Cochrane Library

Cochrane Database of Systematic Reviews

An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes (Review)

Flodgren G, Eccles MP, Shepperd S, Scott A, Parmelli E, Beyer FR

Flodgren G, Eccles MP, Shepperd S, Scott A, Parmelli E, Beyer FR.
An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes.
Cochrane Database of Systematic Reviews 2011, Issue 7. Art. No.: CD009255.
DOI: 10.1002/14651858.CD009255.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Overview of Reviews]

An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes

Gerd Flodgren[1], Martin P Eccles[2], Sasha Shepperd[3], Anthony Scott[4], Elena Parmelli[5], Fiona R Beyer[6]

[1]Department of Public Health, University of Oxford, Headington, UK. [2]Institute of Health and Society, Newcastle University, Newcastle upon Tyne, UK. [3]Department of Public Health, University of Oxford, Oxford, UK. [4]Melbourne Institute of Applied Economic and Social Research, The University of Melbourne, Carlton, Melbourne, Australia. [5]Department of Oncology, Hematology and Respiratory Diseases, University of Modena and Reggio Emilia, Modena, Italy. [6]Centre for Reviews and Dissemination, University of York, York, UK

Contact address: Martin P Eccles, Institute of Health and Society, Newcastle University, Badiley Clark Building, Richardson Road, Newcastle upon Tyne, NE2 4AX, UK. Martin.Eccles@newcastle.ac.uk.

Editorial group: Cochrane Effective Practice and Organisation of Care Group.
Publication status and date: New, published in Issue 7, 2011.

Citation: Flodgren G, Eccles MP, Shepperd S, Scott A, Parmelli E, Beyer FR. An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes. *Cochrane Database of Systematic Reviews* 2011, Issue 7. Art. No.: CD009255. DOI: 10.1002/14651858.CD009255.

ABSTRACT

Background

There is considerable interest in the effectiveness of financial incentives in the delivery of health care. Incentives may be used in an attempt to increase the use of evidence-based treatments among healthcare professionals or to stimulate health professionals to change their clinical behaviour with respect to preventive, diagnostic and treatment decisions, or both. Financial incentives are an extrinsic source of motivation and exist when an individual can expect a monetary transfer which is made conditional on acting in a particular way. Since there are numerous reviews performed within the healthcare area describing the effects of various types of financial incentives, it is important to summarise the effectiveness of these in an overview to discern which are most effective in changing health professionals' behaviour and patient outcomes.

Objectives

To conduct an overview of systematic reviews that evaluates the impact of financial incentives on healthcare professional behaviour and patient outcomes.

Methods

We searched the Cochrane Database of Systematic Reviews (CDSR) (*The Cochrane Library*); Database of Abstracts of Reviews of Effectiveness (DARE); TRIP; MEDLINE; EMBASE; Science Citation Index; Social Science Citation Index; NHS EED; HEED; EconLit; and Program in Policy Decision-Making (PPd) (from their inception dates up to January 2010). We searched the reference lists of all included reviews and carried out a citation search of those papers which cited studies included in the review. We included both Cochrane and non-Cochrane reviews of randomised controlled trials (RCTs), controlled clinical trials (CCTs), interrupted time series (ITSs) and controlled before and after studies (CBAs) that evaluated the effects of financial incentives on professional practice and patient outcomes, and that reported numerical results of the included individual studies. Two review authors independently extracted data and assessed the methodological quality of each review according to the AMSTAR criteria. We included systematic reviews of studies evaluating the effectiveness of any type of financial incentive. We grouped financial incentives into five groups: payment for working for a specified time period; payment for each service, episode or visit; payment for providing care for a patient or specific population; payment for providing a pre-specified level or providing a change in activity or quality of care; and mixed or other systems. We summarised data using vote counting.

Main results

We identified four reviews reporting on 32 studies. Two reviews scored 7 on the AMSTAR criteria (moderate, score 5 to 7, quality) and two scored 9 (high, score 8 to 11, quality). The reported quality of the included studies was, by a variety of methods, low to moderate. Payment for working for a specified time period was generally ineffective, improving 3/11 outcomes from one study reported in one review. Payment for each service, episode or visit was generally effective, improving 7/10 outcomes from five studies reported in three reviews; payment for providing care for a patient or specific population was generally effective, improving 48/69 outcomes from 13 studies reported in two reviews; payment for providing a pre-specified level or providing a change in activity or quality of care was generally effective, improving 17/20 reported outcomes from 10 studies reported in two reviews; and mixed and other systems were of mixed effectiveness, improving 20/31 reported outcomes from seven studies reported in three reviews. When looking at the effect of financial incentives overall across categories of outcomes, they were of mixed effectiveness on consultation or visit rates (improving 10/17 outcomes from three studies in two reviews); generally effective in improving processes of care (improving 41/57 outcomes from 19 studies in three reviews); generally effective in improving referrals and admissions (improving 11/16 outcomes from 11 studies in four reviews); generally ineffective in improving compliance with guidelines outcomes (improving 5/17 outcomes from five studies in two reviews); and generally effective in improving prescribing costs outcomes (improving 28/34 outcomes from 10 studies in one review).

Authors' conclusions

Financial incentives may be effective in changing healthcare professional practice. The evidence has serious methodological limitations and is also very limited in its completeness and generalisability. We found no evidence from reviews that examined the effect of financial incentives on patient outcomes.

PLAIN LANGUAGE SUMMARY

How good are financial incentives in changing health care?

There is a lot of interest in how well financial incentives influence the delivery of health care. Financial incentives are extrinsic sources of motivation and they exist when an individual can expect a monetary transfer which is made conditional on acting in a particular way. Since there are several reviews describing the effects of different types of financial incentives, it is important to bring this together in an overview to examine which are best at changing healthcare professionals' behaviours and what happens to patients. We therefore conducted an overview of systematic reviews that evaluated the impact of financial incentives on healthcare professional behaviour and patient outcomes. We searched a wide range of electronic databases from when they started up to December 2008. We included systematic reviews of studies evaluating the effectiveness of any type of financial incentive. We grouped financial incentives into five groups: payment for working for a specified time period; payment for each service, episode or visit; payment for providing care for a patient or specific population; payment for providing a pre-specified level or providing a change in activity or quality of care; and mixed or other systems. We summarised data using vote counting. We identified four reviews reporting on 32 studies. Two reviews were of moderate quality and two were of high quality. The studies that the reviews reported on were of low to moderate quality. Payment for working for a specified time period was generally ineffective. Payment for each service, episode or visit was generally effective, as were payment for providing care for a patient or specific population and payment for providing a pre-specified level or providing a change in activity or quality of care; mixed and other systems were of mixed effectiveness. When looking at the effect of financial incentives overall across different outcomes, they were of mixed effectiveness on consultation or visit rates; generally effective in improving processes of care, referrals and admissions, and prescribing costs; and generally ineffective in improving compliance with guidelines outcomes. On the basis of these findings, we concluded that financial incentives may be effective in changing healthcare professional practice. The evidence has serious methodological limitations and is also very limited in its completeness and generalisability. We found no evidence that financial incentives can improve patient outcomes.

BACKGROUND

There is considerable interest in the effectiveness of financial incentives in the delivery of health care. Incentives may be used in an attempt to increase the use of evidence-based treatments among healthcare professionals or to stimulate health professionals to change their clinical behaviour with respect to preventive, diagnostic and treatment decisions, or both. For instance, within the UK National Health Service, the Quality and Outcome Framework for Primary Care (NICE 2004) links performance of clinical actions across 12 clinical areas to financial incentives (Roland 2004). The ultimate goal of using financial incentives to change healthcare professionals' behaviours should be increased quality of care and, by extension, improved patient outcomes, reduced costs, or improved access to care. The World Health Organization (WHO) describes six dimensions of the quality of care (WHO 2006). These suggest that for the health care to be considered of high quality it is required to be effective, that is adherent to an evidence base and resulting in improved health outcomes for individuals and communities based on need; efficient, delivered in a manner which maximises resource use and avoids waste; accessible, that is timely, geographically reasonable and provided in a setting where skills and resources are appropriate to medical need; acceptable and patient-centred, delivered in a way which takes into account the preferences and aspirations of individual service users and the cultures of their communities; equitable, delivering health care which does not vary in quality because of personal characteristics such as gender, race, ethnicity, geographical location, or socioeconomic status; and safe, delivered in a way which minimises risks and harm to service users (WHO 2006). It is unclear whether financial incentives will influence all, or any, of these areas in a positive way.

Definition of financial incentives

An incentive is any factor (financial or non-financial) that provides motivation for a particular course of action, or counts as a reason for preferring one choice compared to alternatives.

Intrinsic sources of motivation for clinicians include the likelihood that patients' health will improve as a result of a course of action, and motivation from performing a task well. Other sources of motivation include social and peer group norms, where a particular choice is regarded by others as the right thing to do, as particularly admirable, or where the failure to act in a certain way is condemned.

Financial incentives are extrinsic sources of motivation and exist when an individual receives a monetary transfer which is made conditional on acting in a particular way. Whilst similar incentives exist ('in-kind' transfers of resources, gifts, or lotteries), our focus is on financial incentives, which we defined as changes in the amount or method of monetary transfers to a healthcare provider for the purpose of this review.

Though there are issues with nomenclature and the definitions, the healthcare area uses different types of financial incentives including : (i) salary or sessional payment (payment for working for a specified time period); (ii) fee-for-service (payment for each service, episode or visit); (iii) capitation (payment for providing care for a patient or for a special population); and (iv) target payments and bonuses (payment for providing a pre-specified level or change in a specific behaviour or quality of care).

How the intervention might work

Financial incentives are usually seen as a method of increasing output per unit cost. In public service settings in general, and health care in particular, the context within which financial incentives operate is complex and it is seldom the case that there would be a sole aim to increase quantity or throughput without any consideration of quality of care. In health care there are multiple dimensions of output which are difficult to measure and monitor, and therefore also difficult to attribute to the specific actions of a specific healthcare provider. In addition, there may be tensions between intrinsic motivation and extrinsic financial incentives. Where there is high intrinsic motivation (such as with health professionals), it has been argued that there is less need for high powered financial incentives (Mooney 1993). Others have argued that extrinsic incentives may 'crowd out' or reduce intrinsic motivation, thereby leading to less care being provided (Frey 1997). A financial incentive aimed at increasing the throughput of patients within an out-patient department, or using fee-for-service payment generally, is likely to increase the number of patients seen but may not be compatible with providing high quality care. Similarly, a financial incentive per patient (capitation payment) pays a fixed amount per patient and provides incentives to minimise costs and only treat less complex and resource intensive patients (so called cream-skimming), and so again may not be compatible with improving quality of care.

In health care, authors have argued for blended payment schemes that reduce the impact of 'extreme' incentives in fee-for-service or capitation, alongside an element of pay-for-performance (Robinson 2001). Financial incentives in health care are likely to have the twin aims of increasing the quality and efficiency of care (Ettner 2006). In some cases the strength of the incentive may have a bearing on its effectiveness. For example, a weak incentive to perform a valued behaviour may be more effective than a strong incentive to perform a behaviour that is not seen as important by the physician. Additionally, financial incentives may also produce unpredictable and unintended behaviours or changes in performance in other areas.

Why it is important to do this overview

There is considerable interest in the effectiveness of financial incentives on the delivery of health care. Since there are numerous reviews performed within the healthcare area describing the effects of many various types of financial incentives, it is important to summarise the effectiveness of these in an overview to, if possible, discern which ones are the most effective in changing health professionals' behaviour and patients' outcomes. Furthermore, this overview has potential usefulness for decision-makers in an area where evidence is scattered across many sources. The review will also highlight methodological issues regarding the appropriate conduct of systematic reviews in this area.

OBJECTIVES

To conduct an overview of systematic reviews that evaluated the impact of financial incentives on healthcare professional behaviour and patient outcomes. We have taken into account the type of financial incentive and how it was structured, the healthcare professional being targeted, the behaviour being incentivised, how the behaviour is measured, and the healthcare system.

METHODS

Criteria for considering reviews for inclusion

We included systematic reviews of randomised controlled trials (RCTs), controlled clinical trials (CCTs), controlled before and after studies (CBAs), and interrupted time series (ITSs) evaluating the effect of financial incentives on the quality or efficiency of health care delivered by healthcare providers. Since we wanted to be able to report both size and direction of effect, only reviews presenting results as numerical data on an individual study basis, for all included studies, were included in the overview. If a review reported results as text, with no numerical outcome data, it was excluded.

If there was more than one review covering exactly the same studies, the review that provided the most complete presentation of results (plus information on study designs, characteristics of participants etc.) was chosen to be included in the overview, and the others were excluded.

Reviews rated as having an inadequate search strategy were excluded, defined by having searched less than two sources according to the DARE criteria (www.york.ac.uk/inst/crd/darefaq.htm). We allowed searching both reference lists and 'other relevant sources' as meeting the criteria for one of the sources.

Types of participants

We included physicians, dentists, nurses, and allied healthcare professions (such as physiotherapists, speech therapists etc.) involved in providing direct patient care. Healthcare providers could be targeted individually or at the level of the organisation within which they worked.

Types of intervention

We included systematic reviews of studies evaluating the effectiveness of any type of financial incentive including the following.

- Payment for working for a specified time period (e.g. a salary, sessional payment).
- Payment for each service, episode or visit (fee-for-service).
- Payment for providing care for a patient or specific population (e.g.capitation).
- Payment for providing a pre-specified level or change in activity or quality of care (e.g. target payments, bonuses).
- Mixed and other systems (comprising more than one of the above groups, or not classifiable).

Types of outcome measures

We considered reviews reporting the following objective measures of outcome for inclusion.

Main outcomes

- Measures of health professional clinical behaviour such as rates of performing prevention, diagnosis, and treatment behaviours (e.g. immunisation, blood pressure measurement, prescription, referral).
- Measures of health service utilisation by patients such as participation rates in immunisation schemes or mammography screening programs.
- Healthcare costs, either combined with measures of healthcare professional behaviour, quality of care or health outcomes to produce measures of efficiency or uncombined; including costs of (i) introducing the incentives, (ii) the transaction, (iii) the information systems required to implement the financial incentive, (iv) monitoring.

In the protocol we said that we would also report: measures of health professional non-clinical behaviour such as rates of performing specified non-clinical behaviours (for example education and training); measures of patient outcomes either objectively measured (for example mortality) or patient reported (for example quality of life); and any reported unintended effects of financial incentives. However, the first of these (with hindsight) does not apply to clinical behaviours and we found no data for the other two.

Search methods for identification of reviews

Electronic searches

We searched the following electronic databases for reviews (from their inception dates up to January 2010):

- Cochrane Database of Systematic Reviews (CDSR), Ovid (*The Cochrane Library* Issue 2, 2008);
- Database of Abstracts of Reviews of Effectiveness (DARE), Ovid (*The Cochrane Library* Issue 2, 2008);
- TRIP (June 2008);
- MEDLINE, Ovid (1990 to June Week 4 2008);
- EMBASE, Ovid (1990 to Week 27 2008);
- Science Citation Index, Web of Knowledge (1990 to July 2008);
- Social Science Citation Index, Web of Knowledge (1990 to July 2008).

We later carried out a specific search of the economics literature using the following databases (from their inception dates up to December 2008):

- NHS EED, Ovid (*The Cochrane Library* Issue 4, 2008);
- HEED, Wiley (1990 to December 2008);
- EconLit, OCLC (1990 to December 2008);
- PPD (Program in Policy Decision-Making) (www.researchtopolicy.ca/Search/Reviews.aspx).

The search strategies for systematic reviews incorporated the SIGN filter (www.sign.ac.uk/methodology/filters.html) for systematic reviews in MEDLINE and EMBASE, which was combined with selected index terms and free text terms. We translated the search strategy for each database using the appropriate controlled vocabulary as applicable. There was no language restriction. Studies were included regardless of publication status.

For details of the full search strategy see Appendix 1.

Searching other resources

We searched the reference lists of all included reviews. We carried out a citation search of those papers which cited studies included in

the review. We contacted authors of relevant papers regarding any further published or unpublished work.

Data collection and analysis

We downloaded all titles and abstracts retrieved by electronic searching to the reference management database EndNote. We removed duplicates and two review authors (ME, SS) working independently examined the remaining references. We excluded those studies which clearly did not meet the inclusion criteria and we obtained copies of the full text of potentially relevant references. Two review authors independently assessed the eligibility of retrieved papers (from ME, SS, GF, EP).

Selection of reviews

Two review authors independently applied the inclusion criteria (from ME, SS, GF, EP). Discussion between review authors resolved disagreements.

Data extraction and management

Two review authors independently extracted the data from reviews (ME, SS, GF, EP) into a data extraction form (Appendix 2). Discussion between review authors resolved diasagreements. We contacted the authors of reviews, and in some cases the authors of individual studies, for missing data.

When we were reviewing the studies included within the identified reviews, two review authors independently extracted data (SS, GF) into a data extraction form. Discussion between review authors resolved disagreements. We extracted and reported any relevant data within the trials that were not reported in the review.

We corrected any data errors found in the original reviews.

Assessment of methodological quality of included reviews

We assessed the methodological quality of each systematic review using the AMSTAR (A MeaSurement Tool to Assess Reviews) (Shea 2007). AMSTAR evaluates the methods used in a review against 11 criteria and assesses the degree to which review methods are unbiased. The 11 items and the way they were applied are described in Appendix 3. A review that adequately met all of the 11 criteria was considered to be a review of the highest quality. For this assessment the included reviews were categorised into bottom (score 0 to 3), middle (score 4 to 7), and upper (score 8 to 11) tertiles. Two review authors independently performed quality assessment (from ME, SS, GF, EP). Discussion between review authors resolved disagreements. We had hoped to examine variation in review quality to see if it explained variations in the results of the reviews. However, because we had to use vote counting, this was not possible.

Data synthesis

We presented data from the included reviews in summary tables (Appendix 4).

We organised the review data according to: (i) type of financial intervention(s), and (ii) type of outcomes being assessed. Having viewed the outcomes reported, we categorised them into the five groups: consultation or visit rates; process of care; referral or admission rates; compliance with guidelines; and prescribing costs. We had intended to conduct meta-analysis of the included reviews but as meta-analysis was not possible (due to missing data and heterogeneity in outcome measures) we used vote counting along with individual narrative review summaries to present the results. We reported all outcomes reported by the studies within the relevant category (not preferencing one outcome over a similar or overlapping one). Analyses were then reported as the number of outcomes favouring the intervention out of the total number of outcomes reported, based (as suggested by the Cochrane Handbook for Systematic Reviews of Interventions) on the direction of effect and not statistical significance.

We classified reviews according to the following decision rules (Weir 2010):

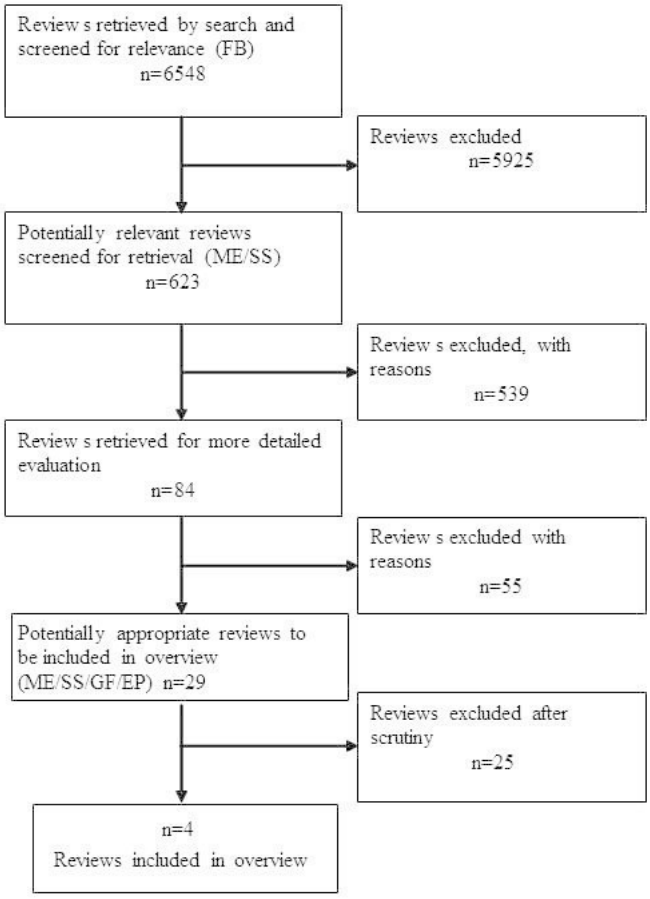
- 0% of studies (outcomes) favour intervention = no effect;
- 1% to 33% of studies (outcomes) favour intervention = generally ineffective;
- 34% to 66% studies (outcomes) favour intervention = mixed effects;
- 67%+ studies (outcomes) favour intervention = generally effective.

RESULTS

Description of included reviews

Figure 1 shows the PRISMA flow chart (Moher 2009). Of the 6548 titles initially identified by the electronic searches, we excluded 5925 titles and we screened 623 potentially relevant titles for retrieval. Of these titles, we excluded 539 and we retrieved 84 titles for more detailed evaluation. Of these, we excluded 55 (see Table 1) and 29 potentially relevant reviews remained. After examining them, we excluded 25, leaving four reviews to be included in this overview: two Cochrane reviews (Akbari 2008; Sturm 2007) and two non-Cochrane reviews (Gosden 2001; Petersen 2006).

Figure 1.



See Table 2 for details on the characteristics of included reviews (Akbari 2008; Gosden 2001; Petersen 2006; Sturm 2007).

Summaries of individual reviews

Akbari 2008

Akbari 2008 searched two databases up to 2007 and included four studies investigating the effectiveness of financial incentives on the appropriateness (and rate) of referrals: one RCT (Davidson 1992), three CBAs (Coulter 1993; Kammerling 1996; Linnala 2001). The studies in the review included more than 94 primary care physicians based in more than 48 primary care practices or municipal health services, and 80 physicians in private office-based practices (the number of practices was not reported). One study was based in the USA (Davidson 1992), two in the UK (Coulter 1993; Kammerling 1996), and one in Finland (Linnala 2001). One study evaluated the effects of change in the remuneration system from a low cost fee-for-service system to either a high cost fee-for-service system or a capitation-based budgetary system (with some degree of risk sharing by the provider) for the management of Medicaid-eligible paediatric care patients (Davidson 1992). The two UK-based studies evaluated the effect of fundholding as compared to non-fundholding on referral patterns (Coulter 1993; Kammerling 1996). Linnala 2001 examined the impact of charging patients the same (lower) cost to be seen by a private specialist as they would have been charged to see a hospital-based specialist, comparing no intervention with patient incentives and a referrals list system.

Gosden 2001

Gosden 2001 searched nine databases up to 1997 and included six studies investigating the impact of payment systems on the behaviour of primary care physicians: three RCTs (Davidson 1992; Hickson 1987; Kouides 1998), one ITS (Ritchie 1992), and two CBAs (Hutchison 1996; Krasnik 1990). The studies in the review included more than 935 primary care physicians or GPs and 149 primary care practices. Three studies were based in the USA (Davidson 1992; Hickson 1987; Kouides 1998) and one in each of Canada (Hutchison 1996), Denmark (Krasnik 1990), and UK (Ritchie 1992). Fee-for-service was the comparison intervention in five of the six studies, and in the sixth it was a mix between fee-for-service and capitation (Krasnik 1990). Davidson 1992 is described above under Akbari 2008. Hickson 1987 compared the effects on improved patient care (continuity of care) of salary reimbursement with fee-for-service (control). Two studies (Kouides 1998; Ritchie 1992) evaluated the effects on immunisation rates of introducing target payments into a fee-for-service system. Hutchison 1996 investigated the effects of changing the payment for primary care physicians who were previously paid by fee-for-service to a mix of capitation with an ambulatory incentive payment (in which the health service organisation received a payment if their

hospitalisation rate was lower than the regional rate). In one studya change from a capitation system to a mixed fee-per-item and capitation system was introduced for face-to-face, telephone and home visit consultations; and repeat prescriptions (Krasnik 1990). Additional fees were payable for a number of special services and laboratory investigations performed in the practice, and for a few preventive services.

Petersen 2006

Petersen 2006 searched one database up to November 2005; in addition the authors searched other sources thereby fulfilling our operationalisation of the DARE criteria for an adequate search. The review included nine RCTs and four CBAs investigating the relationship between explicit financial incentives and the provision of high-quality health care. The studies in the review included more than 212 physicians, 274 primary care practices, and 36 skilled nursing facilities. Four of the included studies were USA based (Beaulieu 2005; Clark 1995; Rosenthal 2005; Shen 2003) but for the remaining nine studies the country of origin was not stated. Four studies evaluated the effectiveness of enhanced fee-for-service payments (Christensen 2000; Clark 1995; Fairbrother 1999; Fairbrother 2001) and 10 studies investigated the effectiveness of target payments and bonuses (Beaulieu 2005; Fairbrother 1999; Fairbrother 2001; Grady 1997; Hillman 1998; Hillman 1999; Kouides 1998; Rosenthal 2005; Roski 2003; Shen 2003).

Sturm 2007

Sturm 2007 searched 14 databases up to April 2006 and included three ITSs and 10 CBAs assessing the effects of financial incentives on prescribing, healthcare utilisation, health outcomes, and prescribing costs. The studies in the review included more than 857 primary care practices. Ten of the studies were based in the UK and evaluated the effects of fundholding on prescriptions, generic drug prescriptions, and prescribing costs (Baines 1997; Bradlow 1993; Burr 1992; Corney 1997; Harris 1996; Kammerling 1996; Rafferty 1997; Whynes 1997; Wilson 1995; Wilson 1999). Two studies were based in Germany and investigated the effects of a German drug budget program on prescriptions and referrals (Guether 1997; Schoffski 1997) . One study was based in Ireland and investigated the effectiveness of Irish indicative drug budgets in decreasing drug prescribing costs (Walley 2000) .

See Appendix 5 for details on the bibliometric analysis.

Summary across reviews

The included reviews reported results from 32 studies, of which three were included in more than one review (Davidson 1992; Kammerling 1996; Kouides 1998). Study designs included: 10 randomised controlled trials (RCTs), 18 controlled before and after studies (CBAs), and four interrupted time series (ITSs). Eleven studies were based in the UK, six in the USA, two in Germany, and one each in Canada, Finland, Denmark, and Ireland. In Petersen 2006, the country of origin was not stated for nine of the included studies.

Population and settings

The target populations of the reviews included primary care physicians (Akbari 2008; Gosden 2001; Sturm 2007), paediatricians (Akbari 2008; Petersen 2006), specialist physicians (Akbari 2008), and other unspecified healthcare providers (working in nursing homes, mental health centres, pharmacies etc.) (Petersen 2006). The settings were mostly primary care clinics (more than 1316 clinics). Other settings included were: 200 pharmacies, 32 skilled nursing facilities, 7 community mental health centres, and 5 health authorities. For some studies included in the reviews, the setting was unclear (Fairbrother 1999: Fairbrother 2001; Schoffski 1997).

Financial incentive interventions

i) Payment for working for a specified time period

One review (Gosden 2001) reported data from one study (Hickson 1987). Hickson 1987 compared the effects of salary reimbursement with fee-for-service (control) on improved continuity of patient care.

ii) Payment for each service, episode or visit

Three reviews (Akbari 2008; Gosden 2001; Petersen 2006) reported data from one, one, and four studies respectively (five studies in total).

Two reviews (Akbari 2008; Gosden 2001) reported data from one study (Davidson 1992). Davidson 1992 compared the effects of paying physicians a high fee-for-service with a low fee-for-service (control) on the number of primary care visits, non-primary care visits, and clinic or emergency department visits.

One review (Petersen 2006) reported four studies that compared the effects of enhanced fee-for-service with control on community treatment time (Clark 1995), the delivery of cognitive services by pharmacists (Christensen 2000), and paediatric immunisation rates (Fairbrother 1999; Fairbrother 2001).

iii) Payment for providing care for a patient or specific population

Three reviews (Akbari 2008; Gosden 2001; Sturm 2007) reported data from 3, 1, and 10 papers respectively (13 papers in total).

One review (Akbari 2008) reported three studies that evaluated a capitation-based payment system (Coulter 1993; Kammerling 1996; Linnala 2001). One of the studies compared the effects on referral rates of providing general physicians (GPs) with a list system and the opportunity to send their patients to private specialists for consultation at a reduced cost for the patient with GPs without a list system (Linnala 2001). Two studies compared the effects of fundholding on referral rates (Coulter 1993; Kammerling 1996) .

One review (Sturm 2007) reported 10 studies (Baines 1997; Bradlow 1993; Burr 1992; Corney 1997; Harris 1996; Kammerling 1996; Rafferty 1997; Whynes 1997; Wilson 1995; Wilson 1999) that evaluated the effects of the UK fundholding system in fundholding practices (compared with non-fundholding practices) on drug use; eight studies reported drug costs (Baines 1997; Bradlow 1993; Burr 1992; Harris 1996; Rafferty 1997; Whynes 1997; Wilson 1995; Wilson 1999); and one referrals (Kammerling 1996).

iv) Payment for providing a pre-specified level or providing a change in activity or quality of care

One review (Petersen 2006) reported 10 studies.

Of the 10 papers included in the review by Petersen 2006, six compared the effects of bonus payments with control for achieving a pre-specified level of quality of care: for diabetes care (Beaulieu

2005); referral rates for mammography (Grady 1997); guideline compliance (Hillman 1998; Hillman 1999); smoking identification and delivery of quitting advice (Roski 2003); and on cervical cancer screening rates, mammography screening rates, and haemoglobin A1c testing (Rosenthal 2005).

Two studies compared the effectiveness of bonus payments with an unspecified control condition on improving immunisation status (Fairbrother 1999; Fairbrother 2001).

One study compared the effects of performance-based contracting with additional funds (level of funding depending on efficiency, effectiveness, and service) with Medicaid payment on health service delivery to special populations in skilled nursing facilities to 5532 clients (Shen 2003). Just under half of these clients (n = 2367) were being treated for substance abuse and the remaining clients received Medicaid.

v) Mixed or other system

Three reviews (Akbari 2008; Gosden 2001; Sturm 2007) reported one, five, and three studies respectively (eight studies in total) that evaluated the effects of mixed or other incentive systems.

Two reviews (Akbari 2008; Gosden 2001) reported one study (Davidson 1992) that compared the effects of paying physicians by capitation (with some degree of financial risk sharing) with low fee-for-service (control) on the number of primary care and non-primary care visits, and the number of clinic or emergency room visits (latter results presented only in Gosden 2001).

In one review (Gosden 2001), Hutchison 1996 compared the effects of a mixed capitation system on the admission rate to hospital and length of stay with fee-for-service payment (control). Kouides 1998 and Ritchie 1992 evaluated the effects on immunisation rates of introducing target payments into a fee-for-service system. Krasnik 1990 compared the impact of introducing fee-for-service on the rate of primary care consultations, delivery of diagnostic and curative services, and renewal of prescriptions into a capitation system with physicians already paid by capitation and fee-for-service (control).

One review (Sturm 2007) reported three studies. Two of these studies compared the effects of German drug budgets on the number of prescribed items per patient (Guether 1997) and on referrals (Guether 1997; Schoffski 1997). The third study compared the effects of Irish Indicative Drug Budgets on 'cost per item' and 'total prescribing cost' (Walley 2000).

Frequency and level of financial incentive

The frequency and level of the financial incentives were described for the 13 studies included in one review (Petersen 2006) but were not systematically described for the studies included in the other three reviews.

Comparisons

Table 3 shows the comparisons made in the different studies.

The types of stated comparisons were: i) salary reimbursement versus fee-for-service (Hickson 1987); (ii) enhanced fee-for-service versus fee-for-service (Christensen 2000; Clark 1995; Davidson 1992) or versus unspecified control (Fairbrother 1999; Fairbrother 2001); iii) fundholding (capitation) versus non-fundholding (control) in 12 studies (Baines 1997; Bradlow 1993; Burr 1992; Corney 1997; Coulter 1993; Harris 1996; Kammerling 1996; Linnala 2001; Rafferty 1997; Whynes 1997; Wilson 1995; Wilson 1999); iv) target payments and bonuses versus control (comparison intervention not specified) in nine studies (Beaulieu 2005; Fairbrother 1999; Fairbrother 2001; Grady 1997; Hillman 1998; Hillman 1999; Rosenthal 2005; Roski 2003; Shen 2003), and versus fee-for-service in two studies (Kouides 1998; Ritchie 1992); and v) mixed incentives (five studies). One study compared capitation plus financial risk sharing with low fee-for-service (Davidson 1992), a second study compared a mix of capitation and an ambulatory incentive payment with fee-for-service (Hutchison 1996), and a third study (Krasnik 1990) compared partial fee-for-service added to a capitation system with capitation and fee-for-service throughout. The last two papers compared the effects of introducing target payments into a fee-for-service system with fee-for-service only (Kouides 1998; Ritchie 1992). For some of the studies the control intervention was not adequately described but fee-for-service was presumed (Christensen 2000; Clark 1995; Fairbrother 1999; Fairbrother 2001).

Outcomes

Table 4 shows the main outcomes of the included reviews. We reported only outcomes for which numerical data was presented. We grouped outcomes into: consultation or visit rates; processes of care; referrals and admissions; compliance with guidelines; and prescribing costs.

Two reviews (Akbari 2008; Gosden 2001) included three studies that assessed the effects of financial incentives on consultation rates (primary and non-primary care consultations, clinic and emergency department visits). Three reviews (Gosden 2001; Petersen 2006; Sturm 2007) and 2, 9, and 10 studies respectively (21 studies in total) assessed the effects on processes of care (for example diagnostic and curative services, prescriptions, healthcare utilisation etc.). All four reviews (Akbari 2008; Gosden 2001; Petersen 2006; Sturm 2007) (three, one, three, and three studies respectively, 10 studies in total) assessed referral rates (for example referral to hospital, outpatient clinics, specialists, and to mammography). Two reviews (Gosden 2001; Petersen 2006) (five studies in total) assessed the effects on compliance with guidelines (that is guidelines for cancer screening, paediatric immunisations, and well-child visits). One review (Sturm 2007) included 10 studies that assessed effects of financial incentives on costs related to prescribing of drugs (that is cost per item, cost per patient, and total prescribing cost).

Methodological quality of included reviews

Quality of included reviews

The AMSTAR scores are presented in Table 5.

Two reviews scored 7 according to the AMSTAR tool and were judged to be of moderate (score 5 to 7) methodological quality (Gosden 2001; Petersen 2006). Two reviews scored 9 and were judged to be of high (score 8 to 11) quality (Akbari 2008; Sturm 2007).

Two Cochrane reviews (Akbari 2008; Sturm 2007) provided an 'a priori' design since previously published protocols reported the research question as well as the inclusion criteria. . For the two non-Cochrane reviews, this was not the case.

All four reviews reported duplicate study selection and data extraction, and we performed a comprehensive literature search for all except one review (Petersen 2006). Petersen 2006 searched only one electronic source (MEDLINE) but also searched other sources.

Only Sturm 2007 clearly stated that their search strategy had been developed without language restriction and that the grey literature had been searched. Petersen 2006 searched only English literature, and for the remaining two reviews the search strategy was unclear (Akbari 2008; Gosden 2001).

All reviews provided a list of included and excluded studies, and described the characteristics of included studies.

All reviews assessed and documented the scientific quality of studies and used it, at least to some extent, in formulating conclusions. However, the methods used to assess the methodological quality of studies varied somewhat between reviews.

Three reviews discussed and chose appropriate methods for combining the findings. In one review, the authors did not discuss the methods and did not assess heterogeneity or report the results from such a test (Sturm 2007).

None of the reviews attempted to assess publication bias, and in only one of the reviews was the issue of possible publication bias brought up in the discussion (Petersen 2006).

All reviews except one (Gosden 2001) stated if there was a conflict of interest and what, if any, sources of funding they had received.

Quality of evidence in included reviews

The tools used to assess the quality of included papers in the four included reviews were as follows. Sturm 2007 reported using the GRADE system (GRADE 2004);Petersen 2006 used a quality assessment checklist (Downs 1998); Gosden 2001 assessed the methodological quality according to 'pre-determined criteria' and referenced two studies as their methodological sources (Bero 1999;Cook 1979;); and Akbari 2008 used an updated Cochrane Effective Practice and Organisation of Care Group (EPOC) tool (Bero 2008).

Sturm 2007 reported serious limitations for all included CBA studies (Baines 1997; Bradlow 1993; Burr 1992; Corney 1997; Kammerling 1996; Rafferty 1997; Whynes 1997; Wilson 1995; Wilson 1999) and the ITS studies were judged to have some limitations (Guether 1997; Schoffski 1997; Walley 2000) as were the three controlled ITS studies (Harris 1996; Rafferty 1997; Wilson 1995).

Petersen 2006 reported the poor quality (score 1) of one study (Beaulieu 2005); six papers scored 2 (Christensen 2000; Clark 1995; Grady 1997; Rosenthal 2005; Roski 2003; Shen 2003), and five papers scored 3 (Fairbrother 1999; Fairbrother 2001; Hillman 1998; Hillman 1999; Kouides 1998). None of the included papers scored 4 (excellent quality). Details of the quality scores and how they were assessed were not provided in the review.

Gosden 2001 reported that two of the included trials (Davidson 1992; Kouides 1998) had unit of analysis errors and that it was unclear if the two included CBA studies (Hutchison 1996; Krasnik 1990) had unit of analysis errors too. In Davidson 1992, there was a possible risk of detection bias since the visits for the capitation group were self-reported. The review authors did not provide a summary of the risk of bias of included studies. In Krasnik 1990, there was a possible risk of selection bias since GPs volunteered to participate in the intervention group, as well as a risk of detection bias since the activity data were obtained in different ways in the intervention and control groups. Ritchie 1992 did not have a concurrent control group. No distribution statistics were provided for the intervention or control group in Hickson 1987, and for the control group in Krasnik 1990.

Akbari 2008 reported that all four included papers (Coulter 1993; Davidson 1992; Kammerling 1996; Linnala 2001) had a unit of analysis error. Of these papers, the three CBAs were judged to be of low quality and the only RCT (Davidson 1992) had a high risk of bias.

Effect of interventions

The results of the intervention effects of different types of financial incentives are presented in Appendix 6 and summarised in Table 6.

Payment for working for a specified time period

Payment for working for a specified time period was generally ineffective, improving 3/11 outcomes from one study reported in one review. One review (Gosden 2001), reporting one randomised controlled trial (Hickson 1987), reported that salary reimbursement improved 3/9 reported consultation or visit rate outcomes and 0/2 compliance with guidelines outcomes. Statistical significance was reported for 11/11 outcomes and reported as significant for 8/11, two of which favoured the intervention.

Payment for each service, episode or visit

Payment for each service, episode or visit was generally effective, improving 7/10 outcomes from five studies reported in three reviews. Three reviews (Akbari 2008; Gosden 2001; Petersen 2006) reported results from one, one, and four studies respectively (five studies in total, four of which were RCTs). These reported that an enhanced fee-for-service improved 2/2 consultation or visit rate outcomes, 4/6 processes of care outcomes, and 0/1 for a referrals and admissions outcome. Statistical significance was reported for 6/10 outcomes and was reported as significant for 4/10, four of which favoured the intervention.

Payment for providing care for a patient or specific population

Payment for providing care for a patient or specific population was generally effective, improving 48/69 outcomes from 13 studies reported in two reviews. Two reviews (Akbari 2008; Sturm 2007) reported results from three and 10 studies respectively (13 studies in total, all non-randomised). These reported that capitation-based payment systems improved 17/30 processes of care outcomes (all drug prescribing related), 4/6 referrals and admissions outcomes, and 28/34 prescribing costs outcomes. Statistical significance was reported for 2/70 outcomes, one of which favoured the intervention.

Payment for providing a pre-specified level or providing a change in activity or quality of care

Payment for providing a pre-specified level or providing a change in activity or quality of care was generally effective, improving 17/20 reported outcomes from 10 studies reported in two reviews. Two reviews (Gosden 2001; Petersen 2006) reported results from one and 10 studies respectively, 10 studies in total, eight of

which were RCTs. These reported that target payments or bonuses improved 16/16 processes of care outcomes, 1/2 admissions and referrals outcomes, and 0/2 compliance with guidelines outcomes. Statistical significance was reported for 20/20 outcomes and was reported as significant for 14/20 outcomes, 14 of which favoured the intervention.

Mixed and other systems

Mixed and other systems were of mixed effectiveness, improving 20/31 reported outcomes from seven studies reported in three reviews. Three reviews (Akbari 2008; Gosden 2001;Sturm 2007) reported results from one, four, and three studies respectively, seven studies in total, one of which was a RCT. These reported an improvement in 4/5 consultation or visit rate outcomes, 4/5 processes of care outcomes, 7/8 referrals and admissions outcomes, and 5/13 compliance with guidelines outcomes. Statistical significance was reported for 2/31 outcomes, and was reported as significant for 0/31 outcomes, none of which favoured the intervention.

Across categories of intervention

When looking at the effect of financial incentives overall, across categories of outcomes, the incentives were of mixed effectiveness on consultation or visit rates (improving 10/17 outcomes from three studies in two reviews), generally effective in improving processes of care (improving 41/57 outcomes from 19 studies in three reviews), generally effective in improving referrals and admissions (improving 11/16 outcomes from 11 studies in four reviews), generally ineffective in improving compliance with guidelines outcomes (improving 5/17 outcomes from five studies in two reviews), and generally effective in improving prescribing costs outcomes (improving 28/34 outcomes from 10 studies in one review).

DISCUSSION

We identified four reviews (reporting 32 studies) investigating the effectiveness of various configurations of financial incentives in changing both the behaviour of health professionals and patient outcomes. The methodological quality of these reviews was moderate to high but their rating of the quality of the studies they included was low to moderate at best.

Summary of main results

Overall, payment for each service, episode or visit; payment for providing care for a patient or specific population; and payment for providing a pre-specified level of care or providing a change in activity or quality of care were generally effective. Mixed or other system interventions produced mixed results and payment for working for a specified time period was generally ineffective. Looking at different groups of outcomes, financial incentives were generally effective at improving processes of care, referrals and admissions, and prescribing costs; they had mixed effects on consultation or visit rates and were generally ineffective in promoting compliance with guidelines.

Overall completeness and applicability of evidence

Whilst we found evidence within each of our categories of financial incentive, there were only a small number of studies within any one category. We also found no studies evaluating patient outcomes and none reporting on any adverse or unexpected outcomes. We found no reviews of studies that systematically examined variable doses of financial intervention. The 'doses' that were evaluated were not adequately described to enable comparison of the effect of differences in the magnitude of the incentives across studies. All of these factors mean that this is a very incomplete body of evidence.

For a majority of studies the comparison intervention was not clearly stated, compromising a reader's ability to understand the context within which the study was conducted and therefore how it might translate to another setting.

Categorising financial incentive interventions was not straightforward. Although we started with a set of a priori groups, when categorising financial interventions it was not clear whether it is best to rely on the original study authors' description and categorisation or that of the review author. Whichever was used, there was still a degree of judgement on the behalf of the overview authors and it is not clear how homogenous the contents of our categories are. For example, GP fundholding in the UK devolved a budget to GPs, calculated on the basis of their population, to cover hospital admissions and prescribing costs. However, these funds could not be directly used as personal income for GPs though 'savings' could be re-invested into patient care. So financial incentives did exist through the ability to make savings or to be faced with losses but this was not a method of personal remuneration for doctors.

We had to group outcomes in order to be able to make sense of the volume of data available. This process was data driven and produced five groups of unequal size though, with the exception of prescribing costs, they all drew on at least two reviews for studies. All four reviews contributed data to one category of outcome and for the other four categories the number of reviews contributing data were three, two, two, and one. With the mainly positive effect of financial incentives on processes of care it is a little difficult to explain the absence of any effect on guideline compliance given that guideline compliance is likely to be composed of (multiple measures of) processes of care. The absence of effect may be due to the methods of calculating compliance (often a composite measure). Equally, it may be that the guideline compliance studies were explicitly referenced against an evidence-based standard of care and that, although measuring changes in processes of care, these other studies may not reflect an improvement in evidence-based care.

Of the 23 studies that reported where the study was conducted, all were conducted in high-income countries (17 in the UK or USA), which may make it difficult to generalise these findings to other high-income countries or the context of any low or middle-income country. Also, none of the studies appeared to be conducted in secondary care settings, limiting the generalisability of the findings to only primary care settings.

Of the 32 original studies included in the four reviews, 25 were performed between 1987 and 1999; only eight studies were performed between 2000 and 2005. Thus, all of the evidence is at least five years old and 75% of it at least 10 years old. It is unclear to what extent the preferences of different cohorts of health professionals towards the importance of financial motivation versus other sources of motivation are likely to change over time.

Quality of the evidence

These results have to be treated with considerable caution. Whilst the methodological quality of the four included reviews was moderate to high, their rating of the quality of the studies they included was low to moderate. The reporting of data within both the reviews and the included studies was poor and precluded a quantitative aggregation, making it impossible to follow our original plan to perform meta-analysis. This forced us to use vote counting to summarise the studies. Thus, although we can report the direction of effect of the reported outcomes we can not say anything of their statistical significance.

Within the categories of financial incentives the most studied was payment for providing care for a patient or specific population, with 13 studies contributing half of the overall reported outcomes. However, this group is dominated by 10 CBA studies (all of which were reported as having serious limitations by the original review author) of a single intervention, the UK NHS fundholding system. As an example of the instability of the results, removing these studies moves this type of financial incentive to mixed effectiveness (based only on improving one of two reported outcomes from one study). The effect of financial incentives on processes of care and referrals and admissions are left unchanged, as generally effective, but all of the evidence for prescribing costs is removed.

Potential biases in the overview process

Two review authors independently applied eligibility criteria and assessed the studies for inclusion, extracted data, and assessed the scientific quality of reviews according to the AMSTAR, which should reduce the risk of bias in the overview process.

The risk of publication bias could not be formally assessed in this overview due to too few included reviews, nor was it assessed in any of the included reviews. This overview may be biased by excluding reviews that did not report quantitative data. Given the sequence of decisions that were made about reviews we decided to only go back to original studies reported in reviews that used quantitative summary methods. We omitted studies that were only included in narrative reviews.

There is the possibility of data reporting errors in the original reviews. Of the eight errors (all corrected) identified by the statistical reviewer, seven of them were accurate transcriptions of the data reported in the original systematic reviews. However, correcting these did not make any difference to the conclusions of this overview.

Agreements and disagreements with other studies or reviews

Whilst there are many reviews of financial incentives, our process of article selection excluded most of them from consideration for inclusion within this overview. We are not aware of any other published overviews of reviews of financial incentives.

AUTHORS' CONCLUSIONS

Four reviews reporting low to moderate quality evidence suggest that financial incentives may be effective in changing healthcare professional practice. The evidence has serious methodological limitations and is also very limited in its completeness and generalisability. We found no evidence that examined the effect of financial incentives on patient outcomes.

Implications for practice

Contrary to popular opinion, there is limited evidence supporting the implementation of financial incentives for changing healthcare professional behaviour. If financial incentives are used as a behaviour change strategy then this should be done in the context of a rigorous evaluation.

Implications for research

Whilst the included reviews were well performed, they still did not report all of the quantitative data that would have allowed a meta-analysis. On some occasions data were available from the primary studies. All relevant data should be reported within reviews.

The studies included in the four reviews did not appear to use robust designs, making it difficult to be confident about the nature of any cause and effect relationship. Future evaluations of financial incentives should make greater efforts to address the following.

- Use an experimental design; if this is not possible then the strongest quasi-experimental design should be used.
- Use an explicit theory-based design of the financial incentive scheme.
- Clearly specify the control group intervention.
- Examine the effect of differing doses of intervention.
- Perform an economic evaluation of the cost-effectiveness of financial incentives (e.g. Clarkson 2008).
- Conduct evaluations in contexts where selection bias is minimised.
- Report all relevant data for intervention and control groups.

ACKNOWLEDGEMENTS

We would like to thank Jeremy Grimshaw for advice about the methods of the overview.

REFERENCES

References to included reviews

Akbari 2008

Akbari A, Mayhew A, Al-Alawi MA, Grimshaw J, Winkens R, Glidewell E, et al. Interventions to improve outpatient referrals from primary care to secondary care. *Cochrane Database of Systematic Reviews* 2008, Issue 1. [DOI: 10.1002/14651858.CD005471.pub2]

Gosden 2001

Gosden T, Forland F, Kristiansen IS, Sutton M, Leese B, Giuffrida A, et al. Impact of payment method on behaviour of primary care physicians: a systematic review. *Journal of Health Services Research & Policy* 2001;**6**(1):44-55.

Petersen 2006

Petersen LA, Woodard LD, Urech T, Draw C, Sookanan S. Does pay-for-performance improve the quality of health care?. *Annals of Internal Medicine* 2006;**145**:265-72.

Sturm 2007

Sturm H, Austvoll-Dahlgren A, Aaserud M, Oxman AD, Ramsay C, Vernby A, et al. Pharmaceutical policies: effects of financial incentives for prescribers. *Cochrane Database of Systematic Reviews* 2007, Issue 3. [DOI: 10.1002/14651858.CD006731]

References to excluded reviews

Achat 1999

Achat H, McIntyre P, Burgess M. Health care incentives in immunisation. *Australian and New Zealand Journal of Public Health* 1999;**23**(3):285-8.

Alshamsan 2010

Alshamsan, R. Majeed A, Ashworth M, Car J, Millett C. Impact of pay for performance on inequalities in health care: systematic review. *Journal of Health Services Research & Policy* 2010;**15**(3):178-84.

Anderson 2009

Anderson P. Overview of interventions to enhance primary-care provider management of patients with substance-use disorders. *Drug and Alcohol Review* 2009;**28**(5):178-84.

Andreae 2006

Andreae MC, Blad K, Cabana MD. Physician compensation programs in academic medical centers. *Health Care Management Review* 2006;**31**(3):251-8.

Armour 2003

Armour BS, Pitts MM. Physician financial incentives in managed care- resource use, quality and cost implications. *Managed Health Outcomes* 2003;**11**(3):139-41.

Barnighausen 2009

Barnighausen T, Bloom DE. Financial incentives for return of service in underserved areas: a systematic review. *BMC Health Service Research* 2009;**9**(86):1-47.

Chaix-Couturier 2000

Chaix-Couturier C, Durand-Zaleski I, Jolly D, Durieux P. Effects of financial incentives on medical practice:results from a systematic review of the literature and methodological issues. *International Journal for Quality in Health Care* 2000;**12**(2):133-42.

Christianson 2007

Christianson J, Leatherman S, Sutherland K. Financial incentives, healthcare providers and quality improvements- a review of the evidence. The Health Foundation 2007:7-105.

Christianson 2008

Christianson JB, Leatherman S, Sutherland K. Lessons from evaluations of purchaser pay-for-performance programs: a review of the evidence. *Medical Care Research Reviews* 2008;**65**:5S-35.

Dudley 1998

Dudley RA, Miller RH, Korenbrot TY, Luft HS. The impact of financial incentives on quality of health care. *The Milbank Quarterly* 1998;**76**(4):649-86.

Frolich 2007

Frolich A, Talavera JA, Broadhead P, Dudley RA. A behavioral model of clinician responses to incentives to improve quality. *Health Policy* 2007;**80**:179-93.

Giuffrida 1999

Giuffrida A, Gosden T, Forland F, Kristiansen IS, Sergison M, Leese B, et al. Target payments in primary care: effects on professional practice and health care outcomes. *Cochrane Database of Systematic Reviews* 1999, Issue 4. [DOI: 10.1002/14651858.CD000531]

Gosden 1997

Gosden T, Torgerson DJ. The effect of fundholding on prescribing and referral costs: a review of the evidence. *Health Policy* 1997;**40**:103-14.

Gosden 1999

Gosden T, Pedersen L, Torgerson D. How should we pay doctors? A systematic review of salary payments and their effect on doctor behaviour. *The Quarterly Journal of Medicine* 1999;**92**:47-55.

Gosden 2000

Gosden T, Forland F, Kristiansen IS, Sutton M, Leese B, Giuffrida A, et al. Capitation, salary, fee-for-service and mixed systems of payment: effects on the behaviour of primary care physicians. *Cochrane Database of Systematic Reviews* 2000 , Issue 3. [DOI: 10.1002/14651858.CD002215]

Johanson 2007

Johansson V, Axtelius B, Soderfelt B, Sampogna F, Lannerud M, Sondell K. Financial systems' impact on dental care: a review of fee-for-service and capitation systems. *Community Dental Health* 2007;**24**:12-20.

Lagarde 2006

Lagarde M, Palmer N. Evidence from systematic reviews to inform decision making regarding financing mechanisms that improve access to health services for poor people. Alliance for Health Policy and Systems Research 2006.

Mason 2008

Mason A. New medicines in primary care: a review of influences on general practitioner prescribing. *Journal of Clinical Pharmacy and Therapeutics* 2008;**33**:1-10.

McDonald 2008

McDonald J, Harris MF, Cumming J, Powell Davies G, Burns P. The implementation and impact of different funding initiatives on access to multidisciplinary primary health care and policy implications. *Medical Journal of Australia* 2008;**188 Suppl**(N88):569-72.

Rosenthal 2006

Rosenthal MB, Frank RG. What is the empirical basis for paying for quality in health care?. *Medical Care Research and Review* 2006;**63**:135-57.

Sabatino 2008

Sabatino SA, Habarta N, Baron RC, Coates RJ, Rimer BK, Kerner J, et al. Interventions to increase recommendation and delivery of screening for breast, cervical, and colorectal cancers by healthcare providers systematic reviews of provider assessment and feedback and provider incentives. *American Journal of Preventive Medicine* 2008;**35 Suppl**(N1):567-74.

Schatz 2008

Schatz M. Does pay-for-performance influence the quality of care?. *Current Opinion in Allergy and Clinical Immunology* 2008;**8**:213-21.

Scott 1995

Scott A, Hall J. Evaluating the effects of GP remuneration: problems and prospects. *Health Policy* 1995;**31**:183-95.

Town 2005

Town R, Kane R, Johnson P, Butler M. Economic incentives and physicians' delivery of preventive care: a systematic review. *American Journal of Preventive Medicine* 2005;**28**(2):234-40.

Van Herck 2010

Van Herck P, De Smedt D, Annemans L, Remmen R, Rosenthal MB, Sermeus W. Systematic review: Effects, design choices, and context of pay-for-performance in health care. *BMC Health Services Research* 2010;**10**:247.

Additional references

Baines 1997

Baines DL, Brigham P, Phillips DR, Tolley KH, Whynes DK. GP fundholding and prescribing in UK general practice: evidence from two rural, English Family Health Services Authorities. *Public Health* 1997;**111**(5):321-5.

Beaulieu 2005

Beaulieu ND, Horrigan DR. Putting smart money to work for quality improvement. *Health Services Research* 2005;**40**(5 Pt 1):1318-34.

Bero 1999

Bero L, Grilli R, Grimshaw JM, Mowatt G, Oxman AD, et al. The Cochrane Effective Practice and Organisation of Care Group (EPOC) module. The Cochrane Library 1999, issue 4.

Bero 2008

Bero L, Eccles M, Grilli R, Grimshaw J, Gruen RL, Mayhew A, et al. Cochrane Effective Practice and Organisation of Care Group. About The Cochrane Collaboration (Cochrane Review groups (CRGs)). Cochrane Effective Practice and Organisation of Care Group 2008, issue 2.

Bradlow 1993

Bradlow J, Coulter A. Effect of fundholding and indicative prescribing schemes on general practitioners' prescribing costs. *BMJ* 1993;**307**(6913):1186-9.

Burr 1992

Burr A, Walker R, Stent S. Impact of fundholding on general practice prescribing patterns. *Pharmaceutical Journal* 1992;**249 Suppl**:R8.

Christensen 2000

Christensen DB, Neil N, Fassett WE, Smith DH, Holmes G, Stergachis A. Frequency and characteristics of cognitive services provided in response to a financial incentive. *Journal of the American Pharmaceutical Association* 2000;**40**(5):609-17.

Clark 1995

Clark RE, Drake RE, McHugo GJ, Ackerson TH. Incentives for community treatment. Mental illness management services. *Medical Care* 1995;**33**(7):729-38.

Clarkson 2008

Clarkson JE, Turner S, Grimshaw JM, Ramsay CR, Johnston M, Scott A, et al. Changing clinicians' behavior: a randomized controlled trial of fees and education. *Journal of Dental Research* 2008;**87**:640-4.

Cook 1979

Cook TD, Campbell DT. Quasi-Experimentation: design and analysis issues for field settings. Wadsworths Publishing, 1979.

Corney 1997

Corney RH, Kerrison S. Fundholding in the south Thames Region. *British Journal of General Practice* 1997;**47**(422):553-6.

Coulter 1993

Coulter A, Bradlow J. Effect of NHS reforms on general practitioners' referral patterns. *BMJ* 1993;**306**(6875):433-7.

Davidson 1992

Davidson SM, Manheim LM, Werner SM, Hohlen MM, Yudkowsky BK, Fleming GV. Prepayment with office-based physicians in publicly funded programs: results from the Children's Medicaid Program. *Pediatrics* 1992;**89**(4 Pt 2):761-7.

Downs 1998

Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology and Community Health* 1998;**52**:377-84. [doi:10.1136/jech.52.6.377]

Ettner 2006

Ettner S, Schoenbaum M. The role of economic incentives in improving the quality of mental healthcare. London: The Elgar Companion to Health Economics, 2006:1-565.

Fairbrother 1999

Fairbrother G, Hanson KL, Friedman S, Butts GC. The impact of physician bonuses, enhanced fees, and feedback on childhood immunization coverage rates. *American Journal of Public Health* 1999;**89**(2):171-5.

Fairbrother 2001

Fairbrother G, Siegel MJ, Friedman S, Kory PD, Butts GC. Impact of financial incentives on documented immunization rates in the inner city: results of a randomized controlled trial. *Ambulatory Pediatrics* 2001;**1**(4):206-12.

Frey 1997

Frey BS. Not Just for the Money: An Economic Theory of Personal Motivation. Cheltenham, UK: Edgar Elgar Publishing, 1997.

GRADE 2004

GRADE working group. Grading quality of evidence and strength of recommendations. *BMJ* 2004;**328**(1490):1-8. [doi: 10.1136/bmj.328.7454.1490]

Grady 1997

Grady KE, Lemkau JP, Lee NR, Caddell C. Enhancing mammography referral in primary care. *Preventive Medicine* 1997;**26**(6):791-800.

Guether 1997

Guether B. [Effects of the Health Structure Regulation on prescribing behavior of established physicians]. *Gesundheitswesen* 1995;**57**(4):185-91.

Harris 1996

Harris CM, Scrivener G. Fundholders' prescribing costs: the first five years . *BMJ* 1996;**313**(7071):1531-4.

Hickson 1987

Hickson GB, Altemeier WA, Perrin JM. Physician reimbursement by salary or fee-for-service: effect on physician practice behavior in a randomized prospective study. *Pediatrics* 1987;**80**(3):344-50.

Hillman 1998

Hillman AL, Ripley K, Goldfarb N, Nuamah I, Weiner J, Lusk E. Physician financial incentives and feedback: failure to increase cancer screening in Medicaid managed care. *American Journal of Public Health* 1998;**88**(11):1699-701.

Hillman 1999

Hillman AL, Ripley K, Goldfarb N, Weiner J, Nuamah I, Lusk E. The use of physician financial incentives and feedback to improve pediatric preventive care in Medicaid managed care. *Pediatrics* 1999;**104**(4 Pt 1):931-5.

Hutchison 1996

Hutchison B, Birch S, Hurley J, Lomas J, Stratford-Devai F. Do physician-payment mechanisms affect hospital utilization? A study of health service organizations in Ontario. *CMAJ* 1996;**154**(5):653-61.

Kammerling 1996

Kammerling RM, Kinnear A. The extent of the two tier service for fundholders. *BMJ* 1996;**312**(7043):1399-401.

Kouides 1998

Kouides RW, Bennett NM, Lewis B, Cappuccio JD, Barker WH, LaForce FM. Performance-based physician reimbursement and influenza immunization rates in the elderly. The Primary-Care Physicians of Monroe County. *American Journal of Preventive Medicine* 1998;**14**(2):89-95.

Krasnik 1990

Krasnik A, Groenewegen PP, Pedersen PA, von Scholten P, Mooney G, Gottschau A, et al. Changing remuneration systems: effects on activity in general practice. *BMJ* 1990;**300**(6741):1698-701.

Linnala 2001

Linnala A, Aromaa A, Mattila K. Specialists as consultants to GPs. Private sector services as an alternative way of organising consultant services in health care. *Scandinavian Journal of Primary Health Care* 2001;**19**(2):90-4.

Moher 2009

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine* 2009;**6**(6):e1000097. doi:10.1371/journal.pmed1000097.

Mooney 1993

Mooney G, Ryan M. Agency in health care:getting beyond the first principles. *Journal of Health Economics* 1993;**12**(2):125-35.

NICE 2004

National Institute of Health and Clinical Excellence. Quality of Outcome Framework. http://www.nice.org.uk/aboutnice/qof/qof.jsp 2004.

Rafferty 1997

Rafferty T, Wilson-Davis K, McGavock H. How has fundholding in Northern Ireland affected prescribing patterns? A longitudinal study. *BMJ* 1997;**315**(7101):166-70.

Ritchie 1992

Ritchie LD, Bisset AF, Russell D, Leslie V, Thomson I. Primary and preschool immunisation in Grampian: progress and the 1990 contract. *BMJ* 1992;**304**(6830):816-9.

Robinson 2001

Robinson JC. Theory and practice in the design of physician payment incentives. *The Milbank Quarterly* 2001;**79**(2):149-77.

Roland 2004

Roland M. Linking physicians' pay to the quality of care- a major experiment in the United Kingdom. *The New England Journal of Medicine* 2004;**351**(14):1448-54.

Rosenthal 2005

Rosenthal MB, Frank RG, Li Z, Epstein AM. Early experience with pay-for-performance: from concept to practice. *JAMA* 2005;**294**(14):1788-93.

Roski 2003

Roski J, Jeddeloh R, An L, Lando H, Hannan P, Hall C, et al. The impact of financial incentives and a patient registry on preventive care quality: increasing provider adherence to evidence-based smoking cessation practice guidelines. *Preventive Medicine* 2003;**36**(3):291-9.

Schoffski 1997

Schoffski O, Graf von der Schulenburg JM. Unintended effects of a cost-containment policy: results of a natural experiment in Germany. *Social Science & Medicine* 1997;**45**(10):1537-9.

Shea 2007

Shea B, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology* 2007;**7**(10):1-7.

Shen 2003

Shen Y. Selection incentives in a performance-based contracting system. *Health Services Research* 2003;**38**(2):535-52.

Walley 2000

Walley T, Murphy M, Codd M, Johnston Z, Quirke T. Effects of a monetary incentive on primary care prescribing in Ireland: changes in prescribing patterns in one health board 1990-1995. *Pharmacoepidemiology & Drug Safety* 2000;**9**(7):591-8.

Weir 2010

Weir MC, Ryan R, Mayhew A, Worswick J, Santesso N, Lowe D, et al. The Rx for Change database: a first-in-class tool for optimal prescribing and medicines use. *Implementation Science* 2010;**5**(89):1-22.

WHO 2006

World Health Organization. Quality of care - a process for making strategic choices in health systems. Geneva: WHO press, 2006.

Whynes 1997

Whynes DK, Baines DL, Tolley KH. GP fundholding and the costs of prescribing: further results. *Journal of Public Health Medicine* 1997;**19**(1):18-22.

Wilson 1995

Wilson RP, Buchan I, Walley T. Alterations in prescribing by general practitioner fundholders: an observational study. *BMJ* 1995;**311**(7016):1347-50.

Wilson 1999

Wilson RP, Hatcher J, Barton S, Walley T. Therapeutic substitution and therapeutic conservatism as cost-containment strategies in primary care: a study of fundholders and non-fundholders. *British Journal of General Practice* 1999;**49**(443):431-5.

ADDITIONAL TABLES

Table 1. Excluded studies

Reviews	Reasons for exclusion
Achat 1999	no numerical data reported or the data reported were incomplete in that tests of significance were reported without numerical outcome data
Alshamsan 2010	does not report data on any of the pre-specified outcomes e.g. health professionals' clinical behaviour such as rates of performing specified clinical behaviours; measures of health professional non-clinical behaviour such as rates of performing specified non-clinical behaviours; measures of health service utilisation by patients such as participation rates in immunisation schemes or mammography screening programs; health care costs; measures of patient outcomes either objectively measured (e.g. mortality) or patient reported (e.g. quality of life)
Anderson 2009	no numerical data reported
Andreae 2006	no numerical data reported or the data reported were incomplete in that tests of significance were reported without numerical outcome data
Armour 2003	data were not reported at an individual study basis
Barnighausen 2009	no outcomes related to quality or efficiency of healthcare provided by healthcare professionals

Table 1. Excluded studies *(Continued)*

Chaix-Couturier 2000	no numerical data reported or the data reported were incomplete in that tests of significance were reported without numerical outcome data
Christianson 2007	no numerical data reported or the data reported were incomplete in that tests of significance were reported without numerical outcome data
Christianson 2008	no numerical data reported or the data reported were incomplete in that tests of significance were reported without numerical outcome data
Dudley 1998	data were not reported at an individual study basis
Frolich 2007	data were not reported at an individual study basis
Giuffrida 1999	reported data from the same studies as two of the included reviews, but with less detail
Gosden 1997	no numerical data reported or the data reported were incomplete in that tests of significance were reported without numerical outcome data
Gosden 1999	no numerical data reported or the data reported were incomplete in that tests of significance were reported without numerical outcome data
Gosden 2000	reported data from the same studies as two of the included reviews, but with less detail
Johanson 2007	data were not reported at an individual study basis
Lagarde 2006	data were not reported at an individual study basis
Mason 2008	data were not reported at an individual study basis
McDonald 2008	no numerical data reported, selected countries only
Rosenthal 2006	no numerical data reported or the data reported were incomplete in that tests of significance were reported without numerical outcome data
Sabatino 2008	data are only reported in terms of a % change
Schatz 2008	no numerical data reported or the data reported were incomplete in that tests of significance were reported without numerical outcome data
Scott 1995	no numerical data reported or the data reported were incomplete in that tests of significance were reported without numerical outcome data
Town 2005	no numerical data reported or the data reported were incomplete in that tests of significance were reported without numerical outcome data
Van Herck 2010	no numerical data reported

Table 2. Characteristics of included reviews

First author, year	Stated aim of review	Details of search (databases+ years of start/ finish)	Number and designs of studies included in the review	Number and type of FI intervention	Number and type of comparison interventions	Number of studies performed in different countries	Number and type of participants in included studies	Number and type of settings in included studies
Akbari 2008	To estimate the effectiveness and efficiency of interventions to change outpatient referral rates or improve outpatient referral appropriateness.	Electronic searches (EPOC) group specialised register (February 2002) and the National Research Register. Updated searches were conducted in MEDLINE and the EPOC specialised register up to October 2007	**Cluster RCT (n=1)** Davidson 1992 **CBA (n=3)** Coulter 1993; Kammerling 1996; Linnala 2001	**Payment for providing care for a patient or specific population (capitation)(n=4)** Coulter 1993; Kammerling 1996 - *UK Fundholding scheme*: The fundholding scheme gives GPs control over budgets to cover prescriptions, specialist outpatients consultations, and elective surgical procedures for their patients *Linnala 2001* - Examined the impact of charging patients the same (lesser) rate to be seen by a private specialist as they would have been charged to see a hospital based specialist **Payment for each service/episode/case (fee-for-service)** *Davidson 1992* - PCPs were paid a fee for comprehensive exams (including treatment), routine office visits, initial and follow-up hospital visits. Fee-for-service (high rate): PCPs were paid a fee for comprehensive exams (including treatment), routine office visits, initial and follow-up visits. Low fee-for-service (control): PCPs paid fees for same services as the high fee-for-service group, but the fees were approximately half size	**Non Fund Holding (n=2)** Coulter 1993; Kammerling 1996 **Low Fee-for-service (n=1)** Davidson 1992 **Unclear patient incentives (n=1)** Linnala 2001	**UK (n=2**) Coulter1993; Kammerling 1996 **US (n=1)** Davidson 1992 **Finland (n=1)** Linnala 2001	**Primary care physicians (n=3)** unclear number in Coulter 1993 and Kammerling 1996 14 GPs in Linnala **Private office based physicians (n=1):** 80 physicians in Davidson 1992 who treated Medicaid children and more than $2000 in Medicaid billings in previous year.	**Primary care practices** n= 16 PCPs in Coulter 1993 (I:10 Fund-holding vs.C:6 non-fund-holding) n=32 PCPs in Kammerling 1996 (I:10FH vs. C:22 NFH) **Private office based practices** unclear number in Davidson 1992 **Municipal health services** n= 4 in Linnala 2001

Table 2. Characteristics of included reviews *(Continued)*

				Mixed incentives *Davidson 1992* - Change in remuneration system from a low cost fee-for-service system to a capitation-based budgetary system (with some degree of risk sharing by the provider for secondary care provision) for the management of Medicaid eligible paediatric care. Capitation + financial risk sharing: PCPs at risk for deficits on referral budget up to a maximum of i) $2000 per child per year; (ii) 25% of total annual capitation payments. PCPs received 40% of budget surplus.			
Gosden 2001	To review the impact of payment systems on the behaviour of primary care physicians.	(1966-1997), BIDS EMBASE (1980-1997), BIDS ISI Social Science Citation Index (1981-1997), ECONLIT (1969-1997), HealthSTAR (1975-1997), HELMIS (1984-1997), the EPOC Register, Psyclit (1987-1997)	**RCT (n=3)** Davidson 1992; Hickson 1987; Kouides 1998 **ITS (n=1)** Ritchie 1992 **CBA (n=2)** Hutchison 1996; Krasnik 1990	**Payment for working for a specified time period (e.g. salary, sessional payment)** Hickson 1987 - Salary group received $20 per month and the control group (i.e. the fee-for-service group) received $2 per visit. **Payment for each service/episode/case (fee-for-service)** *Davidson 1992* - See description under Akbari 2008 *Krasnik 1990* - GPs working inside Copenhagen city were paid by capitation by October 1987.After this date GPs fees were introduced for face to face telephone and home visit consultations, and repeat prescriptions. Additional fees are payable for 40 special services (e.g. cervical smear tests), for 40 special labora-	**Fee-for-service (n=5)** Kouides 1998; Ritchie 1992; Hutchison 1994; Davidson 1992; Hickson 1987 **Fee-for-service / Capitation (n=1)** Krasnik 1990	**USA (n=3)** Davidson 1992;Hickson 1987; Kouides 1998 **Canada (n=1)** Hutchison 1996 **Denmark (n=1)** Krasnik 1990 **UK (n=1)** Ritchie 1992	**Primary care physicians** (n=4 studies) (n=935) 1 study Unknown number of PCPs Kouides 1998 1 study Paediatric residents (n=18) Hickson 1987	**General practices** (n=2 studies) (n=54) Koudes 1998 (n= 95) Ritchie 1992 (n=4) studies number unknown

Table 2. Characteristics of included reviews *(Continued)*

		and the Cochrane Controlled Trials Register (CCTR)		tory investigations performed in the practice (e.g. haemoglobin concentration) and for a few preventive services (e.g. immunisations). **Mixed incentives** Davidson 1992 - See description under Akbari 2008 Hutchisson 1996 - PCPs who were previously paid by fee-for-service changed their payment to a mix of capitation and an ambulatory incentive payment in which the health service organisation received a payment if their hospitalisation rate was lower than the regional rate Kouides 1998 - PCPs in the intervention group received an additional 10% ($0.8) or 20% ($1.6) reimbursement per shot according to whether they immunised 70% or 85%(respectively) of the eligible population Ritchie 1992 - PCPs received a lower or higher payment according to whether they immunised 70% or 90%(respectively) of the eligible population				
Petersen 2006	To assess the relationship between explicit financial incentives and the pro-	PubMed 1 January 1980 - 14 November 2005	**RCT (n=8)** Christensen 2000; Fairbrother 1999; Fairbrother 2001;	**Payment for each service/episode/case (fee-for-service) (n=2)** *Christensen 2000 -* Pharmacists received $4 for cognitive service interventions (<6 min); $6 for >6 min; cognitive services are judgemental or educational services provided by the pharma-	**Medicaid funded Office of Substance Abuse** Shen 2003[Z2]	**US (n=3)** Beaulieu 2005; Rosenthal 2005; Shen 2003	**Primary care Physicians (n=3 studies)** *Fairbrother 1999 -*	**Primary care practices (n=6)** *Beaulieu 2005 -* 21 PCPs contracted with Independent health in upstate New York *Grady 1997 -*

Table 2. Characteristics of included reviews *(Continued)*

	vision of high-quality health care by systematically reviewing empirical studies.	Grady 1997; Hillman 1998; Hillman 1999; Kouides 1998; Roski 2003 (one paper did not report any numerical data:Norton 1992) **CBA (n=4)** Beaulieu 2005; Clark 1995; Rosenthal 2005; Shen 2003 **Other**[Z1] **(n=4)** Cross-sectional surveys (Casalino 2003; McMenamin 2003; Pourat 2005; Safran 2000) - not included in this review	cist to the patient, such as consulting the prescriber about a sub optimal dose *Clark 1995 -* Community mental health centres received $ 15.75 per 15 min spent in community setting delivering mental illness management services. *Fairbrother 1999-* Enhanced Fee-for-service: $5 per vaccine given within 30 d of its coming due; $15 for each visit at which >1 vaccine was due and all were given. *Fairbrother 2001 -* Enhanced Fee-for-service: $5 per vaccine given within 30 d of its coming due; $15 for each visit at which >1 vaccine was due and all were given. **Payment for providing a pre-specified level of quality of care (include target payments, bonuses) (n=11)** *Beaulieu 2005 -* Meeting target CS of >6.23; CS of >6.86; or overall improvement in composite score. CS based on PCPs performance of process and outcome measures for diabetes care (e.g. LDL test, dilated retinal examination; LDL cholesterol level<2.59 mmol/L (<100 mg/dL) Incentive rewards: CS>6.86, $3.00 PMPM (Medicare); $0.75 PMPM (commercial); CS>6.23;$1.50 PMPM (Medicare);$0.37 PMPM (commercial);50% improvements and CS<6.23, $0.75 PMPM (Medicare),$0.18 PMPM (commercial) *Fairbrother 1999 -*	**Primary care/solo or group practices/primary care providers contracted with a particular insurance scheme (e.g. Beaulieu PCP contracted with Independent Health), n=8** Beaulieu 2005; Fairbrother 1999 (feedback only);Fairbrother 2001;Grady 1997; Hillman 1998; Hillman 1999; Kouides 1998; Roski 2003 **Community mental health centre (type of team unclear)** **Not stated (n=9)**	60 physicians (15 bonus; 15 enhanced fee-for-service; 15 feedback only; 15 control) *Fairbrother 2001 -* 57 physicians (24 bonus; 12 fee-for-service; 21 control) Grady 1997 - 95 physicians **Number of participating providers not stated** (n=10) **Participating patients:** *Beaulieu 2005-*	61 community-based primary care practices (20 cue and reward; 18 cue; 23 control) *Hillman 1998 -* 52 primary care-sites (26 intervention: 26 control) *Hillman 1999 -* 49 PC sites (19 feedback plus incentive; 15 feedback only; 15 control) *Kouides 1998 -* 54 practices (27 intervention and 27 control) *Roski 2003 -* 37 primary care sites(13 incentive; 9 incentive +registry and 15 control) **Office of substance abuse-1** *Shen 2003 - unknown number of facilities* **Pharmacies-1** *Christensen 2000* - 200 pharmacies (110 in Intervention group; 90 in control)	

Table 2. Characteristics of included reviews *(Continued)*

	Bonuses: $1000 (20% improvement from baseline); $2500 (40% improvement); $5000 (80% up-to-date immunisation) *Fairbrother 2001 -* Bonuses: $1000 (30% improvement from baseline); $2500 (45% improvement); $5000 (80% up-to-date immunisation); $7500 (90% up-to-date) *Grady 2003 -* "Token" reward, based on the percentage referred for mammography during quarterly audit ($50 for a 50% referral rate). *Hillman 1998 -* Compliance with cancer screening for women age >50 years; aggregate compliance scores and improvement in scores over time; full and partial bonuses (20%; 10% of capitation); range of bonus per site, $570 to $1260. *Hillman 1999 -* Pediatric immunisations; well-child visits; bonuses based on total compliance score for quality indicators; full and partial bonuses (20%; 10% of site's total 6 months capitation for paediatric members age <6 years); 3 highest scoring sites received full bonus; next 3 received partial bonus; most improved sites received partial bonus; average bonus, $2000 (range $772 to $4682). *Kouides 1998-* See description of incentive under Akbari 2008 *Rosenthal 2005-* Incentive payout based on provider's groups ability to reach or exceed target rates for cervical cancer screening, mam-	Clark 1995 **Pharmacies** Christensen 2000 Rosenthal 2005- Provider groups in the Pacific Northwest were the comparison group	476 diabetic patients, n=600 Independent Health diabetic patients were the comparison group *Clark 1995-* 185 clients (95 in traditional case managers and 90 in continuous treatment team) Kouides 1998- active non-nursing home patients 65 years or older who had an office visit in the previous year (21 196 in intervention group and 17 608 in control group) *Shen 2003-* 5552 clients (2367 of-	**Community mental health centre-1** *Clark 1995 -* 7 community mental health centres **Provider groups contracted with PacifiCare-1** *Rosenthal 2005 -* 163 provider groups contracted with PacifiCare Health Systems in California (provider groups in the Pacific Northwest were the comparison group) **Not stated (n=2)** Fairbrother 1999; Fairbrother 2001 (only that they were paediatricians)

Table 2. Characteristics of included reviews *(Continued)*

				mography and haemoglobin A1C testing for diabetic patients. *Roski 2003 -* 75% of patients with smoking status identified/ documented at last visit; 65% of patients with quitting advice documented at last visit (targets set at approximately 15% above the average from 2 Y before study); bonuses, $5000 for sites with 1-7 providers and $10 000 for sites with >8 providers. *Shen 2003 -* Additional funds based on efficiency, effectiveness and service to special pupulations.Improved health status within 90 d (measured by ADL classification); $126 to $370 per case (range of bonus); Discharge incentive: timely discharge and resident did not return within 90 d; $60 to $230 (range of bonus); type A patients not eligible.			fice of substance abuse clients; 3185 Medicaid clients)	
Sturm 2007	To determine the effects on drug use, healthcare utilisation, health outcomes and costs (expenditures) of (pharmaceutical) policies, that intend to affect prescribers by means of financial incentives.	14 databases, 5 web sites up to October 2005 to April 2006	**ITS (n=3)** Guether 1995; Schöffski 1997; Walley 2000 **CITS (n=3**) Harris 1996; Rafferty 1997; Wilson 1995 **CBA (n=10)**	**Payment for providing care for a patient or specific population (capitation) (n=10)** Baines 1997;Bradlow 1993;Burr 1992;Corney 1997;Harris 1996; Kammerling 1996;Rafferty 1997; Whynes 1997;Wilson 1995;Wilson 1999 *UK Fund-holding scheme*:GPs were given financial control over some of their provided services. Besides costs of prescribed drugs, practice staff and a range of secondary care such as specialists services were covered by separate budgets, with the drug budget offering th e greatest saving potential. Overspending in one budget had to be covered by funds from another budget, and savings could be used in other areas of patient care. **Unclear (n=3**)	**Non fund holding (n=10**) **Unclear (n=3)** (the drug budgets)	**UK (n=10)** Baines 1997; Bradlow 1993; Burr 1992; Corney 1997; Harris 1996; Kammerling 1996; Rafferty 1997;Whynes 1997; Wilson 1995; Wilson 1999 **Ireland (n=1)**	**13 studies - Physicians (assumed)** Guether 1995 82 GP's, West Germany Statutory Health insurance Walley 1999 223 GPs from Eastern Health Board Cohort, Ireland	**Primary care practices in 12/13** Not clear in one described as practices (Schöffski 1997) Baines 1997 Fund-holders: Devon (n=19) and Lincolnshire (n=22), Non-FH: Lincolnshire (n=86) and Devon (n=106) Bradlow 1993 FH (1st wave):5; Non-FH:7 practices, Oxford, UK

Table 2. Characteristics of included reviews *(Continued)*

	Baines 1997; Bradlow 1993; Burr 1992; Corney 1997; Harris 1996; Kammerling 1996; Rafferty 1997;Whynes 1997; Wilson 1995; Wilson 1999	Guether 1995;Schöffski1997 *German drug budgets*: Collective budgets for drug expenditure for physicians in private practice. While spending caps were regionally negotiated or nationally set each year and made all physicians in private practice in one region collectively liable, target volumes for each individual practice were only theoretically established. Walley 2000 *Ireland indicative drug budgets*: GPs individual indicative or hypothetical budgets covered prescribing costs and were calculated based on previous spending and the national average. Savings were split between the GP and the local health authority to be used for the development of services. There were no penalties for overspending.	Walley 2000 **Germany (n=2)** Guether 1995; Schöffski 1997	Burr 1992 FH (1st wave): 4; Non-FH: 4 practices Corney 1997 FH (1st wave):4;Non-FH:4 practices (South Thames Region, UK Guether 1995- number of practices not stated Harris 1995 All general practices in England, UK Kammerling 1996 10 fund-holding (study) and 22 non-FH practices Rafferty 1997 FH (1st wave):23; FH (2nd wave):34; FH (3rd wave):9; Non-FH: all in Northern Ireland Schoffski 1997 309-382 practices, (Germany Statutory Sickness Funds)

Table 2. Characteristics of included reviews *(Continued)*

					Walley 1999- number of practices not stated Whynes 1997 FH (4th wave):23; Non-FH:63 practices Wilson 1995 FH (1st wave):20; FH (2:nd wave): 31 and FH (3rd wave):49; Non-FH: 312 practices Wilson 1999 5 health authorities in NW-region, UK

Table 3. Comparisons

Intervention	Fee-for service	Capitation	Unspecified Control	Mixed
Salary reimbursement	Hickson 1987			
Enhanced fee-for-service	Christensen 2000; Clark 1995; Davidson 1992		Fairbrother 1999;Fairbrother 2001	
Capitation			Baines 1997; Bradlow 1993; Burr 1992; Corney 1997; Coulter 1993; Harris 1996; Kammerling 1996; Linnala 2001; Rafferty 1997; Whynes 1997; Wilson 1995; Wilson 1999	
Target payments and bonuses			Beaulieu 2005; Fairbrother 1999; Fairbrother 2001; Grady 1997; Hillman 1998; Hillman 1999; Rosenthal 2005; Roski 2003; Shen 2003	
Mixed Incentives	Davidson 1992;Hutchison 1996; Kouides 1998;Krasnik 1990; Ritchie 1992;			Krasnik 1990
Unclear incentives			Guether 1997;Schoffski 1997;Walley 2000	

Table 4. Main outcomes of included reviews

Author Year	Consultations/visits	Processes of care (prescriptions, drug use vaccinations etc.)	Referrals	Admissions	Guideline compliance	Costs
AKBARI 2009	Primary care and Non-primary care visits (Davidson 1992)		Outpatient referrals (Coulter 1993) Orthopaedic referrals (Kammerling 1996)			
GOSDEN 2001	Primary care and Non-primary care visits , Clinic and emergency department visits (Davidson 1992) Various healthcare visits (Hickson 1987)	Diagnostic services; Curative services; Renewal of prescription (Krasnik 1990) Immunisation rates (Kouides 1998; Ritchie 1992)	Referrals to specialists and to hospitals (Krasnik 1990)	Hospitalisations (Davidson 1992) Admissions and hospital days stayed (Hutchison 1996)	Compliance with CHAP guidelines (Davidson 1992) Compliance with AAP guidelines (Hickson 1987)	

Table 4. Main outcomes of included reviews *(Continued)*

	Face to face and telephone consultations (Krasnik 1990)				
PETERSEN 2005		Diabetic tests (Beaulieu 2005) Time spent in community treatment/care/office based care management time and total care manager time per patient (Clark 1995) Immunisation status (Fairbrother 1999: Fairbrother 2001) Immunisation rate (Kouides 1998) Cancer and mammography screening rates, and HbA1c testing (Rosenthal 2005) Tobacco use identification and quitting advice (Roski 2003) Health services to special populations (Shen 2003)	Referrals to mammography (Grady 1997)	Compliance with cancer screening guidelines (Hillman 1998); Compliance with guidelines concerned with paediatric immunisations and well-child visits (Hillman 1999)	
STURM 2007		Drug use (Items per patient) (Bradlow 1993; Burr 1992; Guether 1997: Harris 1996: Rafferty 1997; Walley 2000; Whynes 1997; Wilson 1995) Generic percentage (Baines 1997; Bradlow 1993; Rafferty 1997; Whynes 1997; Wilson 1995; Wilson 1999)	Referrals to outpatients specialists (Guether 1997) Orthopaedic referrals (Kammerling 1996) Referrals to outpatients clinics and hospitals (Schoffski 1997);		Cost per item (Bradlow 1993; Rafferty 1997; Walley 2000; Wilson 1995; Wilson 1999) Cost per patient (Baines 1997; Bradlow 1993; Burr 1992; Corney 1997; Harris 1996; Rafferty 1997;Whynes 1997; Wilson 1995; Wilson 1999) Total prescribing cost (Harris 1996; Walley 2000) Drug expenditure (Walley 2000)

Table 5. AMSTAR scores

	Akbari 2009	Gosden 2001	Petersen 2006	Sturm 2007
1.Was an 'a-priori' design provided?	YES	CAN'T ANSWER	CAN'T ANSWER	YES
2.Was there duplicate study selection and data extraction?	YES	YES	YES	YES
3.Was a comprehensive literature search performed?	YES	YES	NO	YES
4.Was status of publication (e.g. grey literature) used as an inclusion criterion?	CAN'T ANSWER	CAN'T ANSWER	YES	NO
5.Was a list of studies (included/excluded) provided?	YES	YES	YES	YES
6. Were the characteristics of included studies provided?	YES	YES	YES	YES
7.Was the scientific quality of the included studies assessed and reported?	YES	YES	YES	YES
8. Was the scientific quality of the included studies used appropriately in formulating conclusions?	YES	YES	YES	YES
9. Were the methods used to combine the findings of studies appropriate?	YES	YES	YES	CAN'T ANSWER
10. Was the likelihood of publication bias assessed?	NO	NO	CAN'T ANSWER	NO
11. Was the conflict of interest stated?	YES	NO	YES	YES
AMSTAR SCORE	9	7	7	9

Table 6. Vote counting results

Outcome	Consultation/Visit rates	Processes of care	Referrals/Admissions	Compliance with guidelines	Prescribing costs	Overall effect within intervention
Intervention						
Payment for working for a specified time period	3/9 outcomes from 1 study reported in 1 review favoured the intervention			0/2 outcomes from 1 study reported in 1 review favoured the intervention		3/11 (27%) 1 study 1 review
Payment for each service/episode/visit	3/3 outcomes from 1 study reported in 2 reviews favoured the intervention	4/6 outcomes from 4 studies reported in 1 review favoured the intervention	0/1 outcomes from 1 study reported in 1 review favoured the intervention			7/10 (70%) 5 studies 3 reviews

Table 6. Vote counting results *(Continued)*

Payment for providing care for a patient or a specific population		17/30 outcomes from 8 studies reported in 1 review favoured the intervention	3/5 outcomes from three 3 studies reported in 2 reviews favoured the intervention		28/34 outcomes from 10 studies reported in 1 review favoured the intervention	48/69 (70%) 13 studies 2 reviews
Payment for providing a pre-specified level or providing a change in activity or quality		16/16 outcomes from 5 studies reported in 2 reviews favoured the intervention	1/2 outcomes from three studies reported in 1 review favoured the intervention	0/2 outcomes from 2 studies reported in 1 review favoured the intervention		17/20 (85%) 10 studies 2 reviews
Mixed or other systems	4/5 outcomes from 2 studies (1 outcome unclear) reported in 2 reviews favoured the intervention	4/5 outcomes from 3 studies reported in 2 reviews favoured the intervention	7/8 outcomes from 5 studies reported in 2 reviews favoured the intervention	5/13 outcomes from 2 studies reported in 1 review favoured the intervention		20/31 (65%) 7 studies 3 reviews
Overall effect within outcomes	10/17 (59%) 3 studies 2 reviews	41/57 (72%) 19 studies 3 reviews	11/16 (69%) 11 studies 4 reviews	5/17 (29%) 5 studies 2 reviews	28/34 (82%) 10 studies 1 review	

APPENDICES

Appendix 1. Full search strategy

MEDLINE strategy

(also translated successfully to EMBASE, NHS EED, and CDSR/DARE using .kw instead of .sh)

1. exp Reimbursement Mechanisms/
2. exp Reimbursement, Incentive/
3. exp "Fees and Charges"/
4. exp Fees, Medical/
5. exp Income/
6. exp Capitation Fee/
7. exp Physician Incentive Plans/
8. exp Prospective Payment System/
9. exp "Salaries and Fringe Benefits"/

10. exp Physician's Practice Patterns/ or Physicians/

11. (target* or reimburse* or payment or remunerat* or incentive* or (financ* adj3 penalt*) or financial or salar* or fee or fees or capita* or (pay adj3 perform*) or (payment adj3 perform*) or (pay adj3 reduc*) or (payment adj3 reduc*) or (pay adj3 penalt*) or (payment adj3 penalt*)).ti,ab.

12. or/1-11

13. exp Family Practice/

14. exp Physicians, Family/

15. exp Private Practice/

16. exp Group Practice/

17. exp Institutional Practice/

18. exp Partnership Practice/

19. ((general adj pract*) or (family adj physician*) or (family adj pract*) or (general adj practice) or (primary adj care) or (primary adj health adj care) or (primary adj care adj physician) or physician* or GP* or doctor* or (group adj pract*) or (institutional adj pract*) or (partnership adj pract*) or (private adj pract*) or (primary adj pract*)).ti,ab.

20. or/13-19

21. 12 and 20

Applied SIGN filter[1] for systematic reviews

[1] http://www.sign.ac.uk/methodology/filters.html

Science Citation Index and Social Science Citation Index strategy

#1 TS=(physician financial incentive*)

#2 TS=((payment SAME perform*) or (pay SAME perform*) or (payment SAME incentive*) or (pay SAME incentive*) or (payment SAME reduc*) or (pay SAME reduc*) or (financ* SAME perform*) or (financ* SAME incentive*) or (financ* SAME penalt*) or (pay SAME penalt*) or (payment SAME penalt*) or (reimburs*) or (target* SAME perform*) or (target* SAME incentive*) or (fee SAME service) or (remunerat*) or (capitation))

#3 TS=("general pract*" or "family physician*" or "family pract*" or "primary care" or "primary health care" or "primary care physician" or physician* or GP or GPs or doctor* or "group pract*" or "institutional pract*" or "partnership pract*" or "private pract*" or "primary pract*")

#4 TS=review

#5 (#2 and #3 and #4) or (#2 and #3 (document type = review))

#6 #1 or #5

HEED strategy

#1 Keyword = (general next pract*) or (family next physician*) or (family next pract*) or (general next practice) or (primary next care) or (primary next health next care) or (primary next care next physician) or physician* or GP or GPs or doctor* or (group next pract*) or (institutional next pract*) or (partnership next pract*) or (private next pract*) or (primary next pract*)

#2 Keyword = target* or reimburse* or payment or remunerat* or incentive* or salar* or fee or fees or capitat* or (pay near perform*) or (payment near perform*) or (pay near reduc*) or (payment near reduc*) or (financ* near penalt*) or (pay near penalt*) or (payment near penalt*)

#3 #1 and #2

Econlit strategy

#1 Kw = reimburs* or fee-for-service* or capitation* or incentiv* or salaried or salaries or salary or penalties or penalty

#2 Ti/Ab = target* or reimburse* or payment or remunerat* or incentive* or salar* or fee or fees or capitat* or (pay n3 perform*) or (payment n3 perform*) or (pay n3 reduc*) or (payment n3 reduc*) or (financ* n3 penalt*) or (pay n3 penalt*) or (payment n3 penalt*)

#3 #1 or #2

#4 Kw = physician or physicians

#5 Ti/Ab = (general w pract*) or (family w physician*) or (family w pract*) or (general w practice) or (primary w care) or (primary w health w care) or (primary w care w physician) or physician* or GP or GPs or doctor* or (group w pract*) or (institutional w pract*) or (partnership w pract*) or (private w pract*) or (primary w pract*)

#6 #4 or #5

#7 Kw = review

#8 Ti/Ab = review* or systematic or search* or database*

#9 #7 or #8

#10 #3 and #6 and #9

PPD strategy

#1 Keyword = Funding or Remuneration

#2 Type of review = Systematic reviews of effects or Overview of systematic reviews

Appendix 2. Data extraction form

Data extraction form

1. Reviewer name:

2. Date:

3. First author, Title, Citation

4. Stated aim of review (if stated cut & paste or copy out of the review verbatim)

5. Can the result statement in the review be allocated at an individual study basis?

YES **NO**

N.B. A review must report data on an individual study basis to be included in the overview. If it does not: COLLECT NO FURTHER DATA.

6. Details of search (record dates and databases)

a) Number and design of studies included in the review:

Design (as defined by the author)	Number
Cluster RCT	
RCT	
CCT	
ITS	
CBA	
Other	
Unclear	

b) Number and type of financial incentives interventions (cut and paste from the review verbatim. This may mean that there is more than one pasted description against each category.)

Types of intervention	Number	Text from the review
1. Payment for working for a specified time period (e.g. salary, sessional payment)		
2. Payment for each service/episode/case (fee-for-service)		
3. Payment for providing care for a patient or specific population (capitation)		
4. Payment for providing a pre-specified level of quality of care (include target payments, bonuses)		
5. Other mixed (describe combination)		
6. Unclear		

c) Description of the 'comparison intervention':

Comparison intervention	Number of studies

(Continued)

Unclear

d) Countries within which studies were conducted:

Country	Number of studies
Unclear	

e) Number and type of participants in included studies:

Type of healthcare workers (e.g. doctor, nurse, manager, unclear)	Number of studies

f) Number and type of healthcare settings in included studies:

Type of healthcare setting (e.g. primary care, secondary care, unclear)	Number of studies

g) Total number of patients included in the review:_____

7. Main outcomes and effects addressed in the review:

Outcome	Number of studies measuring outcome	Overall Effect

8. Reviewer's interpretation of the results of the review:

Other comments:

QUALITY ASSESSMENT

Description of the Quality of the review (AMSTAR):

1. Was an 'a priori' design provided? The research question and inclusion criteria should be established before the conduct of the review.	Yes No Can't answer Not applicable
Comment	
2. Was there duplicate study selection and data extraction? There should be at least two independent data extractors and a consensus procedure for disagreements should be in place.	Yes No Can't answer Not applicable
Comment	
3. Was a comprehensive literature search performed? At least two electronic sources should be searched. The report must include years and databases used (e.g. Central, EMBASE, and MEDLINE). Key words and/or MESH terms must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialised registers, or experts in the particular field of study, and by reviewing the references in the studies found.-+	Yes No Can't answer Not applicable
Comment	
4. Was the status of publication (i.e. grey literature) used as an inclusion criterion? The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language etc.	Yes No Can't answer Not applicable
Comment	
5. Was a list of studies (included and excluded) provided? A list of included and excluded studies should be provided.	Yes No Can't answer Not applicable
Comment	
6. Were the characteristics of the included studies provided? In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions and outcomes. The ranges of characteristics in all the studies analysed e.g. age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported.	Yes No Can't answer

(Continued)

	Not applicable
Comment	
7. Was the scientific quality of the included studies assessed and documented?	
'A priori' methods of assessment should be provided (e.g., for effectiveness studies if the author(s) chose to include only randomised, double-blind, placebo controlled studies, or allocation concealment as inclusion criteria); for other types of studies alternative items will be relevant.	Yes No Can't answer Not applicable
Comment	
8. Was the scientific quality of the included studies used appropriately in formulating conclusions? The results of the methodological rigor and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating recommendations.	Yes No Can't answer Not applicable
Comment	
9. Were the methods used to combine the findings of studies appropriate? For the pooled results, a test should be done to ensure the studies were combinable, to assess their homogeneity (i.e. Chi-squared test for homogeneity, I^2). If heterogeneity exists a random effects model should be used and/or the clinical appropriateness of combining should be taken into consideration (i.e. is it sensible to combine?).	Yes No Can't answer Not applicable
Comment	
10. Was the likelihood of publication bias assessed? An assessment of publication bias should include a combination of graphical aids (e.g., funnel plot, other available tests) and/or statistical tests (e.g., Egger regression test).	Yes No Can't answer Not applicable
Comment	
11. Was the conflict of interest stated? Potential sources of support should be clearly acknowledged in both the systematic review and the included studies.	Yes No Can't answer Not applicable
Comment	

Appendix 3. AMSTAR operationalisation

AMSTAR Criteria

1.	Was an 'a priori' design provided? [Yes? the research question and inclusion criteria were established before conducting the review i.e. SRs for which a research protocol is available (i.e.Cochrane), or SRs developed within specific research programs (HTA, U.S, Preventive Services Task Force, AHQR, NICE etc.), No- the authors stated that there's no protocol available, Can't answer- no information about it]
2.	Was there duplicate study selection and data extraction? [Yes? at least two people working independently extracted the data and the method was reported for reaching consensus if disagreements arose. The answer is 'yes' also if only one between selection or extraction is done in duplicate, No-the authors stated that both selection and extraction were performed by one person, Can't answer-no information about it]
3.	Was a comprehensive literature search performed? [Yes? at least two electronic sources were searched; details of the databases, years searched and keywords and/or search strategy were provided; the search was supplemented by searching of the reference lists of included studies, and specialised registers, and by contacting experts, No-only one database searched or used no other sources, Can't answer - partial or no information reported (e.g. databases reported, but keywords or years missing]
4.	Was status of publication (e.g. grey literature) used as an inclusion criterion? [Yes? the authors stated that they excluded studies from the review based on publication status, or language (=0). No? authors searched for reports irrespective of publication type (=1). They did not exclude reports based on their publication type or language from the systematic review, Can't answer- no information about it]
5.	Was a list of studies (included and excluded provided)? [Yes? a list was provided and information given on how many records were found by the search strategy, how many considered for inclusion, how many included/excluded and why, No - no information about it or only the list of included studies provided, Can't answer- partial information (e.g. excluded studies listed in references but not in the text]
6.	Were the characteristics of the included studies provided? [Yes? data on participants, interventions and outcomes were provided, and the range of relevant characteristics reported either in a table or as narrative text, No - no information about it, Can't answer-partial information (e.g. only year of publication and intervention reported, or only some of the included studies described]
7.	Was the scientific quality of the included studies assessed and reported? [Yes? predetermined methods of assessing quality were reported, No-no information about it, Can't answer - the authors stated that a quality assessment was done, but did not describe how it was done]
8.	Was the scientific quality of the included studies used appropriately in formulating conclusions? [Yes? the quality (and limitations) of included studies was used in the analysis, conclusions and recommendations of the review, No-quality assessment was done but not mentioned in analysis, conclusions and recommendations, Can't answer-impact of quality of studies on results unclear or not used for conclusions, Not applicable-the scientific quality of included studies was not assessed in the first place]
9.	Were the methods used to combine the findings of studies appropriate? [Yes? if results were pooled statistically, heterogeneity was assessed and used to inform the decision of statistical model to be used. If heterogeneity was present, the appropriateness of combining studies was considered by review authors, Yes- also if a narrative summary was appropriate to do, No-Heterogeneity present but not discussed, fixed-effect model used by default, Can't answer- heterogeneity test result not reported or model (random vs. fixed) used to combine studies not specified

(Continued)

10.	Was the likelihood of publication bias assessed? [Yes? publication bias was explicitly considered and assessed. Funnel plot or other methods used, No- publication bias was not assessed of some reason or no information about it, Can't answer-mentioned or discussed only in conclusions]
11.	Was the conflict of interest stated? [Yes? conflict of interest and sources of support were clearly acknowledged, No-conflict of interest and sources of funding not reported]

Appendix 4. Results

	AKBARI 2008	GOSDEN 2001
CONSULTATIONS * **(primary and non-primary care visits)**	**DAVIDSON 1992** **Design**: Cluster 3 arm RCT comparing fee-for-service (high rate[1]) vs. capitation +financial risk sharing vs. fee-for-service (low rate[2]) =control **Participants**: primary care providers in private office based practices (n=80) who treated Medicaid children who received welfare benefit under Aid for Families and Dependent Children Programme (n=3770) and had submitted more than $2000 per child in Medicaid billings the previous year. $25 was set aside for each participating child in the capitation physicians practice, PCPs received 40% of any surplus **Type of intervention:** Payment for providing care to a patient or specific population (capitation) and payment for each service/episode/case (fee-for-service) **Description**: Change in remuneration system from a: 1. Low cost fee-for-service system (control group) where PCPs are paid fees for the same services as the high fee-for-service group, but the fees were approximately half size to either a: 2. High cost fee-for-service (market rates): PCPs were paid a fee for comprehensive exams (including treatment), routine office visits, initial and follow-up visits Or 3. Capitation-based budgetary system (with some degree of risk sharing by the provider for secondary care provision) for	**DAVIDSON 1992** See study description under Akbari 2008 **Expected change:** decreased number of clinic/emergency department visits *Mean number (standard error)**[6]** of visits per year per patient for children 5 years and younger* **Clinic/emergency departments visits** *Pre-intervention* : Capitation+ financial risk sharing: 1.25 (0.045) (n=764) Fee-for-service (high rate) 1.37 (0.117) (n=1015) Fee-for-service (low rate)=Control: 1.52 (0.099) (n=1991) *During intervention* : Capitation + financial risk sharing: 0.78 (0.028) (n=764) Fee-for-service (high rate): 0.85 (0.075) (n=1015) Fee-for-service (low rate)=Control: 1.18 (0.062) (n=1991) Absolute changes (intervention minus control): Capitation + financial risk sharing: Six months before: -0.27 During study:-0.30 Fee-for-service (high rate): Six months before: -0.15

(Continued)

the management of Medicaid eligible paediatric care.

Level: not stated

Payment frequency: not stated

Comparison: High cost fee-for–service vs. low cost fee-for-service and capitation-based budgetary system (with some degree of risk-sharing) vs. low cost fee-for-service.

Duration: 6 months pre intervention, post intervention period not specified

Expected change: increased primary care visits and decreased non-primary care visits

Mean number (standard error)[3] of primary care visits per year per patient:

Pre-intervention **:**

Capitation + financial risk sharing: 3.22 (0.116) (n=764)

Fee-for-service (high rate): 3.68 (0.177) (n=1015)

Fee-for-service (low rate)=control: 3.06 (0.129) (n=1991)

During intervention **[4]**:

Capitation+ financial risk sharing: 2.89 (0.143) (n=764)

Fee-for-service (high rate):3.71(0.148) (n=1015)

Fee-for-service (low rate)=control: 2.47(0.101) (n=1991)

Absolute difference (post):

Capitation + financial risk sharing vs. fee-for-service (low rate)=control: +0.42

Fee-for-service (high rate) vs. fee-for-service (low rate)=control: +1.24

Capitation + financial risk sharing vs. fee-for-service (high rate): -0.82

Relative percentage difference:

Capitation +financial risk sharing vs. Fee-for-service (low rate)=control: +17.0%

During study: -0.33

Relative change (percentage difference between intervention and control)

Capitation+ financial risk sharing

Six months before:-18%

During study: 25%.

Fee-for-service (high rate) :

Six months before: -10%

During study: -28%

(See Akbari et al. 2008 for results on PCP and non-PCP visits)

Significance not reported.

KRASNIK 1990

Design: CBA

Participants: 100 GP's in Copenhagen city (intervention) of which 71 were included in the analyses and 326 GPs in Copenhagen county (control)

Type of intervention: Payment for each service/episode/case (fee-for-service)

Description: GPs working inside Copenhagen city were paid by capitation until October 1987.After this date GPs fees were introduced for face to face telephone and home visit consultations, and repeat prescriptions. Additional fees are payable for 40 special services (e.g. cervical smear tests), for 40 special laboratory investigations performed in the practice (e.g. haemoglobin concentration) and for a few preventive services (e.g. immunisations).

Level: not stated

Payment frequency:not stated

Comparison: change from capitation to a mixed fee per item and capitation based system

Duration: pre-intervention prior to March 1987; post-intervention: March 1998 and November 1988

Expected change: increased rate of face-to-face and telephone consultations. Estimated changes in number of contacts per 1000 enlisted patients (95% confidence interval) between six months before (BL=100) versus six and 12 months after intervention over a one week period[7]

Face-to-face consultations

At six months: Intervention: 112.7 (106.8-118.8),(n=71); Control: 105.5 (n=326)(no CI's reported)

(Continued)

Fee-for-service (high rate) vs. Fee-for-service (low rate)=control: +50.2%

Capitation + financial risk sharing vs. Fee-for-service (high rate): -22.1% (relative to fee-for-service (high rate))

Absolute difference from baseline:

-0.33 (Capitation+ financial risk sharing) vs. +0.03 (Fee-for-service (high rate)) vs. -0.59 (Fee-for-service (low rate)= control)

Difference in absolute change from baseline: Capitation + financial risk sharing vs. Fee for service (low rate)=control: +0.26

Fee-for-service (high rate) vs. Fee-for-service (low rate)=control: +0.62

Capitation + financial risk sharing vs.

Fee-for-service (high rate): -0.36

Non primary care visits

*Mean number (standard error)**[5]** of visits per year per patient:*

Pre intervention:

Capitation +financial risk sharing: 0.62 (0.022) (n=764)

Fee-for-service (high rate)

0.67(0.071) (n=1015);

Fee-for-service (low rate)=control

0.61(0.061) (n=1991)

During intervention **:**

Capitation + financial risk sharing

0.57 (0.021)

Fee-for-service (high rate)

0.85 (0,027);

Fee-for-service (low rate)=control

0.80 (0.046)

Absolute difference (post):

Capitation + financial risk sharing vs. Fee-for-service (low rate)=control: -0.23

At 12 months: Intervention: 104.4 (98.9 to 110.2), (n=71); Control: 104.9, (n=326) (no CI's reported)

Absolute changes: intervention minus control : At six months: +7.2 ; At 12 months: -0.5

Relative changes: percentage difference between intervention and control:

At six months:+130.9%; At 12 months: -10.2%

Consultations by telephone

At six months: Intervention: 118.6 (108.5 to 129.7) (n=71); Control: 108.4 (n=326)(no CI's reported)

At 12 months: Intervention:115.4 (105.5 to 126.3) (n=71); Control:104.0 (n=326)

Absolute changes: intervention minus control : At six months: +10.2; At 12 months: +11.4

Relative changes: percentage difference between intervention and control: At six months:+121.4%; At 12 months:+285%

Significance not reported.

HICKSON 1987

Design: RCT (two armed physician randomised trial)

Participants: ten second-year and 8 third-year paediatric residents; and 486 patients (intervention); 395 patients (control)

Type of intervention: Payment for working for a specified time period (e.g. salary, sessional payment)

Description: Salary group received $20 per month and the fee-for-service group (=control group) received $2 per visit.

Level: not stated

Payment frequency:not stated

Comparison: Salary reimbursement vs. fee-for-service (=control)

Duration: September 1983 to June 1984 (9 months)

Expected change: that fee-for-service physicians, compared with salaried physicians, would attend a greater percentage of their patient's visits (i.e. improve the continuity of their patients care), encourage more necessary and unnecessary visits per patient

Average number (over a nine-month period) **[8]** *of:*

Patient visits attended per PCP

Salary: 104.8 (n=9) vs. Fee-for-service (control): 111.6 (n=9)

Absolute changes: -6.8 (NS); Relative changes: -6.1%

Emergency room visits/enrolled patients/ PCP

(Continued)

	Fee-for-service (high rate) vs. Fee-for-service (low rate)=control: +0.05 Capitation + financial risk sharing vs. Fee-for-service (high rate):-0.28 Relative percentage difference: Capitation + financial risk sharing vs. Fee-for-service (low rate)=control: -28.8% Fee for service (high rate) vs. Fee-for-service (low rate)= control: +6.25 Capitation + financial risk sharing vs. Fee-for-service (high rate): -32.9% (relative to Fee-for-service (high rate)) Absolute difference from baseline: -0.05 (Capitation+ financial risk sharing) vs.0.18 (Fee-for-service (high rate)) vs.0.19 (Fee-for-service (low rate)=control) Difference in absolute change from baseline: Capitation + financial risk sharing vs. Fee-for-service (low rates)=control -0.23 Fee-for-service (high rate) vs. Fee-for-service (low rate)=control: -0.01 Capitation + financial risk sharing vs. Fee-for-service (high rate):-0.24 Significance not reported.	Salary: 0.22 (n=9) vs. Fee-for-service (control): 0.12 (n=9) Absolute changes: +10 (S, *P*<0.01); Relative changes: +83.33% **Scheduled visits per enrolled patient/PCP** Salary : 2.83 (n=9) vs. Fee-for-service (control): 3.69 (n=9) Absolute changes: -0.86 (S, *P*<0.01);Relative changes: -23.3% **Completed visits per enrolled patient/PCP** Salary: 2.21 (n=9) vs. Fee-for-service (control): 2.70 (n=9) Absolute changes: -0.49 (S, *P*<0.05); Relative changes: -18.1% **Sick, primary visits per enrolled patient/PCP** Salary: 0.98 (n=9) vs. Fee-for-service (control): 0.95 (n=9) Absolute changes: +0.03 (NS); Relative changes: +3.2% **Sick follow-up visits per enrolled patient/PCP** Salary: 0.24 (n=9) vs. Fee-for-service (control): 0.33 (n=9) Absolute changes: -0.09 (NS); Relative changes: -27.3% **Well-child visits per enrolled patient/PCP** Salary: 0.99 (n=9) vs. Fee-for-service (control): 1.42 (n=9) Absolute changes: -0.43 (S, *P*<0.01); Relative changes:- 30.3% **Patients enrolled per PCP** Salary: 55.1 (n=9) vs. Fee-for-service (control): 43.4 (n=9) Absolute changes: +11.7 (S, *P*<0.05);Relative changes: +27% **Percentage visits attended by patient's primary physician (continuity)-** Salary: 78.3 (n=9) vs. Fee-for-service (control): 86.6 (n=9) Absolute changes: -8.3 (S, *P*<0.051);Relative changes: -9.6%
PROCESSES OF CARE **(vaccinations,** **prescriptions,** **drug use.)**		**KRASNIK 1990** See study description under 'Consultations' **Expected change:** increased provision of diagnostic and curative services and renewal of prescriptions *Differences between six months before intervention (=100) and six and 12 months after in the following outcomes over a one week period per 1000 patients. Mean number of contacts (95% CI)**[9]**:* **Diagnostic services** At six months

(Continued)

Intervention 138.1 (118.7 to 160.5) (n=71); Control:105.3

At 12 months

Intervention 159.5 (137.8 to 184.7) (n=71); Control: 107.6 (n=326)

Absolute changes:

At 6 months :+32.8

At 12 months :+52.2

Relative changes:

At 6 months: +618.9%;

At 12 months: +686.9%

Curative services

At six months

Intervention 194.6 (152.2 to 248.9) (n=71); Control: 106.0 (n=326)

At 12 months

Intervention 194.8 (152.3 to 249.2) (n=71); Control115.0 (n=326)

Absolute changes

At 6 months: +88.6

At 12 months: +79.8

Relative changes

At 6 months +1476.7%

At 12 months +532%

Renewal of prescription

At six months

Intervention 82.5 (68.4 to 99.7) (n=71)

Control 91.5 (n=326)

At 12 months

Intervention 65.2 (53.2 to 79.9) (n=71)

Control 92.6 (n=326)

Absolute changes

At 6 months: -9; At 12 months: -27.4

Relative changes:

At 6 months: 105.8%; At 12 months: -370.3%

Significance not reported.

(Continued)

		KOUIDES 1998 **Design:** RCT **Participants:** 54 practices; n=27 (intervention) and n=27 (control); active non-nursing home patients 65 years or older who had an office visit in the previous year (21 196 in intervention group and 17 608 in control group) **Type of intervention:** Payment for providing a pre-specified level of quality of care (include target payments, bonuses) – mixed financial incentive? **Description:** PCPs in the intervention group received an additional 10% ($0.8) or 20% ($1.6) reimbursement per shot according to whether they immunised 70% or 85% (respectively) of the eligible population. **Level**: provider group **Payment frequency**: one time (end of study) **Comparison:** Salary reimbursement + fee-for-service vs.fee-for-service (=control) **Duration:** September 1991 to January 1992 **Expected change:** increased immunisation rate **Mean influenza vaccination rate in the intervention period** (1991) Absolute changes:+5.9% Intervention:68.6%,(n=27); Control: 62.7% ,(n=27); *P*=0.22 Relative changes: +9.4% **Change in influenza vaccination rate from baseline year** (between 1991 and 1990) Absolute changes: +6.8% Intervention:10.3 %, (n=27); Control: 3.5%, (n=27), *P*=0.03 Relative changes: +194.3% **Overall influenza vaccination rate:** Sum of all immunisations given divided by the sum of eligible patients in the intervention period (1991) Absolute changes: +6.8% Intervention: 66.9%, (n=27); Control:; 60.1%, (n=27); *P*=N/A Relative changes: +11.3%
REFERRALS/ADMISSIONS	**COULTER 1993** **Design:** CBA	**DAVIDSON 1992** see previous study description under ‘Consultations’

(Continued)

Participants: 10 fund-holding (study) and six non-fund-holding practices in Oxford

Type of intervention: Payment for providing care for a patient or specific population (capitation)

Description: The fund-holding scheme gives GPs control over budgets to cover prescriptions, specialist outpatients consultations, and elective surgical procedures for their patients (see detailed description under Sturm 2007)

Level: not stated

Payment frequency: not stated

Comparison: fundholding vs. no intervention (=control)

Duration: 6 months pre-intervention, and six months post-intervention

Expected change: decrease in NHS referral rates and increase in referrals to private clinics

NHS Outpatient referral rates

Standardised mean annual referral rates per 1000 population per year pre- and post-intervention

Pre intervention

Study:109.7; Control: 97.5

Post intervention

Study: 112.1; Control: 122.3 (NS)

Absolute difference (post intervention): -10.3

Relative difference (post intervention): -8.4%

Absolute difference from baseline:

Study: +2.4; Control: + 24.8

Difference in absolute change from baseline: -22.4

Private outpatient referral rate

Pre intervention

Study: 29.4; Control: 27.7

Post intervention:

Expected change: decreased number of hospitalisations

Hospitalisations per year per patient (standard error)[11]

Pre intervention

Capitation + financial risk sharing: 0.0768 (0.003) (n=764)

Fee-for-service (high rate) 0.1440 (0.022) (n=1015)

Fee-for-service (low rate)

Control 0.0864 (0.011) (n=1991)

During intervention:

Capitation+ financial risk sharing: 0.0348 (0.001) (n=764)

Fee-for-service (high rate): 0.0744 (0.007) (n=1015)

Fee-for-service (low rate) (control): 0.0552 (0.005) (n=1991)

Absolute changes-

Capitation+ financial risk sharing:

Six months before: -0.01

During study: -0.02.

Fee-for-service (high rate):

Six months before: +0.06

During study: +0.02

Relative changes-

Capitation+ financial risk sharing:

Six months before: -11%

During study: -37%.

Fee-for-service (high rate):

Six months before: +66%

During study: +35%

Significance not reported.

HUTCHISON 1996

Design: CBA

Participants: primary care physicians (39 in intervention group and 77 in control); 89 148 patients in intervention group; 180 255 in control group

Type of intervention: Payment for providing care to a patient or specific population (capitation) and payment for providing a pre-specified level of quality of care (including target payments, bonuses)

Description: PCPs who were previously paid by Fee-for-service changed their payment to a mix of capitation and an ambulato-

(Continued)

Study: 26.6; Control: 28.8 (NS)

Absolute difference (post intervention): -2.2

Relative difference (post intervention): -7.6%

Absolute difference from baseline:

Study:-2.8; Control: +1.1

Absolute change : -3.9

Significance not reported.

KAMMERLING 1996

Design: CBA

Participants: 10 fund-holding (study) and 22 non-fund-holding practices (control)

Type of intervention: Payment for providing care for a patient or specific population (capitation)

Description: The fundholding scheme gives GPs control over budgets to cover prescriptions, specialist out-patients consultations, and elective surgical procedures for their patients (see detailed description under Sturm 2007)

Level: not stated

Payment frequency: not stated

Comparison: Fund-holding scheme vs. no intervention (control)

Duration: one year pre intervention, and two years post intervention

Expected change: decrease in referrals within the NHS outpatient care

Referral rates for orthopaedic problems

Standardised attendance ratio (95% CI) [10]

Year before FH:

Fund-holders: 84.8 (79.2 to 90.0) (n=862)

Control: 93.5 (88.5 to 98.4) (n=1329)

First year after FH status:

Fundholders: 98.2 (92.1 to 104.2) (n=998)

Control: 110.2 (104.7 to 115.9) (n=1329)

Second year after FH status: Fundholders 95.9 (89.9 to 101.8) (n=975)

ry incentive payment in which the health service organisation received a payment if their hospitalisation rate was lower than the regional rate

Level: not stated

Payment frequency: not stated

Comparison: Capitation + ambulatory incentive payment (in which the health service organisation received a payment if their hospitalisation rate was lower than the regional) vs. Fee-for-service

Duration: three years pre intervention and three years post- intervention

Expected change: decreased admission rate and hospital days stayed

Differences (adjusted for age, sex and social assistance) per 1000 patients between one year before and three years after intervention for:

Hospital separations (admissions).

Absolute changes: -0.2 (*P*=0.312); Relative changes: +0.4%

Hospital length of stay (days)

Absolute changes: +3 (*P*=0.774); Relative changes: -3.7%

KRASNIK 1990

See study description under 'Consultations'

Expected change: reduced referral rate to hospitals and specialists

Differences between six months before intervention (=100) and six and 12 months after per 1000 patients in a one week period

Referrals to specialists

At six months

Intervention: 90.1 (80.7 to 107.5), (n=71); Control: 99.4, (n=326)

At 12 months

Intervention 77.0 (68.6 to 86.4), (n=71)

Control 98.1, (n=326)

Absolute changes:

At six months: -9.3

At 12 months: -21.2

Relative changes:

At six months: -1550%;

At 12 months: -1110.5%

Referrals to hospital

(Continued)

Control: 123.5 (117.7 to 129.5) (n=1754).

Orthopedic referral rates per 1000 population per year

First year after FH status:

Pre intervention:

Study 7.96; Control 8.23

Post intervention:

Study 9.21; Control 9.79

Absolute difference (post intervention): -0.58

Relative percentage difference (post intervention): -5.9%

Absolute difference from baseline:

Study +1.25; Control +1.56

Difference in absolute change from baseline: -0.31

Second year after FH status :

Pre intervention:

Study 7.96; Control 8.23

Post-intervention

Study 9.00; Control 10.97

Absolute difference post intervention -1.97

Relative percentage difference (post intervention): -18.0%

Absolute difference from baseline:

Study +1.04; Control +2.74

Difference in absolute change from baseline: -1.70

Significance not reported.

LINNALA 2001

Design: CBA

Participants: 10 GPs with a list system in 2 municipal health centres with 23,000 residents were given the opportunity to send their patient to private specialists for consultation at reduced cost. 4 GPs without a list system and working in 2 municipal health centres with 10,800 residents were the control group.

Intervention 87.4 (71.1 to 107.5) (n=71); Control 97.1 (n=326)

At 12 months

Intervention 68.4 (54.7 to 85.4), (n=71)

Control 102.1

Absolute changes

At six months: -9.7

At 12 months: -33.7

Relative changes

At six months: -334.5%

At 12 months: -1604.8%

Significance not reported.

(Continued)

Type of intervention: Payment for providing care to a patient or specific population (capitation)

Description: Patients were charged the same (lesser) rate to be seen by a private specialist as they would have been charged to see a hospital based specialist

Level: not stated

Payment frequency: not stated

Comparison: GPs (with a list system) that had the possibility to send their patients to private specialists for consultation at a reduced cost to the patients and GPs (without a list system) that did not have that possibility

Duration: March1991 to December 1993 (34 months)

Expected change: decreased referrals to hospital outpatient clinics and increased referrals to private specialists

Referrals to hospital outpatient clinics
Pre Intervention
Intervention 5.7%; Control 4.4%

Difference 1.3% $p < 0.05$

Post Intervention
Intervention 6.8%; Control 5.5%

Difference 1.3%
Percentage absolute change Intervention and control: 1.1% (P< 0.001

Referrals to private sector

Pre Intervention
Intervention 5.7%; Control 8.8%

Post Intervention
Intervention 33.6%; Control 5.6%

Percentage absolute change
Intervention 27.9%; Control -3.2%
(P<0.001)

COMPLIANCE WITH GUIDELINES (on number of patient visits, on immunisation status)	**DAVIDSON 1992** see study description under 'Consultations' **Expected change:** increased compliance with guidelines Mean number (standard error) of visits per year per patient[12]

(Continued)

Absolute changes: intervention minus control: Relative changes: percentage difference between intervention and control. Note: significance not reported

Percentage compliance with CHAP [13] guidelines for:

Number of PCP visits over a one-year period for children aged:

0-12 months (CHAP=5)

Absolute changes: -10

Relative changes: -12.7%

13-24 months (CHAP=3)

Absolute changes: -12

Relative changes: -15%

25-36 months (CHAP=2)

Absolute changes: -8

Relative changes: -9.4%

3-5 years (CHAP=1)

Absolute changes: -8

Relative changes: -8.7%

PCP and outpatient clinic visits over a one-year period for children aged:

0-12 months (CHAP=5)

Absolute changes: -11

Relative changes -13.3%

13-24 months (CHAP=3)

Absolute changes: -12

Relative changes: -15%

25-36 months (CHAP=2)

Absolute changes: -8

Relative changes: -9.1%.

3-5 years (CHAP=1)

Absolute changes: -10

Relative changes: -10.4

Significance not reported.

HICKSON 1987

See study description under 'Consultations'

Expected change: increased compliance with guidelines

(Continued)

*Average number over a nine-month period **[14]**of:*

Compliance with AAP guidelines:

% of recommended visits missed

Absolute changes:+6.4 (*P*<0.01)

Relative changes: +213%;

% of visits in excess of the recommended

Absolute changes: -13.3 (*P*<0.01)

Relative changes: 73.9%

RITCHIE 1992

Design: ITS

Participants: 95 general practices (313 PCPs) ; all PCPs who have patients with Grampian addresses, Scotland (UK); all children with Grampian addresses, Scotland (UK)

Type of intervention: Payment for providing a pre-specified level of quality of care (include target payments, bonuses)

Description: PCPs received a lower orhigher payment according to whether they immunised 70% or 90% (respectively) of the eligible population.

Level: not stated

Payment frequency: not stated

Duration: January 1990 to September 1991

Comparison: Target payments + fee-for-service vs. Fee-for-service

Expected change: improved paediatric immunisation status

Changes over a 20-month period in the number of practices achieving at least:

95% primary immunisation rates

Absolute changes: +49.5 (from 31% to 80%); Relative changes: +162.1%

90% primary immunisation rates

Absolute changes: +18.9% (from 73% to 92%); Relative changes: +26.1%

95% pre-school immunisation rates

Absolute changes: +41.1% (from 23% to 64%);Relative changes: +177.3%

90% pre-school immunisation rates

Absolute changes: +42.1% (from 38% to 80%); Relative changes: +111.1%

(Continued)

		Proportion of immunisations given by PCPs Absolute changes: +12% (from 86% to 98%);Relative changes: +14.0% Significance not reported.
	PETERSEN 2006	**STURM 2007**
PROCESSES OF CARE **(prescriptions, vaccinations, drug use etc.)**	**BEAULIEU 2005** **Design:** CBA **Participants:** 21 PCPs contracted with Independent health in upstate New York: n=476 diabetic patients, n=600 Independent Health diabetic patients were the comparison group **Type of intervention**: Payment for providing a pre-specified level of quality of care (include target payments, bonuses) **Description:** Meeting target CS of <6.23; CS of <6.86; or overall improvement in composite score. CS based on PCPs performance of process and outcome measures for diabetes care (e.g. LDL test, dilated retinal examination; LDL cholesterol level<2.59 mmol/L (<100 mg/dL) Incentive rewards:CS<6.86, $3.00 PMPM (Medicare); $0.75 (commercial); CS<6.23; $0.75PMPM (Medicare); $0.18PMPM (commercial) **Level:** physician **Payment frequency**: at the conclusion of the study **Duration of intervention:** April 2001 to January 2002 **Comparison:** Bonus vs. Control **Expected change:** improved diabetes care (i.e. improved performance of process and outcome measures) **PCP's performance of process and outcome measures for diabetes care** Patients treated by physicians in the demonstration project had statistically significant improvement (final -baseline performance) on the following process and outcomes measures ($P < 0.001$ unless otherwise noted	**UK Fund-holding scheme:** **Description**: GPs were given financial control over some of their provided services. Besides costs of prescribed drugs, practice staff and a range of secondary care such as specialist services were covered by separate budgets, with the drug budget offering the greatest saving potential. Overspending in one budget had to be covered by funds from another budget, and savings could be used in other areas of patient care. After the introduction of fundholding for GPs with the first wave of voluntary practices, each year practices with initially at least 11 000 registered patients could join the fundholding scheme in 'waves'. With each wave, regulations on requisites' for joining practices were relaxed. **BAINES 1997** **Design:** CBA **Participants:** Fund-holders: Devon (n=19) and Lincolnshire (n=22), Non-FH: Lincolnshire (n=86) and Devon (n=106) **Type of intervention:** Payment for providing care for a patient or specific population (capitation) **Level:** not stated **Payment frequency:** not stated **Duration**: unclear **Comparison:** Fund-holding vs. non-FH **Expected change:** increase in generic drug use **Generic percentage (effects on drug use)** **BAINES Lincolns 1997** Setting – Wave 1-3 Adjusted absolute change: At 12 months: not reported Adjusted relative change [%] – At 12 months: not reported

(Continued)

Second hemoglobin A1c test +25.5%

LDL cholesterol test +18.3%

Diabetic retinal examination +25.6%
Nephropathy test +37.0%

Foot examination +45.4%

Hemoglobin A1c level <9.5%; +13.9% LDL cholesterol level <2.59 mmol/L (<100mg/dL) +10.5% difference

LDL cholesterol level<3.37 mmol/L (< 130 mg/dL) +(23.5%

BP<130/80 mm Hg +6.3%; *P* <0.05).
No significant improvement for the hemoglobin A1c test

CHRISTENSEN 2000

Design: RCT (two arms)

Participants: 200 pharmacies (110 in Intervention group; 90 in control)

Type of intervention: Payment for each service/episode/case (fee-for-service)

Description : $4 for cognitive service interventions (<6 min); $6 for ?6 min; cognitive services are judgemental or educational services provided by the pharmacist to the patient, such as consulting the prescriber about a suboptimal dose

Level: provider group

Payment frequency: fee-for-service

Duration : February 1994 to September 1995 (20 months)

Comparison: Enhanced Fee-for-service vs. Control

Expected change: increased delivery of cognitive services

Cognitive services

Mean rate: 1.59 interventions per 100 Medicaid prescriptions (study pharmacies) vs. 0.67 (controls); *P* < 0.001

CLARK 1995

Design: CBA

Participants: 7 community mental health centers;185 clients (95 in traditional case managers and 90 in continuous treatment team)

At 24 months (3 year f/u): 10.7%

BAINES Devon 1997

Setting – Wave 1-3

Adjusted absolute change:

At 12 months: not reported

Adjusted relative change [%] –

At 12 months: not reported

At 24 months (3 year f/u): 9.5%

BRADLOW 1993

Design: CBA

Participants: FH (1st wave):5; Non-FH:7 practices, Oxford, UK

Type: Payment for providing care for a patient or specific population (capitation)

Level: not stated

Payment frequency: not stated

Duration: unclear

Comparison: Fundholding vs. non-FH

Expected change: reduction in prescribed items per patient and increase in generic drug use

Items per patient (effects on drug use):

Setting – Wave 1

Adjusted absolute change –

At 12 months: 40

Adjusted relative change [%] –

At 12 months: 1.8%

At 24 months: not reported

BRADLOW 1993

Setting – Wave 1

Adjusted absolute change –

At 12 months: not reported

Adjusted relative change [%] –

At 12 months: not reported

At 24 months: - 3,6% (3 year f/u/ Data from Stewart-Brown study)

(Continued)

Type: Payment for each service/episode/case (fee-for-service)

Description: Community mental health centres received $ 15.75 per 15 min spent in community setting delivering mental illness management services.

Level: provider group

Payment frequency: fee-for-service

Duration: July 1992

Comparison: Enhanced Fee-for-service vs. Control[16]

Expected change: increased community treatment time

Time spent in treatment

Average weekly time spent in community treatment per client increased after the payment change (30.71 vs. 38.61 min; *P* <0.05)

Office-based case management weekly time per client decreased (32.96 min vs. 23.31 min; *P* < 0.001).

Total case manager average weekly time per client was not significantly different (63.68 min vs. 61.93 min) after the payment change

MANOVA showed that center-based treatment time decreased (*F*-value = 10.41; *P* <0.001), and community treatment time increased (*F*-value =3.72 (*P* <0.055)

Program type and Medicaid status were not associated with change in time in community vs. mental health center.

FAIRBROTHER 1999

Design: RCT (four arms)

Participants: 60 physicians (15 bonus;15 enhanced Fee-for-service, 15 feedback only; and 15 control)

Type of intervention: Payment for providing a pre-specified level of quality of care (include target payments, bonuses) and Payment for each service/ episode/case (fee-for-service)- mixed financial incentive

Description: Bonuses: $1000 (20% improvement from baseline); $2500 (40% improvement); $5000 (80% up-to-date immunisation)

Generic percentage (effects on drug use)

Setting – Wave 1

Adjusted absolute change: 4.1

Adjusted relative change [%] –

At 12 months: 8.8%

At 24 months: not reported

BRADLOW 1993

Setting – Wave 1

Adjusted absolute change: not reported

Adjusted relative change [%] –

At 12 months: not reported

At 24 months: 17.2% (3 year f/u/ Data from Stewart-Brown study)

BURR 1992

Design: CBA

Participants: FH (1st wave): 4; Non-FH: 4 practices

Type: Payment for providing care for a patient or specific population (capitation)

Level:not stated

Payment frequency:not stated

Duration: unclear

Comparison: Fund-holding vs. non-FH

Expected change: reduction in prescribed items per patient

Items per patient (effects on drug use):

Setting – Wave 1

Adjusted absolute change –

At 12 months: 18

Adjusted relative change [%] –

At 12 months: 0.8%

At 24 months: not reported

HARRIS 1996

Design: CITS

Participants: All general practices in England, UK

(Continued)

Enhanced Fee-for-service: $5 per vaccine given within 30 d of its coming due; $15 for each visit at which >1 vaccine was due and all were given.

Level: physician

Payment frequency: every four months

Duration July 1995 to July 1996 (12 months)

Comparison: Bonus and enhanced Fee-for-service vs. Control[18]

Expected change: improved up-to-date immunisation status

Patient's up-to-date coverage for paediatric immunisations

Bonus group improved significantly in documented up-to-date immunisation status, with an overall change of 25.3% ($P < 0.01$), but none of the other groups improved significantly compared with controls.

FAIRBROTHER 2001

Design: RCT (three arms)

Participants: 57 physicians (24 bonus; 12 Fee-for-service; 12 control)

Type of intervention: Payment for providing a pre-specified level of quality of care (include target payments, bonuses) and payment for each service/ episode/case (fee-for-service)- mixed financial incentive

Description: Bonuses: $1000 (30% improvement from baseline); $2500 (45% improvement); $5000 (80% up-to-date immunisation); $7500 (90% up-to-date)

Enhanced Fee-for-service: $5 per vaccine given within 30 d of its coming due; $15 for each visit at which >1 vaccine was due and all were given.

Level: physician

Payment frequency: every four months

Duration : July 1997 to July 1998

Comparison: Bonus and Fee-for-service vs. Control[19]

Expected change: improved up-to-date immunisation status

Type of intervention: Payment for providing care for a patient or specific population (capitation)

Level: not stated

Payment frequency: not stated

Comparison: Fund-holding vs. non-FH

Duration: unclear

Expected change: reduction in prescribed items per patient

Items per patient (effects on drug use):

Setting – Wave 1

Absolute level effect (95% CI): 0.4 (-1.1 to 1.8)

At 3 months: 0.4 (-1.2 to 2)

At 6 months: 0.7 (-1.3 to 2.7)

At 12 months: 1.4 (-1.5 to 4.2)

At 24 months: 2.6 (-2.1 to 7.2)

Setting – Wave 2

Absolute level effect (95% CI): -0.5 (-1.3 to 0.3)

At 3 months: -0.5 (-1.3 to 0.3)

At 6 months: -0.4 (-1.3 to 0.5)

At 12 months: -0.3 (-1.4 to 0.8)

At 24 months: -0.1 (-1.7 to 1.5)

Setting – Wave 3

Absolute level effect (95% CI) :0.0 (-0.7 to 0.7)

At 3 months: 0.0 (-0.8 to 0.8)

At 6 months: 0.0 (-0.8 to 0.9)

At 12 months: 0.2 (-0.7 to 1.2)

At 24 months: 0.4 (-0.7 to 1.6)

Setting – Wave 4

Absolute level effect (95% CI): 0.3 (-0.4 to 1)

At 3 months: 0.3 (-0.4 to 1.1)

At 6 months: 0.1 (-0.6 to 0.9)

At 12 months: -0.4 (-1.2 to 0.5)

At 24 months: not reported

Setting – Wave 5

Absolute level effect (95% CI): -0.2 (-1 to 0.5)

(Continued)

Patient's up-to-date coverage for paediatric immunisations

Both the bonus and the enhanced Fee-for-service groups improved significantly in documented up-to-date immunisation status, with an overall change of 5.9% ($P < 0.05$) and 7.4% ($P < 0.01$), respectively, compared with the control group.

KOUIDES 1998

See study description under 'Processes of care'

Expected change: increased immunisation rate

Immunisation rate

Absolute increase in immunisation rates (from 1990 [baseline] to 1991) was 6.8%; $P = 0.03$

ROSENTHAL 1995

Design: CBA

Participants: 163 provider groups contracted with PacifiCare Health Systems in California (provider groups in the Pacific Northwest were the comparison group)

Type of intervention: Payment for providing a pre-specified level of quality of care (include target payments, bonuses)

Description: Incentive payout based on provider's groups' ability to reach or exceed target rates for cervical cancer screening, mammography and haemoglobin A1C testing for diabetic patients.

Level: provider group

Payment frequency: quarterly

Duration : October 2001 to April 2004 (10 months)

Comparison: Bonus vs. Control [17]**Expected change:** increased cervical cancer screening rates, mammography screening rates and hemoglobin A1c testing

Cervical screening rate, mammography screening rates, hemoglobin A1c testing

Improvement in cervical cancer screening rates was statistically significant between the intervention and comparison groups (difference, 3.6%; $P = 0.02$).

At 3 months: -0.2 (-1 to 0.5)

At 6 months: -0.2 (-1 to 0.6)

At 12 months: not reported

At 24 months: not reported

RAFFERTY 1997

Design: Controlled Interrupted Time Series (CITS)

Participants: FH (1st wave):23; FH (2nd wave):34; FH (3rd wave):9;Non-FH: all in Northern Ireland

Type of intervention: Payment for providing care to a patient or specific population (capitation)

Level: not stated

Payment frequency: not stated

Comparison: Fundholding vs. non-FH

Duration: unclear

Expected change: reduction in prescribed items per patient and increase in generic drug use

Items per patient (effects on drug use):

Setting – Wave 1

Absolute level effect (95% CI): -63.6 (-249.3 to 122.1)

Relative change (95% CI)

At 3 months: -2.5 (-9.8 to 4.9)

At 6 months: -1 (-8.8 to 6.8)

At 12 months: -2.8 (-11.5 to 5.9)

At 24 months**:** 0.2 (-10.3 to 10.7)

Setting – Wave 2

Absolute level effect (95% CI): -43.6 (-257 to 169.8)

At 3 months: -1.6 (-9.2 to 6)

At 6 months: -2.4 (-10.3 to 5.5);

At 12 months: -3.6 (-12.1 to 4.8)

At 24months: -4.2 (-13.7 to 5.4)

Setting – Wave 3

Absolute level effect (95% CI) 44.3 (-280.1 to 191.4)

At 3 months: -1.4 (-9.9 to 7)

At 6 months: 1.5 (-7.2 to 10.1)

At 12 months: 1.5 (-7.5 to 10.5

(Continued)

ROSKI 2003

Design: RCT (3 arms)

Participants: 37 primary care sites (13 incentive; 9 incentive +registry and 15 control)

Type of intervention: Payment for providing a pre-specified level of quality of care (include target payments, bonuses)

Description:75% of patients with smoking status identified/ documented at last visit; 65% of patients with quitting advice documented at last visit (targets set at approximately 15% above the average from 2 years before study); bonuses, $5000 for sites with 1-7 providers and $10 000 for sites with ?8 providers

Level: provider group

Payment frequency; one time (end of study)

Duration : May 1999 to June 2000

Comparison: Bonus vs. Control (=no intervention)

Expected change: more patients with smoking status identified and given quitting advice

Changes in tobacco use status identification:

Incentive group +14%

Incentive + registry group +8.1% Control group + 6.2%; $P < 0.009$

Providing quitting advice to patients:

Incentive group +24.2%

Incentive + registry + 18.3%

Control +18.3%

No significant difference across the study groups

Quitting rate (7-days sustained abstinence)

Incentive group:+22.4%

Incentive +registry group:+21.7% Control group:+19.2%

At 24 months: not reported

Generic percentage (effects on drug use)

Setting – Wave 1

Absolute level effect (95% CI): 2.8 (1.5 to 4.1)

At 3 months: 10.8 (5.6 to 16)

At 6 months: 12.7 (7.1 to 18.2);

At 12 months: 15.8 (9.4 to 22.2)

At 24 months: 23 (15 to 31)

Setting – Wave 2

Absolute level effect (95% CI):1.3 (-0.2 to 2.9)

At 3 months: 5.1 (-0.9 to 11.1)

At 6 months: 5.9 (-0.4 to 12.2)

At 12 months: 8.5 (1.6 to 15.5)

At 24months: 13.6 (5.4 to 21.7)

Setting – Wave 3

Absolute level effect (95% CI):0.5 (-1 to 1.9)

At 3 months: 1.8 (-3.9 to 7.4)

At 6 months:5.7 (-0.1 to 11.5)

At 12 months: 14.2 (8.1 to 20.4)

At 24 months: not reported

RAFFERTY 1997

Design: CBA, see study description under 'Processes of care'

Expected change: reduction in prescribed items per patient and increase in generic drug use

Items per patient (effects on drug use)

Setting – Wave 1

Adjusted absolute change –

At 12 months: -461

Adjusted relative change [%] –

At 12 months: -4%

At 24 months: -5.2%

Setting – Wave 2

Adjusted absolute change –

At 12 months: -218

(Continued)

No significant difference across the study groups.

SHEN 2003

Design: CBA

Participants: 36 skilled nursing facilities (18 study facilities and 18 control facilities); 5552 clients (2367 office of substance abuse clients; 3185 Medicaid clients)

Type of intervention:: Payment for providing a pre-specified level of quality of care (include target payments, bonuses)

Description: Additional funds based on efficiency, effectiveness and service to special populations.[15]

Level: payment system

Payment frequency: yearly

Comparison: Performance based contracting (PBC) with additional funds based on efficiency, effectiveness and service to special populations vs. Medicaid (=control)

Duration: 1991 to 1995

Expected change: improved health service to special populations

Health service utilisation

After PBC implementation there was a significant decrease in the likelihood that an 'office of substance abuse' patient was a "most severe user" compared with the likelihood of a Medicaid (control) patient; coefficient = -0.74; *t*-value = 3.26; $P < 0.01$

Adjusted relative change [%] –

At 12 months: -1.8%

At 24 months: -2.6%

Setting – Wave 3

Adjusted absolute change –

At 12 months: -211

Adjusted relative change [%] –

At 12 months: -1.7%

At 24 months: not reported

Generic percentage (effects on drug use)

Setting – Wave 1

Adjusted absolute change:

At 12 months: 3.2

Adjusted relative change [%] –

At 12 months: 12.7%

At 24 months:16.1%

Setting – Wave 2

Adjusted absolute change:

At 12 months: 2.4

Adjusted relative change [%] –

At 12 months: 9.5%

At 24 months:13.6%

Setting – Wave 3

Adjusted absolute change:

At 12 months: 3.4

Adjusted relative change [%] –

At 12 months: 13.4%

At 24 months):not reported

WHYNES 1997

Design: CBA

Participants: FH (4^{th} wave):23; Non-FH:63 practices

Type of intervention: Payment for providing care for a patient or specific population (capitation)

(Continued)

Level: not stated

Payment frequency: not stated

Duration: not clear

Comparison: Fund-holding vs. non-FH

Expected change: reduction in prescribed items per patient and increase in generic drug use

Items per patient (effects on drug use):

Setting – Wave 4

Adjusted absolute change –

At 12 months: not reported

Adjusted relative change [%] –

At 12 months:-1.2

At 24 months: not reported

Generic percentage (effects on drug use)

Setting – Wave 4

Adjusted absolute change:

At 12 months: 3.5

Adjusted relative change [%] –

At 12 months: not reported

At 24 months: not reported

WILSON 1995

Design: CITS

Participants: FH (1st wave):20; FH (2:nd wave): 31 and FH (3rd wave):49; Non-FH: 312 practices

Type of intervention: Payment for providing care for a patient or specific population (capitation)

Level: not stated

Payment frequency: not stated

Duration: unclear

Comparison: Fund-holding vs. non-FH

Expected change: reduction in prescribed items per patient and increase in generic drug use

Items per patient (effects on drug use):

Setting – Wave 1

Absolute level effect (95% CI): 1.4 (-6.6 to 9.4)

(Continued)

At 3 months: 1.9 (-9.3 to 13.1)

At 6 months: -4.1 (-4.3 to -4)

At 12 months: -10.2 (-10.4 to -10)

At 24 months: not reported

Setting – Wave 2

Absolute level effect (95% CI): 2.7 (-9.5 to 14.9)

At 3 months: 7.1 (-25.1 to 39.2)

At 6 months: -15.8 (-16.1 to -15.5)

At 12 months: -14.5 (-15.2 to -13.9);

At 24 months: not reported

Setting – Wave 3

Absolute level effect (95% CI): 4.8 (-4.8 to 14.4)

At 3 months: 16.8 (-17.1 to 50.8)
At 6 months: -21.3 (-21.6 to -20.9)

At 12 months: -28.9 (-29.4 to -28.3)

At 24 months: not reported

Generic percentage (of items per patient)

Setting – Wave 1

Absolute level effect (95% CI): 1.7 (0.8 to 2.7)

At 3 months: 345.7 (151.8 to 539.6)

At 6 months: 342.7 (341.1 to 344.4)

At 12 months: 190.5 (189 to 192)

At 24 months):not reported

Setting – Wave 2

Absolute level effect (95% CI): 1.0 (-0.1 to 2.1)

At 3 months: 45.4 (-2.4 to 93.2)

At 6 months: 66.5 (66.1 to 66.8)

At 12 months: 68.1 (67.6 to 68.7

At 24months):not reported

Setting - Wave 3

Absolute level effect (95% CI): 1.9 (0.8 to 3)

At 3 months: 35.5 (15.1 to 55.9)

At 6 months: -12.2 (-12.4 to -12.1)

At 12 months: -43.7 (-44.0 to -43.5)

(Continued)

At 24 months): not reported

WILSON 1995

Design: CBA

Participants (practices) FH (1st wave):20; FH (2nd wave): 31 and FH (3rd wave):49; Non-FH: 312 practices

Type: Payment for providing care for a patient or specific population (capitation)

Level: not stated

Payment frequency:not stated

Duration : unclear

Comparison: Fund-holding vs. non-FH

Expected change: reduction in prescribed items per patient and increase in generic drug use

Items per patient (effects on drug use):

Setting – Wave 1(3 year f/u)

Adjusted absolute change –

At 12 months: not reported

Adjusted relative change [%] –

At 12 months: -5.7% (Median)

At 24 months: not reported

Setting – Wave 2 (Median)

Adjusted absolute change –

At 12 months: not reported

Adjusted relative change [%] –

At 12 months: not reported

At 24 months (3 year f/u): 0.8%

Setting – Wave 3 (Median)

Adjusted absolute change –

At 12 months: not reported

Adjusted relative change [%] –

At 12 months: not reported

At 24 months (3 year f/u): -5.6%

Generic percentage (effects on drug use)

(Continued)

Setting – Wave 1

Adjusted absolute change:

At 12 months: not reported

Adjusted relative change [%] –

At 12 months: not reported

At 24 months (3 year f/u):12.1%

Setting – Wave 2

Adjusted absolute change:

At 12 months: not reported

Adjusted relative change [%] –

At 12 months: not reported

At 24 months (3 year f/u): 10,1%

Setting – Wave 3

Adjusted absolute change:

At 12 months: not reported

Adjusted relative change [%] –

At 12 months: not reported

At 24 months (3 year f/u):10.3%

WILSON 1999

Design: CBA

Participants: 5 health authorities in NW-region, UK

Type of intervention: Payment for providing care for a patient or specific population (capitation)

Level: not stated

Payment frequency: not stated

Duration: unclear

Comparison: Fund-holding vs. non-FH

Expected change: reduction in prescribed items per patient and increase in generic drug use

Items per patient (effects on drug use):

Setting – Wave 3/4

Adjusted absolute change –

At 12 months: not reported

(Continued)

Adjusted relative change [%] –

At 12 months: not reported

At 24 months:39.2% (Combined wave 4:1 year f/u; wave 3: 2 year f/u)

Generic percentage (effects on drug use)

Setting – Wave 3/4

Adjusted absolute change:

At 12 months: not reported

Adjusted relative change [%] –

At 12 months: not reported

At 24 months: 4% (Combined wave 4:1 year f/u; wave 3: 2 year f/u)

All anti-ulcer drugs (DDD)

Setting – Wave 3/4

Adjusted absolute change:

At 12 months: not reported

Adjusted relative change [%] –

At 12 months: not reported

At 24 months; -6.7% (Combined wave 4:1 year f/u; wave 3: 2 year f/u)

Percentage PPI of all anti-ulcer drugs (DDD)

Setting – Wave 3/4

Adjusted absolute change:

At 12 months: not reported

Adjusted relative change [%] –

At 12 months: not reported

At 24 months: -7.9% (Combined wave 4:1 year f/u; wave 3: 2 year f/u)

All anti-depressant drugs (DDD)

Setting – Wave 3/4

Adjusted absolute change:

At 12 months: not reported

Adjusted relative change [%] –

At 12 months: not reported

(Continued)

At 24 months: -7.9% (Combined wave 4:1 year f/u; wave 3: 2 year f/u)

Percentage SSRIs of all anti-depressant drugs

Setting – Wave 3/4

Adjusted absolute change:

At 12 months: not reported

Adjusted relative change [%] –

At 12 months: not reported

At 24 months: -0.8%(Combined wave 4:1 year f/u; wave 3: 2 year f/u)

Drug-budgets:

GUETHER 1997

Design: ITS

Participants: 82 GP's, West Germany Statutory Health insurance:

Type of intervention: unclear

Description: Collective budgets for drug expenditure for physicians in private practice. While spending caps were regionally negotiated or nationally set each year and made all physicians in private practice in one region collectively liable, target volumes for each individual practice were only theoretically established.

Level: not stated

Payment frequency:not stated

Duration: unclear

Comparison: Drug-budgets vs. no intervention

Expected change: reduction in prescribed items per patient

Items per patient

Setting - Social insurance

Absolute level effect (95% CI):

-34552 (-99896 to 30791)

Relative change (95%CI)

At 3 months: -11.2 (-32.3 to 10.0)

At 6 months: -12.1 (-37.8 to 13.7)

At 12 months: -13.4 (-48.9 to 22.1)

At 24months: – n.a.

WALLEY 2000

(Continued)

Design: ITS

Participants: 223 GPs from Eastern Health Board Cohort, Ireland

Type of intervention:: unclear

Description: GPs individual indicative or hypothetical budgets covered prescribing costs and were calculated based on previous spending and the national average. Savings were split between the GP and the local health authority to be used for the development of services. There were no penalties for overspending.

Level: not stated

Payment frequency: not stated

Duration: unclear

Comparison: Drug-budgets vs. no intervention (Ireland indicative drug targeting saving scheme, IDTSS)

Expected change: reduction in prescribed items per patient

Items per patient (effects on drug use):

Absolute level effect (95% CI): -0.8 (-1.4 to -0.2)

Relative change (95%CI) –

At 3 months: not reported

At 6 months: not reported

At 12 months: -8.2 % (-14.4 to -2.0)

At 24 months: -10.1% (-17.5 to -2.7)

REFERRALS	**GRADY 1997** **Design:** RCT (3 arms) **Participants:**61 community-based primary care practices (20 cue and reward; 18 cue; 23 control) **Type:** Payment for providing a pre-specified level of quality of care (include target payments, bonuses) **Description:** Individualised feedback was provided showing the particular physicians' percentage of referrals and patient completions and mammography compliance rates for each physician's patients compared with the averages for all physicians in the study. The physicians also received a token reward: a check based on the percentage referred during each quarterly audit period, i.e. $50 for 50% referral rate. The mammography compliance rates for the physicians in the cue and reward group was compared with the averages of all physicians in the study. In order to have	**GUETHER 1997** see the study description above under 'Processes of care' **Expected change:**increase in referrals for socially insured patients **Referrals to outpatient specialists** Absolute level effect (95% CI) 1543 (-5095.6 to 8181.7) Relative change (95%CI) At 3 months: 3.4 (-11.3 to 18.1) At 6 months: -3.5 (-21.9 to 14.9) At 12 months: -15.4 (-40.3 to 9.5) At 24 months: not reported **KAMMERLING 1996** **Design:** CBA, see study description under Akbari 2008

(Continued)

comparison data the feedback and token rewards were not begun until the second half of the first year. The only difference from the cue group was the monetary reward.

Level: physician

Payment frequency: 1 per quarterly audit; rewards given last two quarters

Duration:6 months

Comparison: Bonus vs. control[20]

Expected change: increased referral rates for mammography

Mammography referral rates

*Mean referral rates (SD)**[21]** for the first year of the intervention*

Control: Baseline 17.0 (11.6) (n=23); 1st quarter 1: 23.0(14.0);

2nd quarter 19.1(10.4);

3rd quarter 20.7(13.7);

4th quarter 4: 16.8(11.6);

Over the whole year: 36.5 (15.1)

Cue : Baseline 25.8 (15.7) (n=18)

1st quarter 37.0 (19.0)

2nd quarter 29.8 (18.7)

3rd quarter 32.8 (19.4)

4th quarter 29.0 (14.9)

Over the whole year 50.8 (20.5)

Cue and reward: Baseline 19.0 (14.5) (n=20)

1st quarter 32.7(14.5)

2nd quarter 24.6(15.2)

3rd quarter 26.9(14.6)

4th quarter 24.7(14.9)

Over the whole year: 45.0(17.8)

The financial incentive arm was not significantly different from the control arm.

Expected change: decrease in referrals within the NHS outpatient care

Referrals to NHS outpatient care

Adjusted absolute change –

At 12 months: -18.9

Adjusted relative change [%] –

At 12 months: not reported

Over 24 months: -15.3%

(see Akbari 2008 for more detailed results)

SCHOFFSKI 1997

Design: ITS

Participants: 309-382 practices, (Germany Statutory Sickness Funds)

Type of intervention: unclear

Description: see description of the drug-budgets under Guether 1995 and 'Processes of care'

Level:not stated

Payment frequency:not stated

Duration: unclear

Comparison: Drug-budgets (German) vs. no intervention

Expected change:increase in referrals for socially insured patients

Referrals to outpatient clinics *:*

Absolute level effect (95% CI) 7.5 (-2 to 17)

Relative change (95%CI) –

At 3 months: 22.8 (-6 to 51.6)

At 6 months: 8.4 (-25 to 41.8)

At 12 months: 13.2 (-59.3 to 85.7)

At 24 months: not reported

Referrals to hospitals (health care utilisation)

Absolute level effect (95% CI)

0.1 (0 to 0.2)

Relative change (95%CI)

(Continued)

	At 3 months: 13.3 % (1.2 to 25.5) At 6 months:10.8 % (-3.1 to 24.7) At 12 months:13.3 % (-16.6 to 43.2) At 24 months: not reported
PRESCRIBING COSTS	
	BAINES 1997 **Design:** CBA, see previous study description under 'Processes of care' **Expected change:** reduced drug expenditure (£) per patient **Cost per patient (effects on drug expenditure)** *Setting – Wave 1-3, Lincolns* Adjusted relative change [%] – At 24 months: -18.5 % (3 year f/u) Setting – Wave 1-3, *Devon* Adjusted relative change [%] – At 24 months: -16.4% (3 year f/u) **BRADLOW 1993** **Design:** CBA, see study description under 'processes of care' **Expected change:** reduced drug expenditure (£) per item and per patient **Cost per item** Setting – Wave 1 Adjusted absolute change – At 12 months: -0.5 Adjusted relative change [%] – At 12 months: -6.3% At 24 months: not reported **Cost per item** *Setting – Wave 1* Adjusted absolute change – At 12 months: not reported Adjusted relative change [%] – At 12 months**:** not reported

(Continued)

At 24 months: -5.2 (3 year f/u/ Data from Stewart Brown study)

Cost per patient (effects on drug expenditure)

Setting – Wave 1

Adjusted absolute change –

At 12 months**:** -0.8

Adjusted relative change [%] –

At 12 months**:** -4.6

At 24 months: not reported

Cost per patient (effects on drug expenditure)

Setting – Wave 1

Adjusted absolute change –

At 12 months**:** -1.1

Adjusted relative change [%]:

At 12 months:-6.2

At 24 months:0.4 (3 year f/u/ Data from Stewart Brown study)

BURR 1992

Design: CBA, see description of study under 'Processes of care'

Expected change: reduced drug expenditure (£) per patient

Cost per patient (effects on drug expenditure)

Setting – Wave 1

Adjusted absolute change -

At 12 months: -0.6

Adjusted relative change [%] –

At 12 months: -4.5%

At 24 months: not reported

CORNEY 1997

Design: CBA

Participants: FH (1st wave):4;Non-FH:4 practices (South Thames Region, UK)

Type of intervention: Payment for providing care for a patient or specific population (capitation)

Level: not stated

Payment frequency: not stated

Duration: unclear

Comparison: Fundholding vs. non-FH

(Continued)

Expected change: reduced drug expenditure (£) per patient

Cost per patient (effects on drug expenditure)

Setting – Wave 2

Adjusted absolute change -

At 12 months**:** 0.2

Adjusted relative change [%] –

At 12 months: 0.5%

At 24 months: -4.8%

HARRIS 1996

All Harris outcomes: percentages of non-fundholders

See study description under 'Processes of care'

Expected change: reduced drug expenditure (£) per patient and reduced total prescribing cost

Cost per patient (effects on drug expenditure)

Setting – Wave 1

Absolute level effect (95% CI): -1.2 (-3 to 0.7)

Relative change (95%CI)

At 3 months: -1.2 (-3.1 to 0.7)

At 6 months: -0.8 (-3.3 to 1.7)

At 12 months: 0.1 (-4 to 4.2)

At 24 months:2 (-5.9 to 10)

Setting – Wave 2

Absolute level effect (95% CI): -2.9 (-4.1 to -1.7)

Relative change (95%CI)

At 3 months: -2.9 (-4.1 to -1.7)

At 6 months: -2.8 (-4.1 to -1.4)

At 12 months: -2.5 (-4.1 to -0.9)

At 24 months:-2 (-4.3 to 0.3)

Setting – Wave 3

Absolute level effect (95% CI): -0.6 (-2 to 0.7)

Relative change (95%CI)

At 3 months: -0.6 (-2 to 0.7)

At 6 months: -0.6 (-2 to 0.9)

(Continued)

At 12 months: -0.5 (-2.3 to 1.4)

At 24 months: -0.3 (-3.4 to 2.8)

Setting – Wave 4

Absolute level effect (95% CI): -1.5 (-2.9 to 0)

Relative change (95%CI)

At 3 months: -1.5 (-3 to 0)

At 6 months: -1.9 (-3.4 to -0.5)

At 12 months: -2.8 (-4.5 to -1.2)

Setting – Wave 5

Absolute level effect (95% CI) : -1.2 (-2.3 to -0)

Relative change (95%CI)

At 3 months: -1.2 (-2.4 to -0)

At 6 months: -2.1 (-3.1 to -1)

At 12 months: not reported

At 24 months: not reported

Change in total prescribing cost

Setting – Wave 2

Absolute level effect (95% CI): -1.4 (-3.6 to 0.9)

Relative change (95%CI)

At 3 months: 37.6 (-24.1 to 99.3)

At 6 months: 13.4 (-57.2 to 84.1)

At 12 months: -27.3 (-109.4 to 54.9)

At 24 months:-89.6 (-183.6 to 4.4)

Setting – Wave 3

Absolute level effect (95% CI): 1(-1.5 to 3.4)

Relative change (95%CI)

At 3 months: -18.8 (-65.6 to 28.4

At 6 months: -35.9 (-87.6 to 15.8)

At 12 months: -69.6 (-127.4 to -11.9)

At 24 months: -97 (-160.7 to -33.3)

Setting – Wave 4

Absolute level effect (95% CI): -0.3 (-3.7 to 3)

Relative change (95%CI)

(Continued)

At 3 months: 10.3 (-90.6 to 111.2)

At 6 months: -14.2 (-121.6 to 93.3)

At 12 months: -50.6(-166.2 to 65.1)

At 24 months: not reported

Setting – Wave 5

Absolute level effect (95% CI) : -0.9 (-3 to 1.2)

Relative change (95%CI)

At 3 months: 38.7 (-50.5 to 127.9)

At 6 months: 21.2 (-63.9 to 106.2)

At 12 months: not reported

At 24 months: not reported

HARRIS 1996

Design: CBA, see study description under 'Processes of care'

Expected change: reduced drug expenditure (£) per patient

Cost per patient (effects on drug expenditure)

Setting – Wave 1

Adjusted absolute change -

At 12 months: -1.2

Adjusted relative change [%]:

At 12 months:-3.2%

At 24 months:-7.7%

Setting – Wave 2

Adjusted absolute change -–

At 12 months: -1.7

Adjusted relative change [%]:

At 12 months: -4%

At 24 months**:** -6.4%

Setting – Wave 3

Adjusted absolute change -

At 12 months**:** -1.8

Adjusted relative change [%]:

At 12 months: -3.7%

At 24 months: -4.4%

Setting – Wave 4

(Continued)

Adjusted absolute change –

At 12 months: -1.8

Adjusted relative change [%] :

At 12 months: -3.4%

At 24 months: -5.6%

Setting – Wave 5

Adjusted absolute change -

At 12 months: -1.9

Adjusted relative change [%]:

At 12 months: -3.4%

At 24 months: not reported

RAFFERTY 1997

All Rafferty outcomes: difference of mean (cost per item results for year 3 were not re-analysable)

Design: CITS- see study description under 'Processes of care'

Expected change: reduced drug expenditure (£) per item and per patient

Cost per item (effects on drug expenditure)

Setting – wave 1

Absolute level effect (95% CI): -0.4 (-0.8 to 0)

Relative change (95%CI)

At 3 months: -4.9 (-10.1 to 0.4)

At 6 months: -5.8 (-11.3 to -0.3)

At 12 months: -7 (-13 to -1)

At 24 months: -9.2 (-16.1 to -2.3)

Setting – Wave 2

Absolute level effect (95% CI): -0.3 (-0.8 to 0.2)

Relative change (95%CI)

At 3 months: -3.5 (-9.2 to 2.2)

At 6 months: -4.2 (-10.1 to 1.6)

At 12 months: -6.2 (-12.4 to 0)

At 24 months: -9.8 (-16.7 to -3)

Cost per patient (effects on drug expenditure)

Setting – wave 1

(Continued)

Absolute level effect (95% CI): -922.7 (-2045.8 to 200.4)

Relative change (95%CI)

At 3 months: -4.9 (-10.8 to 1.1)

At 6 months: -4 (-10.2 to 2.3)

At 12 months: -7.3 (-14.2 to -0.4)

At 24 months: -9.1 (-17.1 to -1.1)

Setting – Wave 2

Absolute level effect (95% CI): -566.6 (-1594.6 to 461.4)

Relative change (95%CI)

At 3 months: -2.6 (-7.3 to 2)

At 6 months: -3.4 (-8.2 to 1.4)

At 12 months: -6.7 (-11.7 to -1.6)

At 24 months: -11 (-16.5 to -5.5)

Setting – Wave 3

Absolute level effect (95% CI): -192.6 (-1482.6 to 1097.5)

Relative change (95%CI)

At 3 months; -0.6 (-6 to 4.9)

At 6 months: -2.3 (-7.9 to 3.3)

At 12 months:-5.6 (-11.3 to 0.2)

At 24 months: - n.a.

RAFFERTY 1997

Design: CBA, see study description under 'Processes of care'

Expected change: reduced drug expenditure (£) per item and per patient

Cost per item

Setting – Wave 1

Adjusted absolute change -

At 12 months **:** -0.4

Adjusted relative change [%]

At 12 months: -5.5%

At 24 months: -8.1%

Setting – Wave 2

Adjusted absolute change -

At 12 months : -0.5

(Continued)

Adjusted relative change [%]:

At 12 months**:** -5.3%

At 24 months**:** -9.9%

Setting – Wave 3

Adjusted absolute change -

At 12 months**:** -0.5

Adjusted relative change [%]:

At 12 months: -5.3%

At 24 months: n.a.

Cost per patient (effects on drug expenditure)

Setting – Wave 1

Adjusted absolute change -

At 12 months: -8.4

Adjusted relative change [%]:

At 12 months: -9.5%

At 24 months: -15.3%

Setting – Wave 2

Adjusted absolute change -

At 12 months: -7.2

Adjusted relative change [%]:

At 12 months: -7.2%

At 24 months: -13.9%

Setting – Wave 3

Adjusted absolute change -

At 12 months: -7.7

Adjusted relative change [%] –

At 12 months: -7%

At 24 months: not reported

WALLEY 2000

Design: ITS, see previous study description under 'Processes of care'

Expected change: reduced drug expenditure (£) per item and total prescribing cost

Cost per item (British £)

Absolute level effect (95% CI): 0.1 (-2.5 to 2.8)

(Continued)

Relative change (95%CI) –

At 3 months: not reported

At 6 months: not reported

At 12 months: 0.6 (-10.1 to 11.7)

At 24 months: 1.2 (-12.9 to 15.3)

Change in total prescribing cost

Absolute level effect (95% CI): -5.2 (-10 to -0.4)

Relative change (95%CI) –

At 3 months: not reported

At 6 months: not reported

At 12 months: -18.0 (-34.6 to -1.4)

At 24 months : -21.7 (-41.7 to -1.8)

WHYNES 1997

Design: CBA, see study description under 'Processes of care'

Expected change: reduced drug expenditure (in British £) per patient

Cost per patient (effects on drug expenditure)

Setting – Wave 4

Adjusted absolute change –

At 12 months: -0.7

Adjusted relative change [%] –

At 12 months: not reported

At 24 months: not reported

WILSON 1995

All Wilson outcomes are reported as a difference between medians

Design: CITS

Participants: FH (1st wave):20' FH (2nd wave): 31 and FH (3rd wave):49; Non-FH: 312 practices

Type: Payment for providing care for a patient or specific population (capitation)

Duration: unclear

Comparison: Fund-holding vs. non-FH

Expected change: reduced drug expenditure (£) per item and per patient

(Continued)

Cost per item

Setting – Wave 1

Absolute level effect (95% CI): -0.2 (-0.3 to -0.1)

Relative change (95%CI)

At 3 months: -31.4 (-50 to -13.1)

At 6 months: -41.6 (-41.8 to -41.4)

At12 months: -47.8 (-48.2 to -47.5)

At 24 months: not reported

Setting – Wave 2

Absolute level effect (95% CI):-0.2 (-0.4 to -0)

Relative change (95%CI)

At 3 months: -36.9 (-71.1 to -2.7)

At 6 months: -45.1 (-45.5 to -44.7)

At12 months: -49.2 (-49.9 to -48.5)

At 24 months: not reported

Setting – Wave 3

Absolute level effect (95% CI):-0.3 (-0.5 to -0.1)

At 3 months: -99.6 (-157.4 to -41.8)

At 6 months: -85.3 (-86 to -84.6)

At 12 months: -44.3 (-49.9 to -42.9)

At 24 months: not reported

Cost per patient (effects on drug expenditure)

Setting – Wave 1

Absolute level effect (95% CI): -0 (-0.1 to 0.1)

Relative change (95%CI)

At 3 months: -6 (-26.5 to 14.6)

At 6 months: 6.7 (6.5 to 6.9)

At 12 months: 1 (0.6 to 1.3)

At 24 months: not reported

Setting – Wave 2

Absolute level effect (95% CI): -0.1 (-0.2 to -0)

Relative change (95%CI)

(Continued)

At 3 months: -166.8 (-306.9 to -26.5)

At 6 months:128.6 (127.9 to 129.4)

At 12 months: 66.8 (65.6 to 67.9)

At 24 months: not reported

Setting – Wave 3

Absolute level effect (95% CI): -0 (-0.1 to 0.1)

Relative change (95%CI)

At 3 months: -1.2 (-42.4 to 39.9)

At 6 months: -61.5 (-61.8 to -61.2)

At 12 months: -79.7 (-80.2 to -79.3);

At 24 months: not reported

WILSON 1995

Design: CBA, see description of study under 'Processes of care'

Expected change: reduced drug expenditure (£) per item and per patient

Cost per item

Setting – Wave 1-

Adjusted absolute change -

At 12 months: not reported

Adjusted relative change [%] –

At 12 months: not reported

At 24 months (3 year f/u): -0.9%

Setting – Wave 2

Adjusted absolute change -–

At 12 months: not reported

Adjusted relative change [%] –

At 12 months: not reported

At 24 months (3 year f/u):0.3%

Setting– Wave 3

Adjusted absolute change -

At 12 months: not reported

Adjusted relative change [%] –

At 12 months: not reported

At 24 months (3 year f/u): -0.3%

Cost per patient (effects on drug expenditure)

(Continued)

Setting – Wave 1

Adjusted relative change [%] :

At 24 months (3 year f/u): -7.9

Setting – Wave 2

Adjusted relative change [%]:

At 24 months (3 year f/u) -7.1

Setting – Wave 3

Adjusted relative change [%] :

At 24 months (3 year f/u): -2.7

WILSON 1999

Design: CBA

Participants: 5 health authorities in the NW-region

Type: Payment for providing care for a patient or specific population (capitation)

Level: not stated

Payment frequency: not stated

Duration: unclear

Comparison: Fund-holding vs. non-FH

Expected change: reduced drug expenditure (£) per item, per item PPIs (proton pump inhibitors), per item SSRIs (selective serotonin reuptake inhibitors), reduced drug expenditure per patient for all anti ulcer drugs and all anti depressant drugs

Cost per item (British £)

Setting – Wave ¾

Adjusted absolute change –

At 12 months: not reported

Adjusted relative change [%] –

At 12 months: not reported

At 24 months**:** -2.8% (Combined wave 4: 1 year f/u; wave 3:2 year f/u)

Cost per item PPIs (in British £)

Setting – Wave ¾

Adjusted absolute change -

At 12 months: not reported

(Continued)

	Adjusted relative change [%] – At 12 months: not reported At 24 months**:** -1% (*Combined wave 4: 1year f/u, wave 3: 2 year f/u)* **Cost per item SSRIs** (in British £) Setting – Wave 3/4 Adjusted absolute change - At 12 months: not reported Adjusted relative change [%] – At 12 months: -1.9% At 24 months: -2.7% (*Combined wave 4: 1year f/u, wave 3: 2 year f/u)* **Cost per patient - all anti ulcer drugs** (in British £) Setting – Wave 3/4 Adjusted relative change [%] **–** At 24 months**:** -10.6% (*Combined wave 4: 1year f/u, wave 3: 2 year f/u)* **Cost per patient - all anti-depressant drugs** (in British £) Setting – Wave 3/4 Adjusted relative change [%] **–** At 24 months**:** -1.9% (*Combined wave 4: 1year f/u, wave 3: 2 year f/u)*
COMPLIANCE WITH GUIDELINES	**HILLMAN 1998** **Design:** RCT (2 arms) **Participants:** 52 PC-sites (26 intervention: 26 control) **Type of intervention**: Payment for providing a pre-specified level of quality of care (include target payments, bonuses) **Description**: Compliance with cancer screening for women age >50 y; aggregate compliance scores and improvement in scores over time; full and partial bonuses (full bonus=20%; partial bonus=10% of capitation); 3 highest scoring sites received full bonus; next 3 received partial bonus; most improved sites received par-

(Continued)

tial bonus;range of bonus per site, $570 to $1260.

Level: provider group

Payment frequency: every six months

Duration: 1993 to 1995 (18 months)

Comparison: Bonus vs. control[23]

Expected change: increased guideline compliance score

Guideline compliance

Absolute increase in total mean compliance scores for intervention group from baseline was 26.3%; control group was 26.4%.

No significant differences between the groups.

HILLMAN 1999

Design: RCT (3 arms)

Participants: 49 PC sites (19 feedback plus incentive; 15 feedback only; 15 control)

Type: Payment for providing a pre-specified level of quality of care (include target payments, bonuses)

Description: Pediatric immunisations; well-child visits; bonuses based on total compliance score for quality indicators; full and partial bonuses (full bonus=20%; partial=10% of site's total 6 months capitation for pediatric members age ?6 y); 3 highest scoring sites received full bonus; next 3 received partial bonus; most improved sites received partial bonus; average bonus, $2000 (range $772 to $4682).

Level: provider group

Payment frequency: every six months

Duration: 1993 to 1995

Comparison: Feedback **+** Bonus or Feedback only vs. control[24]

Expected change: increased compliance score

Guideline compliance

(Continued)

	Absolute increase in total mean compliance scores from baseline: Feedback + bonus: 17.2%; Feedback only: 22.6% Control: 22.6% Differences in compliance score improvement between groups: Feedback+ Incentive vs. Control, 5.9%; Feedback only vs. Control, 11.3% No significant differences between the groups.

* Any data errors found in the original reviews were corrected.

[1] Fees set as approximately twice the normal Medicaid fees.

[2] Low rate as defined by normal Medicaid fees. The fees depend on type of visit.

[3] Standard errors retrieved from the original paper (Davidson 1992)

[4] Referred to as 'post-intervention' in the review by Akbari 2008, but as 'during demonstration' in the original paper by Davidson 1992, which we have changed to 'during intervention' to avoid confusion.

[5] See footnote 3

[6] See footnote 3

[7] Data on estimated changes in number of contacts and confidence intervals retrieved from the original paper (Krasnik 1990)

[8] Data on mean number of visits retrieved from the original study (Hickson 1987)

[9] See footnote 7

[10] Standardised attendance ration (CI) retrieved from the original paper (Kammerling 1996)

[11] See footnote 3

[12] See Footnote 3

[13] CHAP, New York Child Health Assurance Program periodicity schedule, based on the AAP Guidelines for Health Supervision

[14] See Footnote 7

[15] Special populations: female; 0 to 19 years; >50 years; corrections; homeless; concurrent psychological problems; history of IV drug use; polydrug use

[16] No intervention for non-Medicaid enrollees

[17] No intervention for comparison group consisiting of Pacificare, Pacific Northwest network

[18] The control condition is not specified

[19] The control condition is not specified

[20] The no intervention group received physician education, but so did all the intervention groups.

[21] Means and standard deviations are retrieved from the original paper (Grady 1997)

[22] It is not defined what constitute a type D or E patient, just that these patients are 'sicker', neither do they explain what constitutes an A patient

[23] The control group recieive printed educational material

[24] The control condition is not specified

Appendix 5. Bibliometric analysis

Review	Included study designs	Included interventions	Excluded interventions	Outcomes	Cut-off date of search	Included studies
Akbari 2008	RCT, CCT, CBA, ITS	Interventions (not necessarily financial) to change outpatient referral rates or improve outpatient referral appropriateness (studies had to specify that influencing referral was primary objective)	Interventions to change or improve referrals for open access radiological or laboratory diagnostic investigations (eg radiology)	Objectively measured provider performance (e.g. referral rates or appropriateness of referral), or health outcomes	October 2007	n=4 (Coulter 1993; Davidson 1992; Kammerling 1996; Linnala 2001)
Sturm 2007	RCT, CCT, CBA, ITS, RM	Policies that intend to affect prescribing by means of financial incentives for prescribers. Included are management of drug-budgets by prescribers, indicative prescribing schemes, and other financial policies for prescribers such as pay-for-performance, if they are specifically targeted at prescribing or drug utilisation	Interventions at the level of a single facility; educational interventions; remuneration of physicians; restriction of reimbursement for patients	Drug use (prescribed, dispensed or used); health care utilisation; health outcome; costs (drug, health care, administration)	2003 or 2005 (depending on the database)	n=13 (Baines 1997; Bradlow 1993; Burr 1992; Corney 1997; Guether 1995 ; Harris 1996 ; Kammerling 1996 ; Rafferty 1997 ; Schoffski 1997; Walley 2000;Whynes 1997; Wilson 1995; Wilson 1999)

(Continued)

Petersen 2006	RCT, CBA, cross sectional studies	Explicit financial reward for improving health care quality, targeted at the level of physician, provider group, or at the level of the payment system, such as performance based contracting	Studies with no concurrent comparison group; no baseline analysis of groups on quality measure	Quantitative measure of healthcare quality (access to care, structure of care, process of care, outcome of care, experience of care)	November 2005	n=12 {Beaulieu 2005; Christensen 2000 ; Clark 1995 ; Fairbrother 1999 ; Fairbrother 2001; Grady 1997 ; Hillman 1998; Hillman 1999 ; Kouides 1998 ; Rosenthal 2005; Roski, 2003; Shen 2003}
Gosden 2001	RCT, CBA, ITS	Salary vs capitation vs fee-for-service vs target payment (physician level payment only)	Changes within an existing system; patient-level randomisation; CBAs without contemporaneous data collection in different arms	Objective measurement of patient outcomes, health services utilisation, health care costs, equity of care and PCP satisfaction with working environment	October 1997	n=6 (Davidson 1992; Hickson 1987; Hutchison 1996; Kouides 1998; Krasnik 1990; Ritchie 1992)

Four systematic reviews (Akbari 2008; Gosden 2001; Petersen 2006; Sturm 2007) met the inclusion criteria. A total of 50 studies were included in these four reviews, but this report investigates only those 32 studies which had a financial incentive as intervention (other studies used educational or organisational interventions). Most of the included studies were unique to the review in which they appeared; only three were included in more than one review. This bibliometric analysis investigates the reasons for the differences in included studies across these four systematic reviews of financial incentives.

Characteristics of reviews included in the overview

Although the included systematic reviews in our overview all fitted our inclusion criteria in terms of financial incentive interventions, they differed from each other in their aims and inclusion criteria. Akbari 2008 was a review of any type of intervention (including but not limited to financial incentives) designed to affect the frequency or appropriateness of referrals, with objectively measured provider performance or health outcomes. Sturm 2007 investigated financial incentive-related policies directed at prescribing practices, and looked at healthcare utilisation, health outcomes and healthcare costs as well as drug use. Petersen 2006 was a review of "explicit financial incentives for quality" targeted at the physician or the provider group, with quantitative quality-related outcomes. Gosden 2001 looked at financial incentives affecting the personal income of a primary care physician Gosden 2001 and was a summary of two Cochrane reviews (Giuffrida 1999; Gosden 2000). The most recent review (Akbari 2008) incorporated a search to October 2007, and the oldest review reported a search cut-off date of October 1997. All reviews included RCTs, CBAs and ITS; one review (Petersen 2006) also included cross sectional studies, and one (Sturm 2007) repeated methods studies.

Characteristics of financial intervention studies included in the reviews

Of the 33 included studies of financial interventions, 10 were RCTs (Christensen 2000; Davidson 1992; Fairbrother 1999; Fairbrother 2001; Grady 1997; Hickson 1987; Hillman 1998; Hillman 1999; Kouides 1998; Roski 2003), 15 were CBAs (Baines 1997; Beaulieu 2005; Bradlow 1993; Burr 1992; Clark 1995; Corney 1997; Coulter 1993; Hutchison 1996; Kammerling 1996; Krasnik 1990; Linnala 2001; Rosenthal 2005; Shen 2003; Whynes 1997; Wilson 1999), four ITS (Guether 1997; Ritchie 1992; Schoffski 1997; Walley 2000), and three were characterised as CBA/CITS (Harris 1996; Rafferty 1997; Wilson 1995). In terms of outcomes[1], 11 studies recorded drug use and/or costs (Baines 1997; Bradlow 1993; Burr 1992; Guether 1997; Harris 1996; Krasnik 1990; Rafferty 1997; Walley 2000; Whynes 1997; Wilson 1995; Wilson 1999), including prescriptions. Six studies recorded referrals (Coulter 1993; Grady 1997; Kammerling 1996; Krasnik 1990; Linnala 2001; Schoffski 1997), and ten measured outcomes relating to other GP activity (Beaulieu 2005; Christensen 2000; Fairbrother 1999; Fairbrother 2001; Hillman 1998; Hillman 1999; Kouides 1998; Krasnik 1990; Ritchie 1992; Rosenthal 2005) (such as completion of immunisations or compliance

with guidelines). Two studies measured patient-related outcomes such as patient satisfaction or smoking (smoking status identification, provision of quitting advice and quitting rate) status (Hutchison 1996; Roski 2003). Four studies recorded patient visits (Davidson 1992; Hickson 1987; Schoffski 1997 Shen 2003) (to GP or hospital or emergency department), and one measured case management time (Clark 1995). One study (Clark 1995) looked at the setting of care.

Studies included in more than one review

Only three studies were included in more than one of the four reviews. Davidson 1992 reported an RCT looking at capitation versus fee-for-service and their impact on visits to the GP or to hospital (emergency or non-emergency, as outpatient or inpatient). This study was included in Akbari 2008 and Gosden 2001.

Akbari 2008 and Sturm 2007 both included Kammerling 1996, a report of a CBA investigating the effect of fundholding on referrals.

Kouides 1998 reported an RCT investigating the impact of target payments on vaccination rates, and was included in Gosden 2001 and Petersen 2006.

Reasons for inclusion/exclusion of studies in reviews

Why are so many of these 33 studies included in only one of our four systematic reviews of financial incentives? Only three of these 33 studies are explicitly listed as excluded[1] in any of the reviews, and so we have necessarily made assumptions about authors' reasons for not including other studies.

Studies explicitly excluded from reviews (studies mentioned by review authors and listed among exluded studies)

Krasnik 1990 reported a CBA investigating the effect of changing GPs' capitation-based remuneration system to a mixed capitation/fee-for-service model. It was included in Gosden 2001, but excluded from Akbari 2008 and Petersen 2006 because there was only one intervention and one control group.

Ritchie 1992 described an ITS looking at the effect of target payments and fee-for-service on immunisation rates. Gosden 2001 included this study, but Petersen 2006 excluded because of its ITS design.

Coulter 1993 reported results from a CBA describing the impact of fundholding on the number of outpatient referrals that would incur a charge. This study was included in Akbari 2008 but excluded from Sturm 2007 due to an inadequate control group.

Studies matching review exclusion criteria

Gosden 2001 specifically excluded studies of fundholding in the UK and organisational level payment systems in the US, "…since the payments made to PCPs do not affect PCP personal income directly". This accounts for thirteen non-included studies (Baines 1997; Bradlow 1993; Burr 1992; Clark 1995; Corney 1997; Coulter 1993; Guether 1997; Harris 1996; Kammerling 1996; Rafferty 1997; Schoffski 1997; Whynes 1997; Wilson 1995) in this review.

Petersen 2006 states that "Because the evidence regarding the relationship between the financial incentives embedded in fee-for-service and capitation arrangements and the quality of health care has been thoroughly reviewed in previous work (Dudley 1998), we focused our review on literature that addresses explicit financial rewards for improving health care quality".Three studies (Davidson 1992; Hickson 1987; Hutchison 1996) not appearing in Petersen 2006 but included in the other reviews compare fee-for-service and capitation, although an alternative explanation for the non-inclusion of one of these (Hutchison 1996) is that it does not appear in the databases searched for the Petersen review.

Akbari 2008 the review of referral outcomes, explicitly excluded "interventions to change or improve referrals for open access radiological or laboratory diagnostic investigations (eg radiology)", and consequently did not include a study investigating referrals for mammography (Grady 1997).

None of these studies were explicitly listed as excluded studies.

Publication date of studies

Thirteen of the studies (Beaulieu 2005; Christensen 2000; Fairbrother 1999; Fairbrother 2001; Grady 1997; Hillman 1998; Hillman 1999; Linnala 2001; Rosenthal 2005; Roski 2003; Shen 2003; Walley 2000; Wilson 1999) were published too recently to be included in Gosden 2001. The publication date (2005) of the most recent studies (Beaulieu 2005; Rosenthal 2005) included in the latest review (Akbari 2008) was also the cut-off date of the searches of the other two reviews. Since these searches ran late in the year, it's possible that these studies were located for assessment, and alternative reasons are given for their non-inclusion below.

Financial incentives interventions not included (i.e. studies not mentioned by review authors)

Akbari 2008 specified that the intervention must aim to change or optimise referral rates; 16 studies did not satisfy this criterion (Beaulieu 2005; Christensen 2000; Clark 1995; Fairbrother 1999; Fairbrother 2001; Guether 1997; Hickson 1987; Hillman 1998; Hillman 1999; Hutchison 1996; Kouides 1998; Ritchie 1992;Rosenthal 2005 Roski 2003; Shen 2003; Schoffski 1997). Similarly, Sturm 2007 specified that interventions

should be policies that aim to affect prescribing by means of financial incentives, and consequently did not include 18 studies (Beaulieu 2005; Christensen 2000; Clark 1995; Davidson 1992; Fairbrother 1999; Fairbrother 2001; Grady 1997; Hickson 1987; Hillman 1998; Hillman 1999; Hutchison 1996; Kouides 1998; Krasnik 1990; Linnala 2001; Ritchie 1992; Rosenthal 2005 ; Roski 2003; Shen 2003).

Ineligible outcomes and settings

Akbari 2008 defined required outcomes as "objectively measured provider performance (e.g. referral rates or appropriateness of referral), or health outcomes."

Ten studies (Baines 1997; Bradlow 1993; Burr 1992; Corney 1997; Harris 1996; Rafferty 1997; Walley 2000; Whynes 1997; Wilson 1995; Wilson 1999) reported only drug use or costs as outcome measures, and so were not eligible for inclusion in this review. Petersen 2006 specified quality of care as outcome, defining domains of quality as access to care, structure of care, process of care, outcomes of care, and patient experience of care. Fourteen studies (Baines 1997; Bradlow 1993; Burr 1992; Corney 1997; Coulter 1993; Guether 1997; Harris 1996; Kammerling 1996; Rafferty 1997; Schoffski 1997; Walley 2000; Whynes 1997; Wilson 1995; Wilson 1999) of the impact of fundholding or drug budgets were not included in Petersen 2006. Since the review does not explicitly exclude these studies, it's likely that this is because the drug use and cost outcomes reported in these studies do not match the specified quality outcome, although three of them (Coulter 1993; Kammerling 1996; Schoffski 1997) reported measures of health care utilisation (such as referrals or hospitalisation rate).

Gosden 2001 specified that the interventions should be directed at primary care physicians, and consequently did not include one study (Clark 1995) which took place in the community and in mental health centres.

Appendix 6. Results grouped by type of financial incentive

Intervention/ Review	Consultation /Visit rates	Processes of care	Referrals/Admissions	Compliance with guidelines	Prescribing costs
Payment for working for a specified time period					
Gosden 2001	3/9 outcomes favoured the intervention Continuity of care (Hickson 1987)			0/2 outcomes favoured the intervention Compliance with AAP guidelines (P<0.01) (Hickson 1987)	
VOTE COUNTING SUMMARY:					
	3/9 outcomes from 1 study reported in 1 review favoured the intervention. Statistical significance was reported for 9/9 outcomes, and reported as significant for 6/9 outcomes, two of which favoured the intervention.			0/2 outcomes from 1 study reported in 1 review favoured the intervention. Statistical significance was reported for 2/2 outcomes. 2/2 outcomes were statistically significant, none of which favoured the intervention.	
Payment for each service/episode/visit					

(Continued)

Gosden 2001	2/3 outcomes favoured the intervention PCP and Non-PCP visits, and Health / emergency department visits (Davidson 1992) Statistical significance not reported.		0/1 outcomes favoured the intervention Admission to hospital (Davidson 1992) Statistical significance not reported.
Akbari 2008	2/2 outcomes favoured the intervention PCP and Non-PCP visits (Davidson 1992) Statistical significance not reported.		
Petersen 2006		4/6 outcomes favoured the intervention (1/1) Cognitive services (Christensen 2000),(P<0.001) (2/3) Treatment time (Clark 1995) (P<0.05-0.001) Immunisation status (0/1)Immunisation status (Fairbrother1999), NS (1/1)Immunisation status (Fairbrother 2001), (P<0.05)	
VOTE COUNTING SUMMARY:			
	3/3 outcomes from 1 study reported in 2 reviews favoured the intervention. Statistical significance was reported for 0/3 outcomes.	4/6 outcomes from 4 studies reported in 1 review favoured the intervention. Statistical significance was reported for 6/6 outcomes and reported as significant for 4/6 outcomes, 4 of	0/1 outcomes from 1 study reported in 1 review favoured the intervention. Statistical significance was reported for 0/1 outcomes.

(Continued)

	which favoured the intervention.		
Payment for providing care for a patient or a specific population			
Akbari 2008		3/5 outcomes favoured the intervention (1/2) NHS outpatient referrals and private outpatient referrals (Coulter 1993), Statistical significance not reported (1/2) Referrals to hospital outpatient clinics and to private sector, P<0.001) (Linnala 2001) (1/1) Orthopedic referrals (Kammerling 1995), Statistical significance not reported.	
Sturm 2007	17/30 outcomes from 8 studies favoured the intervention Items per patient: (0/1) Bradlow 1993 (0/1) Burr 1992 (2/5) Harris 1996 (3/3) Rafferty 1997 (0/3) Wilson 1995 (1/1 adj. rel. change) Whynes 1997 (0/1 adj. rel. change) Wilson 1999 Generic percentage: (1/1) Bradlow 1993 (3/3) Rafferty 1997 (1/1) Whynes 1997 (3/3) Wilson 1995 (2/2 adj.rel. change) Baines 1997 (1/5 adj.rel. change) Wilson 1999	1/1 outcomes favoured the intervention Referrals to NHS outpatient care (Kammerling 1995), Statistical significance not reported	28/34 outcomes from 10 studies favoured the intervention Cost per item and per patient (2/2) Bradlow 1993 (5/5) Rafferty 1997 (3/6) Wilson 1995 (5/5 adj. rel. difference) Wilson 1999 Cost per item and total prescribing cost (1/2) Walley 2000

(Continued)

			Cost per patient and total prescribing cost (8/9) Harris 1995 Cost per patient (1/1) Burr 1992 (0/1) Corney 1997 (1/1) Whynes 1997 (2/2 adj. rel. change) Baines 1997
VOTE COUNTING SUMMARY:			
	17/30 outcomes from 8 studies Statistical significance was reported for 0/30 outcomes.	3/5 outcomes from 3 studies reported in 2 reviews favoured the intervention Statistical significance was reported for 2/5 outcomes, and was reported significant for 1/5, which favoured the intervention.	28/34 outcomes from 10 studies reported in 1 review favoured the intervention. Statistical significance was reported for 0/34 outcomes.
Payment for providing a pre-specified level or providing change in activity or quality			
Gosden 2001	1/1 outcomes favoured the intervention Change in Immunisation rate, S, P<0.03 (Kouides 1998)		

(Continued)

Petersen 2006		16/16 outcomes favoured the intervention (9/9) Diabetes care (Beaulieu 1995), (P<0.05-0.001) (for eight of nine outcomes, the ninth NS) (1/1) Immunisation status (Fairbrother 1999), (P<0.01) (1/1) Immunisation status (Fairbrother 2001), (P<0.05) (1/1) Change in Immunisation rate, P<0.03, (Kouides 1998) (1/1) Cervical screening rate,(P<0.02) (no data reported for two non-significant outcomes) (Rosenthal 2005) (3/3) Smoking identification (P=0.009), quitting advice (NS) and quitting rate (NS), (Roski 2003)	1/2 outcomes favoured the intervention (0/1) Referrals to mammography (Grady 1997), NS (1/1) Health service to special populations, (P<0.01) (Shen 2003)	0/2 outcomes favoured the intervention (0/1) Compliance with cancer screening guidelines, NS (Hillman 1998) (0/1) Compliance with immunisation and well child visit guidelines, NS (Hillman 1999)
VOTE COUNTING SUMMARY:				
		16/16 outcomes from 6 studies reported in 2 reviews favoured the intervention. Statistical significance was reported for 16/16 outcomes across two reviews, and was reported as significant for 13/16 outcomes,13 of which favoured the intervention.	1/2 outcomes from 2 studies reported in 1 review favoured the intervention. Statistical significance was reported for 2/2 outcomes across one review, and was reported as significant for 1/2, which favoured the intervention.	0/2 outcomes from 2 studies reported in 1 review favoured the intervention. Statistical significance was reported for 2/2 outcomes across one review, of which none were reported as significant and none favoured the intervention.
Mixed or other system				
Gosden 2001	4/5 outcomes favoured the intervention	2/3 outcomes favoured the intervention	4/5 outcomes favoured the intervention (1/1) Admission to hospital Statistical significance not reported.	5/13 outcomes favoured the intervention

(Continued)

	(3/3) PCP and non-PCP visits, and Clinic/ Emergency department visits, Statistical significance not reported (Davidson 1992) (1/2) Face-to-face consultation and telephone consultations Stattistical significance not reported. (Krasnik 1990)	Diagnostic and Curative services, Renewal of prescriptions Statistical significance not reported. (Krasnik 1990)	(Davidson 1992) (1/2) Admission to hospital (NS, P=0.312), and hospital length of stay (NS, P=0.774) (Hutchison 1996) (2/2) Referrals to specialists and to hospitals Statistical significance not reported. (Krasnik 1990)	(0/8) Compliance with CHAP guidelines Statistical significance not reported. (Davidson 1992) (5/5) Immunisation guidelines Statistical significance not reported. (Ritchie 1992)
Akbari 2008	2/2 outcomes favoured the intervention PCP and Non-PCP visits: Statistical significance not reported. (Davidson 1992)			
Sturm 2007		2/2 outcomes favoured the intervention Items per patient Statistical significance not reported. (Guether 1997; Walley 2000)	3/3 outcomes favoured the intervention (1/1) Referrals to outpatient specialists Statistical significance not reported. (Guether 1997) (2/2) Referrals to outpatient clinics and to hospital clinics Statistical significance not reported (Schoffski 1997)	
VOTE COUNTING SUMMARY:				
	4/5 outcomes from 2 studies reported in two reviews favoured the intervention. Statistical significance was reported for 0/5 outcomes.	4/5 outcomes from 3 studies reported in two reviews favoured the intervention. Statistical significance was reported for 0/5 outcomes.	7/8 outcomes from 5 studies reported in two reviews favoured the intervention. Statistical significance was reported for 2/8 outcomes across two reviews, and reported as significant for none of the outcomes. 7/8 outcomes favoured the intervention.	5/13 outcomes from 2 studies reported in 1 review favoured the intervention Statistical significance was re-

(Continued)

	ported for 0/13 outcomes.

CONTRIBUTIONS OF AUTHORS

FB developed and ran the search strategy. MPE and SS sifted the titles and abstracts and assessed the eligibility of studies. GF,EP, MPE and SS extracted data and assessed the scientific quality of the reviews. FB performed the bibliometric analysis. AS was involved in determining the definitions of financial incentives and writing the protocol. GF led the writing of the review, and all other authors commented on and contributed to sequential drafts and approved the final version.

DECLARATIONS OF INTEREST

Anthony Scott is an author of a review on financial incentives (Scott 1995).

SOURCES OF SUPPORT

Internal sources

- Newcastle University, UK.
- Department of Public Health, University of Oxford, UK.

External sources

- NIHR EPOC Program Grant, UK.
- Australian Research Council, Future Fellowship (Anthony Scott), Australia.

INDEX TERMS

Medical Subject Headings (MeSH)

*Motivation; Capitation Fee; Delivery of Health Care [*economics]; Fee-for-Service Plans; Professional Practice [*economics]; Review Literature as Topic; Salaries and Fringe Benefits [*economics]; Treatment Outcome

MeSH check words

Humans