

The Impact of Vouchers on the Use and Quality of Health Goods and Services in Developing Countries

A systematic review



by Carinne Meyer
Nicole Bellows
Martha Campbell
Malcolm Potts

May 2011

The EPPI-Centre reference number for this report is 1905.

This report should be cited as: Meyer C, Bellows N, Campbell M, Potts M (2011) The Impact of Vouchers on the Use and Quality of Health Goods and Services in Developing Countries: A systematic review. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
ISBN: 978-1-907345-10-4

© Copyright

Authors of the systematic reviews on the EPPI-Centre website (<http://eppi.ioe.ac.uk/>) hold the copyright for the text of their reviews. The EPPI-Centre owns the copyright for all material on the website it has developed, including the contents of the databases, manuals, and keywording and data extraction systems. The centre and authors give permission for users of the site to display and print the contents of the site for their own non-commercial use, providing that the materials are not modified, copyright and other proprietary notices contained in the materials are retained, and the source of the material is cited clearly following the citation details provided. Otherwise users are not permitted to duplicate, reproduce, re-publish, distribute, or store material from this website without express written permission.

Table of Contents

List of Abbreviations	iii
Abstract	1
Executive Summary.....	2
1. Background	6
1.1 Aims and rationale for current review	6
1.2 Definitional and conceptual issues.....	6
1.3 Policy and research background.....	11
2. Methods.....	14
2.1 User involvement.....	14
2.2 Identifying and describing studies	14
2.3 Methods for synthesis.....	18
2.4 Deriving conclusions and implications.....	21
3. Search Results.....	23
3.1 Studies included from searching and screening.....	23
3.2 Details of included studies	25
4. Synthesis Results.....	30
4.1 Further details of studies and outcomes included in the synthesis	30
4.2 Synthesis of evidence findings.....	39
4.3 Sensitivity analyses.....	42
4.4 Summary of synthesis findings.....	43
5. Discussion	44
5.1 Implications of findings.....	44
5.2 Strengths of the research approach	45
5.3 Limitations of the research approach	45
6. Conclusions and Recommendations	47
6.1 Conclusions from the synthesis.....	47
6.2 Areas for further analysis and future research.....	47
References	49
Appendices	55
Appendix 1.1: Authorship	55
Appendix 2.1: Inclusion and exclusion criteria	56
Appendix 2.2: Search strategy for electronic databases	58
Appendix 2.3: Journals hand-searched	61
Appendix 2.4: Supplemental search terms.....	62
Appendix 2.5: Key experts contacted to review obtained information	65
Appendix 2.6 Data Extraction Form	66
Appendix 2.7: Coding tool, Confidence in Findings Assessment (CFA)	67
Appendix 2.8: Decision tree for making conclusions regarding outcome categories.....	70
Appendix 3.1: Voucher programmes with quantitative evaluation data included in the review.....	71

List of Abbreviations

ACCP	Alliance for Cervical Cancer Prevention
ANC	Antenatal Care
BPL	Below the Poverty Line
CCT	Conditional Cash Transfer
CFA	Confidence in Findings Assessment
DFID	Department for International Development, UK
DHS	Demographic and Health Surveys
GPOBA	Global Partnership on Output-Based Aid
HDI	Human Development Index
HEF	Health Equity Funds
HH	Household
HIV	Human Immunodeficiency Virus
3ie	International Initiative for Impact Evaluation
ICAS	Instituto CentroAmerica de la Salud
ITN	Insecticide-Treated Nets
IUD	Intrauterine Device
KfW	German Development Bank
LLIN	Long Lasting Insecticide-treated Nets
MSH	Management Sciences for Health
NBCCEDP	National Breast and Cervical Cancer Early Detection Programme
PNC	Postnatal Care
PPV	Positive Predictive Value
RBF Health	Results-Based Financing for Health
RCT	Randomised Controlled Trial
RH Vouchers	Reproductive Health Vouchers
RTI	Research Triangle International
SES	Socio-Economic Status
SRH	Sexual and Reproductive Health
STI	Sexually Transmitted Infection
TPHA	Treponemal pallidum particle agglutination
USAID	United States Agency for International Development
VDRL	Venereal Disease Research Laboratory

Abstract

Background: One approach to delivering health assistance to developing countries is the use of health voucher programmes, where vouchers are distributed to a targeted population for free or subsidised health goods/services. Theoretically, vouchers are expected to successfully target specific populations, increase utilisation, improve quality, enhance efficiency, and ultimately improve the health of populations.

Objectives: The primary objective of this systematic review is to assess whether voucher programmes thus far have been successful in achieving these desired outcomes.

Methods: Using explicit inclusion/exclusion criteria, a search of bibliographic databases, key journals, and organisational websites were conducted in September - October 2010. Other search strategies used include bibliographic back-referencing, supplemental keyword searches using specific programme information, and contacting key experts in the field. A narrative synthesis approach was taken to qualitatively summarise the identified quantitative outcome variables in five categories (targeting, utilisation, efficiency, quality, and health impact). Using the direction of effect of outcome variables and the confidence in the study findings, the findings for each category of outcomes were aggregated and assigned to one of five pre-established conclusion categories: (1) insufficient evidence; (2) evidence of no effect; (3) conflicting evidence; (4) modest evidence of effect; or (5) robust evidence of effect. Sub-group and sensitivity analyses were also performed. A quantitative meta-analysis was not conducted due to the heterogeneous natures of the outcome variables reviewed.

Results: A total of 24 studies evaluating 16 different health voucher programmes were identified in this review. The findings from 64 outcome variables informed five main conclusions: (1) there is modest evidence that voucher programmes effectively target voucher for health goods/services to specific populations (based on four programmes); (2) there is insufficient evidence to determine whether voucher programmes deliver health goods/services more efficiently than competing health financing strategies (based on one programme); (3) there is robust evidence that voucher programmes increase utilisation of health goods/services (based on 13 programmes); (4) there is modest evidence that voucher programmes improve the quality of health services (based on three programmes); and (5) the evidence indicates that voucher programmes do not have an impact on the health of populations (based on six programmes); however, this last conclusion was found to be unstable in a sensitivity analysis.

Conclusions: The evidence indicates that health voucher programmes have been successful in increasing utilisation of health goods/services, targeting specific populations, and improving the quality of services. While these results are encouraging, the subsequent link that voucher programmes improve the health of the population is not evident in the data analysed in this review. The methodology used in this analysis allows policy-makers to synthesise evidence from heterogeneous studies and therefore include more data than could be used in a standard meta-analysis. However, vouchers are still relatively new and the number of published studies evaluating vouchers is a limitation. Future reviews using this methodology can compare health voucher programmes to competing financing techniques and incorporate new evidence on voucher programmes for evaluations currently underway; however, the synthesis tools used in this review should be validated.

Executive Summary

Background

Donors of health assistance to developing countries have emphasised the need for efficient and transparent spending of aid funds. One approach is the use of health voucher programmes, where vouchers are distributed to a targeted population for free or subsidised health goods/services. The theoretical context for voucher programmes can be found in the basic economic theories of supply and demand, where voucher programmes aim to inject market mechanisms into the delivery of health aid in order to improve efficiency and improve health.

One advantage of voucher programmes is the potential to grant purchasing power to low-income individuals who might otherwise be ignored in the market due to their lack of funds or knowledge of goods and services. By removing financial and knowledge barriers to a targeted population, utilisation of the specific health goods/services is expected to increase. Additionally, most voucher programmes also aim to improve the supply of goods/services available by reimbursing providers based on the volume of goods/services delivered.

Another important aspect of most voucher programmes is the process of contracting with providers who meet minimum standards for quality, of providing some form of provider training and of competition between providers with the expectations that voucher programmes can increase quality for voucher goods/services. Providers are incentivised to meet the obligations of their contract in the most efficient manner in order to capture more of the set payment established in the voucher reimbursement. Furthermore, competitive contracting of providers can reduce prices of goods and services. As such, voucher programmes are also expected to be able to improve the efficiency of health goods/services.

By targeting specific populations, increasing utilisation, and enhancing quality and efficiency, it is expected that voucher programmes can improve the health of populations.

Objectives

The overall objective of this systematic review is to assess whether voucher programmes thus far have been successful in achieving their objectives and should therefore be considered as a mechanism for further health aid. Five hypotheses are tested in this review:

1. Voucher programmes effectively target specific populations.
2. Voucher programmes increase utilisation of specified health goods/services.
3. Voucher programmes allow for more efficient distribution of health goods/services compared to other forms of aid distribution.
4. Voucher programmes improve the quality of health goods/services.
5. Voucher programmes result in the improved health of a population.

Methods

This systematic review sought studies on voucher programmes that provided health goods/services to populations in developing countries. Included studies were limited to those that: (1) evaluated some aspect of a health voucher programmes; (2) contained some quantitative evidence with an observable contrast such as time or control group; and (3) examined at least one outcome variable that addressed one of the five outcome categories of interest (targeting, utilisation, quality, efficiency, and health impact). Study designs such as randomised controlled trials

(RCTs), non-randomised trials (time series), case-control, cohort, before and after with and without controls, and cross-sectional analyses with a comparison group were eligible for inclusion.

The search for studies was conducted in two phases. In the first phase bibliographic databases were searched using specified keywords, as were publishers' pages of key journals. The "grey literature" was also examined by hand-searching key organisational and network websites. In the second phase of the search, three additional activities took place. First, the reference lists of all the reviewed full-text studies were examined to identify additional, secondary resources. Next, a supplemental keyword search in PubMed and google.com was conducted, where search terms were generated from voucher programmes identified in phase one, such as specific programme names and locations. Third, key contacts were consulted to make further suggestions of publications for consideration.

Two reviewers applied specified inclusion/exclusion criteria to full-text studies and a third reviewer arbitrated discrepancies. The same strategy applied to characterising and synthesising the data extracted from the studies.

In synthesising the data, a narrative synthesis approach was taken, where qualitative conclusions were made with regards to the five hypotheses. Since evaluation studies of health voucher programmes may contain several outcomes, the unit of analysis used was the outcome variable, nested within individual studies. Each study was reviewed for relevant evaluation outcomes that fit within the five outcome categories, and was data-extracted for synthesis. Two reviewers independently assessed the level of confidence in the study findings as either low confidence, medium confidence or high confidence.

For each outcome variable used for synthesis, information was recorded on the overall direction of effect as either: no effect of statistical significance; positive effect indicating good voucher programme performance; or negative effect of poor voucher performance. Using the results from confidence in the study findings and the direction of effect for each outcome variable, the findings for each category of outcomes were aggregated and assigned to one of five pre-established conclusion categories: (1) insufficient evidence; (2) evidence of no effect; (3) conflicting evidence; (4) modest evidence of effect; or (5) robust evidence of effect. In order to assess the stability of the conclusions, a sensitivity analysis was performed.

The outcome variables were reviewed to determine whether a statistical meta-analysis could provide further insight into the effectiveness of voucher programmes. Due to the heterogeneity of outcome variables and study designs, however, it was determined that a meta-analysis would not be particularly useful for this report.

Details of the included studies

From an initial 1,031 abstracts and 119 grey literature studies, a total of 24 quantitative evaluation studies were included in the analysis for the purposes of synthesising the data. The 24 studies evaluated 16 different health voucher programmes.

Two main types of voucher programmes had a substantial amount of evaluation findings. Six programmes were for insecticide-treated bed net (ITN) distribution programmes and nine voucher programmes addressed some aspect of reproductive health (e.g. maternity services, family planning, and treatment for sexually transmitted infections (STIs)). One general health services programme was also evaluated. Of the 16 studies, eight evaluated voucher programmes were located in

Africa, five were in Asia, and three were in Latin America (all located in Nicaragua).

The majority of the 24 evaluation studies included in this review are observational studies with a pre/post design, a cross-sectional intervention/control design or a combination of the two. One case-control study and two economic modelling studies were included, as well as one clinical record review. Finally, one simulated patient before-during-after evaluation was included.

Synthesis results

For the *targeting* outcome, evidence from four programmes, six studies, and nine total outcome variables were synthesised and it was concluded that there was modest evidence indicating that voucher programmes were able to effectively target specific populations for health goods/services.

There was only one *efficiency* outcome variable that examined whether a voucher programme was able to deliver health goods/services more efficiently than a competing financing strategy. While the finding from this study was positive, there is insufficient evidence to make a conclusion on this outcome category.

Almost half of the total outcome variables were classified as *utilisation* outcomes. The evidence from 13 programmes, 16 studies and 30 outcome variables found robust evidence that voucher programmes were able to increase utilisation of health goods/services.

For the *quality* outcome category, the evidence from three voucher programmes, six studies, and 13 outcome variables found modest evidence that voucher programmes were able to impact some dimension of quality for health goods/services. A conclusion of modest positive evidence was made because there were only three voucher programmes evaluated instead of the required four for robust evidence.

Six voucher programmes with six studies and eleven outcome variables were used in synthesising the evidence on *health impact*. The synthesis found that voucher programmes did not have a significant effect on health outcomes; however, this conclusion was found to be unstable in a sensitivity test.

Conclusions and recommendations

The strongest finding from this review is that the evidence indicates that health voucher programmes have been successful in increasing utilisation of health goods/services. There is also modest evidence that voucher programmes can effectively target specific populations and can improve the quality of services. While these results are encouraging, the subsequent link that voucher programmes improve the health of the population is not evident in the data analysed in this review.

The narrative synthesis approach and methodology used in this analysis allows for the synthesising of evidence from heterogeneous studies and the inclusion of cross-sectional studies that would likely be excluded from a more standard meta-analysis. As such, this approach allows inclusion of more information; however, the synthesis tools developed for this review require validation. Additionally, further limitations exist, such as the lack of statistical compilation and the reliance on reviewer judgement for grading the evidence.

Future research in this area should focus on four areas. First, a consistent systematic review methodology should be applied to several health financing strategies so that policy-makers can examine the relative effectiveness of different strategies. Second, programme managers of current and future voucher

programmes would benefit from a review of lessons learned when implementing voucher programmes. Third, there is a clear need for more evidence on the efficiency of voucher programmes and a different approach for analysing data on efficiency. Finally, an update to this review should take place in approximately three years' time to incorporate new evaluations currently underway.

1. Background

Outline of chapter

Section 1.1 introduces the aims and rationale for the current review on health voucher programmes in developing countries.

Section 1.2 details how voucher programmes work, the theoretical background on expected outcomes associated with voucher programmes, and the hypotheses to be evaluated.

Section 1.3 discusses the policy and research background for voucher programmes in health care.

Section 1.4 clarifies the primary objectives of the systematic review.

1.1 Aims and rationale for current review

Measures of infant mortality, maternal mortality, and disease-specific deaths such as those related to malaria and HIV vary widely between countries (Satcher 2000). Developing countries bear 93% of the world's disease burden and account for only 11% of the world's health spending (Satcher 2000). As a result of this gap between burden of disease and funding, the health sector in developing countries has been an important recipient of international aid. Over the last three decades, the world has seen an increased focus on global health partly due to the identification of health as a key determinant of economic growth and poverty reduction (Greene and Merrick 2005, Deaton 2003). As a result, official development assistance from bilateral and multilateral agencies towards health has increased from \$4.5 billion in 1996 to \$7.9 billion in 2004 (OECD 2006).

In addition to allocating increased funds to address health inequalities, donors have emphasised the need for efficient and transparent spending of aid funds (World Bank 2007). A variety of strategies exists for distributing health aid. One strategy that is growing in popularity is the use of voucher programmes, where vouchers are distributed to a targeted population for free or subsidised health goods/services. While there is much discourse in the literature on how voucher programmes work and why they are potentially important, the literature lacks a systematic assessment of the existing evidence on whether vouchers yield value for donors in the form of efficient spending of health aid. As such, the overall objective of this systematic review is to assess whether voucher programmes thus far have been successful in achieving their objectives and should therefore be considered as a mechanism for further health aid. Additionally, this review aims to identify conditions in which voucher programmes are more or less successful and to specify gaps in the literature that require further research.

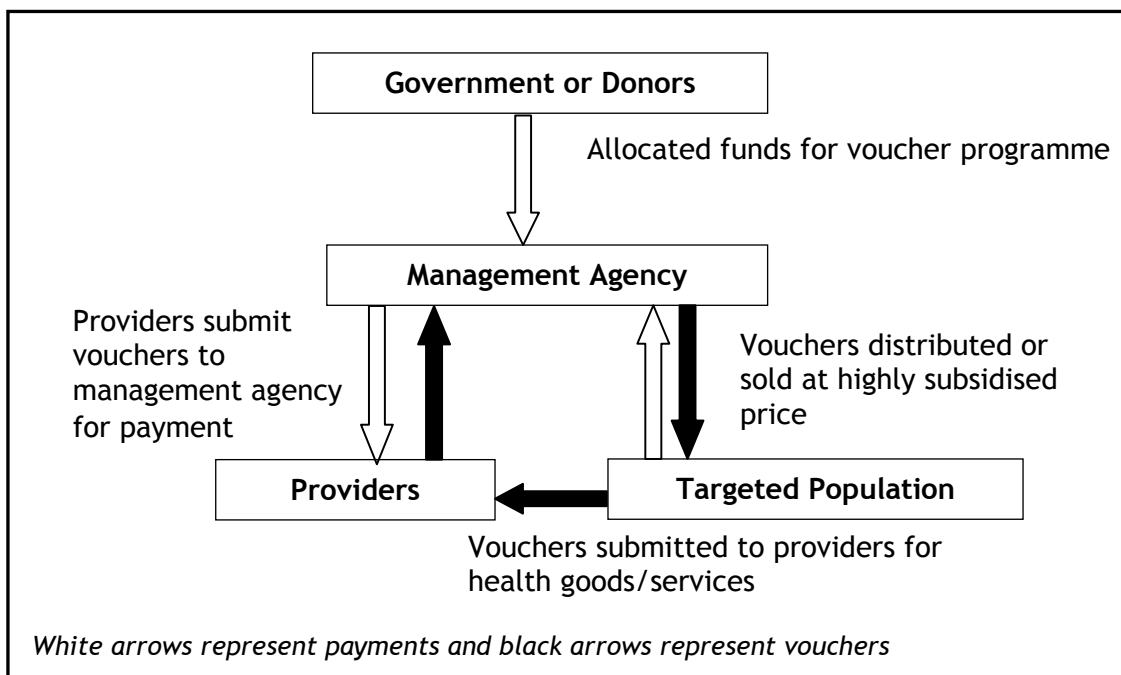
1.2 Definitional and conceptual issues

1.2.1 *How voucher programmes work*

In a voucher programme there are typically four major actors: (1) the government or donors who provide the funding; (2) a management agency that administers the programme; (3) providers who deliver the health goods/services; and (4) the voucher recipients who are in need of health goods/services. Vouchers are usually competitive with multiple providers; however, it is possible for them also to be non-competitive as well.

The management agency plays an important role in contracting providers to deliver health goods/services to voucher holders, distributing the vouchers to the targeted population, and overseeing the delivery of care by the providers. The targeted voucher recipients may be selected on income status, whether they were living in a geographic region, whether they have certain risk factors, or other relevant characteristics depending on the programme. Once vouchers are distributed, recipients bring the vouchers to participating providers. After the specified health goods/services are delivered by the provider, the provider submits the vouchers to the management agency for reimbursement. Figure 1.2.1 describes how monies and vouchers flow between the primary participants in voucher programmes.

Figure 1.2.1 Flow diagram of payments and vouchers



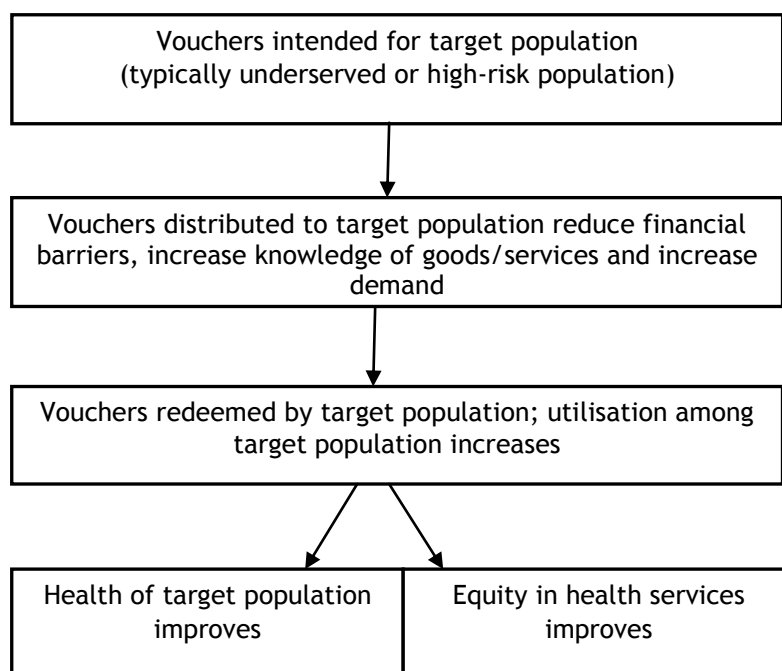
1.2.2 Theoretical context for vouchers

The theoretical context for voucher programmes can be found in the basic economic theories of supply and demand where voucher programmes aim to inject market mechanisms into the delivery of health aid in order to improve efficiency and improve health (Bhatia et al. 2006). The end goal of voucher programmes is to improve the health of the population.

One rationale for subsidising health care is the inequitable distribution of wealth and health (Sandiford et al. 2005). Low-income individuals may have the need for health goods/services; however, without financial resources or knowledge of these health goods/services they do not have the ability to access them, particularly in the private sector. Voucher programmes are a form of output-based aid, where aid monies are used to stimulate demand for health goods/services, contrasting with more traditional supply-side strategies, which often focus on providing the inputs for health aid such as construction of facilities or provision of supplies.

One advantage of voucher programmes is the potential to grant purchasing power to low-income individuals who might otherwise be ignored in the market due to their lack of funds or knowledge of health goods/services (Mumssen et al. 2010). By *targeting* the benefit towards low-income and/or high-risk individuals, voucher programmes are expected to increase demand among those most in need, increase utilisation of necessary health goods/services, improve health within the targeted population and improve equity of health distribution. Figure 1.2.2.1 depicts the causal expectations with regards to targeting.

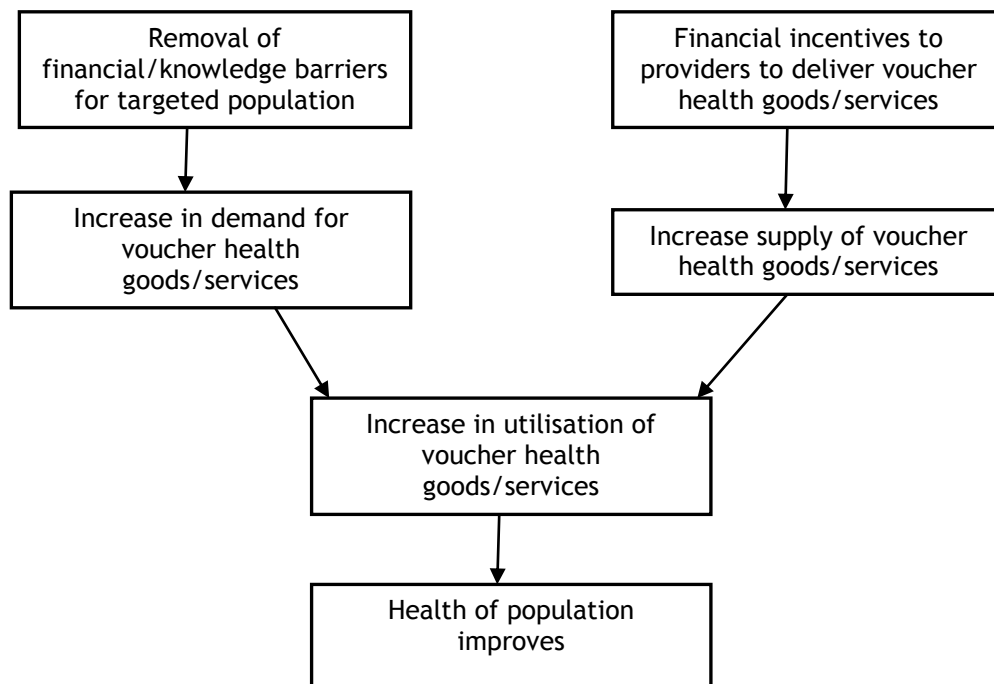
Figure 1.2.2.1: Causal expectations for targeting



As indicated in Figure 1.2.2.1, by removing financial or knowledge barriers in a targeted population, utilisation of the voucher health goods/services is expected to increase. Additionally, voucher programmes aim to improve the supply of goods/services available. The theoretical basis for supply enhancement in voucher programmes can be found in the principal-agent model of the economics literature where the principal delegates a task to an agent via an inducement embedded in a contract (Azam and Laffont 2003). Inducements (sanctions or incentives) are

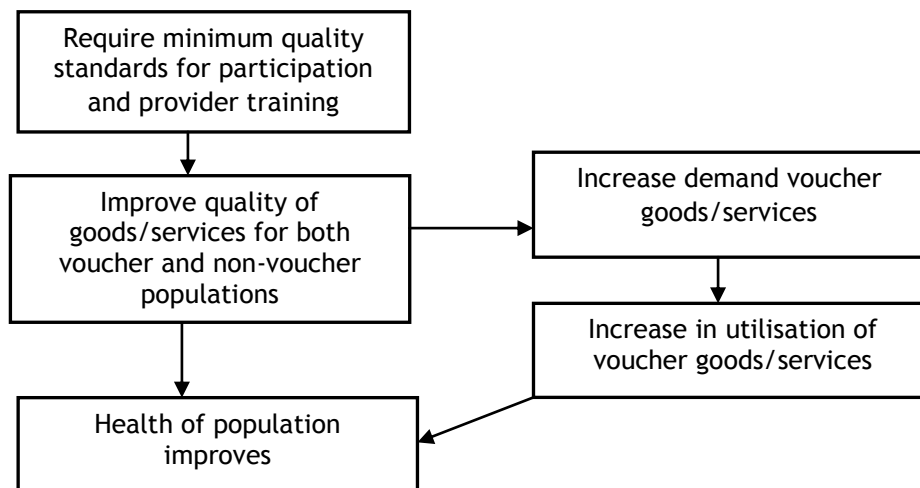
designed to ensure that the task is completed satisfactorily. In voucher programmes, the funder or management agency serves as the principal and the providers (either public or private) are the contracted agents. Providers are given a financial incentive to deliver health goods/services and this incentive coupled with increased purchasing power is expected to yield increased *utilisation* of voucher goods/services. Figure 1.2.2.2 indicates the causal expectations with regards to utilisation

Figure 1.2.2.2: Causal expectations for utilisation



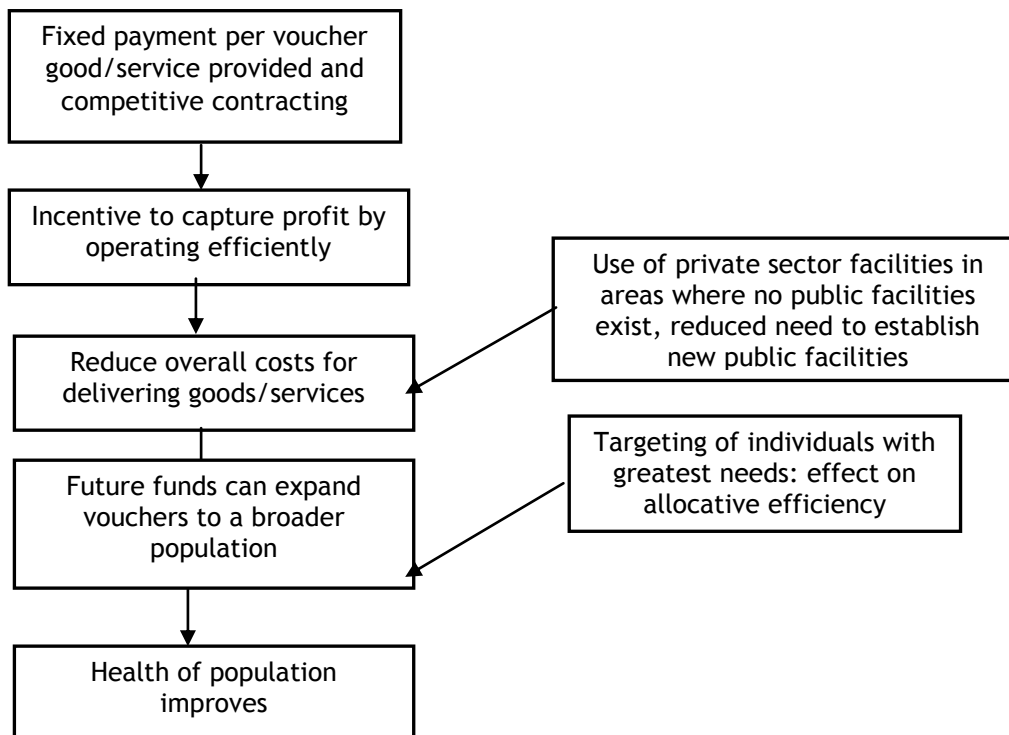
Another important aspect of most voucher programmes is the process of contracting with providers who meet minimum standards for quality and providing some form of provider training (Gorter et al. 2003). Some voucher programmes may also have additional financial incentives for quality outcomes, further incentivising providers to increase quality. By establishing quality standards and introducing some form of competition between providers for the market, providers are expected to improve the delivery of health goods/services for both voucher and non-voucher populations and thus improve the health of the population. Also, an increase in the quality of health goods/services is expected to increase the demand for voucher health goods/services. Figure 1.2.2.3 depicts the causal expectations for quality.

Figure 1.2.2.3: Causal expectations for quality



Providers not only have the incentive to provide health goods/services up to the quality standards established in the voucher programme, they also have the incentive to meet the obligations of their contract in the most efficient manner in order to capture more of the set payment established in the voucher programme (Gorter et al. 2003). Furthermore, competition between providers for the market may result in reduced prices of health goods/services. As such, voucher programmes are also expected to be able to improve the *efficiency* of health goods/services. If health goods/services can be provided more efficiently, then future funds can expand voucher services to a broader population and therefore improve the health of the population. Figure 1.2.2.4 indicates the causal expectations for efficiency. Furthermore, targeting of health goods/services to individuals for which the greatest health gain will be obtained has effects on the allocative efficiency (Gorter et al. 2003). In addition, vouchers make it possible to use existing private sector facilities in areas where there are few public facilities, and may reduce the need to establish new public facilities, which is more costly (Gorter et al. 2003).

Figure 1.2.2.4: Causal expectations for efficiency



1.2.3 Hypotheses examined in this review

For the purpose of this systematic review, the claims around voucher programmes that are of high interest to policy makers include: whether voucher programmes effectively target populations of interest; whether vouchers increase utilisation of health goods/services; whether the costs of delivering health goods/services in voucher programmes are more efficient than other forms of health aid distribution; whether vouchers result in improved quality of care; and whether voucher programmes improve the health of populations. These questions can be stated as hypotheses to be tested in voucher programme evaluations. The systematic review will examine the evaluation research to test the following hypotheses:

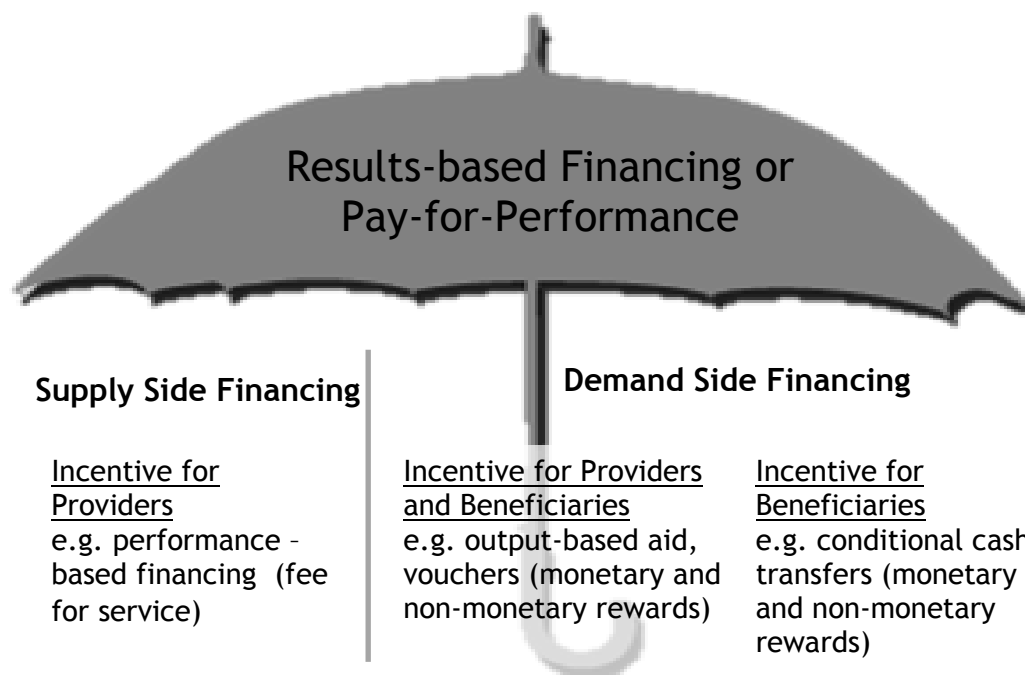
1. Voucher programmes effectively target specific populations.
2. Voucher programmes increase utilisation of specified health goods/services.
3. Voucher programmes allow for more efficient distribution of health goods/services compared to other forms of aid distribution.
4. Voucher programmes improve the quality of health goods/services.
5. Voucher programmes result in the improved health of a population.

1.3 Policy and research background

Vouchers exist within the context of a number of financing strategies aimed at rewarding both providers and beneficiaries for specific outcomes or outputs. The umbrella terms used to categorise these types of financing strategies is “results-based financing” and “pay-for-performance”. Under these terms there are three main types of schemes: those that incentivise provider outputs (supply side

financing), those that reward both provider and beneficiary outputs and those that focus rewards primarily on beneficiaries (demand side financing) (Musgrove 2010). Vouchers are considered to be a type of demand side output-based aid - rewarding both providers and beneficiaries.

Figure 1.3.1: Situating vouchers within results-based financing



To date, vouchers have been used for basic health services, maternal and child health services, and sexual and reproductive health services (STI treatment, family planning, prevention of gender-based violence and safe abortion and post abortion care). Health voucher programmes have been used throughout the developing world including countries in Asia, Africa, and Latin America.

Much of the literature on voucher programmes describes the potential benefits of delivering health aid via vouchers, as described in the previous section. There is not, however, a definitive consensus on whether voucher programmes achieve their goals or potential benefits. Within the peer review literature there is some evaluation information on individual voucher programmes. For example, voucher programmes for reproductive health services in Nicaragua have been evaluated in several papers examining and data on costs, utilisation, quality measures, and population health impact (Borghini et al. 2005, McKay et al. 2006, Meuwissen et al. 2006a-d). In Tanzania, numerous studies detail the results of a voucher programme to increase the use of ITNs (Hanson et al. 2009, Khatib et al. 2008, Kileen et al. 2007, Marchant et al. 2010, Mulligan et al. 2008). In general, these studies have found favourable results for the Nicaragua and Tanzanian programmes on the variables presented.

One systematic review examined private for-profit interventions for the poor and examined three voucher programmes: insecticide treated bed nets in Tanzania and Zambia and reproductive health services in Nicaragua (Patouillard et al. 2007). The review identified 52 impact evaluations on interventions targeting private for-

profit providers in low and middle-income countries including five studies on voucher programmes. The quality of the evidence of all included studies was classified as low. The authors concluded that while stronger evidence is needed to make any recommendations for policy, the evidence suggests that interventions involving private-for-profit providers are feasible, and the fact that many interventions have been successfully implemented in poor regions indicates that they may benefit the poor significantly.

A number of health voucher programmes have been implemented in more recent years, with evaluation data now available in both the peer-review and grey literature. Still, it is particularly important to include the grey literature, such as agency and funder reports which may show more outcomes of limited effects than those selected for peer-review publications. To date, no systematic review has assessed whether voucher programmes have achieved their specified goals.

1.4 Objectives

The primary objective of this review is to summarise the evidence on the impact of voucher programmes on health services in developing countries. The literature on health voucher systems is evaluated to determine the extent to which voucher systems have successfully targeted specific groups, improved the quality, efficiency and use of health services, and improved the health of populations in developing countries.

2. Methods

Outline of chapter

Section 2.1 describes the user involvement in the review process and plans for distribution of findings to important users.

Section 2.2 details the process on how relevant studies were identified and the inclusion/exclusion criteria used to determine whether studies would be included in the analysis.

Section 2.3 explains the methods used for synthesising the results of various studies.

Section 2.4 describes the process used for deriving conclusions from the included studies.

2.1 User involvement

The research question guiding this review was developed by research and policy staff at the UK Department for International Development (DFID) in an effort to strengthen evidence-based decision-making.

The intended users of this review are primarily policy-makers within donor government agencies and recipient government ministries of health and finance who may consider establishing health voucher programmes in developing countries. Additional potential users include programme staff at non-governmental organisations, health care providers, and academics involved in health access and health systems strengthening programmes. This report aims to provide these stakeholders with an impartial review of the evidence on health vouchers, and under what circumstance they have been successful, to help inform decisions on whether health voucher programmes should be considered for future applications.

Using existing publications and government reports, key personnel with experience working on voucher programmes were identified and contacted to help inform this report. These key contacts helped review both the study protocol and the preliminary list of studies identified and included in the analysis, in order to identify gaps in the findings and other programmes/studies that should be examined.

2.2 Identifying and describing studies

2.2.1 *Defining relevant studies: inclusion and exclusion criteria*

Eight different criteria were used to determine whether a study was included or excluded from the review. These criteria are also available in Appendix 2.1.

1. Health voucher programmes - included studies were limited to those regarding health voucher programmes that provided health goods/services. Examples of relevant health goods/services are: skilled provider care; hospital and clinic services; health insurance; pharmaceuticals; family planning products; ITNs for the prevention of malaria; and vaccinations. Publications and studies not related to voucher programmes or concerning voucher programmes delivering food, clean water, and non-health education were not included in this review, even though they may have a health impact.

2. Language - included studies were limited to those with an abstract published in English. Non-English publications with an English abstract were reviewed for relevance and an appropriate translation was sought when necessary.
3. Population - included studies were limited to those located in developing countries at the time the voucher programme was operating. The Human Development Index (HDI) was used to determine development, and voucher programmes located in a country assessed as “very high human development” by HDI were excluded from the analysis.
4. Time frame - included studies were limited to those published from 1960 to 2010. The 1960 cut-off date was chosen because the background literature indicates that the earliest health voucher programmes for which there are evaluation data (reproductive health care in Taiwan and Korea) occurred during the 1960s. No programmes prior to the 1960s were identified and no studies indicated that earlier voucher programmes existed.
5. Type of study - included studies were limited to those that evaluated some aspect of a health voucher programme and contained some quantitative evidence. General descriptions and opinion pieces on voucher programmes were not included in the analysis. Additionally, summaries that discussed evaluation findings but did not provide a methodology section were excluded unless they specifically cited a description of the methods in a separate publication.
6. Study designs - included studies evaluating voucher programmes required an observable comparison such as measurements at two points in time (e.g. before and after programme implementation); control group (e.g. non-voucher control areas, or non-voucher patients); control programme (e.g. supply-side programme delivering the same health goods/services in the same location); or comparison with accepted benchmarks of success (e.g. national screening programme statistics). Studies reporting implementation statistics such as increase in the number of services delivered were excluded unless they measured proportions that could be tested for statistically significant levels of change. Relevant study designs included: RCTs, non-randomised trials (time series), case-control, cohort, pre-post with and without controls.
7. Voucher characteristics - included studies were limited to voucher programmes that operate where health aid is distributed to a population of potential users (either for free or at subsidised price) through a physical voucher or a voucher-like targeting mechanism, such as a “poverty card”, and vouchers are used for provider reimbursement. Studies were included if the use of vouchers played a substantial role in the intervention or if evaluation outcomes were specific to the role of vouchers. Evaluation studies of a broader intervention where the use of vouchers was a minor component of the intervention were excluded.
8. Study outcomes - included studies examined at least one variable that fit into one of the five categories of interest: targeting; utilisation; quality; efficiency; and health impact. Studies without any relevant outcomes were excluded, however; no such studies were identified.

2.2.2 Identification of potential studies: search strategy

The search was conducted in two phases. In the first phase the following sources were searched with a start date of 1960:

- *Bibliographic databases:* PubMed, POPLINE, ELDIS, Inter-Science (Wiley), ScienceDirect, The Cochrane Central Register of Controlled Trials (CENTRAL), The Database of Abstracts of Reviews of Effectiveness and the EPOC Register, Web of Science, Bioline International, World Health Organisation Library Information System (WHOLIS), African Healthline.
- *Publishers' pages of key journals:* Journal of Development Effectiveness Health Policy, The Lancet, Global Public Health, Health Policy and Planning, International Journal of Health Planning and Management (Appendix 2.3).
- *Organisation and network websites:* UNICEF, US Agency for International Development (USAID), DFID, KfW (German Development Bank), the World Bank, Partnerships for Health Reforms, Abt Associates, Instituto CentroAmerica de la Salud (ICAS), Management Sciences for Health (MSH), Oxford Policy Management, the Private Sector Partnerships-One, Marie Stopes International, Population Council, Global Partnership on Output-Based Aid (GPOBA), Results-Based Financing for Health (RBFHealth), Reproductive Health (RH) Vouchers, International Initiative for Impact Evaluation (3ie), and Research for Development.

The basic search terms used were: (voucher* OR coupon* OR output-based* OR "output based" OR "result based" OR "results based" OR results-based* OR "performance based" OR performance-based* OR pay-for-performance OR "pay for performance" OR "demand side" OR demand-side) AND (developing countr* OR "poor countr*" OR "low-income countr*" OR "low-resource countr*" OR "low and middle income") modified as necessary according to database (See Search Strategy - Appendix 2.2). These terms were selected based on a preliminary review of the literature as well as a scan of the current terms used to refer to these types of programmes by the major funders and implementing organisations.

The results of the bibliographic database searches were saved in separate Endnote libraries for abstract review. The findings from publishers' pages, organisations and network websites were saved in a separate file.

In the second phase of the search, three research activities took place. First, bibliographic back-referencing was conducted; reference lists of all the reviewed full-text studies were examined to identify additional, secondary resources. General articles describing, but not evaluating voucher programmes were also reviewed for other potential references. Secondary references were sought if the text indicated they contained a health voucher programme or if the title of the reference indicated a potential health voucher programme evaluation. The full-text of the obtained secondary references were also reviewed for further studies until no new studies were identified.

Next, a supplemental keyword search in PubMed and google.com was conducted based on leads generated by the search described above. For example, when the first phase of the search identified an included programme for maternal health services in India called *Chiranjeevi Yojana*, searches were conducted using "India AND maternal AND voucher" and "*Chiranjeevi Yojana*" to identify any additional information on the voucher programme that may include evaluation information relevant to the analysis. Supplemental searches were conducted for all identified health voucher programmes suspected of being relevant for this review. Appendix

2.4 details the search terms used for the supplemental searches and the references identified through the supplemental search.

Additionally, experts in the field as identified in the literature (listed in Appendix 2.5) were consulted to make further suggestions of publications for consideration, particularly for unpublished studies. A list of preliminary identified health voucher programmes and evaluation references was provided to the experts and they were asked to comment on whether they were aware of additional health voucher programmes and/or evaluation findings that should be considered for the systematic review.

2.2.3 Screening studies: applying inclusion and exclusion criteria

In phase one, the application of the inclusion and exclusion criteria took place in three rounds. In the first round, one team member reviewed abstracts by using the first five criteria specified in section 2.2.1: health voucher programme; language; population; time frame; and type of study. If no abstract was provided, studies were sought for retrieval if the study title contained the words “voucher(s)” “coupons” or “output-based” or if the study title contained the name of a known voucher programme. Full text copies of studies that were not excluded in round one were sought for further analysis.

In round two, two team members independently applied all of the specified criteria listed in section 2.2.1 to the retrieved studies and additional studies identified in the grey literature to determine whether the study should be included for analysis based on the full text of the study. In the case of a discrepancy between the two reviewers’ assessments, a third team member reviewed the study for a decision.

In round three, further studies retrieved through the reference review, supplemental search, and key expert contacts were reviewed by two team members, applying all the specified criteria. After the three rounds of review, the included studies were entered into a separate workbook in Microsoft Excel.

For phase two, where references were obtained through reference lists, supplemental searches or key expert contacts, one reviewer identified potentially relevant studies, as described above. Two reviewers evaluated all selected studies to determine inclusion or exclusion from the systematic review.

2.2.4 Characterising included studies

Information from the studies included in the review was extracted using a data extraction form (see Appendix 2.6) and data were entered into a Microsoft Excel file. Information extracted on voucher programmes included: location of voucher programme; time period of voucher programme; type of health voucher programme; funders; targeted population; targeting mechanism; private or public providers; programme management agency; and programme scale. Information on included evaluation studies related to: study design; study time period; relevant study outcome(s); authors; and publication date.

2.2.5 Identifying and describing studies: quality assurance process

Pilot testing of key word searches was conducted to ensure that the keyword list and limits were neither too broad nor too narrow. The round one criteria for inclusion/exclusion at the abstract level was straight-forward; however, any hesitation by the reviewing team member as to whether the abstract should be included or excluded defaulted to inclusion so that two individuals could assess the criteria in round two.

The two team members applying inclusion/exclusion criteria in round two did so independently and discrepancies were arbitrated by a third team member to discuss and decide on whether the study was to be included or excluded. The same strategy applied to characterising and synthesising the data extracted from the studies.

2.3 Methods for synthesis

2.3.1 Overall approach to synthesis

This systematic review follows a narrative synthesis approach, which aims to provide a narrative summary as opposed to a purely statistical summary of the findings due to the heterogeneity of studies and outcome variables examined (Rodgers et al. 2009). In particular, this review focuses on a synthesis of the evidence found in multiple studies (Popay et al. 2006). A narrative synthesis such as this one can serve to combine results from heterogeneous studies and outcomes when a meta-analysis is not feasible and can also strengthen the arguments and evidence from meta-analyses.

This review summarises and makes conclusion around five themes with regard to expectations about voucher programmes and health goods/service delivery (targeting; utilisation; quality; efficiency; health impact). Sections 1.2.2 and 1.2.3 describe the causal logic and hypotheses regarding health voucher programmes. Additionally, the synthesis uses sub-group analyses to review results by type of voucher programme, geographic location, and voucher programme characteristics.

Evaluation studies of health voucher programmes may contain several outcomes, covering multiple themes per publication. As a result, for the purpose of synthesising the results, the unit of analysis used is the outcome variable, nested within individual studies.

2.3.2 Extracting outcome variables

Each study was reviewed for relevant outcomes in the following areas:

- **Targeting** - the extent that vouchers reach and are redeemed by the intended recipients (e.g., low-income individuals). Example: income level of voucher users compared to non-voucher users.
- **Utilisation** - the extent that voucher programmes change the utilisation of health good/services. Example: use of an ITN among residents in voucher areas compared to residents in non-voucher areas.
- **Efficiency** - the extent that voucher programmes deliver health goods/services efficiently. Example: cost of STI cured through a voucher system compared to pre-voucher system.
- **Quality** - the extent that voucher programmes increase the quality of health goods/services being provided. Outcome variables associated with quality may include several different measures. Typically, quality measures are classified as structure, process, or outcome measures. Example: satisfaction with care ratings among voucher patients compared to non-voucher patients.
- **Health Impact** - the extent that voucher programmes improve the health of the population. Example: prevalence of STIs before and after voucher programme intervention.

Outcomes addressing these areas were selected for analysis if they fit the following criteria:

1. The outcome was designed by the study authors to assess voucher programme performance. This was determined based on understanding of study objectives and the discussion of the outcome in the text of the study.
2. The outcome variable reflected a comparison that could be measured through statistical significance tests or against established benchmarks or competing programmes within the same country. If statistical significance was not given but the information was available to compute a test of significance, the outcome was included.
3. The outcome variable was related to the evaluation of a voucher programme. In instances where the use of vouchers was part of a broader intervention where vouchers played a fairly minor role, only outcomes that specifically address the voucher component were included.
4. The outcome variable addressed the overall sample in the study or a specific sub-group of interest. For example, if a study on utilisation of ITNs gave before-and-after utilisation statistics for multiple groups (e.g. the overall population; pregnant women; children under five; men; women; adolescents) then the outcomes selected will be the specified targeted populations of interest in the study (e.g. the overall population, pregnant women, and children under five).

One reviewer initially selected the outcomes for synthesis and classified the outcomes into one of the five categories. A second reviewer examined the data extraction and reviewed whether the outcomes were appropriate, categorised correctly, and whether any additional outcomes should be considered for analysis. The two reviewers discussed any discrepancies and a third reviewer refereed any disagreements.

2.3.3 *Assessing quality and confidence in study findings*

Due to the heterogeneous nature of outcomes and study designs examined in this review, it is important to assess the quality of included studies and the confidence in the study findings in order to assess the risk of bias in synthesising the literature. The original intent for assessing quality and risk of bias was to rely on the guidelines recommended by the Cochrane Collaboration for assessing study quality using the CONSORT checklist (Schulz et al. 2010) for RCTs, cluster RCTs, controlled before and after, and interrupted time series and the Newcastle-Ottawa Scale (Wells et al. 2010) for case-control and cohort studies. However, we found very few studies that could be adequately assessed using these tools and therefore adapted these tools and others identified in the literature to better serve the purposes of this review. The tool utilised for this review, the “Confidence in Findings Assessment” (CFA) checklist is found in Appendix 2.7.

The foundation for the CFA was the Newcastle-Ottawa scale, which divided the checklists for quality evaluation into three categories: selection, comparability, and exposure (case-control studies) or outcome (cohort studies). The CFA also uses the “selection” and “comparability” categories and a third category of “measurement”, capturing questions about exposure and outcome measurement. Next, the CFA was augmented based on suggestions in the literature of steps to take when assessing the quality of studies, particularly relying on an article by Fowkes and Fulton (1991).

The assessment tool was named the “Confidence in Findings Assessment” instead of a quality assessment tool because the purpose of the tool was not just to judge the quality of the research performed, but rather to assist in gauging one’s confidence that the findings presented accurately assess the results of the voucher

programme. For example, cross-sectional studies based on non-random samples may be executed well (e.g. little missing data, control for important confounding variables, thoroughly described methodology) and be deemed a fairly high-quality execution. However, there is only so much confidence one can place in a cross-sectional study of a non-random sample of the population and therefore the findings are less likely to be regarded with “high confidence”. As such, questions regarding overall study design, magnitude of effect, and statistical significance were added to the CFA.

After going through the four categories on the CFA (study design; selection; comparability; measurement), a set of “final consideration” questions were used to summarise one’s assessment of the study. Next, the user of the CFA selected an “overall judgement” in one of the following categories:

- I have some MAJOR concerns about the methods used or the lack of information available on this study and therefore seriously question the findings. (*low confidence*)
- I have some MINOR concerns about the methods used or the lack of information available on this study and therefore would consider the findings with some caution. (*medium confidence*)
- I do not have concerns about the methods used or information provided on this study and consider the findings with confidence. (*high confidence*)

Using the CFA, two researchers independently evaluated each included study and provided an overall judgement. The overall judgements were compared and a third reviewer refereed any discrepancies between the two reviewers. Next, the studies were grouped by low, medium and high CFA scores and reviewed to ensure that the classification groupings held together.

2.3.4 Selection of outcome variables for synthesis

A maximum of three outcome variables were allowed from the same study in any given outcome category. The selection of three outcome variables was based on a review of preliminary studies identified prior to the protocol development, where it was noted that some studies had multiple outcome variables that could potentially bias the results if all outcome variables were included. Therefore, the protocol established a limit of three variables which would allow for the key outcome variables to be included, but limit the influence of studies that reported multiple outcome variables in the same category.

Thus, for the six studies with greater than three outcomes in a category, a two-stage process was used to determine variables for synthesis. First, outcome variables were reviewed to determine if they could logically be condensed into fewer outcomes. For example, when a list of quality outcomes around physician knowledge of several separate family planning services all had the same finding (no statistically significant effect), they were condensed into “provider knowledge of family planning”. If no logical combination occurred or there were different summary results among related items, then the three outcome variables that were deemed most relevant to the outcome category and most representative of the overall findings were selected. Two researchers independently selected outcome variables for synthesis using this strategy and a third researcher refereed any disagreements.

2.3.5 Process used to combine/ synthesise data

For each outcome variable used for synthesis, information was recorded on the overall direction of effect as: (1) no effect; (2) positive effect; or (3) negative effect. A finding of “no effect” was typically one of no statistical difference from the point of comparison; for example, a non-statistically significant difference in disease prevalence between individuals located in areas with voucher programmes and control areas. A finding of “positive effect” was one that reflected positively on the effectiveness of voucher programmes. For example, voucher users having statistically significantly fewer complications after childbirth compared to non-voucher users was considered a positive finding. Finally, a “negative effect” was recorded for findings that found voucher programmes performing worse than their comparison group. For example, voucher recipients being wealthier than non-voucher recipients when the targeted population was specified as the poor.

In addition to the direction of effect, the following information for each study outcome was recorded: first author last name and year (e.g. Hatt et al. 2010); location and type of health voucher programme (e.g. Bangladesh maternity services); comparison (e.g. time and control areas); and CFA assessment (e.g. high confidence). This information was then used to derive conclusions about the overall findings, as described below.

The outcome variables were reviewed to determine whether a statistical meta-analysis could provide further insight into the effectiveness of voucher programmes. Due to the heterogeneity of the outcome variables and study designs, however, it was determined that a meta-analysis would not be particularly useful for this report.

2.4 Deriving conclusions and implications

In deriving conclusions and implications, we established a system for evaluating the evidence for each of the conclusion categories during the protocol stage of the review, based on the preliminary list of voucher programmes and evaluations identified. The system reflects internal a priori discussions on what the research team felt would be convincing evidence on the success or failure of health voucher programmes, in addition to a review of relevant literature on systematic reviewing methodology.

The synthesis system used the four primary measures: (1) the total number of outcome variables examined; (2) the total number of voucher programmes evaluated; (3) the direction of effect of the outcome variables; (4) and the CFA scores. Based on the aggregated outcome variables for each outcome category, we concluded that the evidence supported one of five conclusion categories:

1. Insufficient evidence - indicating that there was not enough evidence available to determine the relationship between voucher programmes and the outcome category. A conclusion of insufficient evidence was made if there were fewer than four variables in a particular outcome category, if all outcomes only derived from one voucher programme (e.g. Nicaragua cervical cancer programme), or if all outcomes derived from studies with a low CFA.
2. No effect - indicating that the evidence suggested that vouchers did not have an effect on the outcome category. A conclusion of no effect was made if more than 50% of outcomes within a category indicated there was no effect. A pre-established exception to this rule was when all of the “no

effect” conclusions come from low CFA studies and at least 25% of the outcomes from medium/high CFA studies found a significant effect.

3. Conflicting evidence - indicating that vouchers have had both positive and negative effects on the outcome category and may signal a need for sub-analysis to indicate under what conditions voucher programmes have positive or negative findings. A conclusion of “conflicting evidence” was drawn if two different high CFA studies or sets of medium/low CFA studies (25% or greater) had findings in opposing directions.
4. Modest evidence - indicating that there was modest evidence that voucher programmes had an impact on the outcome category. A conclusion of “modest evidence” was made if there was evidence indicating a positive or negative relationship; however, the evidence was not strong enough to be called robust. The outcomes may have derived from fewer than four voucher programmes or the confidence in the study findings may not have been adequate to qualify for robust evidence.
5. Robust evidence - indicating that there was clear and convincing evidence that voucher programmes had a significant positive or negative impact on the outcome category. A conclusion of “robust evidence” was drawn if four or more voucher programmes were reviewed, 50% of the findings (in the same direction) derived from medium or high CFA studies, and no conflicting evidence from medium/high CFA studies was found.

Appendix 2.8 shows a decision tree that depicts how the conclusion categorisation occurred for the overall outcome categories and sub-group analyses. Calculations of percentages relevant to the synthesis were conducted in Microsoft Excel.

In order to assess the stability of the conclusions, a sensitivity analysis was performed. A sensitivity analysis measures the impact on the results of an adjustment of one or more characteristics of the studies. The strength of inference is greater if the results are unchanged under varying conditions. For each outcome category with sufficient evidence to assess stability, the conclusions were examined under three hypothetical situations: (1) if one additional outcome variable of “no effect” was added; (2) if one additional outcome variable of “positive effect” was added; and (3) if the median outcome was deleted. The hypothetical “additional outcome” variable added to assess stability was assigned a CFA of “medium confidence” and did not represent an additional study or voucher programme. Any conclusions that changed under these situations were deemed as unstable results, to be considered with caution.

3. Search Results

Outline of chapter

Section 3.1 describes the results from the search and screening process.

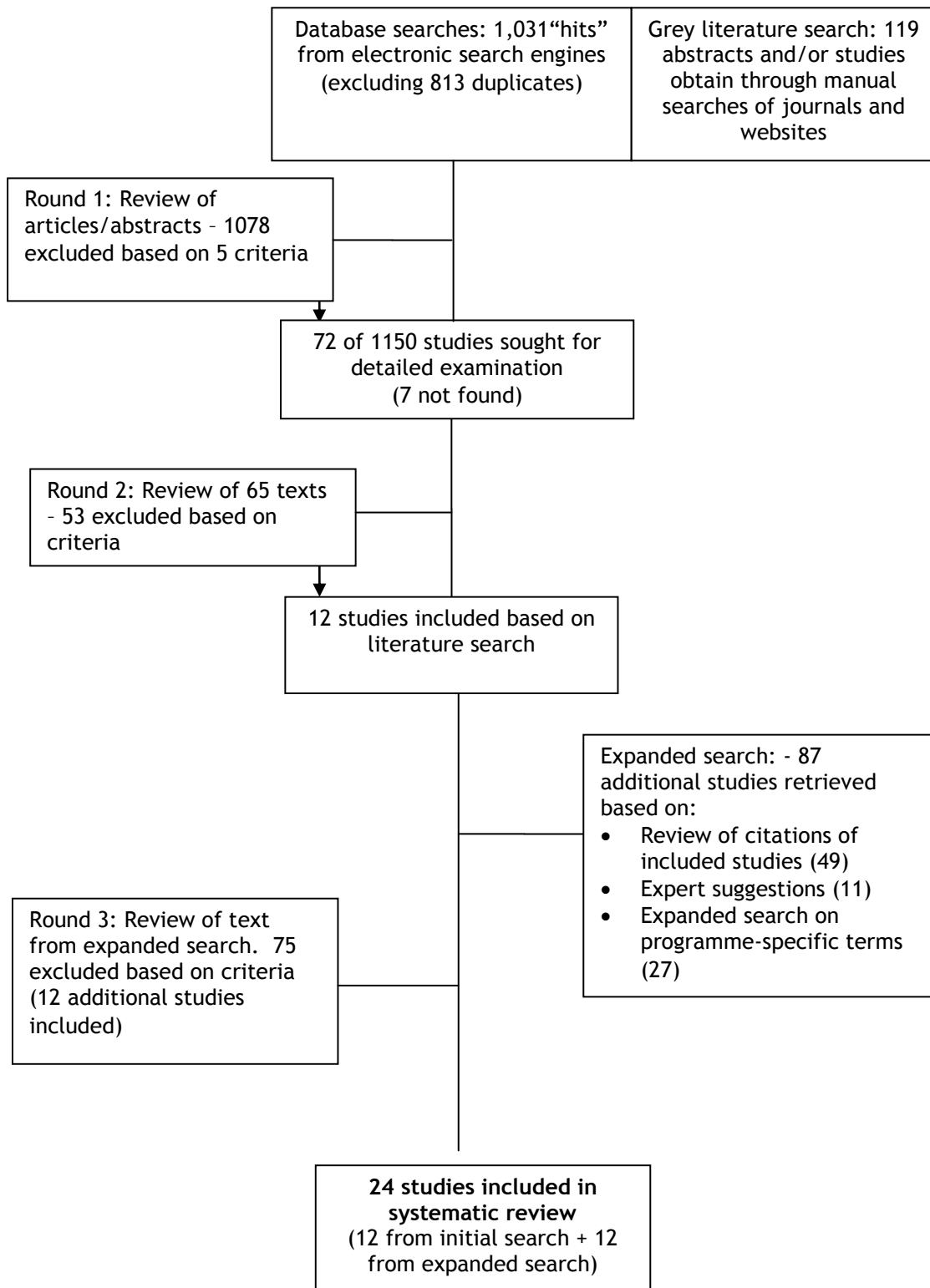
Section 3.2 reports on the voucher programmes identified and the quantitative evaluation studies included in the analysis.

3.1 Studies included from searching and screening

The search was conducted from September to November 2010 (Appendix 2.2). A total of 24 studies were included in the analysis for the purposes of synthesising the data. Figure 3.1 details the flow diagram of the filtering process used to identify the included studies. Initially 1,031 abstracts were reviewed from electronic database searches and an additional 119 studies or abstracts were reviewed from the grey literature search. Of this 1,150 total, 1,078 were eliminated based on the first five criteria in round one. Of the 72 selected for full text review, 65 (90%) were obtained while seven studies could not be located. Some of the sources that could not be found were government documents that were over 20 years old, such as reports on voucher programmes in Iran and Korea. Two researchers independently applied the review criteria and came to the same conclusion for 95% of the studies. A third team member decided on the inclusion/exclusion for three studies.

In round two, based on the full-text of the 65 studies, 53 were eliminated when applying all of the criteria listed in section 2.2.1 and 12 studies were included. The expanded search processes yielded an additional 87 studies for review, 49 identified through the reference lists of studies reviewed in round two, 27 identified through supplemental searches, and 11 identified through expert contacts listed in Appendix 2.5. Of the 87 studies from the expanded search studies, 75 were excluded and 12 were included based on all eight criteria with 98% agreement between the two reviewers. In total, 24 studies were included in the review.

Figure 3.1: Flow diagram of identifying included studies



3.2 Details of included studies

3.2.1 Identified health voucher programmes

A total of 43 health voucher programmes located in 27 countries were identified in reviewing the abstracts and studies as described in section 3.1. Table 3.2.1 details the 43 programmes identified. Common health goods and services provided through vouchers included maternity services, other reproductive health services such as family planning and STI care, and ITNs.

For 17 of the voucher programmes identified in Table 3.2.1, only descriptions of the voucher programmes were found -- no evaluation data were identified. A further ten programmes had some evaluation information; however, the evaluation study or studies did not meet the inclusion criteria established in section 2.2.1.

Table 3.2.1: Identified Health Voucher Programmes

Location and time frame	Health Goods/Services	Targeted Population	References
Antigua (1973-1974)	Contraception	Poor men and women	Isaacs 1975
Armenia (2008 - present)	Maternity services	Pregnant women	Truzyan et al. 2010
Bangladesh* (2006-present)	Maternity services	Poor pregnant women	Schmidt et al. 2010; Rahman 2009; Hatt et al. 2010
Cambodia* (2007 - present)	Maternity services	Poor pregnant women	Ir et al. 2010
China (1995 - 2002)	Maternal and child health services	Poor women and children	Gorter 2003
China/Vietnam (2002 - 2006)	Condoms and needles	Injection drug users	Hammett et al. 2003
Costa Rica (1970s)	Contraception - birth control pills	Poor women	Isaacs 1975
El Salvador (2002 -)	Cervical cancer screening	At-risk women	Gorter 2003
Ethiopia (2005 - 2006)	ITNs	Pregnant women	NetMark 2010
Ghana (2004-2008)	ITNs	Pregnant women and mothers of young children	NetMark 2010
India (Agra) (2007-2009)	Maternity services, Family Planning, STI services	Poor women	Donaldson et al. 2008
India (Janani Suvidha Yojana) (2006-present)	Maternity Services	Poor women	Gupta et al. 2010

Location and time frame	Health Goods/Services	Targeted Population	References
India (MAMTA)* (2008-present)	Maternity Services	Poor women	Nandan et al. 2010
India (Sambhav) (2009-present)	Family Planning	Poor women	Gupta et al. 2010
India (Seva Mandir) (2003-present)	Maternity Services	Poor women	Gupta et al. 2010
India (Gujarat)* (2005 -present)	Maternity services	Poor pregnant women	Bhat et al. 2009
India (Kolkata slums) (1999-2002)	Reproductive and child health care	Poor women and children in slums	Gorter 2003
Indonesia (Aceh) (2005)	Midwifery services post tsunami	Pregnant tsunami refugees	JEPIEGO 2005
Indonesia (Pemalong district) (1998 - 2003)	Midwifery services	Poor pregnant women	Nachuk 2006
Iran (1970s)	Family planning	Married women	Treadway et al. 1976
Kenya (2006 - present)	Maternity, family planning, and gender violence recovery services	Poor persons in need of services	Janisch et al. 2010
Kenya (Nyeri) (1997 - 2001)	Reproductive health services	Youth ages 10-24	Erulkar et al. 2004
Korea (1965-1979)	Intrauterine devices (IUDs) and sterilization	Women at risk of pregnancy	Robey 1987
Mali (Koutiala circle) (2004)	ITNs	Pregnant women	NetMark 2010
Mozambique* (2005)	ITNs	Mothers of young children	De Oliveira et al. 2010
Nicaragua (Managua)* (1995 - 2009)	STI services	High-risk groups (i.e. sex workers, clients, transvestites, glue sniffers)	Borghini et al. 2005; McKay et al. 2006
Nicaragua (Managua)* (2000 - 2005)	Sexual and reproductive health services	Poor urban adolescents	Meuwissen et al. 2006a-d
Nicaragua* (1999 - 2009)	Cervical cancer screening	Poor women 30-65 years	Howe et al. 2005
Niger* (2005-2006)	Long-lasting insecticide-treated bed nets (LLINs)	Children under 5-years-old	Thwing et al. 2008
Nigeria (Lagos state) (Dates unclear)	ITNs	Pregnant women and children under 5	NetMark 2010
Pakistan (DG Khan) (2008-2009)	Family Planning and Maternity services	Low income women in need of services	Bashir et al. 2009; MSI 2010

Location and time frame	Health Goods/Services	Targeted Population	References
Senegal* (2005-2009)	ITNs	Pregnant women and children attending vaccinations	NetMark 2009
Sierra Leone (2010 - present)	Reproductive health services	Low income women	MSI 2010
Taiwan* (1964-1967)	IUDs and sterilization	Couples in need of contraception	Chang et al. 1969
Tanzania (KINET)* (1996 - 2008)	ITNs	Pregnant women and children under 5	Kikumbih et al. 2005
Tanzania (National)* (2003-2006)	ITNs	Pregnant women and children under 5	Hanson et al. 2009; Khatib et al. 2008; Marchant et al. 2008
Uganda* (2006 - 2010)	STI testing and treatment	Poor persons with STI symptoms	Bellows et al. 2009
Uganda (2007 - present)	Maternity services	Poor pregnant women	Arur et al. 2009
Uganda (2004-2006)	ITNs	Pregnant women and under five year old children	Worrall et al. 2005
Vietnam	Reproductive health services	Low income women	MSI 2010
Zambia* (2002 - 2004)	ITNs	Pregnant women and children under 6 months	Grabowsky 2005
Zambia (Lusaka) (2000)	Emergency contraception	Women in need, especially young women	Skibiak 2001
Zambia* (1998-1999)	Health services	Low income individuals	Kondo and McPake 2007

* Included in review

3.2.2 Health voucher programmes with included quantitative evaluations

Of the 43 health voucher programmes listed in Table 3.2.1, 24 quantitative evaluations meeting the inclusion criteria established in section 2.2.1 were identified for 16 voucher programmes. Voucher programmes with evaluation findings included mostly ITN distribution programmes (six programmes) and reproductive health programmes (eight programmes). One general health services programme was included. Appendix 3.1 gives details on each of the voucher programmes that had at least one included evaluation study.

Information provided in Appendix 3.1 includes: location of voucher programme; time period of programme operation; the type of health goods/services provided; description of the role of vouchers in the programme; whether vouchers were for free or subsidised health goods/services; targeted population; mechanism used for targeting; role of public and private providers/retailers; programme scale; external funders; programme managers; and associated evaluation studies included in the review.

3.2.3 Included evaluation studies

Table 3.2.3 provides detail on the 24 included evaluation studies. The majority of evaluations included in this review are observational studies that had either a

pre/post design, a cross-sectional intervention/control design or a combination of the two. One case control study was included (Chang et al. 1969). Two economic modelling studies were included (Borghi et al. 2005 and Kondo and McPake 2007) as well as one clinical record review (Howe et al. 2005). Finally, one simulated patient before-during-after evaluation was included (Meuwissen et al. 2006b).

Of the 16 voucher programmes where quantitative evidence was found, eight were located in Africa, five were in Asia, and three were in Latin America (all located in Nicaragua).

Table 3.2.3: Included evaluation studies

Voucher Programme First Author, Year (Study Time Period)	Study Design	Study Comparison
Bangladesh Maternity Hatt, 2010 (2005-2009)	HH surveys of rural poor pregnant women in districts with the longest duration of the programme and women in matched control districts	Women in matched control districts
Bangladesh Maternity Rahman, 2009 (2007-2008)	Pre/post intervention survey of service providers, field workers and poor pregnant women	Pre-programme data providers, field workers and pregnant women
Bangladesh Maternity Schmidt, 2008 (March 2008)	Data extraction on caesarean sections from district health facilities	Data from non-voucher district health facilities
Cambodia Maternity Ir, 2010 (2006-2008)	Extraction of delivery rates and outcomes from routine health information systems at total hospitals; expected births estimated from DHS data	Nearby rural districts without the voucher scheme
India Maternity Bhat, 2009 (March 2007)	Cross-sectional survey on demographic information of randomly selected rural pregnant women	Rural pregnant women who were not beneficiaries of the voucher scheme
India Maternity Nandan, 2010 (2008-2009)	In-depth interviews with women delivering at MAMTA Friendly Hospitals in five districts	Women delivering at non-MAMTA Friendly Hospitals
Mozambique ITNs De Oliveira, 2010 (Feb. 2006)	Cross-sectional survey on net ownership of HHs in rural districts	Information recalled from before the intervention
Nicaragua Cervical Cancer Screening Howe, 2005 (2003-2005)	Clinical records review of Pap screenings and pathology reports for >25 and < 25 yr olds	Expected values from international and US data sources
Nicaragua STIs Borghi, 2005 (1999-2000)	Cost-effectiveness analysis using data for 1 year of a voucher programme	Costs in the absence of the programme using baseline data, reports and literature
Nicaragua STIs McKay, 2006 (1995-2005)	Medical record extraction on timing of treatments and prevalence of STIs	Change over time through time series analysis
Nicaragua Reproductive Health Meuwissen, 2006a (2000-2002)	Female simulated patient visits before, during and after intervention	Simulated patient scores from before the intervention

Voucher Programme First Author, Year (Study Time Period)	Study Design	Study Comparison
Nicaragua Reproductive Health Meuwissen, 2006b (2000-2002)	Pre/post intervention interviews with providers at SRH clinics	Providers' responses from before the intervention
Nicaragua Reproductive Health Meuwissen, 2006c (2000-2002)	Cross-sectional survey of female adolescents' knowledge and use of contraceptives	Female adolescents who did not receive a voucher
Nicaragua Reproductive Health Meuwissen, 2006d (2002-2004)	Self-administered questionnaires to a random sample of adolescent girls	Adolescent female respondents who did not use the voucher
Niger ITNs Thwing, 2008 (Jan. -Sept. 2006)	National-wide cross sectional survey of HH before and after the campaign	Data from survey conducted prior to campaign
Senegal ITNs Netmark, 2009 (Nov. 2008 - Jan. 2009)	Cross-sectional survey of urban and rural HHs in 2008	Data from 2000 and 2004 surveys
Taiwan IUD Chang, 1969 (1968)	Case-control study using government HH registries to identify characteristics of IUD acceptors and non-acceptors	Individual "matched" controls based on similar demographic characteristics at time of first insertion
Tanzania ITNs Hanson, 2009 (2005-2007)	Cross-sectional HH and facilities surveys early, midway and at the end of the programme	Data from early and midway surveys
Tanzania ITNs Khatib, 2008 (June -Aug. 2006)	Cross-sectional survey on ITN use in randomly selected HH in rural districts	Different distribution arms (vaccination campaign, commercial market, etc.)
Tanzania ITNs Kikumbih, 2005 (1999)	Survey of randomly selected HHs on net use and a costing exercise using secondary data	HHs in districts without voucher programme
Tanzania ITNs Marchant, 2008 (July-Sept. 2008)	Cross sectional surveys on ITN ownership and use of HH in 2008	Data from the 2005- 2007 surveys
Uganda STIs Bellows, 2009 (2006-2007)	HH survey of STI prevalence, knowledge and treatment utilisation before and after the intervention	Data from before the intervention
Zambia ITNs Grabowsky, 2008 (Feb. 2006)	Cross-sectional survey on ITN use in HH with children >5 yrs in rural districts 6-months post- campaign	Responses recalled from before the campaign
Zambia Health Services Kondo, 2007 (1998-2000)	Economic models of characteristics that may predict choice of payment for health services	User charges or prepayment

HH=household; DHS=Demographic and Health Surveys; SRH=Sexual and Reproductive Health; IUD=Intrauterine Device

4. Synthesis Results

Outline of chapter

Section 4.1 provides details on the outcomes included in the synthesis

Section 4.2 reports the findings from the synthesis

Section 4.3 describes the sensitivity analysis for the synthesis findings

Section 4.4 summarises the findings of the synthesis

4.1 Further details of studies and outcomes included in the synthesis

4.1.1 Confidence in Findings Assessment results

Using the CFA described in Section 2.3.3 and shown in Appendix 2.7, two reviewers independently made confidence in findings assessments (high, medium, and low confidence) for the included studies¹. The consensus rate between the two reviewers was 95%. For the one study where the reviewers disagreed on the confidence score, a third reviewer decided on the final score.

A majority of the studies, approximately 70%, were assessed as “medium confidence in the findings”. A score of “medium confidence” means that there were some minor concerns about the study design, methods or the lack of information presented in the study, but overall the findings were believable with regards to the impact of the voucher programme on the outcome variables.

Five studies were assessed as “low confidence”, meaning that the reviewers had some major concerns about the study findings. The rationale for giving a score of “low confidence” was different for different studies and could reflect concerns about the selection of the data, statistical analysis (e.g. lack of controlling for important confounding variables), or the study authors’ own stated concerns about the validity of the findings. The remaining two studies were given a score of “high confidence”, meaning the reviewers were confident that the findings reflected the impact of the voucher programme.

4.1.2 Outcome variables used for synthesis

A total of 64 outcome variables were extracted from 24 studies. In six instances, there were more than three outcomes from the same study in a particular outcome category and therefore the outcomes used in the synthesis were reduced. In three of these instances, the outcome variables were easily combined into three broader outcomes (e.g. two quality outcomes related to physician knowledge of family planning with the same overall finding were reduced to “physician knowledge of family planning”). In the three other instances, the outcomes were not easily combined; however, the overall findings within the larger group of variables were all the same and therefore the selection of any three did not make a difference in the synthesis.

For the targeting outcome category, nine health outcome variables were identified and are detailed in Table 4.1.1. The nine outcome variables used for synthesising the targeting findings derive from six studies and four voucher programmes: India maternity services (MAMTA); Nicaragua cervical cancer screening; Tanzania ITNs; and Zambia ITNs.

¹ For one included study conducted by Venture Strategies for Health and Development (VSHD), a third party not associated with VSHD was given the CFA tool and assessed the confidence in findings score in order to avoid bias.

Table 4.1.1: Targeting - outcome variables included in the synthesis

Study Citation	Study Design	CFA Grade	Outcome Variables (sample size)	Direction of Effect (Statistics)
India maternal health care				
Nandan et al. 2010	In-depth interviews with beneficiary and non-beneficiary women delivering at facilities	low	% of women with a BPL card using maternity services (n=312)	Positive (16.7% vs. 32.4%, p=0.004)
Nicaragua cervical cancer screening				
Howe et al. 2005	Clinical records review of Pap screenings and pathology reports	medium	% of high-risk women screened compared to standard benchmarks (n=1448)	Positive (3.7% compared to 1-5% ACCP; 0.2-1.5% NBCCEDP; 0.45% Bethesda)
Tanzania ITNs				
Hanson et al. 2009	Cross sectional HH and facilities surveys early, midway and at the end of the programme	high	Increase in % of ITNs purchased with a voucher - infant under one (n=1115) Increase in % of ITNs purchased with a voucher - children under five (n=3410) Increase in % of ITNs purchased with a voucher - pregnant women (n=752)	Positive (7% to 50%, p<0.001) Positive (3.5% to 33.5%, p<0.001) Positive (6.3% to 23.6%, p<0.001)
Khatib et al. 2008	Cross-sectional community survey of ITN utilisation comparing various bed net programmes	medium	% of ITNs obtained through a voucher - infants (n=422)	Positive (highest proportion of infant use is through voucher, SS)
Marchant et al. 2008	Cross sectional surveys on ITN ownership and use of HH	medium	SES level of pregnant women receiving and redeeming vouchers (n=1203) SES level of infants receiving and redeeming vouchers (n=1203)	Negative (least poor at 60% to most poor at 39%, p=0.01) No effect (least poor at 23% to most poor at 35%, p=0.5)
Zambia ITNs				

Study Citation	Study Design	CFA Grade	Outcome Variables (sample size)	Direction of Effect (Statistics)
Grabowsky et al. 2007	Cross-sectional survey on ITN ownership and use with retrospective component	medium	Equity ratio between high and low quintiles pre-intervention vs. post intervention (n=369)	Positive (change from SS difference to NSS difference)

SS = statistically significant at $p < 0.05$; NSS = not statistically significant at $p < 0.05$
SES=Socio-economic status; BPL=Below the Poverty Line

Only one outcome variable fitting into the efficiency category was identified in this review and is described in Table 4.1.2. The efficiency outcome variable derived from one study on the voucher programme for STI care for high-risk groups in Nicaragua.

Table 4.1.2 Efficiency outcome variables included in the synthesis

Study Citation	Study Design	CFA Grade	Outcome Variable (sample size)	Direction of Effect (Statistics)
<i>Nicaragua STI care for high-risk groups</i>				
Borghi et al. 2005	Cost-analysis using data for one year of a voucher programme and costs in the absence of the programme using baseline data, reports and literature estimates	medium	Costs per STI case cured (n=1543 patients)	Positive (estimated cost of \$118 STI effectively cured vs. \$200 for competing programme)

Table 4.1.3 provides details on the 30 outcome variables in the utilisation category. The 30 variables are derived from a total of 16 studies of 13 voucher programmes including: maternity services in Bangladesh, Cambodia and India; ITN programmes in Mozambique, Niger, Senegal, Tanzania, and Zambia; reproductive health/STI care in Nicaragua and Uganda; cervical cancer screening in Nicaragua; and general health services in Zambia.

Table 4.1.3 Utilisation - outcome variables included in synthesis

Study Citation	Study Design	CFA Grade	Outcome Variables (sample size)	Direction of Effect (Statistics)
Bangladesh maternity				
Hatt et al. 2010	Pre/post intervention HH surveys of rural poor pregnant women in districts with the longest duration of the programme and women in matched control districts 2	high	% of deliveries attended by skilled providers (n=2,028) % facility-based deliveries (n=2,028) % any ANC (n=2,028)	Positive (p<0.001) Positive (p<0.001) Positive (p<0.001)
Rahman et al. 2009	Pre/post intervention survey of service providers, field workers and poor pregnant women 3	medium	% women using 1+ ANC visit(n=850) % women with trained provider at delivery (n=850) % facility-based delivery (n=850)	Positive (79% to 89%, SS) Positive (5.5% to 21.6%, SS) Positive (2.3% to 18.3%, SS)
Cambodia maternity				
Ir et al. 2010	Actual versus expected facility-based births using administrative information and census data	low	% facility-based deliveries (n=5611 deliveries)	Positive (estimated 16% to 45%, SS not given) ⁴
India maternity (Gujarat)				
Bhat et al. 2009	Cross-sectional survey on demographic information of randomly selected rural pregnant women	low	% delivery conducted by private OB/GYN (n=656) % utilisation of PNC (n=656)	No effect (39% vs. 32%, NSS) No effect (28% vs. 31%, NSS)
India maternity (MAMTA)				
Nandan et al. 2010	In-depth interviews with beneficiary and non-beneficiary women delivering at facility	low	% of women using ultrasound services	Positive (83.3% vs. 97.6%, p<0.001)

² Due to the limit of three variables per study per outcome category, an outcome % in PNC was not used in synthesis. The finding for this variable was a positive effect at p=0.004

³ Due to the limit of three variables per study per outcome category, a series of outcome variables showing the increase in proportion of women experiencing a complication who consulted with a trained provider were not included. Additionally, an outcome on the percent of women using PNC was not used for the synthesis. All of the omitted variables had the same overall finding of a statistically significant positive effect.

⁴ Although SS was not given, the increase appeared dramatic enough that a conclusion of positive effect was given.

Study Citation	Study Design	CFA Grade	Outcome Variables (sample size)	Direction of Effect (Statistics)
Mozambique ITNs				
De Oliveira 2010 et al.	Cross-sectional survey on net ownership of HHs in rural districts	medium	% household ownership of net - entire sample (n=947)	Positive (21% pre to 55% post, p=0.0012)
Nicaragua cervical cancer screening				
Howe et al. 2005	Clinical records review of Pap screenings and pathology reports	medium	% of women who received diagnostic work-ups and treatment compared to standard benchmarks (n=1,448)	Positive (94% compared to ACCP>90%; 86% NBCCEDP)
Nicaragua reproductive health (adolescents)				
Meuwissen et al. 2006(c)	Cross-sectional community-based survey comparing voucher receivers to non-receivers	medium	% use of reproductive health (n=5114)	Positive (34% vs.19%, SS)
			% use of modern FP methods (n=853)	No effect (50% vs. 50%, NSS)
			% use of condoms in last sex act (n=835)	No Effect (23% vs. 20%, NSS)
Niger bed nets				
Thwing 2008	National-wide cross sectional survey of HH before and after the campaign	medium	% household own any net (n=4251)	Positive (67% to 87%, p<.05)
			% household own ITN (n=4251)	Positive (6% to 65%, p<.05)
Senegal bed nets				
NetMark 2009	Cross sectional surveys of urban and rural HHs	medium	% nets used among those who own - voucher vs. free and purchase without voucher (n=2,998)	Positive (77% use with voucher compared to 61% and 53%, p <0.001)
Tanzania bed nets				
Hanson et al. 2009	Cross sectional HH and facilities surveys early, midway and at the end of the programme ⁵	high	% own ITN - all (n=5951)	Positive (18% to 36%, p<0.001)
			% own any net - target groups (n=5951)	Positive (3 groups combined for target groups; all SS at p<0.05)
			% own ITN - target groups (n=5951)	Positive (same as above)

⁵ Due to the limit of three variables per study per outcome category, an outcome % own any ITN (all) was not used in synthesis. The finding for this variable was a positive effect at p<0.001.

Study Citation	Study Design	CFA Grade	Outcome Variables (sample size)	Direction of Effect (Statistics)
Kikumbih et al. 2005	Survey of randomly selected HHs on net use in intervention and control districts	medium	Knowledge of voucher scheme as a predictor of net ownership (n=268) ITN coverage among children under 5 (n=268) Bed net coverage among pregnant women (n=268)	Positive (1.104, p<0.05) Positive (84% vs. 32%, p<0.001) Positive (44% vs. 89%, p=0.004)
Marchant et al. 2008	Cross sectional surveys on ITN ownership and use of HH	medium	ITN ever treated - voucher as predictor (n=7160) ITN effectively treated - voucher as predictor (n=6344)	Positive (p<0.001) Positive (p<0.001)
Uganda STI services				
Bellows 2009	HH survey of STI prevalence, knowledge and treatment utilisation before and after the intervention	medium	% seeking treatment for STI symptoms post intervention - men and women (n=2125)	No effect (34% to 38% for women and 37% to 37% for men, both NSS)
Zambia general health services				
Kondo and McPake 2007	Economic models of characteristics that may predict choice of payment for health services	Low	Health attendances per one year rates of vouchers versus pre-payment (n=5598) Health attendances per one year rates of vouchers versus user fees (n=574)	Negative (1.12 vs. 1.23, SS) Positive (1.75 vs. 1.29, SS)
Zambia ITNs				
Grabowsky et al. 2005	Cross-sectional survey on ITN ownership and use with retrospective component	medium	% of households with an ITN (n=406)	Positive (51% to 76.2%, SS)

SS = statistically significant at p<0.05; NSS = not statistically significant at p<0.05
ANC=Antenatal Care; PNC=Postnatal Care; FP=Family Planning

The quality outcome category includes 13 outcome variables stemming from six studies of three voucher programmes (Bangladesh maternity, Nicaragua cervical cancer screening, and Nicaragua adolescent reproductive health). Table 4.1.4 summarises the outcome variables identified for synthesis.

Table 4.1.4: Quality - outcome variables included in the synthesis

Study Citation	Study Design	CFA Grade	Outcome Variables (sample size)	Direction of Effect (Statistics)
Bangladesh maternity				
Schmidt et al. 2010	Data extraction on caesarean sections from district health facilities	low	% of facility deliveries as caesarean section (poor quality - so fewer or no difference between controls considered positive) (sample size not given)	Positive (SS not given, however graph indicates c-section rates below controls over time)
Rahman et al. 2009	Pre/post intervention survey of service providers, field workers and poor pregnant women	medium	% of services delivered during ANC visit (n=714) % providers performing well on quality indicators at ANC visits	Positive (series of variables that all had SS increases) Positive (series of variables that all had SS increases)
Nicaragua cervical cancer screening				
Howe et al. 2005	Clinical records review of Pap screenings and pathology reports	medium	Reliability of detection PPV (n=1448)	Positive (68% compared to 53.2%-59.7% NBCCEDP; Bethesda 70-75% ⁶)
Nicaragua reproductive health (adolescents)				
Meuwissen et al. 2006(a)	Female simulated patient visits before, during and after intervention	medium	% of providers with appropriate family planning treatment (n=33) % of providers with appropriate STI/HIV prevention treatment (n=33) % of providers with appropriate organisation of clinic (n=33)	No effect (NSS difference from before to during) No effect (NSS difference from before to during) No effect (NSS difference from before to during)

⁶ Although the 68% finding was less than the 70-75% Bethesda range, the effect was still deemed “positive” since the 68% exceeded the National Breast and Cervical Cancer early Detection Programme in the USA and the intent of the authors was to show that the Nicaragua programme could have comparable results to those in developed countries.

Study Citation	Study Design	CFA Grade	Outcome Variables (sample size)	Direction of Effect (Statistics)
Meuwissen et al. 2006(b)	Before and after interviews with physicians assessing knowledge and attitudes	medium	Mean scores of doctors' knowledge of contraceptive use and STI prevention and treatment (n=37)	Positive (6.4 to 8.0; 2.7 to 5.2, SS)
			Mean scores of doctors' attitudes towards SRH accessibility and contraceptive use (n=37)	No effect (5.4 to 7.2; 3.9 to 4.9, NSS combined)
			Mean scores of experiencing medical barriers due to erroneous knowledge or socio-cultural values (n=37)	Positive (4.9 to 6.9; 5.5 to 7.0, SS)
Meuwissen et al. 2006(d)	Cross-sectional community-based survey comparing voucher users to non-voucher users	medium	Patient satisfaction with voucher as explanatory variable (n=700)	Positive (AOR 2.23, SS)
			Satisfaction with reception, with voucher as explanatory variable (n=700)	Positive (UOR ⁷ 1.99, SS)
			Stated that doctor's explanations were clear with voucher as explanatory variable (n=700)	No effect (AOR 1.37, NSS)

SS = statistically significant at $p < 0.05$; NSS = not statistically significant at $p < 0.05$
 PPV=Positive Predictive Value

A total of 11 outcome variables were identified in the health impact category. Table 4.1.5 details the outcomes found in six studies examining six voucher programmes (Bangladesh maternity; India maternity; Nicaragua STI care for high-risk groups; Taiwan IUDs; Tanzania ITNs; and Uganda STI care).

⁷ Adjusted odds ratio not given for this variable.

Table 4.1.5: Health Impact - outcome variables included in the synthesis

Study Citation	Study Design	CFA Grade	Outcome Variables (sample size)	Direction of Effect (Statistics)
Bangladesh maternity				
Rahman et al. 2009	Pre-post survey with no controls	medium	% experiencing NO life-threatening complications during pregnancy (n=463)	Positive (27% to 75%, SS)
			% experiencing NO life-threatening complications during delivery (n=387)	Positive (41% to 75%, SS)
			% experiencing NO life-threatening complications post-partum (n=850)	Positive (44% to 70%, SS)
India maternity				
Bhat et al. 2009	Cross-sectional survey on demographic information of randomly selected rural pregnant women	low	% complications during postnatal period (n=656)	No effect (10% in voucher users and 26% in non voucher users; NSS difference)
Nicaragua STI care for high-risk groups				
McKay, et al.. 2006	Time series analysis of STI prevalence	medium	Prevalence of STIs with time lag between voucher distributions as explanatory variable (n=20 treatment rounds)	Positive (linear relationship, SS p=0.004)
Taiwan IUDs				
Chang et al.1969	Case-control study using government registries and matching IUD acceptors and to non-acceptors	medium	Reduction in live births per 1,000 voucher acceptors compared to matched controls (n=6362)	Positive (cases 381 to 77; controls 376 to 195)
Tanzania ITNs				
Marchant et al. 2008	Multiple cross-sectional surveys	medium	Malaria prevalence among target groups (children < 5, pregnant women) with use of voucher ITN as predictor (n=6051)	No effect (p=0.3, p=0.1)
			Mean haemoglobin levels among target groups with use of voucher ITN as predictor (n=6096)	No effect (NSS, p value not given)
			Anaemia prevalence among target groups (n=6096)	No effect (p=0.6, p=0.4)

Study Citation	Study Design	CFA Grade	Outcome Variables (sample size)	Direction of Effect (Statistics)
Uganda STI services				
Bellows 2009	Community household survey before and after the intervention	medium	Prevalence of syphilis among all respondents ⁸ (n=5184) Self-report STI symptoms	No effect (3% to 3%, NSS) No effect (42% to 40% ,NSS)

4.2 Synthesis of evidence findings

4.2.1 Overall findings

Using the criteria established in Section 2.3.4 and the flow chart in Appendix 2.8, the synthesis process evaluated the impact of vouchers on the outcome categories and assigned a conclusion. Table 4.2.1 summarises the findings for the overall synthesis conclusions, using all 64 extracted outcome variables.

For **targeting**, evidence from four programmes, six studies, and nine total outcome variables was synthesised and it was concluded that there was modest evidence indicating that voucher programmes were able to effectively target specific populations for health goods/services. The evidence was modest instead of robust because one outcome indicated a negative effect, where voucher receivers were wealthier than non-voucher receivers, when the intended targeted population was low-income individuals.

There was only one **efficiency** outcome variable that examined whether a voucher programme was able to deliver health goods/services more efficiently than a competing financing strategy. While the finding from this study was positive, there is insufficient evidence to make a conclusion on this outcome category.

Almost half of the total outcome variables were classified as **utilisation** outcomes. The evidence from 13 programmes, 16 studies, and 30 outcome variables found robust evidence that voucher programmes were able to increase utilisation of health goods/services.

For the **quality** outcome category, the evidence from three voucher programmes, six studies, and 13 outcome variables found modest evidence that voucher programmes were able to impact some dimension of quality for health goods/services. A conclusion of modest evidence was made instead of robust evidence because only three voucher programmes were reviewed in this synthesis and a criterion of at least four voucher programmes is required for a robust conclusion.

Six voucher programmes with six studies and 11 outcome variables were used in synthesising the evidence on **health impact**. The synthesis found that voucher programmes did not have a significant effect on health outcomes.

⁸ Prevalence between 2006 and 2007 did appear to be reduced significantly when using the TPHA and VDRL tests alone, however, the author states that these tests alone have high false positive rates and therefore only the combined test results were used for this analysis.

Table 4.2.1: Overall synthesis conclusions ^a

Outcome Category	# of Programmes	# of Studies	# of Outcomes	% with Effect	% Effect Positive ^b	Conclusion Category
Targeting	4 (2 services 2 goods)	6	9	88	87	Modest evidence
Efficiency	1 (services)	1	1	100	100	Insufficient evidence
Utilisation	13 (8 services 5 goods)	16	30	83	96	Robust evidence
Quality	3 (services)	6	13	62	100	Modest evidence
Health impact	6 (5 services 1 goods)	6	11	45	100	Evidence of no effect

^a The CFA scores were also used in determining the conclusion categories, see Appendix 2.8

^b Of the total variables that found an effect

The protocol for this review indicated that a quantitative analysis would be considered if there were three or more studies with the same outcome variable measured. Out of the 24 studies, three pre/post studies measured utilisation of ITNs. The reviewers and external experts felt that a quantitative synthesis of the data from these three studies would add little value to the overall review. Therefore, no quantitative analysis was conducted.

4.2.2 Sub-group analysis findings

In addition to examining the overall outcome categories, the outcomes were analysed according to sub-groups of interest, including:

- The type of voucher health goods/services - all of the voucher goods were for ITNs and almost all of the voucher services were for reproductive health (maternity services, family planning, etc.).
- Location of the voucher programme - vouchers programmes were classified as being located in Africa, Asia, or Latin America.
- Characteristics of the voucher programme - voucher programmes were examined on whether they provided free health goods/services or a subsidy for health goods/services and whether voucher programmes used public, private or a mix of public/private providers.

Since only one outcome was found in the entire efficiency category, efficiency was not examined for sub-group analysis. Table 4.2.2 presents the findings for the sub-group analyses, excluding all findings of “insufficient evidence” where there were not enough voucher programmes or outcome variables to sufficiently make conclusions. Within the sub-group analysis, almost half of all outcomes are “insufficient evidence” because they examine fewer than four variables or the outcomes represent only one voucher programme.

Table 4.2.2: Synthesis conclusions for sub-groups^a

Sub-group	# of Programmes	# of Studies	# of Outcomes	% with Effect	% Effect Positive ^b	Conclusion Category
TARGETING						
<i>Type of voucher health goods/services</i>						
Health goods (ITNs)	2	4	7	86	83	Modest evidence
<i>Location of voucher programme</i>						
Africa	2	4	7	86	83	Modest evidence
<i>Voucher programme characteristics</i>						
Private providers	2	4	7	86	83	Modest evidence
UTILISATION						
<i>Type of voucher health goods/services</i>						
Health goods (ITNs)	5	7	13	100	100	Robust evidence
Health services	8	9	17	76	92	Robust evidence
<i>Location of voucher programme</i>						
Africa	7	9	16	94	93	Robust evidence
Asia	4	5	10	80	100	Robust evidence
Latin America	2	2	4	75	100	Modest evidence
<i>Voucher programme characteristics</i>						
Vouchers for free goods/services	10	11	19	84	100	Robust evidence
Vouchers for subsidised goods/services	3	5	11	91	90	Modest evidence
Public providers	3	3	5	100	80	Modest evidence
Private providers	5	7	13	77	100	Robust evidence
Public/private providers	4	5	11	91	100	Robust evidence
QUALITY						
<i>Type of voucher health goods/services</i>						
Health services	3	6	13	62	100	Modest evidence
<i>Location of voucher programme</i>						
Located in Latin America ^c	2	4	10	50	100	Modest evidence
<i>Voucher programme characteristics</i>						
Vouchers for free health goods/services	3	6	13	62	100	Modest evidence
Public/private providers	3	6	13	62	100	Modest evidence

Sub-group	# of Programmes	# of Studies	# of Outcomes	% with Effect	% Effect Positive ^b	Conclusion Category
HEALTH IMPACT						
<i>Type of voucher health goods/services</i>						
Health services	5	5	8	62	100	Robust evidence
<i>Location of voucher programme</i>						
Asia	3	3	6	83	100	Modest evidence
<i>Voucher programme characteristics</i>						
Vouchers for free services	3	3	6	83	100	Modest evidence
Vouchers for subsidised services	2	2	5	0	NA	Evidence of no effect
Private providers	2	2	5	0	NA	Evidence of no effect
Public/private providers	3	3	5	80	100	Modest evidence

^a The CFA scores were also used in determining the conclusion categories, see Appendix 2.8

^b Of the total variables that found an effect

^c Finding considered unstable per the sensitivity test described below.

When there is sufficient data, the outcomes for the subgroup analyses often reflect the overall findings listed in 4.2.1. Comparing the results from Tables 4.2.1 and 4.2.2, two main differences are evident. First, the utilisation conclusions change from *robust* to *modest* for voucher programmes in Latin America, those where the voucher is a subsidy as opposed to a free health good/service, and for voucher programmes using only public providers. This change from robust to modest evidence occurs because less than four voucher programmes are represented in these categories.

Second, the overall conclusion of *no effect* for health impact changed when looking exclusively at health services to *robust* evidence. Additionally, *modest* evidence of a health impact effect was found in voucher programmes located in Asia, with free health goods/services, and the use of both public and private providers.

When comparing sub-groups, most differences appear to stem from the availability of evidence, with the difference between *modest* and *robust* evidence stemming from the number of programmes reviewed in the literature rather than reflecting a different outcome in the evaluation findings. One outcome category where clear differences are seen between sub-groups is the health impact category. The findings on health impact indicate that free health goods/services have *modest* evidence of an effect whereas subsidised health goods/services have a finding of *no effect*. Similarly, the use of exclusively private providers has a finding of *no effect*, while the use of public and private providers finds *modest* positive evidence. However, the findings around health impact appear to be unstable, as discussed below.

4.3 Sensitivity analyses

As described in Section 2.4, a sensitivity analysis was conducted to determine whether the synthesis results were unstable or fragile, meaning that small changes in the outcome data could easily change the conclusions. For the overall

conclusions listed in Table 4.2.1, the conclusions for targeting, utilisation, and quality all withstood the three step sensitivity test of adding one positive outcome variable, adding one *no effect* outcome variable, and deleting the median outcome variable. The conclusion of *no effect* for health impact, however, appeared to be unstable with one additional positive outcome variable changing the conclusion to *robust* evidence.

In examining the conclusions listed in Table 4.2.2, one additional variable was found to be unstable, the *modest* effect for quality finding for voucher programmes in Latin America changed to *no effect* when a no effect variable was added.

4.4 Summary of synthesis findings

- Overall, there is *modest* evidence that voucher programmes effectively target vouchers for health goods/services to specific populations.
- Overall, there is *insufficient* evidence to determine whether voucher programmes deliver health goods/services more efficiently than competing health financing strategies
- Overall, there is *robust* evidence that voucher programmes increase utilisation of health goods/services
- Overall, there is *modest* evidence that voucher programmes improve the quality of health services. The evidence was *modest* instead of *robust* because there were only three voucher programmes reviewed instead of the required four.
- Overall, the evidence indicates that voucher programmes do not have an impact on the health of populations; however, this conclusion was identified as unstable through a sensitivity analysis since one additional positive outcome variable would change the conclusion to *robust* evidence.

5. Discussion

Outline of chapter

Section 5.1 discusses the implications of the findings for policy-makers

Section 5.2 discusses the strengths of the research approach

Section 5.3 reviews the limitations to the research approach

5.1 Implications of findings

The findings from this systematic review can assist governments, funding agencies, and other development policy-makers interested in knowing whether current evidence indicates that health voucher programmes have been successful in delivering health aid effectively. In terms of efficiency, however, this review could identify only one study and therefore the evidence on the relative cost-effectiveness of health voucher programmes is lacking and the review was unable to provide a conclusion beyond *insufficient evidence*.

The strongest finding from this review is the indication that health voucher programmes have been successful in increasing utilisation of health goods/services. This finding holds true for health goods/services, in the three examined geographic regions, and for voucher programmes with different characteristics. While some of the sub-group analyses indicate that the *robust* conclusion changes to *modest*, this is only because within these sub-categories fewer than four voucher programmes are being reviewed.

In addition to increasing utilisation, there is also *modest* evidence that voucher programmes can effectively target specific populations and can improve the quality of services. While these results are encouraging, the subsequent link that voucher programmes improve the health of the population is not evident in the data analysed in this review.

The findings regarding health impact are unstable and therefore one should hesitate to make firm conclusions based on the overall conclusion of *no effect*. Based on the sub-group analyses, the evidence suggests that voucher programmes delivering health services, those located in Asia, those offering free health goods/services, or those using public and private providers have a positive health impact. Still, due to the instability of the conclusions around health impact, it is too early to place a substantial amount of confidence in these findings.

It is not particularly surprising that the link between voucher programmes and health impact is more tenuous than other outcomes since it is markedly easier to measure changes in utilisation, success in targeting, and improvements in quality than it is to measure health improvements. Health improvements often take longer to realise and sometimes these improvements may not be evident in the time frame allotted for the evaluation.

With the exception of health impact, the sub-group analyses were limited in their ability to offer new information on the effectiveness of voucher programmes. The large majority of the sub-group analyses findings either did not have enough data to generate a finding beyond *insufficient* evidence, or the sub-group conclusion mirrored the findings from the overall conclusion.

5.2 Strengths of the research approach

There are several strengths to the research approach taken for this systematic review. The first strength is that this review takes a relatively broad view of including evidence in order to maximise the amount of information available for consideration. This is a strength of the narrative synthesis approach. A less inclusive approach, such as a traditional meta-analysis, would substantially reduce the amount of information available for synthesis and thus the results would not be representative of all the available evidence. It was therefore necessary to combine voucher programmes for health goods with those for services, even though these programmes are rather different. For example, the impact of health goods might be much more difficult to measure than the impact of health services.

In this review, evidence is collected from a heterogeneous group of health voucher programmes, types of research publications and types of study designs. For emerging policy issues such as vouchers, for which there are important sources that do not appear in the peer-reviewed literature or do not meet the rigor standards of some review strategies, this approach is particularly beneficial. Additionally, a wide-ranging spectrum of outcome variables was considered in this analysis in order to maximise the amount of data collected. No quantitative evaluation studies of voucher programmes that met our inclusion/exclusion criteria were eliminated because their outcome variables did not fall into one of our outcomes categories.

A second strength of this research approach is that it is adaptable to a wide spectrum of research or policy questions, where the criteria can be tailored to analyse specific sub-groups and where the evidence can be weighed along a number of dimensions (e.g. health goods versus health services, Asia versus Africa, free versus subsidised voucher systems). Additionally, a similar analysis approach can be applied to a competing question, such as “what is the impact of conditional cash transfer (CCT) programmes on the use and quality of health goods/services in developing countries” and one could compare the synthesis results of a voucher approach versus CCT approach on the same outcome categories.

Another strength of this research approach is the use of the sensitivity analysis, which allows one to gauge whether the synthesis results are unstable and if so, provides caution in placing too much emphasis on the unstable findings.

5.3 Limitations of the research approach

Numerous limitations to this research approach exist. First, the focus of this research is on effectiveness evidence with the intent to assist policy-makers in deciding whether or not to consider using voucher programmes when deciding on financing mechanisms. However, this review does not provide insight into those “lessons learned” on implementing voucher programmes that would be useful for policy-makers who have decided to move forward with a voucher programme. Additionally, the findings presented in this synthesis rely primarily on observational studies and do not provide statistical estimates of how much targeting, utilisation, and quality may improve with a voucher programme.

Another limitation of this research is that it relies on subjective judgements and therefore is not completely objective. While assessment tools, explicit criteria, and the use of two independent reviewers serve as quality controls for the process, ultimately this approach does rely on the judgement of the reviewers to extract data and evaluate the confidence of the outcomes.

The lack of validation of the tools used in this analysis is also an important limitation. In particular, the CFA was adapted for this review when other tools that were initially designed for the quality assessment process were not

appropriate for this analysis of such heterogeneous studies and outcomes. While the CFA was useful and appeared effective for this study with a 95% consensus rate, it has yet to be validated by other experts and may be improved to avoid homogeneous results, such as the nearly 70% deemed “medium confidence” in this review. Similarly, the five conclusion categories and the criteria that determine them also need validation. While these categories were developed a priori and generally seem to reflect the overall impressions of the reviewers, they may be improved through further review and enhancing the role of the CFA score in weighting the evidence.

Finally, this review is limited to those evaluations that could be retrieved from the peer-reviewed and grey literature during the study period through the previously defined search strategy. There may be important and relevant evaluations that are not published in the peer-reviewed literature, are not available in libraries, or are not available on the Internet, and are thus not represented in this review. In addition, vouchers are a relatively new form of financing and many programmes have not yet undergone evaluation.

6. Conclusions and Recommendations

Outline of chapter

Section 6.1 summarises the main conclusions from the synthesis

Section 6.2 identifies areas for further analysis and future research

6.1 Conclusions from the synthesis

The overall message from this systematic review is that voucher programmes appear successful in the intermediary goals of targeting specific populations, increasing utilisation, and improving quality. The overall findings on health impact indicate that vouchers do not have an impact on the health of populations; however, this conclusion was found to be unstable and one positive finding on health outcomes would change this conclusion to *robust* evidence. Future evidence from voucher programmes currently being evaluated (Population Council 2010) may change the balance of this assessment, either from more definitive evidence of *no effect* or one of *modest/robust* effect.

6.2 Areas for further analysis and future research

We recommend four main areas for further analysis and future research. First, this review has assessed the evidence around vouchers but not other demand-side financing approaches such as CCTs, user-fee exemptions, and health insurance programmes. Applying the same systematic review methodology to a variety of financing mechanism would allow policy-makers to compare different health financing strategies to determine which strategies have shown to be more effective in reaching certain goals, such as increasing utilisation and improving health status, and under what conditions are certain financing strategies preferable to others.

Second, as stated in the previous section, this systematic review does not evaluate the implementation of voucher programmes and does not systematically collect evidence on the lessons learned in programme implementation. A review of this type would be very helpful for programme managers who are designing voucher programmes and would like information on the relative merits of specific programme design. Questions that could be addressed in a systematic review focused on implementation include: what forms of targeting (e.g. individual-level, district-level, service-level) are most effective for specific health goods/services; what are the expected administrative costs associated with the programme; what are the complicating factors that occur with regards to provider reimbursement?

In terms of new research, it is clear that there is a need for more evidence on the efficiency of voucher programmes and a different approach for analysing data on efficiency. With 38 voucher programmes identified during the course of this review, there is likely sufficient information existing to analyse whether health voucher programmes have been efficient thus far. During the course of this review, some information on costs and efficiency was available; however, it did not meet the criteria for inclusion in the analysis. One possible strategy for analysing efficiency is to focus on collecting cost information for specific health goods/services under a variety of health financing models and to determine how voucher programmes compare to non-voucher programmes.

A final important area for future research and analysis is the incorporation of more evaluation information into this report as it becomes available. At present, there is a multi-site voucher programme evaluation research underway on at least four voucher programmes that should be completed in the next two years (Population

Council 2010). These findings may change whether voucher programmes are considered effective and including this information will be essential in understanding the extent to which voucher programmes are successful in achieving their goals. Additionally, as new and future voucher programmes are established, it is important that funders of health voucher programmes and health voucher programme managers include mechanisms for monitoring and evaluation so that more information can be included into the synthesis.

References

Studies included in review

- Bellows, B, Hamilton M (2009). Vouchers for Health: Increasing Utilisation of Facility-Based STI and Safe Motherhood Services in Uganda. USAID, Bethesda, MD.
- Bhat, R, Mavalankar DV, Singh PV, Singh N. (2009). "Maternal healthcare financing: Gujarat's Chiranjeevi Scheme and its beneficiarie." Journal of Health, Population, and Nutrition **27**(2): 249-258.
- Borghi, J, Gorter A, Sandiford P, Segura Z (2005). "The cost-effectiveness of a competitive voucher scheme to reduce sexually transmitted infections in high-risk groups in Nicaragua." Health Policy and Planning **20**(5): 222-231.
- Chang, M, Liu TH, Chow LP (1969). "Study by Matching of the Demographic Impact of an IUD Programme: A Preliminary Report." The Milbank Memorial Fund Quaterly **47**(2): 137-157.
- De Oliveira, A, Wolkon A, Krishnamurthy R, Erskine M, Crenshae DP, Roberts J, Saute F. (2010). "Ownership and usage of insecticide-treated bed nets after free distribution via a voucher system in two provinces in Mozambique." Malaria Journal **9**(22).
- Grabowsky, M, Farrell N, Hawley W. Chimumbwa J, Hoyer S (2005). "Integrating insecticide-treated bednets into a measles vaccination campaign achieves high, rapid and equitable coverage with direct and voucher-based methods." Tropical Medicine and International Health **10**(11): 1151-1160.
- Hanson, K, Marchant T, Nathan R, Mponda, Jones C, Bruce J, Mshinda H, Schellenberg JA (2009). "Household ownership and use of insecticide treated nets among target groups after implementation of a national voucher programmeme in the United Republic of Tanzania: plausibility study using three annual cross sectional household surveys." BMJ **339**: b2434. doi: 10,1136/bmj.b2434.
- Hatt, L, Nguyen H, Sloan N, Miner S, Magvanjav O, Shrama A, Chowdury J, Chowdury R, Paul D, Islam M, Wang H (2010). Economic Evaluation of Demand-side Financing (DSF) Programme Maternal Health in Bangladesh. Bethesda, MD, Abt Associate Inc.
- Howe, SL, D Vargas, Granada D, Smith J. (2005). "Cervical cancer prevention in remote rural Nicaragua: A programme evaluation." Gynecologic Oncology **99**(3, Supplement 1): S232-S235.
- Ir, P, Horemans D, Souk N, Van Damme W. (2010). "Using targeted vouchers and health equity funds to improve access to skilled birth attendants for poor women: a case study in three rural health districts in Cambodia." BMC Pregnancy Childbirth **10**: 1.
- Khatib, R, Killeen GF, Abdulla SM, Kahigwa E, McElroy PD (2008). "Markets, voucher subsidies and free nets combine to achieve high bed net coverage in rural Tanzania." Malaria Journal **7**: 98.
- Kikumbih, N, Hanson K, Mills S, Mponda H, Shellenberg JA. (2005). "The economics of social marketing: the case of mosquito nets in Tanzania." Social Science & Medicine **60**(2): 369-381.

- Kondo, M and B McPake (2007). "Making choices between prepayment and user charges in Zambia. What are the results for equity?" Health Econ **16**(12): 1371-1387.
- Marchant, T, Bruse J, Nathan R, Mponda H, Sedekia Y, Hanson K (2008). "Monitoring and Evaluation of the TNVS: Report on 2008 Household, Facility services and Facility users surveys." Ifakara Health Institute
- McKay, J, Campbell D, Gorter AC (2006). "Lessons for Management of Sexually Transmitted Infection Treatment as part of HIV/AIDS Prevention Strategies." American Journal of Public Health **96**(10): 1760-1761.
- Meuwissen, L, Gorter AC, Kester AD, Knottnerus JA (2006a). "Does a competitive voucher programme for adolescents improve the quality of reproductive health care? A simulated patient study in Nicaragua." BMC Public Health **6**: 224.
- Meuwissen, L, AC Gorter, Kester AD, Knottnerus JA. (2006b). "Can a comprehensive voucher programme prompt changes in doctors' knowledge, attitudes and practices related to sexual and reproductive health care for adolescents? A case study from Latin America." Tropical Medicine & International Health **11**(6): 889-898.
- Meuwissen, L, AC Gorter, Knottnerus JA. (2006c). "Impact of accessible sexual and reproductive health care on poor and underserved adolescents in Managua, Nicaragua: a quasi-experimental intervention study." Journal of Adolescent Health **38**(1): 56.e51-56.e59.
- Meuwissen, L, AC Gorter, Knottnerus JA. (2006d). "Perceived quality of reproductive care for girls in a competitive voucher programme. A quasi-experimental intervention study, Managua, Nicaragua." International Journal of Quality in Health Care **18**(1): 35-42.
- Nandan D, Nair KS, Tiwari VK (2010). Evaluation of MAMTA Scheme in National Capital Territory of Delhi, Report, January 2010, Department of Planning & Evaluation, National Institute of Health and Family Welfare, New Delhi – 110 067.
- Netmark (2009). "2008 Households Survey on Insecticide-treated Nets (ITNs) in Senegal." Academy for Educational Development.
- Rahman, M, Ubaidur R, Tasnima K (2009). Implementation of the Maternal Health financial Scheme in Rural Bangladesh, Population Council, Bangladesh.
- Schmidt, JO, T Ensor, Hossain A, Khan S. (2010). "Vouchers as demand side financing instruments for health care: A review of the Bangladesh maternal voucher scheme." Health Policy **96**(2): 98-107.
- Thwing, J, Hochberg N, Eng JV, Issifi S, Eliades J, Minkoulou E, Wolkon A, Gado H, Ibrahim O, Newmans R, Lama M (2008). "Insecticide-treated net ownership and usage in Niger after a nationwide integrated campaign." Tropical Medicine & International Health **13**(6): 827-834.

References of other identified voucher programmes

- Arur, A, Nelson Gitonga, Barbara O'Hanlon, Francis Kundu, Maggie Senkaali and Richard Ssemujju (2009). Insights From Innovations: Lessons From Designing

- and Implementing Family Planning/Reproductive Health Voucher Programmes in Kenya and Uganda. USAID. Bethesda, MD.
- Bashir H, Kazmi S, Eicher R, Beith A, Brown E (2009) "Pay for Performance: Improving Maternal Health Services in Pakistan. Health Systems 2020, United States Agency for International Development.
- Donaldson, D, Sethi H, Sharma S. (2008). Vouchers to improve access by the poor to reproductive health services: Design and early implementation experience of a pilot voucher scheme in Agra District, Uttar Pradesh, India. USAID, Washington, D.C., Futures Group International, Health Policy Initiative.
- Erulkar, A., Etyyang LIA, Onoka C, Nyagah FK, Muyonga A (2004). "Behavioural change evaluation of a culturally consistent reproductive health programme for young Kenyans." International Family Planning Perspectives **30**: 58-67.
- Gorter, AC (2003). Review paper on evidence for using competitive voucher schemes and related methods for ensuring young people have access to health service interventions designed to prevent or provide care for HIV / AIDS. Background paper prepared for the Consultation on the Health Services Response to the Prevention and Care of HIV / AIDS among Young People. WHO with UNICEF, UNFPA, UNAIDS, and YouthNet, Montreux, Swiss.
- Gorter AC, Sandiford P, Rojas Z, Salvetto M (2003) Competitive Voucher Schemes for Health. Background Paper. ICAS/Private Sector Advisory Unit of The World Bank Group, Washington, DC.
- Gupta I, Joe W, Rudra S (2010) "Demand Side Financing in Health: How far can it address the issue of low utilization in developing countries?" World Health Report, Background Paper 27.
- Hammett, TM, DC Des Jarlais, Liu W, Ngu D, Tung D, Hoang TV, Van LK, Donghua M. (2003). "Development and implementation of a cross-border HIV prevention intervention for injection drug users in Ning Ming County (Guangxi Province), China and Lang Son Province, Vietnam." International Journal of Drug Policy **14**(5-6): 389-398.
- Isaacs, S (1975). "Nonphysician Distribution of Contraception in Latin America and the Caribbean." Family Planning Perspectives **7**(4): 158-164.
- Janisch, C, Albrecht M, Wolfshuetz A, Kundu F, Klein S. (2010). "Vouchers for Health A Demand Side Output Based Aid Approach to reproductive health services in Kenya " Global Public Health: 1-17.
- JHPIEGO (2005). Progress Made in Rebuilding Indonesia's Health Care System after the Tsunami. Medical News Today (December 20, 2005).
- MSI (2010). "Voucher Schemes for Sexual and Reproductive Health Services: a Marie Stopes International Perspective." Marie Stopes International. London, United Kingdom.
- Nachuk, S. (2006). "Making Service Work for the Poor: Nine Case Studies from Indonesia." INDOPOV - World Bank.
- Netmark. (2010). "Netmark Country Activities." AED Netmark: Public-Private Partnerships for Sustainable Malaria Prevention. Available at: <http://www.netmarkafrica.org/countries/>. Retrieved November 18, 2010.
- Robey, B. (1987). "Community-based Contraceptive Distribution: A Korean Success Story." Asia-Pacific Population & Policy **4**, 1-4.

- Skibiak J P, Chambeshi-Moyo M, Ahmed Y. (2001) Testing Alternative Channels for Providing Emergency Contraception to Young Women. Population Council. July 2001, Washington.
- Treadway, R, Gillespie RW, Loghmani M (1976). "The Model Family Planning Project in Ishfahan, Iran." *Studies in Family Planning* 7(11): 308-321.
- Truzyan, N, Grigoryan R, Avetisyan T, Crape B, Petrosyan V. "Protecting the Right of Women to Affordable and QQuality Health Care in Armenia: Analysis of the Obstetric Care State Certificate Program." Center for Health Services Research and Development of the American University of Armenia, Yerevan, 2010.
- Worrall, E, Hill J, Webster J, Mortimer J (2005). "Expderience of targeting subsidies on insecticide-treated nets: what do we knoe and what are the knowledge gaps?" *Tropical Medicine & International Health* 10(1): 19-31.

Other references used in the text of the technical report

1. Satcher D (2000) Eliminating Global Health Disparities. *JAMA*, 284:2864.
2. Greene M, Merrick T (2005) Poverty Reduction: Does Reproductive Health Matter? World Bank.
3. Deaton A (2003) Health, inequality, and economic development. *Journal of economic literature*. Volume 41, Issue 1, Pages 113-158.
4. OECD (2006) Recent Trends in Official Development Assistance to Health. Organization for Economic Cooperation and Development. Available at: <http://www.oecd.org/dataoecd/1/11/37461859.pdf>. Retrieved December 3, 2010.
5. World Bank (2007) Healthy Development: The World Bank Strategy for Health, Nutrition, and Population Results. Washington DC: World Bank.
6. Bhatia M, Yesudian CAK, Gorter A, Thankappan KR (2006) Demand side financing for reproductive and child health services in India. *Economic and Political Weekly* XLI(3): 279–284.
7. Sandiford P, Gorter A, Rojas Z and Salvetto M (2005) A guide to competitive vouchers in health. Private Sector Advisory Group, The World Bank Group, Washington D.C.
8. Mumssen Y, Johannes L, Kumar G (2010) Output-Based Aid: Lessons Learned and Best Practices. The International Bank for Reconstruction and Development / The World Bank, Washington DC 20433.
9. Azam J, Laffont J (2003) Contracting for Aid. *Journal of Development Economics* Volume 70, Issue 1, February, Pages 25-58.
10. Gorter AC, Sandiford P, Rojas Z, Salvetto M (2003) Competitive Voucher Schemes for Health. Background Paper. ICAS/Private Sector Advisory Unit of The World Bank Group, Washington, DC.

11. Musgrove, Philip (2010) Financial and Other Rewards for Good Performance or Results: A Guided Tour of Concepts and Terms and a Short Glossary. World Bank.
12. Borghi J, Gorter A, Sandiford P, Segura Z (2005) The cost-effectiveness of a competitive voucher scheme to reduce sexually transmitted infections in high-risk groups in Nicaragua. *Health Policy Plan*. Jul;20(4):222-31.
13. McKay JE, Campbell DJ, Gorter AC (2006) Lessons for management of STI treatment programmes as part of HIV/AIDS prevention strategies. *American Journal of Public Health* 2006; 96:7-9.
14. Meuwissen LE, Gorter AC, Knottnerus JA (2006) Perceived quality of reproductive care for girls in a competitive voucher programme. A quasi-experimental intervention study, Managua, Nicaragua. *Int J Qual Health Care*. Feb;18(1):35-42. Epub 2006 Jan 18.
15. Hanson, K, Marchant T, Nathan R, Mponda, Jones C, Bruce J, Mshinda H, Schellenberg JA (2009). "Household ownership and use of insecticide treated nets among target groups after implementation of a national voucher programme in the United Republic of Tanzania: plausibility study using three annual cross sectional household surveys." *BMJ* **339**: b2434.
16. Khatib, R, Killeen GF, Abdulla SM, Kahigwa E, McElroy PD (2008). "Markets, voucher subsidies and free nets combine to achieve high bed net coverage in rural Tanzania." *Malaria Journal* **7**: 98.
17. Killeen GF, Tami A, Kihonda J, Okumu FO, Kotas ME, Grundmann H, Kasigudi N, Ngonyani H, Mayagaya V, Nathan R, Abdulla S, Charlwood JD, Smith TA, Lengeler C (2007) Cost-sharing strategies combining targeted public subsidies with private-sector delivery achieve high bednet coverage and reduced malaria transmission in Kilombero Valley, southern Tanzania. *BMC Infectious Diseases*; 7:121.
18. Marchant T, Hanson K, Nathan R, Mponda H, Bruce J, Jones C, Sedekia Y, Mshinda H, Schellenberg J (2010) Timing of delivery of malaria preventive interventions in pregnancy: results from the Tanzania National Voucher program. *J Epidemiol Community Health*. Jun 7.
19. Mulligan JA, Yukich J, Hanson K (2008) Costs and effects of the Tanzanian national voucher scheme for insecticide-treated nets. *Malaria Journal*, 7:32.
20. Patouillard E, Goodman CA, Hanson KG, Mills AJ (2007) Can working with the private for-profit sector improve utilization of quality health services by the poor? A systematic review of the literature. *Int J Equity Health*, 2007; 6:17.
21. Rodgers M, Arai L, Popay J, Britten N, Roberts H, Petticrew M, Sowden A (2009) Testing methodological guidance on the conduct of narrative synthesis in systematic reviews: effectiveness of interventions to promote smoke alarm ownership and function. *Evaluation*; 15:49-73.
22. Popay J, Roberts H, Soeden A, Petticrew M, Arai L, Rodgers M, Britten N (2006) Guidance on the Conduct of Narrative Synthesis in Systematic Reviews. Economic and Social Research Council, National Centre for Research Methods. Southampton, UK.

23. Schulz KF, Altman DG, Moher D, for the CONSORT Group (2010) CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMC Medicine* 2010, 8:18. (24 March 2010).
24. Wells GA, B Shea, D O'Connell, J Peterson, V Welch, M Losos, P Tugwell (2010) The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa Hospital Research Institute.
25. Fowkes FGR and Fulton PM (1991) Critical Appraisal of Published Research: Introductory Guidelines. *BMJ*; 302:1136-40.
26. Population Council. Evaluating Voucher-and-Accreditation Interventions to Improve Reproductive Health Service Delivery Business Models. 2010 [cited 2010 November 18]; Available at:
http://www.popcouncil.org/projects/231_EvalVouchAccredRH.asp

Appendices

Appendix 1.1: Authorship

Authorship

This technical report was conducted by Venture Strategies for Health and Development (VSHD). The primary authors of this report are Carinne Meyer, MPH/MA and Nicole Bellows, PhD. Additional authors who contributed to the report are Martha Campbell, PhD and Professor Malcolm Potts, MD/PhD/FRCOG. Carinne Meyer is a doctoral student in the School of Public Health at the University of California, Berkeley and Nicole Bellows is an independent consultant working in Nairobi, Kenya. Martha Campbell is the Presidents/CEO of VSHD and Malcolm Potts is a Professor of Public Health and director of the Bixby Center for Population, Health and Sustainability at the University of California, Berkeley.

Reviewers

Anna Gorter, MD/MSc/PhD and Andrew Anglemeyer PhD/MPH reviewed the protocol for this systematic review and will review this draft of the report. Anna Gorter is an expert on voucher programmes and is the project director for vouchers at the ICAS. Andrew Anglemeyer is an expert on systematic review methodology and is the Cochrane HIV/AIDS Group Editor at the Institute for Global Health at the University of California, San Francisco.

Acknowledgements

In addition to the reviewers, four other individuals assisted in the preparation of this review. Sylvia Bereknyei MS, conducted the confidence in finding assessment for an included study associated with VSHD (see conflicts of interest statement below). Additionally, key contacts Micol Salvetto, Ben Bellows PhD, and Claus Janisch contributed information for consideration in this report.

Conflicts of interest

VSHD has previously been involved in evaluating health voucher programmes in Uganda. While VSHD does not have any financial interests in the results of this report, steps were taken to eliminate any bias in reviewing evidence generated through VSHD. An external reviewer, Sylvia Bereknyei MS, who is not associated with VSHD, conducted the “confidence in findings assessment” (see Appendix 2.7) for the study conducted under VSHD.

Appendix 2.1: Inclusion and exclusion criteria

Inclusion/exclusion of studies

Eight primary inclusion/exclusion criteria were used to determine whether a study would be included in the analysis. These criteria are described in section 2.2.1 of the report and are summarised as follows:

1. Health voucher programme - included studies were limited to those regarding a health voucher programmes that provide health goods/services.
2. Language - included studies were limited to those with an abstract published in English.
3. Population - included studies were limited to those located in developing countries at the time the voucher programme was operating, as defined by the HDI.
4. Time frame - included studies were limited to those published from 1960 to 2010.
5. Type of study - included studies were limited to those that evaluated some aspect of a health voucher programmes and contained some quantitative evidence.
6. Study designs - included studies evaluating voucher programmes required an observable contrast such as time, control group, control programme, or comparison with accepted benchmarks of success.
7. Voucher characteristics - included studies were limited to voucher programmes that operated where health aid was distributed to a population of potential users (either for free or at subsidised price) through a physical voucher or a voucher-like targeting mechanism. Studies were included if the use of vouchers played a substantial role in the intervention or if evaluation outcomes were specific to the role of vouchers.
8. Study outcomes - included studies examined at least one variable that fit into one of the five categories of interest: targeting; utilisation; quality; efficiency; and health impact.

There were some instances where abstracts were not available, either due to there being no abstract from a peer-review database or when identifying secondary references from the citations of included studies. In these instances, studies were sought for retrieval if the study title contained the words “voucher(s)” “coupons” or “output-based” or if the study title contains the name or characteristics of a known voucher programme.

Inclusion/exclusion of outcome variables in synthesis

Outcome variables relevant to one of the five outcome categories (targeting; utilisation; efficiency; quality; and health impact) were selected for analysis if they fit the following criteria:

1. The selected outcome was designed by the study authors to assess voucher programme performance.
2. The outcome variable reflected a comparison that could be measured through statistical significance tests or against established benchmarks or competing programmes within the same country. If statistical significance was not given but the information was available to compute a test of significance, the outcome was included. Statistical tests were only considered if the sample size (n) was greater than 30.
3. The outcome variable was related to the evaluation of a voucher programme. In instances where the use of vouchers was part of a broader intervention where vouchers play a fairly minor role, only outcomes that specifically addressed the voucher component were included.
4. The outcome variable addressed the overall sample in the study or a specific sub-group of interest. For example, if a study on utilisation of ITNs gives before-and-after utilisation statistics for multiple groups (e.g. the overall population; pregnant women; children under five; men; women; adolescents) then the outcomes selected would be the specified targeted populations of interest in the study (e.g. the overall population, pregnant women, and children under five).

A maximum of three outcome variables were allowed from the same study in any given outcome category. When there were more than three outcome variables from the same study in a given outcome category, a two-stage process was used to determine variables for synthesis. First, outcome variables were reviewed to determine if they could logically be condensed into fewer outcomes. If not, then the three outcome variables that were deemed most relevant to the outcome category and most representative of the overall findings were selected. Two researchers independently selected outcome variables for synthesis and a third researcher refereed any disagreements.

Appendix 2.2: Search strategy for electronic databases

PubMed

September 12, 2010

search 1- voucher* OR coupon* OR output-based* OR "output based" OR "result based" OR "results based" OR results-based* AND developing countr* OR "poor countr*" OR "low-income countr*" OR "low-resource countr**"

search 2- "performance based" OR performance-based* OR pay-for-performance OR "pay for performance" OR "demand side" OR demand-side AND developing countr* OR "poor countr*" OR "low-income countr*" OR "low-resource countr**"

RESULTS: 132

Web of Science

September 12, 2010

search 1- voucher* OR coupon* OR output-based* OR "output based" OR "result based" OR "results based" OR results-based* AND developing countr* OR "poor countr*" OR "low-income countr*" OR "low-resource countr**"

search 2- "performance based" OR performance-based* OR pay-for-performance OR "pay for performance" OR "demand side" OR demand-side AND developing countr* OR "poor countr*" OR "low-income countr*" OR "low-resource countr**"

RESULTS: 0

Library of Congress

September 12, 2010

search 1- voucher* OR coupon* OR output-based* OR "output based" OR "result based" OR "results based" OR results-based* AND developing countr* OR "poor countr*" OR "low-income countr*" OR "low-resource countr**"

search 2- "performance based" OR performance-based* OR pay-for-performance OR "pay for performance" OR "demand side" OR demand-side AND developing countr* OR "poor countr*" OR "low-income countr*" OR "low-resource countr**"

RESULTS: 0

LISTA

September 12, 2010

search 1- voucher* OR coupon* OR output-based* OR "output based" OR "result based" OR "results based" OR results-based* AND developing countr* OR "poor countr*" OR "low-income countr*" OR "low-resource countr**"

search 2- "performance based" OR performance-based* OR pay-for-performance OR "pay for performance" OR "demand side" OR demand-side AND developing countr* OR "poor countr*" OR "low-income countr*" OR "low-resource countr**"

RESULTS: 26

POPLINE

September 13, 2010

search 1 - voucher* /coupon* /output-based* /"output based" /"result based"
/"results based" /results-based* /"performance based" /performance-based* /pay-
for-performance /"pay for performance" /"demand side" /demand-side &
developing countr* /"poor countr*" /"low-income countr*" /"low-resource
countr*" /"low and middle income"

RESULTS: 39

ELDIS Search

September 14, 2010

NOTE: Not able to use Boolean operators. Since the focus of database is already on developing countries, this was the search strategy used:

voucher* AND health

RESULTS: 30

coupon* AND health

RESULTS: 4

"output based aid"

RESULTS: 32

"pay for performance" health

RESULTS: 55

performance-based health

RESULTS: 4

"results-based" AND health

RESULTS: 9

TOTAL: 134

ScienceDirect

September 14, 2010

search1 - (voucher* OR coupon* OR output-based* OR "output based" OR "result based" OR "results based" OR results-based* OR "performance based" OR performance-based* OR pay-for-performance OR "pay for performance" OR "demand side" OR demand-side) AND (developing countr* OR "poor countr*" OR "low-income countr*" OR "low-resource countr*" OR "low and middle income")

RESULTS: 28,276

search 2 - voucher* AND health AND "developing countr*"

RESULTS: 749

Inter-Science (Wiley)

September 14, 2010

search 1 - voucher* in All Fields OR coupon* in All Fields OR "output-based aid" in All Fields OR "results-based financing" in All Fields OR "performance-based financing" in All Fields OR "pay-for-performance" in All Fields OR "demand-side financing" in All Fields AND "developing countr*" in All Fields OR "low income countr*" in All Fields OR "low resource countr*" in All Fields AND health in All Fields

RESULTS: 4737

search 2 - voucher* in All Fields OR coupon* in All Fields OR "output-based aid" in All Fields OR "results-based financing" in All Fields OR "performance-based financing" in All Fields OR "pay-for-performance" in All Fields OR "demand-side financing" in All Fields AND "developing countr*" in All Fields OR "low income countr*" in All Fields OR "low resource countr*" in All Fields AND "health care" in All Fields between years 1960 and 2010

RESULTS: 2431

Search 3 - voucher* AND health AND developing countr*

Between 1960 and 2010

Results: 761

Cochrane EPOC

September 19, 2010

search 1 - Searching multiple times using different keyword terms did not result in any hits.

voucher* AND health AND developing countr*

RESULTS: 1

WHOLIS Search

September 19, 2010

search 1 - Searching multiple times using different keyword terms did not result in any hits.

search 2 - "voucher\$" and got 2 hits. See word document with abstracts.

African Healthline

September 19, 2010

No hits when searching any of the keywords individually or combined.

Appendix 2.3: Journals hand-searched

1. Journal of Development Effectiveness
2. Health Policy
3. The Lancet
4. Global Public Health
5. Health Policy and Planning
6. International Journal of Health Planning and Management
7. Journal of Development Studies

Appendix 2.4: Supplemental search terms

Supplemental searches were conducted October 28, 2010 - October 29, 2010. Additional references obtained from searches indicated in bullets underneath search terms

PubMed

“Antigua AND coupon”
“Antigua AND family planning”

- Pillai 1973

“Bangladesh AND voucher”
“Bangladesh AND safe motherhood”
“Cambodia AND voucher”
“Cambodia AND matern* AND demand”
“China AND voucher*”
“China AND matern* AND Yunnan AND demand”
“China AND Vietnam AND voucher*”
“China AND Vietnam AND needles”

- Hammett et al. 2006
- Hammett et al. 2007
- DeJarlais et al. 2007

“China AND Vietnam AND condoms”
“Costa Rica AND coupon* AND pill”
“Costa Rica AND coupon*”
“Costa Rica AND pills”
“El Salvador AND voucher AND cervical”
“El Salvador AND cervical”
“Ethiopia AND voucher* AND net*”
“Ghana AND voucher”

- Kweku et al. 2007

“India AND voucher*”
“India AND Agra AND matern*”
“Chiranjeevi”

- Malvalankar et al. 2009

“India AND Kolkata slum* AND reproduct*”
“voucher AND Calcutta”
“Indonesia AND voucher AND midwif*”
“Indonesia AND midwife* AND TPC”
“Indonesia AND midwife* AND Aceh”
“Indonesia AND midwife* AND Johnson”
“Indonesia AND midwife AND Pemalong”
“Iran AND ‘family planning’ AND coupon”
“Iran AND coupon”
“Kenya AND vouchers”
“Kenya AND Nyeri AND Youth”
“Kenya AND matern* AND demand”
“Korea AND voucher*”
“Korea AND coupons”

- Hong et al. 1978

“Mali AND voucher* AND net*”
Added Oct 30 - “Mozambique AND voucher”
“Nicaragua AND voucher AND sex workers”

- Gorter et al. 2006
- “Nicaragua AND voucher AND adolescents”
- “Nicaragua AND voucher AND cervical”
- “Nigeria AND voucher* AND net*”
- “Senegal AND voucher*”
- “Senegal AND nets”
- Yukich et al. 2008
- “Taiwan AND voucher*”
- “Taiwan AND coupon*”
- “Tanzania AND voucher* AND bednet*”
- “Tanzania AND KINET AND eval*”
- Schellenberg 1999
- “Uganda AND voucher*”
- “Uganda AND safe motherhood”
- “Uganda AND bednet*”
- “Zambia AND voucher”
- “Zambia AND emergency contraception”
- “Zambia AND ‘discount card’ AND health services AND 1998”
- “Zambia AND ‘discount card’ AND Mwekera”

Google

- “Antigua AND family planning AND coupon”
 - “Antigua AND contraception AND coupon* AND Westinghouse”
 - “Bangladesh AND matern* AND voucher”
 - “Bangladesh AND “safe motherhood” AND voucher”
 - “Bangladesh AND voucher AND Hatt”
 - Hatt et al. 2010
 - “Cambodia AND voucher”
 - “Cambodia AND Health equity funds AND voucher*”
 - “China AND maternal child health AND voucher*”
 - Rhaman et al. 2009 (Bangladesh)
 - “China AND matern* AND Yunnan AND voucher”
 - “China AND Yunnan AND Kelin AND voucher”
 - “China AND Yunnan AND MCHPAF”
 - “China AND Vietnam AND voucher* AND needles AND evaluation”
 - “Costa Rica AND coupon* AND contraception”
 - “Costa Rica AND family planning AND coupon*”
 - “El Salvador AND voucher AND cervical”
 - “El Salvador AND “Edgardo Platero” AND cancer cervico-uterino”
 - “Ghana AND bednets AND voucher”
 - De Oliveira 2010
 - “NetMark AND bednets”
 - Indication of bednet programme in Nigeria - added to list
 - “NetMark AND bednets AND evaluation”
- [Hand searching of NetMark website due to use of vouchers for ITNs]. Found information on:
- Ghana
 - Nigeria
 - Senegal - survey might be applicable to review (NETMARK_2008)
 - Mali
 - Zambia
- “India AND Agra AND voucher*”
 - “India AND Agra AND voucher AND evaluation”

“Iran AND family planning AND coupon*”

- Treadway et al. 1976.

“Chiranjeevi scheme”

“Chiranjeevi Yojana AND evaluation”

“Chiranjeevi AND UNFPA AND evaluation”

- Found BHAT_INVOLVING
- Found ACHARA_ASSESSING
- Found JEGA_CONTRACTING

[Hand searched www.gujhealth.gov.in]

- Found UNFPA_RAPID

“India AND Kolkata slum AND voucher”

“India AND Calcutta AND voucher”

“India AND Kolkata AND CINIASHA”

“India AND Kolkata slum* AND reproduct*”

“Indonesia AND midwife* AND voucher”

- Medical news today article discussing voucher programme in Aceh post tsunami
- World Bank case study on vouchers in Pemalong district

“Indonesia AND midwife* AND TPC AND evaluation”

“Indonesia AND midwife* AND Aceh”

“Kenya AND vouchers AND matern*”

“Kenya AND vouchers AND evaluation”

“Kenya AND Nyeri AND Youth AND voucher”

“Korea AND voucher* AND reproduct*”

“Korea AND coupon* AND reproductive health”

Added Oct 30 “Mozambique AND voucher AND nets”

“Nicaragua AND voucher AND sex workers”

“Nicaragua AND voucher AND adolescents”

“Nicaragua AND voucher AND cervical”

“Senegal AND voucher* AND nets”

“Taiwan AND coupon* AND reproduct*”

“Taiwan AND coupon* AND IUD”

- Chow 1969
- Hermalin 1971

“Tanzania AND voucher* AND bednet*”

“Tanzania AND KINET AND eval*”

- Mulligan_costs

“Hati Pungozo”

- Marchant_Monitoring

“Uganda AND voucher* AND STI”

- BELLOWS_EVALUATION

“Uganda AND voucher AND matern*”

“Uganda AND voucher AND ‘safe delivery’”

“Uganda AND voucher* AND net*”

“Zambia AND voucher* AND net*”

“Zambia AND voucher AND ‘emergency contraception’”

Appendix 2.5: Key experts contacted to review obtained information

Micol Salvetto	Was responsible for developing some of the early voucher schemes for cervical cancer prevention in at least three different countries.
Peter Sandiford	Has conducted evaluations and cost-effectiveness assessments of competitive voucher schemes
Anna Gorter	Has implemented and evaluated many voucher programme in Central America and worldwide
Claus P. Janisch	Medical Advisor, KfW Banking Group (Funder of Voucher Programmes)
Amarjit Singh	Commissioner Health and Secretary Family Welfare, Government of Gujarat who implemented the Chiranjeevi Yojana Programme
Kara Hanson	Conducted a number of evaluations of the Tanzania ITN evaluations at London School of Hygiene and Tropical Medicine
K.R. Thankappan	Involved in demand-side financing implementation and evaluation in India
Darin Dorkin	Task Team Leader for the Health Results Innovation Trust Fund at the World Bank
Ben Bellows	Programme Manager of voucher programme implementation in 4 countries for Population Council

Appendix 2.6: Data Extraction Form

File Name:

Authors:

Publication Date:

Location of Voucher Programme:

Targeted Population:

Time Period of Programme:

Type of Voucher:

Funders:

Programme Management Agency:

Scale of Programme:

Study Design:

Study Time Period:

Outcomes:

Outcome Variables (for each category):

Direction of Effect:

Significance Level:

Other Details:

Notes:

Appendix 2.7: Coding tool, Confidence in Findings Assessment (CFA)

File name:	Dates of study period:
Population:	Intervention:
Comparison:	Outcomes:

- What are the main study objectives? Are they to evaluate the programme?
- What is the overall study design?
- Does the study design allow one to assess causation?
- Is the study design acceptable for synthesis according to Cochrane standards?
- Do the authors express confidence in the findings, or are they discussed as “preliminary”?

Selection

- Is the sampling methodology appropriate given the study objectives?
- Are there any concerns about bias with regard to the sampling methodology?
- Is the sample size sufficient to detect statistical differences?
- Does the sample appear representative of the broader population of interest?
- For survey-based studies, is the response rate sufficiently high or is there evidence that responders/ non-responders have similar characteristics?
- If randomisation was used, does it appear to have been done correctly?
- Are there any other concerns about selection?

Comparability

- What comparison was used (e.g. before and after, control districts, control facilities, individual controls, benchmarks, etc...)? Is this comparison appropriate given the study objectives?
- Does the study fail to control for any important confounding factors that could reasonably explain findings?
- For studies with control groups, were the criteria for defining controls appropriate?
- For studies with control groups, was the control group examined to ensure there were no cases or no leakage of the intervention?

- For studies with control groups, were the controls groups comparable in terms of demographic characteristics? If not, were findings controlled for these differences?
- If matching was used, did this process seem to have been carried out correctly?
- For before-and-after studies without controls, is the time period reasonable to observe a change due to the intervention?
- For before-and-after studies without controls, could the results be due to changes over time, regardless of the intervention?
- Are there any other concerns about comparability?

Measurement (Outcomes and Exposure)

- Do the measures (explanatory and outcome variables) seem appropriate given the study objectives?
- Have the measures been validated or used in previous studies?
- Are the measures subjective or tenuous? To what extent would a different observer or different day of observation change the results?
- When controls are used, was the observer/interviewer blind to intervention status of the observation?
- When controls are used, was the same methodology used to assess both treatment and control groups?
- How much data are missing? Could missing data introduce bias?
- If assumptions are made in the analyses, are these reasonable and/or based in the literature?
- If a sensitivity analysis is appropriate, was this conducted? Were important factors included?
- If the main findings indicate an effect, what is the statistical significance level?
- Are there any other concerns about measurement?

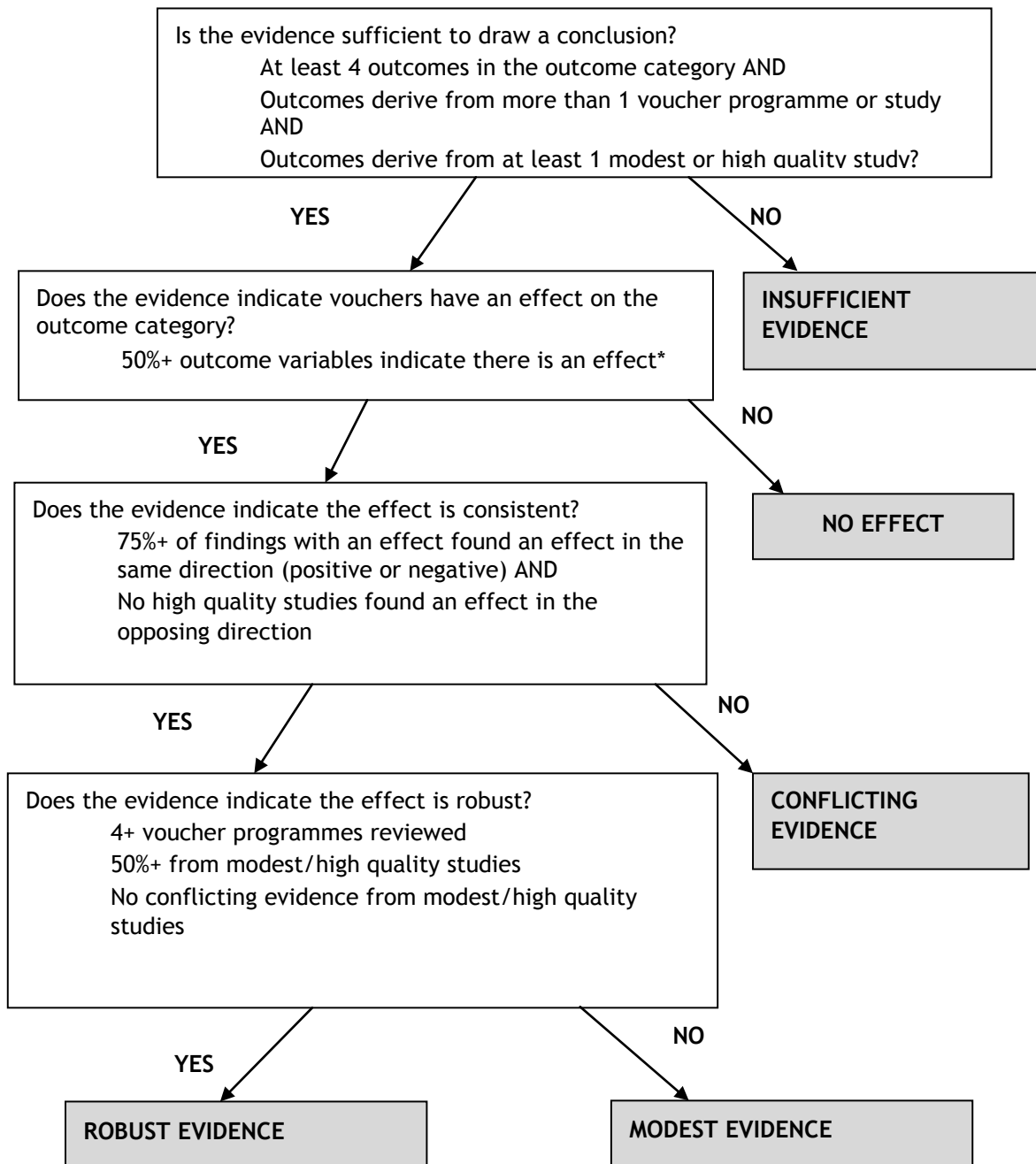
Final Considerations

- Is the study design appropriate given the study objectives?
- Are there any concerns about bias of the study in a particular direction?
- Are any confounding or distorting influences not accounted for?
- Is it likely that the findings occurred by chance?
- Are there any other major quality concerns not addressed in this assessment?

Overall Judgment - choose 1

- I have some MAJOR concerns about the methods used or the lack of information available on this study and therefore seriously question the findings. (*low confidence*)
- I have some MINOR concerns about the methods used or the lack of information available on this study and therefore would consider the findings with some caution. (*medium confidence*)
- I do not have any concerns about the methods used or information provided on this study. (*high confidence*)

Appendix 2.8: Decision tree for making conclusions regarding outcome categories



* See exception described in text.

Appendix 3.1: Voucher programmes with quantitative evaluation data included in the review

BANGLADESH MATERNITY
<p><i>Location:</i> Bangladesh</p> <p><i>Time period:</i> 2006 - present</p> <p><i>Type of health goods/or services:</i> Maternity services, transport, conditional cash transfer and gift box after delivery</p> <p><i>Description of the role of vouchers in the programme:</i></p> <p><i>Free or subsidised:</i> Free</p> <p><i>Targeted population:</i> Poor pregnant women with first or second pregnancy</p> <p><i>Targeting mechanism:</i> Two forms: 1. Geographic - all pregnant women in low-income areas are eligible. 2. Government field workers identify and distribute vouchers to eligible women, using means testing</p> <p><i>Role of public and private providers/retailers:</i> Public, NGO and private clinics contracted</p> <p><i>Scale of programme:</i> Administered in 33 sub-districts, covering approximately 7% of the population</p> <p><i>External funders:</i> World Bank, DFID, European Community, Germany, Sweden, Canada, Netherlands, and the United Nations Population Fund</p> <p><i>Management of programme:</i> Ministry of Health and Family Welfare of Bangladesh</p> <p><i>Evaluations:</i></p> <ul style="list-style-type: none"> • Schmidt et al. 2010; Cross-sectional analysis of data in intervention and control areas • Hatt et al. 2010; Cross-sectional analysis of intervention and control areas • Rahman et al. 2009; Before and after implementation study
CAMBODIA MATERNITY
<p><i>Location:</i> Cambodia, three rural districts in Kampong Cham (Cheung Prey, Prey Chhor and Chamkar Leu)</p> <p><i>Time period:</i> 2007 - present</p> <p><i>Type of health goods/services:</i> Maternity services and transportation</p>

Description of the role of vouchers in the programme: Voucher programme is complementary to established “Health Equity Funds” (HEF) which function at hospital level. Vouchers assist low-income individuals to access maternity services at health centre level and link them to HEF in case of complications. Qualifying pregnant women get further subsidy for free maternity services

Free or subsidised: Free

Targeted population: Poor pregnant women in rural districts

Targeting mechanism: Individual - pre-defined questionnaire answers determine eligibility, using same criteria as HEF. Local health volunteers or voucher programme managers identify qualifying women in the population

Role of public and private providers/retailers: Public health centres

Scale of programme: Voucher area includes three referral hospitals and 42 health centres, serving a total population of approximately 538,000

External funders: Belgium Technical Cooperation

Management of programme: Subcontracted to NGOs running the complementary Health Equity Funds programme

Evaluations:

- Ir et al. 2010 - before and after study with control areas

INDIA MATERNITY (MAMTA)

Location: Delhi, India

Time period: 2008-present

Programme Name: The MAMTA Scheme

Type of health goods/services: Maternity services and transport

Description of the role of voucher in the programme: Private hospitals (MAMTA-friendly hospitals) provide comprehensive package of maternal health services in underserved areas and are paid a fixed amount for each institutional delivery

Free or subsidised: Free

Targeting population: Poor pregnant women in Delhi

Targeting mechanism: 1) Geographic - resident of Delhi with residential proof and 2) those who are Below the Poverty Line (BPL)/Scheduled Caste/ Scheduled Tribe with a caste certificate or BPL card

Role of public and private providers/retailers: Private hospitals and health centres

Scale of programme: 36 private hospitals in Delhi - 4220 women delivered under the

scheme in the first year

External Funders: none specified, Government of India internal funders

Management of programme: Government of India

Evaluation:

- Nandan et al. 2010 - cross-sectional survey of beneficiaries and non-beneficiaries

INDIA MATERNITY (GUJARAT)

Location: India (Gujarat, pilot districts: Banaskantha, Dahod, Kutch, Panchmahals, Sabankantha, currently state-wide)

Time period: December 2005 - present

Programme name: Chiranjeevi Yojana Scheme

Type of health goods/services: maternity services (ANC, delivery, transport)

Description of the role of vouchers in the programme: Poor families can avail themselves of private maternity services - private providers are reimbursed by the government

Free or subsidised: Free

Targeted population: Pregnant women below the poverty line (BPL)

Targeting mechanism: Individual BPL status

Role of public and private providers/retailers: The government contracted with private providers

Scale of programme: Programme covered 18.6% of all deliveries in the pilot district Dahod).

External funders: State and Central Government funding

Management of programme: Government of Gujarat

Evaluations:

- Bhat et al. 2009 - cross sectional survey

MOZAMBIQUE ITNs

Location: Manica and Sofala provinces of Mozambique

Time period: 2005

Type of health goods/services: ITNs

Description of the role of vouchers in the programme: Vouchers for ITNs were combined with a vaccination campaign for measles and polio. Vouchers were distributed during the

second dose of the polio vaccine

Free or subsidised: Free

Targeted population: Young children

Targeting mechanism: Vouchers administered through vaccination campaign

Role of public and private providers/retailers: Fixed distribution posts in each district (not clear if public or private)

Scale of programme: Approximately 250,000 vouchers distributed and over 350,000 ITNs distributed after a policy change that did not require a voucher for getting an ITN

External funders: Canadian International Development Agency, Canadian Red Cross

Management of programme: Mozambique's Ministry of Health

Evaluations:

- De Oliveira et al. 2010 - Cross sectional survey analysis

NICARAGUA REPRODUCTIVE SERVICES TO ADOLESCENTS

Location: Nicaragua (Managua- capital city, Chinandega, Rivas)

Time period: 2000 - 2005

Type of health goods/services: Reproductive health services (counselling on sexual and reproductive health, contraception, STI testing and treatment, pregnancy testing, ANC)

Description of the role of vouchers in the programme: Vouchers are provided to adolescents for free reproductive health care clinics. Clinics receive reimbursement for each adolescent attending with a voucher

Free or subsidised: Free

Targeted population: Poor adolescents aged 12-20

Targeting mechanism: Geographic - vouchers distributed to adolescents outside schools, markets and in poor neighbourhoods,

Role of public and private providers/retailers: Public and private and NGO clinics contracted to provide voucher services

Scale of programme: Approximately 80,000 vouchers distributed from 2000 to 2005, around 15,000 used (11,000 by girls, 4,000 by boys)

External funders: DFID, USAID, Dutch government

Management of programme: NGO - Instituto CentroAmericano de la Salud (ICAS)

Evaluations:

- Meuwissen et al. 2006a - Simulated patients before, during, and after intervention, no controls
- Meuwissen et al. 2006b - Before and after intervention interviews with participating physicians, no controls
-
- Meuwissen et al. 2006c - Cross-sectional community sample of voucher receivers and non-receivers 3-15 months post intervention
- Meuwissen et al. 2006d - Cross-sectional survey of voucher users and non-users 3-15 months post intervention
- Meuwissen et al. 2006 - Cross-sectional analysis of medical files

NICARAGUA STI TREATMENT TO HIGH-RISK GROUPS

Location: Nicaragua (Managua-capital city, Chinandega, Leon, Masaya, Granada, Rivas)

Time period: 1995 - 2009

Type of health goods/services: STI services (testing for STIs, HIV and cervical cancer, health education, condoms, treatment for STIs)

Description of the role of vouchers in the programme: Vouchers distributed to targeted population and used for free STI testing and treatment at contracted clinics. Vouchers are good for three months and return vouchers are given to those testing positive for STIs and pregnant women

Free or subsidised: Free

Targeted population: High-risk groups for sexual infections, particularly female and male sex workers, regular clients, transvestites, and glue sniffers

Targeting mechanism: Geographic - vouchers distributed in areas where high-risk populations reside

Role of public and private providers/retailers: Public, private and NGO clinics

Scale of programme: From 1996-2009, over 60,000 vouchers distributed in Managua, provided over 22,000 consultations, and treated more than 6,000 STIs

External funders: DFID, USAID, Dutch government, several NGOs (NOVIB, Elton John Aids Foundation, ICCO) and Global Fund to fight AIDS, TB and Malaria

Management of programme: NGO - Instituto CentroAmericano de la Salud (ICAS)

Evaluations:

- Borghi et al. 2005 - Cross-sectional cost analysis of STI treatment
- McKay et al. 2006 - Interrupted time series of STI infections

NICARAGUA CERVICAL CANCER SCREENING

Location: Nicaragua (Siuna, Mulukukú, Ocotol, Jalapa, Somoto, Estelí, Matagalpa, Jinotega, Juigalpa, Rivas, Granada, Masaya, Chinandega, Leon, Boaco, Sebaco, San Marco,

Ticuantepe, Tipitapa, Carazo, and Managua)

Time period: 2003 - 2009

Programme name: The Ginecobono Programme

Type of health goods/services: Cervical cancer screening

Description of the role of vouchers in the programme: The programme distributed donated vouchers to women at risk of cervical cancer throughout the country (with emphasis on older women in remote rural communities) covering Pap tests and any necessary diagnostics and treatment. Vouchers are delivered to women where they live or interact

Free or subsidised: Free

Targeted population: Women at risk of cervical cancer with low access to Pap screenings.

Targeting mechanism: Geographical (with emphasis on remote rural areas identified by the Ministry of Health)

Role of public and private providers/retailers: Vouchers redeemed at public and private (for profit and not-for-profit) facilities

Scale of programme: The programme has screened 38,564 women in Nicaragua from all over the country

External funders: Foundation Gurdían Ortiz, Dutch government and DFID

Management of programme: Instituto CentroAmericano de la Salud (ICAS)

Evaluations:

- Howe et al. 2005 - Clinical records review

NIGER LLINs

Location: Niger

Time period: 2005-2008

Type of health goods/services: LLINs

Description of the role of vouchers in the programme: The second stage of a mass distribution campaign (December 2005) integrated polio vaccination, vitamin A administration and LLIN distribution in the seven regions of Niger. Mobile teams provided polio vaccinations and vitamin A to children at their homes. Any mother or caretaker who had one or more children vaccinated was given a single voucher and / or nail marking for later exchange for an LLIN at a distribution post

Free or subsidised: Free

Targeted population: Children under five years old

Targeting mechanism: Geographical (along with a vaccination campaign)

Role of public and private providers/retailers: Vouchers are redeemed at LLIN distribution posts

Scale of programme: Two million LLINs were distributed

External funders: World Health Organisation, International Federation of Red Cross and Red Crescent Societies, Canadian Red Cross, and Rotary International

Management of programme: Ministry of Health

Evaluations:

- Thwing et al. 2008 - before and after survey

SENEGAL LLINs

Location: Senegal (Thiès, Louga, Kaolack, Ziguinchor, Kolda, Dakar, and Fatick)

Time period: 2005-2009

Type of health goods/services: LLINs

Description of the role of vouchers in the programme: As part of a broader programme by NetMark aimed at increasing supply and demand, vouchers were implemented for pregnant women and children, starting in 2005

Free or subsidised: Free

Targeted population: pregnant women and children under five years

Targeting mechanism: Individual (delivered door to door to pregnant women and mothers of children under five)

Role of public and private providers/retailers: Vouchers are redeemed at private retail outlets, pharmacies and supermarkets as well as district health committees' pharmacies.

Scale of programme: 2.9 million LLINs delivered as of 2009 using the voucher scheme.

External funders: USAID

Management of programme: NetMark

Evaluations:

- NetMark 2008 - 2008 survey with additional survey data from 2000 and 2004.

TAIWAN IUDs

Location: Taiwan, island-wide

Time period: 1964 - 1970s

Programme name: Taiwan Population Studies Center Coupon System

Type of health goods/services: IUD insertion (later also sterilisations)

Description of the role of vouchers in the programme: Coupons are distributed by a pre-pregnancy health field worker who visits women in their townships or at home

Free or subsidised: Subsidised

Targeted population: Pre-pregnancy women and women interested in birth spacing or limiting

Targeting mechanism: Individual (identified by field workers)

Role of public and private providers/retailers: Public and private providers

Scale of programme: Island-wide (10,000 IUD acceptors per month by 1967)

External funders: USAID (first years), Taiwan Ministry of Health

Management of programme: Maternal and Child Health Association (in 1969 renamed to Planned Parenthood Association of China), in collaboration with the Committee on Family Planning and the Taiwan Population Studies Center (last two merged in 1969 to the Institute of Family Planning)

Evaluations:

- Chang et al. 1968 - matched case control

TANZANIA ITNs (KINET)

Location: Kilombero and Ulanga districts of Tanzania

Time period: 1996 - 2003

Type of health goods/services: ITNs

Description of the role of vouchers in the programme: Vouchers are one component of a larger social marketing programme designed to strengthen the supply and demand of ITNs in Tanzania through strengthening distribution systems, subsidising the purchase price, and marketing the ITNs. Within this programme, vouchers are used to further reduce costs to pregnant women by offering a discount of approximately 20% to the already subsidised price

Free or subsidised: Subsidised

Targeted population: Pregnant women and young children

Targeting mechanism: Health facilities - distributed to all pregnant attending an antenatal clinic

Role of public and private providers/retailers: Vouchers are used at private retailers

Scale of programme: Pilot project in two rural districts of southern Tanzania; 7344 vouchers were redeemed that had been issued between January 2000 and March 2003.

External funders: Swiss Agency for Development and Co-operation

Management of programme: Implemented by the Ifakara Health Research and Development Centre (IHRDC) and the Swiss Tropical Institute

Evaluations:

- Kikumbih et al. 2005 - Market analysis of treatment and control districts

TANZANIA ITNs (NATIONAL)

Location: Tanzania

Time period: 2003-present

Type of health goods/services: ITNs

Description of the role of vouchers in the programme: The KINET programme discussed above was adapted and expanded nationally. Vouchers are one component of a larger programme designed to strengthen the supply and demand of ITNs in Tanzania through strengthening distribution systems, subsidising the purchase price, and conducting social marketing. Within this programme, vouchers are used to further reduce costs to pregnant women by offering a discount of approximately 20% to the already subsidised price

Free or subsidised: Subsidised

Targeted population: Pregnant women and young children

Targeting mechanism: Health facilities - distributed to all pregnant women attending an antenatal clinic

Role of public and private providers/retailers: Vouchers are used at private retailers

Scale of programme: Implemented nationwide, representing approximately 5% of the Ministry of Health's total budget; approximately 2.5 million treated net years

External funders: Global Fund to fight AIDS, Tuberculosis and Malaria, President's Malaria Initiative

Management of programme: National Malaria Control Programme

Evaluations:

- Hanson et al. 2009 - Before and after study of ITN use
- Khatib et al. 2008 - Cross-sectional analysis of voucher-obtained ITNs compared to ITNs obtained through other means
- Marchant et al. 2008 - Multiple survey (4 data points) analysis

UGANDA STI SERVICES

Location: Uganda (Mbarara, Kirahura, Ibanda, Isingiro and Bushyeni)

Time period: July 2006 - present

Type of health goods/services: STI testing and treatment

Description of the role of vouchers in the programme: Vouchers are sold by community distributors and redeemed at accredited facilities

Free or subsidised: Subsidised

Targeted population: Poor people who are exposed to or suffering from an STI

Targeting mechanism: Geographical targeting in sub counties with poverty levels over 35%

Role of public and private providers/retailers: Private providers are contracted

Scale of programme: Six districts in southwest Uganda

External funders: KfW and GPOBA (World Bank)

Management of programme: Marie Stopes International - Uganda

Evaluations:

- Bellows et al. 2009 - Before and after study with treatment and control districts

ZAMBIA HEALTH SERVICES

Location: Zambia (Lusaka and Kitwe)

Time period: 1991 - 2000

Programme name: The Discount Card Scheme

Type of health goods/services: Basic Health Services

Description of the role of vouchers in the programme: The discount cards were introduced as a way to address some of the problems with the national pre-purchase scheme

Free or subsidised: Subsidised

Targeted population: Urban and rural poor

Targeting mechanism: Geographical

Role of public and private providers/retailers: Vouchers are used in lieu of user fees at public facilities

Scale of programme: Two districts

External funders: Does not specify

Management of programme: Government of Zambia

Evaluations:

- *Kondo and McPake 2007 - economic modelling*

ZAMBIA ITNs

Location: Rural and urban district of Zambia (Chilubi, Kaputa, Mambew, Nyimba and Kalalushi)

Time period: June 2003

Type of health goods/services: ITNs

Description of the role of vouchers in the programme: Vouchers combined with measles vaccination campaign for free bed net and insecticide to treat the net. Vouchers distributed at time of vaccination

Free or subsidised: Free

Targeted population: Children under five years of age

Targeting mechanism: Distributed during a measles vaccination campaign

Role of public and private providers/retailers: Private retailers and one public clinic used for distribution

Scale of programme: Districts has approximately 450,000 persons with 81,000 children under five

External funders: International Federation of Red Cross and Red Crescent Societies

Management of programme: National Malaria Control Centre

Evaluations:

- Grabowsky et al. 2005 - Cross-sectional survey with before and after questions

This document is available in a range of accessible formats including large print.

Please contact the Institute of Education for assistance:

telephone: +44 (0)20 7947 9556

email: info@ioe.ac.uk

This material has been funded by the Department for International Development. However the views expressed do not necessarily reflect the department's official policies. The report was first published in 2011 by:

Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre)
Social Science Research Unit
Institute of Education, University of London
18 Woburn Square
London WC1H 0NR

Tel: +44 (0)20 7612 6397
<http://eppi.ioe.ac.uk>
<http://www.ioe.ac.uk/ssru>

ISBN: 978-1-907345-10-4

The authors of this report were supported by the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre).

The Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) is part of the Social Science Research Unit (SSRU), Institute of Education, University of London. Since 1993, we have been at the forefront of carrying out systematic reviews and developing review methods in social science and public policy. We are dedicated to making reliable research findings accessible to the people who need them, whether they are making policy, practice or personal decisions. We engage health and education policy makers, practitioners and service users in discussions about how researchers can make their work more relevant and how to use research findings.

Founded in 1990, the Social Science Research Unit (SSRU) is based at the Institute of Education, University of London. Our mission is to engage in and otherwise promote rigorous, ethical and participative social research as well as to support evidence-informed public policy and practice across a range of domains including education, health and welfare, guided by a concern for human rights, social justice and the development of human potential.

The views expressed in this work are those of the authors and do not necessarily reflect the views of the EPPI-Centre or the funder. All errors and omissions remain those of the authors.