IMPLEMENTING PREVENTIVE SERVICES: TO WHAT EXTENT CAN WE CHANGE PROVIDER PERFORMANCE IN AMBULATORY CARE? A REVIEW OF THE SCREENING, IMMUNIZATION, AND COUNSELING LITERATURE

Lynda A. Anderson, Ph.D. and Gail R. Janes, Ph.D.
Centers for Disease Control and Prevention

Carolyn Jenkins, M.S.N., Dr.P.H.
Medical University of South Carolina

ABSTRACT

Strategies to improve the delivery of preventive care often consist of office-based interventions, which are designed to modify provider behaviors or practice patterns. We report on a meta-analysis of 117 behavioral outcomes extracted from 43 studies. Meta-analytic techniques were used to express the results in a common metric, which allowed quantitative comparisons across outcomes. Studies were examined by domains of preventive care (screening, immunization, and counseling) and divided into two groups based on unit of analysis (provider or patient categories). The mean effect size reflects the difference in proportion of physicians providing the targeted behavior between the experimental and control groups. In the provider category, the weighted mean effect size for screening was .14, for immunization was .18, and for counseling was .28. In the patient category, the weighted means for screening and immunization were .12 and .15, respectively, but were smaller for the counseling (.08). Because tests for homogeneity of effect sizes were rejected in the patient category, caution in interpreting mean effect sizes is warranted because of variability across individual values. In summary, office-based interventions were found to have positive effects on providers’ adherence to preventive recommendations. We discuss the methodological issues and needs for future work to enhance the delivery of preventive services.


INTRODUCTION

Many studies document that the delivery of preventive services in ambulatory care falls significantly below recommended guidelines for care (1-3). Calls to improve the delivery of preventive services have escalated in recent years, driven in part by increasing interests in controlling health care costs (4,5). Health Plan Employer Data and Information Set (HEDIS) measures, used to assess managed care systems, emphasize clinical preventive services such as breast and cervical cancer screening in establishing health care performance standards (6). The underlying rationale for this shift in emphasis is that prevention is cost-effective, especially when long-term costs of disease and disability are considered.

Strategies to improve the delivery of preventive care frequently focus on the health care provider. Office-based interventions have been studied since the late 1970s. Office-based interventions refer to strategies, such as reminders and computer monitoring systems, directed at changing or modifying provider behaviors or practice (7). These strategies are distinct from continuing medical education, such as print materials or didactic courses, that try to alter behavior by changing knowledge and attitudes (8). Davis and colleagues (9) found evidence that short continuing medical education courses have little direct impact on changing practice but that physician-directed interventions such as reminders may result in positive changes when used alone. That study did not examine the extent to which these strategies changed provider behavior but rather counted the total number of studies that showed statistically significant differences between intervention and comparison groups.

An important and practical means of providing quantitative information about office-based strategies is to apply meta-analytic techniques to the available literature (10). Meta-analysis not only assembles and summarizes the relevant literature, as in a traditional review, but also expresses the results of each study in a common metric, which allows quantitative comparison across studies, as well as construction of a single summary measure of effect across similar studies. Meta-analysis of office-based interventions is quite recent. Austin et al. (11), for example, conducted a meta-analysis of four studies assessing the impact of various interventions on the delivery of immunization. Snell and Buck (12) examined patient- and provider-directed interventions to improve the delivery of cancer screening. However, many such studies are limited by scant information about their search strategy or they do not account for methodological issues such as unit of analysis differences across studies. Recognizing the need to determine the extent of change that can be expected from office-based interventions and the need to assess these effects across different types of preventive outcomes, we performed a meta-analysis of office-based interventions to determine their impact on three domains of preventive care: screening, immunization, and counseling.

MATERIALS AND METHODS

The research proceeded in three stages: (a) development of criteria for inclusion and literature search; (b) abstraction and coding of study characteristics and findings; and (c) data analysis and aggregation of findings, as appropriate.

Inclusion Criteria and Literature Search

Our search of relevant published studies was restricted to those with a replicable description of an office-based intervention.
We selected a study for inclusion if it: (a) was published in English, (b) included a control or comparison group, (c) evaluated an intervention in an ambulatory care setting, (d) involved a research setting within the United States (to avoid possible confounding by health care system differences), and (e) provided an objective assessment of the provider’s adherence to preventive services. Medical chart reviews and patients’ reports of providers’ performance were included, whereas self-reports by providers of their own performance were excluded from this review.

We obtained articles from database searches of AIDSLINE, CANCERLIT, CINAHL, ERIC, HealthSTAR, HEALTH & WELLNESS (formerly known as HEALTH PERIODICALS), Mantis (formally known as CHIROLARS), MEDLINE, NAIL, NTIS, and PsycINFO. Major search terms included more than 50 categories and phrases (e.g. preventive health services, prompting, cuing, reminders, flowsheet, prescription pads). Combinations of key words were used, and no exclusions were made for year of publication; the search period went through January 1, 1997. Additional sources included reference lists from published studies and personal contacts with several authors. A list of major search terms is available from the authors. To be included in the review, a study was required to meet all inclusion criteria. If the original two raters were uncertain or disagreed about the study’s eligibility, a third rater independently rated the article’s eligibility and consensus was obtained among the three raters.

We identified and reviewed 739 citations that dealt with some aspect of provider behavior. If no abstract was available, we retrieved and reviewed the original article. We excluded 567 articles that were not published in English or did not involve an intervention with a control or comparison group. Ninety-four of the remaining articles were eliminated because they did not evaluate an office-based intervention in an ambulatory care setting or were set outside the United States. Of the 78 remaining articles, 36 were eliminated because they did not provide an objective assessment of the providers’ adherence to preventive services (i.e. self-reported performance) or did not study preventive services that could be individually identified (e.g. use of a composite performance score in which individual prevention behaviors could not be isolated and assessed). Finally, one article duplicated the findings of another and was excluded from this review. Thus, a total of 41 articles met all of the inclusion criteria.

Classification of Studies

The first and third authors independently reviewed the 41 articles, which represented 43 separate studies. Using a standard form, the reviewers classified the studies by preventive care area, randomization procedures, research site, type and number of providers and patients (study population), treatment intervention, type of control or comparison group, intervention implementation, assessment period, and unit of analysis.

Treatment interventions targeted outcomes in the areas of screening, immunization, or counseling, as defined by the U.S. Preventive Services Task Force (3). Screening is defined as a special test (e.g. pap and fecal occult blood tests) or a standardized examination procedure (e.g. clinical breast and rectal examinations) used to identify patients requiring special intervention (3). Immunizations include vaccines against childhood and adult diseases (3). Counseling is defined as provision of information and advice concerning behavior that affects the risk of subsequent illness or injury (e.g. nutrition, physical activity, safe sex practices, and smoking cessation) (3). For this study, the unit of analysis was the individual preventive behaviors (i.e. behavioral outcomes) rather than the individual study. A total of 117 behavioral outcomes are presented in this review (13-53).

We coded interventions by the type of information provided to alter the delivery of care rather than by the instruments used to convey the information (e.g. flags, tags, stickers, computerized prompts). Based on criteria suggested in the literature (54), interventions were classified hierarchically into one of four categories: (a) feedback, (b) prompting, (c) prompting and monitoring, and (d) combined interventions. Feedback refers to information given to providers about how their practice patterns or patient outcomes compare with those of peers, their prior performance, or an external standard (8). Prompting refers to information given to providers about a desired action concurrent with a patient encounter or event, targeting a specific clinical decision point. Prompting and monitoring refers to ongoing or recent event information concerning the delivery of a service. When a study included more than two interventions, the most theoretically powerful intervention (e.g. prompting and monitoring versus feedback alone) was selected to compare to the control/comparison group. Using the chance-corrected κ coefficient among a randomly selected sample of 15 articles, interrater agreement for intervention classification was good; κ was .82 (1 indicates complete agreement) (55). All data reported herein were established by consensus of the authors.

Measurement and Analytical Techniques

To synthesize findings, it was necessary to convert the various summary statistics for each outcome into a single common metric or effect size, which assesses the strength of the relationship of interest. To ensure consistency of approach, we analyzed simple comparisons of outcome for all studies, unadjusted for potential confounders regardless of the availability of tests of adjusted effect.

Some studies used the provider as the unit of analysis whereas others used the patient as the unit of analysis. For methodologic reasons, we examined these two sets of studies separately. When the patient was the unit of analysis, the Pearson product–moment correlation coefficient (r) was the calculated measure of effect, as recommended by Rosenthal (56). When the provider was the unit of analysis, the patient level data on adherence in the experimental and control groups often were not available. Instead, the average adherence rate per provider was most frequently reported as the outcome. Average adherence rate was calculated as the percentage of each provider’s patients who were both eligible for and received the targeted procedure (e.g. screening). Effect size was calculated as d’, the difference between average adherence rates in the intervention and control groups. To combine estimates of effect across multiple independent outcomes, a common or adjusted (hereafter referred to as weighted) mean estimate of effect was calculated for each intervention area by unit of analysis. Each effect size was weighted by its estimated variance (57). The weighted mean of the correlation coefficient was calculated using the Fisher z transformation of r, to correct for non-normality. A χ² test for heterogeneity of effect size was performed for each subgrouping, and if it was non-significant, the mean estimate of effect was considered valid (57).

Rosenthal (56) notes that r and d’ are equivalent when the sample sizes are of equal size and the proportions in the control and the experimental groups are the same amount above and below .50. Furthermore, he notes that r (and d’) can be interpreted as the difference in success rate between experimental and control groups (e.g. an intervention that increased an adherence rate from .45 to .55 would have an effect size of .10). Cohen classifies effect size