Selection bias resulting from the requirement for prior consent in observational research: a community cohort of people with ischaemic heart disease.

Brian Buckley
Andrew W Murphy
Molly Byrne
Liam Glynn

1 Department of General Practice, National University of Ireland, Galway, Ireland
2 Department of Psychology, National University of Ireland, Galway, Ireland.

Address for correspondence

Brian Buckley
Department of General Practice
National University of Ireland
Galway
Ireland

bsbuckley@iol.ie

Key words:

Selection bias
Consent
Observational research
Ischaemic heart disease
Abstract

Objective
To evaluate differences between adults who consent to participate in observational research and those who do not.

Design
Prospective, population based cohort study.

Setting
35 randomised Irish general practices.

Participants
1,609 adults with ischaemic heart disease identified in 2000/1

Intervention
Medical records search, postal questionnaire and consent form in 2005/6.

Main outcome measures
Differences in demographic and prognostic risk factors between consenters and non-consenters.

Results
At follow up charts were located for 1,592 patients (98.9%). Questionnaires were sent to 1,269 patients and 876 were returned (69%). Of these, 574 (65.5%) gave consent for participation in further research.

Logistic regression identified four characteristics as independently positively predictive of consent to participation in further research amongst questionnaire responders: having undergone percutaneous transluminal coronary angioplasty (PTCA) was associated with increased odds of consent, with an Odds Ratio (OR) of 1.77 (95%CI 1.09 to 2.86), as was a last recorded blood pressure < 140/90 mm/Hg (OR 1.45 [1.00 to 2.09]), a last recorded total cholesterol level < 5 mmol/l (OR 1.71 [1.16 to 2.54]) and being an ex-smoker rather than a current smoker or non-smoker (OR 1.73 [1.17 to 2.57]).

Conclusions
This research demonstrates the potential impact of consent bias in observational research on ischaemic heart disease (IHD), a disease area of everyday clinical importance in Europe. It demonstrates that clinically important prognostic variables may be associated with consent preferences. Future cohorts, dependent upon prior written consent, may contain disproportionate numbers of those who have made healthy lifestyle decisions, have previously benefited from treatment or whose clinical risk factors are already well managed. As a result, the generalisability of such research may be diminished and the effects of treatments over- or under-estimated.
Selection bias resulting from the requirement for prior consent in observational research: a community cohort of people with ischaemic heart disease.

**What is known on this subject**
There is worrying but limited evidence that requiring prior written consent may affect the validity of observational research. Those with “sensitive” conditions are less likely to consent to the use of their medical data. In the relatively rare context of intracranial vascular malformation, it has been demonstrated that the requirement for prior consent could affect a study’s ability to identify the prognostic importance of a factor which in clinical practice often influences the decision to treat.

**What this paper adds**
The potential impact of consent bias is demonstrated amongst a large European community-dwelling cohort of people with ischaemic heart disease, a disease of everyday clinical importance. In future, observational research consent bias may mean that determination of the effects of treatments becomes unreliable. The potential of consent bias to effect the validation of clinically important prognostic predictors is demonstrated in a sample large enough to allow other predictors to be controlled for.
Selection bias resulting from the requirement for prior consent in observational research: a community cohort of people with ischaemic heart disease.

Background

In recent years patients’ rights to privacy and confidentiality have been increasingly acknowledged both in law and in the considerations of research ethics committees. Patients’ prior written consent is required almost universally where identifiable data are sought for research purposes 1-6.

The implications for research of the requirement for prior written consent have been widely discussed, but the debate “has largely been confined to professional circles” 7-10. Research evidence considering public attitudes towards the use of personal medical information in research is limited 1,7,9. While the Medical Research Council (MRC) reported that the use of personal data in publicly funded health research is viewed positively by the public “if it will advance medical practice” and research in Leicestershire reported patients’ “altruistic views about participation in research” 1,11, a number of other studies have found that the public would prefer their consent to be sought before access to personal medical data is granted 9,12,13. Research also suggests that there may be only limited understanding amongst the public of what information is contained in medical records and of the value of such data to research 1,11,13. The promotion of greater understanding appears worthwhile: a study on attitudes to the retention and research use of tissue samples has demonstrated a striking increase in the proportion of patients willing to consent to the use of their personal data in research when the potential value to of their doing so was explained to them 14.

The threat posed to observational research by current requirements for prior written consent is grave. Concerns have been expressed amongst researchers that there is a danger that constraints may have become too strict: that in order to prevent limited and largely theoretical harm to individuals, work may be prevented which offers very large benefits to society 7,13,15-20.

Amongst the most serious consequences of the requirement for prior written consent is the threat to the validity of observational research posed by “consent bias”, a term coined to describe the selection bias resulting from the loss of non-consenters to any cohort. It has been suggested that “patients, the public and professional organisations must consider the implications … before epidemiology and health services research are regarded as too biased to rely on.” 15

However, demonstrating the potential seriousness of this phenomenon has only been possible in a small number of studies with access to data for non-consenters as well as consenters. In Rochester, Minnesota, Jacobsen et al found in their medical records research – to which 79% consented - that those refusing consent were more likely to be female and younger than 60 years. Those with “sensitive” diagnoses such as reproductive disorders, mental disorders or infectious diseases were also less likely to consent to participation 18. A study based on the Registry of the Canadian Stroke Network supported the American finding that women were more likely to refuse consent than men but found that those who did not consent to interview and medical
records review (49%) were more likely to be older. In the UK, Angus et al found that different proportions of men and women consented to research participation in different age-groups and that consenters were less likely to live in deprived areas\textsuperscript{21}. Al Shahi et al demonstrated that clinically important prognostic variables can be vulnerable to consent bias: amongst a cohort of people with intracranial vascular malformation, the positive association between one prognostic variable – which often influences the decision to treat in clinical practice - and an important outcome was confirmed when data relating to their whole cohort were analysed, but not so when those who did not give consent (41%) were excluded\textsuperscript{15}.

**Methods**

**The CoHeart study**
The CoHeart study is a five-year follow up of a representative cohort of 1,609 people with established ischaemic heart disease (IHD) in 35 randomised and stratified general practices in the west and north-west of Ireland\textsuperscript{22}. The cohort was established in a cross-sectional study in 2000-2001 and presented a valuable opportunity to conduct a follow up study of the Secondary Prevention of IHD amongst a representative community cohort at a time when the future of such observational work is at risk. IHD was defined as a history of previous acute myocardial infarction (AMI), angina pectoris, cardiac artery bypass grafting (CABG) or percutaneous transluminal coronary angioplasty (PTCA). With original data available for cohort members irrespective of their consent preferences regarding research participation, the study also presented an opportunity to test hypotheses relating to consent bias amongst a large community based cohort of people with established IHD.

**Ethical approval for the study**
Ethical approval for both baseline and follow up studies was granted by the Irish College of General Practitioners (ICGP). Ethical approval at baseline allowed researchers to identify the practice populations of people with IHD through review of general practice records, to establish practice IHD registers, to collect medical and demographic data and to send each member of the cohort a patient questionnaire. The baseline database was subsequently anonymised. Follow up data collection was dependent upon the consent and participation of the individual general practitioners (GPs), who are the legal data controllers of patient records in Ireland. Ethical approval allowed for collection of anonymous data from patient records, linked in the practices to baseline data by unique patient identification codes. A protocol was established which enabled the sending of patient questionnaires by agents of the practices without the necessity for disclosure of patients’ identities to follow up researchers. In order to facilitate future follow-up, a consent form was enclosed with the patient questionnaire which requested patients’ written consent and identification for involvement in further research.

**Analysis of consent bias**
Patient questionnaires could be returned with or without the consent form – or not returned at all. Thus the study identified three subgroups of patients with different consent preferences: “consenting responders” (who completed the questionnaire and gave their consent and contact details for participation in further research), “non-consenting responders” (who completed the questionnaire but did not consent) and “non-responders”.
These subgroups represented different preferences relating to participation in research and consent for access to personal medical data and thus afforded an opportunity to consider associations between these preferences and demographic, medical and health care data.

**Statistical methods**

The selection of variables considered in this study was informed by sources of consent bias identified in the literature: age (by ten-year age bands), sex, socioeconomic status (measured by eligibility or ineligibility for free General Medical Services within the Irish health system – at the time of the follow up study just under 40% of the population in the area in which the study was based were GMS eligible, representing the least affluent members of society), and prognostic risk factors. Prognostic risk factors for subsequent IHD events which were considered in analysis were previous AMI, previous CABG or previous PTCA (whether or not patients had ever experienced these events), blood pressure and cholesterol management (last reading at follow up ≤ or > level recommended by the Second Joint Task Force of European and other Societies on Coronary Prevention 23, 24), Body Mass Index (≤ or > 25kg/ m²), self-reported smoking status (current, ex or non-smoker at follow up) and exercise behaviour (≤ or > level recommended by the Second Joint Task Force, measured by the Godin Leisure Time Exercise Questionnaire at follow up 25).

Univariate analysis used to consider associations between demographic and clinical risk factor variables were $\chi^2$ (for sex, socioeconomic status, IHD status, blood pressure and cholesterol management, body mass index, smoking status and exercise behaviour) and $\chi^2$ for trend (for age ). Multiple logistic regression was subsequently used to determine which variables independently predicted questionnaire response and consent preference when the effect of other variables was taken into account. Those who died since baseline and those who were excluded from receipt of the patient questionnaire and consent form were excluded from analyses.

**Results**

Baseline chart data were available for 1,609 patients: 65.4% (n=1053) male and 34.6% female (n=556); mean age was 66 (standard deviation 9.1); 79% were GMS eligible. Patients could be excluded from receipt of the baseline patient questionnaire by their GPs for reasons such as very poor health status or illiteracy. Baseline patient questionnaires were sent to 1,577 patients and returned by 1,084 (response rate 68.7%).

At follow up records were located for 1,592 patients (98.9%). The mean age of the surviving cohort was 69.5 (SD 9.2). Patients could be excluded from receipt of the patient questionnaire at follow up for reasons such as death since baseline, serious illness, having moved away from the practice, illiteracy: 340 patients (including all of those for whom no records could be located) were excluded by their general practitioners for these reasons as outlined in Figure 1. Questionnaires were sent to the remaining 1,269 patients and were returned by 876 (response rate 69%). Of responders, 574 (65.5%) completed and signed the form which gave consent for participation in further research.
Responders and non-responders – demographics and risk factors

The levels of non-response were almost identical between sexes, with 30.9% of women and 31.0% of men not responding at all. Univariate analysis identified a significant association between response and cholesterol management: of patients whose last serum cholesterol reading was under the recommended 5mmol/l, 71.8% (n=522) returned the completed questionnaire compared with 66% (n=285) of those whose readings were above this level ($\chi^2$ 4.4; df 1; p<0.05). However, logistic multiple regression identified no variable as independently predictive of questionnaire response.

Consenter and non-consenter – demographics and risk factors

Amongst responders univariate analysis detected no significant association between consent preferences and either age or socioeconomic status. Neither was any significant association between consent preference and BP management, previous AMI, BMI or exercise behaviour identified. Gender was found to be significantly associated with consent preference, with 68.1% of men (n=388) compared with 60.8% of women (n=186) consented to participation in further research ($\chi^2$ 4.7; df 1; p<0.05). Of those whose last serum cholesterol reading was below the recommended 5 mmol/l, 69.2% (n=361) consented to further participation compared with 61.4% (n=175) of those whose reading was above this level ($\chi^2$ 5; df 1; p<0.05). Having had surgical cardiac interventions was found to be significantly associated with consent, with 72.5% (n=150) of those who had had a previous CABG compared with 64% (n=419) of those who had not consented ($\chi^2$ 5.1; df 1; p<0.05) and 74.2% (n=138) of those who had had
PTCA consenting compared with 63.5% (n=427) of those who had not ($\chi^2$ 7.5; df 1; p<0.01). Smoking status was associated with consent, with smoking cessation in particular influencing consent preferences: 73.4% (n=312) of ex-smokers consented to participation compared with 61.4% (n=62) of smokers and 61.6% (n=162) of non-smokers ($\chi^2$ 13.2; df 3; p<0.01).

**Multiple regression**

In order to evaluate how IHD risk factor variables predicted consent preferences when other variables were controlled for, a logistic multiple regression model was fitted to the 540 cases for whom complete data were available and is summarized in Table 1.

<table>
<thead>
<tr>
<th>Table 1: Risk factors as predictors of consent to further participation in research - summary of logistic regression (N=540)</th>
</tr>
</thead>
<tbody>
<tr>
<td>n (% of N)</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td><strong>Consenters</strong></td>
</tr>
<tr>
<td><strong>Gender: male</strong></td>
</tr>
<tr>
<td><strong>Age: increasing</strong></td>
</tr>
<tr>
<td><strong>GMS eligible</strong></td>
</tr>
<tr>
<td><strong>Previous AMI</strong></td>
</tr>
<tr>
<td><strong>Previous CABG</strong></td>
</tr>
<tr>
<td><strong>Previous PTCA</strong></td>
</tr>
<tr>
<td><strong>Last BP &lt; 140/90 mmHg</strong></td>
</tr>
<tr>
<td><strong>Last total cholesterol &lt; 5 mmol/l</strong></td>
</tr>
<tr>
<td><strong>Ex-smoker</strong></td>
</tr>
<tr>
<td><strong>BMI &gt; 25kg/m^2</strong></td>
</tr>
<tr>
<td><strong>Adequate exercise</strong></td>
</tr>
</tbody>
</table>

*Cox & Snell $R^2$ .056 Nagelkerke $R^2$ .078

Although the explained variation was relatively small, analysis identified four characteristics as independently and significantly positively predictive of consent to further research participation when the effect of other predictors was taken into account: having undergone PTCA, with an Odds Ratio (OR) of 1.77 (1.09 to 2.86); previous smoking cessation (that is to say being an ex-smoker rather than a current smoker or non-smoker), with an OR of 1.73 (1.17 to 2.57); a last recorded BP within recommended levels (< 140/90 mm/Hg), OR 1.45 (1.00 to 2.09); and a last recorded total cholesterol level within recommended levels (< 5 mmol/l), OR 1.71 (1.16 to 2.54). Associations were considered amongst all responders between the variables identified as predictors of consent preference: smoking cessation was significantly associated with PTCA ($\chi^2$ 19.8; df 1; p<0.01).

The cases included in the analysis represented 61.6% of responders. The loss of 38.4% resulted for the most part from incomplete patient questionnaire data, most notably failure to supply either height or weight so that BMI could not be calculated. No significant difference in demographic or clinical characteristics was identified between those included and those not included.


Discussion

This study adds to the small but important body of evidence which can demonstrate the potential of increasing consent requirements to create selection biases which could undermine observational research. The study has benefited from having access to a large community based cohort for whom a rich database based on medical records review and patient questionnaire was available at both baseline and five year follow up. The request for patient’s contact details and consent for participation in research was included with patient questionnaires to facilitate subsequent follow up, but also afforded the opportunity to consider consent and research participation preferences amongst a large representative cohort of people with IHD.

This study’s main finding is that consent bias threatens the validity of future observational research amongst people with IHD. In this respect, the research supports the findings in the existing literature on consent bias. Like Al Shahi et al, this research demonstrates that clinically important prognostic variables can be associated with consent preferences and this has serious implications for future research. However, this study extends these implications from the disease area in the previous research - which is potentially catastrophic but relatively rare - to one which is of everyday clinical importance in Europe.

The size of the cohort in this study allowed sufficient power to conduct multivariate analysis. Unlike the existing literature, this research found that although gender was identified as significantly associated with consent preferences by univariate analysis, this was not the case when other predictors were controlled for.

However the association between clinically important prognostic variables and consent preferences was shown to be significant: patients who had ceased smoking, whose cholesterol and blood pressure were well managed or who had previously benefited from PTCA were more likely to consent for research participation. Previous PTCA was shown to be associated with smoking cessation. While it seems likely that the former drives the latter, this cannot be determined from the data; if so, then it might be concluded that it is the healthy lifestyle decision (smoking cessation) rather than the medical intervention which is the true predictor of consent to research participation.

The implication is that if cohorts in the future are dependent upon prior written consent they are likely to contain disproportionate numbers of those who have made healthy lifestyle decisions, who have previously benefited from health care or those whose clinical risk factors are already well managed. This may have two serious consequences: first, the generalisability of observational research will be reduced; second, the effects of treatments may be variously overestimated or underestimated if those who are most unwell or are not making healthy lifestyle decisions are under-represented in study populations.

Policy implications

All the stakeholders in health care research, including the public, need to understand and address the very serious implications for research - and ultimately for medicine - of over-zealous interpretation and implementation of confidentiality laws and
guidance. The directive approved by the European Parliament which prompted the introduction of new national data protection legislation permitted member states to make exceptions in the case of health related research where the benefits to society outweigh any harm attributable to the invasion of privacy. In some countries these provisions were not incorporated in legislation so that some forms of observational research have become nearly impossible. In others – such as the UK - the use of identifiable data without prior consent in medical research is permitted in law under certain circumstances. But the associated laws and bureaucracy are complex and public attitudes towards the use of personal medical information in research are unclear. As a result, the requirement for prior consent has emerged as a default position of safety amongst regulatory bodies and research ethics committees.

If the effectiveness of observational research, of epidemiology and ultimately of some elements of medicine are not to be diminished, every effort must be made to ensure that the provisions made available to health related research by the European Parliament are adopted and implemented where this is not already the case.

Acknowledgements
We wish to particularly acknowledge the help of the Health Research Board, Noel Scott and Diarmuid O’Donovan of the HSE (Western Area) and the practice staff and participating general practitioners: Drs. Marcus Allen; Desmond Bluett; Sean Bourke; Martin Brennan; Vivian Brennan; James Brogan; Marian Brogan; Charles Bourke; Sean Bourke; Declan Clinton; Seamus Cryan; Martin Daly; Anthony Delap; John- Mark Dick; Ken Egan; Noel Farrell; Mary Feerick; Brendan Forkan; Margaret Gilligan; Enda Harhen; Edward Harty; Ciara Kelly; Richard Joyce; Bernard McGuire; Pauric Mitchell; Paul Money; Kay Moran; Daniel Murphy; Kieran O’Reilly; John O’ Sullivan; Roddy Quinn; John Regan; Michael Regan; Eamonn Shea; John Sheerin; Richard Tobin; David Townley; Kieran Whyte. One GP requested to remain anonymous.

Conflicts of interest
None. All authors declare that the answer to all the questions on your competing interest form is No and therefore have nothing to declare.

Sources of funding
Both baseline and follow up studies have been funded by the Health Research Board, Dublin and the Health Services Executive (Western Area).

Copyright
The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd and its Licensees to permit this article (if accepted) to be published in HEART editions and any other BMJPLG products to exploit all subsidiary rights.
References

2. An information guide to the data protection acts for general practitioners: Irish College of General Practitioners / National General Practice Information Technology Group; 2003.