

• RAPID COMMUNICATION •

Internet-based data inclusion in a population-based European collaborative follow-up study of inflammatory bowel disease patients: Description of methods used and analysis of factors influencing response rates

Frank L Wolters, Gilbert van Zeijl, Jildou Sijbrandij, Frederik Wessels, Colm O'Morain, Charles Limonard, Maurice G Russel, Reinhold W Stockbrügger

Frank L Wolters, Reinhold W Stockbrügger, Department of Gastroenterology and Hepatology, University Hospital Maastricht, P. Debyeplein 25, 6202 AZ Maastricht, The Netherlands
Gilbert van Zeijl, Jildou Sijbrandij, Charles Limonard, MEMIC, Centre for Data and Information Management, P. Debyeplein 1, 6229 HA Maastricht, The Netherlands
Frederik Wessels, Global Vitis, Wattstraat 52, 2171 TR Sassenheim, The Netherlands

Colm O'Morain, Adelaide and Meath Hospital, Department of Gastroenterology, Trinity College, Tallaght, SE-41345 Dublin 24, Ireland

Maurice G Russel, Department of Gastroenterology and Hepatology, Medisch Spectrum Twente, Haaksbergerstraat 55, 7513 ER Enschede, The Netherlands

Supported by the European Commission as a fifth framework shared cost action (QLG4-CT-2000-01414)

Correspondence to: Dr FL Wolters, Department of Gastroenterology and Hepatology, PO box 5800, 6202 AZ Maastricht, The Netherlands. frankwolters@ace-on-air.nl

Telephone: +31-43-3875021

Received: 2005-01-13 Accepted: 2005-02-18

Abstract

AIM: To describe an Internet-based data acquisition facility for a European 10-year clinical follow-up study project of a population-based cohort of inflammatory bowel disease (IBD) patients and to investigate the influence of demographic and disease related patient characteristics on response rates.

METHODS: Thirteen years ago, the European Collaborative study group of IBD (EC-IBD) initiated a population-based prospective inception cohort of 2 201 uniformly diagnosed IBD patients within 20 well-described geographical areas in 11 European countries and Israel. For the 10-year follow-up of this cohort, an electronic patient questionnaire (ePQ) and electronic physician per patient follow-up form (ePpPFU) were designed as two separate data collecting instruments and made available through an Internet-based website. Independent demographic and clinical determinants of ePQ participation were analyzed using multivariate logistic regression.

RESULTS: In 958 (316 CD and 642 UC) out of a

total number of 1 505 (64%) available IBD patients, originating from 13 participating centers from nine different countries, both ePQ and ePpPFU were completed. Patients older than 40 years at ePQ completion (OR: 1.53 (95%CI: 1.14-2.05)) and those with active disease during the 3 mo previous to ePQ completion (OR: 3.32 (95%CI: 1.57-7.03)) were significantly more likely to respond.

CONCLUSION: An Internet-based data acquisition tool appeared successful in sustaining a unique Western-European and Israeli multi-center 10-year clinical follow-up study project in patients afflicted with IBD.

© 2005 The WJG Press and Elsevier Inc. All rights reserved.

Key words: Internet; Questionnaire; IBD; Cohort study; Population-based

Wolters FL, van Zeijl G, Sijbrandij J, Wessels F, O'Morain C, Limonard C, Russel MG, Stockbrügger RW. Internet-based data inclusion in a population-based European collaborative follow-up study of inflammatory bowel disease patients: Description of methods used and analysis of factors influencing response rates. *World J Gastroenterol* 2005; 11(45): 7152-7158
<http://www.wjgnet.com/1007-9327/11/7152.asp>

INTRODUCTION

Data acquisition in multi-center epidemiological follow-up studies traditionally has been dependent on postal paper form questionnaires filled out by study subjects and on paper form extraction files for relevant medical information completed by physicians and/or research assistants. Great advantages of the use of the Internet as a platform for health status measurements have been reported since the mid-1990s^[1]. Internet-based facilities supporting data acquisition in studies concerning patients with inflammatory bowel disease (IBD) have been performed from then onwards^[2-4]. Traditionally, in multi-center studies, questionnaires in paper format are forwarded from patients and centers to central database-hosting facilities, where data are manually transformed

into electronic databases. This multi-step process is time- and labor-consuming and prone to errors. Paper form data acquisition utilities normally contain question flow instructions for study subjects and for physicians that can easily be misinterpreted resulting in missing or wrong answers. Furthermore, the physical transformation of large numbers of paper format data into an electronic database adds to the risk of human errors.

For the purpose of a large multi-center follow-up study of a European population-based cohort of patients with IBD^[5], we created an Internet-based data acquisition facility, in order to avoid the above-mentioned methodological imperfections and to assure data quality. The design of the mentioned study as well as practical and technical details and implications of the new electronic instrument are described and discussed. The influence of demographic and disease-related patient characteristics on the patient questionnaire response rates is also analyzed.

History of the cohort

After 3 years of preparation, between October 1991 and September 1993, the European Collaborative study group of IBD (EC-IBD) gathered a population-based prospective inception cohort of 2 201 uniformly diagnosed IBD patients within 20 well-described geographical areas in 12 European countries^[6]. Initially it was hypothesized that a North-South incidence gradient in IBD existed, which could be proved^[5]. Later, in the context of this registry, a 1- and 4-year clinical follow-up^[7,8] were studied, as well as the occurrence of rheumatologic manifestations^[9] and the relationship between quality of care and quality of life^[10].

The currently executed 10-year clinical follow-up study project was planned since 1998, granted by the European Commission, and was started in 2001.

MATERIALS AND METHODS

Materials

In order to optimize data acquisition in this European multi-center study project, electronic data acquisition tools were designed for Internet-based access. The development of the tools and the subsequent data management leading to an analyzable dataset involved a five-step process; (1) a "patient questionnaire" (PQ) and "physician per patient follow-up form" (PpPFU) were prepared as two separate data collecting instruments and originally constructed for the purpose of this study. The information obtained by the PQ was derived from direct questioning of the patients, whereas physicians and research assistants extracted data from patient files into the PpPFU. PQ items were demographics, life style factors, disease activity and use of medication, family history, data on pregnancy and fertility, health care consumption, disability, and quality of life (SF-36 was used). PpPFU items, as observed during the 10-year follow-up period, were vital status, cause of death, disease activity (change of immunosuppressive medication and surgical events were used as indicators due to the retrospective nature of data acquisition), disease location

and behavior, use of medication, surgery, health care consumption, colorectal cancer, and dysplasia. Members of the original study group and additional investigators formed "working groups" to deal with various research topics of the follow-up, including genotype-phenotype, course of disease, pregnancy and fertility, cancer, health care consumption and costs, and dissemination of messages to the medical and lay public. A central body of senior representatives from the EC-IBD group constructed final comprehensive versions of both instruments based on the input of all the members of the working groups. The electronic patient questionnaire (ePQ) questions were translated into the nine languages of the participating countries by a professional translation agency and subsequently checked for correctness by a representative from every participating country from within the EC-IBD study group. For the PpPFU, the language used was English; (2) a tailor-made innovative high performance data entry application was designed, assuring maximum flexibility for programmers and users. The ePQ was designed by the construction of a questionnaire design module, where questions and answers could be directly introduced in different languages with needed routings. The final ePQ contained 747 questions and could be used in nine different languages including Hebrew. The electronic physician per patient follow-up form (ePpPFU) was designed to receive an unlimited number of entries on multiple topics on a 10-year time scale. Graphic design using simple colors in order to visually discriminate users' choices improved user friendliness; (3) the PQ questions and the PpPFU items, originally constructed on paper, were introduced into the electronic facilities. A pre-designed compulsive question flow was implemented in the ePQ. This implicated that patients filling in the ePQ were automatically guided towards a subsequent question, based on a given previous answer thus skipping irrelevant questions. Electronic guidance facilities concerning mandatory and optional fields were embedded in the ePpPFU. In this way, physicians and/or research assistants completing the ePpPFU of a particular patient were not allowed to close and finalize the document concerning an individual patient before having collected all the minimally required information; (4) the electronic data acquisition facilities were made available to all data collecting centers through a username-password secured and firewall protected Internet-based website. Links to ePQ statistics and ePpPFU statistic pages displayed real time counts of both instruments for all centers together showing instantaneous, comparative and cumulative progress of data inclusion, accessible to every member involved in the data inclusion of this project; (5) after completion of the data acquisition phase, export files from the Internet questionnaire program were imported into the database environment. The database was situated in Maastricht. Database validation took place at this center before the distribution of the final database to all working groups. The validation was done by logical comparison of the crude database with existing databases of previous EC-

Table 1 Cumulative number of IBD (CD+UC) patients per center followed-up or lost to follow-up for ePQ and ePpPFU. Centers indicated in gray had not reached the 60% response rate threshold

Center	Total number of CD+ UC patients	EPQ response (%)	EPQ response without deaths and No IBD %	Reason for non response EPQ					PpPFU response (%)	PpPFU response without No IBD %	Reason for non response PpPFU				EPQ and PpPFU (%)
				Not willing	Untrace-able	No IBD	Death	Not started			Not willing	Untrace-able	No IBD	Not started	
Beer Sheeva	60	41 (68)	76	5	8	3	3	0	46 (77)	77	3	11	0	0	42 (70)
Copenhagen	147	103 (70)	84	12	7	0	25	0	132 (90)	90	6	9	0	0	120 (82)
Cremona	51	35 (68)	76	4	7	0	5	0	43 (84)	84	1	7	0	0	40 (78)
Heraklion	62	38 (61)	68	2	13	0	5	4	49 (79)	79	3	17	0	0	40 (65)
Ioannina	43	36 (84)	88	2	3	0	2	0	39 (91)	91	0	4	0	0	38 (88)
Oslo	378	226 (60)	69	65	34	6	45	2	332 (88)	89	21	25	1	0	271 (72)
Reggio Emilia	84	61 (73)	79	4	12	0	7	0	69 (82)	82	5	11	0	0	68 (81)
S-Limburg	216	156 (72)	76	28	20	0	11	1	194 (90)	90	0	23	0	0	159 (74)
Vigo	100	73 (73)	77	4	18	0	5	0	81 (81)	81	3	16	0	0	77 (77)
SUBTOTAL	1 141	769 (67)	75	126	122	9	108	7	985 (86)	86	42	123	1	0	855 (75)
Almada	21	11 (52)	58	1	7	0	2	0	12 (57)	57	0	9	0	0	11 (52)
Dublin	186	57 (30)	34	10	58	3	17	41	80 (43)	43	1	26	2	78	56 (30)
Firenze	115	48 (42)	42	0	0	0	0	67	39 (34)	34	0	4	0	72	36 (31)
Milano	42	19 (45)	58	1	6	0	9	7	0 (0)	0	0	0	0	42	0 (0)
SUBTOTAL	364	135 (37)	41	12	71	3	28	115	131 (36)	36	1	39	2	192	103 (28)
Total	1 505	904 (60)	67	138	193	12	136	122	1 116 (74)	74	43	162	3	192	958 (64)

IBD projects in terms of numbers of patients per center with known diagnosis, sex, age, and disease behavioral characteristics. The results of this data validation effort were summarized in a logbook that was distributed to all the project members for possible revisions that subsequently were implemented as changes in the database. Final data sets were available for analysis promptly after closing the data acquisition phase.

Methods

Before the start of the data collection, a minimally required response rate per center was set at 60% identification of the patients to have both ePQ and ePpPFU completed. Centers that did not comply with this minimum response rate were analyzed separately from those that did.

UC patients were grouped for disease location at diagnosis as: (1) rectum only; (2) rectum and sigmoid and/or descending colon; and (3) disease location beyond the splenic flexure. The CD patients were classified for disease location at diagnosis, according to the Vienna classification^[11], into: (1) upper gastrointestinal (esophagus and/or stomach and/or duodenum and/or jejunum); (2) (terminal) ileum; (3) ileocolonic; and (4) colon only.

Independently, both demographic and clinical determinants of ePQ participation (diagnosis, gender, center, disease location at diagnosis, age at ePQ completion, and disease activity during 3 mo period previous of ePQ completion) were analyzed using univariate and multivariate logistic regression models. All potential predictors were included in both models. A $P \leq 0.05$ was considered to be statistically significant in the multivariate analysis. The statistical package for the social sciences (SPSS 11.5.1 for Windows; SPSS Inc., Chicago, IL, USA) was used for the analysis.

The Ethics Committees of all the participating centers approved the study protocol, and all subjects gave their informed consent before the start of the study.

RESULTS

During the data collection period from August 1st 2002 to January 1st 2004, the electronic data entry facilities appeared to be well-equipped and flexible tools for data entry and storage. However, the data export facility used in order to enable fast provision of stored data into formats prepared for analysis showed some initial deficiencies. Questions, answer categories and formal control rules, used for feedback during insertion, were not always automatically incorporated in the export process. This could, however, be repaired by the information management team.

All centers that originally participated in the EC-IBD cohort had been approached to take part in the follow-up study. Thirteen out of the original twenty centers distributed over nine countries participated. In 958 (316 CD and 642 UC) out of a total number of 1 505 IBD patients eligible for follow-up (64%), both ePQ and ePpPFU were completed (Table 1). When considering the entire cohort of both CD and UC patients, nine centers complied with the minimal 60% response threshold rendering 855 out of 1 141 patients (75%). Table 1 summarizes the details of the total number of IBD (both CD and UC) patients in the cohort. In the CD category, 10 of the participating 13 centers complied with the minimal 60% response threshold leaving 378 patients of whom 288 (76%) had completed both ePQ and ePpPFU (Table 2). In the UC category, nine centers complied leaving 776 patients of whom 575 (74%) had completed both instruments (Table 3).

Table 2 shows two Beer Sheeva patients indicated as "no IBD" in the ePQ, who were not classified as such in the ePpPFU. Review of available clinical data revealed that both patients have CD. An identical discrepancy concerning one Dublin patient in the "no IBD" category could not be clarified. This patient was considered to have CD. In the UC category (Table 3) divergences between ePQ and ePpPFU for patients indicated as "no IBD"

Table 2 Number of CD patients per center followed-up or lost to follow-up for ePQ and ePpPFU. Centers indicated in gray had not reached the 60% response rate threshold

Center	Total number of CD patients	EPQ response (%)	EPQ response without deaths and No IBD %	Reason for non response EPQ					PpPFU response (%)	PpPFU response without No IBD %	Reason for non response PpPFU				EPQ and PpPFU (%)
				Not willing	Untrace-able	No IBD	Death	Not started			Not willing	Untrace-able	No IBD	Not started	
Almada	13	8 (62)	67	0	4	0	1	0	8 (62)	62	0	5	0	0	8 (62)
Beer Sheeva	21	13 (62)	72	2	3	2	1	0	15 (71)	71	1	5	0	0	14 (67)
Copenhagen	58	41 (71)	85	5	2	0	10	0	52 (90)	90	6	0	0	0	47 (81)
Cremona	10	7 (70)	78	2	0	0	1	0	9 (90)	90	1	0	0	0	8 (80)
Heraklion	14	7 (50)	70	0	3	0	3	1	11 (79)	79	0	5	0	0	9 (64)
Ioannina	6	4 (67)	80	0	1	0	1	0	5 (83)	83	0	1	0	0	5 (83)
Oslo	110	69 (63)	70	15	13	1	11	1	97 (88)	89	4	8	1	0	79 (73)
Reggio Emilia	34	24 (71)	77	0	7	0	3	0	27 (79)	79	1	6	0	0	27 (79)
S-Limburg	77	60 (78)	82	8	5	0	4	0	69 (90)	90	0	8	0	0	61 (79)
Vigo	35	27 (77)	84	1	4	0	3	0	30 (86)	86	1	4	0	0	30 (86)
SUBTOTAL	378	260 (69)	77	33	42	3	38	2	323 (85)	85	14	42	1	0	288 (76)
Dublin	63	17 (27)	30	2	25	2	5	12	25 (40)	40	0	6	1	31	17 (27)
Firenze	29	13 (45)	45	0	0	0	0	16	13 (45)	45	0	1	0	15	11 (38)
Milano	13	4 (31)	50	0	2	0	5	2	0 (0)	0	0	0	0	13	0 (0)
SUBTOTAL	105	34 (32)	36	2	27	2	10	30	38 (36)	37	0	7	1	59	28 (27)
TOTAL	483	294 (61)	69	35	69	5	48	32	361 (75)	75	14	49	2	59	316 (65)

Table 3 Number of UC patients per center followed-up or lost to follow-up for ePQ and ePpPFU. Centers indicated in gray had not reached the 60% response rate threshold

Center	Total number of CU patients	EPQ Response (%)	EPQ response without deaths and No IBD %	Reason for non response EPQ					PpPFU response (%)	PpPFU response without No IBD %	Reason for non response PpPFU				EPQ and PpPFU (%)
				Not willing	Untrace-able	No IBD	Death	Not started			Not willing	Untrace-able	No IBD	Not started	
Beer Sheeva	39	28 (72)	78	3	5	1	2	0	31 (79)	79	2	6	0	0	28 (72)
Copenhagen	89	62 (70)	84	7	5	0	15	0	80 (90)	90	0	9	0	0	73 (82)
Cremona	41	28 (68)	76	2	7	0	4	0	34 (83)	83	0	7	0	0	32 (78)
Heraklion	48	31 (65)	67	2	10	0	2	3	38 (79)	79	3	12	0	0	31 (65)
Ioannina	37	32 (86)	89	2	2	0	1	0	34 (92)	92	0	3	0	0	33 (89)
Oslo	268	157 (59)	69	50	21	5	34	1	235 (88)	88	17	17	0	0	192 (72)
Reggio Emilia	50	37 (74)	80	4	5	0	4	0	42 (84)	84	4	5	0	0	41 (82)
S-Limburg	139	96 (69)	72	20	15	0	7	1	125 (90)	90	0	15	0	0	98 (71)
Vigo	65	46 (70)	73	3	14	0	2	0	51 (78)	78	2	12	0	0	47 (72)
SUBTOTAL	776	517 (67)	74	93	84	6	71	5	670 (86)	86	28	86	0	0	575 (74)
Almada	8	3 (38)	50	1	3	0	1	0	4 (50)	50	0	4	0	0	3 (38)
Dublin	123	40 (33)	36	8	33	1	12	29	55 (45)	45	1	20	1	47	39 (32)
Firenze	86	35 (41)	41	0	0	0	0	51	26 (30)	30	0	3	0	57	25 (29)
Milano	29	15 (51)	60	1	4	0	4	5	0 (0)	0	0	0	0	29	0 (0)
SUBTOTAL	246	93 (38)	41	10	40	1	17	85	85 (35)	35	1	27	1	133	67 (27)
TOTAL	1 022	610 (60)	67	103	124	7	88	90	755 (74)	74	29	113	1	133	642 (63)

regarding one patient from Beer Sheeva and five patients from Oslo could not be clarified. These six patients were considered to have UC. Discrepancies between total numbers of patients for ePQ and ePpPFU also occurred because of Internet disruptions during the insertion process in the CD category regarding two patients from Heraklion (Table 2), and in the UC category regarding five patients from Heraklion, one from Oslo, one from Reggio Emilia, one from South-Limburg and one from Dublin (Table 4). The answers given by these patients were lost and these patients were excluded from the database. Disease location at diagnosis was known for 739 of 766 UC patients and 363 of 378 CD patients.

In the 60% of complying centers, 126 of 1 141 patients [11%] refused to participate without a large difference

between the age groups (79/707 [11.2%] in patients older than 40 years at ePQ completion and 47/434 [10.8%] in patients equal to or younger than 40 years at ePQ completion). 10.7% of the patients lost to follow-up for the ePQ were indicated as untraceable. Of patients aged >40 years at ePQ completion, 56/707 [7.9%] could not be traced. In those ≤40 years at ePQ completion, 66/434 [15.2%] of patients could not be traced.

In the UC patient group, age >40 years and in the combined CD and UC patient groups active disease recorded within 3 mo previous to ePQ completion were significant positive predictors of ePQ response. Gender, center, diagnosis, age at diagnosis and disease location showed no significant differences in ePQ response in centers that had complied with the 60% minimal ePQ

Table 4 Results of univariate and multivariate logistic regression analyses for identification of predictors of ePQ completion in the entire cohort

CD+UC					
<i>N</i> total = 1 141	<i>n</i>	Response (%)	Unadjusted Odds Ratio (95%CI)	Adjusted Odds Ratio (95%CI)	<i>P</i>
CD/UC	365/776	289 (79)/588 (76)	1.22 (0.90-1.64)	1.24 (0.91-1.70)	0.171
Female/male	552/589	429 (77)/448 (76)	1.10 (0.84-1.44)	1.09 (0.82-1.44)	0.564
Age at ePQ completion >40/≤40 yr	707/434	559 (79)/318 (73)	1.38 (1.04-1.82)	1.53 (1.14-2.05)	0.004
Flare/no flare ¹	88/1 053	80 (91)/797 (76)	3.21 (1.53-6.73)	3.32 (1.57- 7.03)	0.002

¹Flare within 3 mo period previous of the moment of ePQ completion. *N*: number of patients; CI: confidence interval.

Table 5 Results of univariate and multivariate logistic regression analyses for identification of predictors of ePQ completion in the UC population separately

UC					
<i>N</i> total = 776	<i>n</i>	Response (%)	Unadjusted Odds Ratio (95%CI)	Adjusted Odds Ratio (95%CI)	<i>P</i>
Female/male	370/406	280 (76)/308 (76)	0.99 (0.71-1.38)	1.01 (0.72-1.42)	0.956
Age at ePQ completion >40/≤40 yr	520/256	409 (78)/179 (70)	1.59 (1.13-2. 23)	1.70 (1.19-2.41)	0.003
Flare/no flare ¹	55/721	48 (87)/540 (75)	2.30 (1.02-5.17)	2.35 (1.03-5.37)	0.043
Left sided/rectum	333/221	252 (76)/169 (77)	0.96 (0.64-1.43)	0.99 (0.65-1.51)	0.965
Pancolitis/rectum	185/221	140 (76)/169 (77)	0.96 (0.61-1.51)	1.04 (0.65- 1.67)	0.875
Unknown/rectum	37/221	27 (73)/169 (77)	0.83 (0.38-1.83)	0.71 (0.31-1.64)	0.426

¹Flare within 3 mo period previous of the moment of ePQ completion. *N*: number of patients; CI: confidence interval; Left sided: rectum and sigmoid and/or descending colon; Unknown: unknown disease location.

Table 6 Results of univariate and multivariate logistic regression analyses for identification of predictors of ePQ completion in the CD population separately

CD					
<i>N</i> total = 378	<i>n</i>	Response (%)	Unadjusted Odds Ratio (95%CI)	Adjusted Odds Ratio (95%CI)	<i>P</i>
Female/male	190/188	154 (81)/144 (77)	1.31 (0.80-2.15)	1.35 (0.79-2.30)	0.270
Age at ePQ completion >40/≤40 yr	196/182	156 (80)/142 (78)	1.10 (0.67-1.80)	1.34 (0.79-2.30)	0.281
Flare/no flare ¹	34/344	33 (97)/265 (77)	9.84 (1.32-73.1)	10.9 (1.43-82.4)	0.021
Ileo colon/upper	136/20	110 (81)/16 (80)	1.06 (0.33-3.42)	0.91 (0.26-3.19)	0.879
Ileum/upper	54/20	41 (76)/16 (80)	0.79 (0.22-2.78)	0.95 (0.27-3.35)	0.631
Colon/upper	153/20	122 (80)/16 (80)	0.98 (0.31-3.15)	0.71 (0.18-2.77)	0.622
Unknown/upper	15/20	9 (60)/16 (80)	0.38 (0.083-1.69)	0.38 (0.075-1.90)	0.237

¹Flare within 3 mo period previous of the moment of ePQ completion. *n*: number of patients; CI: confidence interval; Upper: esophagus and/or stomach and/or duodenum and/or jejunum; Unknown: unknown disease location.

response in neither the uni- nor multivariate analyses. Details of the results of these analyses are displayed in Tables 4, 5, and 6.

Results of identical analyses performed in the patients of non-complying centers revealed no different viewpoints.

DISCUSSION

An Internet-based data acquisition tool sustaining a unique European multinational multi-center 10-year clinical follow-up study in patients afflicted with IBD appeared successful. User-friendly tailor-made software applications facilitated remote data inclusion, making it a one-step process, thereby minimizing the risk of human error, optimizing efficiency and convenience of the study effort and rendering continuous transparency of the project progress. The data management procedure made

a reliable analyzable data set available promptly after the completion of the data acquisition phase. Overall 67% and 74% follow-up rates were observed for ePQ and ePpPFU, respectively. UC patients older than 40 years at the moment of ePQ completion and UC and CD patients with active disease in the 3 mo period previous to ePQ completion were significantly more likely to respond.

Does the use of Internet-based data acquisition instruments compare to or even outweigh the traditionally available methods? Some comparative studies are available in this field. In one study comprising UC patients, it appeared possible to use direct electronic mail contact to conduct follow-up research; response rates appeared to be related to the number of messages sent, the age of the recipients and the time since the initial contact^[12]. The use of Internet tools in psychological research and Web-based assessments of personality constructs have proven to render outcome scores comparable to those of traditional

methods^[13,14]. In one study the traditional paper-and-pencil questionnaire resulted in a higher number of errors compared with the Web-based questionnaire^[15]. Two other studies compared several student populations that were similar in terms of age, gender, and racial backgrounds with regard to differences in outcome for Web-based and traditional questionnaires^[16,17]. In these studies outcome scores of the Internet-based questionnaire and the traditional paper-and-pencil based questionnaires were not different. The time- and cost-effectiveness, convenience for both patients and researchers, the possibilities of immediate availability of data for analysis, and the structural prevention of missing data and incomplete responses by a priori programmed configuration, are important advantages related to the use of Internet-based data inclusion instruments as already discussed elsewhere^[18,19].

The type of survey population addressed is of importance in this context. In open Web-based surveys, selection bias occurs due to the non-representative nature of the Internet population, and - more importantly - through self-selection of participants, also called as the 'volunteer effect'^[18]. Furthermore, exact measurement of response rates is not possible, because the original size of the patient population is unknown. In quantitative epidemiological research, averages are measured of pre-defined patient populations, and therefore representative patient samples and adequate response rates are needed^[20]. In the present study a finite 'closed' patient sample with known socio-demographic details was approached for follow-up for the fourth time in its disease history. In this way the exact response rate could easily be calculated. The patients could not log on to the username-password secured web site by themselves, but were guided by a project representative who assisted with the insertion procedure according to the patients' individual needs.

The patient group under study, being a population-based cohort, was representative by its nature. High follow-up rates were observed in the majority of the centers complying with the minimally required 60% response rate. This reduced selection bias. Thirteen of the original twenty EC-IBD centers participated in this FU study. The remaining seven had to refrain because of technical and/or logistic reasons. This did not jeopardize the population-based character of the study, since all participating centers had individually met the criteria for population based patient inclusion when the cohort was formed in the period of 1991-1993. In the present study, centers not reaching the minimally required 60% follow-up rate were analyzed separately acknowledging their effort and also securing the intent of overall population based methodology. The original population size of almost 930 000 inhabitants negatively influenced the follow-up rates in Dublin. The magnitude of the Irish project covering six university and six private hospitals made sufficient follow-up percentages, 10 years after the diagnosis, unrealistic. Most areas that reached the 60% follow-up rate threshold had original population sizes between 300 000 and 500 000 inhabitants, suggesting that this size enables

successful long-term follow-up. Considering the broad European multi-center nature of this study and the fact that the entire patient population had been diagnosed 10 years earlier, the high response rates must be regarded as excellent.

UC patients who responded to the ePQ were significantly older than those who did not. This reflects most probably the higher dedication and availability of this age group compared to the younger patients. In the CD group, there was no difference between the age groups. The expected drop of response of the elderly because of supposed computer anxiety could not be confirmed in this study. Eleven percent of the patients indicated to have been lost to follow-up for the ePQ were in fact traced, but refused to participate without a difference between the age groups. The need of physical presence to complete the ePQ could explain at least part of the observed refusal rate. Furthermore, active disease within 3 mo period from ePQ completion was shown to be a clear positive predictor of ePQ participation in both UC and CD, and the majority of patients (1 073/1 162 [92.3%]) were in clinical remission during this period. Most probably, disease activity accompanies increased concern and increases motivation to participate. This observation could be an important factor of bias that has to be controlled for in the quality of life analyses planned in this project. 10.7% of the patients lost to follow-up for the ePQ were indicated to be untraceable. Patients equal to or younger than 40 years at ePQ completion were less likely to be traced, compared to those patients being older than 40 years at ePQ completion. This difference could be explained by the high mobility characteristic of younger people.

Apart from 11 occasions of system failure and some initial inconsistencies of the data export facility, this electronic facility appeared to be a well-equipped flexible tool for data entry and storage. Possibilities for an efficiency gain in terms of a better fit between Internet questionnaire application and the data management process had been described before^[21] and could be confirmed in this study.

In conclusion, an Internet-based electronic data acquisition application successfully sustained a unique multinational multi-center study project in patients afflicted with IBD. The application has high potential for future use in follow-up studies of this cohort and could serve as a template for other multi-national follow-up studies.

ACKNOWLEDGMENTS

This study was initiated and carried out by all members of the European Collaborative study Group of IBD (EC-IBD).

REFERENCES

- 1 Bell DS, Kahn CE. Health status assessment via the World Wide Web. *Proc AMIA Annu Fall Symp* 1996: 338-342
- 2 Soetikno RM, Mrad R, Pao V, Lenert LA. Quality-of-life

- research on the Internet: feasibility and potential biases in patients with ulcerative colitis. *J Am Med Inform Assoc* 1997; **4**: 426-435
- 3 **Hilsden RJ**, Meddings JB, Verhoef MJ. Complementary and alternative medicine use by patients with inflammatory bowel disease: An Internet survey. *Can J Gastroenterol* 1999; **13**: 327-32
- 4 **Soetikno RM**, Provenzale D, Lenert LA. Studying ulcerative colitis over the World Wide Web. *Am J Gastroenterol* 1997; **92**: 457-460
- 5 **Shivananda S**, Lennard-Jones J, Logan R, Fear N, Price A, Carpenter L, van Blankenstein M. Incidence of inflammatory bowel disease across Europe: is there a difference between north and south? Results of the European Collaborative Study on Inflammatory Bowel Disease (EC-IBD). *Gut* 1996; **39**: 690-697
- 6 **Stockbrügger RW**, Russel MG, van Blankenstein M S. EC-IBD: a European effort in inflammatory bowel disease. *Eur J Intern Med* 2000; **11**: 187-190
- 7 **Lennard Jones JE**, Shivananda S. Clinical uniformity of inflammatory bowel disease a presentation and during the first year of disease in the north and south of Europe. EC-IBD Study Group. *Eur J Gastroenterol Hepatol* 1997; **9**: 53-59
- 8 **Witte J**, Shivananda S, Lennard-Jones JE, Beltrami M, Politi P, Bonanomi A, Tsianos EV, Mouzas I, Schulz TB, Monteiro E, Clofent J, Odes S, Limonard CB, Stockbrügger RW, Russel MG. Disease outcome in inflammatory bowel disease: mortality, morbidity and therapeutic management of a 796-person inception cohort in the European Collaborative Study on Inflammatory Bowel Disease (EC-IBD). *Scand J Gastroenterol* 2000; **35**: 1272-1277
- 9 **Salvarani C**, Vlachonikolis IG, van der Heijde DM, Fornaciari G, Macchioni P, Beltrami M, Olivieri I, Di Gennaro F, Politi P, Stockbrügger RW, Russel MG. Musculoskeletal manifestations in a population-based cohort of inflammatory bowel disease patients. *Scand J Gastroenterol* 2001; **36**: 1307-1313
- 10 **van der Eijk I R**, Russel M. Influence of quality of care on quality of life in inflammatory bowel disease (IBD): literature review and studies planned. *Eur J Intern Med* 2000; **11**: 228-234
- 11 **Gasche C**, Scholmerich J, Brynskov J, D'Haens G, Hanauer SB, Irvine EJ, Jewell DP, Rachmilewitz D, Sachar DB, Sandborn WJ, Sutherland LR. A simple classification of Crohn's disease: report of the Working Party for the World Congresses of Gastroenterology, Vienna 1998. *Inflamm Bowel Dis* 2000; **6**: 8-15
- 12 **Treadwell JR**, Soetikno RM, Lenert LA. Feasibility of quality-of-life research on the Internet: a follow-up study. *Qual Life Res* 1999; **8**: 743-747
- 13 **Riva G**, Teruzzi T, Anolli L. The use of the internet in psychological research: comparison of online and offline questionnaires. *Cyberpsychol Behav* 2003; **6**: 73-80
- 14 **Cronk BC**, West JL. Personality research on the Internet: a comparison of Web-based and traditional instruments in take-home and in-class settings. *Behav Res Methods Instrum Comput* 2002; **34**: 177-180
- 15 **Pettit FA**. A comparison of World-Wide Web and paper-and-pencil personality questionnaires. *Behav Res Methods Instrum Comput* 2002; **34**: 50-54
- 16 **Fouladi RT**, McCarthy CJ, Moller NP. Paper-and-pencil or online? Evaluating mode effects on measures of emotional functioning and attachment. *Assessment* 2002; **9**: 204-215
- 17 **Davis RN**. Web-based administration of a personality questionnaire: comparison with traditional methods. *Behav Res Methods Instrum Comput* 1999; **31**: 572-577
- 18 **Eysenbach G**, Wyatt J. Using the Internet for surveys and health research. *J Med Internet Res* 2002; **4**: E13
- 19 **Schleyer TK**, Forrest JL. Methods for the design and administration of web-based surveys. *J Am Med Inform Assoc* 2000; **7**: 416-425
- 20 **Braithwaite D**, Emery J, De Lusignan S, Sutton S. Using the Internet to conduct surveys of health professionals: a valid alternative? *Fam Pract* 2003; **20**: 545-551
- 21 **Wright G**. The triple-s standard. Presented at the Association for Survey Computing conference "Open Standards: Breaking down the barriers" at Imperial College 2002