doi:10.1093/fampra/cmp067 Family Practice Advance Access published on 13 October 2009

The course of newly presented unexplained complaints in general practice patients: a prospective cohort study

H Koch[a], MA van Bokhoven[b], PJE Bindels[c], T van der Weijden[b], GJ Dinant[b] and G ter Riet[a,d]

Koch H, van Bokhoven MA, Bindels PJE, van der Weijden T, Dinant GJ and ter Riet G. The course of newly presented unexplained complaints in general practice patients: a prospective cohort study. *Family Practice* 2009; **26**: 455–465.

Objective. Newly presented unexplained complaints (UCs) are common in general practice. Factors influencing the transition of newly presented into persistent UCs have been scarcely investigated. We studied the number and the nature of diagnoses made over time, as well as factors associated with UCs becoming persistent. Finally, we longitudinally studied factors associated with quality of life (QoL).

Methods. Prospective cohort study in general practice of patients presenting with a new UC. Data sources were case record forms, patient questionnaires and electronic medical registries at inclusion, 1, 6 and 12 months. Presence of complaints and diagnoses made over time were documented. Potential risk factors were assessed in mixed-effect logistic and linear regression models.

Results. Sixty-three GPs included 444 patients (73% women; median age 42) with unexplained fatigue (70%), abdominal complaints (14%) and musculoskeletal complaints (16%). At 12 months, 43% of the patients suffered from their initial complaints. Fifty-seven percent of the UCs remained unexplained. UCs had (non-life-threatening) somatic origins in 18% of the patients. QoL was often poor at presentation and tended to remain poor. Being a male [odds ratio (OR) 0.6; 95% confidence interval (CI) 0.4–0.8] and GPs' being more certain about the absence of serious disease (OR 0.9; 95% CI 0.8–0.9) were the strongest predictors of a diminished probability that the complaints would still be present and unexplained after 12 months. The strongest determinants of complaint persistence [regardless of (un)explicability] were duration of complaints >4 weeks before presentation (OR 2.6; 95% CI 1.6–4.3), musculoskeletal complaint at baseline (OR 2.3; 1.2–4.5), while the passage of time acted positively (OR 0.8 per month; 95% CI 0.78–0.84). Musculoskeletal complaints, compared to fatigue, decreased QoL on the physical domain (4.6 points; 2.6–6.7), while presence of psychosocial factors decreased mental QoL (5.0; 3.1–6.9).

Conclusion. One year after initial presentation, a large proportion of newly presented UCs remained unexplained and unresolved. We identified determinants that GPs might want to consider in the early detection of patients at risk of UC persistence and/or low QoL.

Keywords. Cohort study, course, general practice, unexplained complaints.

Introduction

Unexplained complaints (UCs) have been defined as complaints of which the origins remain unclear to a GP after adequate history taking, physical examination and careful consideration of the psychosocial context.[1] On average, 13% of GP consultations involve UCs.[2] It is often assumed that most of the (medical) problems that give rise to newly presented UCs are self-limiting because patients often do not

Received 7 July 2008; Revised 22 July 2009; Accepted 10 September 2009.
[a]Department of General Practice, Division of Clinical Methods and Public Health, Academic Medical Center-University of Amsterdam, Amsterdam, The Netherlands, [b]Department of General Practice, School for Public Health and Primary Care (CAPHRI), Maastricht University, Maastricht, The Netherlands, [c]Department of general practice, Erasmus Medical Center, Rotterdam, The Netherlands and [d]Horten Centre, University of Zurich, Zurich, Switzerland. Correspondence to: H Koch, Department of General Practice, Division of Clinical Methods and Public Health, Academic Medical Centre-University of Amsterdam, PO Box 22660, 1100 DD Amsterdam, The Netherlands; Email: h.koch@amc.uva.nl

revisit their GPs on account of that specific UC.[3–5] For example, Kroenke *et al.* found that ~90% of UC patients did not revisit their GPs within 1 month.

Newly presented UCs may be a sign of underlying somatic or psychosocial pathology. Furthermore, they may remain unexplained and may be transformed into what may then be called persistent(ly) UCs to emphasize that the complaints as well as the inexplicability continue. In the literature, percentages of patients for whom these complaints are the reason to consult their GPs, vary from 3 to 39%.[6] Patients with persistent UCs (also called MUS for medically unexplained symptoms) often use many drugs and other health care facilities and frequently show considerable psychological distress.[7,8] Furthermore, persistent UCs are associated with a low quality of life (QoL).[9] Moreover, as these patients often do not feel taken seriously or helped and sometimes feel treated as malingerers, they tend to lose trust in their GPs.[10]

So far, the transition of newly presented UCs into persistent UCs and its risk factors have scarcely been investigated although a better understanding of the mechanisms involved may eventually lead to better prevention. Our objectives were to find how many and what type of diagnoses are made within 1 year after first presentation of a UC, which proportion of newly presented UCs persist, and to identify factors associated with becoming persistent. Finally, we longitudinally studied risk factors of poor QoL.

Materials and Methods

General design

Between February 2002 and December 2003, GPs from the southern and the western parts of The Netherlands recruited patients for a prospective 12-month follow-up cohort study on newly presented UCs. The Medical Ethics Committees of the University of Amsterdam and Maastricht University approved the study protocol. The study was part of a randomized trial on the value of blood test ordering for patients with UCs.[11,12]

Patients and complaints

GPs included patients >18 years of age, presenting with a new complaint that the GP designated as 'unexplained' at the end of the first consultation. GPs were given the Dutch College of General Practitioners' definition of UCs: those complaints that remain of unclear origin to the GP after adequate history taking, physical examination and careful consideration of the patient's psychosocial context. Eligible were UCs that concerned fatigue, abdominal complaints, musculoskeletal complaints, weight changes or itch. A UC was called new if it had not been presented to the GP within the 6 months prior to the visit. Patients were instructed to revisit their GPs when their complaints had not resolved after 4 weeks. Each patient presented no more than a single UC. Patients gave written informed consent after having read the study information at home. They received this information from their GPs after the first consultation and returned it by mail.

Measurements

GPs filled out structured case record forms (CRFs) immediately after the entry consultation and again when patients revisited with the same complaint after 4 weeks as instructed. CRFs included detailed questions about the history, such as type of complaint, duration of complaint or symptoms, findings on physical examination (if performed) and some general questions about how the GP appraised the consultation [degree to which s/he felt the complaint was unexplained, degree of certainty about the absence of serious disease, (certainty about) working hypothesis and his/her satisfaction]. After the entry consultation (T0), patients were given a patient questionnaire enquiring after demographic characteristics, intensity of complaints, satisfaction with care, anxiety, QoL and health care utilization. Patients filled out the same questionnaires at 4–6 weeks (T1), 6 months (T6) and 12 (T12) months. These follow-up questionnaires included a question on the presence of the complaint.

Generic health-related QoL was measured by means of the RAND-36. The RAND-36[13] is a Dutch version of the SF-36[14] and consists of 36 questions and standardized response choices, organized into eight multi-item domains: physical functioning, role limitations due to physical health problems, social functioning, general mental health, role limitations due to emotional problems, vitality, bodily pain and general health perception. Domain scores may vary from 0 to 100 (where 0 indicates 'very poor' and 100 'excellent'). From the eight RAND-36 domains, summary physical component scale (PCS) and mental component scale (MCS) were constructed.[15] The EuroQol's thermometer was assessed as a proxy for QoL as well. The thermometer provides a single index value (0–100) for health status.[16]

At 12 months, electronic medical records (EMRs) were retrospectively searched for all complaint-related entries such as diagnoses, working hypotheses and complaint level evaluations in the previous year. The final categorization of the complaints was carried out in three steps. First, two of the authors (HK and MAvB) independently searched the EMRs for all complaint-related GP entries and summarized these into one diagnosis per patient. To obtain a single diagnosis, we first individually gathered a maximum of three complaint-related entries throughout the year and from these selected the one that seemed most important. Discordances were resolved by discussion.

Unresolved discordances were settled by one of two experienced GPs (PJEB and GJD). Second, these diagnoses were categorized in 'complaint not otherwise specified' (e.g. fatigue and abdominal complaints; final evaluation identical to initial complaint), 'psychosocial illness' (e.g. depression and burn out), 'somatic illness' (e.g. diabetes mellitus and hypothyroidism), 'syndrome' (e.g. irritable bowel syndrome and chronic fatigue syndrome) or 'otherwise' (for all evaluations in the EMR with a question mark or listings of differential diagnoses). In this step, discordances were resolved as described above. Finally, the categories 'complaint not otherwise specified' and 'otherwise' were collapsed into 'unexplained', whereas the psychosocial illness, somatic illness and syndrome categories were collapsed into 'explained'.

Statistical analysis

We summarized the diagnoses made and course of complaints over time. We assessed determinants of (i) the probability that the initial complaint was still unexplained and present at 12 months; (ii) the longitudinal course of the initial complaint at 1, 6 and 12 months, regardless of whether it became explained and (iii) the longitudinal course of the QoL, separately for summary scales PCS and MCS of the RAND-36. Table 1 shows which risk factors were assessed for each of these analyses.

To study the effect of potential risk factors, we used longitudinal maximum likelihood mixed-effect logistic and linear regression analysis with complaint present 'yes/no' and RAND-36 PCS or RAND-36 MCS as the dependent variables, respectively. Prior to these, we familiarized ourselves with the data using dedicated descriptive commands for longitudinal data and line graphs to study the unadjusted course of the dependent variables over time or after stratification for single determinants (see Fig. 1 for an example). In the longitudinal analyses, the time variable had the values 1, 6 and 12 for the follow-up times at 1, 6 and 12 months, respectively.

Missing data were multiply ($m = 10$) imputed using iterative chained equations. Briefly, for each variable in turn missing values are filled in with randomly chosen observed values. Then, 'filled-in' values in the first variable are removed, leaving the original missing values. These missing values are then imputed using regression imputation on all other variables. Next, the filled-in values in the second variable are removed. These missing values are then imputed using regression imputation on all other variables. This process is repeated for each variable. Once each variable has been

TABLE 1 *Potential determinants used in the models*

Determinant	Source	Answer categories	Used in		
			Model 1	Model 2	Model 3
Patient's age	CRF/PQ	Years	x	x	x
Patient's sex	CRF/PQ	Male/female	x	x	x
Type of complaint	CRF	Fatigue/abdominal/ musculoskeletal	x	x	x
Duration of complaints before presentation	CRF	Less/more than 4 weeks	x	x	x
Patient's perception of complaint intensity	PQ	Bearable/unbearable	x	x	x
RAND-36 PCS at inclusion	PQ	0 through 100	x	x	
RAND-36 MCS at inclusion	PQ	0 through 100	x	x	
EuroQol's thermometer at inclusion	PQ	0 through 100	x	x	x
Intensity of complaints according to GP	CRF	Bearable/unbearable	x	x	x
Presence of psychosocial factors	CRF	Present/absent	x	x	x
Presence of previous episodes	CRF	Present or absent/unknown	x	x	x
Performance of physical examination	CRF	Yes/no	x	x	x
Presence of abnormal findings on physical examination	CRF	Yes/no	x	x	x
Degree to which complaint is unexplained	CRF	0 (totally unexplained) to 6 (a little unexplained)	x	x	x
GP's certainty on absence of serious disease	CRF	0 (totally uncertain) to 6 (totally certain)	x	x	x
GP's certainty on working hypothesis	CRF	0 (totally uncertain) to 6 (totally certain)	x	x	x
GP's satisfaction with the consultation	CRF	0 (very dissatisfied) to 10 (very satisfied)	x	x	x
Diagnosis made within 12 months	EMR	Yes/no	x	x	x
Time	CRF/PQ	0 (baseline); 1 (1 mo ab); 6 (6 mos ab); 12 (12 mos ab)	x	x	x
Intervention arm		Immediate BTO/postponement of BTO	x	x	x

Model 1: assess determinants of probability that, 12 months after baseline, the initial complaint was still unexplained and present. Model 2: assess determinants of the longitudinal course of the main complaint (regardless whether it became explained or remained unexplained). Model 3: assess determinants of the longitudinal course of the QoL. PQ, patient questionnaire; BTO, blood test ordering; mo(s), month(s); ab, after baseline.

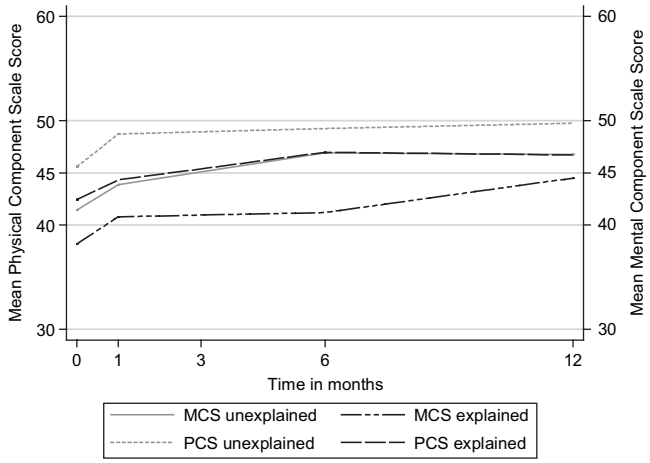

FIGURE 1 *Line graph showing the (unadjusted) course of mean QoL [physical (left y-axis) and mental (right y-axis)] over 12 months. Note that after an initial and modest rise between baseline and 1 month, the mean QoL scores remain stable and fairly low in all subgroups. At baseline all complaints are unexplained. Note that the GP finding an apparent explanation for the initial complaint during the 12-month follow-up had no clear impact on the QoL differences that were already present between the groups at baseline for PCS and MCS, respectively; this was confirmed in the multivariable analysis. MCS unexplained/explained, mean MCS scores of patients with newly presented UCs that remained unexplained/became explained over time; PCS unexplained/explained, mean PCS scores of patients with newly presented UCs that remained unexplained/became explained over time.*

TABLE 2 *Background characteristics (N = 444).*

Median age in years (p10, p25, p75, p90)	42.3 (22.2, 30.6, 53.5, 66.5) *n* (%)
Complaint group	
Fatigue	308 (69)
Abdominal	63 (14)
Musculoskeletal	73 (17)
Sex	
Male	119 (27)
Female	325 (73)
Highest educational level	
None	3 (1)
Elementary	35 (8)
Secondary	287 (65)
Higher	93 (21)
Missing	26 (6)

p, percentile; e.g. p10 means that 10% of the patients were <22.2 years old.

imputed, we have completed one 'cycle'. The process is continued for 10 cycles.[17] Before imputation, we changed the 'long' data file format into a 'wide' format to ensure that one patient had one record while imputing, as recommended by Allison.[18–20] After imputation, the long data format needed for longitudinal analyses was restored.

Multiple imputation assumes that the mechanisms responsible for the missing values are captured with the variables available in the dataset,[21] the so-called missing-at-random (MAR) assumption. In accordance with the advice to use a rich model for imputation, we used 29 variables for imputation—seven patient-reported variables: sex, age, complaint type, duration of complaints at baseline, patients' perception of complaint intensity, presence of previous episodes, EuroQol's thermometer at inclusion; eight GP-reported variables: performance of physical examination, presence of abnormal findings at physical examination, degree to which the complaint was unexplained, certainty on absence of serious disease, certainty on working hypothesis, satisfaction with the consultation, perception of complaint intensity, presence of psychosocial factors; two EMR-based variables: group assignment in randomized trial, whether a diagnosis had been made within 12 months; nine outcome variables and the three time indicators. The MAR assumption is far weaker than its missing completely at random counterpart, which assumes that the missing data are a random sample of all data. For each patient, the number of missing values on time-independent covariates ranged between 0 and 12 (median number 0; p10 = 0; p90 = 4.5). In total, 541 missing values were imputed out of 5328 values (10%) (444 patients × 12 time-independent covariates).

We assessed the differences between those with complete and incomplete data using a logistic regression model (incomplete coded as 1 and complete as 0) with baseline variables as predictors. Under the null hypothesis of no differences between these groups, the odds ratios (ORs) associated with the predictors should be not significantly different from unity. We studied the relationships between continuous predictors and the dependent variables using scatter plots of locally weighted regressions (lowess plots) to check linearity assumptions and avoid model misspecification. We chose not to impute missing values on dependent variables in accordance with recommendations from the statistical literature.[22,23] However, dependent variables were used to impute missing values on the independent variables.[24]

The final regression models were fit using the Li–Raghunathan–Rubin formula to estimate the mean coefficients and the correct variance across the 10 imputed sets.[19,20] Likelihood ratio (chunk) tests ($P < 0.05$) were used to reduce the models including all predictors to leaner models[25] provided that the regression coefficients of the variables remaining in the model did not change by >10%.[26] This was also applied to decide on the necessity to keep the 63 GPs as random intercepts; the random intercepts for patients were always retained. Random slopes were not considered since we deemed the theory in this field of research insufficient to underpin potential random

slopes with good explanations. We considered *P*-values <0.05 as statistically significant. All analyses were performed using Stata software (College Station, TX, version 10.1).

Results

Study population

In total, 63 GPs in 57 practices included 513 patients with newly presented UCs. The median number of included patients per GP was 6 (interquartile range 3–10). We excluded the complaint categories for which few patients were included (weight changes, $n = 27$; itch, $n = 18$) to avoid the impression that our findings could be generalized to cover these poorly represented complaints. This left 467 patients with newly presented unexplained fatigue ($n = 324$), abdominal ($n = 66$) or musculoskeletal ($n = 77$) complaints. Twenty-three of these dropped out because of moving house or for other personal reasons. The background characteristics of the remaining 444 patients are presented in -Table 2. Patient classification in terms of complaint-related diagnoses/evaluations and presence of complaints 12 months after baseline are presented in Figure 2. Logistic regression analysis showed that no baseline

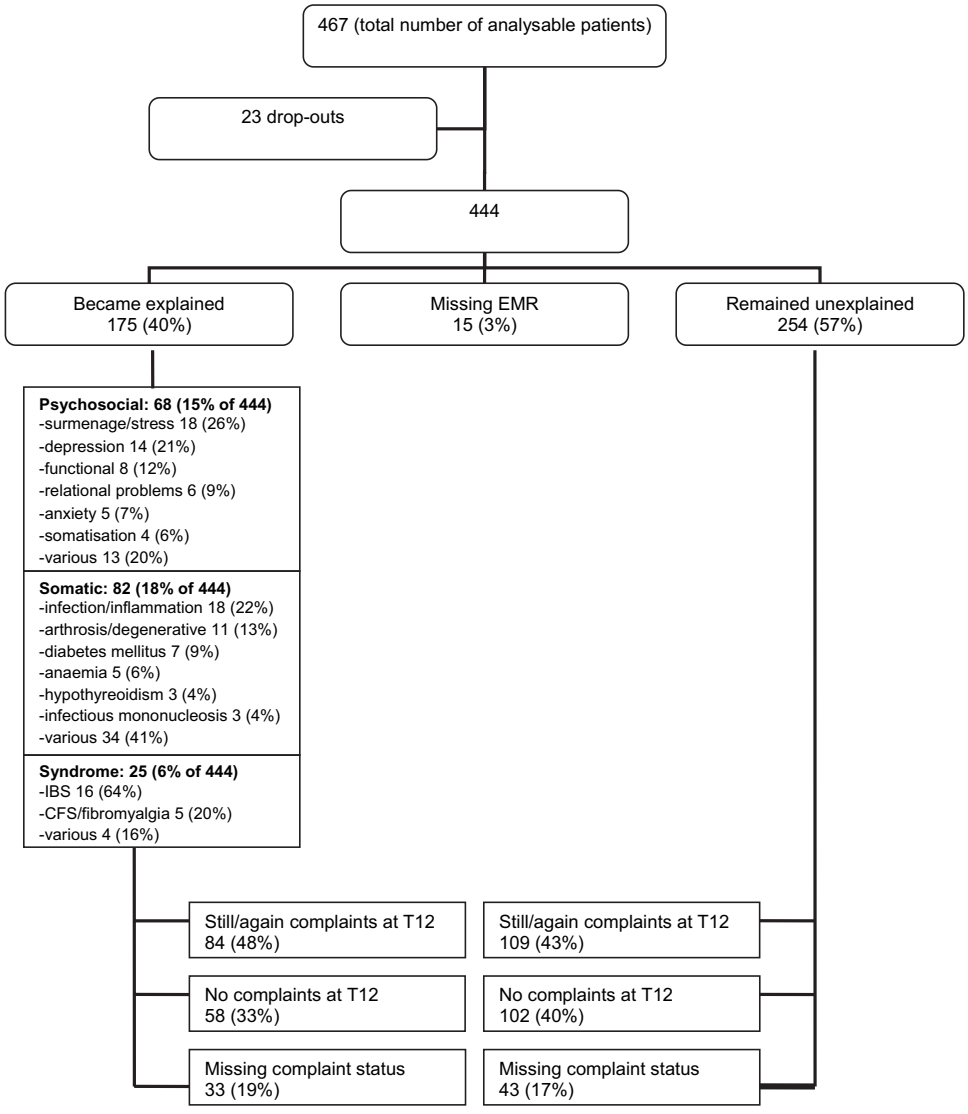

FIGURE 2 *Patient classification in terms of complaint-related diagnoses or evaluations and the presence or absence of complaints 12 months after inclusion.*

variables were associated with whether data for a patient were complete. Table 3 shows a comparison between patients with complete data and those with at least one missing value for age, sex and complaint group. Table 4 shows that the distributions of the observed values were very similar to those after imputation.

Complaint-related diagnoses/evaluations

In total, 254 [57%; 95% confidence interval (CI) from 53 to 62] of the newly presented UCs remained unexplained after 1 year. Out of 175 (39%; 95% CI 35 to 44) newly presented UCs that became explained during the 12-month follow-up, 82 explanations were of somatic origin (18% of the total of 444; 95% CI 15–22) and 68 of psychosocial origin (15% of 444; 95% CI 12–18). Overall, in 15 (3%; 95% CI 2 to 6) patients, data from the EMRs were missing. Figure 2 shows complaint-related diagnoses made within the psychosocial, somatic and syndrome categories.

Presence/absence of complaints over time

Regardless of whether their complaints became explained or not, 193 of 444 patients (44%; 95% CI 39–48) with newly presented UCs claimed that they suffered from these complaints at 12 months. This was the case in 109/254 (43%; 95% CI 37–49) patients with complaints that remained unexplained and 84/175 (48%; 95% CI 41–55) of patients with complaints that became explained during the follow-up. In 120/444 patients (27%; 95% CI 23–31), the complaints were present at all four times points. Forty-eight patients (11% of 444; 95% CI 8–14) complained only at inclusion. In total, 276 patients (62% of 444; 95% CI 58–67) had variable patterns of presence/absence of complaints at 1, 6 and 12 months.

Determinants of the course of complaints and QoL

There were four time points for follow-up, implying 1776 (4 × 444) potential observations. Since some dependent variables were missing, the number of observations in the regression models varied. The final models are based on a minimum of 1426 (81%) to 1628 (92%) observations. Full and lean models are presented in Tables 5, 6, 7 and 8.

Four predictors were found of the probability of newly presented UCs remaining unexplained and present at 12 months. Being a male (OR 0.6; 95% CI 0.4–0.8), having abdominal complaints (OR 0.5; 95% CI 0.3–0.8), the GP's perception of the complaint intensity being unbearable (OR 0.4; 95% CI 0.2–0.7) and GPs being more certain about the absence of serious disease (OR 0.9 per point higher on the scale; 95% CI 0.8–0.99) all diminished the probability that 12 months after baseline, the complaint would still be unexplained and present. Table 5 shows the OR and 95% CI of all determinants in the model.

Strong predictors of the probability that complaints continued to be present over time were as follows: having musculoskeletal complaints (OR 2.3; 95% CI 1.2–4.5), >4 weeks duration of the complaint before first presentation (OR 2.6; 95% CI 1.6–4.3), the MCS of the RAND-36 at presentation (OR 0.95 per point higher on the scale; 95% CI 0.93–0.97), the PCS of the RAND-36 at presentation (OR 0.9 per point higher on the scale; 95% CI 0.88–0.92), presence of

TABLE 3 *Comparison of age, sex and complaint group for patients with complete data (n = 190) and those with at least one missing value (n = 254)*

	Complete cases (n = 190)	Incomplete cases (n = 254)
Age (SD)	44.4 (14.81)	42.7 (16.54)
Sex (male, %)	26.8	26.8
Complaint group (%)		
Fatigue	72.6	66.9
Abdominal complaints	14.7	13.8
Muscle and joint	12.6	19.3

In total, 190/444 (42.8%) of all patients had complete data; 254/444 (57.2%) had at least one missing value. The median number of missing values in the latter group was 6 (interquartile range from 2 to 13).

TABLE 4 *Distributions of variables as observed and after multiple (10-fold) imputation*

Variable	Observed	Imputed (n = 4440)
Duration of complaint before presentation >4 weeks (%)	76.1 (n = 402)	76.2
Patient's perception of complaint intensity as unbearable (%)	25.3 (n = 384)	25.8
EuroQol's thermometer at inclusion, mean (SD)	58.9 (19.11) (n = 425)	58.8 (19.15)
Intensity of complaints unbearable according to GP	9.0 (n = 389)	10.6
Presence of psychosocial factors (%)	36.5 (n = 395)	36.2
Presence of previous episodes (%)	27.3 (n = 406)	27.4
Performance of physical examination (%)	72.7 (n = 406)	72.5
Presence of abnormal findings at physical examination (%)	20.7 (n = 406)	21.8
Degree to which complaint is unexplained, mean (SD)	3.69 (1.80) (n = 397)	3.68 (1.83)
GP's certainty on absence of serious disease, mean (SD)	5.61 (1.31) (n = 394)	5.57 (1.32)
GP's certainty on working hypothesis, mean (SD)	5.39 (1.32) (n = 385)	5.35 (1.35)
GP's satisfaction with the consultation, mean (SD)	7.29 (1.50) (n = 398)	7.27 (1.52)

TABLE 5 *Determinants of probability that, 12 months after baseline, the initial complaint was still unexplained and present*

Determinant (reference value)	Full model OR (95% CI)	Lean model OR (95% CI)
Patient's age (1 year older)	1.00 (0.99 to 1.01)	
Patient's sex (male)	0.48 (0.32 to 0.71)	0.60 (0.43 to 0.83)
Type of complaint		
Fatigue (reference category)	1 (n.a.)	1 (n.a.)
Abdominal complaints	0.51 (0.29 to 0.88)	0.54 (0.34 to 0.84)
Musculoskeletal complaints	0.75 (0.44 to 1.27)	1.20 (0.81 to 1.76)
Duration of complaints before presentation (over 4 weeks)	1.03 (0.69 to 1.53)	
Patient's perception of complaint intensity (unbearable)	0.86 (0.54 to 1.37)	
RAND-36 PCS at inclusion (1 point higher)	1.00 (0.98 to 1.02)	
RAND-36 MCS at inclusion (1 point higher)	1.00 (0.98 to 1.02)	
EuroQol's thermometer at inclusion (1 point higher)	0.99 (0.98 to 1.00)	
GP's perception of complaint intensity (unbearable)	0.38 (0.18 to 0.78)	0.39 (0.22 to 0.71)
Presence of psychosocial factors (yes)	1.16 (0.80 to 1.68)	
Presence of previous episodes (yes)	0.94 (0.65 to 1.36)	
Performance of physical examination (yes)	0.83 (0.56 to 1.23)	
Presence of abnormal findings on physical examination (yes)	0.86 (0.55 to 1.35)	
Degree to which complaint is unexplained (1 point less unexplained)	1.09 (0.97 to 1.22)	
GP's certainty on absence of serious disease (1 point more certain)	0.77 (0.66 to 0.90)	0.88 (0.78 to 0.99)
GP's certainty on working hypothesis (1 point more certain)	1.06 (0.90 to 1.23)	
GP's satisfaction with the consultation (1 point more satisfied)	1.10 (0.96 to 1.25)	
Intervention arm (immediate BTO)	1.02 (0.31 to 3.33)	
Time (1 measurement time further)	n.a.	
Establishment of explanation (diagnosis/evaluation) (yes)	n.a.	
	Lean model's variance:	
	Variance at GP level	9.07 (3.03 to 80.11)

BTO, blood test ordering; n.a., not applicable.

previous episodes >6 months prior to presentation (OR 2; 95% CI 1.2–3.4), passage of time (OR 0.81; 95% CI 0.8–0.84) and the GP documenting an explanation for the complaint (OR 0.5; 95% CI 0.3–0.8) (Table 6).

Longitudinal analysis showed that musculoskeletal complaints (–4.6; 95% CI –6.7 to –2.6), presence of psychosocial factors (OR 2; 95% CI 0.5–3.5), previous episodes of the same complaints >6 months prior to presentation (–2; 95% CI –3.7 to –0.4) and age (–0.09 per year; 95% CI –0.1 to –0.05) negatively influenced the course of physical QoL. Passage of time (0.3; 95% CI 0.2–0.4), the GP documenting an explanation for the complaint (OR 1.1; 95% CI 0.06–2.2) and the EuroQol thermometer score at presentation (0.1 for each point; 95% CI 0.08–0.2) positively influenced this course (Table 7).

Previous episodes of the same complaints >6 months before presentation (–2.8; 95% CI –4.9 to –0.7) and the presence of psychosocial factors (–5; 95% CI –6.9 to –3.1) negatively influenced the course of mental QoL. Musculoskeletal complaints (5.5; 95% CI 3–7.9) in comparison to being tired (fatigue), passage of time (0.4; 95% CI 0.3–0.5) and the EuroQol thermometer score at presentation (0.2 for each point; 95% CI 0.1–0.21) positively influenced this course (Table 8).

Discussion

In this prospective cohort study among patients with newly presented UCs in general practice, we found that 43% suffers from these complaints 12 months after initial presentation. Overall, almost 60% of newly presented UCs remained unexplained during 1 year. In many patients, UCs followed a varying pattern of presence and absence. About 40% of UCs received an explanation. Male sex, abdominal complaints, as compared to fatigue, and increased GPs' certainty about the absence of serious disease diminished the probability that the complaints would be unexplained and still present after a year. Longer duration of the complaint before presentation was associated with lower probability of it resolving, regardless of it remaining unexplained. Obviously, first presentation of a problem to the GP does not necessarily imply that this problem is new to the patient. Patients with new UCs tended to have a poor QoL; we found no strong determinants of variation in this course. Especially musculoskeletal complaints and previous episodes of the same complaints were associated with a negative course of the QoL on the physical dimension. The course of the mental summary scale of QoL was negatively influenced by fatigue, previous episodes of the same complaints and the presence of psychosocial problems.

TABLE 6 *Determinants of an unfavourable longitudinal course of the initial complaint in terms of presence of complaints (regardless whether they were explained or remained unexplained)*

Determinant (reference value)	Full model OR (95% CI)	Lean model OR (95% CI)
Patient's age (1 year older)	0.99 (0.97 to 1.00)	
Patient's sex (male)	1.08 (0.65 to 1.81)	
Type of complaint		
Fatigue (reference category)	1 (n.a.)	1 (n.a.)
Abdominal complaints	1.73 (0.88 to 3.40)	1.67 (0.88 to 3.15)
Musculoskeletal complaints	2.44 (1.16 to 5.13)	2.29 (1.16 to 4.51)
Duration of complaints before presentation (over 4 weeks)	2.56 (1.52 to 4.34)	2.62 (1.59 to 4.33)
Patients perception of complaint intensity (unbearable)	1.12 (0.60 to 2.08)	
RAND-36 PCS at inclusion (1 point higher)	0.89 (0.86 to 0.93)	0.90 (0.88 to 0.92)
RAND-36 MCS at inclusion (1 point higher)	0.94 (0.92 to 0.97)	0.95 (0.93 to 0.97)
EuroQol's thermometer at inclusion (1 point higher)	1.01 (1.00 to 1.02)	
GP's perception of complaint intensity (unbearable)	0.95 (0.36 to 2.48)	
Presence of psychosocial factors (yes)	1.22 (0.73 to 2.04)	
Presence of previous episodes (yes)	1.95 (1.14 to 3.34)	2.04 (1.21 to 3.43)
Performance of physical examination (yes)	0.86 (0.51 to 1.46)	
Presence of abnormal findings on physical examination (yes)	1.10 (0.59 to 2.05)	
Degree to which complaint is unexplained (1 point less unexplained)	0.96 (0.84 to 1.10)	
GP's certainty about absence of serious disease (1 point more certain)	0.90 (0.73 to 1.11)	
GP's certainty about working hypothesis (1 point more certain)	1.03 (0.85 to 1.25)	
GP's satisfaction with the consultation (1 point more satisfied)	0.97 (0.82 to 1.14)	
Intervention arm (immediate BTO)	0.96 (0.56 to 1.65)	
Time (1 measurement time further)	0.81 (0.78 to 0.84)	0.81 (0.78 to 0.84)
Establishment of explanation (diagnosis/evaluation) (yes)	0.50 (0.31 to 0.80)	0.50 (0.31 to 0.80)
	Lean model's variance:	
	Variance at GP level	1.29 (1.06 to 2.89)
	Residual error	4.17 (2.29 to 11.65)

BTO, blood test ordering; n.a., not applicable.

Little research on the course of UCs and their determinants has been performed so far. A recently published systematic review by Olde Hartman *et al.*[27] concludes that due to the limited numbers of studies and their heterogeneity, there is a lack of rigorous empirical evidence to identify relevant prognostic factors in patients presenting persistent medically unexplained symptoms (MUS). They also conclude that it seems that a more serious condition at baseline is associated with a worse outcome. When 'longer duration of complaints before presentation' and 'having had previous episodes of the same complaints' are considered as a 'more serious condition at baseline', this fits in with our findings that a longer duration of complaints before presentation is associated with a poorer prognosis of complaint duration and that having had previous episodes of a complaint negatively influences the course of the QoL.

Some limitations in our study deserve attention. First of all, inclusion may not have been consecutive. Sixty-three participating GPs identified between 3 and 10 patients with newly presented UCs each during a 23-month period. Based on earlier studies, we suspect that they must have seen many more patients with UCs. In a non-inclusion study in the participating general practices, we searched the electronic medical files by means of text words for eligible but not-included patients with UCs. This study did not show major sex and age differences between included (n = 513) and not-included (n = 507) patients. However, differences may exist on other characteristics and this may limit the generalizability of our findings. To clarify, complete non-selection of, for example, men simply limits the formal generalizability of our findings to women. In reality, opinion on scientific generalizability, as opposed to statistical, may vary with opinions on the importance of sex as an effect modifier of the associations reported. So selection as such should be distinguished from selection bias, which has been defined as conditioning on factor that is affected by a determinant of interest and by the outcome.[28] If non-selection is only partial, for example, if some patients with fatigue were excluded, incorporation of fatigue in the regression model (through the complaint group variable) deals with this issue by adjusting for fatigue. Residual confounding is possible since the predictors on psychosocial problems and single items of the SF-36 (mental scale) may not have captured full information on anxiety, depression and self-efficacy.

We showed that UCs are not as frequently self-limiting as is sometimes assumed. Whereas Kroenke and Jackson[5] and Kenter *et al.*[3], for example, found that 90% of UCs are self-limiting, we found that >40% of patients experience their UCs after 1 year. One

TABLE 7 *Determinants of an unfavourable longitudinal course of the physical QoL in terms of lower scores on the summary scales*

Determinant (reference value)	Full model Coefficient (95% CI)	Lean model Coefficient (95% CI)
Patient's age (1 year older)	–0.09 (–0.14 to –0.04)	–0.10 (–0.14 to –0.05)
Patient's sex (male)	0.002 (–1.70 to 1.70)	
Type of complaint		
Fatigue (reference category)	1 (n.a.)	1 (n.a.)
Abdominal complaints	–1.41 (–2.58 to 2.30)	0.39 (–1.80 to 2.59)
Musculoskeletal complaints	–4.55 (–6.89 to –2.22)	–4.65 (–6.70 to –2.60)
Duration of complaints before presentation (over 4 weeks)	–1.40 (–3.23 to 0.44)	
Patient's perception of complaint intensity (unbearable)	–1.04 (–3.09 to 1.01)	
RAND-36 PCS at inclusion (1 point higher)	n.a.	n.a.
RAND-36 MCS at inclusion (1 point higher)	n.a.	n.a.
EuroQol's thermometer at inclusion (1 point higher)	0.12 (0.08 to 0.17)	0.13 (0.08 to 0.17)
GPs' perception of complaint intensity (unbearable)	0.98 (–2.11 to 4.07)	
Presence of psychosocial factors (yes)	1.80 (0.14 to 3.46)	1.96 (0.45 to 3.46)
Presence of previous episodes (yes)	–1.98 (–3.71 to –0.26)	–2.04 (–3.71 to –0.38)
Performance of physical examination (yes)	0.93 (–0.88 to 2.74)	
Presence of abnormal findings on physical examination (yes)	–1.22 (–3.36 to 0.93)	
Degree to which complaint is unexplained (1 point less unexplained)	0.12 (–0.35 to 0.60)	
GPs' certainty on absence of serious disease (1 point more certain)	0.35 (–0.36 to 1.06)	
GPs' certainty on working hypothesis (1 point more certain)	–0.32 (–1.00 to 0.36)	
GPs' satisfaction with the consultation (1 point more satisfied)	0.14 (–0.39 to 0.67)	
Intervention arm (immediate BTO)	–0.62 (–2.45 to 1.20)	
Time (1 measurement time further)	0.29 (0.21 to 0.36)	0.29 (0.22 to 0.36)
Establishment of explanation (diagnosis/evaluation) (yes)	1.13 (0.03 to 2.22)	1.11 (0.06 to 2.16)
	Lean model's variance:	
	Variance at GP level	1.48 (1.00 to 4.60)
	Variance at pt level	30.97 (22.25 to 43.85)
	Residual error	19.01 (16.37 to 22.16)

BTO, blood test ordering; n.a., not applicable; pt, patient.

explanation for this difference might be that the other studies consider complaints to be resolved when patients do not revisit their GP with the same complaints. Our findings were based on patient questionnaires on top of the EMR entries. Other data from our study show that patients do not frequently revisit their GP for their UCs, but when asked personally after 1 year, >40% of them reported to perceive complaints.

In the group of patients with established explained complaints, there are patients with somatic diagnoses of chronic diseases, such as diabetes mellitus and arthrosis. A small fraction of this group stated not to suffer from their complaints after 12 months. This may be an effect of their treatment. When treated properly, patients with diabetes mellitus may no longer suffer from fatigue, for example. Our data did not allow unambiguous analysis of the role of treatment effects.

We chose to categorize syndrome-type diagnoses (chronic fatigue syndrome, irritable bowel syndrome etc.) as 'explained complaints'. Moreover, the few somatoform disorder diagnoses were categorized as 'explained'. Had we categorized these diagnoses as 'unexplained', the percentage of persistently UCs would have been even higher (64% instead of 57%).

Surprisingly, if the GP perceived the complaints as being unbearable to the patient, this diminished the chance of the complaints being present and unexplained after 12 months. We do not have a plausible explanation for this finding and replication in future studies seems warranted.

Our study sheds some light on the transition of newly presented to persistent UCs and on determinants of the course of complaints and QoL. A complaint's duration of ⩾4 weeks before presentation appears to signal a lower probability that the complaints will resolve, as do poor perceived physical and mental well-being. If a diagnosis is made, the probability that the complaint will resolve increases, perhaps as a consequence of the treatment. The simple passing of time also has a beneficial effect. The few determinants influencing the QoL were relatively weak and their potential to change the unfavourable course appears limited.

In our opinion, our findings should stimulate GPs to take UCs more seriously as from their initial presentation. Paying attention to the determinants we found may be helpful in detection of patients with an unfavourable course of complaints and QoL among the patients with newly presented UCs in general practice.

Acknowledgements

The authors would like to thank all participating GPs and patients for their willingness to take part in this study.

TABLE 8 *Determinants of an unfavourable longitudinal course of the mental QoL*

Determinant (reference value)	Full model Coefficient (95% CI)	Lean model Coefficient (95% CI)
Patient's age (1 year older)	0.02 (–0.04 to 0.08)	
Patient's sex (male)	1.25 (–0.82 to 3.32)	
Type of complaint		
Fatigue (reference category)	1 (n.a.)	1 (n.a.)
Abdominal complaints	1.80 (–1.14 to 4.74)	2.46 (–0.25 to 5.17)
Musculoskeletal complaints	4.91 (2.11 to 7.72)	5.45 (3.00 to 7.90)
Duration of complaints before presentation (over 4 weeks)	–0.44 (–2.65 to 1.76)	
Patients' perception of complaint intensity (unbearable)	–1.77 (–4.26 to 0.72)	
RAND-36 PCS at inclusion (1 point higher)	n.a.	n.a.
RAND-36 MCS at inclusion (1 point higher)	n.a.	n.a.
EuroQol's thermometer at inclusion (1 point higher)	0.16 (0.10 to 0.21)	0.16 (0.11 to 0.21)
GPs' perception of complaint intensity (unbearable)	–3.02 (–6.72 to 0.69)	
Presence of psychosocial factors (yes)	–5.18 (–7.18 to –3.17)	–5.02 (–6.90 to –3.15)
Presence of previous episodes (yes)	–2.79 (–4.89 to –0.69)	–2.81 (–4.88 to –0.74)
Performance of physical examination (yes)	0.07 (–2.09 to 2.24)	
Presence of abnormal findings on physical examination (yes)	0.01 (–2.58 to 2.60)	
Degree to which complaint is unexplained (1 point less unexplained)	0.47 (–0.08 to 1.01)	
GPs' certainty on absence of serious disease (1 point more certain)	–0.55 (–1.42 to 0.31)	
GPs' certainty on working hypothesis (1 point more certain)	–0.32 (–1.12 to 0.49)	
GPs' satisfaction with the consultation (1 point more satisfied)	0.18 (–0.44 to 0.80)	
Intervention arm (immediate BTO)	0.50 (–1.35 to 2.35)	
Time (1 measurement time further)	0.39 (0.30 to 0.48)	0.37 (0.29 to 0.45)
Establishment of explanation (diagnosis/evaluation) (yes)	–0.72 (–2.07 to 0.64)	
	Lean model's variance:	
	Variance at GP level	86.22 (60.38 to 124.96)
	Residual error	44.82 (37.82 to 53.32)

BTO, blood test ordering; n.a., not applicable.

Declaration

Funding: Dutch Health Care Insurance Board (CVZ); Centraal Ziekenfonds health insurance company; The Netherlands Organization for Scientific Research (NWO) (920-03-127).
Ethical approval: Medical Ethics Committees of the University of Amsterdam and Maastricht University.
Conflict of interest: None of the sources of funding influenced the study design or the writing of the manuscript or the decision to submit the manuscript for publication.

References

[1] Dinant GJ, Wijk van MAM, Janssens HJEM *et al*. Dutch College of General Practitioners guideline 'Blood testing. General principles and use by GP' [NHG-standaard Bloedonderzoek. Algemene principes en uitvoering in eigen beheer]. *Huisarts Wet* 1994; **37:** 202–11.
[2] van der Weijden T, van Velsen M, Dinant GJ, van Hasselt CM, Grol R. Unexplained complaints in general practice: prevalence, patients' expectations, and professionals' test-ordering behavior. *Med Decis Making* 2003; **23:** 226–31.
[3] Kenter EG, Okkes IM, Oskam SK, Lamberts H. Tiredness in Dutch family practice. Data on patients complaining of and/or diagnosed with ''tiredness''. *Fam Pract* 2003; **20:** 434–40.
[4] Kenter EG, Okkes IM. [Patients with fatigue in family practice: prevalence and treatment]. *Ned Tijdschr Geneeskd* 1999; **143:** 796–801.
[5] Kroenke K, Jackson JL. Outcome in general medical patients presenting with common symptoms: a prospective study with a 2-week and a 3-month follow-up. *Fam Pract* 1998; **15:** 398–403.
[6] Burton C. Beyond somatisation: a review of the understanding and treatment of medically unexplained physical symptoms (MUPS). *Br J Gen Pract* 2003; **53:** 231–9.
[7] Katon WJ, Walker EA. Medically unexplained symptoms in primary care. *J Clin Psychiatry* 1998; **59:** 15–21.
[8] Olde Hartman TC, Lucassen PL, van de Lisdonk EH, Bor HH, van WC. Chronic functional somatic symptoms: a single syndrome? *Br J Gen Pract* 2004; **54:** 922–7.
[9] Feder A, Olfson M, Gameroff M *et al*. Medically unexplained symptoms in an urban general medicine practice. *Psychosomatics* 2001; **42:** 261–8.
[10] Hartz AJ, Noyes R, Bentler SE *et al*. Unexplained symptoms in primary care: perspectives of doctors and patients. *Gen Hos Psychiatry* 2000; **22:** 144–52.
[11] van Bokhoven MA, Koch H, van der WT *et al*. Blood test ordering for unexplained complaints in general practice: the VAMPIRE randomised clinical trial protocol. [ISRCTN55755886]. *BMC Fam Pract* 2006; **7:** 20.
[12] Koch H, van Bokhoven MA, ter RG *et al*. What makes general practitioners order blood tests for patients with unexplained complaints? A cross-sectional study. *Eur J Gen Pract* 2009; **15:** 22–8.
[13] van der Zee KI, Sandermann R. *Measuring General Health Status with RAND-36: A Manual. [Het meten van de algemene gezondheidstoestand met de RAND-36; een handleiding.].* Groningen, The Netherlands: Noordelijk Centrum voor Gezondheidsvraagstukken, NCG reeks meetinstrumenten, 1993.
[14] Ware JE, Sherbourne CD. The Mos 36-item short-form health survey (Sf-36) .1. Conceptual framework and item selection. *Med Care* 1992; **30:** 473–83.

[15] Ware JE Jr, Kosinski M, Keller SD. *SF-36 Physical and Mental Health Summary Scales: A User's Manual.* Boston, MA: Health Assessment Lab, New England Medical Center, 1994.
[16] EuroQol group. EuroQol—a new facility for the measurement of health-related quality of life. The EuroQol Group. *Health Policy* 1990; **16:** 199–208.
[17] Carpenter J, Kenward M. Missing data. http://www.lshtm.ac.uk/msu/missingdata/ (accessed on 19 July 2009).
[18] Allison PD. Missing Data. [36]. *Quantitative Applications in the Social Sciences.* London: Sage University Paper, 2002.
[19] Raghunathan TE, Lepowski JM, van Hoewyk J, Solenberger P. A multivariate technique for multiply imputing missing values using a sequence of regression models. *Sur Methodol* 2001; **27:** 85–95.
[20] Royston P. Multiple imputation of missing values. *Stata J* 2004; **3:** 227–41.
[21] Little RJA, Rubin DB. *Statistical Analysis with Missing Data.* New York: John Wiley & Sons, 2002.
[22] Albert PS. Longitudinal data analysis (repeated measures) in clinical trials. *Stat Med* 1999; **18:** 1707–32.
[23] Omar RZ, Wright EM, Turner RM, Thompson SG. Analysing repeated measurements data: a practical comparison of methods. *Stat Med* 1999; **18:** 1587–603.
[24] Sterne JA, White IR, Carlin JB *et al*. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009; **338:** b2393.
[25] Kleinbaum DG. *Logistic Regression: A Self-learning Text.* New York: Springer Verlag, 1994.
[26] Greenland S. Invited commentary: variable selection versus shrinkage in the control of multiple confounders. *Am J Epidemiol* 2008; **167:** 523–9.
[27] Olde Hartman TC, Borghuis MS, Lucassen PL *et al*. Medically unexplained symptoms, somatisation disorder and hypochondriasis: course and prognosis. A systematic review. *J Psychosom Res* 2009; **66:** 363–77.
[28] Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004; **15:** 615–25.