Original article

doi:10.1111/j.1463-1318.2011.02724.x

The effect of seeking consent on the representativeness of patient cohorts: iron-deficiency anaemia and colorectal cancer

S. Damery, R. Ryan, R. J. McManus, S. Warmington, H. Draper, S. Wilson and on behalf of the Improving Colorectal Outcomes (ICOS) Group

Primary Care Clinical Sciences, School of Health and Population Sciences, University of Birmingham, Edgbaston, West Midlands B15 2TT, UK

Received 24 January 2011; accepted 7 June 2011; Accepted Article online 11 August 2011

Abstract

Aim The study aimed to establish the level of selection bias that may occur should individual patient consent be sought, by comparing characteristics of consenters and nonconsenters to a request for access to medical records within a cohort of patients diagnosed with iron-deficiency anaemia (IDA).

Method A cohort study and cross-sectional survey was carried out of consent preferences that compared the sociodemographic characteristics of patients providing or not providing consent for access to their records, the consent rates by participant subgroup and the predictors of consent/nonconsent.

Results Of 599 patients mailed requesting consent for access to their medical records, 425 (71.0%) responses were received. Of the valid responses, explicit consent was granted by 371 (62.7%) respondents, with 47 (7.9%) refusals. The characteristics of consenters and nonconsenters differed with regard to age, gender and deprivation quartile. Nonconsent was associated with younger age (40–60 years *vs* 60+ years; bivariate OR = 2.84; 95% CI = 2.01–4.02), female gender (OR = 1.62; 95% CI = 1.13–2.34) and being socioeconomically deprived (OR = 1.61; 95% CI = 1.15–2.26).

Conclusion The current research governance framework demonstrates a conflict between protecting the rights of the individual and the development of a sound research base to improve the delivery of healthcare services for society as a whole. If epidemiological research includes data only from individuals who have given consent for access to their records, the resulting selection bias may have consequences for the scientific validity and generalizability of research findings, and ultimately the quality of patient care.

Keywords Consent, bias, colorectal cancer

What is new in this paper

In showing statistically and clinically significant differences between consenters and nonconsenters for access to medical records, this study suggests that if epidemiological research studies include data only from individuals who have given consent for their records to be accessed, the resulting selection bias may reduce the scientific validity of the research.

Introduction

For many years, there have been calls to improve the evidence base supporting the delivery of healthcare in the UK [1]. In particular, there has been considerable investment in cancer research, driven in part by a need to address the relatively poor cancer survival rates in the UK in comparison with other European countries [2,3]. There has been a growing investment in the evaluation of methods for improving the early diagnosis of cancer, much of it being focused on the role of primary care in facilitating prompt and appropriate patient referral where cancer is suspected [4]. Such research depends on the collection and use of patient data to assess cancer incidence, prevalence and survival, and to allow evaluation of the extent to which recommended care pathways are followed [5].

Epidemiological research which focuses on care pathways, where data relating to presentation or referral are typically obtained from different sources to those for outcomes data [6], can only succeed if the same individuals can be reliably identified in different data sets [7–9]. However, regulatory requirements governing

Correspondence to: Sue Wilson, Professor of Clinical Epidemiology, Primary Care Clinical Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.
E-mail: s.wilson@bham.ac.uk

access to, and use of, patient data have become increasingly stringent, particularly with regard to the need for informed consent to be obtained prior to patients becoming research subjects. [10,11]. The existing governance framework aims to support research in the public interest, to protect patients and research participants from risk, to promote good clinical and research practice and to ensure that individual rights to privacy and confidentiality are upheld [6].

Patients' written consent is necessary in almost all instances where identifiable data are sought for research purposes [12,13]. If it is not feasible to secure informed consent, or if the seeking of informed consent introduces significant bias, this requirement may prevent research that may be in the public interest, in order to avoid potential harm to the individual that may be felt to occur as a result of their loss of privacy [14,15]. However, a growing body of evidence demonstrates that studies which rely solely on data obtained from individuals who have provided consent may produce findings that are unrepresentative [16,17] or misleading [14,18]. This makes it unlikely that scientifically valid and generalizable conclusions can be drawn from such studies, thereby reducing the potential value of the research findings in informing clinical practice and improving healthcare delivery and patient outcomes [19].

This paper outlines the findings of a study designed to establish the level of selection bias that may occur should individual patient consent for access to medical records be sought. The study aimed to inform the methodology for subsequent research to describe care pathways for patients with iron-deficiency anaemia (IDA) and to determine the predictors of delayed diagnosis of colorectal cancer (CRC), which was developed because of concerns that IDA, a presenting symptom in 50% of cases of CRC [20,21], was being poorly managed by general practitioners (GPs) despite National Institute for Health and Clinical Excellence (NICE) referral guidelines [4]. An analysis of care pathways requires identification of a representative cohort of patients with IDA, yet no single existing data set includes complete information on presentation and investigations in primary care, secondary care referrals and data on outcomes (cancer). Large, anonymized primary care databases have incomplete referral information [22], and neither hospital episode statistics nor cancer registry data sets detail the investigation of IDA prior to CRC diagnosis.

In the absence of data sets that share common anonymized or pseudo-anonymized identifiers, record linkage is needed, which requires either individual informed consent or approval from the National Information Governance Board Ethics and Confidentiality Committee (NIGB ECC) to set aside the common-law duty of confidentiality under Section 251 of the NHS Act (2006). This study aimed to determine whether there was a need for Section 251 approval by comparing the characteristics of consenters and nonconsenters to a request for access to medical records within a cohort of patients with a primary care diagnosis of IDA, and investigated the extent, direction and wider implications of any bias that would occur should individual informed consent be required.

Method

Study design and setting

This cohort study was carried out using patient records from one large urban general practice in Birmingham, with a list size of 6750 patients.

Participants

All currently registered adult patients, ≥ 40 years of age, who received a haematological or clinical diagnosis of IDA between 2001 and 2006, were identified via a computerized search of practice records undertaken in November 2008. IDA was defined as a haemoglobin (Hb) measurement below 10 g/dl in women and below 11 g/dl in men. The record level data held by the GP were aggregated to produce a list of patients with the lowest Hb measurement and the date of that estimation was recorded for each patient. Practice partners checked the list of patients in order to identify and exclude any who should not be contacted owing to a lack of capacity or other reasons, and any patients who had died were removed from the list. The list was ranked according to each individual's Hb measurement (lowest to highest), and a sample of patients with the lowest Hb measurements was selected to receive an invitation to participate in the study.

A sample size of 600 patients was chosen based on an assumed 60% response rate, in order to yield 360 responses. Assuming that half of these respondents would provide consent for access to their records, this would generate a sample of 180 patients, enabling a 30% effect size to be detected in analytical variables in a comparison of consenters and nonconsenters with 80% power and 5% significance.

Selected patients were sent a patient information sheet and a letter describing the study. These outlined the purpose of the research (to use information contained within the medical records of patients who had a low blood count to help improve the understanding of the reasons for anaemia). A consent form and a reply slip were also enclosed, to be completed and

returned in a FREEPOST envelope. One reminder was sent to those who had not responded to the initial mailing after 2 weeks. Permission was sought for the University research team to examine medical records. The patient information leaflet stated that participation in the research would incur no personal benefits although the research may help other patients in the future.

Analysis

The characteristics of patients providing consent for access to their records were compared with those of nonconsenters with respect to age, gender and deprivation quartile (according to the 2004 Index of Multiple Deprivation) [23], and whether or not an individual was recorded as having a diagnosis of CRC. Categorical variables were compared using chi-square tests. An anonymized data set, which included the characteristics of all eligible patients at the same general practice (excluding those who had been mailed asking for consent), was also analysed to allow a comparison between the characteristics of consenters and the rest of the eligible patient population. The characteristics of those refusing consent were determined by univariate and multivariate logistic regression. All data were analysed using the Statistical Package for the Social Sciences (SPSS) version 15.0 (SPSS Inc., Chicago IL, USA).

Results

A total of 2895 patients, ≥ 40 years of age, were registered at a general practice, had a valid postal address and a recorded Hb measurement indicative of IDA between 2001 and 2006. A sample of patients with the lowest recorded Hb values were contacted in writing requesting consent for access to their medical records. An administrative error resulted in 599 rather than the proposed 600 patients being mailed. In total, 425 (71.0%) responses were received (Fig. 1). This included responses on behalf of 2 (0.3%) patients who had died in the period between obtaining the practice data and conducting the survey, and 5 (0.8%) patients whose consent forms were classed as spoiled as a result of both the 'consent' and 'no consent' fields being completed. These seven patients were removed from subsequent analyses.

Consent rates

Patient preferences fell into three categories: those who provided explicit consent (n = 371; 62.7%); those who explicitly refused consent (n = 47; 7.9%); and nonresponders (n = 174; 29.4%). The simplest derivation of a consent rate is the proportion of responders who provide explicit consent: in this study 88% (371 out of 418). If an 'opt-in' consent model is adopted, whereby nonresponders are assumed to implicitly refuse consent, the consent rate in this study can be estimated at 62.7% (371 out of 592). If an 'opt-out' model is followed, in which nonresponders are assumed to be consenters, the 'consent' rate in this study is 92.1% (545 out of 592). Analyses investigating the characteristics of those providing and refusing consent will focus on the most conservative estimate of the consent rate and adopt an opt-in model. Hereafter, any reference to 'consenters' includes individuals who provided explicit consent, and any reference to 'nonconsenters' includes all individuals who either refused consent or did not respond.

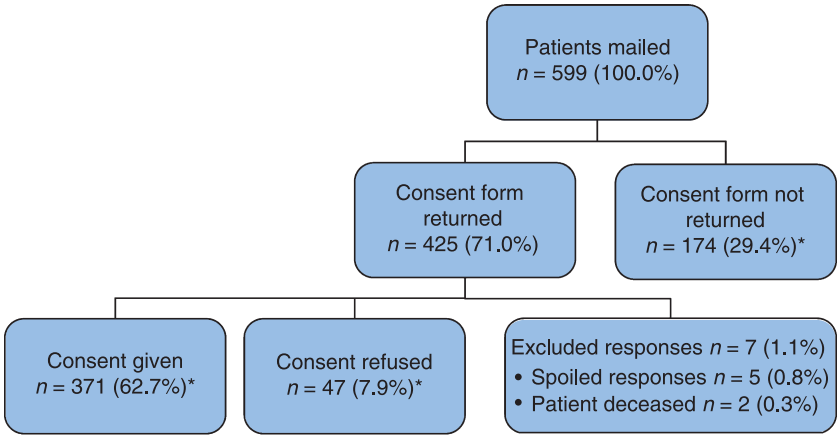

Figure 1 Schematic of the response to consent request mailing. *Corrected proportions after removal of the seven excluded responses (two patients had died and five patients returned spoiled responses) from the denominator.

Characteristics of survey recipients

Table 1 shows the characteristics of the practice population (all practice patients ≥ 40 years of age who received a Hb measurement indicative of IDA between 2001 and 2006; n = 2895), with the characteristics of those mailed (n = 599). The mean age of practice patients with IDA was 60.6 years, whilst the mean age of those mailed was slightly higher, at 65.2 years. More women received the survey than men (n = 391; 65.2% *vs* n = 208; 34.7%, respectively). The majority of practice patients and mailing recipients were from more deprived areas. In both populations, fewer than 10% lived in either of the two most affluent deprivation quartiles.

Comparison between consenters and nonconsenters

The characteristics of consenters and nonconsenters were compared (Table 2). In total, 72.1% of those ≥ 60 years of age gave consent (n = 263), compared with 47.6% of those < 60 years of age (n = 108). More men than women consented (n = 142; 69.9% *vs* n = 229; 58.9%). Individuals in affluent quartiles were more likely to consent than those in deprived quartiles (quartiles 1 and 2: 68.5% consent, n = 191 *vs* quartiles 3 and 4: 57.4% consent, n = 179). Univariate logistic regression confirmed the predictors of consent refusal to be younger age, female gender and socioeconomic deprivation. Those between 40 and 60 years of age were significantly more likely than those > 60 years of age to refuse consent for access to their medical records (bivariate OR = 2.84; 95% CI: 2.01–4.02). Women were significantly more likely to refuse consent than men (bivariate OR = 1.62; 95% CI: 1.13–2.34). Patients living in more deprived areas were also significantly more likely to refuse consent for access to medical records than patients living in more affluent areas (bivariate OR = 1.61; 95% CI: 1.15–2.26). Only CRC status did not show a statistically significant difference between consenting and nonconsenting groups.

All factors that had been significant in the univariate model remained significant in a multivariate regression analysis, with the exception of gender (Table 2). The inclusion of interaction terms to test for confounding effects did not reveal any statistically significant interactions between predictor variables and did not improve the predictive ability of the regression model.

Consent rates by sociodemographic characteristics

The distribution of consenters and nonconsenters varied across sociodemographic variables. Figure 2 shows consent rates stratified by age, gender and deprivation category (quartiles 1 and 2 were grouped as 'Affluent' and quartiles 3 and 4 were grouped as 'Deprived'). The subgroup with the highest consent rate contained older (60+ years of age) deprived men (78.1% provided consent; n = 57). Least likely to respond positively were younger (40–60 years of age), deprived women (38.1% consent rate; n = 37). Within the older age group, both affluent men and women had high consent rates (74.0%; n = 54 and 75.2%; n = 82 respectively).

Table 1 Demographic characteristics of the eligible practice population and of the sample.

Characteristic	Practice population*	Sample
Age		
40–60 years	1509 (52.1)	227 (37.9)
60+ years	1386 (47.9)	372 (62.1)
Total	2895 (100.0)	599 (100.0)
Mean age (years)	60.6	65.2
Median age (years)	59.0	65.0
Gender		
Male	1404 (48.5)	208 (34.7)
Female	1491 (51.5)	391 (65.2)
Total	2895 (100.0)	599 (100.0)
Index of Multiple Deprivation†		
Quartile 1	43 (1.5)	9 (1.5)
Quartile 2	238 (8.2)	49 (8.3)
Quartile 3	680 (23.6)	136 (23.0)
Quartile 4	1925 (66.7)	404 (68.2)
Total	2886 (100.0)	598 (100.0)
Index of Multiple Deprivation within practice‡		
Quartile 1 + Quartile 2	1434 (49.5)	281 (46.9)
Quartile 3 + Quartile 4	1452 (50.2)	317 (52.9)
Not recorded	9 (0.3)	1 (0.1)
Total	2895 (100.0)	599 (100.0)
Colorectal cancer status		
Yes	34 (1.2)	13 (2.2)
No	2861 (98.8)	586 (97.8)
Total	2895 (100.0)	599 (100.0)

Data are given as n (%).

*The practice population consisted of all patients at the practice, ≥ 40 years of age, who had a recorded haemoglobin measurement indicative of iron-deficiency anaemia (IDA) (defined as < 11 g/dl for men and < 10 g/dl for women) between 2001 and 2006.

†Index of Multiple Deprivation (2004); quartile 1 = least deprived, quartile 4 = most deprived.

‡Index of Multiple Deprivation within practice: uses the Index of Multiple Deprivation score of each of the 2895 potential study patients to construct a measure of deprivation for the patients of the practice.

Table 2 Comparison between consenters and nonconsenters.

Characteristic	Consent preference (percentage of individuals mailed in each group)*			Bivariate OR, likelihood of refusing consent (95% CI)†	Multivariate OR, likelihood of refusing consent (95% CI)†
	Yes	No opinion expressed‡	No		
Age group					
40–60 years	108 (47.6)	106 (46.7)	13 (5.7)	Younger *vs* older	Younger *vs* older
60+ years	263 (72.1)	68 (18.6)	34 (9.3)	2.84 (2.01–4.02);	2.77 (1.94–3.96);
Total	371 (62.7)	174 (29.4)	47 (7.9)	$P < 0.0001$	$P < 0.0001$
Gender					
Male	142 (69.9)	47 (23.2)	14 (6.9)	Females *vs* males	Females *vs* males
Female	229 (58.9)	127 (32.6)	33 (8.5)	1.62 (1.13–2.34);	1.46 (0.99–2.12);
Total	371 (62.7)	174 (29.4)	47 (7.9)	$P = 0.008$	$P = 0.051$
Deprivation					
Affluent (Q1 + Q2)	191 (68.5)	66 (23.7)	22 (7.8)	Deprived *vs* affluent	Deprived *vs* affluent
Deprived (Q3 + Q4)	179 (57.4)	108 (34.6)	25 (8.0)	1.61 (1.15–2.26);	1.57 (1.11–2.23);
Not recorded	1 (100.0)	0 (0.0)	0 (0.0)	$P = 0.005$	$P = 0.011$
Total	371 (62.7)	174 (29.4)	47 (7.9)		
CRC status					
Ever had CRC	8 (61.5)	5 (38.5)	0 (0.0)	CRC *vs* no CRC	CRC *vs* no CRC
Never had CRC	363 (62.7)	169 (29.2)	47 (8.1)	0.99 (0.65–1.50);	1.64 (0.52–5.20);
Total	371 (62.7)	174 (29.4)	47 (7.9)	$P = 0.945$	$P = 0.402$

CRC, colorectal cancer; Q, quartile.
Data are given as *n* (%) or as OR (95% CI).
*Seven patients in the cohort of 599 who were mailed were removed from these analyses. Two recipients had died and five returned spoiled responses.
†Nonconsenters are defined as those who either explicitly refused consent or did not return a reply slip.
‡Questionnaire not returned.

Discussion

This study has shown significant sociodemographic differences between the characteristics of patients who gave explicit consent for access to their medical records and those who did not. Consent was found to be associated with gender, age and deprivation in univariate analyses, and with age and deprivation in multivariate analyses. When sociodemographic variables were considered independently, the patients refusing consent tended to be younger, female and less affluent than those providing consent. When taken together, older, male, deprived individuals were most likely to respond positively, whereas younger, female, deprived patients were the group least likely to give consent. In showing statistically (and clinically) significant sociodemographic differences between consenters and nonconsenters, this study adds to a growing body of evidence which suggests that if epidemiological research studies include data only from individuals who have given consent for their records to be accessed, the resulting selection bias may reduce the scientific validity of the research [12,18].

If the direction and magnitude of consent bias is predictable, it would be possible to perform some statistical adjustment to compensate for anticipated bias, weighting responses from groups that were under-represented in any analysis. However, whilst this approach may be effective in population surveys, where the primary issue is lack of representativeness, the biases introduced by requiring consent in the present study (and in all similar studies) may be caused by unknown factors that cannot be adjusted for. Other studies have found inconsistency in the nature and extent of differences between consenters and nonconsenters for all outcomes [9,18,24]. A systematic review of prospective observational studies requiring informed consent for access to medical records identified seven studies with no age-related differences, one in which consenters were more likely to be younger than nonconsenters and seven with significant differences across age strata. Similarly, six

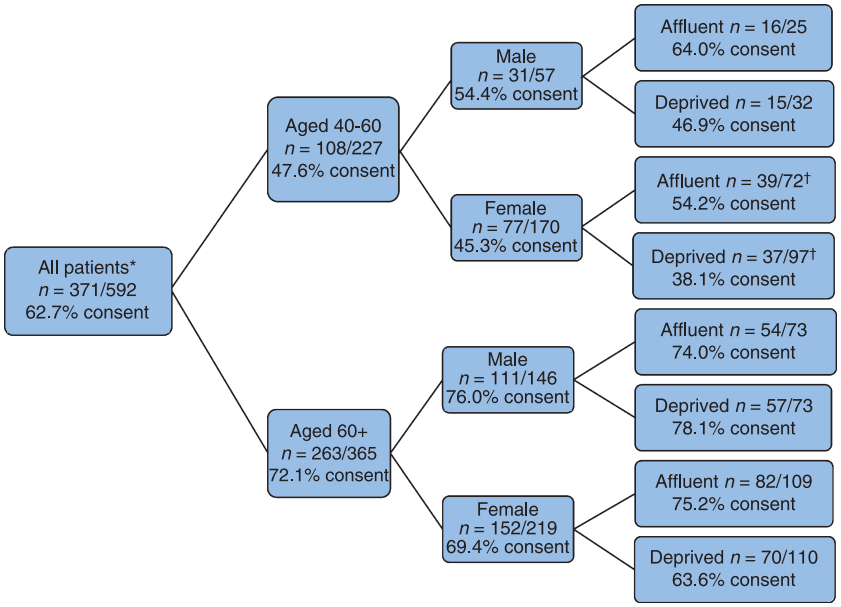

Figure 2 Consent rates stratified by sociodemographic variables. *Seven patients in the cohort of 599 mailed were removed from these analyses: two recipients had died, and a further five returned spoiled responses. †The deprivation quartile could not be derived for one patient.

studies found significant differences in consent preferences according to gender: two where women were more likely to consent; and four where men were more likely to consent [25].

The factors associated with consent preferences may also be significantly related to health outcomes, and clinically important prognostic variables may be associated with consent preferences [12]. Analyses in our study adopted an 'opt-in' consent model, combining nonresponders with those who refused consent, rather than an 'opt-out' model combining nonresponders with explicit consenters. An opt-in approach is typically favoured by research ethics committees, who may consider an opt-out model to undermine the principle that consent should be freely given. Although the number of explicit nonconsenters in this study was too small to allow a robust statistical analysis of the implications of adopting an opt-out model, it is likely that adherence to an opt-in consent model introduces bias. Opt-in models lead to lower response rates and may skew samples towards those with less acute health problems or lower levels of material disadvantage, thereby excluding individuals whose data could most usefully contribute to the analysis [26]. In the proposed study of IDA, it is possible that selection bias would lead to an over-representation of healthier patients and an underestimation of the number of cancers diagnosed following IDA. If those who had experienced delays in diagnosis refused consent, analyses would fail to determine the factors predictive of delayed diagnosis and may inappropriately conclude that delays in referral do not exist, and that any refinement of care pathways is unnecessary.

Whilst few would argue against consent for medical research being sought from participants wherever possible, there are instances where the seeking of consent will result in a sample that is so biased (selection bias) that the analyses will produce unreliable and misleading results. There are also instances where the seeking of consent may cause distress. In this example, seeking informed consent would require us to approach patients with IDA and inform them that we had reason to believe that the care and investigations undertaken by their GP may have been suboptimal. This would potentially cause significant anxiety for the majority of patients who had received appropriate care for their IDA; may unnecessarily lead some patients to believe that they have CRC; may make some patients who had a diagnosis of CRC believe that their GP had delayed the diagnosis; and may damage doctor–patient relationships.

The findings of this and other studies demonstrate a conflict between the rights of the individual and the development of a sound research base to establish appropriate patient care for society as a whole. Patient autonomy and the right to privacy need to be balanced against the selection bias and potential distress that may be the consequences of seeking consent. In the absence of comprehensive, reliable, linked and anonymized data sets, access to patient identifiable information is essential if

accurate record linkage and valid determination of outcomes are to be achieved. Assuming that research staff are properly trained, data are handled confidentially and appropriate data-security policies are in place, the potential harm resulting from the analysis of routinely collected medical data without seeking consent is likely to be less than the harm that could result from seeking consent.

The current research governance framework with regard to patient consent is based on the assumption that access to medical records for research purposes without consent is unacceptable to the majority of the public [6]. This is largely because of contradictory findings from research into public preferences. The Medical Research Council (MRC) has reported that the use of personal data in publicly funded health research is seen positively by the public [27], whereas others have found that individuals would prefer consent to be sought before medical records are used [28]. However, when asked, few patients explicitly object to their records being used or deny consent [29]. For most research studies, the risk of disclosure of sensitive information is extremely low, particularly where patient identifiers are required only as a temporary step to ensure accurate record linkage, with identifiable information being deleted after linkages have been made and anonymized data sets created prior to analysis. The recently published Academy of Medical Sciences report [30] makes recommendations that may improve the situation with regard to research governance, data linkage and patient consent; however, far-reaching reforms are required.

Limitations

The general practice chosen for the study may have been atypical because of its location in a relatively economically deprived area. This may have resulted in a different consent profile than would have been obtained from a practice in a more affluent location. The sample was unrepresentative of the general population, with approximately twice as many women as men, and around twice as many in the older age group *vs* the younger age group. However, the sample population is broadly representative of the subpopulation that suffers from IDA [31]. Other disease groups and patient cohorts may have a different age, gender and deprivation mix and therefore a different consent profile.

Conclusions

The requirement for explicit patient consent for the use of identifiable data may introduce biases that reduce the validity of research. Consent bias has potentially serious consequences for the quality and wider generalizability of epidemiological research findings, the assessment of the delivery of healthcare services and the quality of patient care. Data from all patients with IDA require analysis if we are to better understand the incidence of cancer subsequent to IDA, describe care pathways, establish the time to cancer diagnosis and ultimately identify the reasons for delays in diagnosis and develop interventions to reduce these delays. Despite a growing body of evidence which highlights the effects and implications of selection bias in the identification of patient cohorts and assessment of care pathways, requirements governing access to, and use of, identifiable patient information are becoming more stringent owing to the need to conform with the Data Protection Act (1998). No resolution has yet been made to ensure a balance between the rights of the individual and the rights of the population, in spite of an increasing recognition that these may often be in conflict with respect to the conduct of high-quality healthcare research.

Acknowledgements

The Improving Colorectal Outcomes (ICOS) Group: the writing committee and Professor William Hamilton, Dr Jim Murray, Professor Greg Rubin, Mr Tariq Ismail, Dr Tariq Iqbal, Professor Richard Hobbs, Mrs Christine Gratus.

Conflict of interests

No author has any competing interests.

Funding

This study was funded by the Cancer Research UK Population and Behavioural Sciences funding programme (Ref C9783/A8).

Ethical approval

The study was approved by the South Birmingham Research Ethics Committee in August 2008 (Ref: 08/7H1207/166), and Research Governance approval was given by Birmingham and the Black Country Primary Care Trust RM&G Consortium (Ref 1080). Approval was also obtained from the Patient Information and Advisory Group (PIAG, now National Information Governance Board), Ref: PIAG 4-05(k)/2007.

Author contributions

S.D. analysed the data and wrote the first draft of the manuscript. Sa.W. and R.R. were responsible for data

collection and manipulation. R.M. was responsible for identifying eligible participants. H.D. provided input into the ethical aspects of study design and evaluation. All authors were involved in the design of the study, critically commented on early drafts of the manuscript and approved the final version.

References

1 Department of Health (2000). *The NHS Cancer Plan*. London: NHS Executive.

2 Sant M, Capocaccia R, Coleman MP, Berrino F, Gatta G, Micheli A. Cancer survival increases in Europe, but international differences remain wide. *Eur J Cancer* 2001; **37:** 1659–67.

3 Berrino F, Verdecchia A, Lutz JM, Lombardo C, Michell A, Capocaccia R; EUROCARE Working Group. Comparative cancer survival information in Europe. *Eur J Cancer* 2009; **45:** 901–8.

4 National Institute for Health and Clinical Excellence (2005). *Referral Guidelines for Suspected Cancer*. London: NICE.

5 Department of Health (2007) *The Cancer Reform Strategy*. London: Department of Health.

6 Academy of Medical Sciences (2006) *Personal Data for Public Good: Using Health Information in Medical Research*. London: The Academy of Medical Sciences.

7 Iversen A, Liddell K, Fear N, Hotopf M, Wessely S. Consent, confidentiality and the Data Protection Act. *BMJ* 2006; **332:** 165–9.

8 Ward HJ, Cousens SN, Smith-Bathgate B *et al*. Obstacles to conducting epidemiological research in the UK general population. *BMJ* 2004; **329:** 277–9.

9 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:** 21.

10 Dixon-Woods M, Yeung K (2010) *Governing Health Services Research: Is it Working?* London: Health Services Research Network, NHS Confederation.

11 Cassell J, Young A. Why we should not seek individual informed consent for participation in health services research. *J Med Ethics* 2002; **28:** 313–7.

12 Buckley B, Murphy AW, Byrne M, Glynn L. Selection bias resulting from the requirement for prior consent in observational research: a community cohort of people with ischaemic heart disease. *Heart* 2007; **93:** 1116–20.

13 Al-Shahi R, Warlow C. Using patient-identifiable data for observational research and audit. *BMJ* 2000; **321:** 1031–2.

14 Al-Shahi R, Vousden C, Warlow C; for the Scottish Intracranial Vascular Malformation Study (SIVMS) Steering Committee. Bias from requiring explicit consent from all participants in observational research: prospective, population based study. *BMJ* 2005; **331:** 942.

15 Regidor E. The use of personal data from medical records and biological materials: ethical perspectives and the basis for legal restrictions in health research. *Soc Sci Med* 2004; **59:** 1975–84.

16 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:** 336–8.

17 Dunn KM, Jordan K, Lacey RJ, Shapley M, Jinks C. Patterns of consent in epidemiologic research: evidence for over 25,000 responders. *Am J Epidemiol* 2004; **159:** 1087–94.

18 Macleod U, Watt GCM. The impact of consent on observational research: a comparison of outcomes from consenters and non consenters to an observational study. *BMC Med Res Methodol* 2008; **8:** 15.

19 Hansson MG. Need for a wider view of autonomy in epidemiological research. *BMJ* 2010; **340:** c2335.

20 Jiwa M, Hamilton W. Referral of suspected colorectal cancer: have guidelines made a difference? *Br J Gen Pract* 2004; **54**(505)**:** 608–10.

21 Hamilton W, Lancashire R, Sharp D, Peters TJ, Cheng KK, Marshall T. The importance of anaemia in diagnosing colorectal cancer: a case-control study using electronic primary care records. *Br J Cancer* 2008; **98:** 323–7.

22 Damery S, Ryan R, Wilson S, Ismail T, Hobbs FDR; on behalf of the Improving Colorectal Outcomes Group (ICOS). Iron deficiency anaemia and delayed diagnosis of colorectal cancer: a retrospective cohort study. *Colorectal Dis* 2011; **13**(4)**:** e53–60.

23 Office of the Deputy Prime Minister (2004) *The English Indices of Deprivation 2004 – Summary (revised)*. London: Office of the Deputy Prime Minister.

24 Huang N, Shih S-F, Chang H-Y, Chou Y-J. Record linkage research and informed consent: who consents? *BMC Health Serv Res* 2007; **7:** 18.

25 Kho ME, Duffet M, Willison DJ, Cook DJ, Brouwers MC. Written informed consent and selection bias in observational studies using medical records: systematic review. *BMJ* 2009; **338:** b866.

26 Junghans C, Jones M. Consent bias in research: how to avoid it. *Heart* 2007; **93:** 1024–5.

27 Medical Research Council (2000) *Personal Information in Medical Research*. London: Medical Research Council.

28 Willison DJ, Keshavjee K, Nair K *et al*. Patients' consent preferences for research uses of information in electronic medical records: interview and survey data. *BMJ* 2003; **326:** 373–7.

29 Barrett G, Cassell JA, Peacock JL *et al*. National survey of British public's views on use of identifiable medical data by the National Cancer Registry. *BMJ* 2006; **332:** 1068–72.

30 Academy of Medical Sciences (2011) *A New Pathway for the Regulation and Governance of Health Research. Academy of Medical Sciences*. London: Academy of Medical sciences.

31 Goddard AF, James MW, McIntyre AS, Scott BB; on behalf of the British Society of Gastroenterology (2005). *Guidelines for the Management of Iron Deficiency Anaemia*. London: British Society of Gastroenterology.