MEDICAL CARE
Volume 41, Number 3, pp 368–374


# Antidepressant Use

## Concordance Between Self-Report and Claims Records

ANNE KWON, MS,* KATHLEEN M. BUNGAY, PHARMD,* YU PEI, MPA,* WILLIAM H. ROGERS, PHD,* IRA B. WILSON, MD,* QI ZHOU, MD,† AND DAVID A. ADLER, MD*

**BACKGROUND. Researchers need valid methods to assess whether patients are taking their antidepressant medications. Two important sources of data on drug exposure are patients' self-reports and pharmacy claims.**

**OBJECTIVE. To compare self-report and claims data for antidepressant exposure.**

**RESEARCH DESIGN. Cross-sectional analysis.**

**SUBJECTS. This study comprised 422 contemporaneous self-report and claims data points obtained from 164 unique patients in a longitudinal depression study in which patients completed up to five surveys during an 18-month period.**

**MEASURES. For the self-report measure, the following question was asked: Do you now take any prescription medicines for depression? Using claims data, patients were considered to be using an antidepressant if they had filled at least one antidepressant prescription in the 90 days before survey dates.**

**RESULTS. Self-report and claims agreed in 85% (358/422) of cases, with a kappa of 0.69. Eighty-eight percent (56/64) of discrepant cases using other study data sources was resolved. Reasons for discrepancies included the use of medications for conditions other than depression (32/64), recent AD discontinuations (6/64), samples usage (3/64), and low-frequency/PRN use (7/64).**

**CONCLUSIONS. Self-report and claims showed good concordance, but they reflect different truths. Self-report identifies medications intended primarily for the treatment of depressive disorders, whereas claims data identify use of medicines with antidepressant effects. Our assessment of discordant cases showed self-report to be more valid than claims to assess current antidepressant use for depression therapy.**

**Key words: Antidepressant exposure; medication use; self-report; pharmacy claims. (Med Care 2003;41:368–374)**

Researchers need valid methods of assessing medication use and drug exposure, because medication-taking behavior is a critical part of the success of most medical treatment. Data on medication exposure are typically obtained from patient interview, self-report surveys, pill counts, medical records or claims databases, but there is no currently accepted gold standard for determining drug exposure.[1–7] Consequently, validation of one approach to medication exposure requires that it be compared with a different approach.

Specific comparisons of self-report and claims data in the assessment of current drug exposure are rare, and especially in regard to antidepressant

*From the Health Institute, Division of Clinical Care Research, Tufts-New England Medical Center, Boston, Massachusetts.

†From Tufts Health Plan, Watertown, Massachusetts.

Supported by grant NIMH RO1 MH56214 and the New England Medical Center Research Fund.

Address correspondence and reprint requests to: Kathleen M. Bungay, PharmD, The Health Institute, Division of Clinical Care Research, Tufts-New England Medical Center, 750 Washington Street, Boston, MA 02111. E-mail: kbungay@lifespan.org

Received April 23, 2002; initial review June 14, 2002; accepted September 12, 2002.

use. Existing studies show moderate concordance, but are hampered by several important limitations including small sample sizes and long claims exposure windows. For example, Sjahid et al[8] focused on cardiovascular drugs in European elderly with a 6-month window for exposure in claims data, whereas Enlund et al[9] concentrated on antihypertensive drugs in a sample of patients aged 25 to 29 with a year-long window for exposure. Klungel et al[10] and Lau et al[11] examined all drug classes, but again, Lau's study used only a small sample of elderly in a 1-year claims exposure window and Klungel[10] used survey self-report data in a small sample of hypertensive patients. Saunders et al[12] is the only published study of claims data and self-report for the measurement of antidepressant exposure and adherence of which we are aware. They reported kappa values of 0.33 and 0.72 at 1 and 4 months in a sample of 164 patients. Their self-report measure consisted of two telephone interview questions and their claims measure assessed the possibility of continuous use of medications beyond the patients' interview date.

We therefore compared self-report and claims data on drug exposure in a longitudinal study of patients with major depression, dysthymia, and subthreshold depression. We utilized self-report data from a mailed questionnaire and claims data from a local health maintenance organization. In addition, we used other study data to examine discrepancies between the two data sources and to further understand the strengths and weaknesses of each.

## Materials and Methods

### Patients

Patients were participants in a longitudinal randomized trial of depression detection and treatment. We recruited participants from nine primary care sites in the metropolitan Boston area, including five general medical practices at an academic medical center, one urban and two suburban community Boston medical practices, and a community health center. Patients had many different types of insurance coverage.

At the nine sites, we screened 16,707 consecutive patients for depression during a 2-year period using the Primary Care Screener for Affective Disorders (PC-SAD). The PC-SAD is a self-administered health survey that provides a DSM-IV symptom count for major depressive disorder (MDD) and dysthymia.[13] Patients who screened positive for MDD and/or dysthymia, and gave informed consent approved by the two institutional review boards representing all sites, were enrolled in a RCT of a clinical pharmacist intervention versus usual care. Patients whose score indicated sub-threshold depression (positive response to prescreener depression/dysthymia questions and at least 2 MDD symptoms or 1 dysthymia symptom) were offered concurrent enrollment into a separate nonrandomized follow-up group that also received usual care. Eligible patients were aged 18 or older, could read and understand English, were not pregnant or had not delivered a baby in the last 6 months, and did not present with evidence of current alcoholism, bipolar disorder, or psychotic disorder. We did not exclude patients with lifetime alcoholism, long-term/chronic depression, anxiety, personality disorder, or comorbid medical conditions. We mailed questionnaires to all patients at the initial, 3-, 6-, 12-, and 18-month time points. Fifty-five percent of patients in this analysis were from the RCT, and 45% were from the subthreshold observational study. For this analysis we studied the subset of study participants (n = 200, 21.3%) who were members of Tufts Health Plan.

### Study Measures

We constructed three dichotomous measures of current antidepressant use, two from self-report and one from claims data. All these variables produced a yes/no indicator of whether the patient was currently taking an antidepressant.

### Self-Report Measures

The "direct" self-report measure used two redundant questions on antidepressant use: (1) "Do you now take any prescription medicines for depression?" and (2) "Are you now taking a prescription medicine for depression?" These items were both included to facilitate a skip pattern and were separated by an interval of 10 items on the questionnaire. Response options were "yes" or "no." We defined the "direct" self-report measure from the response to question (1); if (1) was not answered, we used the response from question (2).

TABLE 1. Direct Self-report vs. Claims Records: Agreement on Antidepressant Use

| | | Pharmacy Claims | | |
|---|---|---|---|---|
| | | Yes | No | Totals |
| Self report | Yes | 138 | 15 | 153 |
| | No | 49 | 220 | 269 |
| | Totals | 187 | 235 | 442 |

n = 422 episodes.

Patients were also asked: "Please take a minute to write down the names of ALL your medications. If you are unsure, get the prescription bottle(s) and write down the name(s) that you see on the label." Responses were coded according to medication type using the American Hospital Formulary System,[14] where codes beginning with 28:16.04 or 28:12.92 were considered antidepressant medicines. Mood stabilizers (code 28:12.92) were included because of their use in augmenting antidepressant response, and analyses of the data with and without these medications were not significantly different.

The "list" self-report measure categorized patients as currently taking antidepressants if any of their listed medications had a code indicating an antidepressant. Patients whose listed medications did not contain a single match for the required codes were categorized as not taking antidepressant medicines, whereas patients who did not list any medications were considered to have missing values for this measure.

## Pharmacy Claims Data

Pharmacy utilization and cost data were available for a subset of patients in our study who were members of Tufts Health Plan, which is one of the largest nonprofit HMOs servicing the greater Boston area with a membership of approximately 900,000 individuals. These claims records were examined to make sure that patients had valid member files that extended at least 90 days before their survey return dates. Antidepressant use within the claims database was then defined as the appearance of any claim for an antidepressant within the 90 days preceding a patient survey date, using the same set of antidepressants generated earlier for the "list" measure.

## Additional Data Sources

An intervention protocol that was based on principles of clinical pharmacy practice and the AHCPR guidelines[15] provided the framework for pharmacists' intervention activities. The intervention involved a series of seven in-person or phone visits during 18 months by a clinical pharmacist trained in depression care who assisted in the choice of medication, optimized the medication regimen, and insured adherence to AHCPR depression guidelines. Pharmacists consulted automated hospital-based medical records as needed and relayed information to the responsible physician based on the PCP's preference for frequency and format. Medical records were also available for patients who were not in the pharmacist intervention arm of the trial.

Analyses of the experimental group receiving the pharmacist intervention did not reveal greater concordance over time. Thus, all study groups were combined for the final analyses.

## Statistical Analyses

Our primary analysis compared the direct self-report and claims measures. Secondary analyses investigated agreement between the list self-report and claims measure, and between list and direct self-report measures.

The individual episode, or case, was the unit of analysis. We defined an episode as an instance where data were available for two of our measures; eg, a self-report matched with a pharmacy claims data measure. Out of 527 usable claims data points, 422 had matched "direct" report measures and 482 had matched "list" measures. Two-by-two contingency tables were used to generate comparisons between the three derived measures of antidepressant use, with concordance between measures assessed using percentage agreement and

the kappa statistic. Percentage agreement was obtained by calculating the proportion of episodes where both measures agreed, out of all possible episodes. The kappa statistic was used to determine the amount of agreement that would exist beyond chance levels. Kappa values of approximately 0.70 have been characterized as good and substantial.[16,17] Significance tests were corrected for the intercorrelation of responses from the same individuals at multiple points over time.[18] We computed a design effect for using multiple episodes per patient by comparing logistic regressions for one measure on the other with and without correction for multiple episodes. The ratio of corrected to uncorrected standard error for kappa was taken to be the same as the ratio for the logistic regression.

Discrepant episodes from the two-way tables between self-report and pharmacy claims records were also investigated in detail. For episodes with patients in the experimental arm of the study, extensive histories from pharmacist and hospital medical records were used to further assess drug exposure at each time point. For episodes involving nonexperimental patients, we analyzed patterns of medication use in the claims data and looked for corresponding survey responses to questions about current medical conditions. We analyzed all discrepant episodes in terms of patients' entire medication histories surrounding each time point as defined in the claims database, taking into account prescription days' supply and strength. Decisions on explanations were based on clinical experience and the patients' self-reported medical history. All discrepant episodes were reviewed independently by at least two reviewers. All analyses were conducted using STATA versions 6.0 and 7.0.

## Results

The mean age of our sample was 45 years, 66% were female, 81% were white, 34% were married, and 75% had at least some college education. Nineteen percent reported incomes greater than $80,000 per year and 77% worked 20 or more hours per week. Patients were using an average of 2.9 medications, and 35% were using at least one antidepressant. SSRIs accounted for 70% of instances where a claim for an antidepressant appeared within 90 days of the survey return date. There was an average of 3.1 ± 1.0 episodes per patient.

### Agreement Between "Direct" Self-Report and Claims Records

Overall agreement on antidepressant use between the "direct" self-report measure and claims records was 85%, with a kappa of 0.69. In this comparison, there were 422 distinct episodes from 164 unique respondents (Table 1). Relative to the claims measure, the "direct" self-report measure had a sensitivity of 74% and a specificity of 94%. Conversely, the claims measure had a sensitivity of 90% and a specificity of 82% against self-report.

### Discrepant Responses: "Direct" Self-Report = NO, Claims = YES

Examination of the discrepant cases in the "direct" self-report versus claims record comparison revealed several trends (Table 2). Among the 49 episodes where the "direct" method reported no antidepressant use whereas the Claims method showed antidepressant fills, more than half of the cases were accounted for by the claims measure capturing medications that were being used for conditions other than depression (n = 32, 65% of 49 episodes). These indications ranged from anxiety disorders to smoking cessation and migraine prophylaxis, with insomnia comprising the largest subcategory (n = 21 of 32). The remainder of the 49 discrepant cases involved patients who appeared to have discontinued their medication, eg, an abrupt lack of refills after the survey return date by patients who had been consistently filling their antidepressant prescriptions. Other cases revealed single medication fills that were not continued beyond the patient's survey return date, which was interpreted as a decision not to undertake an antidepressant regimen.

### Discrepant Responses: "Direct" Self-Report = YES, Claims = NO

There were substantially fewer episodes for which the "direct" measure classified patients as using an antidepressant and the claims measure disagreed (n = 15) (Table 2). The largest category in this set consisted of episodes of possible low-

TABLE 2. Direct Self-report vs. Claims Records: Description of Disagreements

| | |
|---|---|
| Patient says NO, Claims say YES (n = 49) | |
| Medications in claims records used for conditions other than depression | 65% |
| Medications were recently discontinued | 12% |
| Failure to continue with a newly started antidepressant | 8% |
| Unknown | 14% |
| Patient says YES, Claims say NO (n = 15) | |
| Low frequency or PRN use | 47% |
| Patient was using samples | 20% |
| Possible irregularities within claims record | 27% |
| Unknown | 7% |

frequency use (n = 7, 47%). Other episodes could be explained by the use of medication samples by patients newly started on antidepressant treatment (n = 3), and by possible irregularities within the claims database (n = 4).

### Comparisons With List Self-Report

"List" data were also available for those with "direct" self-report and claims data, in 96% of self-report and 96% of claims records. Percentage agreement between the "list" self-report measure and the claims measure (n = 482) was 91% with a kappa of 0.82. The comparison between the two self-report measures was also high, with 92% agreement and a kappa of 0.82 (n = 421) (Table 3).

## Discussion

There were several important findings from this research. First, concordance between "direct" self-report and pharmacy claims data for antidepressant drug use was high. Although good percentage agreement has been reported for nonspecific classes of medications,[10,11] many studies have presumed that agreement would not be as high for antidepressant medications.[19,20] In the only other similar study of antidepressant use, Saunders et al[12] studied patients just starting antidepressants. They found fair concordance (kappa = 0.33) at 1 month, but substantial concordance (kappa = 0.72) at 4 months. Their approach differed from ours in that they studied patients newly started on antidepressants, and used a list of only seven antidepressants. Our study measures are also slightly different; we define our claims measure to utilize a broader time window to capture longer prescriptions, and we use self-report instead of a phone interview to assess medication use, because phone-based survey data may be susceptible to interview bias. Our two dichotomous self-report questions comply with several important rules of good self-administration: they are short, simply worded, and do not require the respondent to recall behavior over a time interval, which may be especially difficult for depressed subjects. Their separated placement within the questionnaire is also helpful to minimize missing data.

Second, we found that most of the discordant cases could be resolved. That is, we found that discordant cases were related to characteristics of the measurement approach rather than to "errors" in either self-report or the claims data. For example, the "direct" self-report method identified use of medications intended primarily for the treatment of depressive disorders while claims records measured the purchase of medicines with antidepressant effects. Consequently, each measurement approach may be thought of as reflecting different situations. This understanding is important in deciding which data source to use and in part depends on the nature of the study to be undertaken.

Ideally, we would like to know if a patient was taking a drug that was prescribed to treat depression. In that light, the modest sensitivity of self-report to claims data (74%) could be viewed as a limitation of the self-report or the claims or both. To our surprise, our investigations revealed that self-report reflected our ideal more accurately than claims. Of 49 patients who denied taking antidepressants but had claims for antidepressants, 42 (86%) were determined to have responded accurately, with their antidepressants primarily prescribed for conditions other than depression.

Fifteen cases demonstrated the reverse error, where patients reported taking antidepressants but lacked a corresponding exposure in claims records. Only one of these cases exhibited the expected pattern of no explanation of the difference. The others were due primarily to low frequency (PRN) use, claims irregularities, and availability of samples. Although low-frequency use constitutes a classic adherence problem, the patient is still exposed to the drug as intended by the physician.

Third, most patients were willing to provide a list of their medications, possibly because of their

TABLE 3. Comparisons Among Direct Self-Report, Claims, and List Self-Report Measures

| Comparison | n | Percentage Agreement | Kappa (95% CI) |
|---|---|---|---|
| Direct self-report vs. claims records | 422 | 84.8 | 0.69 (0.56, 0.77) |
| List self-report vs. claims records | 482 | 91.3 | 0.82 (0.74, 0.92) |
| Direct self-report vs. list self-report | 421 | 91.7% | 0.82 (0.73, 0.90) |

participation in a research study. We had initially emphasized analyses of our direct questions because we believed that the time required to fill out such a list would be detrimental to its success as a survey item. Yet, the list proved to be a useful source of drug information. Face validity was shown by excellent agreement (92%) with "direct" self-report data, and 91% agreement with claims records. "List" reports answer the question of whether the patient is taking a drug that has antidepressant activity, even if there is another intention. In this respect, it is an accurate check on claims records. The "list" method offers detailed, current information and provides perspective on both claims and "direct" reports.

There are several relevant study limitations. First, classification of discrepancies was a subjective judgment. This was minimized by the use of multiple data sources and reviewers. Secondly, concordance may be lower for patients new to antidepressant medication.[12] However, we did not see any trends toward improved concordance over time or in the experimental group. Third, the study was conducted in primary care settings in the Boston metropolitan area, (an urban setting with high mental health utilization) as part of a research study (which may have influenced patients' knowledge of their medications). Concordance may differ elsewhere. Fourth, the results are influenced by the definition of the claims window, which was 0 to 90 days before the questionnaire date in this paper. Finally, our results could be mildly affected by our use of a broad set of antidepressants, which includes some medications with potential use as adjunctive depression treatment in primary care.

## Conclusion

We found that the patient is the best source of antidepressant drug exposure information. Although our data show that current antidepressant usage can be reliably determined from patients' responses to both questionnaire data and pharmacy claims records, patient self-report was determined to be more accurate than suspected. Investigations of disagreements between measures also suggest that these measures have differential validity. Self-report was found to identify medications intended primarily for the treatment of depressive disorders, whereas claims data identified use of medicines with antidepressant effects. An awareness of multiple indications is important, as claims were often shown to over-report antidepressant use in that regard. Thus, although the paradigm we established for analyzing discrepancies in antidepressant exposure applies broadly to other types of medications, the extent to which drugs such as antidepressants or beta-blockers have multiple uses may limit the usefulness of claims to assess drug exposure or adherence in many pharmaco-epidemiologic studies. Finally, percentage adherence with medications, especially among depressed patients, is best assessed through multiple sources and the same considerations that lead to discrepancies for drug exposure may also arise when considering adherence using multiple data sources.

## Acknowledgments

The authors thank the members of Tufts Health Plan, Clinical Quality Measurement Unit, namely, Diane Lancaster, RN, PhD, Andrew Bergman, Director, Quality Management, Tufts Health Plan for their support with the claims data. The project could not have been completed without the assistance of Doris Hernandez, Project Assistant.

## References

1. **Steiner JF, Prochazka AV.** The assessment of refill compliance using pharmacy records: methods, validity, and applications. J Clin Epidemiol 1997;50:105–116.

2. **Farmer KC.** Methods for measuring and monitoring medication regimen adherence in clinical trials and clinical practice. Clin Ther 1999;21:1074–1090.

3. **Rudd P, Byyny RL, Zachary V, et al.** The natural history of medication compliance in a drug trial: limitations of pill counts [see comments]. Clin Pharmacol Ther 1989;46:169–176.

4. **Cramer JA, Mattson RH, Prevey ML, et al.** How often is medication taken as prescribed? A novel assessment technique [published erratum appears in JAMA 1989 Se;15:262:1472] [see comments]. JAMA 1989; 261:3273–3277.

5. **Rudd P, Byyny RL, Zachary V.** Pill count measures of compliance in a drug trial: Variability and suitability. Am J Hypertens 1988;1:309–312.

6. **Fletcher SW, Pappius EM, Harper SJ.** Measurement of medication compliance in a clinical setting. Comparison of three methods in patients prescribed digoxin. Arch Intern Med 1979;139:635–638.

7. **Hugen PW, Langebeek N, Burger DM.** Assessment of adherence to HIV protease inhibitors: comparison and combination of various methods, including MEMS (electronic monitoring), patient and nurse report, and therapeutic drug monitoring. J Acquir Immune Defic Syndr 2002;30:324–334.

8. **Sjahid SI, van der Linden PD, Stricker BH.** Agreement between the pharmacy medication history and patient interview for cardiovascular drugs: the Rotterdam elderly study. Br J Clin Pharmacol 1998;45:591–595.

9. **Enlund H, Tuomilehto J, Turakka H.** Patient report validated against prescription records for measuring use of and compliance with antihypertensive drugs. Acta Med Scand 1981;209:271–275.

10. **Klungel OH, de Boer A, Paes AHP.** Influence of question structure on the recall of self-reported drug use. J Clin Epidemiol 2000;53:273–277.

11. **Lau HS, de Boer A, Beuning KS, et al.** Validation of pharmacy records in drug exposure assessment. J Clin Epidemiol 1997;50:619–625.

12. **Saunders K, Simon G, Bush T, Grothaus L.** Assessing the feasibility of using computerized pharmacy refill data to monitor antidepressant treatment on a population basis: a comparison of automated and self-report data. J Clin Epidemiol 1998;51:883–890.

13. **Rogers WH, Wilson IB, Bungay KM, et al.** Assessing the performance of a new depression screener for primary care (PC-SAD). J Clin Epidemiol 2002;55:164–175.

14. **American Hospital Formulary Service.** AHFS Drug Information 2001. Bethesda, MD: American Society of Health-System Pharmacists, Inc., 2001.

15. **Depression Guidelines Panel.** Depression in Primary Care: Volume 1. Detection and diagnosis. Volume 2. Treatment of major depression. Clinical Practice Guidelines Number 5. Rockville: US Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research 1993.

16. **Fleiss JL.** Measuring agreement between two judges on the presence or absence of a trait. Biometrics 1975;31:651–659.

17. **Landis JR, Koch GG.** The measurement of observer agreement for categorical data. Biometrics 1977;33:159–174.

18. **StataCorp.** Stata Reference Manual Release 6. 6.0 ed. College Station, TX: Stata Press, 1999.

19. **Rogers WH, Wells KB, Meredith LS, et al.** Outcomes for adult outpatients with depression under prepaid or fee-for-service financing [see comments]. Arch Gen Psychiatry 1993;50:517–525.

20. **Stewart AL, Greenfield S, Hays RD, et al.** Functional status and well-being of patients with chronic conditions: Results from the Medical Outcomes Study. JAMA 1989;262:907–913.