenia [*therapy]; Schizophrenic Psychology

MeSH check words

Adult; Humans