Practice of Epidemiology

Heightened Attention to Medical Privacy: Challenges for Unbiased Sample Recruitment and a Possible Solution

Ann B. Nattinger*, Liliana E. Pezzin, Rodney A. Sparapani, Joan M. Neuner, Toni K. King, and Purushottam W. Laud

* Correspondence to Dr. Ann B. Nattinger, Division of General Internal Medicine, Medical College of Wisconsin, General Internal Medicine—CLCC, 9200 West Wisconsin Avenue, Suite 5400, Milwaukee, WI 53226 (e-mail: anatting@mcw.edu).

Subject recruitment for epidemiologic studies is associated with major challenges due to privacy laws now common in many countries. Privacy policies regarding recruitment methods vary tremendously across institutions, partly because of a paucity of information about what methods are acceptable to potential subjects. The authors report the utility of an opt-out method without prior physician notification for recruiting community-dwelling US women aged 65 years or older with incident breast cancer in 2003. Participants ($n = 3,083$) and possibly eligible nonparticipants ($n = 2,664$) were compared using characteristics derived from billing claims. Participation for persons with traceable contact information was 70% initially (2005–2006) and remained over 90% for 3 follow-up surveys (2006–2008). Older subjects and those living in New York State were less likely to participate, but participation did not differ on the basis of socioeconomic status, race/ethnicity, underlying health, or type of cancer treatment. Few privacy concerns were raised by potential subjects, and no complaints were lodged. Using opt-out methods without prior physician notification, a population-based cohort of older breast cancer subjects was successfully recruited. This strategy may be applicable to population-based studies of other diseases and is relevant to privacy boards making decisions about recruitment strategies acceptable to the public.

breast neoplasms; confidentiality; data collection; ethics; Health Insurance Portability and Accountability Act; jurisprudence; privacy; quality of health care

Abbreviations: CMS, Centers for Medicare and Medicaid Services; HIPAA, Health Insurance Portability and Accountability Act.

Editor's note: Invited commentaries on this article appear on pages 645 and 648, and the authors' response appears on page 651.

Many barriers to clinical health-care research have been identified, and these barriers have impeded the translation of new research into the clinical practice, decision-making, and organizational changes needed to improve health (1). One of the major challenges affecting investigators conducting clinical, epidemiologic, and health-services research in the past 5 years has been the heightened attention to medical privacy and resulting restrictions imposed by the Health Insurance Portability and Accountability Act (HIPAA) in the United States (2), the Data Protection Act in the United Kingdom (3), and similar privacy laws in other countries (4–7). There is a strong belief on the part of clinical investigators that these rules have made research more difficult, have delayed time to study completion, and have led to nonrepresentative samples, studies' being abandoned, and added cost (8–11). The prestigious US Institute of Medicine recently agreed with these concerns, concluding that the HIPAA Privacy Rule impedes important health research and recommending that the federal government develop a new approach to protecting privacy in health research (12).

Privacy-related recruitment issues which have been consistently noted by researchers to be problematic include the desirability of requiring physician consent or notification to initiate subject contact and whether recruitment procedures should utilize a more restrictive ''opt-in'' approach (after

receiving an introductory letter, investigators are permitted to follow up only with potential subjects who actively indicate interest, such as by returning a postcard or telephoning an office) or whether an "opt-out" approach (after initial contact, investigators are permitted to follow up with all subjects except those who opt out of further contact) is satisfactory.

Investigators are concerned about specific methods for identifying and recruiting a patient sample partly because of the desirability of having a satisfactory participation rate, which is often considered a surrogate measure of a high-quality study (13). Even more important is the question of whether persons agreeing to participate in a study are representative of the larger group of persons eligible for participation—that is, whether nonparticipation bias is present in the study sample. Nonparticipation bias refers to systematic errors introduced into the study when reasons for study participation are associated with the health aspects of interest (13). A better understanding of the participation rates and generalizability of samples that can be achieved with various recruitment procedures is of interest to policymakers, institutional review boards, privacy boards, investigators, and the public—all of whom have a vested interest in balancing the protection of privacy with the desire to have medical research conducted in such a way as to yield valid results.

In this paper, we report on the participation rate, acceptability, and generalizability of a novel method for generating a representative sample by using an opt-out approach without prior physician approval to recruit older subjects with a chronic illness (breast cancer) for a study involving longitudinal survey and claims review elements. Although this study focuses on breast cancer, the sample selection and recruitment methods employed are of interest for recruiting subjects with other health conditions.

MATERIALS AND METHODS

Study cohort

The overall study goals were to relate initial breast cancer care processes to 5-year outcomes of mortality, recurrence, and quality of life. We aimed to recruit a representative sample of community-dwelling women from 4 large and diverse US states who were between the ages of 65 and 89 years when they underwent surgery for incident breast cancer in 2003. To identify a potentially eligible sample of women with breast cancer, we applied a previously described and validated claims-based prediction algorithm (14) to administrative claims from the Medicare program, a government-funded health insurance program for US residents aged 65 years or older. Subjects considered for inclusion were required to be enrolled in Medicare Part A (hospital coverage) and Part B (physician coverage) and not enrolled in a Medicare health maintenance organization (a managed care organization providing care through contracted providers; individual service claims are not submitted to Medicare) for calendar year 2003, to have had a breast cancer operation in 2003 according to the prediction algorithm, and to have an associated Medicare surgeon claim.

Potentially eligible subjects were excluded if their address or telephone number could not be traced or if they were deceased by the time of contact, had a diagnosis of dementia or a long-term-care facility stay of 100 days or more in 2003, were physically unable to participate, were residing in a long-term-care facility at the time of contact, did not speak English or Spanish, or did not confirm a diagnosis of incident breast cancer in 2003 once contacted.

Potential subjects identified through the breast cancer algorithm ($n = 8,742$) were sent a letter via US mail by a mailing contractor selected by the Centers for Medicare and Medicaid Services (CMS), the US government agency that administers Medicare. This letter, printed on CMS letterhead, explained that our research team wished to contact them regarding possible participation in a research study involving breast surgery, clarified that participation was entirely voluntary and would not affect Medicare benefits, and provided a telephone number that could be called if the beneficiary did not wish to participate, as well as a form that could be returned by mail if the beneficiary did not wish to participate or had questions. The letter was accompanied by a trifold brochure that outlined the study protocol. All materials were sent in both Spanish and English. Batch mailings occurred in September and December 2005.

Telephone recruitment began in October 2005 and was concluded by October 2006. An academically affiliated survey center conducted the field work, which consisted of telephone tracing, subject recruitment, and survey interviewing. Protocols were developed for each field procedure, including establishment of a toll-free respondent line and a call history notes database and development of protocols for structured verbal consent, proxy response, and refusal. A computer-assisted telephone interviewing system allowed for precoded questions as well as open-ended questions. Completed interviews averaged 32 minutes in length.

Subjects enrolled in the study were offered participation in 3 follow-up telephone survey waves conducted in June 2006–March 2007 (wave 2), June–November 2007 (wave 3), and May–September 2008 (wave 4). Subjects were mailed a $20 check after participation in each wave of the survey. The study protocol was approved by the Medical College of Wisconsin institutional review board and the CMS Privacy Board. Access to Medicare data and approval for the participant contact methods was facilitated by the Research Data Assistance Center, a CMS contractor which receives funding to assist researchers in using Medicare data for their work (http://www.resdac.umn.edu).

Measures of respondent generalizability

The study team had access to Medicare claims with certain patient identifiers in order to identify potentially eligible subjects. Once it was determined which subjects were not participating in the study, an appropriately deidentified data set was employed for the purpose of comparing participants and nonparticipants. This data set contained previously collected Medicare-based study data elements, including demographic factors, type of breast surgery (mastectomy or breast-conserving surgery), comorbidity score (determined using the method of Klabunde et al. (15)),
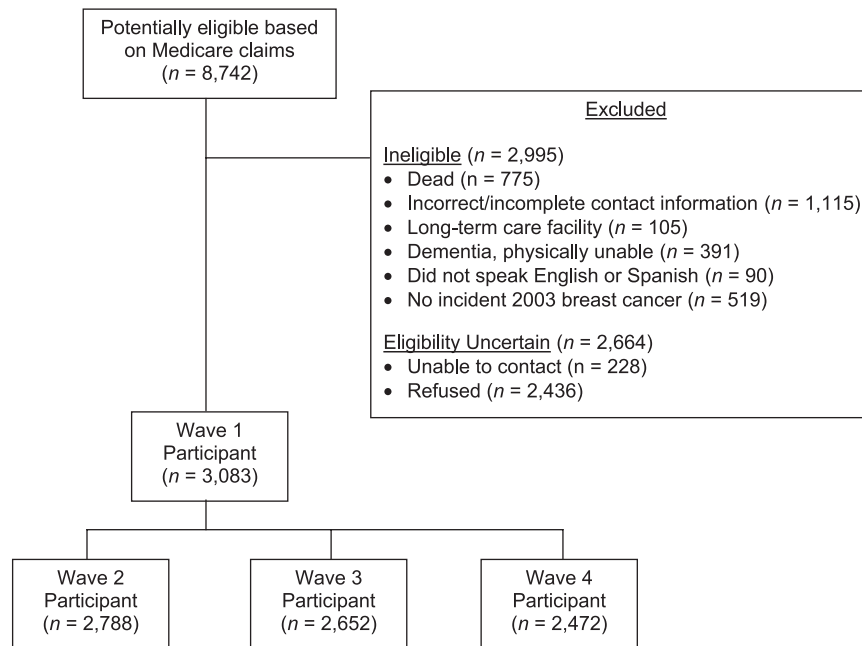
Figure 1.   Distribution of 8,742 community-dwelling US women aged 65 years or older who were initially identified as potentially eligible for a breast cancer study based on Medicare claims, 2005–2006. Women who could not be contacted or gave soft refusals during wave 2 (June 2006–March 2007) or wave 3 (June–November 2007) were offered the opportunity to participate in future waves. In wave 2, 93 women were ineligible, 109 could not be reached, and 93 refused. In wave 3, 196 women were ineligible, 123 could not be reached, and 112 refused. In wave 4 (May–September 2008), 310 women were ineligible, 172 could not be reached, and 129 refused. The numbers of persons who were ineligible, could not be contacted, and refused in waves 2–4 are each incremental to wave 1.

hospital volume (annual number in 2002–2003 of Medicare breast cancer operations for the hospital at which the patient underwent breast surgery), and surgeon volume (annual number in 2002–2003 of Medicare breast cancer operations conducted by the surgeon who operated on the subject).

Statistical analysis

The participation rates proposed by the American Association for Public Opinion Research (Deerfield, Illinois) were computed, using estimates that adjust for eligibility (16). The "response rate 4" (which counts complete and partial responses) was computed according to the following formula:

$$\frac{n_P}{n_P + p_E(n_R + n_U)},$$

where p_E is the probability of eligibility, calculated as

$$\frac{n_P}{n - (n_R + n_U)},$$

and where n_P = the number participating, n_R = the number of refusers, n_U = the number with unknown eligibility, and n = the number of subjects in the sample initially considered for contact. For estimating the proportion eligible among persons with unknown eligibility, the same proportion was estimated as among those with eligibility determined (51%), but a sensitivity analysis was conducted using a 20% higher and 20% lower proportion eligible among persons with unknown eligibility.

Standard univariate descriptive statistics were used to compare participants with nonparticipants. A multiple logistic regression model was constructed to determine independent predictors of response status. Although the claims-based algorithm had been previously validated (14), its use to generate a survey cohort was, to our knowledge, novel. For this reason, we assessed the positive predictive value of the algorithm when used for developing a sampling frame, by computing the percentage of otherwise-eligible subjects who reported that their breast surgery in 2003 was not for an incident breast cancer that year.

RESULTS

Participation and acceptability

Of the 8,742 subjects initially mailed a letter by the CMS contractor, 2,995 subjects were determined to be ineligible, 228 had contact information identified but were never reached by telephone, and 2,436 declined participation (Figure 1). Of those declining participation, 836 (34.3% of refusers and 9.6% of the potentially eligible cohort) contacted the CMS contractor to decline based on the letter; the remainder declined when contacted by telephone. Those participating numbered 3,083 women, of whom 3,008 provided complete interviews for the wave 1 survey (2005–2006).

The response rate was 70% for complete or partial interviews and 68% for complete interviews only. These response rate calculations assumed that participants and nonparticipants were equally likely to be eligible for study participation. The sensitivity analysis yielded a 66% response rate if 20% more of the nonparticipants were actually eligible and 74% if 20% fewer of the nonparticipants were actually eligible. If subjects with untraceable contact information had not been excluded, the response rate would have been 57%.

By waves 2–4, some subjects had become ineligible because of death or movement into a long-term-care facility. Among the remaining subjects, participation in waves 2–4 (adjusted for eligibility) remained at 93%, 92%, and 90%, respectively (Figure 1). All materials provided to subjects and potential subjects included the contact information of the principal investigator. During the first 3 years of the study, the research team fielded 91 communications from or on behalf of subjects, including 3 from treating physicians. Of these, 22 communicants had concerns about privacy (including 2 who wished to terminate participation), 6 wished to terminate participation for other reasons, 35 wished to offer additional information or ask for medical advice, and 28 had questions related to logistical issues, such as the reimbursement for participation. Although 2 physicians expressed concern about the lack of their consent to contact subjects, the subjects did not raise this issue explicitly. The concerns appeared to be satisfied in the great majority of cases by the provision of more information about the study procedures or approvals. No subject lodged a complaint with the institutional review board, the contact information for which was also provided.

Comparison of participants and nonparticipants

In Table 1, the 3,083 participants are compared with the 2,664 possibly eligible nonparticipants (2,436 refusers and 228 persons never reached). In univariate analyses, participants were modestly more likely to be of white race/ethnicity, to reside in a state other than New York, to reside in a less urban area, and to be cared for by a high-volume surgeon. There were no differences in participation based on the per capita income of the zip code (US Postal Service address code) of residence, the underlying health conditions of the subject, the type of breast surgery undergone, or the case volume of the hospital at which the surgery was performed.

In a multiple regression model predicting participation, the only factors that remained significant independent predictors of participation were younger age and state of residence, as seen in the last column of Table 1.

Validation of claims-based breast cancer algorithm

Although the sensitivity of the breast cancer algorithm as applied to this cohort could not be determined, it was possible to estimate the positive predictive value, defined as the percentage of algorithm-positive subjects who confirmed during the wave 1 survey the presence of an incident breast cancer in 2003. When the algorithm was developed (14), the specificity was estimated at 99.95%, and the positive predictive value based on a 1995 tumor registry breast cancer cohort was 88.1% (95% confidence interval: 85.6, 90.7). Based on the otherwise-eligible patients who responded to survey questions to confirm an incident breast cancer in 2003, the algorithm positive predictive value for this study was 85.7%. Of those persons who were "false-positive" on the basis of the algorithm, 60.1% stated that their first occurrence of breast cancer was in a year prior to 2003 and 39.9% reported that they had never had breast cancer.

DISCUSSION

In this article, we report on a method for using an opt-out approach without prior physician approval to develop a population-based sample of older breast cancer subjects for a longitudinal study. The recruitment methods yielded satisfactory participation rates, which is notable, since older subjects and retired subjects tend to have lower rates of participation in clinical studies (17, 18). Based on the modest numbers of potential subjects refusing participation initially, the very small numbers declining follow-up participation, and the small numbers of concerns raised, the method we employed appears to have been acceptable to the target population. The comparison of participants with possibly eligible nonparticipants showed differences with respect to age and state of residence but not for race/ethnicity, socioeconomic status of the postal code area in which the person resided, measures of initial treatment, or underlying health status. Although these comparisons did not exclude differences in unmeasured characteristics (19), the results are certainly encouraging with regard to evidence of an unbiased sample.

It is a particular challenge to recruit such unbiased samples for population health studies in this era, in which government privacy rules have become common internationally. Institutions face substantial challenges in attempting to maintain follow-up of patients they have treated (20), and it can be even more challenging to gain the trust of subjects when attempting to recruit a population-based sample of individuals, who may not have had any contact with the researcher's institution. While legislatively mandated registries exist in many industrialized countries for a few diseases such as cancer (9, 21), this is not true for most diseases. In addition, the research that can be conducted using deidentified or limited data sets from these registries is restricted to questions that can be answered using the set of general variables that is typically collected. Conducting more detailed studies by recruiting subjects already identified through these registries presents challenges that are similar to those faced by clinical investigators more generally, including heterogeneity in the policies of the relevant institutional review boards and privacy boards with regard to data access and recruitment procedures (5, 9, 22). The methods that were successful in this study may provide a helpful precedent for such oversight entities.

In this study, we employed an "opt-out" approach to patient recruitment, as opposed to an "opt-in" approach. The opt-out approach was deemed acceptable by the relevant

Table 1.   Available Data on Characteristics of Women Aged 65 Years or Older Who Had Incident Breast Cancer in 2003, by Survey Participation, United States, 2005–2006

Characteristic	Participants (*n* = 3,083)		Nonparticipants (*n* = 2,664)		Multiple Logistic Regression	
	No.	%	No.	%	Odds Ratio	95% Confidence Interval
Age group, years						
65–69	972	31.5	570	21.4	4.1	3.0, 5.5
70–74	984	31.9	663	24.9	2.3	1.7, 3.1
75–89	1,127	36.6	1,431	53.7		
Race/ethnicity						
White	2,900	94.1	2,443	92.0		
Black	111	3.6	106	4.0	0.9	0.7, 1.2
Other	70	2.3	107	4.0	0.7	0.5, 1.0
State						
California	821	26.6	688	25.8	1.3	1.1, 1.6
Florida	995	32.3	754	28.3	1.3	1.1, 1.5
Illinois	602	19.5	500	18.8	1.3	1.1, 1.5
New York	665	21.6	722	27.1		
Size of US Census Metropolitan Statistical Area						
≥250,000 persons	2,011	65.2	1,832	68.8	0.9	0.8, 1.1
≤250,000 persons or rural area	1,072	34.8	832	31.2		
Per capita income[a]						
Poorest tertile	972	31.5	890	33.4		
Middle tertile	1,046	33.9	823	30.9	1.1	0.9, 1.3
Wealthiest tertile	1,065	34.5	951	35.7	1.0	0.9, 1.2
Breast surgery						
Mastectomy	2,028	65.8	1,764	66.2		
Breast-conserving surgery	1,055	34.2	900	33.8	1.0	0.8, 1.1
Comorbidity score[b]						
0	2,019	65.5	1,727	64.8		
1	760	24.7	639	24.0	1.0	0.9, 1.2
≥2	304	9.9	298	11.2	0.9	0.7, 1.0
Annual Medicare hospital volume[c], no. of cases						
0–11	564	23.0	551	25.1		
12–23	821	33.5	706	32.1	1.1	0.9, 1.3
≥24	1,065	43.5	941	42.8	1.0	0.9, 1.2
Annual Medicare surgeon volume[c], no. of cases						
0–5	931	32.5	928	36.4		
6–11	889	31.0	773	30.3	1.1	0.9, 1.3
≥12	1,049	36.6	848	33.3	1.2	1.0, 1.4

[a] Based on the mean per capita income of the subject's zip code.
[b] Calculated using the method of Klabunde et al. (15).
[c] The treating hospital could not be determined for all subjects. In addition, the volume of cases could not be reliably determined for 4 hospitals and 209 surgeons located in states other than the 4 study states.

oversight boards and appeared acceptable to this patient population, as demonstrated by the satisfactory participation rate, the outstanding follow-up rates, and the low number of concerns raised. However, some privacy boards have required an opt-in approach in similar circumstances (9, 10). Opt-in requirements have been shown to yield lower participation rates (12). For example, a newly instituted opt-in requirement for a well-established registry of acute coronary syndrome patients led to a decline in participation rates from 96% to 34% (20). A similar requirement in the United Kingdom led to a 16% participation rate in a survey study (3). In a randomized trial of opt-in versus opt-out recruitment in Australia, the participation rates were 47% for opt-in recruitment as compared with 67% for opt-out recruitment (4). In addition, the opt-in sample was biased toward persons with greater baseline risk and those preferring an active role in health decision-making.

We considered the 12.7% of subjects for whom we had no traceable address or telephone contact to be ineligible for this telephone-based study. Some might believe that these persons should be considered nonrespondents, although they did not have the opportunity to be contacted for study participation. In this case, the study response rate would be 57%, which is still a reasonable rate for elderly patients with whom we had no prior connection. Persons for whom we had incomplete or inaccurate contact data were the same age as others in the larger sample but were more likely to be of black or another race/ethnicity, to live in a larger urban area, and to reside in California. Therefore, the factors associated with having traceable contact information were not the same as the factors associated with study participation once a person had been contacted.

The Medicare claims algorithm employed to identify potentially eligible subjects performed well, with only a modest number of ''false-positive'' cases being identified. The algorithm had been developed and validated on claims data dating to the mid-1990s, yet the positive predictive value of the algorithm to identify 2003 subjects remained within the prior confidence limits. This was true despite the fact that we had available, on average, only 3 years of prior Medicare claims with which to identify and remove prevalent cases from the potentially eligible sample. Therefore, the majority of ''false-positive'' cases identified through the algorithm were women with incident breast cancer in some prior year. The use of 4 or more prior years of data to remove prevalent cases would further improve the positive predictive value of the algorithm and may be appropriate for future studies.

The use of Medicare claims to develop the sampling framework for a study poses strengths and weaknesses. The most obvious weaknesses are the limitation to subjects aged 65 years or older and the fact that claims-based algorithms are not available for identifying patients with every type of health condition. In addition, the sample selected of necessity excluded subjects enrolled in Medicare health maintenance organizations, which skewed the sample slightly toward older and sicker beneficiaries (23). These weaknesses were offset by several advantages. Because 97% of US residents aged 65 years or older are eligible for Medicare coverage and because almost all US hospitals

Table 2. Distribution of a 5% Sample of Cases From the Medicare Chronic Condition Warehouse[a], United States, 2002 and 2007[b]

Condition	2002		2007	
	No.	%	No.	%
Bone/joint disease				
Hip/pelvic fracture	13,032	0.8	12,705	0.8
Osteoporosis	172,008	10.8	190,470	12.1
Rheumatoid arthritis or osteoarthritis	298,184	18.8	323,005	20.5
Cancer				
Colorectal	15,802	1.0	14,955	0.9
Endometrial	2,232	0.2	2,344	0.3
Female breast	27,551	3.0	29,711	3.3
Lung	14,156	0.9	15,253	1.0
Prostate	44,851	6.6	45,453	6.0
Cardiovascular disease				
Acute myocardial infarction	18,152	1.1	15,215	1.0
Atrial fibrillation	108,564	6.8	122,890	7.8
Heart failure	281,699	17.8	272,398	17.2
Ischemic heart disease	500,563	31.6	508,181	32.2
Neuropsychiatric illness				
Alzheimer's disease	68,062	4.3	82,559	5.2
Alzheimer's disease and related disorders or senile dementia	156,177	9.8	171,645	10.9
Depression	165,528	10.4	196,845	12.5
Stroke/transient ischemic attack	78,098	4.9	69,031	4.4
Miscellaneous				
Cataract	375,731	23.7	343,305	21.7
Chronic kidney disease	109,218	6.9	185,664	11.8
Chronic obstructive pulmonary disease	162,997	10.3	172,701	10.9
Diabetes mellitus	337,470	21.3	411,315	26.0
Glaucoma	145,396	9.2	159,415	10.1
Overall 5% sample				
Total	1,586,139		1,579,292	
Females	905,726	57.1	888,407	56.3
Males	680,413	42.9	690,885	43.7

[a] The Chronic Condition Warehouse is a database created to make Medicare claims data more readily available to researchers studying 21 predefined conditions. More details are available from the Centers for Medicare and Medicaid Services (http://www.ccwdata.org/datadoc.php).

[b] Data were available for a 5% random sample from 2000 forward and for a 100% sample from 2005 forward. Sample sizes provided are for a 5% random sample for 2 representative years. Beneficiaries may have been counted in more than 1 condition category.

and physicians accept Medicare insurance, a sample of Medicare beneficiaries is more representative than samples derived from individual hospitals or medical practices. Another advantage of a Medicare-based sample is the fact that the claims data from a given calendar year are typically

available approximately 6 months after the year ends. Unless a registry has rapid case ascertainment capabilities (9), registry data are often not available for 3–4 years after an index event. While neither time frame is rapid enough for studies requiring patient contact during the acute phase of an illness, the Medicare claims may permit somewhat quicker availability of data for analysis or subject contact.

Methods for utilizing Medicare data to conduct patient-oriented research continue to slowly improve (24). The CMS contracts with the Research Data Assistance Center, which assists investigators in using Medicare data for their research (http://www.resdac.umn.edu). A Chronic Condition Data Warehouse (25) has been created to make Medicare claims data more readily available to researchers studying 21 predefined conditions (Table 2). For subjects identified in the Chronic Condition Data Warehouse, claims data from inpatient, physician, outpatient, hospice, home health agency, and durable medical equipment bills are available. The imminent availability of Medicare Part D (pharmaceutical coverage) claims records for prescription drugs (http://www.resdac.umn.edu), beginning with 2006 claims for all beneficiaries, will further enhance the ability of researchers to use Medicare data to study health outcomes and conduct related types of clinical research.

In summary, the heightened attention to medical privacy, especially as manifested by privacy laws which have been passed in many countries, has led to major challenges in subject recruitment for many types of studies. Recruitment issues are often handled heterogeneously by different privacy boards. Requirements for physician notification or consent to initiate subject contact and requirements for an opt-in approach to recruitment have led to particular concern and can result in more biased samples. The present study demonstrated the successful use of an opt-out approach without prior physician notification in the recruitment of a representative sample of subjects treated by multiple different hospitals and physicians. This recruitment approach proved quite acceptable to the target population. This method may offer advantages for investigators studying other conditions, and also should assist privacy boards in determining what recruitment methods are acceptable to the public more generally.

The authors thank Gerald F. Riley of the Centers for Medicare and Medicaid Services for his helpful contribution to this study.

Conflict of interest: none declared.

REFERENCES

1. Sung NS, Crowley WF Jr, Genel M, et al. Central challenges facing the national clinical research enterprise. *JAMA*. 2003;289(10):1278–1287.
2. Annas GJ. Medical privacy and medical research—judging the new federal regulations. *N Engl J Med*. 2002;346(3):216–220.
3. Ward HJ, Cousens SN, Smith-Bathgate B, et al. Obstacles to conducting epidemiological research in the UK general population. *BMJ*. 2004;329(7460):277–279.
4. Trevena L, Irwig L, Barratt A. Impact of privacy legislation on the number and characteristics of people who are recruited for research: a randomised controlled trial. *J Med Ethics*. 2006;32(8):473–477.
5. Willison DJ, Emerson C, Szala-Meneok KV, et al. Access to medical records for research purposes: varying perceptions across research ethics boards. *J Med Ethics*. 2008;34(4):308–314.
6. Sheikh AA. The Data Protection (Amendment) Act, 2003: the Data Protection Directive and its implications for medical research in Ireland. *Eur J Health Law*. 2005;12(4):357–372.
7. Conti A. The recent Italian Consolidation Act on privacy: new measures for data protection. *Med Law*. 2006;25(1):127–138.
8. Ness RB; Joint Policy Committee, Societies of Epidemiology. Influence of the HIPAA Privacy Rule on health research. *JAMA*. 2007;298(18):2164–2170.
9. Beskow LM, Sandler RS, Weinberger M. Research recruitment through US central cancer registries: balancing privacy and scientific issues. *Am J Public Health*. 2006;96(11):1920–1926.
10. Greene SM, Geiger AM, Harris EL, et al. Impact of IRB requirements on a multicenter survey of prophylactic mastectomy outcomes. *Ann Epidemiol*. 2006;16(4):275–278.
11. Goss E, Link MP, Bruinooge SS, et al. The impact of the privacy rule on cancer research: variations in attitudes and application of regulatory standards. *J Clin Oncol*. 2009;27(24):4014–4020.
12. Institute of Medicine, National Academy of Sciences. *Beyond the HPIAA Privacy Rule: Enhancing Privacy, Improving Health Through Research*. Washington, DC: National Academies Press; 2009.
13. Galea S, Tracy M. Participation rates in epidemiologic studies. *Ann Epidemiol*. 2007;17(9):643–653.
14. Nattinger AB, Laud PW, Bajorunaite R, et al. An algorithm for the use of Medicare claims data to identify women with incident breast cancer. *Health Serv Res*. 2004;39(6):1733–1749.
15. Klabunde CN, Legler JM, Warren JL, et al. A refined comorbidity measurement algorithm for claims-based studies of breast, prostate, colorectal, and lung cancer patients. *Ann Epidemiol*. 2007;17(8):584–590.
16. American Association for Public Opinion Research. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. Revised 2008*. Deerfield, IL: American Association for Public Opinion Research; 2008. (http://www.aapor.org/standarddefinitions). (Accessed November 22, 2009).
17. Dunn KM, Jordan K, Lacey RJ, et al. Patterns of consent in epidemiologic research: evidence from over 25,000 responders. *Am J Epidemiol*. 2004;159(11):1087–1094.

18. Eagan TM, Eide GE, Gulsvik A, et al. Nonresponse in a community cohort study: predictors and consequences for exposure-disease associations. *J Clin Epidemiol*. 2002;55(8): 775–781.

19. McLeod TG, Costello BA, Colligan RC, et al. Personality characteristics of health care satisfaction survey non-respondents. *Int J Health Care Qual Assur*. 2009;22(2):145–156.

20. Armstrong D, Kline-Rogers E, Jani SM, et al. Potential impact of the HIPAA privacy rule on data collection in a registry of patients with acute coronary syndrome. *Arch Intern Med*. 2005;165(10):1125–1129.

21. Sanson-Fisher R, Carey M, Mackenzie L, et al. Reducing inequities in cancer care: the role of cancer registries. *Cancer*. 2009;115(16):3597–3605.

22. Gostin LO, Nass S. Reforming the HIPAA Privacy Rule: safeguarding privacy and promoting research. *JAMA*. 2009; 301(13):1373–1375.

23. Riley G, Zarabozo C. Trends in the health status of Medicare risk contract enrollees. *Health Care Financ Rev*. 2006;28(2): 81–95.

24. Lamont EB, Herndon JE II, Weeks JC, et al. Measuring disease-free survival and cancer relapse using Medicare claims from CALGB breast cancer trial participants (companion to 9344). *J Natl Cancer Inst*. 2006;98(18):1335–1338.

25. Research Data Assistance Center. *Available CMS Data*. Minneapolis, MN: School of Public Health, University of Minnesota; 2009. (http://www.resdac.umn.edu/Available_CMS_Data.asp). (Accessed November 22, 2009).