The use of polygenic risk scores to identify phenotypes associated with genetic risk of bipolar disorder and depression: a systematic review

Sumit Mistry¹, Judith R. Harrison¹, Daniel J. Smith², Valentina Escott-Price¹ and Stanley Zammit¹,³

¹Institute of Psychological Medicine and Clinical Neurosciences, MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, UK; ²Institute of Health and Wellbeing, I Lilybank Gardens, University of Glasgow, UK; ³Centre for Academic Mental Health, Population Health Sciences, Bristol Medical School, University of Bristol, UK

Abstract

Background

Identifying the phenotypic manifestations of increased genetic liability for depression (MDD) and bipolar disorder (BD) can enhance understanding of their aetiology. The polygenic risk score (PRS) derived using data from genome-wide-association-studies can be used to explore how genetic risk is manifest in different samples.

Aims

In this systematic review, we review studies that examine associations between the MDD and BD polygenic risk scores and phenotypic outcomes.
Methods

Following PRISMA guidelines, we searched EMBASE, Medline and PsycINFO (from August 2009 – 14th March 2016) and references of included studies. Study inclusion was based on predetermined criteria and data were extracted independently and in duplicate.

Results

Twenty-five studies were included. Overall, both polygenic risk scores were associated with other psychiatric disorders (not the discovery sample disorder) such as depression, schizophrenia and bipolar disorder, greater symptom severity of depression, membership of a creative profession and greater educational attainment. Both depression and bipolar polygenic risk scores explained small amounts of variance in most phenotypes (<2%).

Limitations

Many studies did not report standardised effect sizes. This prevented us from conducting a meta-analysis.

Conclusions

Polygenic risk scores for BD and MDD are associated with a range of phenotypes and outcomes. However, they only explain a small amount of the variation in these phenotypes. Larger discovery and adequately powered target samples are required to increase power of the PRS approach. This could elucidate how genetic risk for bipolar disorder and depression is manifest and contribute meaningfully to stratified medicine.

Key words: polygenic risk score, bipolar disorder, depression, genetic, phenotypes
1. Introduction

Mood disorders (major depressive disorder (MDD) and bipolar disorder (BD)) are common, highly heritable psychiatric conditions as evidenced by twin, adoption and family studies (Bienvenu et al., 2011; Oswald et al., 2003; Shih et al., 2004; Sullivan et al., 2000). Family history of a mood disorder is a strong risk factor for future development of a mood disorder. High-risk studies of offspring of parents with a mood disorder have shown that offspring have an increased lifetime risk for BD (Duffy et al., 2014) or MDD when compared to controls (Rice et al., 2002). Additionally, they are also at increased risk for developing other psychiatric disorders (Rasic et al., 2014). Previously, information about how genetic risk for these disorders was manifest in the population was obtained from high-risk studies that followed up offspring of parents with these disorders. However, a major limitation of these high-risk studies was the relatively small sample sizes, which meant studies were inadequately powered to detect small effect sizes.

The advent of genome-wide association studies (GWAS) has revolutionised identification of genetic variants contributing to psychiatric disorders. GWAS technology can examine many genetic variants in the genome simultaneously, without an a priori hypotheses. Approaches have been developed to harness this information, allowing us to study how individuals with different burdens of genetic risk differ from one another. Through GWAS, several risk variants associated with MDD and BD have been identified (Major Depressive Disorder Working Group of the Psychiatric Genetics Consortium et al., 2017; Sklar et al., 2011; Stahl et al., 2017; Sullivan, 2013). Though the number of risk variants is fewer than for schizophrenia (SZ) (Purcell et al., 2009; Ripke et al., 2014), this is likely due to the smaller discovery samples used in the GWAS analyses, and the lower heritability of MDD compared to schizophrenia (SZ). Whilst individual single nucleotide polymorphisms (SNPs) have very small effect on disease risk, summing the weighted allelic dosage across all SNPs, creating a single polygenic risk score (PRS) has enabled the exploration of how genetic risk is manifest directly in individuals across different populations (Figure 1) (Wray et al., 2014).

Since the first GWAS were published, many studies have examined whether the MDD-PRS or BD-PRS are associated with a range of different phenotypes and outcomes. Though other methods such as linkage disequilibrium (LD) score regression or genomic residual maximum likelihood (GREML) analysis exist to examine genetic correlation, the focus of this systematic review was to identify and
summarise studies that used the PRS approach to examine how genetic risk for MDD and BD is manifest in clinical and population-based samples.
2. Methods

We undertook a systematic review following Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009) (Supplementary Table 1).

2.1 Search strategy

2.1.1 Inclusion/exclusion criteria

We included articles that examined associations between the PRS (derived from GWAS data of participants with a diagnosis of MDD or BD) and a measurable phenotype. We excluded neuroimaging outcomes as the complexity of this field is beyond the scope of this systematic review. Articles reporting associations between the MDD-PRS/BD-PRS with a diagnosis of MDD or BD respectively were not included. We required articles to be in peer reviewed journals and written in English (see Supplementary Table 2 for inclusion/exclusion criteria).

2.1.2 Data sources

We searched EMBASE, Medline and PsychINFO databases from 06/08/2009 to 14/03/2016. We also hand searched references of the articles included.

2.1.3 Search terms and delimiters

We searched for articles using the terms “depression/bipolar (or variations of)” AND “polygenic (or variations of)”. Our full search strategy terms are listed in Supplementary List 1.
2.1.4 Data collection and analysis

2.1.4.1 Selection of studies

In keeping with our previous systematic review (Mistry et al., 2017), after duplicates were removed, 1043 articles remained (see Figure 2 for flow diagram). One author, (S.M.) screened the title and abstract. If it was unclear whether the paper was relevant, or if the abstract was unavailable, we retrieved the full-text article. Two authors independently reviewed each full-text article and checked them against the inclusion criteria. A third author resolved disagreements. We extracted relevant information using a data extraction form (see Supplementary Table 3). We used a narrative approach to summarise results as most studies did not report standardised effect sizes (or provide data that would allow us to calculate these) as required for conducting a meta-analysis.
3. Results

This review includes 25 articles that examined associations between the MDD-PRS/BD-PRS and a measurable phenotype (see Table 1 for a summary of associations with broad phenotypic outcomes). Most studies derived their PRS from the same discovery sample: the first GWAS from the Psychiatric Genetics Consortium (PGC) for MDD and BD (Sklar et al., 2011; Sullivan, 2013). Individual studies used different p-value thresholds (P Ts) to assess the relationship between genetic risk for the disorder and phenotype or outcome (see Supplementary Tables 4-6). Most studies were of White/Caucasian adults of European ancestry.

3.1 Studies examining phenotypic associations with both the MDD-PRS and BD-PRS

Several studies used both the MDD and BD polygenic risk scores to examine associations between genetic risk and a measurable phenotype.

3.1.1 Associations with psychiatric disorders

Using data from the PGC, both the MDD and BD polygenic risk scores were strongly associated with SZ (strongest P T<0.4, p=1x10^-16; strongest P T<0.3, p=1x10^-50 respectively). However, there was no association between the MDD-PRS and ADHD or ASD across any P T (no p-values reported), and weak evidence for associations between the BD-PRS and ADHD or ASD (strongest P T<0.4, p<0.05; strongest P T<0.001, p<0.05 respectively). The MDD-PRS was associated with BD (strongest P T<0.5, p=1x10^-12) and the BD-PRS was associated with MDD (strongest P T<0.5, p=1x10^-16) (Cross-Disorder Group of the Psychiatric Genomics, 2013).

A study using data from the Australian Twin Registry (ATR) and Midwest Alcohol Research Centre found no evidence that the MDD-PRS was associated with seasonal affective disorder (SAD) (strongest at P T<0.01, p=0.25), but there was some evidence of association between the BD-PRS and SAD (strongest P T<0.1, p=0.004) (Byrne et al., 2015).
Both the BD and MDD polygenic scores were associated with SZ, either when comparing those with SZ and a positive history of psychotic illness to controls (strongest at $P_T<1$, 1-sided $p=7.11\times10^{-149}$ and $P_T<1$, 1-sided $p=1.49\times10^{-59}$ respectively) or when comparing those with SZ and a negative family history of psychotic illness to controls (strongest at $P_T<1$, 1-sided $p<1\times10^{-300}$ and $P_T<1$, 1-sided $p=9.95\times10^{-254}$ respectively) (Bigdeli et al., 2015).

Finally, a higher BD-PRS was associated with post-traumatic stress disorder (PTSD) (strongest $P_T<0.3$, $p=0.028$), but the MDD-PRS was not associated at any $P_T$ (no statistics provided) in the Marine Resilience Studies (Nievergelt et al., 2015).

### 3.1.2 Association with other phenotypes

Both the MDD and BD polygenic risk scores were also associated with: i) greater episode count of MDD in individuals with MDD (strongest $P_T<0.4$, $p<0.001$; strongest $P_T<0.2$, $p=0.015$ respectively), particularly in individuals with MDD and a positive family history of MDD (strongest $P_T<0.4$, $p=0.01$; strongest $P_T<0.2$, $p=0.004$ respectively) using data from the RADIANT sample (Ferentinos et al., 2014), and ii) a family history of psychotic illness in individuals with SZ from the PGC-1-SZ GWAS (strongest $P_T<0.4$, one-sided $p=0.011$; strongest $P_T<0.2$, one-sided $p=0.04$ respectively) (Bigdeli et al., 2015).

In a combined analysis using data from four different samples (Mullins et al., 2014), the MDD-PRS was associated with a greater number of suicide attempts (strongest $P_T<0.3$, $p=0.013$), whilst the BD-PRS was not (strongest $P_T<0.1$; no $p$-value reported).

Using data from the National Institute of Neurological Disorders and Stroke (NINDS) sample, the MDD-PRS was higher in Parkinson’s disease (PD) cases than controls (no $P_T$ reported; $p=0.001$; $AUC=0.52$), but the BD-PRS was not (no $P_T$ or $p$-value reported; $AUC=0.5$) (Schulze et al., 2014).

Finally, whilst neither the MDD or BD polygenic risk scores were associated with measures of latent inhibition (ASSR, P3 latency or P50 ratio), there was strong evidence of an association between the BD-PRS and P3 amplitude (strongest $P_T<1\times10^{-5}$; $p=0.005$) in a sample of individuals with psychotic illness and controls (Hall et al., 2015).
3.2 Studies examining phenotypic associations with the MDD-PRS

3.2.1 MDD-PRS and psychiatric phenotypes

Studies reported associations between the MDD-PRS and: i) greater chronicity of depressive symptoms (no p-value reported) in the Health and Retirement Study (HRS) sample (Levine et al., 2014), ii) higher Hospital Anxiety and Depression Scale-Anxiety subscale (HADS-A) scores in both the Rotterdam Study of Anxiety Disorder Case-Control sample (strongest $P_t < 0.5$, $p=0.0025$) and the Erasmus Ruchpen Family (ERF) sample (strongest $P_t < 0.01$, $p=0.01$) (Demirkan et al., 2011) and iii) higher General Health Questionnaire (GHQ) scores (strongest $P_t < 0.05$; $p=9x10^{-6}$) and neuroticism (strongest $P_t < 1$, $p=1x10^{-4}$) in the Generation Scotland: The Scottish Family Health Study (GS:SFHS) (Clarke et al., 2015).

3.2.2 MDD-PRS and other phenotypes

There was no consistent evidence of associations between the MDD-PRS and stressful life events (SLE) in the RADIANT samples, varying by both type of SLE (dependent or independent) and MDD case/control status (Mullins et al., 2016). There was also inconsistent evidence of associations between the MDD-PRS and childhood trauma (CT) in RADIANT, with opposite patterns of association in MDD cases compared to the controls. Furthermore, the pattern of association across MDD cases and controls in RADIANT was in the opposite direction to that within MDD cases and controls in the combined Genetics of Recurrent Early-Onset Depression (GenRED) and Depression Genes Networks (DGN) (Peyrot et al., 2014).

Finally, the MDD-PRS was associated with higher body mass index (BMI) (strongest $P_t < 0.01$, $p=0.01$) in the GS:SFHS, which did not survive correction for multiple testing (no statistics reported) (Clarke et al., 2015).
3.3 Studies examining phenotypic associations with the BD-PRS

3.3.1 BD-PRS and BD symptoms/severity

The BD-PRS was associated with a manic symptom factor (strongest $P_T<0.3$, $p=0.003$), but not with depressive, positive or negative symptom factors across any $P_T$ (no statistics reported) in individuals with SZ drawn from the PGC-1-SZ GWAS (Ruderfer et al., 2014).

There was very weak evidence that the BD-PRS was associated with a positive history of psychosis in individuals with BD (strongest $P_T<0.05$, $p=0.079$) using data from the Thematically Organised Psychosis sample (Aminoff et al., 2015), and with decreased paranoia (strongest $P_T<0.5$, one-sided $p=0.064$) and decreased anhedonia (one-sided $p=0.048$). However, the BD-PRS was not associated with parent-rated negative symptoms, grandiosity, cognitive/disorganization and hallucinations in adolescents from the population-based Longitudinal Experiences and Perceptions (LEAP) study (Sieradzka et al., 2014).

The BD-PRS was not associated with antidepressant response in individuals with MDD treated with Selective Serotonin Reuptake Inhibitors (SSRIs) (strongest $P_T<0.05$, $p=0.613$), Noradrenaline Reuptake Inhibitors (NRIs) (strongest $P_T<0.1$, $p=0.168$), or either antidepressant (strongest $P_T<0.3$, $p=0.623$) in the New Medications in Depression and Schizophrenia sample. Neither was the BD-PRS associated with response to citalopram (strongest $P_T<0.3$, $p=0.826$) in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) sample (Tansey et al., 2014).

When combining the NESDA and Netherlands Twin Registry (NTR) samples, the BD-PRS was associated with MDD (strongest $P_T<0.5$, $p=0.001$) and both severe typical (strongest $P_T<0.5$; $p=0.018$) and atypical MDD (strongest $P_T<0.1$, $p=0.032$) (Milaneschi et al., 2016).

There was inconsistent evidence of associations between the BD-PRS and characteristics of depression (including severity, age of onset, history of suicide attempt, recurrence, and atypicality) within a multivariate framework, or with subclinical mania in the STAR*D, Mannheim and NESDA samples (Wiste et al., 2014).
3.3.2 BD-PRS and other psychopathology

In data from a Norwegian sample, the BD-PRS was i) higher in SZ-spectrum cases compared to controls without SZ, BD or MDD (strongest $P_T<0.05; p=0.01$), but not in SZ-spectrum cases compared to BD-spectrum cases (strongest $P_T<0.05; p=0.13$), and ii) not associated with a lifetime history of psychosis in BD-spectrum cases compared to controls (no statistics provided) (Tesli et al., 2014).

Using data from i) the combined NESDA/NTR samples and ii) the Queensland Institute of Medical Research (QIMR) sample, a study reported associations between the BD-PRS and post-partum depression (strongest $P_T<0.001, p=0.004$ and $P_T<0.1, p = 3.04\times10^{-5}$ respectively) (Byrne et al., 2014).

In the GAIN MDD sample, the BD-PRS was higher in MDD cases compared to controls (no $P_T$ reported, $p = 7.32\times10^{-7}$) with modest diagnostic accuracy AUC=0.55, and was also higher in SZ cases compared to controls in the GAIN-SZ sample (no $P_T$ reported, $p=2.9\times10^{-9}$; AUC=0.56) (Schulze et al., 2014).

Finally, a higher BD-PRS was found in ADHD cases compared to controls (strongest $P_T<0.5; p=0.052$) using combined data from UK community child psychiatry and paediatric clinics and from Dublin Ireland (Hamshere et al., 2013).

3.3.3 BD-PRS and other phenotypes

In an Icelandic general population study, the BD-PRS was associated with creativity (strongest at $P_T<0.2; p=5.2\times10^{-6}$), number of years in school ($p=4.8\times10^{-9}$) and having a university degree ($p=5.2\times10^{-7}$) (Power et al., 2015).
4. Discussion

To the best of our knowledge, this is the first paper to systematically review how genetic risk for MDD and BD is associated with a broad range of phenotypic outcomes. Other reviews have focused on the application of PRS methodology to psychiatric disorders, particularly SZ (Wray et al., 2014). More recently we have reported phenotypes associated with genetic risk for SZ (Mistry et al., 2017).

Higher MDD and BD polygenic risk scores were associated with increased risk of different psychopathologies. Typically, $R^2$ values for other psychiatric disorders (0.5-2%) were greater than for other phenotypes examined (most <1%). One exception is a study that reported that approximately 6% to 13% of the variance for SZ was explained by the MDD-PRS, and 17% to 22% by the BD-PRS (depending on the presence or absence of a family history of psychotic illness. However, these estimates seem likely to be over-estimates given that the SZ-PRS only explained 12% to 13% of the variance for SZ in this study (Bigdeli et al., 2015).

4.1 Psychiatric outcomes

Associations between genetic risk for MDD/BD and psychiatric disorders was stronger for disorders typically presenting in adulthood rather than in childhood. Associations between the MDD-PRS and BD-PRS with SZ is consistent with evidence of pleiotropy across these disorders (Gale et al., 2016). Those with MDD who also have psychotic symptoms or a family history of psychotic illness likely index individuals with more severe forms of MDD (Park et al., 2014; Parker et al., 2013). There was no evidence the MDD-PRS was associated with SAD (Byrne et al., 2015). However, associations between both the BD and SZ polygenic scores and SAD in the same study have been found using risk scores derived from larger GWAS discovery data sets than for MDD, suggesting that low power may be a likely explanation for this lack of association (Dudbridge, 2013).

4.2 Other psychopathology

Both the MDD-PRS and BD-PRS were associated with greater episode count of MDD, the latter particularly in those who also had a family history of MDD (Ferentinos et al., 2014). Furthermore,
associations between the BD-PRS and early age of onset of MDD (<26 years) has been reported more recently (Power et al., 2017; Verduijn et al., 2017). These genetic study findings are also consistent with previous epidemiological studies showing that greater recurrence of MDD, presence of a family history of MDD, and early age of onset might index individuals who have an underlying ‘latent BD’ rather than having MDD (i.e. are at higher risk of transitioning from MDD to BD diagnoses than those without these characteristics) (Benazzi, 2006).

The MDD-PRS was also associated with neuroticism in the Generation Scotland: The Scottish Family Health Study (Clarke et al., 2015), a finding also shown in another population based study, the UK Biobank (Gale et al., 2016). Neuroticism is robustly associated with increased risk for developing MDD (Hakulinen et al., 2015). Taken together, data from both genetic and epidemiological studies would suggest that within the general population, genetic risk for MDD can manifest as personality traits such as neuroticism.

Genetic risk for BD was not associated with antidepressant response (though effects on response to tricyclics or monoamine oxidases has not yet been investigated). More recently, a study reported no association between the MDD-PRS and antidepressant response in individuals with MDD using the same classes of antidepressant used by Tansey and colleagues (Garcia-Gonzalez et al., 2017). Taken together, this suggests that neither genetic risk for BD or MDD is currently helpful in determining whether an individual will respond better to a particular class of antidepressant. It is likely that polygenic risk scores based on gene pathways are more likely to determine whether antidepressant response can be determined using genetic risk for BD or MDD, or exploration of whether the MDD and BD PRSs are associated with tricyclics or monoamine oxidases.

4.3 Other phenotypes

Associations between the MDD-PRS and CT were inconsistent, with studies reporting opposing patterns of association (Mullins et al., 2016; Peyrot et al., 2014), likely reflecting the problems inherent in studying how patterns differ across sub-groups (Zammit et al., 2010).

Increased genetic risk for BD was associated with creativity and greater educational attainment (Power et al., 2015). Some epidemiological studies report positive associations between higher childhood IQ and BD (Koenen et al., 2009; Smith et al., 2015) and greater educational attainment and BD (MacCabe et al., 2010; Vreeker et al., 2016). However, others looking at adolescent or adult
IQ report no association with BD (Reichenberg et al., 2002; Zammit et al., 2004) and poorer educational attainment with BD (Glahn et al., 2006; Mojtabai et al., 2015; Vonk et al., 2012).

Similarly, cognitive impairments throughout all mood states of BD manifest in young adults (Kurtz and Gerraty, 2009; Martino et al., 2016), rather than in the premorbid phase (Martino et al., 2015). However, BD is a heterogeneous disorder: not all of those affected go on to have cognitive deficits (Martino et al., 2008).

Assessments of creativity in those with BD can be studied using neuroimaging. During creative tasks, those with BD have increased functional connectivity in the fronto-parietal network (Goya-Maldonado et al., 2016), including dorosolateral prefrontal (Boccia et al., 2015) and medial prefrontal cortices (Jung et al., 2013) compared to controls. As these areas of the brain are typically involved in executive functioning and planning, these neuroimaging findings might explain increased openness to experience, elevated impulsivity and increased desire to engage in more complex activities in those with BD. Whilst currently limited by refinement of neuroimaging methods, sample sizes, and validity of BD-PRS, it is feasible that the PRS approach, especially if based on gene-pathways, could inform the biology of specific cognitive outcomes and imaging parallels of these.

Similar approaches such as LD score regression also suggest similar findings when examining genetic correlations between adult psychiatric disorders, educational attainment and personality traits such as neuroticism (Bulik-Sullivan et al., 2015; Verneri et al., 2016).

4.4 Implications

Overall, the results presented in this systematic review highlight that the polygenic risk scores for MDD and BD explain a small proportion of the variance for all phenotypes (<2%), with the exception of the study by Bigdeli et al. (2015). The variance explained will likely increase as discovery sample sizes increase (Dudbridge, 2013), though this may still be limited as the PRS does not capture rare SNP or copy number variants contributions to variance.

At the present time, the PRS approach is not useful for informing clinical practice or pharmacological interventions. However, as predictive utility increases with larger discovery and adequately powered target samples, the PRS, alongside other approaches, could help inform stratified medicine.

Our review also shows that genetic risk for MDD and BD is manifest in the adult population as a broad range of psychopathologies. As the literature on genetic risk scores for MDD and BD
associations with childhood psychopathologies is extremely limited, it is difficult to determine how or whether genetic risk influences phenotypes earlier in development.

4.5 Strengths and limitations

A major strength to this systematic review is that we used a comprehensive search strategy to reduce the likelihood of omitting eligible studies. In addition, we did not limit this review to a specific sampling framework or research design e.g. clinical samples or longitudinal studies. We also included studies investigating a range of phenotypes and outcomes, providing a broad overview of how genetic risk for bipolar and depression can manifest. However, the inconsistency in the reporting of results across studies meant that only a narrative approach to this review was feasible, and assessment of publication bias was not possible. We excluded neuroimaging results as they are significantly more complex to summarise and are therefore not able to comment on how depression or bipolar genetic risk relates to imaging phenotypes. Studies not published in English-language journals will also have been missed.

We identified a number of limitations of studies included in this review. This made it difficult to put their findings in context. Problems included inadequate descriptions of sample ascertainment, the strength of evidence incorrectly interpreted based on one-sided p-values, and failure to report standardised effect sizes or confidence intervals to permit comparison across studies. In order to improve interpretation and comparison between studies, we have developed a reporting framework for researchers to use (Mistry et al., 2017).

An important limitation of the PRS approach is that it is based on the assumption that both target and discovery samples are independent of each other: any overlap will inflate the risk estimates. Furthermore, when interpreting the findings from PRS analyses, both an adequately powered target sample size and a large discovery sample size are required to optimise association testing and risk prediction (Dudbridge, 2013); so whilst most studies were adequately powered to detect small-moderate effect sizes (OR=1.2-1.5 per SD of the PRS), this is based on the assumption of no measurement error in the PRS.
4.6 Conclusions

At present, the PRS approach explains very little of the variance in phenotypes, with the greatest variance explained being for other psychiatric disorders. This is the first review attempting to collate information on how the PRS approach has informed our understanding of a variety of phenotypes in terms of association with depression and bipolar disorder genetic risk. Future studies will benefit from larger sample sizes from which to derive PRS, adequately powered target sample sizes, well-validated phenotypic measures and more robust ways of reporting PRS associations with phenotypes.

5. Funding

SM is funded by Mental Health Research UK. DJS is a Lister Institute Prize Fellow (2016-2021). JRH is funded by the Wellcome Trust GW4 Clinical Academic Training Fellowship. This study was supported by the NIHR Biomedical Research Centre at University Hospitals Bristol NHS Foundation Trust and the University of Bristol. The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health.

6. Contributors

Author S.M carried out the literature search, writing of the first draft of the manuscript and authors S.M., J.R.H., D.S. and S.Z. checked studies to be included against inclusion criteria. Authors S.M. and J.R.H. extracted data from included studies and all authors checked the final manuscript.

All authors have no conflicts of interest to declare.
7. Acknowledgements

We would like to thank the Cardiff University librarians who assisted with our search strategy.
References


