HHS Public Access

Author manuscript

J Eval Clin Pract. Author manuscript; available in PMC 2018 December 01.

Published in final edited form as:

J Eval Clin Pract. 2017 December ; 23(6): 1322–1328. doi:10.1111/jep.12781.

The Hawthorne Effect in Direct Observation Research with Physicians and Patients

Meredith A. Goodwin, PhD[1], **Kurt C. Stange, MD,PhD**[1,2,3,4,5], **Stephen J. Zyzanski, PhD**[1,2,5], **Benjamin F. Crabtree, PhD**[6], **Elaine A. Borawski, PhD**[2,4,5], and **Susan A. Flocke, PhD**[1,2,4,5]

[1]Department of Family Medicine and Community Health, Case Western Reserve University, 11000 Cedar Ave, Suite 402, Cleveland, Ohio, 44106-7136, USA

[2]Department of Epidemiology and Biostatistics, Case Western Reserve University

[3]Department of Sociology, Case Western Reserve University

[4]Prevention Research Center for Healthy Neighborhoods at Case Western Reserve University

[5]Case Comprehensive Cancer Center, Case Western Reserve University

[6]Department of Family Medicine, Rutgers University

Abstract

Rationale, Aims, and Objectives—This study examines the degree to which a "Hawthorne effect" alters outpatient-visit content.

Methods—Trained research nurses directly observed 4454 visits to 138 family physicians. Multiple data sources were used to examine the Hawthorne effect including: differences in medical record documentation for observed visits and the prior visit by the same patient, time use during visits on the first versus the second observation day of each physician, and report by the patient, physician, and observer of the effect of observation.

Results—Visits on the first versus the second observation day were longer by an average of 1 minute ($p<.001$); there were time-use differences for 4 of 20 behavior categories evaluated. No effect of the observer on the interaction was reported by 74% of patients and 55% of physicians. The majority of those that reported an affect indicated it was slight. Patients with non-white race, lower educational level and poorer health were more likely to report being affected by the observer.

Conclusions—In a study that was designed to minimize the Hawthorne effect, the presence of an observer had little effect on most patient-physician visits, but appeared to at least slightly effect a subgroup of vulnerable patients.

Keywords

primary care; direct observation; multi-method; methodology; chart review; survey; Hawthorne effect; primary care practice; practice-based research network

Corresponding author: Susan A. Flocke, PhD, Case Western Reserve University. 11000 Cedar Ave., Suite 402, Cleveland, OH 44106. Phone 216 368-3887. Fax 216 368 4348.

INTRODUCTION

This study examines the degree to which a Hawthorne Effect alters the content of a directly observed outpatient visit. Direct observation is likely to be more accurate than clinician report, which tends to over represent performance of service delivery[1,2] and may be subject to social desirability bias. Patient report tends to under represent delivery of services that are not well understood by the general public[3], whereas medical record review tends to under represent health behavior counseling[4]. Administrative data may be biased toward services that are important for billing[5].

While direct observation should overcome many of these biases, its use is limited in part by concern about the Hawthorne effect. The Hawthorne effect was first described in industrial experiments[6] and represents the concern that participants may knowingly or unknowingly alter their behavior as a result of being observed. Despite significant controversy regarding the validity of the interpretations of the study findings by the investigators[7,8], the concern that direct observation as a data collection method can affect the very thing that one is trying to measure is common.

Despite the common use of the term and the reasonableness of the concern, surprisingly few rigorous studies have been conducted to evaluate the impact of a Hawthorne effect[9–14]. Some prior research suggests that observation may not result in significant changes in behavior[8,15–18]. A systematic review of 19 studies assessing a Hawthorne effect found mixed results by study type[19]. The clinical trials and the quasi-experimental studies included in this review showed no effect; four observational studies showed a positive overall effect, but the reviewers noted substantial potential for observer bias. Overall, the studies included for evaluation were highly heterogeneous making the data difficult to interpret[19]. Most previous studies of the Hawthorne effect on health care visits compare different modes of measuring the content of visits but do not specifically examine the effect of observation[2,20–33]. The few relevant empirical studies are limited by small samples of patients or physicians[17,18], comparison only of an observation period vs. a non-observation period[7,33], and the examination of a single characteristic or the use of a single measure of observer effect[33,34].

Therefore, we took advantage of the unique opportunity to evaluate the Hawthorne effect in a large, multimethod direct observation study of outpatient service delivery. The study design attempted to minimize any Hawthorne effect, but also assessed the possible effect of the observer by 1) report of all participants (patient, physician, observer) of the degree to which the observer's presence changed physician or patient behavior; 2) comparing the observed visit vs. a prior visit across key variables assessed from the medical record; and 3) comparing time use during the first day of observation vs. a second day of observation. This study also aims to identify if sub-groups of patients were more likely to be affected by observation.

METHODS

Study Design and Data Collection

The Direct Observation of Primary Care Study (DOPC)[3,4] was a multimethod cross-sectional observation study designed to examine the content and context of outpatient visits to family physicians in Northeast Ohio, with a particular emphasis on clinical preventive service delivery. Participating physicians were observed by a trained research nurse while providing outpatient care on two separate days. The data collection days were separated by approximately four months.

A nurse researcher completed a direct observation checklist of visit characteristics. After the observed encounter, patients were asked to complete a patient exit questionnaire. Subsequently, medical records of the observed patients were abstracted to assess patient characteristics and services delivered during the observed visit. For a random sample of 5% of patients for each physician the medical record also was reviewed for the visit prior to the observed visit. Following the conclusion of the two observation days, physicians completed a questionnaire and the nurse researcher completed a checklist of practice characteristics. The study protocol was approved by the University Hospitals of Cleveland Institutional Review Board.

Design Features to Minimize the Potential for a Hawthorne Effect

Several study design features were implemented to minimize any Hawthorne effect. First, no specific hypotheses were shared with participants. Participants were told that this was a study of the content of outpatient visits to family physicians. Second, consecutive outpatient visits were observed in order to minimize deviations from usual routine. Third, the observation was conducted by nurses, whose presence is common in outpatient encounters. Fourth, the research nurses observed from the least obtrusive corner of the room and attempted to behave as a "fly on the wall." Patients and physicians were also instructed to ignore the nurse observer.

Measures

Assessment of effect of observer—Multiple data sources were used to examine the Hawthorne effect including: patient- and physician-report of the effect of observation, the nurse observer's impression of how much her presence affected the visit, differences in medical record documentation for observed visits and the prior visit by the same patient, and time use during visits on the first observation day versus the physician's second observation day.

The patient and physician questionnaires each included an item to assess the effect of the presence of the observer. The patient questionnaire asked: "How much would you say the presence of the nurse-observer changed what you said or did today?" The physician questionnaire asked: "How much did the presence of the nurse observer change what you did when you were with patients?" In addition, on the practice environment checklist completed for each physician, the observing research nurse reported her response to the question: "How much do you think your presence changed the physician's behavior during

patient visits?" The response format for these questions was a on a 5 point Likert scale: 1 'a lot', 2 'quite a bit', 3 'moderately', 4 'slightly' and 5 'not at all'.

Assessment of the content of the visit

Direct Observation: Two aspects of the content of the visit were assessed by direct observation: the delivery of preventive services and time use. Delivery of preventive services recommended by the US Preventive Services Task Force[3,35,36], was assessed by direct observation for screening (15 services such as blood pressure, Pap smear, and cholesterol screening) and counseling (24 services such as discussion or advice about diet, contraception, and tobacco cessation).

Time use during the visit was measured using the Davis Observation Code (DOC)[37], which classifies time spent during each 15 second interval into 20 different behavioral categories relevant to delivering primary health care. Behavioral categories include history taking, physical exam, and health education, for example.

Medical record: The effect of the observer on what was documented regarding the visit was assessed by comparing the medical record notes for the observed visit and the visit just prior to the observed visit. The following items were assessed: whether a drug was prescribed, whether a referral was made, the number of problems addressed, the number of preventive services delivered. The Evaluation and Management Current Procedure Terminology (E&M CPT) code as well as the extent of history taking and the thoroughness of the physical exam[38] were also assessed based on the documentation in the medical record.

Other Co-variates—Other patient, visit, physician and practice factors were assessed through direct observation, chart review or questionnaire. Patient gender, visit type (acute, chronic, well care, etc.) and presence of a family member were measured using direct observation. Patient age, whether they were a new or established patient, the number of chronic illnesses, the number of medications, the number of years the patient has been with the physician and the number of visits to the practice within the previous year were assessed by medical record review. Patient race was measured by questionnaire for the 75% of patients that returned a questionnaire, and by direct observation for others (Kappa=.90 for direct observation and patient report)[39]. Patient health status was measured by a 5-item index (α=.81)[40] modified from the MOS 6-item General Health Survey[41]. Patients' educational level was ascertained by patient questionnaire.

The physician questionnaire assessed age, sex, years in current practice, residency training, and practice ownership. The number of patients seen by the physician on the observation day (total including those who chose not to participate), the practice type (solo or group) and the practice location (urban, suburban, rural) were assessed by the nurse researcher using a practice characteristic checklist at the conclusion of the observation day.

Analyses

First, for a 5% random subsample of visits, visit characteristics documented in the medical record were compared for the observed vs. the previous visit, using paired t-tests for

continuous and ordinal variables, Wilcoxon's matched-pairs signed ranks test for categorical variables, and McNemar's test for dichotomous variables.

Next, differences in time use during visits for the first day of observation vs. the second day of observation for each physician were assessed using a series of generalized linear models with each DOC time use category as an outcome variable and day of observation as the independent variable. The generalized linear models were used to account for the clustered design effect (i.e., patients clustered with in physician practices). Subsequent analyses also controlled for patient and visits characteristics that differed between the two observation days. Further, given that the DOC consisted of 20 behavioral categories, the analyses were limited to those that occurred in the majority of visits (>50%) and the Benjamini-Hochberg procedure was used to adjust for multiple testing[42]. **The proportional reduction in variance (PRV) statistic is reported as an indicator of effect size for these multilevel generalized linear models[43].**

The degree to which the presence of an observer affected behaviors during the visit also was reported as frequencies from three perspectives: the patient, the physician, and the observing nurse. Patient characteristics associated with any patient reported effect of the observer (not at all vs. all other responses) were assessed using multilevel logistic regression. Patient characteristics in the model included age, sex, race, education level, reason for visit, other family member present, number of chronic illnesses and health status. As a final step, physician report of effect was included as a variable in the model to assess the degree to which the physician report of effect was associated with patient report. **Significant change in model deviance (tested using Chi Square and p <.05) from the null model to the full predictor model was used as an indicator of model fit.**

RESULTS

The 138 participating physicians in 84 practices were similar to a nationally representative sample in age, percent in rural locations and number of patients seen per week[3,44], but were slightly more likely to be female and residency trained[3]. Among 4994 patients presenting for care, 4454 (89%) consented to participate and 3062 (69%) completed the questionnaire item assessing the effect of observer on their behavior. Patient questionnaire responders were older, more likely to be female, white, married, have a greater number of chronic illnesses and a longer relationship with the practice (all p<.01).

The comparison of the visit characteristics documented in the medical record for the observed visit versus the previous visit for the 5% random sample per physician (n=242) is presented in Table 1. None of the eight visit characteristics significantly differed indicating that there was no effect of the presence of the observer on the documentation of drugs prescribed, referral, number of problems addressed or the number of preventive services delivered. Likewise, there was no effect on the documentation regarding the extent of history taking and the thoroughness of the physical exam, as documented in the medical record.

Table 2 shows the mean duration of time for the visit and for the behavioral categories assessed using the DOC for Day 1 and Day 2 of observation. The means are adjusted for

patient and visit characteristics that differed between the two observation days ($p < .1$) as well as for the clustered design effect. Overall, the duration of the visit was 10% longer on average for the first day of observation compared to the second day. Four specific behavioral categories were significantly longer on the first day of observation compared to the second day: history taking, planning treatment, structuring interaction and physical examination. The actual magnitude of difference in time is small for all categories except history taking which differed by about 30 seconds on average.

The report of the effect of observation by patients, physicians and the nurse observers is shown in Table 3. A total of 3062 (69%) patients completed the questionnaire item assessing the effect of observer on their behavior. Seventy-four percent of patients reported being 'not at all' affected by the presence of the observer. Among the 138 physicians in the study, 128 completed the questionnaire and provided data regarding the effect of the nurse observer on interaction with patients. None of the physicians reported that the nurse observer affected their interactions with patients 'a lot'; 54.7% reported that they were not at all affected. Interestingly, the observing nurse was the least likely to report having a significant effect on the encounter. The nurse research report was completed on 137 of the 138 participating physicians; for only 16.7% of physicians did the nurse report any effect of her presence on physician behavior.

In order to identify characteristics associated with patient report of any effect of the observer on the visit, twelve patient and visit characteristics were examined. Table 4 presents the multivariable association of each characteristic with patient report of any effect of observation (i.e. not at all vs. all other responses). Most of the patient characteristics examined were not associated with a patient report of any effect of the observer on the interaction with the physician. However, poorer health status, being non-white and having a lower education level were associated with an increased odds of reporting an effect of the observer on the visit. For example, each one point improvement in health status was associated with a 0.84 decreased odds of the patient reporting an effect of the observer on the visit. This multivariable model was also evaluated including physician report of any effect of the observer. The association of physician report with patient report of any effect of the observer is not statistically significant (OR=1.01 (0.845, 1.21)); and inclusion of this variable in the model does not alter the observed associations reported in Table 4. (data not shown)

DISCUSSION

Our direct observation study was designed to minimize the potential for a Hawthorne effect. In this setting, little or no effect of observation by a research nurse during the outpatient visit was reported by most patients and physicians, or by the nurse observers. The lack of difference in medical record review of observed and unobserved visits provides some confirmation of this self-report. However, differences in time use between the first and second days of observation may be evidence of acclimation by the physician to the presence of the nurse-observer. The observed differences (longer visits and more time spent on history taking, physical examination, planning treatment and structuring the interaction) are consistent with physicians attempting to be more thorough in the presence of an observer.

Three patient characteristics were found to be associated with patient-reported effect: health status, patient race and educational level. The finding that sicker individuals more often reported an effect is consistent with other studies that have found poorer health status to be associated with greater patient discomfort during encounters[8,15,16,45].

The association of race with the Hawthorne effect is consistent with previous literature showing that African Americans (the most common minority group in our sample) have higher levels of mistrust of medical research[46–51]. Mistrust may also explain why individuals with lower education levels were more likely to report being affected. A study by Fiscella et al. (1998) demonstrated that lower education was associated with increased skepticism toward health care providers[49], perhaps also attributable to prior exploitation in medical research[48,52]. An additional explanation for these three groups of individuals reporting being affected is that they may have perceived the observer to be very different from themselves, that is, being healthy, predominantly white, and educated[53]. Matching observers to the predominate characteristics of the study population may diminish the perceived effect of the observer.

The rate at which physicians reported being unaffected in our study is similar to the 71% rate reported by Ram et al[54] in a video assessment study. Our study did not identify physician characteristics that define who was affected; however, it was clear from self-report and the time use data that not all physicians were affected in the same way. The physician-reported effect appeared to have little influence on patients' report of being affected as evidenced by the multilevel modeling. This independence of effects on the patient and the physician may indicate that any effect on physicians was subtle enough not to affect their patients. This hypothesis is supported by the findings of two previous studies using videotaping of patient encounters[49,50].

There are several limitations that need to be considered in terms of the present analyses. First, except for the sample of visits for which the prior visit served as a comparison, there was no control group. Due to voluntary participation and missing data, it is possible that the study selected patients and physicians less likely to be affected by observation. Patients who completed the survey question about the influence of an observer were different on several characteristics compared to the non-responders. Our analyses included co-variables to adjust for these differences. Finally, these data apply only to studies that take strong precautions to try to minimize the potential for a Hawthorne effect, such as avoiding sharing specific hypotheses and observing consecutive visits from unobtrusive vantage points.

To our knowledge, this is the first study to thoroughly examine the Hawthorne effect in medical outpatient visits using multiple measures and methods. It helps to inform whether, when, and for whom direct observation during research protocols affects behavior under study and other outcomes[19]. The findings of this study indicate that with proper attention paid to minimizing the Hawthorne effect, direct observation can be used to gather valuable data that cannot be collected accurately in other ways. This finding is consistent with a small number of studies in other settings that suggest that the Hawthorne effect may be more of a pragmatic concern[8,15,16] in well-designed studies.

Future studies using direct observation may want to account for a potential acclimation period by examining carefully the initial observed encounters and potentially even excluding the first several encounters for each physician. Studies that target “vulnerable” individuals such as sick, minority and lower educated individuals need to pay particular attention to the potential effects of observers and may want to consider alternative methods. If direct observation is still the method of choice, efforts should be made to strengthen trust of research personnel by ensuring privacy and confidentiality.

Acknowledgments

This research was supported by grants (R01 CA60862, K24 CA81031, R01 CA105292) from the National Cancer Institute, by a Family Practice Research Center Grant from the American Academy of Family Physicians, a Generalist Physician Faculty Scholars Award (RWJ 23851) from the Robert Wood Johnson Foundation, and by a Clinical Research Professorship (CRP-07-236-01) from the American Cancer Society.

The authors are grateful for the clinician, staff and patient participants of this study; we wish to thank the Family Medicine writing workgroup member for helpful feedback on early drafts of this manuscript.

References

1. McPhee SJ, Richard RJ, Solkowitz SN. Performance of cancer screening in a university general internal medicine practice: Comparison with the 1980 American Cancer Society Guidelines. J Gen Intern Med. 1986; 1:275–281. [PubMed: 3772615]
2. Montano DE, Phillips WR, Montaño DE. Cancer screening by primary care physicians: A comparison of rates obtained from physician self-report, patient survey, and chart audit. Am J Public Health. 1995; 85(6):795–800. [PubMed: 7762712]
3. Stange KC, Zyzanski SJ, Jaén CR, et al. Illuminating the “black box”. A description of 4454 patient visits to 138 family physicians. J Fam Pract. 1998; 46(5):377–389. [PubMed: 9597995]
4. Stange KCC, Zyzanski SJJ, Smith TFF, et al. How valid are medical records and patient questionnaires for physician profiling and health services research? Med Care. 1998; 36(6):851–867. [PubMed: 9630127]
5. Cooper GS, Yuan Z, Stange KC, Dennis LK, Amini SB, Rimm AA. Agreement of Medicare claims and tumor registry data for assessment of cancer-related treatment. Med Care. 2000; 38(4):411–421. [PubMed: 10752973]
6. Roethlisberger, FJ., Dickson, WJ. Management and the Worker. Cambridge, MA: Harvard University Press; 1961.
7. Gillespie, R. Manufacturing Knowledge: A History of the Hawthorne Experiments. Cambridge, MA: Cambridge University Press; 1991.
8. Parsons HM. What happened at Hawthorne? Science. 1974; 183(4128):922–930. [PubMed: 17756742]
9. Chiesa M, Hobbs S. Making sense of social research: How useful is the Hawthorne Effect? Eur J Soc Psychol. 2008; 38(1):67–74.
10. Coleman T. Using video-recorded consultations for research in primary care: advantages and limitations. Fam Pract. 2000; 17(5):422–427. [PubMed: 11021903]
11. Verstappen WHJM, van der Weijden T, ter Riet G, Grimshaw J, Winkens R, Grol RPTM. Block design allowed for control of the Hawthorne effect in a randomized controlled trial of test ordering. J Clin Epidemiol. 2004; 57(11):1119–1123. [PubMed: 15567627]
12. McCarney R, Warner J, Iliffe S, van Haselen R, Griffin M, Fisher P. The Hawthorne Effect: a randomised, controlled trial. BMC Med Res Methodol. 2007; 7:30. [PubMed: 17608932]
13. Fernald DH, Coombs L, DeAlleaume L, West D, Parnes B. An assessment of the Hawthorne Effect in practice-based research. J Am Board Fam Med. 2012; 25(1):83–86. [PubMed: 22218628]

14. Srigley JA, Furness CD, Baker GR, Gardam M. Quantification of the Hawthorne effect in hand hygiene compliance monitoring using an electronic monitoring system: a retrospective cohort study. BMJ Qual Saf. 2014; 23(12):974–980.
15. Adair JG. The Hawthorne effect. A reconsideration of the methodological artifact. J Appl Psychol. 1984; 69:334–345.
16. Jones SR. Was there a Hawthorne Effect? Am J Sociol. 1992; 98(3):451–468.
17. Pringle M, Stewart-Evans C. Does awareness of being video recorded affect doctors' consultation behavior. Br J Gen Pract. 1990; 40(340):455–458. [PubMed: 2271278]
18. Martin E, Martin PML. The reactions of patients to a video camera in the consulting room. J R Coll Gen Pract. 1984; 34(268):607–610. [PubMed: 6502570]
19. McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: New concepts are needed to study research participation effects. J Clin Epidemiol. 2014; 67(3):267–277. [PubMed: 24275499]
20. Rohrbaugh M, Rogers JC. What did the doctor do? When physicians and patients disagree. Arch Fam Med. 1994; 3:125–129. [PubMed: 7994433]
21. Gerbert B, Hargreaves WA. Measuring physician behavior. Med Care. 1986; 24(9):838–847. [PubMed: 3762247]
22. Suarez L, Goldman DA, Weiss NS. Validity of Pap smear and mammogram self-reports in a low-income Hispanic population. Am J Prev Med. 1995; 11(2):94–98. [PubMed: 7632456]
23. McKenna MT, Speers M, Mallin K, Warnecke R. Agreement between patient self-reports and medical records for Pap smear histories. Am J Prev Med. 1992; 8(5):287–291. [PubMed: 1419128]
24. Sawyer JA, Earp JA, Fletcher RH, Daye FEF, Wynn TM. Accuracy of women's self-report of their last Pap smear. Am J Public Health. 1989; 79(8):1036–1037. [PubMed: 2751021]
25. Silagy C, Muir J, Coulter A, Thorogood M, Yudkin P, Roe L. Lifestyle advice in general practice: Rates recalled by patients. Br Med J. 1992; 305:871–874. [PubMed: 1422405]
26. Johnson CS, Archer J, Campos-Outcalt D. Accuracy of Pap smear and mammogram self-reports in a southwestern native American tribe. Am J Prev Med. 1995; 11(6):360–363. [PubMed: 8775656]
27. Harlow SD, Linet MS. Agreement between questionnaire data and medical records: The evidence for accuracy of recall. Am J Epidemiol. 1989; 129(2):233–247. [PubMed: 2643301]
28. Etzi S, Lane DS, Grimson R. The use of mammography vans by low-income women: The accuracy of self-reports. Am J Public Health. 1994; 84(1):107–109. [PubMed: 8279594]
29. Degnan D, Harris R, Ranney J, Quade D, Earp JA, Gonzalez J. Measuring the use of mammography: Two methods compared. Am J Public Health. 1992; 82(10):1386–1388. [PubMed: 1415867]
30. Zapka JG, Bigelow C, Hurley T, et al. Mammography use among sociodemographically diverse women: The accuracy of self-report. Am J Public Health. 1996; 86(7):1016–1021. [PubMed: 8669504]
31. Zuckerman AE, Starfield B, Hochreiter C, Kovasznay B. Validating the content of pediatric outpatient medical records by means of tape-recording doctor-patient encounters. Pediatrics. 1975; 56(3):407–411. [PubMed: 1161397]
32. Romm FJ, Putnam SM. The validity of the medical record. Med Care. 1981; 19(3):310–315. [PubMed: 7218896]
33. Starfield B, Steinwachs D, Morris I, Bause G, Siebert S, Westin C. Presence of observers at patient-practitioner interactions: Impact on coordination of care and methodologic implications. Am J Public Health. 1979; 69(10):1021–1025. [PubMed: 484755]
34. Block MR, Schaffner KF, Coulehan JL. Ethical problems of recording physician-patient interactions in family practice settings. J Fam Pract. 1985; 21(6):467–472. [PubMed: 4067528]
35. U.S. Preventive Services Task Force. Guide to Clinical Preventive Services. Vol. 263. Baltimore: Agency for Healthcare Research and Quality; 1996. http://www.uspreventiveservicestaskforce.org/
36. Stange KC, Flocke SA, Goodwin MA, Kelly RB, Zyzanski SJ. Direct observation of rates of preventive service delivery in community family practice. Prev Med (Baltim). 2000; 31:167–176.
37. Callahan EJ, Bertakis KD. Development and validation of the Davis Observation Code. Fam Med. 1991; 23(1):19–24. [PubMed: 2001776]

38. Kirschner CG, Burkett RC, Coy JA. Physicians' Current Procedural Terminology: CPT '94. 1994
39. Williams RL, Flocke SA, Stange KC. Race and preventive service delivery among African Americans and whites. Med Care. 2001; 38:1260–1267.
40. Flocke SA, Stange KC, Zyzanski SJ. The association of attributes of primary care with preventive service delivery. Med Care. 1997; 36:AS21–AS30.
41. Ware, J., Nelson, E., Sherbourne, C., Stewart, A. Preliminary tests of a 6-item general health survey: a patient application. In: Ware, ASJ., editor. Measuring Functioning and Well-Being. Durham, NC: Duke University Press; 1992.
42. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. JSTOR. 1995; 57(1):289–300.
43. Peugh JL. A practical guide to multilevel modeling. J Sch Psychol. 2010; 48(1):85–112. [PubMed: 20006989]
44. American Academy of Family Physicians. Facts about Family Practice. Kansas City, MO: American Academy of Family Physicians; 1996.
45. Hall JA, Roter DL, Milburn MA, Daltroy LH. Patients' health as a predictor of physician and patient behavior in medical visits. A synthesis of four studies. Med Care. 1996; 34(12):1205–1218. [PubMed: 8962587]
46. Shavers VL, Lynch CF, Burmeister LF. Racial differences in factors that influence the willingness to participate in medical research studies. Ann Epidemiol. 2002; 12:248–256. [PubMed: 11988413]
47. Doescher MP, Saver BG, Franks P, Fiscella K. Racial and ethnic disparities in perceptions of physician style and trust. Arch Fam Med. 2000; 9(10):1156–1163. [PubMed: 11115223]
48. Corbie-Smith G, Thomas SB, Williams MV, Moody-Ayers S. Attitudes and beliefs of African Americans toward participation in medical research. J Gen Intern Med. 1999; 14(9):537–546. [PubMed: 10491242]
49. Fiscella K, Franks P, Clancy CM. Skepticism toward medical care and health care utilization. Med Care. 1998; 36(2):180–189. [PubMed: 9475472]
50. Quam K. Evaluation of patient comfort during a videotaped office visit. Eval Program Plann. 1990; 13(4):395–398.
51. Hargreaves PN, Peppiatt R. Is videotaping of consultations acceptable to patients attending a hospice day centre? Palliat Med. 2001; 15(1):49–54. [PubMed: 11212468]
52. Fairchild AL, Bayer R. Uses and abuses of Tuskegee. Science. 1999; 284(5416):919–921. [PubMed: 10357678]
53. Robson S, Wardle J. Who's watching whom? A study of the effects of observers on group discussions. J Mark Res Soc. 1988; 30(3):333–359.
54. Ram P, Grol R, Rethans JJ, Schouten B, van der Vleuten CPM, Kester A. Assessment of general practitioners by video observation of communicative and medical performance in daily practice: issues of validity, reliability and feasibility. Med Educ. 1999; 33(6):447–454. [PubMed: 10354322]

Table 1

Medical Record Compared for the Observed Visit and Previous Visit for the Same Patient (n=242 patients)

Visit Characteristic	Observed visit	Previous visit	p
Drug prescribed (%)	63.4	60.0	.46
Referral made (%)	5.5	5.4	1.00
Number of problems addressed (mean)	1.8 (1.0)	1.7 (1.0)	.11
Extent of examination * (median)	1	1	.59
Extent of history taking * (median)	1	1	1.00
Nurse estimate of CPT code ** (median)	3.0	2.9	.10
Number of preventive services delivered			
Screening (mean)	2.2 (1.2)	2.1 (1.4)	.26
Counseling (mean)	0.5 (0.9)	0.5 (1.2)	.39

* 0=none, 1=problem focused, 2=expanded problem focused, 3=detailed and 4=comprehensive

** CPT codes 99201 and 99211 =1; 99202 and 99212=2; 99203 and 99213=3; 99204 and 99214=4; and99205 and 99215=5.

Table 2

Time Use Compared for the Patient Visits on the First and Second Days of Observation

	Adjusted Means*			
Time use behavior, minutes	**First Observation Day n=2106**	**Second Observation Day n=2348**	**P**	**PRV**
Length of visit	10.0	9.0	.001	**14.5%**
History taking	6.5	5.9	.001	**12.1%**
Planning treatment	3.1	2.9	.002	**6.2%**
Health education	2.1	2.1	.86	**5.3%**
Feedback on evaluation results	1.3	1.2	.10	**6.4%**
Gathering family information	1.6	1.5	.23	**6.6%**
Chatting	1.1	1.0	.09	**6.9%**
Structuring the interaction	0.9	0.8	.001	**9.8%**
Answering patient questions	0.6	0.6	.65	**3.6%**
Physical examination	1.6	1.4	.001	**2.3%**

* More than one behavior could be recorded per interval. Analyses were adjusted for patient age, sex, race, reason for visit, other family member present, number of medications, number of chronic illnesses, number of years as a patient of the physician and number of visits to the practice within the past year. Analyses were also adjusted for clustered design effects. **The degrees of freedom for these multilevel models were (137 for level 2 and 4171 for level 1).**

PRV is the proportional reduction in variance statistic and represents the proportion of reduction in variance from the model including the predictor variables compared to the null model.

Table 3

Patient, Physician and Observing Research Nurse Report of the Effect of the Observer

	Not at all	Slightly	Moderately	Quite a bit	A lot
	%	%	%	%	%
Patient self-report (n=3062)					
Degree to which observer's presence changed their behavior	73.7	9.5	7.8	3.4	5.5
Physician self-report (n=128)					
Degree to which observer's presence changed their behavior	54.7	36.7	7	1.6	0
Observing research nurse report (n=137)					
Degree that presence changed physician's behavior	83.2	13.1	3.6	0	0

Table 4

Patient and Visit Characteristics Associated with Patient-Report of any Effect of the Nurse Observer (n=3062)

Patient and Visit Characteristics	B[†]	SE	OR[*]	95 % CI	p
Age	−.00001	.003	.99	(0.99, 1.00)	0.98
Patient sex					
Male (reference)			1.00		
Female	−0.06	0.10	0.94	(0.77, 1.15)	0.55
Patient race					
White (reference)			1.00		
Non-white	0.48	0.18	1.61	(1.13, 2.30)	0.01
Educational level attained					
High school or less (reference)			1.00		
More than high school	−0.37	0.10	0.69	(0.57, 0.84)	<.001
Reason for visit					
Other			1.00		
Acute illness	0.28	0.20	1.32	(089, 1.97)	0.17
Chronic illness	0.19	0.20	1.21	(0.81, 1.80)	0.34
Well care	0.25	0.24	1.28	(0.79, 2.07)	0.31
Family member present					
No (reference)			1.00		
Yes	−0.24	0.15	0.78	(0.58, 1.05)	0.10
Number of chronic illnesses	0.02	0.02	1.03	(0.98, 1.07)	0.23
Health status[**]	−0.17	0.06	0.84	(0.75, 0.96)	0.009

[†] B represent unstandardized coefficients

[*] OR indicates odds associated with patient reporting any effect of the nurse observer.

[**] 1–5 score, low scores indicate poor health status