

# It takes patience and persistence to get negative feedback from patients

---

Elizabeth West<sup>1</sup>, Professor of Applied Social Science and Director of Research, Mary Seacole Building, School of Health and Social Care, University of Greenwich, Avery Hill Road, Eltham, London SE9 2UG  
[e.west@gre.ac.uk](mailto:e.west@gre.ac.uk)

David N. Barron, University Reader in Organizational Sociology, Saïd Business School and Jesus College, University of Oxford, Park End Street, Oxford OX1 7HP. [David.barron@jesus.ox.ac.uk](mailto:David.barron@jesus.ox.ac.uk)

Rachel Reeves, Principal Research Fellow, School of Health and Social Care, University of Greenwich, Avery Hill Road, Eltham, London SE9 2UG [R.Reeves@gre.ac.uk](mailto:R.Reeves@gre.ac.uk)

Denise Hawkes, Senior Lecturer in Economic Sociology, Business School, University of Greenwich, Royal Naval College, Park Row, Greenwich. [D.Hawkes@gre.ac.uk](mailto:D.Hawkes@gre.ac.uk)

---

<sup>1</sup> Corresponding author

## ABSTRACT

**Background:** Patient experience surveys are increasingly used to gain information about the quality of healthcare. This paper investigates whether patients who respond early or late, and before and after reminders, to a large national survey of in-patient experience differ in systematic ways in how they evaluate the care they received.

**Methods:** The English national in-patient survey of 2009 obtained data from just under 70,000 patients. We analyse their responses to the question “Overall, how did you rate the care you received” in relation to the time they took to respond and whether or not they had had a reminder, using statistical models designed to examine the length of time taken for an event to occur, known as “failure time regression models”.

**Results:** 41 per cent of patients responded after the first questionnaire and 11 per cent after reminders. Those who were least positive in their evaluation of care replied on average 3.1 days later than the most positive. However, the main dividing line was between patients who responded to the initial mailing or to the reminders. Even controlling for other factors that influence the likelihood of an early response, those who respond after the initial mailing were more likely to be positive about the care they received.

**Conclusion:** This study, using a large national dataset, shows that bias towards a positive evaluation of care could be introduced if the length of time that patients are allowed to respond is truncated or if reminders are omitted. Both patience (time) and persistence (reminders) are required to achieve unbiased results. Quality improvement efforts depend on having accurate data and negative evaluations are particularly valuable. The relevance of these findings for recent developments in patient evaluation and quality improvement are drawn out, as well as the implications for practitioners, managers and policy makers.

(Limit of 350 words)

KEY WORDS: Patient satisfaction/statistics and numerical data; Hospitals/ standards; health care surveys/methods; Bias (epidemiology); Questionnaires

## Introduction

The national patient survey programme was first proposed in England in *The National Health Service: Modern, Dependable* (1997) [1] as a way of assessing patients' experiences of care and how they change over time. The survey was part of a more general commitment to make the NHS more responsive to patients. The reasoning was—and still is—that if hospital Trusts are given information about how patients evaluate the quality of the care they received as in-patients, managers and clinicians in the Trust will be able to respond to any identified shortcomings, leading to a general improvement in the quality of care.

The first Trust-based national survey of adult in-patients was conducted in 2001, when each hospital Trust in England was asked to conduct a postal survey of 850 patients using a questionnaire designed to elicit the patient's evaluations of, among other things, access to information, hospital cleanliness, communication by clinical staff, responsiveness of hospital staff, information on discharge and relationships among clinicians. Since the first survey was reported (Bullen and Reeves 2003) [2] inpatient surveys have been repeated almost every year. They are a potentially important resource for NHS Trusts as they provide information on experiences of care from probability samples of recent patients. However, their usefulness depends on the representativeness of those who respond. A number of studies have reported that response rates to postal surveys have declined in recent years (Hazel et al 2009; Anseel et al 2010) [3 and 4] and the NHS inpatient survey is no exception.

In England, Hospital Trusts are required to adopt a methodology that attempts to increase response rates by sending a reminder letter followed by a second reminder with a duplicate questionnaire to non-responders. Trusts usually send out a reminder around 21 days after the first questionnaire and a second questionnaire around 21 days after the reminder. Response to the first mailing without the need for a reminder in 2009 was 41 per cent, so the reminder and second questionnaire added eleven per cent to that initial response. This raises the important question of whether the eleven

per cent of patients who responded later or after reminders differed in some systematic way from those who responded at the first invitation. If there are systematic differences, this suggests that closing the survey after the first questionnaire and/or failing to send out reminders would have led to bias in the survey methodology leading to erroneous conclusions being drawn about patients' experiences in the NHS. The purpose of this paper then is to test whether there are significant differences between early and late responders and whether it is simply time or reminders that play the key role in achieving higher participation rates. This information could be used to determine whether the expense of carrying out repeat mailings and the use of survey methods which allow reminders to be sent is justified.

The research questions are:

1. Is there a significant difference in the way that early and late respondents to the survey evaluate the care they received in the NHS?
2. Is there a significant difference between patients who respond before and after the two reminders in how they evaluate their care?

The main purpose of the paper is to examine a potential source of bias that could be introduced if the scientific approach to gathering patient evaluations of care is not followed and to investigate whether it is time (patience) and/or reminders (persistence) that is most important in reducing the potential for bias.

The significance and timeliness of this study lies in the fact that since the publication of the NHS Operating Framework in 2008 there has been increasing emphasis in the English NHS on obtaining "real-time patient feedback". Although this concept is not defined, it seems to involve gathering data very quickly from patients, giving them little time to respond, and does not provide for sending reminders. We want to investigate whether there are potential sources of bias built into truncating the approach to gathering patient evaluations by examining one year of data from the National In-

Patient Survey before discussing the possible implications for current developments in this area and any messages for the wider audience of policy makers, managers and practitioners.

## Previous research

Concerns about quality of healthcare have led to the proliferation of patient satisfaction surveys. Researchers have been concerned about potential sources of bias and have examined whether those patient who do not respond differ in important ways from those who do respond. Many studies have shown that patient characteristics such as age, sex, ethnic background and own health status are related to the propensity to respond to healthcare surveys. The evidence that some patient groups are more likely to return the English national in-patient survey, for example, is summarised in a recent systematic review of research using data gathered from 2002 to 2009 (DeCourcy, West and Barron 2012) [5]. Patient surveys are designed to increase the “voice” that patients have about health care services, so when there are systematic differences in the rates at which groups respond, this inevitably raises concerns about equity.

In addition to studies focusing on patients socio-demographic characteristics, a few studies have investigated whether there is an association between the patients’ level of satisfaction with their care and whether or not they respond to surveys. Mazor (2002) [6] examined patient satisfaction survey data to assess the effect of response bias and found that there was a correlation between response rate and mean satisfaction score—that is, more satisfied patients were more likely to respond than those who were less satisfied. They then used this analysis as the basis of a simulation study which showed that non-response bias would most likely lead to patient satisfaction being overestimated. Further, as they were dealing with data about patient satisfaction with individual physicians, they were able to conclude that the scores for the physicians with whom the patients were least satisfied, would have the greatest magnitude of error.

In addition to non-response, one study did investigate whether there were systematic differences between early and late responders. Perenger, Chamot and Boyer (2005) [7] used items from the Picker Experience Questionnaire to show that early responders reported significantly fewer problems with the healthcare they received in one hospital in Switzerland than late responders or non-responders. While they concluded that there was a negative association between propensity to participate in the survey and problems during hospitalisation, they also found that increasing participation from 30 to 70 per cent had only a slight impact on the conclusions that could be drawn from the survey. More recently, Kwon et al (2010) [8] sent a postal questionnaire to 387 patients who had had knee replacement surgery 12 months previously and showed that the majority of the patient who responded to the questionnaire (91.4%) were satisfied with the results of the surgery. They showed that the non-responders were much more likely to have lower scores on scales to measure how their knee was functioning than those who were satisfied but their scores were better than those who responded but were dissatisfied with the surgery.

These studies suggest that if the response rates to a patient survey are low, we may over-estimate the extent to which patients are satisfied with the care they have received. However, the studies cited above are fairly small scale, involving a few hundred patients in a restricted number of locations (often just one) and simulated data in the first case. It is urgent therefore that we seek a greater understanding of how response bias might affect large scale patient surveys, such as the English national in-patient survey, as this is used to compare among hospitals and will have important consequences for patients and staff and for the organisations involved. Unlike the studies cited above, our focus is on the time taken to respond to the questionnaire and whether the reminders played a key role in eliciting data from patients who were less satisfied with their care. Although it would have been good to have included non-responders in this study, this was not possible using the large national dataset on which this study was based.

## Methods

This study uses the 2009 English national inpatient survey data. Questionnaires were sent to 137,360 patients, of whom 69,348 returned usable responses. Excluding undelivered questionnaires and deceased patients, this corresponds to a response rate of 52 per cent. We used two types of dependent variable. The first, Response Time, was the time in days from the dispatch of a mailing until the receipt of a response from the surveyed patient. It was measured either from the time of the first mailing or from the time of the most recent mailing. Where no questionnaire was returned we used the day after the date on which the latest questionnaire was received in a Trust as the censoring time. The second dependent variable was a binary indicator of whether or not a response was received before a reminder was sent out or not.

In one set of analyses our outcome variables were the times in days patients took to return their survey. The first variable of this type uses the total time since the initial questionnaire was mailed. The other three variables of this type use the time since the most recent mailing. By way of example, a hospital sent out reminders 21 and 42 days after the initial mailing. Three patients returned their questionnaires after 10, 25, and 45 days, respectively. This would give the values on the response time variables shown below.

Total response time	Time since 1 <sup>st</sup> mailing	Time since 2 <sup>nd</sup> mailing	Time since 3 <sup>rd</sup> mailing
10	10	n/a	n/a
25	n/a	4	n/a
45	n/a	n/a	3

The duration dependent variables were analysed using accelerated failure time regression assuming a hazard rate with a log-logistic distribution. This distribution was chosen as non-parametric estimates of the hazard rate showed a clearly non-monotonic form, as indeed would be expected. Robust standard errors that account for clustering within trusts were used. Binary logistic regression with random intercept was used for the analysis of the binary dependent variable.

The main explanatory variable was the patient's response to the question "Overall, how would you rate the care you received?" Responses to this variable are shown in table 1. We can see that a



large majority (79.3%) rated their care to be at least very good. We therefore divided responses to this question into three categories: Excellent, Very Good, and Other.

We also controlled for other factors that previous research has suggested may be associated with satisfaction with care. A systematic review (DeCourcy, West and Barron 2012) [5] of all the published research outputs produced using the patient survey data showed that several patient characteristics are associated with their evaluation of care. In this study therefore we control for these factors including age, sex, length of stay in hospital, and whether the person was admitted as an emergency or not.

## Results

Table 2 shows descriptive statistics for the key variables. We can see that there is a significant association between patients' evaluations of the quality of care with age, sex, whether a patient was an emergency admission, as well as a significant association between the length of stay and whether a patient replied after the first or subsequent mailing. Patients who rated their care as "Excellent" were more likely to respond to the first mailing than those who rated their care to be "Very Good" or worse. On the other hand, those patients who rated their care as worse than "Very Good" were more likely to respond to the second reminder. This suggests already that the repeated mailings are fulfilling a worthwhile function by eliciting responses from patients with a wider range of experiences. We can also see in the tables that those who responded to the first mailing were more likely to be older, male, have had an emergency admission and had a shorter stay in hospital than those who responded after the first or second reminder.

Figure 1 shows how the response rate changes over time, with the most common times of the first and second reminders shown as vertical lines. Of the questionnaires that were returned, 63.5 per cent were received before the first reminder was sent. A further 17.2 per cent were received before the second reminder was mailed, with the rest arriving after that.

Table 3 shows results of the accelerated failure time regressions. The first column shows results where the outcome variable is time since the first mailing was sent out. The other columns show the time that elapsed since the mailing that immediately preceded the return of the questionnaire. Predicted response times for a typical respondent are shown in Table 4. In the first column, the least satisfied respondents reply 3.1 days later on average than the most satisfied respondents. However, when we analyse response times from the most recent mailing these differences virtually disappear. This suggests that the key issue is not the length of time that patients take to respond, but whether or not they respond to the initial mailing or only reply after receiving one or two reminders. In part this may be an artefact of the data, as there is much less variation in these response times given the relatively short times between the mailings. Presumably some of the people who responded after the second or third mailings would have replied to the original questionnaire without having a reminder, but such long response times are lost in these analyses.

We next carried out two logistic regressions designed to estimate the probability that a survey respondent will reply before the first reminder is sent out. These results are shown in Table 5; they show a significantly higher probability of respondents who were less happy with their care replying after at least the first reminder was received. Predicted probabilities are shown in Table 6. It is clear that, even after controlling for other factors that influence the likelihood of an early response, a higher proportion of later respondents were less satisfied with their care than is true of people who responded before a reminder was sent out.

## **Conclusion**

In the past, some have questioned whether or not repeat mailings are important. There are of course costs associated with sending out tens of thousands of reminders every year, and there may even be ethical concerns. This paper shows that there is a relationship between a patient's overall evaluation of their care and the length of time it takes them to respond. Less satisfied patients are

more likely to fail to respond to the initial mailing, but significant numbers of them do respond to subsequent reminders. This demonstrates that repeat mailings reduce response bias in patient surveys. Without the repeat mailings, the proportion of people reporting their care was Excellent or Very Good would be significantly higher. This study suggests that both patience—giving patients time to respond, and persistence—sending reminders, is required to ensure that the survey data do not exclude patients who have had a more negative experience of care. The implications of this paper are that bias can be introduced through small changes to the survey protocol. As the NHS becomes more and more dependent on patient evaluations of their care it is essential that we work to produce data that gives a true picture of patients' experiences, rather than data that is misleading. In a paper titled "25 Years of Health Surveys: Does more data mean better data?", Berk, Schur and Feldman (2007) [9] reflected that, in the US "...survey designers are the victims of their own success; as policy makers understand the value of survey data in assessing policy changes, growing demands for data force agency budgets to emphasize short-term efforts while postponing longer term investments in data quality." One of their main recommendations is that more be invested in methods research. Perhaps if we were more confident about the quality of the data that we already collect, then we would have to collect less of it.

## References

- [1] Department of Health: **The new NHS Modern and Dependable: A National Framework for Assessing Performance**. 1998.
- [2] Bullen N, Reeves R: **Acute Inpatient Survey: National Overview 2001-02**. Department of Health. 2003.
- [3] Hazell M, Morris J, Linehan M, Frank P, Frank T: **Factors influencing the response to postal questionnaire surveys about respiratory symptoms**. *Primary Care Respiratory J*. **18**:3:165-170.
- [4] Anseel F, Lievens F, Schollaert E, Choragwicka B: **Response Rates in Organizational Science 1995-2008: A Meta-analytic Review and Guidelines for Survey Researchers**. *Journal of Business and Psychology* **25**, 335-349, 2010. doi:10.1007/s10869-010-9157-6
- [5] DeCoursey A, West E, Barron D: **National Adult Inpatient Survey conducted in the English National Health Service from 2002-2009: How have the data been used and what do we know as a result?** *BMC Health Services Research*. 12:71 (21<sup>st</sup> March 2012). DOI: 10.1186/1472-6963-12-71.
- [6] Mazor KM, Clauser BE, Field T, Yood RA, Gurwitz JH. **A demonstration of the impact of response bias on the results of patient satisfaction surveys**. *Health Services Research* 37:5 (October 2002).
- [7] Perneger T V, Chamot E, Bovier PA. **Nonresponse Bias in a Survey of Patient Perceptions of Hospital Care**. *Medical Care* **43**, 374-380 (2005).
- [8] Kwon SK, Kang YG, Chang CB, Sung SC, Kim TK: **Interpretations of the clinical outcomes of the nonresponders to mail surveys in patients after total knee arthroplasty**. *J Arthroplasty* 2010, jan;25(1):133-7.
- [9] Berk ML, Schur CL, Feldman J: **Twenty-five years of health surveys: Does more data mean better data?** *Health Affairs*, 26, no 6, October 2006. 1403-1417.

## CONTRIBUTION

EW: initiated the paper, devised the research questions, contributed to the interpretation of the findings, drafted, finalised and submitted the manuscript.

DNB: paid for the data, designed the study, conducted the data analysis, interpreted the findings and contributed to drafting the paper.

RR: contributed to the conception of the study, interpretation of findings and drafting of the manuscript.

DH: contributed to the research design, interpretation of findings and drafting the manuscript.

All 4 authors gave consent for the final document to be submitted.

## Acknowledgements

The authors would like to acknowledge the research assistance of Anna DeCourcy who assisted in the acquisition of the data and with the literature search.

## Sources of Funding

EW, and DB are full-time staff at University of Greenwich and RR is part-time (.2). DNB is a full-time member of staff of the University of Oxford.

## Tables

Table 1. Responses to the question: "Overall how would you rate the care you received?"

Excellent	Very good	Good	Fair	Poor
30038 (44.7%)	23228 (34.6%)	8821 (13.1%)	3532 (5.3%)	1527 (2.3%)

Table 2. Descriptive statistics

Quality of care	Replied to 1 <sup>st</sup> mailing	Replied after 2 <sup>nd</sup> or 3 <sup>rd</sup> mailing	Total
Excellent	20043 (67.2%)	9804 (32.8%)	29847
Very good	14449 (62.6%)	8622 (37.4%)	23071
Other	7950 (57.8%)	5803 (42.2%)	13753
Total	42442 (63.7%)	24229 (36.3%)	66671

Pearson chi-square: 371.7, df: 2, p < .05

Age group	Replied to 1 <sup>st</sup> mailing	Replied after 2 <sup>nd</sup> or 3 <sup>rd</sup> mailing	Total
16-35	3295 (48.8%)	3461 (51.2%)	6756
36-50	6279 (55.7%)	5001 (44.3%)	11280
51-65	12685 (62.9%)	7484 (37.1%)	20169
66+	27614 (63.7%)	15757 (36.3%)	43371
Total	49873 (61.1%)	31703 (38.9%)	81576

Pearson chi-square: 720.2, df: 3, p < .05

Sex	Replied to 1 <sup>st</sup> mailing	Replied after 2 <sup>nd</sup> or 3 <sup>rd</sup> mailing	Total
Female	23477 (63.2%)	13690 (36.8%)	37167
Male	20279 (64.0%)	11408 (36.0%)	31687
Total	43756 (63.5%)	25098 (36.5%)	68854

Pearson chi-square: 5.11, df: 1, p < .05

	Replied to 1 <sup>st</sup> mailing	Replied after 2 <sup>nd</sup> or 3 <sup>rd</sup> mailing	Total
Emergency	19705 (65.3%)	10465 (34.7%)	30170
Planned	22684 (62.5%)	13617 (37.5%)	36301
Total	42389 (63.8%)	24082 (36.2%)	66471

Pearson chi-square: 56.9, df: 1, p < .05

	Length of stay (days)
Replied to 1 <sup>st</sup> mailing	6.14
Replied after 2 <sup>nd</sup> or 3 <sup>rd</sup> mailing	6.65

t test: 6.20, df: 63117, p < .05

**Table 3. Log-logistic regression estimates. The first column is time from first mailing regardless of when received, others time from most recent mailing. (Robust standard errors in parentheses)**

	All replies	First Mailing	Second mailing	Third mailing
Intercept	2.90* (.024)	2.29* (.024)	2.09* (.033)	2.37* (.026)
Very Good	.088* (.009)	.026* (.007)	-.007 (.018)	.011 (.012)
Other	.169* (.011)	.040* (.011)	.006 (.020)	.039* (.014)
Emergency admission	.047* (.014)	.006 (.014)	-.004 (.017)	.002 (.015)
Age 36-50	-.132* (.016)	-.025* (.011)	.023 (.027)	-.047* (.020)
Age 51-65	-.265* (.017)	-.040* (.011)	-.015 (.032)	-.085* (.020)
Age 66+	-.333* (.018)	-.041* (.014)	-.036 (.031)	-.119* (.019)
Log length of stay (days)	.010* (.005)	.006 (.005)	.011 (.008)	.004 (.006)
Male	.009 (.008)	-.018* (.006)	.015 (.015)	-.021 (.011)
Scale	.477	.256	.381	.346
N	64847	41181	10334	12267
Log likelihood	-260414	-119231	-31732	-39410

Table 4. Predicted response times in days, based on the estimates shown in table 3.

	All replies	First Mailing	Second mailing	Third mailing
Excellent	17.1	9.62	8.51	10.0
Very Good	18.6	9.87	8.45	10.2
Other	20.2	10.0	8.56	10.4

Table 5. Odds ratios from random intercept logistic regression estimates of the probability of a respondent returning his or her questionnaire without the need for a reminder. (Standard errors in parentheses)

	I	II
Intercept	.500* (.024)	.745* (.112)
Very Good	1.21* (.118)	1.19* (.129)
Other	1.46* (.084)	1.37* (.099)
Emergency admission		1.08* (.239)
Age 36-50		.807* (.129)
Age 51-65		.626* (.043)
Age 66-80		.541* (.027)
Log length of stay (days)		1.00 (1.82)
Male		1.03 (.525)
SD of intercept	.389	.379
N	66671	64512
Log likelihood	-42872	-41125

Table 6 Using column I of Table 5, predicted probabilities of being a late responder for different levels of overall satisfaction with care.

Excellent	Very Good	Other
.333	.377	.422



**Figure 1**

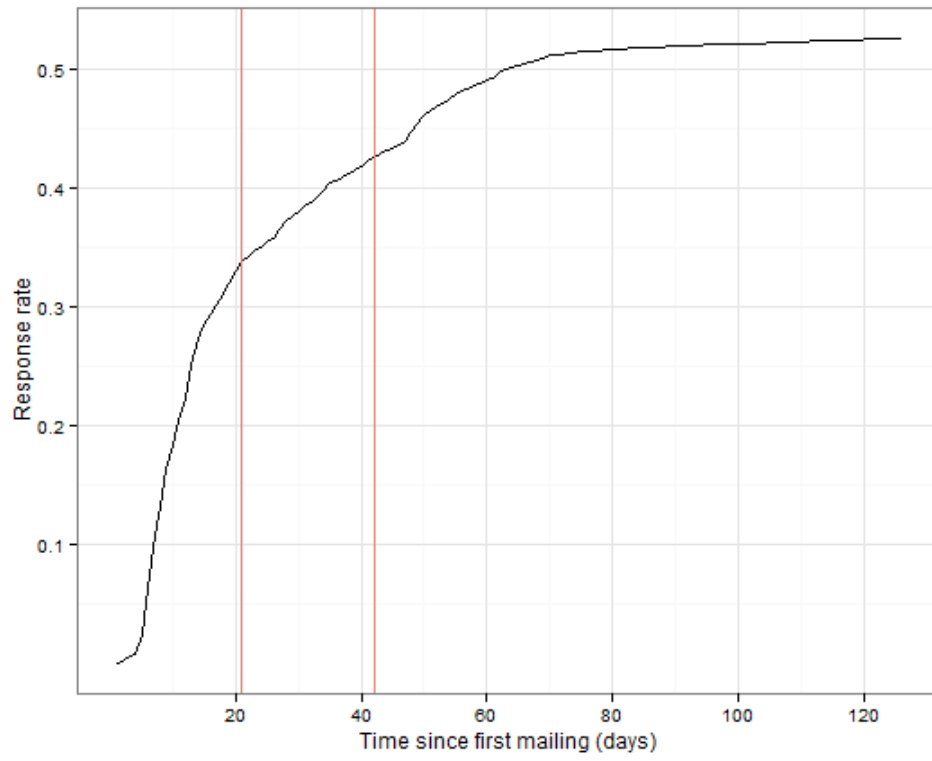


Figure 1. Response rate increases over time.