Published in final edited form as:

Ann Epidemiol. 2011 September ; 21(9): 706–709. doi:10.1016/j.annepidem.2011.03.007.

Who Doesn't Authorize the Linking of Survey and Administrative Health Data? A General Population-based Investigation

Timothy J. Beebe, Ph.D.^{1,2}, Jeanette Y. Ziegenfuss, Ph.D.^{1,2}, Sarah Jenkins, M.S.³, Lindsey R. Haas, B.A.^{1,2,3}, and Michael E. Davern, Ph.D.⁴

¹Division of Health Care Policy and Research, Mayo Clinic, Rochester Minnesota

²Survey Research Center, Department of Health Sciences Research, Mayo Clinic, Rochester Minnesota

³Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, Minnesota

⁴Department of Public Health Research, NORC at the University of Chicago

Keywords

Community surveys; Survey methods; Personal Health Records; Authorization; HIPAA; Health Insurance Portability and Accountability Act

INTRODUCTION

Linking survey and administrative data is becoming increasingly popular among epidemiologists and other health researchers because it brings about a confluence of virtues that each individual method cannot supply on its own and broadens the range of research questions that can be explored (1). Germane health data regulations such as the Health Insurance Portability and Accountability Act (HIPAA) of 1996 require individual authorization before such linking can take place. Individual authorization for linking is likely to be less than complete, raising concerns over what some have deemed “authorization bias” whereby those who authorize differ systematically from those who do not in terms of attributes important to health-related investigations (2). While there is evidence that authorizers and non-authorizers differ in terms of age, gender, income, and education, (3-10) past investigations are limited by their use of specialized patient populations and/or specific socio-demographic sub-groups (11), potentially compromising the generalizability of their findings. No one has attempted to characterize non-authorizers in terms of their health and health care seeking behavior in the context of an area-based population survey to our knowledge. The aim of this research is to describe who is willing to complete a survey but not willing to authorize use of their protected information in terms of their sociodemographic characteristics, medical and surgical diagnoses, comorbidities, and health care utilization using linked survey and administrative health data from a general population-based study.

© 2011 Elsevier Inc. All rights reserved.

Corresponding Author: Timothy J. Beebe, Ph.D., Associate Professor of Health Services Research, Mayo Clinic College of Medicine, 200 First Street SW, Rochester, MN 55905, beebe.timothy@mayo.edu Tel : 507-538-4606, Fax : 507-284-1180.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

METHODS

The data on authorization status (authorized vs. not) come from a sequential mixed mode, mail and telephone survey of non-institutionalized residents of Olmsted County (home to Mayo Clinic), Minnesota aged 18 and older as identified in a purchased list-based sample of residents. An experiment was embedded within this survey to determine the impact of including a HIPAA Authorization Form (HAF) on response rates where half of the sample was randomized to be sent a HAF and the other half were not sent a HAF. The findings from this study and further details of the methodology are available elsewhere (12). For the present study, we focus just on the half of the sample that was randomized to be sent a HAF and who responded to the survey request. Authorizers were defined as those who completed a survey and returned a completed HAF while non-authorizers were those who filled out a survey but did NOT return a completed HAF.

The sampling frame for the HAF Experiment Study was linked to the Rochester Epidemiology Project (REP), the medical record system for health care providers to residents of Olmsted County (13). Each health care provider in Olmsted County uses a unit medical record system whereby all data collected on an individual are assembled in one place. Each participating site also solicits and documents permission from patients for their records to be used. Currently, 95% of patients have granted this permission. In addition to medical diagnoses, hospital admissions and surgical procedures, the REP includes demographic information such as age, gender, and race. Overall, almost 97% of the cases in the sample file were matched to members in the REP database. A more detailed account of the linked survey and administrative health data has already been published(14).

For purposes of analysis and with the exception of age, the background and clinical variables were dichotomized as shown in Table 1. The dichotomization cut-offs were chosen to facilitate analysis and interpretation, informed by the items' marginal distributions to identify natural breaks, and designed to accord with prior authorization studies in Olmsted County using the REP (5). Associations between individual characteristics and whether or not a HAF was signed and returned were estimated with logistic regression models. Odds ratios and 95% confidence intervals were reported. All analyses were performed using SAS v. 9.1 software. A p -value of < 0.05 was regarded as statistically significant.

RESULTS

A total of 1,574 surveys or telephone interviews were completed. Of these, 1,357 returned a completed HAF for an overall authorization rate of 86.2%. Comparing authorizers ($n=1,357$) to non-authorizers ($n=217$) demonstrated that subjects in the latter category were significantly younger, healthier (with respect to comorbidities), and less likely to have had three or more clinic office visits in 2005 and 2006 (see Table 1). In a logistic model including all variables, age remained significant along with clinic office visits and ER admissions; however, the Charlson score was no longer significant.

DISCUSSION

Overall, 86.2% of those that completed a mailed survey or a telephone interview authorized linking to administrative health data, suggesting that some element of the HAF (evoked privacy concerns or signature requirement) was compelling some to not authorize, as observed by Singer (15-17). We found non-authorizers to be significantly younger, healthier (with respect to comorbidities), and less likely to have had three or more recent clinic office visits than authorizers. However, once we controlled for age in our multivariate analyses, we saw no differences in recorded comorbidities between authorizers and non-authorizers and

past ER use was found to significantly diminish the likelihood of authorization. For the most part, our findings accord with those observed in other studies of this nature, albeit with more specialized populations (3, 5, 7, 8, 10, 18-20).

Researchers should be aware of the potential bias brought about by the exclusion of younger subjects who use different types of health care services at various levels when analyzing their population-based linked survey and administrative data. It may be that our finding with respect to age is due to a cohort effect. Respondents in the category least likely to authorize (18-34) were born between 1975 and 1993 and came of age during a period of increased commodification of personal information, an explosion in the use of the Internet and social networking sites where personal data are shared as a matter of course, the emergence of HIPAA regulations, and the reporting of several high profile privacy breaches. Thus, people in this age cohort may be more attuned to the risks of authorizing data linkages. With respect to health care utilization, our research presents a more complicated picture. Individuals more connected to clinics (3+ visits) were more likely to authorize in both the bivariate and multivariate analyses while those who use the ER were found to be less likely to authorize once we controlled for age and other factors. Clearly, the presence of authorization bias in the realm of health care utilization depends on how this concept is operationalized and how the analyses are undertaken.

References

1. Klassen AF, Lee SK, Barer M, Raina P. Linking survey data with administrative health information: characteristics associated with consent from a neonatal intensive care unit follow-up study. *Can J Public Health*. 2005; 96(2):151-4. [PubMed: 15850038]
2. Yawn BP, Yawn RA, Geier GR, Xia Z, Jacobsen SJ. The impact of requiring patient authorization for use of data in medical records research. *J Fam Pract*. 1998; 47(5):361-5. [PubMed: 9834771]
3. Bolcic-Jankovic D, Clarridge BR, Fowler FJ Jr, Weissman JS. Do characteristics of HIPAA consent forms affect the response rate? *Med Care*. 2007; 45(1):100-3. [PubMed: 17279027]
4. Harris T, Cook DG, Victor C, Beighton C, Dewilde S, Carey I. Linking questionnaires to primary care records: factors affecting consent in older people. *J Epidemiol Community Health*. 2005; 59(4):336-8. [PubMed: 15767389]
5. Jacobsen SJ, Xia Z, Champion ME, et al. Potential effect of authorization bias on medical record research. *Mayo Clin Proc*. 1999; 74(4):330-8. [PubMed: 10221460]
6. Laaksonen M, Aittomaki A, Lallukka T, et al. Register-based study among employees showed small nonparticipation bias in health surveys and check-ups. *J Clin Epidemiol*. 2008; 61(9):900-6. [PubMed: 18486445]
7. Murdoch M, Pietila DM, Partin MR. Randomized trial showed an 'embedded' survey strategy optimized authorization rates compared to two 'after survey' strategies in veterans with PTSD. Accepted for publication.
8. Partin MR, Burgess DJ, Halek K, et al. Randomized trial showed requesting medical records with a survey produced a more representative sample than requesting separately. *J Clin Epidemiol*. 2008; 61(10):1028-35. [PubMed: 18550333]
9. Tate AR, Calderwood L, Dezateux C, Joshi H. Mother's consent to linkage of survey data with her child's birth records in a multi-ethnic national cohort study. *Int J Epidemiol*. 2006; 35(2):294-8. [PubMed: 16373376]
10. Young AF, Dobson AJ, Byles JE. Health services research using linked records: who consents and what is the gain? *Aust N Z J Public Health*. 2001; 25(5):417-20. [PubMed: 11688619]
11. Huang N, Shih SF, Chang HY, Chou YJ. Record linkage research and informed consent: who consents? *BMC Health Serv Res*. 2007; 7:18. [PubMed: 17291357]
12. Beebe TJ, Talley NJ, Camilleri M, Jenkins SM, Anderson KJ, Locke GR 3rd. The HIPAA authorization form and effects on survey response rates, nonresponse bias, and data quality: a randomized community study. *Med Care*. 2007; 45(10):959-65. [PubMed: 17890993]

13. Melton LJ 3rd. History of the Rochester Epidemiology Project. *Mayo Clin Proc.* 1996; 71(3):266–74. [PubMed: 8594285]
14. Beebe TJ, Ziegenfuss JY, St Sauver JL, et al. HIPAA authorization and survey nonresponse bias. *Med Care.* 2011 forthcoming.
15. Singer E. Informed consent: consequences for response rate and response quality in social surveys. *Am Sociol Rev.* 1978; 43(2):144–62. [PubMed: 655499]
16. Singer E, Mathiowetz NA, Couper MP. The impact of privacy and confidentiality concerns on survey participation. The case of the 1990 U.S Census. *Public Opin Q.* 1993; 57(4):465–a-482.
17. Singer E, Van Hoewyk J, Neugebauer RJ. Attitudes and Behavior: The Impact of Privacy and Confidentiality Concerns on Participation in the 2000 Census. *Public Opin Q* %R 10.1086/377465. 2003; 67(3):368–384.
18. Angus VC, Entwistle VA, Emslie MJ, Walker KA, Andrew JE. The requirement for prior consent to participate on survey response rates: a population-based survey in Grampian. *BMC Health Serv Res.* 2003; 3(1):21. [PubMed: 14622444]
19. 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):1125–9. [PubMed: 15911725]
20. Dunn KM, Jordan K, Lacey RJ, Shapley M, Jinks C. Patterns of consent in epidemiologic research: evidence from over 25,000 responders. *Am J Epidemiol.* 2004; 159(11):1087–94. [PubMed: 15155293]

Table 1

Effects of Selected Respondent Sociodemographic and Clinical Characteristics by Authorization Status.

Variable	Authorizers n (%)	Non-authorizers n (%)	Univariate models ¹ OR (95% CI)	Full model ² OR (95% CI)
N	1357	217		
<u>Background</u>				
Gender				
Female (ref)	677 (49.9%)	103 (47.5%)	ref	ref
Male	680 (50.1%)	114 (52.5%)	0.91 (0.68, 1.21)	0.91 (0.67, 1.23)
Age				
18 to < 35 (ref)	181 (13.3%)	51 (23.5%)	ref	ref
35 to < 50	427 (31.5%)	85 (39.2%)	1.42 (0.96, 2.09)	1.49 (1.00, 2.23) *
50 to < 65	402 (29.6%)	63 (29%)	1.80 (1.20, 2.71) **	1.92 (1.24, 2.97) **
65 +	347 (25.6%)	18 (8.3%)	5.43 (3.08, 9.57) **	5.79 (3.11, 10.76) **
Race				
Other (ref)	145 (10.7%)	30 (13.8%)	ref	ref
White	1212 (89.3%)	187 (86.2%)	1.34 (0.88, 2.05)	1.22 (0.79, 1.91)
<u>Clinical</u>				
Charlson Score ^{3, 5}				
<2 (ref)	987 (72.7%)	175 (80.6%)	ref	ref
2 or More	370 (27.3%)	42 (19.4%)	1.56 (1.09, 2.23) *	0.95 (0.63, 1.43)
ER Admission ⁴				
No (ref)	963 (71%)	142 (65.4%)	ref	ref
Yes	394 (29%)	75 (34.6%)	0.78 (0.57, 1.05)	0.60 (0.42, 0.86) **
Hospital Admission ⁴				
No (ref)	1059 (78%)	169 (77.9%)	ref	ref
Yes	298 (22%)	48 (22.1%)	0.99 (0.70, 1.40)	0.99 (0.66, 1.48)
Clinic office visit ⁴				
<3 (ref)	412 (30.4%)	96 (44.2%)	ref	ref
≥ 3	945 (69.6%)	121 (55.8%)	1.82 (1.36, 2.44) **	1.81 (1.30, 2.52) **
Procedures ⁵				
None (ref)	531 (39.1%)	92 (42.4%)	ref	ref
≥1	826 (60.9%)	125 (57.6%)	1.15 (0.86, 1.53)	0.76 (0.55, 1.06)

NOTE: Authorization status is coded 1 for "Authorizers" and 0 for "Non-authorizers" in the logistic regression models.

¹ Univariate logistic regression models.² Multivariable logistic regression model including all predictors.³ Severity weighted.⁴ During 2005-2006.⁵ Considering data 1995 – 2006.

*
 $p < 0.05$;

**
 $p < 0.01$