Chapter 2. Methods

Solicitation of Input and Data

We began the review process conferencing with AHRQ to clarify the scope of the project and other background information. Six experts also agreed to serve as members of a technical expert panel group (TEP, see Appendix A). The comments and suggestions provided by the TEP were helpful in clarifying the conceptual framework used for the project. The research strategy consulting staff at the University of Minnesota Biomedical Library was invaluable in developing the search strategies, which the TEP members also reviewed.

Literature Search

We conducted the searches for both the provider and the consumer incentives simultaneously, as most search terms were in common to both. We identified MEDLINE®, the Cochrane Library, EconLit, Business Source Premier, and PsychInfo as the literature sources for this review. We searched MEDLINE® for relevant articles published between 1966 and 2002 using the strategy shown in Appendix B. Results of this search, transferred to an EndNote database and cleaned of nine duplicate records, identified 306 articles. PsychInfo, EconLit and Business Source Premier, were approached with a very simple strategy of combining keyword searches for "incent$" and "health". The results of these keyword searches added an additional 76 articles to the database for a total of 382 articles.

The 382 articles were subjected to a review of the abstracts. For the provider incentive review, experimental, quasi-experimental, including observational with controls, and simple pre-post designs were included in this stage of the screening to avoid premature loss of potentially useful information. From this number, eight articles for provider incentives were pulled for full text review. Given the sparsity of the literature, we also conducted a title scan of the 2,834 entries tagged merely as "journal articles" from the original MEDLINE® search results to identify potentially relevant articles. This resulted in an additional eight articles for full text review.

Reference lists from previous systematic reviews,[26, 28-37] including the Cochrane Library, were examined as well, resulting in further four articles. In all, including six articles culled from reference lists of articles pulled for full text review, 26 articles were identified for possible inclusion in the structured review for provider incentives.

For the consumer incentives, 50 articles were identified from MEDLINE®, six from PsychInfo, 11 were pulled from review articles. An additional 17 were culled from reference lists of identified studies pulled for review. In total, 84 articles were identified for possible inclusion in the structured review for consumer incentives.

Articles for both provider and consumer incentives were subjected to the inclusion and exclusion criteria using the screening tool found in Appendix C.

Note: Appendixes and Evidence Tables cited in this report are provided electronically at http://www.ahrq.gov/clinic/epcindex.htm

Inclusion Criteria

- Published between 1966 and Oct 2002.
- Address primary preventive care defined as vaccination, screening, and health promotion behaviors such as smoking cessation and weight loss.
- Be a primary study.
- Take place in an industrialized country.
- Be written in English.
- Examine preventive care as at least one primary outcome.

Exclusion Criteria

- Patient adherence to drug therapy. For example, the treatment of tuberculosis, which has been the subject of many adherence studies and could be considered a preventive step because it prevents the spread of the disease, is excluded. For our purposes, primary prevention is defined as occurring pre-diagnosis.
- Financial rewards for participating in a research study. The economic incentive must be conceived as part of the intervention.
- Multiple component interventions in which the economic incentive is only one component and the study design precluded analyzing the effect of the incentive separate from the other components. Education and outreach efforts are examples of components in addition to economic incentives. This was most strongly noted with community and education-based prevention programs and worksite health promotion programs.
- Studies examining payment forms provided by more than one payment system, i.e., HMO vs. FFS were excluded because there are too many potential confounding factors. It was therefore too difficult to isolate the effect of specific economic incentives.

Selection bias is a major concern for outcomes research on economic incentives. The opportunity for self-selection of consumers, in particular, is high, given all of the unmeasured economic and contextual issues that may come into play for any specific economic incentive intervention. A strong research design including randomization greatly minimizes this concern. Therefore, we included only RCTs, time series, and prospective quasi-experimental designs for the structured literature reviews. However, we also provided information from relatively well-designed econometric cross-sectional studies for the provider incentives as another perspective for consideration.

A list of the identified articles was distributed to the TEP members and other colleagues, asking for whether they were aware of other relevant studies that were not represented. No

further articles were identified from this query. Figure 2 provides a tree diagram of included and excluded references.

Abstraction

A single data abstraction form for both the provider and consumer incentive reviews was devised during the initial stages of the literature search. Formal meta-analysis of the incentive literature was not possible because there were not a large number of studies that examined the same incentive type, research outcome measures, and similar populations. The abstraction tool was created with the purpose of facilitating the ability to capture emergent themes from the heterogeneous literature. The form was reviewed and commented on by the TEP members, piloted, and subsequently revised. Abstraction of the articles was performed by two independent reviewers. Disagreements were resolved by consensus of the group. See Appendix D for the full abstract forms.

Emergent Themes

Given the lack of consensus and implied, rather than specifically reported, concepts and assumptions underlying explicit economic incentives, we characterized the incentives based on what appeared to emerge as the predominant, if implicit, conceptualizations within the identified studies. We used three different approaches for this process: psychological, economic, and functional views. These characterizations, imposed upon the studies as they are, do not fit perfectly and certainly there is room for debate as to which category each is slotted into.

The psychological approach established four major incentive categories of reward, negative reinforcement and/or punishment, structural barrier removal, and attitudinal barrier removal. The approach was informed by several theories and frameworks that are standard fare in psychological or health services research literature and were directly cited or implied in at least one study article.

One set of psychological theories are the non-motivational operant and information-processing theories. Basic operant or behavioral theory implies a causal model of Stimulus → Organism → Response (SOR) which essentially states an external stimulus acts upon the organism, or human, through reinforcements, and the human responds with a behavior. In the operant view, people respond to stimuli because they were reinforced for responding to those stimuli in the past. Cognitive models introduced into the SOR models a focus on expectations of future reinforcements or rewards, and people's attributions about why they engaged in certain behaviors in the past.

Many of the health psychology theories applied in health services research, such as social learning theory and the Health Belief Model, were outgrowths of SOR models.[38, 39] The incentive categories of rewards and negative reinforcement/punishments, then, are incentives or disincentives that act as consequences, reinforcements, or behavioral goals.

A second set of psychological theories views the causal model as Organism → Stimulus → Organism → Response (OSOR). Here, the human is seen as selecting and interpreting stimuli or information based on their individual drives, needs, and orientations. The environment and events surrounding the human are seen as ambiguous enough that a person can apply considerable discretion as to the salience placed on, and interpretation of, these events.

Frame's framework of barriers to the practice of preventive care and health promoting behavior could be viewed as limited attempts to introduce this broader view into the discussion of contributions to preventive health behavior.[40] From this perspective, one might apply an incentive as a device to attract attention to what is believed an appropriate stimulus, such as the healthy effects of exercise, increasing the salience of the stimulus, an attitudinal barrier removal.

Several studies also clearly stated the incentive was intended to remove a specific economic structural barrier, such as transportation costs, or lower the cost of a health care service. The majority of structural barrier removals are found within simple preventive care, seven of ten, and both punishment disincentives. Rewards were well represented within both simple (58 percent) and complex (78 percent) preventive care categories.

The economic approach established two major categories based on whether the incentive was triggering a purchasing behavior or income-generating behavior on the part of the participant. "Purchasing behavior" was defined as behavior characterized by a money-saving orientation. In such cases, the participant would realize the benefit of the incentive only through a "purchase" of the targeted behavior or service. Examples would include free postage if a package was mailed, free flu shots if the participant showed at the clinic for immunization, or free nicotine patches if the participant filled the prescription.

"Income-generating behavior" was defined as behavior that leads to a payment of some kind (cash or in-kind transfer). That is, how could the participant maximize the potential money in his or her pocket for discretionary use, either in cash or goods. Examples of this approach would include cash, vouchers and coupons, proceeds from a lottery, and gifts. We note all of the mixed intervention studies were of this category. We found a relatively clean correlation between structural barrier removal and purchasing behavior, and rewards/punishment and income-generating behavior.

The functional approach established categories based on what appeared to be the problem the researchers were intending to address. We identified four major types: facilitating, participating, adhering, and outcome.

"Facilitating incentives" included incentives aimed at making it easier for the consumer to engage in the behavior. This category includes reducing the costs of specific preventive medical services or necessary steps in the process to seek or complete such care.

"Participating incentives" included those incentives aimed at increasing the rate at which consumers participated in a desirable behavior. These incentives attempted to draw people into the specific preventive care or health promotion process in question, to join in, and would include such situations as enticing people to attend educational sessions. It is generally assumed participation in education or other experiential programs that increase understanding of and exposure to a desired behavior will lead to downstream effects of adapting positive health and lifestyle changes. The measure of the effectiveness of the incentives would be reflected in direct measures of the targeted behavior.

"Adhering incentives" are incentives aimed at increasing the rate at which consumers adhered to the specific preventive care guidelines or change in health or lifestyle behaviors. In the case of complex preventive care behaviors, the incentives may be applied to reinforce a behavior until such time as a person's internal motivation is activated and sufficient to sustain the behavior. Direct measures of adhering incentives would then focus on the maintenance of the target behavior.

"Outcome incentives" are incentives that promote achievement of a particular outcome or goal. As both simple and complex preventive care concerns may involve possible multiple paths

to achieve the target goal, the researcher would need to take care to assess whether the desired behaviors are actually increased or whether the participant had employed unintended, or even potentially unhealthy, behavior to gain the incentive. For example, a person may choose to attain weight loss through healthy means of exercise and moderate calorie intake or through behaviors such as skipping meals or extreme calorie reduction—behaviors which undermine weight loss in the long run.

Rating Strength of the Evidence

Given the presence of both experimental and quasi-experimental designs within the relevant literature, the checklist for assessing methodological quality devised by Downs and Black[41] and recommended by the EPC Technology Report #47, "Strength of the Evidence" was chosen for this project.[42] After assessing the checklist for each article, a grade of 1 (poor) to 4 (excellent) was assigned to each article. A minimum level of study validity was assured in part by the selection criteria listed above.

It would be good to remember the social science experiments involving economic incentives differ from clinical trials in that it is not possible to mask the intervention from the patient or provider, nor always blind the researchers to the participant's group assignment. Empirical research into the critical components of social science experimentation is not at the level of clinical trials.[42] There is, therefore, necessarily subjectivity and expert opinion involved in the grading process.

In addition, this review is at the edge of evidence-based practice in that it includes econometric studies of the type usually performed to guide policy. Study quality and strength of evidence literature for evidence-based reviews has not addressed econometric research. Thus, we created an assessment tool to assist in this review. (See Appendix E.) The application of the assessment tool was performed by two independent reviewers. Disagreements were resolved by consensus of the group.

Econometric Research

Studies that rely on non-randomized trials (we will refer to these as observational studies) can teach us a great deal about health care provider and consumer behavior.[b] In fact, data from observational studies in which subject are not randomized into "treatment" categories form the foundation of the empirical body of knowledge for both health economics and medical sociology. However, because the subjects are not randomized, the data analysis often requires special statistical attention to address potential biases induced by the nonrandomization. If the data meet certain criteria and appropriate statistical methods are applied given the nature of the data, observational studies can yield inferences that are equally valid (and in some circumstances the inferences are more easily generalized to real world practice) than those inferences made from a randomized trial. While observational studies can yield valid inferences if the appropriate statistical methods are applied, often those methods are not applied, or there are no methods that can correct for the biases given the nature of the data, and the resulting inferences are incorrect. Thus, in assessing the validity of the conclusions from a body of observational studies, careful

[b] In our experience, it is rare that a so-called randomized clinical trial is truly randomized. While subjects may be successfully randomized into different treatment arms, subject attrition often unravels the efficacy of the initial randomization.

consideration of the nature of the data and the appropriateness of the statistical methods is necessary. In order to provide a systematic method for assessing the strength of the evidence for literatures in which observational studies comprise an important component of the body of work, we developed an instrument to assess the validity of a given study's conclusions. This algorithm is provided in Appendix E.

The instrument attempts to jointly assess the quality of the data used in the analysis and, given the nature of the data, the appropriateness of the statistical methods applied to the data. Specifically, the instrument places studies into one of three categories: Very Informative, Informative, and Not Informative. The underlying logic of the instrument is that the greatest risk in making biased inferences from observational studies is from poor identification of the variables of interest from the lack of appropriate exogenous variation in the variables of interest. Thus, the instrument measures two categories of characteristics: ones that may bias the coefficient estimates, and dimensions of the study that affect the generalizability of the findings.

In order to place a given study in one of these three categories the algorithm first assesses if there is the appropriate variation in the data to measure the coefficients of interest. Then, the instrument assesses the nature of the variation in the data. That is, observational studies are at risk for "endogenous" variation in the variables of interest. An endogenous variable is one in which there is correlation between the error term and the right hand side variables. That is, there is some unobserved factor that impacts both the outcome of interest and the intervention thereby confounding inferences regarding the independent impact of the intervention on the outcome. The instrument ascertains whether the data generating process can be viewed as exogenous of the unobserved factors. It is often the case in observational data the variables of interest are endogenous and the study uses an appropriate statistical method to correct for any bias that would result from Ordinary Least Squares analysis. In a given study, if data is likely endogenous, then the instrument asks if appropriate statistical methods have been employed to correct for this potential bias. For a study to be categorized as 'Very Informative' it must have both appropriate variation in the data and either exogenous variation in the data or employ an appropriate statistical method to correct for the potential bias (e.g. instrumental variables with appropriate instruments). In addition, for the study to score in the 'Very Informative' category it must also score well in the other study criteria.

The other criteria attempt to measure the generalizability of the findings to populations outside of the sample population analyzed in the study and the appropriateness of the statistical methods.

Peer Review Process

Several individual experts in the fields of health economics and preventive medicine independently reviewed an earlier draft of this report. They are acknowledged in Appendix A. In addition we asked our TEP members to review the draft report. We revised the report in response to the suggestions of all these individuals.
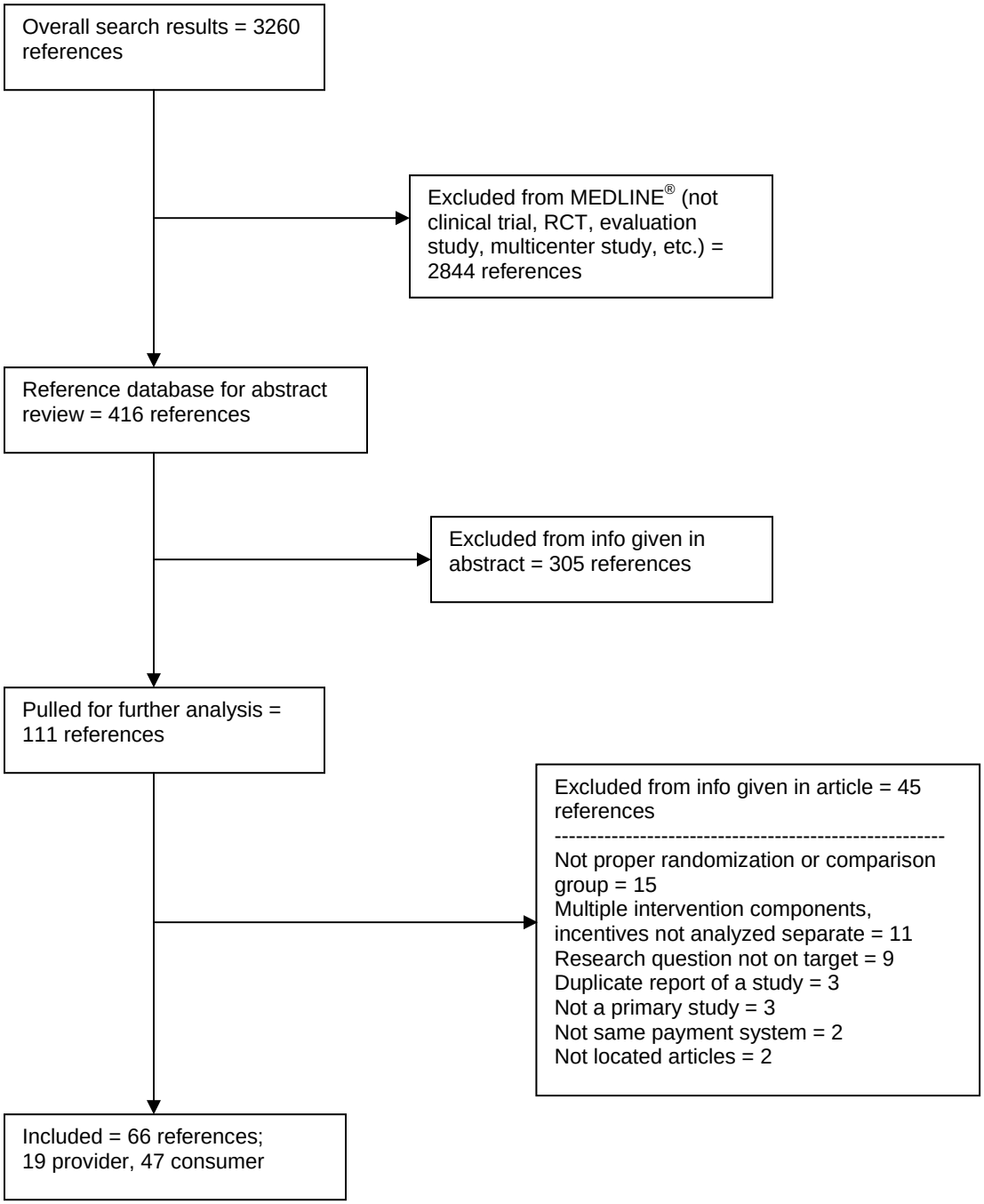
Figure 2. Economic incentives search result flow chart