Agreement and accuracy between telephone self-report of drug use in older people and pharmaceutical claims data

Running title: Telephone self-report of drug use

Sabrina Winona Pit, PhD[1], Julie Ellen Byles, PhD[1], Jill Cockburn, PhD[1,2]

[1]School of Medicine and Public Health, Faculty of Health, University of Newcastle, Australia , [2]Professor Cockburn has passed away.

Corresponding author:

Dr Sabrina Pit, School of Medicine and Public Health, Faculty of Health, University of Newcastle, Australia, Room 266 DMB, Royal Newcastle Hospital, Newcastle NSW 2300, Australia. Ph: 0061 2 87 57 20 Fax: 0061 2 6687 5740

Email address: sabrina.pit@newcastle.edu.au/ sabrinapit@gmail.com

Alternate Corresponding Author:

Professor Julie Byles. Email address: Julie.byles@newcastle.edu.au

ABSTRACT

Objectives: To determine agreement between two measures of medication use in an elderly population: telephone interview self-report and pharmaceutical claims data. Classes of drugs commonly used in the elderly were selected for comparison.

Design: Agreement study.

Setting: General practice.

Participants: 566 community dwelling general practice patients aged 65 and over.

Measurements: Self-reported use of medicines, pharmaceutical claims data for different retrieval periods.

Results: 1094 people were eligible for the main study. Of these 697 people completed a follow up survey.Of these, 625 consented to the release of pharmaceutical claims data. A futher 59 participants were excluded from the analysis because they had a home visit instead of a telephone interview. The proportion of observed agreement between telephone self-report and the various retrieval periods was consistently high. Kappas showed good to very good agreement (0.75 and above) with retrieval periods of 30, 60 and 90 days for benzodiazepines, low risk NSAIDs and thiazide diuretics and most other drugs. Specificity of self-reported medication use compared with claims data, was consistently high across all drug classes, suggesting that people usually did not mention drugs which were not in the claims data. Sensitivity values varied according to drug class and according to retrieval period. Values were lower for NSAIDs than for benzodiazepines and thiazide diuretics. Decline in sensitivity with increased retrieval periods was most marked for benzodiazepines,

NSAIDs and low risk NSAIDs which are often used on an as needed basis. Positive predictive values increased with longer retrieval periods.

Conclusion: High agreement and accuracy was demonstrated for self-report of medicines use when asked over the telephone, when compared with pharmaceutical claims data. The telephone inventory method can be used in future studies for accurately measuring older peoples' drugs use when claims data are not available.

INTRODUCTION

Participant self-report methods are commonly employed to collect data on use of medicines[1-3]. These methods can include mail, telephone, face-to-face interviews, and touch-screen computers[4], Advantages and downsides of each method has been well documented[5-7]. In many cases, self-report methods are the only methods available as claims data may be incomplete or inaccessible. Methods using self-report can also measure additional information about the drugs, including directions and indications for use and compliance issues, and can measure use of OTC and complementary drugs which is not always available in prescription and pharmaceutical claims databases. The problem common to all self-report methods to measure drug use, is the reliance on accurate recall and reporting[7, 8].

Accuracy and recall depend on a number of issues. First, the method and manner in which questions are asked, interviewer skills, and the wording of questions impact on accuracy and participant recall[3, 7, 9]. Second, care should be taken when comparing methods of data collection because longer recall periods may lead to less accuracy. Some studies do not report the time frame, but ranges from four days[10] to one month[11, 12] are typical. Third, correspondence between self-report and reference standards varies per drug category[10, 13]. Consequently, it is imperative to examine each drug class separately and consider the features of each individual drug and its impact on recall and accuracy. Other factors which may impact on recall include the number of drugs used[11], regularity of drugs used[14], recency of drugs used[15], and participant demographics. Contrary to expectations, several studies have shown that older age does not seem to have a negative impact on recall accuracy[11, 12, 16, 17]. Finally, Stewart[11] found that patients receiving repeat prescriptions rather than new prescriptions were less accurate when self-reporting medication use (40% versus 80% sensitivity).

Since the 1960s, researchers have investigated methods for assessing the way in which people use medicines[7], especially in the area of compliance. Hence, there is a plethora of studies looking at self-report but few studies which have specifically measured the agreement and accuracy of telephone self-report when collecting data on medicines use. Moreover, most studies comparing telephone self-report with other methods do not focus on older people and focus on one particular drug class[3, 18] The latter has consequences for the development of questionnaires about medicine use because surveys can be adapted to a particular drug class which is likely to lead to better recall for that particular drug class.

Several studies have evaluated the validity of pharmaceutical claims databases among an elderly population and concluded that pharmaceutical records can be a useful source for measuring drug

exposure.[17, 19] In Australia, pharmaceutical claims data are processed through Medicare Australia under the Pharmaceutical Benefits Scheme (PBS) and the Repatriation PBS (for Department of Veterans' Affairs cardholders). PBS data do not cover all drugs used, and omits prescribed drugs, that are not subsidised through the PBS scheme, drugs provided in hospital, and OTC and complementary medicines[20]. Further, once people reach a certain monetary threshold (safety net), they are allocated a new identifying number[21], which applies to families and not individuals. In the case of the medications selected for this study, all are covered under the PBS scheme, but some usage may not be recorded for individuals who have reached the safety net or for individuals who bought NSAIDs over-the-counter. Thus PBS/RPBS data provide an appropriate and objective, but not perfect, comparison for self-reported use of medications.

A benefit of using pharmaceutical claims data is the lack of interviewer and recall bias once people have consented to the release of their data. However, access to this data is not always available to researchers. It can also be difficult to get a high consent rate for access to pharmaceutical claims databases. People may be happy to consent to an interview but providing ones personal Medicare number and other health services card numbers can be seen as a threat to some people. For example, an Australian study found that only 53% of women aged 70-75 consented to the release of their Medicare data[22]. Consenters tended to have higher levels of education and were in better health than non-consenters. Hence, relying on claims data can potentially lead to lower response rates, introduce bias, and decrease the power of studies. Another benefit of using patient self-report rather then prescription claim databases is that a person may collect prescription drugs but they may not actually take the drug which means they are not actually exposed to the drug.

The aim of this study was to determine the agreement and accuracy of medicines use of self-report data from a telephone interview in an elderly population by using pharmaceutical claims data as the reference standard. Classes of prescription drugs that are commonly used in the elderly were selected for comparison. Since some people will be currently taking drugs that were prescribed up to one year prior to interview, a second aim was to investigate what effect different retrieval periods prior to the telephone interview date have on the accuracy of telephone self-report as determined by the RPBS/PBS dataset.

METHODS

Participants were recruited from 20 general practices in the Hunter Region of Australia as part of a larger randomised controlled trial[23] which aimed to improve medicines use among older people through a multifaceted intervention. General practitioners were eligible if they had been based at their current practice for 12 months or more and practised 10 hours or more per week. Briefly, the intervention consisted of education (academic detailing, prescribing information and feedback), medication risk assessment, facilitation of medication review, and financial incentives. Community-dwelling people aged 65 years and over, attending participating general practices, were approached

by general practice office staff to seek consent to participate in the study. Participants who completed a telephone survey at at four month follow-up were eligible for participation in this comparison study.

Participants provided their Medicare number, Department of Veterans' Affairs card and other health concession cards numbers (if any), and gave consent for the release of claim details for the period of January 2000 until August 2003.

DATA COLLECTION

Trained interviewers asked participants to bring to the phone all the medicines they had taken in the previous seven days and prompted in advance: "Have you included eye drops, headache/pain killers, any medicines that help you sleep, fluid tables, laxatives or bowel medicines, antacids/ digestion, vitamins and herbal things". Participants were then asked to spell out the brand name of each medication. For each drug, the person was asked the strength and form of the medication, how they usually used it, how long they had used it, the medical condition they were using the drug for and whether the drug was prescribed by their own GP, other GP, specialist or the hospital, or whether it was bought "over-the-counter" at the chemist (without prescription), supermarket, naturopath, health food store, family or friend or elsewhere. When the participant had exhausted their list of medicines, the interviewer prompted again for the above mentioned drugs but also for creams, medicines used weekly or monthly, asthma puffers, eye and ear drops, cough or cold medicines and injections. For this analysis, only information on the medication name and the source of the medication were used. All drugs were classified according to the ATC classification System 2001[24].

Linking individual information to claims history for RPBS/PBS occurred by the combination of Medicare card number, date of birth and gender or if that failed by the combination of surname, first name, date of birth and gender. Medicare Australia was also provided with Department of Veterans' Affairs card numbers and health concession card numbers when available to improve matching with the datasets. Identifiers, date of birth and gender were cross checked with the data collected from the telephone survey to ensure data quality.

Ethics approval was received from the National Department of Veterans' Affairs, the University of Newcastle Human Research Ethics Committee and Hunter Health Research Ethics Committee.

STATISTICAL ANALYSES

The analysis was partly based on a combination of methods used by two previous studies[14, 21] Agreement was assessed by determining whether self-report and claims data identified drugs in the same ATC class. If a person used two benzodiazepines in the same ATC class, then the person was recorded as "positive" for using this class of drug. The observed agreement (proportion of

participants for whom claims data and self-report data agreed) was given for each drug category. Bias and prevalence adjusted Kappa coefficients (BPAC) and their 95% confidence intervals were calculated because kappa is affected in complex ways by the existence of bias between methods and by the distributions of data across the groupings that are used[25]. Sensitivity, specificity, and positive and negative predictive value of the information collected from telephone interviewing were also calculated, with claims data being employed as the reference standard. Several different claim periods were used to identify medication use from PBS claims (see Figure 1). In this way, the agreement and accuracy were assessed by searching claims data for periods of one month prior to interview to three months prior to interview for all drugs, and up to 12 months prior to interview for benzodiazepines, NSAIDS and thiazides which were used in the larger trial as outcome measures.

RESULTS

The doctors' participation rate was low (20/195). Of 1094 eligible people , 849 (78%) took part in the main study. Reasons for not taking part were: 139 were not interested, 36 were missed by office staff, 32 were mentally or physically unwell, 10 had language difficulties, 9 had no glasses or poor vision, and 19 had other reasons. 697 participated in the four month follow-up survey. Reasons for main study participants not taking part in the follow up survey were: 38 refused, 25 were non-contactable, three moved to a nursing home, four had died and 82 had other reasons. Of the 697 who were eligible for this comparison study, 625 (90%) consented to data linkage and 59 were excluded from further analysis because they had a home visit instead of a telephone interview due to poor hearing or other physical problems. leaving 566 individuals who provided data for this analysis. Nine individuals were not identified in the PBS dataset, of whom three did not report use of prescription medicines during the telephone interview. The median age of participants was 74 years (interquartile range: 70 to 78) and 41% were male.

ACCURACY OF TELEPHONE SELF-REPORT

The proportion of observed agreement between telephone self-report and the various retrieval periods was consistently high, with agreement reaching at least 81% for all drug categories at the 30 day retrieval period (Tables 1 and 2). Kappas showed good to very good agreement (0.75 and above) with retrieval periods of 30, 60 and 90 days for benzodiazepines, low risk NSAIDs and thiazide diuretics (Table 1) and for most medications in Table 2. Good agreement was observed for NSAIDs at 30 and 60 retrieval periods, but Kappas tended to be lower for NSAIDs at the 180 and 365 days retrieval period. Highest Kappas were observed for low risk NSAIDs and for Warfarin which showed perfect agreement with 60 day and 90 day retrieval.

Specificity was consistently high across all drug classes, suggesting that people did not usually report drugs which were not in the claims data. Sensitivity values varied according to drug class and according to retrieval period. Values were lower for NSAIDs than for benzodiazepines and thiazide

diuretics. Participants were least able to accurately report the use of low risk NSAIDs with a sensitivity of 56% when using a 90 day retrieval period. However these estimates were imprecise (with wide confidence intervals). Decline in sensitivity with increased retrieval periods was most marked for benzodiazepines, NSAIDs and low risk NSAIDs which are often used on an as needed basis. A prescription, filled a month prior to the telephone survey, was more likely to be in current use at the time of interview than a prescription that was filled a year ago.

Positive predictive values increased with longer retrieval periods. For example, a person reporting use of a benzodiazepine at telephone interview had a 53% percent chance of this use being verified in the claims data for the past 30 days, but had a 90% chance of their use being verified if claims retrieval was extended to include the entire past year.

DISCUSSION

This study estimated the agreement and accuracy of self-report medicines use data obtained from telephone interviews in an elderly population by comparing with pharmaceutical claims data as the reference standard. Good to excellent agreement was found between telephone self-report compared to pharmaceutical claims data and specificity values were high for all retrieval periods. Johnson and Vollmer[19] also reported high specificity for 18 prescription drug classes (>88%) among older people. This indicates that people usually did not mention drugs which were not in the claims data.

Sensitivity and positive predictive values were variable across different drug categories and retrieval periods, which is also supported by other studies[14, 19]. For example, two other studies found when comparing claims data with home-inventory data[14] or nursing home records[21] as the reference standard, that the number of weeks preceding data collection affected the validity of the claims data.

Sensitivity and predictive values for benzodiazepines were high for most retrieval periods, and there was very good overall observed agreement between self-reported use of benzodiazepines and claims data. Other studies have also reported high agreement and accurate self-reporting of benzodiazepines.[10, 14, 19, 21] The relatively low sensitivity rate for a 90 day retrieval (74%) can be explained by the fact that benzodiazepines can be used when needed and may not have been used in the week preceding the telephone interview even if they had been prescribed within the past 3 months. Likewise the low positive predictive value for reported benzodiazepine use may be explained by current use of medications that had been prescribed up to 12 months earlier.

NSAIDs are also commonly used on a when needed basis and can be purchased over the counter and hence do not show up in claims data. This effect was explored by excluding NSAIDs bought

over the counter to see what impact this restriction had on the outcome measures. As expected, the measures improved when OTC medicines were excluded (results not shown).

Cardiovascular drugs are commonly used to evaluate various methods of measuring drugs[16, 26, 27]. One study reported that a home inventory method was more likely to identify the use of diuretics than a prescription database[19]. However the retrieval period was 90 days for that study and higher utilisation may have been identified with longer retrieval. Many other studies found high to very high agreement and accuracy levels for self-reported use of thiazide diuretics[14] and diuretics[16, 17, 19].

The generalisability of the findings of this study is confined to older people living in the community. Because study subjects were participating in an underlying randomized controlled trial, participants could potentially be quite different from elders recruited from the general population. Participating in the broader study may have had an impact on the recall accuracy in this study by increasing patients awareness of drugs used. Also, since the underlying randomized controlled trial involved a multifaceted intervention to improve medicines use among older people, the intervention may have had an effect on the outcome of this study. The study shows that older people can find prescriptions drugs in their home and read and present the labels accurately. A major limitation is that the study could not measure whether older people could also do this for OTC and complementary drugs because these are not included in the prescription databases.

People who could not take part in a telephone interview due to physical problems such as hearing problems, and who's data were collected by home visit instead, were excluded from this analysis (about 10%).This may have biased the results. A sensitivity analysis was carried out to test whether including home visit data from these participants made a difference to the results and conclusions. All agreement and accuracy measures changed only a few percentage points The low GP response rate may also limit the generalisability of the findings as it is likely that participating doctors were those with an interest in medicine-related issues in elderly patients. However, practitioners demographic and practice characteristics were similar to national data (data available on request) and was in line with other high intensity intervention studies.[28, 29] A limitation of the measure of agreement is that medications that have the lowest prevalence of use have the highest level of agreement because the concordance is being dominated by non-users. Weighted Kappas only partially address this problem.

Participants did not need to know whether they were taking a specific drug nor did they need to recognise the brand or generic names of the drugs. An open-ended question was used to identify medicines use by asking people to bring all their medicines to the phone and spell out the names of each drug, with a general prompt for the types of drugs that people commonly forget. This is likely to have an impact on the accuracy of self report. Previous studies have found that open-ended questions lead to low sensitivity rates and hence have encouraged use of more specific questions to increase sensitivity rates[3, 9]. For example, one study[3] found that 50% of antibiotics were not

reported during an open-ended question, but a list of drugs and condition specific questions increased the sensitivity to 73%. These results seem at odds with this study since relatively high sensitivity rates were found in this study despite using only an open-ended question, with very general prompts. Two factors might explain this.

First, if the patient preferred, the interviewer would ring back once the patient had all the drugs collected. It may be that this gave people time to think about which drugs they were actually using and to collect all the drugs they had in the house. Second, other studies[3, 9] focused on one drug class, which creates a more difficult cognitive task for individuals who need to know either the drug type or indication. People do not always know the reasons for taking their drug. Therefore, an open-ended question asking people whether they have taken any antibiotics in the last six months[3] might lead to under-reporting if people are unsure about what type of drugs they are actually using. Hence, while asking people to list all their drugs might be more time consuming, it may also be easier as they simply list all the drugs they use.

The variability in results when using different claims retrieval periods could be seen as a cause of concern. However, the differences can be explained by the nature of the data captured in the PBS dataset. For example, sensitivity declined with longer retrieval intervals. Increasing retrieval periods leads to an increase in the number of drugs detected by the PBS data-set, however not all these drugs were used the week prior to the telephone survey. A prescription which was filled a month prior to the telephone survey was more likely to be in the week prior to the telephone survey than a prescription filled a year ago. The prescription filled a year ago may have run out or be no longer required, so these drugs would not be reported during the telephone survey and it would appear the patient was less accurate in reporting.

The completeness of the claims dataset also contributes to discrepancies between the two methods. It has been compulsory since April 2002 for patients to show their Medicare card when collecting prescriptions. Some pharmaceutical claims data from patients were collected prior to this date, especially when retrieval periods of 180 and 365 days were used, hence it may be that this had a negative impact on the completeness of the PBS dataset. However, it is more likely that the dataset had relative high data capture of prescriptions medicines for two reasons. First, most pharmacies were already asking patients to show their Medicare card as this has been encouraged by the government since 1 January 2001[30]. Second, almost all participants in this study were concession card holders (96%) which improves the completeness of the dataset.

Lau and colleagues[14] used similar methods as this study. Both studies found high specificities for most drugs. Sensitivities were similar across both studies for the 90 days retrieval period. The main difference in results between both studies was that Lau et al found lower sensitivities during the 30 day window periods ranging between 29% and 73% whereas this study found higher sensitivity rates ranging between 75% and 100%. The difference lies in the reference standard used: Lau and co-workers validated a pharmaceutical claims database with a home inventory as the reference

standard. By using the same reference standards, the results would be reasonably similar for both studies. Lau and co-workers argued that a retrieval period of 30 days was too short to accommodate the average supply of drugs in the Netherlands because in the Netherlands drugs were usually dispensed for a maximum period of 90 days due to reimbursement regulations. Johnson and Vollmer[19] used a 90 day window period to retrieve medications dispensed prior to the home inventory date. The authors also state that this may have been too short for analysis of medicines used for chronic conditions and too long for medicine used to treat acute conditions. Lau and colleagues argue that their results show the consequences for not adequately taking into account local dispensing policies when defining retrieval periods and measuring drug use. This study supports their statement. One Australian study recommends that claims data should be retrieved for the previous 12 weeks when estimating current medication use[21]. Although this finding was generally confirmed in this study, some drug categories might need a different retrieval period. The same authors[21] recommended that various drug groups may require different approaches when measuring drug use and suggested that medications which are used regularly, which should be consumed in a month, and are PBS benefits appear to be the most appropriate for studies using claims data. Different retrieval periods may be required in other countries because of prescribing differences. Also, care should be taken when comparing studies because as shown above, using different reference standards to test the validity of drug use can be confusing. The validity of using claims data as the reference standard may also change over-time if drug policies and prices are changed, prescribing practices change and more drugs become available over-the-counter.

In conclusion, high agreement and accuracy was demonstrated for self-report of medicines use when asked over the telephone, when compared with pharmaceutical claims data. The telephone inventory method can be used in future studies for accurately measuring older peoples' drugs use, in particular in situations where no claims data are available or where these are likely to be biased or incomplete.

REFERENCES

1. Australian Bureau of Statistics. National Health Survey 1995: Use of Medications. Canberra: Australian Bureau Of Statistics; 1999.

2.Hancock L, Walsh R, Henry DA, Redman S, Sanson-Fisher R. Drug use in Australia: a community prevalence study. Medical Journal of Australia. 1992 //;156:759-64.

3. Metlay JP, Hardy C, Strom BL. Agreement between patient self-report and a Veterans Affairs national pharmacy database for identifying recent exposures to antibiotics. Pharmacoepidemiology & Drug Safety. 2003;12(1):9-15.

4. Bonevski B, Sanson-Fisher RW, Campbell EM, Ireland MC. Do general practice patients find computer health risk surveys acceptable? A comparison with pen and paper method. Health Promotion Journal of Australia. 1997 //;7(2):100-6.

5.Silman A. Epidemiological studies: a practical guide. Cambridge: Cambridge University Press; 1995.

6.Bonevski BA. Increasing preventive care in general practice: An Examination [Doctoral dissertation]. Newcastle: University of Newcastle (NSW); 1996.

7.Farmer KC. Methods for measuring and monitoring medication regimen adherence in clinical trials and clinical practice. Clinical Therapeutics. 1999;21(6):1074-90; discussion 3.

8.Cockburn J. Variables related to antibiotic compliance in general practice patients: the application of behavioural science methodologies [Doctoral dissertation]. Newcastle: University of Newcastle (NSW); 1986.

9. Klungel OH, de Boer A, Paes AH, Herings RM, Seidell JC, Bakker A. Influence of question structure on the recall of self-reported drug use. Journal of Clinical Epidemiology. 2000;53(3):273-7.

10.Hancock L. Drug use in the Australian community: prevalence sociodemographic characteristics of users, and context of use. [Doctoral dissertation]. Newcastle: University of Newcastle (NSW); 1991.

11.Stewart M. The validity of an interview to assess a patient's drug taking. American Journal of Preventive Medicine. 1987;3(2):95-100.

12.Kehoe R, Wu SY, Leske MC, Chylack LT, Jr. Comparing self-reported and physician-reported medical history. American Journal of Epidemiology. 1994;139(8):813-8.

13.Paganini-Hill A, Ross RK. Reliability of recall of drug usage and other health-related information. American Journal of Epidemiology. 1982;116(1):114-22.

14.Lau HS, de Boer A, Beuning KS, Porsius A. Validation of pharmacy records in drug exposure assessment. Journal of Clinical Epidemiology. 1997;50(5):619-25.

15.West SL, Savitz DA, Koch G, Strom BL, Guess HA, Hartzema A. Recall accuracy for prescription medications: self-report compared with database information. American Journal of Epidemiology. 1995;142(10):1103-12.

16.Klungel OH, de Boer A, Paes AH, Herings RM, Seidell JC, Bakker A. Agreement between self-reported antihypertensive drug use and pharmacy records in a population-based study in The Netherlands. Pharmacy World & Science. 1999;21(5):217-20.

17.Sjahid SI, van der Linden PD, Stricker BH. Agreement between the pharmacy medication history and patient interview for cardiovascular drugs: the Rotterdam elderly study. British Journal of Clinical Pharmacology. 1998;45(6):591-5.

18.Wang PS, Benner JS, Glynn RJ, Winkelmayer WC, Mogun H, Avorn J. How well do patients report noncompliance with antihypertensive medications?: A comparison of self-report versus filled prescriptions. Pharmacoepidemiology & Drug Safety. 2004;13(1):11-9 .

19.Johnson RE, Vollmer WM. Comparing sources of drug data about the elderly. Journal of the American Geriatrics Society. 1991;39(11):1079-84.

20. Robertson J, Fryer JL, O'Connell DL, Smith AJ, Henry DA. Limitations of Health Insurance Commission (HIC) data for deriving prescribing indicators. Medical Journal of Australia. 2002;176(9):419-24.

21.King MA, Purdie DM, Roberts MS. Matching prescription claims with medication data for nursing home residents: implications for prescriber feedback, drug utilisation studies and selection of prescription claims database. Journal of Clinical Epidemiology. 2001;54(2):202-9.

22.Young AF, Dobson AJ, Byles JE. Health services research using linked records: who consents and what is the gain? Australian & New Zealand Journal of Public Health. 2001;25(5):417-20.

23.Pit S, Byles J, Henry D, Holt L, Hansen V, Bowman D. Effectiveness of an educational Better Use of Medicines Program on medicines use, falls and quality of life for older patients in general practice: Cluster randomised controlled trial. Under editorial review MJA. 2006.

24.WHO Collaborating Centre for Drug Statistics Methodology. Anatomical therapeutic chemical (ATC) classification index with defined daily doses (DDDs). Oslo: WHO Collaborating Centre for Drug Statistics Methodology; 2001 2001.

25.Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol. 1993;46(5):423-9.

26.Choo PW, Rand CS, Inui TS, Lee ML, Cain E, Cordeiro-Breault M, et al. Validation of patient reports, automated pharmacy records, and pill counts with electronic monitoring of adherence to antihypertensive therapy. Medical Care. 1999;37(9):846-57, 1999 Sep.

27.Smith NL, Psaty BM, Heckbert SR, Tracy RP, Cornell ES. The reliability of medication inventory methods compared to serum levels of cardiovascular drugs in the elderly. Journal of Clinical Epidemiology. 1999;52(2):143-6.

28. Kable S. An application of interactive computer programs to promote adherence to clinical practice guidelines for childhood asthma in general practice [Doctoral dissertation]. Newcastle: University of Newcastle (NSW); 2004.

29.Reeve JF, Peterson GM, Rumble RH, Jaffrey R. Programme to improve the use of drugs in older people and involve general practitioners in community education. J Clin Pharm Ther. 1999;24(4):289-97.

30.Department of Health and Ageing. Pharmaceutical Benefits Scheme. Medicare cards and fairer benefits. 2003 8 May 2002. [cited 2003 18 October 2003]; Available from: http://www.health.gov.au/haf/ime/

31. Robertson J, Fryer JL, O'Connell DL, Sprogis A, Henry DA. The impact of specialists on prescribing by general practitioners. Med J Aust. 2001;175(8):407-11.

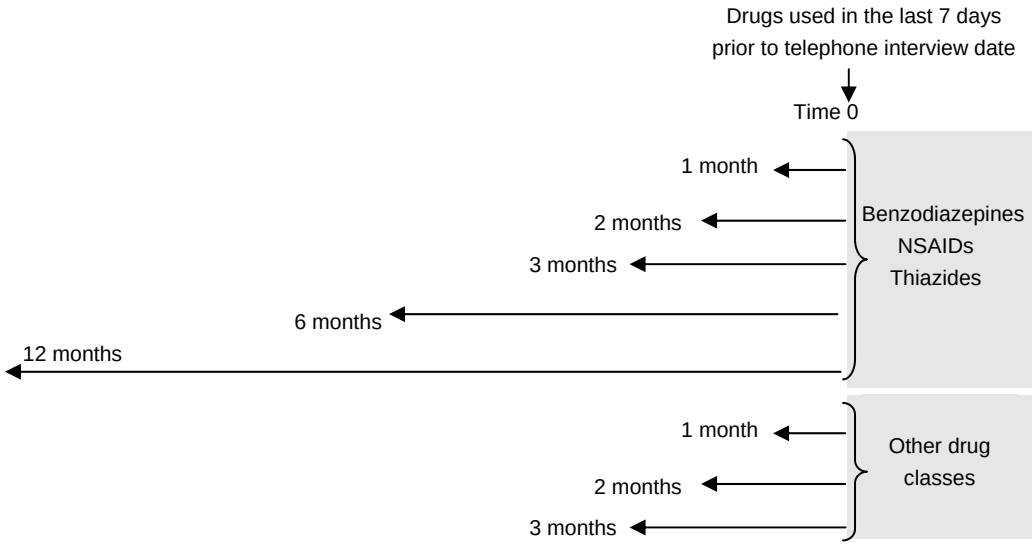

Figure 8.1 Fixed window periods (cut-off points) used to determine which prescriptions were picked up from the pharmacist by patient in months preceding telephone interview date:

- *30 days, because PBS prescriptions are usually prescribed for one month (or a standard course of treatment) [20, 31]*
- *60 days, because when patients are not compliant a prescription may last longer, for example six weeks rather than the intended four week period*
- *90 days, because previous research has shown that a 90 day retrieval period yields high sensitivity, specificity and positive predictive validity[14, 21].*

For benzodiazepines, thiazide diuretics, and NSAIDs which were the primary outcome measures in the main study, the following time periods were also explored to gain an improved understanding of the self-report measure:

- *26 weeks, because repeat prescriptions are usually intended to last up to six months*
- *52 weeks, because patients can claim repeat prescriptions up to one year after prescription date.*

Table 1 Prevalence of telephone self-report and pharmaceutical claims data, prevalence and bias adjusted weighted Kappa coefficient, sensitivity, specificity, and positive and negative predictive value of telephone self-report compared with pharmaceutical claims data for prescription benzodiazepines, NSAIDs and thiazide diuretics among 566 patients

Prescription drugs picked up from chemist:	Prevalence				Observed agreement	Prevalence and bias adjusted kappa		Sensitivity		Specificity		Positive predictive value	Negative predictive value
	Telephone self-report		Pharmaceutical claims database										
	95% CI		95%CI			K	95%CI	%	95%CI	%	95%CI	%	%
Benzodiazepines	10%	(8 to 13)											
- 30 days prior to telephone survey			6%	(5 to 9)	94%	0.88	(0.80 to 0.97)	88%	(75 to 97)	95%	(93 to 97)	53%	99%
- 60 days prior to telephone survey			9%	(7 to 12)	95%	0.90	(0.82 to 0.98)	80%	(69 to 91)	96%	(95 to 98)	68%	98%
- 90 days prior to telephone survey			10%	(8 to 13)	94%	0.89	(0.81 to 0.97)	74%	(62 to 85)	97%	(95 to 98)	71%	97%
- 180 days prior to telephone survey			13%	(10 to 16)	94%	0.88	(0.80 to 0.95)	66%	(55 to 77)	98%	(97 to 99)	83%	95%
- 365 days prior to telephone survey			16%	(14 to 20)	92%	0.84	(0.76 to 0.92)	57%	(47 to 67)	99%	(98 to 100)	90%	92%
NSAIDs (including COX-2s)	22%	(19 to 26)											
- 30 days prior to telephone survey			16%	(13 to 20)	88%	0.75	(0.67 to 0.84)	81%	(73 to 89)	89%	(86 to 92)	58%	96%
- 60 days prior to telephone survey			25%	(21 to 28)	88%	0.75	(0.68 to 0.84)	71%	(64 to 79)	93%	(91 to 96)	78%	91%
- 90 days prior to telephone survey			29%	(25 to 33)	86%	0.71	(0.63 to 0.79)	64%	(56 to 71)	95%	(92 to 97)	83%	86%
- 180 days prior to telephone survey			37%	(33 to 41)	80%	0.61	(0.53 to 0.69)	54%	(47 to 61)	96%	(94 to 98)	89%	78%
- 365 days prior to telephone survey			48%	(44 to 52)	73%	0.45	(0.37 to 0.53)	45%	(39 to 51)	98%	(96 to 100)	95%	66%
Low risk NSAID(ibuprofen/diclofenac	2%	(1 to 4)											
- 30 days prior to telephone survey			2%	(1 to 4)	99%	0.98	(0.90 to 1.06)	75%	(51 to 100)	99%	(99 to 100)	69%	100%
- 60 days prior to telephone survey			3%	(2 to 4)	99%	0.97	(0.89 to 1.05)	67%	(43 to 91)	100%	(99 to 100)	77%	99%
- 90 days prior to telephone survey			3%	(2 to 5)	98%	0.96	(0.88 to 1.04)	56%	(33 to 79)	100%	(99 to 100)	77%	99%
- 180 days prior to telephone survey			5%	(3 to 7)	97%	0.94	(0.85 to 1.02)	41%	(22 to 59)	100%	(99 to 100)	85%	97%
- 365 days prior to telephone survey			6%	(5 to 9)	96%	0.91	(0.83 to 0.99)	33%	(18 to 49)	100%	(99 to 100)	92%	96%
Thiazide diuretics	19%	(16 to 23)											
- 30 days prior to telephone survey			10%	(8 to 13)	89%	0.77	(0.69 to 0.85)	89%	(81 to 97)	88%	(86 to 91)	46%	99%
- 60 days prior to telephone survey			16%	(13 to 19)	91%	0.82	(0.73 to 0.90)	82%	(74 to 90)	93%	(90 to 95)	68%	97%
- 90 days prior to telephone survey			19%	(16 to 23)	94%	0.87	(0.79 to 0.95)	83%	(76 to 90)	96%	(94 to 98)	84%	96%
- 180 days prior to telephone survey			22%	(19 to 26)	95%	0.89	(0.81 to 0.98)	81%	(74 to 88)	99%	(98 to 100)	95%	95%
- 365 days prior to telephone survey			27%	(23 to 31)	92%	0.84	(0.76 to 0.92)	71%	(64 to 78)	100%	(99 to 100)	99%	90%

Table 2 Prevalence of telephone self-report and pharmaceutical claims data, prevalence and bias adjusted weighted Kappa coefficient, sensitivity, specificity, and positive and negative predictive value of telephone self-report compared with pharmaceutical claims data for some common prescription drugs among 566 patients

Prescription drugs picked up from chemist:	Prevalence				Observed agreement	Prevalence and bias adjusted kappa		Sensitivity		Specificity		Positive predictive value	Negative predictive value
	Telephone self-report		Pharmaceutical claims database										
	95% CI		95%CI			K	95%CI	%	95%CI	%	95%CI	%	%
Psychotropic	22%	(19 to 26)											
- 30 days prior to telephone survey			16%	(13 to 19)	89%	0.79	(0.70 to 0.87)	88%	(81 to 94)	89%	(87 to 92)	61%	98%
- 60 days prior to telephone survey			21%	(17 to 24)	92%	0.83	(0.75 to 0.92)	85%	(78 to 91)	94%	(91 to 96)	77%	96%
- 90 days prior to telephone survey			23%	(20 to 27)	92%	0.83	(0.75 to 0.92)	81%	(74 to 88)	95%	(93 to 97)	82%	95%
Any diuretic	32%	(28 to 36)											
- 30 days prior to telephone survey			16%	(13 to 19)	81%	0.62	(0.54 to 0.70)	90%	(84 to 96)	79%	(76 to 83)	46%	98%
- 60 days prior to telephone survey			24%	(21 to 28)	87%	0.74	(0.65 to 0.82)	88%	(83 to 94)	86%	(83 to 90)	67%	96%
- 90 days prior to telephone survey			30%	(26 to 33)	91%	0.81	(0.73 to 0.90)	88%	(83 to 93)	92%	(89 to 94)	82%	95%
Any antihypertensive	70%	(66 to 73)											
- 30 days prior to telephone survey			56%	(52 to 61)	86%	0.72	(0.64 to 0.80)	99%	(98 to 100)	69%	(63 to 75)	81%	98%
- 60 days prior to telephone survey			67%	(63 to 71)	94%	0.87	(0.79 to 0.95)	97%	(95 to 99)	86%	(81 to 91)	93%	94%
- 90 days prior to telephone survey			69%	(65 to 72)	95%	0.90	(0.81 to 0.98)	97%	(96 to 99)	90%	(86 to 94)	95%	93%
H_2 antagonist/proton pump inhibitors	34%	(30 to 38)											
- 30 days prior to telephone survey			26%	(22 to 30)	89%	0.79	(0.70 to 0.87)	95%	(91 to 98)	87%	(84 to 91)	75%	98%
- 60 days prior to telephone survey			33%	(29 to 37)	94%	0.88	(0.80 to 0.96)	92%	(88 to 96)	95%	(93 to 97)	90%	96%
- 90 days prior to telephone survey			36%	(32 to 40)	94%	0.87	(0.79 to 0.95)	88%	(84 to 93)	96%	(95 to 98)	93%	94%

Prescription drugs picked up from chemist:	Prevalence				Observed agreement	Prevalence and bias adjusted kappa		Sensitivity		Specificity		Positive predictive value	Negative predictive value
	Telephone self-report	95% CI	Pharmaceutical claims database	95%CI		K	95%CI	%	95%CI	%	95%CI	%	%
Hypoglycaemics	11%	(9 to 14)											
- 30 days prior to telephone survey			8%	(6 to10)	96%	0.93	(0.84 to 1.00)	98%	(93 to102)	96%	(95 to 98)	68%	100%
- 60 days prior to telephone survey			10%	(8 to 13)	98%	0.97	(0.89 to 1.05)	98%	(95 to 102)	98%	(97 to 100)	87%	100%
- 90 days prior to telephone survey			11%	(8 to 13)	99%	0.98	(0.89 to 1.06)	98%	(95 to 102)	99%	(98 to 100)	91%	100%
Digoxin	7%	(5 to 9)											
- 30 days prior to telephone survey			2%	(1 to 4)	95%	0.91	(0.82 to 0.98)	100%	(100 to 100)	95%	(94 to 97)	33%	100%
- 60 days prior to telephone survey			4%	(3 to 6)	97%	0.93	(0.85 to 1.01)	91%	(80 to 103)	97%	(95 to 98)	54%	100%
- 90 days prior to telephone survey			6%	(4 to 8)	98%	0.95	(0.87 to 1.03)	91%	(81 to 101)	98%	(97 to 99)	74%	99%
Allopurinol	9%	(7 to 12)											
- 30 days prior to telephone survey			3%	(2 to 5)	94%	0.87	(0.78 to 0.95)	86%	(71 to 101)	93%	(92 to 96)	35%	99%
- 60 days prior to telephone survey			6%	(4 to 8)	96%	0.91	(0.82 to 0.98)	86%	(78 to 99)	96%	(94 to 98)	60%	99%
- 90 days prior to telephone survey			8%	(6 to 10)	97%	0.95	(0.87 to 1.02)	91%	(83 to 99)	98%	(97 to 99)	79%	99%
Warfarin	6%	(4 to 8)											
- 30 days prior to telephone survey			4%	(3 to 6)	98%	0.97	(0.88 to 1.05)	100%	(100 to 100)	98%	(97 to 99)	70%	100%
- 60 days prior to telephone survey			5%	(4 to 8)	100%	0.98	(0.90 to 1.06)	100%	(100 to 100)	99%	(99 to 100)	91%	99%
- 90 days prior to telephone survey			6%	(4 to 8)	100%	1.00	(0.91 to 1.08)	97%	(91 to 103)	100%	(100 to100)	100%	100%

ACKNOWLEDGEMENTS

We thank all the general practitioners (in particular Dr Parker Magin and Dr Peter Hopkins) and their patients for participating; Vibeke Hansen for managing and assisting in data collection, Deborah Bowman for research assistance; Professor Tony Smith, Professor David Henry, Lucy Holt and general practitioners for providing extensive advise on developing the content of the Medication Review Checklist; Chris Bonner for advise and providing the Therapeutic Flags manual; Professor Paul Glasziou for developing the grant; the peer reviewers of the thesis for their comments; The Royal Australian College of General Practitioners for approving parts of the research as a clinical audit and as a Continuing Professional Development activity; The National Prescribing Service for approving Practice Incentive Payments; Medicare Australia for providing data and all the advisory group members (in particular Judith Mackson (National Prescribing Service) and Professor Kichu Nair) for advise, interviewers and statisticians.

Financial Disclosures: National and Health Medical Research Council, Australia funded the study. SP received a University of Newcastle Postgraduate Scholarship as a PhD student. The researchers are independent from the funders. The researchers in the Faculty of Health are members of the Hunter Medical Research Institute.

Author Contributions: Professor Jill Cockburn (deceased) conceived, designed, analysed and interpreted the results of the study. SWP and JEB designed, analysed, interpreted the results of the study. JC and JB developed the grant. SWP

managed and conducted the study and drafted the manuscript. JEB conceived the study, and revised the manuscript. All authors (SWP, JEB) contributed to the final version. SWP and JEB are guarantors.

Sponsor's Role: The funding body played no role in the formulation of the design, methods, subject recruitment, data collection, analysis, or preparation of this paper.