BMJ Open

The impact of referral templates on patient experience of the referral and care process. A cluster randomized trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-011651
Article Type:	Research
Date Submitted by the Author:	03-Mar-2016
Complete List of Authors:	Wåhlberg, Henrik; Universitetet i Tromso Institutt for Samfunnsmedisin ISM, ; Sykehuset Østfold Kalnes, Medical department Braaten, Tonje; University of Tromsø, Department of Community Medicine Broderstad, Ann Ragnhild; Uit the Arctic University of Norway, Faculty of Health; University Hospital of North Norway, medical department
Primary Subject Heading:	Health services research
Secondary Subject Heading:	General practice / Family practice, Patient-centred medicine
Keywords:	PRIMARY CARE, Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

Title: The impact of referral templates on patient experience of the referral and care process. A cluster randomized trial

Henrik Wåhlberg – corresponding author

Department of Community Medicine, UiT The Arctic University of Norway

Faculty of Health Sciences, UiT The Arctic University of Norway

9037 Tromsø, Norway

University Hospital of North Norway Harstad

St. Olavsgate 70, 9480 Harstad, Norway

henrik.wahlberg@uit.no/henrik_waahlberg@hotmail.com

Tonje Braaten

Department of Community Medicine, UiT The Arctic University of Norway

Faculty of Health Sciences, UiT The Arctic University of Norway

9037 Tromsø, Norway

tonje.braaten@uit.no

Ann Ragnhild Broderstad

Centre for Sami Health Research, UiT The Arctic University of Norway

UiT The Arctic University of Norway, 9037 Tromsø, Norway

University Hospital of North Norway Harstad

St. Olavsgate 70, 9480 Harstad, Norway

ann.ragnhild.broderstad@uit.no

Keywords: Patient experience. Referral. General practitioner.

Wordcount: 4002 words (excluding abstract, references, tables and author contributions)

ABSTRACT

Objectives: To evaluate if a referral intervention improves the patient experience of the referral and treatment process.

Setting: Interface between fourteen primary care surgeries and a district general hospital.

Participants: The fourteen GP surgeries (seven intervention, seven control) in the area around the University Hospital of North Norway Harstad were randomised and all completed the study. Consecutive individual patients were recruited at their hospital appointment. A total of 500 patients were recruited with 281 in the intervention and 219 in the control arm.

Interventions: Dissemination of referral templates for four diagnostic groups (dyspepsia, suspected colorectal cancer, chest pain and COPD) coupled with intermittent surgery visits by study personnel. Control arm continued standard referral practice. The intervention was in use for 2.5 years.

Outcome: Patient experience as measured by self-report questionnaires. General practitioners in the intervention group could not be blinded. Patients were blinded to intervention status. Analysis was based on single question comparison with a total score used to assess the effect of clustering.

Results: On the individual questions overall satisfaction was very high with only minor differences between the intervention and control group. Interestingly the most negative responses, in both groups, were concerned questions relating to patient interaction and information. In the regression model used to assess the effect of clustering being in the intervention group predicted a small 0.71 (95% CI -0.33, 1.74, p=0.180), non-significant, increase in the patient experience score with little evidence of clustering (The intracluster correlations coefficient was estimated at $2.19e^{-09}$).

Conclusion: In total this indicates no clear effect of the implementation of referral templates on the patient experience, in a setting of generally high patient satisfaction.

Trial registration: This trial has been registered at ClinicalTrials.gov. The trial registration number is NCT01470963.

ARTICLE SUMMARY

Strengths and limitations of the study

- Clinically relevant research in a regular district hospital setting
- High response rate
- Newly developed questionnaire hampers wider generalisation

INTRODUCTION

Evaluation of patient experience and satisfaction is widespread with a wealth of literature concerning the development and use of questionnaires[1-5]. The evaluation of patient experience can help drive quality improvement[6] and improved patient experience is associated with safety and clinical effectiveness[7].

Care coordination is an important aspect of a well-functioning high quality health service. It has been defined as "the deliberate integration of patient care activities between two or more participants involved in a patient's care to facilitate the appropriate delivery of health care services."[8]. In the US the National Quality Forum (NQF) has published preferred practices for care coordination, including transitions of care[9]. This report includes clear recommendations for participation of the patient, or his/her designee, in the decision, planning and execution of a care transition. This is important, as exemplified by a recent Australian article, where patients with colorectal cancer perceived that poor information exchange led to suboptimal care[10]. Hence assessing patient experience of the referral process may be beneficial in assessing the effect of a referral intervention.

This article presents the patient experience aspect of a cluster randomized study evaluating the effect of the implementation of referral templates for four diagnostic groups – dyspepsia, suspected colorectal cancer, chest pain and COPD - in the patient referral pathway[11]. Previously we have shown that the referral templates led to increased referral quality[12], and assessment is underway to evaluate whether the

templates have led to increase quality of care at the hospital. This publication aims to assess whether the implementation of a referral template in the transition of care from the general practitioner to the hospital has affected the patient experience of the care process.

METHODS

Study setting

In Norway the health care system is quite uniformly organised throughout the country. GPs act as gatekeepers to secondary care[13] with specialist health services delivered by governmentally owned regional health authorities, mainly through public hospitals. Some specialist outpatient care is delivered by private specialists, but this is mainly purchased by the regional health authorities. The access to private specialists in the geographical area of the current study is very limited.

Study design

The study was designed as a cluster randomized study with the general practitioner (GP) surgery as the clustering unit. A total of 14 surgeries were randomized, seven to the intervention and seven to the control group. The clustered design was chosen to avoid possible spill-over effect from the intervention to control GPs. Randomization was done by simple drawing by a person not connected to the research team, stratified by town vs countryside location of surgery.

As the intervention was to be actively used by the GPs the referring GP could not be blinded. Patients, hospital doctors and outcome evaluators were blinded to the intervention status of the patients. Due to the design of the intervention the referral letter would sometimes reveal the intervention status. No separate sample size calculation was performed for the patient experience outcome. The full study details are published in the methods paper[11].

Intervention

The intervention consisted of the distribution of four separate referral templates to the intervention surgeries. These templates covered four clinical areas (dyspepsia, suspected colorectal malignancy, chest pain and chronic obstructive pulmonary disease). The templates were to be use when initiating a new referral to the medical outpatient clinic at the University Hospital of North Norway, Harstad (UNN Harstad). The templates were distributed by the corresponding author (HW) during educational and/or lunch meetings and where provided as laminated reference sheets or in electronic form. In addition follow up visits were conducted regularly during the study period and intermittent mail leaflets and reminders were distributed to the intervention offices. Control offices continued standard referral practice.

Outcomes

The main outcome in the project was the quality of care delivered to each individual patient. In addition health process indicators such as correct prioritization

were recorded and referral quality was also compared between the intervention and control group. The current paper presents the patient experience aspect of the study, as measured by self-report questionnaires.

Participants

The GP surgeries primarily served by UNN Harstad were included in the randomization process. In 2013 these surgeries had a total list size of 39,523 patients. The individual patients were recruited from consecutive new patients referred to the medical outpatient clinics at UNN Harstad. Study information and a consent form were sent to each individual patient together with their appointment letter. Further information, including a new consent form if appropriate, was provided at their hospital appointment.

Recruitment

The study recruited patients for approximately 2.5 years and a total of 538 patients were included with 290 in the intervention arm and 227 in the control arm. Thirty-eight patients were excluded as they did not fill the inclusion criteria. In total this left 281 patients in the intervention arm and 219 patients in the control arm.

Questionnaire development

Multiple tools exist for measuring different aspects of care coordination[14] and patient experience, however no complete questionnaire were located that covered the area in the current study completely. Therefore a questionnaire was developed by combining two validated questionnaires regarding patient experiences and care coordination. The questions used were the full version of the Generic Short Patient Experiences Questionnaire (GS-PEQ) [15;16] and the two questions about health interaction from the Commonwealth Fund Survey 2010[17]. Three further questions were added to assess (1) who referred the patient, (2) if the referral was seen as appropriate and (3) an overall evaluation of the institution. Table 1 presents the questions in the questionnaire. GS-PEQ uses Likert style response categories. The health interaction questions had a yes/no response. Table 2 presents the answer categories for all items with numerical coding. The full questionnaire, including the demographic questions, is available upon request.

Table 1 – Questionnaire details

Question No	Wording of question
1	Did the clinicians[a] talk to you in a way that was easy to understand?
2	Do you have confidence in the clinicians' professional skills?
3	Did you get sufficient information about how examinations and tests were to be performed?
4	Did you get sufficient information about your diagnosis/conditions?
5	Did you perceive the treatment to be adapted to you situation?
6	Were you involved in decisions regarding your treatment?
7	Did you perceive the institution work practices to be well organized?
8	Did you perceive the equipment at the institution to be in good working order?
9	Overall, was the help and treatment you received at the institution satisfactory?
10	Do you believe that you were in any way given incorrect treatment (according to your own judgment)?
11	Did you have to wait before you were given an appointment at the institution?
12	Overall, what benefit have you had from the care at the institution?
13	Did the hospital specialist not have basic medical information from your GP about the reason for your visit or test results?
14	After your saw the hospital specialist did your GP not seem informed and up-to-date

	about the care you got from the specialist?
15	Was the referral to the outpatient department necessary (according to your own judgment)?
16a	Were you referred by your GP for the outpatient appointment?
16b	If no in question 16a; who referred you?
17	If you take an overview of your entire treatment process, how would you evaluate the institution?

[a] With 'clinicians' we mean those who had the main treatment responsibility. This is linguistically clearer in the Norwegian wording.

Table 2 – Answer categories with numerical coding

Question	Answer categories	Numerical coding
Q1-10	Not at all	1
	To a small extent	2
	To a moderate extent	3
	To a large extent	4
	To a very large extend	5
	Not applicable	
Q11	No	5
	Yes, but not too long	3.67
	Yes, quite long	2.33
	Yes, too long	1
Q12	No benefit	1
	Little benefit	2
	Some benefit	3
	Large benefit	4
	Very large benefit	5
	Not applicable	
Q13	Yes	1
	No	5
Q14	Yes	1
	No	5
	Have not visited the GP after the consultation	
Q15	Yes	5
	No	1
Q17	Much poorer than expected	1
	Somewhat poorer than expected	2
		3

As expected	4
Somewhat better than expected	5
Much better than expected	
Not applicable	

The questionnaire was piloted for content validity with four local health professionals; these felt that it covered the important aspects of patient experience and care coordination. It was then piloted with five outpatients with a median age of 72 years (average 68.8 years) to ensure face validity and acceptability. Two patients needed clarification on one of the questions before they felt they could answer, and the wording of this question was adjusted accordingly. The patients felt the questionnaire was acceptable, with logical response categories and that the questions covered their clinical path during the referral process well. These patients did not take part when the project was later initiated. No further formal evaluation of the questionnaire was carried out.

Questionnaire distribution

The questionnaire was mailed to patients who had consented to take part in the referral project presented above. To increase response rates a pre-paid response envelope was included, addresses were handwritten, the questionnaire was kept short and association with research bodies was indicated[3]. For non-responders one reminder was sent approximately one month after the first questionnaire, with a new questionnaire and pre-paid response envelope.

Ethics

The study followed the principles of the Helsinki Declaration. Before recruitment started it was presented to the Regional Ethical Committee for Medical Research in Northern Norway, who determined it not to be within the scope of the Health Research Act (REK NORD 2010/2259). The project has been approved by the Data Protection Officer for Research. The study is registered at www.clinicaltrials.gov with trial registration number NCT01470963. All patients provided written informed consent.

Imputation

Missing data hampers the analysis of full datasets, for analysis beyond simple single item comparison. To further aid comparison between the intervention and control group data was therefore imputed. For the imputation answers set as 'not applicable' were counted as missing. Missing data was seen to be random and multiple imputation using chained equations was employed. This has been shown to perform well for a variety of variable scaling types[18]. Every variable used in further statistical analysis was entered into the imputation model, as failure to do so may bias estimates towards the null[19]. The ordinal response scales for each single question were to be combined into a continuous score, and as such it was determined that imputation with predictive mean matching was appropriate. As shown by van Buuren et al the number of iterations can usually be quite low, between 5 and 20[19]. In this study the Stata standard of 10 iterations as burn-in period was used.

Statistical analysis

Results are presented on single question basis with comparison between the two groups using the Mann-Whitney U test. No correction for clustering was made as the estimated ICC was very low (shown below). Aggregation of scores was postulated in the methods paper[11], but discarded as a main outcome as the psychometric properties of the questionnaire were not suitable for such an approach. However, to further assess the relationship between the intervention and control group an overall scores were calculated. Sum of scores from the GS-PEQ part of the questionnaire was calculated to assess patient experience. Scoring was done following the scoring rules outlined above. This was validated against a question asking for an overall evaluation of the institution (Question 17 in Table 1). Overall experience with health interaction was calculated from the health interaction part of the questionnaire. As described multiple imputation was used to account for missing data.

To obtain an idea of the effect of clustering a total crude score was calculated and included the GS-PEQ part, health interaction part and Question 15. A multilevel regression model was built with total crude score from the questionnaire as the dependent variable. Intervention status was included in the model as this is the main outcome. Based on prior knowledge, gender, age, education level and self-perceived health were included in the model, as these tend to influence patient experience[20-22]. Age was centered to ease interpretation in a mixed model analysis[23]. Self-perceived health was reported on a five level Likert style scale and education level in four categories. Both were included as dummy variables in the analysis. Other confounding variables were assessed in the model and included if their inclusion led to a >10%

change in the main outcome when added to the base model (main outcome + intervention status). Relevant interactions were checked for relevant variables, where $p<0.10$ was set as the significance level. As imputation was used, Monte Carlo error estimates were employed to assess the level of simulation error, as suggested by White and Royston[24]. Normal evaluation of multilevel models with log likelihood ratio tests were not carried out, as this is not well defined for multilevel models with imputed data. However, very little of the variance was explained by the clustered design and the ICC was very low (further data available on request). A slope for group status (intervention/control) was added, but did not explain much of the variance in the data and was therefore not kept in the model. The analysis employed restricted maximum likelihood techniques throughout, as suggested when the number of clusters is small[25].

RESULTS

Response rate

Response rate was 69.4% before reminders were sent out, rising to 82.0% after reminders (Figure 1). Mean age for responders were 61 years and for non-responders 47 years (t-test <0.0001). There was no significant gender difference between the responders and non-responders, and the response rate did not differ significantly between the intervention and control group (χ^2-test).

Missing data

Missing data for most questions was low, ranging from 0 to 11 out of 410 answered questionnaires. Statistically these were considered missing completely at random (MCAR) with no clear relation to either age, gender, self-perceived health, disease severity or other variables[26]. However, Questions 6, 10 and 12 in the general part of the questionnaire had higher amounts of not applicable ranging from 14 to 34 representing 3.4% to 8.3% of returned questionnaires. In these questions the word 'treatment' was used. This was intended to cover the medical examination and interventions during the outpatient visit. However, it seems that this has been misunderstood by several patients. It seems reasonable to assume that patients who underwent 'only' diagnostic evaluation felt that they had received no 'treatment', and hence felt unable to answer the question. This was also highlighted by one patient in a free-text response in the questionnaire. "Not applicable" to Question 6, 10 and 12 did not vary significantly with age (t-test), intervention group status (χ^2-test), gender (χ^2-test) or self-reported health (χ^2-test). This was treated as missing at random for imputation purposes (MAR)[26]. Question 14 had a missing rate of 15.9% but also yielded a high level of not applicable responses, at 46.0% of returned questionnaires. This was expected, as many people will not have had a new appointment with their GP following the hospital outpatient evaluation. It is also reasonable to assume that the high amount of missing was related to the same concept. The response "not applicable" did not significantly vary with age (t-test p=0.868), intervention group status (χ^2-test p=0.064) or self-reported health (χ^2-test p=0.459). However, due to the high "not applicable" rate, this question was neither included in the final score nor the multi-level model.

A histogram of responses to questions with five categories showed all response sets to be skewed to the left. However, earlier work has indicated that multiple imputation can perform well, even when the categorical variable is non-normally distributed, as long as MAR does not exceed 10%[27]. In a 2010 article Finch argues that multiple imputation performs well for imputation of missing categorical questionnaire data[26]. There was no association between levels of missing data and the multilevel structure of the data.

Baseline characteristics

Baseline characteristics are presented in Table 3. There was no major difference between the intervention and control group with regard to gender, age, urban or rural residency or questionnaire response. The effect of the referral intervention on referral quality has previously been shown to be clinically significant with an effect of 18% (95% CI 11, 25 p<0.001)[12]. However this was for the full dataset of 500 patients. To ensure that this was also representative of the subpopulation who answered the questionnaire, the multi-level regression model was employed using data from only the 410 patients who answered it. This showed an intervention effect of 19% (95% CI 12, 27 p<0.001) on referral quality, well within the 95% CI of the full analysis.

Table 3: Selected patient baseline characteristics by intervention status

	Intervention group	Control group	p-value
Female/male[a]	140 (59.3)/96 (40.7)	102 (58.6)/72 (41.4)	0.887
Age, year[b]	60.9 ±12.5	60.3 ± 13.5	0.628
Urban/rural[a]	145 (61.4)/91 (38.6)	95 (54.6)/79 (45.4)	0.165

Clinical group[a]			
- dyspepsia	117 (49.6)	96 (55.2)	0.293
- suspected colorectal malignancy	75 (31.8)	57 (32.8)	
- chest pain	40 (17.0)	18 (10.3)	
- chronic obstructive pulmonary disease	4 (1.7)	3 (1.7)	
Hospital appointment with senior house officer/specialist[a]	107 (45.3)/129 (54.7)	78 (44.8)/96 (55.2)	0.918
Questionnaire returned promptly/after mailed reminder[a]	202 (85.6)/34 (14.4)	145 (83.3)/29 (16.7)	0.531

[a] Data are number (%)
[b] Data are mean (±SD)

Questionnaire results

Overall satisfaction with services was high and as Figure 2a-c depicts there was little difference between the intervention and control group for the individual questions. Using the Mann-Whitney U test no significant differences between the intervention and control groups were found for any of the questions. All response sets were skewed to the left, that is, towards more positive responses. However for single answer categories some interesting differences emerged. Significantly more patients in the intervention than the control group felt that they, to a very large extent, received sufficient information about their diagnosis/condition (Question 4) (χ^2-test p=0.007). There were also more patients in the intervention than the control group who felt that the help and treatment received at the institution was satisfactory to a very large extent, although this was not significant (Question 9) (χ^2-test p=0.08). Also fewer patients in the intervention

group felt that their GP lacked information after their hospital appointment (χ^2-test p=0.037), although here the actual numbers are very small (8 vs. 4 patients) (Question 14). In addition no patient in the intervention group evaluated the institution as somewhat or much poorer than expected, whereas 6 patients (3%) in the control group indicated this level of dissatisfaction (Question 17).

Interestingly the highest numbers of scores indicating dissatisfaction were for Questions 4 and 6, both for the intervention and control group patients. These questions concern patient interaction and information in the treatment process. Also approximately 5% in each group felt that they had no or little benefit from their care at the institution (Question 12). If the three lowest response categories are included, approximately 34% felt they had received no, little or some benefit.

An interesting, but not significant, difference was also seen between the intervention and control group regarding fulfillment of care expectation, where 46% of intervention patients experienced the care as better than expected, whereas for the control group this figure was 38%. None in the intervention group and 3.7% of the control group felt their care to be below expectations.

GS-PEQ

Average score, corrected for clustering, for the GS-PEQ component in the intervention group was 50.71 (95% CI 50.02, 51.41) vs. 50.01 (95% CI 48.94, 51.07) in the control group. The maximum possible score was 60. An internal validation against

the overall hospital evaluation in Question 17 using linear regression analysis, corrected for intervention status, showed a significant trend as presented in Table 4. Monte Carlo error estimates were within those recommended in the literature[24].

Table 4: Association between patient satisfaction (GS-PEQ) and overall patient hospital satisfaction (N=410)

Variable	Regression coefficient	95% CI	p-value[b]
Much poorer than expected[a]			
Somewhat poorer than expected	4.04	-6.61, 14.69	0.456
As expected	8.46	-1.25, 18.18	0.088
Somewhat better than expected	10.44	0.65, 20.23	0.037
Much better than expected	14.12	4.38, 23.86	0.005

[a] reference [b] p-value for trend <0.0001

Health interaction

Of the two questions about health interaction, one (Question 14) was not included in the imputation process due to high numbers of "not applicable" as discussed above. Average score, corrected for clustering for the remaining interaction question in the intervention group was 4.93 (95% CI 4.85, 5.01) vs. 4.86 (95% CI 4.74, 4.98) in the control group. Maximum score on this part of the questionnaire was 5 when Question 14 was left out of the analysis.

Assessment of clustering effect

The total crude questionnaire score was 60.61 (95% CI 59.90, 61.33) in the intervention group and 59.85 (95% CI 59.04, 60.66) in the control group out of a maximum of 70. In the regression model, being in the intervention group predicted a

small, non-significant increase in patient experience score, as illustrated in Table 5. No significant interaction was found and the result was not confounded by GP specialist status, GP gender, specialist status of hospital doctor or seriousness of final diagnosis. The Monte Carlo error estimates were within the limits recommended[24].

Table 5: Effect estimate for intervention on patient experience with care

Variable	Regression coefficient	95% CI	p-value
Crude	0.76	-0.32, 1.85	0.169
Adjusted	0.71	-0.33, 1.74	0.180
- Age (centered)	0.04	-0.00, 0.08	0.070
- Gender (male)	0.28	-0.77, 1.33	0.600
- Self-perceived health			
- poor[a]			<0.001[b]
- quite good	-1.71	-3.85, 0.43	
- good	0.98	-1.09, 3.05	
- very good	3.61	1.26, 5.95	
- excellent	2.67	-2.40, 7.76	
- Highest educational level			
- compulsory schooling[a]			0.545[b]
- sixth form college	-0.91	-2.26, 0.44	
- college/university 1-4 years	0.14	-1.53, 1.80	
- college/university 4 years or more	0.08	-2.03, 2.20	

[a] reference [b]overall p-value

Intraclass cluster coefficient

Initial multilevel analysis of the data revealed virtually no variance of the intercepts. The ICC was estimated at $2.19e^{-09}$. Hence very little of the variation in the data was related to the clustered design.

DISCUSSION

In the presentation in Figure 2 of the data from each question, it is quite clear that, for the most part, patients in this project report positive experiences, with no large differences between the intervention and control group. It hence seems that although the intervention has increased the referral quality significantly[12], this has not translated into a more positive patient experience with the referral process and treatment, as measured by self-report questionnaires. In the current study in depth data analysis with imputation and multilevel regression modeling was employed to further explore the intervention effect and the effect of clustering. No clear differences were found between the intervention and control group and no clear effect of clustering was found. Care must be taken in the interpretation of the in depth analysis of the imputated data. In the imputation process the answer "not applicable" was treated as a missing value, and its interpretation when imputed is difficult. However, as shown above, the number of "not applicable" was low and it was hence seen that imputation could be a useful way to ensure that no relevant differences had been missed in the single question analysis.

A strength of the current study is the fairly high response rate (82.0%) compared to other mail response studies[28]. However, the potential for non-response bias is always present. Others have previously shown the effect of this to be small[29;30]. Earlier Norwegian studies have suggested only minor differences between answers provided by responders and non-responders, when the latter have been obtained through telephone follow-up interviews[31-33]. A clear limitation is the use of short form questionnaires with single items, which may be less valid than longer forms[34]. However, shorter forms will increase the response rate[4;35]. The current project aimed to assess the effect of a health system intervention and the patient experiences with this intervention. We hence

decided to keep the questionnaire short to enable a high response rate and keep the patient and staff workload manageable. In addition it is possible that the simple fact that a research interest is taken in the patient's experience may lead to a positive perception of the services and consequently more positive ratings[36], similar to a Hawthorne effect. On the other hand it could be argued that the distribution of patient experience questionnaires should give patients an opportunity to respond by voicing negative concerns about the institution. However the issue of a 'self-interest bias' could also be raised, whereby patients respond positively to continue to receive appropriate services[36]. The current questionnaire aimed to measure patient experience with defined aspects of care to try and reduce some of these limitations, and hence reduce some of the more open ended aspects of the evaluation, as has been discussed by others[37]. There is no indication that these potential biases varied between the intervention and control group.

Overall the experience with health interaction between the hospital and the GP was quite good with few people indicating the information went missing in the transfer of care. 1.7% in the intervention group and 3.5% in the control group felt the hospital specialist lacked information from the GP. 2.1% in the intervention and 5.3% in the control group felt the GP lacked information from the specialist. In the Norwegian part of the 2010 Commonwealth Fund Survey, the same questions gave much higher negative ratings, with 12.1% indicating that the specialist lacked information from the GP and 38.3% indicating that the GP lacked information from the hospital[17]. Data from the 2013 Commonwealth Fund Survey suggest similar ratings as in 2010, although the wording of the questions is slightly different[38]. A Norwegian report concerning patient experience as inpatients also suggests higher dissatisfaction with co-operation between

the hospital and the GP[39] than in the current study. In total this clearly suggests that the patient experience of the GP/specialist communication is better in a small district hospital than the country average suggests. It is therefore possible that the effect of the intervention on patient experience could have been higher if the level of dissatisfaction with the health care cooperation had been higher in the local population. However, this also may suggest that although the hospital consultants often feel information is lacking in the referrals[40;41], this is not necessarily experienced as a problem by patients.

Earlier publications have shown that patients are generally satisfied with their care[36;42;43], even if they identify problems during their clinical encounter[44]. In light of this, others have argued that dissatisfaction may be more interesting than satisfaction in research into patients experiences [43]. In the current study, two questions were answered more negatively than others. These questions therefore probably provide the most interesting points for further quality improvement at the local facility. These two questions represent areas where communication is the main concept, namely patient involvement in the treatment process and information from doctors to patients. Others have previously shown communication and information errors as a cause for dissatisfaction[45], and in other jurisdictions even malpractice claims[46;47]. However, there was no clear difference between the intervention and the control group with regard to these questions.

In total 34% felt no, little or some benefit from their referral and contact with the hospital, although no difference was seen between the intervention and control group. This may indicate that the doctor and patient had different perceptions of the purpose of

the referral in the first place, or that the expectations from the patients were higher than the service was able to provide. It may be that the GP referred in order to rule out serious illness, whereas the patient in fact might have wanted a clear explanation of his/hers symptoms. To increase patient experience further, this area would need to be more closely examined than the current data allows.

CONCLUSION

In this project, patient satisfaction, as measured by patient experience questionnaires, was generally high, with no major differences between the intervention and control group. No clear effect of the implementation of referral templates on patient satisfaction was evident.

Interestingly, the most negative feedback, from both intervention and control group, was concerning patient interaction, involvement and information. Effecting communication and involving patients in decision making may help to increase patient satisfaction to an even higher level.

AUTHOR CONTRIBUTIONS:

The administration and daily running of the study was performed by HW, who was also the grant holder. ARB and HW developed the questionnaire. All authors participated in the analysis and interpretation of the data. All authors revised drafts of the manuscript and approved the final version.

FUNDING STATEMENT:

This work was funded by a research grant from the Northern Norway Regional Health Authority (Helse Nord RHF) with grant number HST1026-11.

COMPETING INTERESTS:

The authors declare that they have no competing interests.

DATA SHARING

No additional data available.

FIGURE LEGENDS:

Figure 1: Questionnaire response

Figure 2a-c: Graphical depiction of response to individual questions for intervention vs control group

Reference List

1. Crow R, Gage H, Hampson S et al. The measurement of satisfaction with healthcare: implications for practice from a systematic review of the literature. *Health Technol Assess* 2002; 6(32):1-244.
2. Garrat A, Solheim E, Danielsen K. National and cross-national surveys of patient experiences: a structured review. Oslo: Norwegian Knowledge Centre for the Health Services 2008.
3. Edwards PJ, Roberts I, Clarke MJ, DiGuiseppi C, Wentz R, Kwan I et al. Methods to increase response to postal and electronic questionnaires. *Cochrane Database Syst Rev* 2009;(3):MR000008.

4. Edwards P. Questionnaires in clinical trials: guidelines for optimal design and administration. *Trials* 2010; 11(1):2.

5. Fitzpatrick R, Davey C, Buxton M et al. Evaluating patient-based outcome measures for use in clinical trials: a review. *Health Technol Assess* 1998; 2(14):74.

6. Ellins J. Great expectations? Reflections on the future of patient and public involvement in the NHS. *Clin Med* 2011; 11(6):544-547.

7. Doyle C, Lennox L, Bell D. A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. *BMJ Open* 2013; 3(1).

8. McDonald K, Sundaram V, Bravata D, et al. Closing the Quality Gap: A Critical Analysis of Quality Improvement Strategies (Vol. 7: Care Coordination). Rockville, MD: Agency for Healthcare Research and Quality (US); 2007.

9. National Quality Forum. Preferred Practices and Performance Measures for Measuring and Reporting Care Coordination: A Consensus Report. Washington, DC: National Quality Forum. 2010

10. Pascoe S, Veitch C, Crossland L et al. Patients' experiences of referral for colorectal cancer. *BMC Fam Pract* 2013; 14(1):124.

11. Wahlberg H, Valle PC, Malm S et al. Practical health co-operation - the impact of a referral template on quality of care and health care co-operation: study protocol for a cluster randomized controlled trial. *Trials* 2013; 14(1):7.

12. Wahlberg H, Valle P, Malm S et al. Impact of referral templates on the quality of referrals from primary to secondary care: a cluster randomised trial. *BMC Health Serv Res* 2015; 15(1):353.

13. Carlsen B, Norheim OF. Introduction of the patient-list system in general practice. Changes in Norwegian physicians' perception of their gatekeeper role. *Scand J Prim Health Care* 2003; 21(4):209-213.

14. Schultz E, Pineda N, Lonhart J et al. A systematic review of the care coordination measurement landscape. *BMC Health Serv Res* 2013; 13(1):119.

15. Sjetne IS, Bjertnes OA, Iversen HH et al. Pasienterfaringer i spesialisthelsetjenesten. Et generisk, kort spørreskjema. Oslo: Nasjonalt kunnskapssenter for helsetjenesten 2009.

16. Sjetne I, Bjertnaes O, Olsen R et al. The Generic Short Patient Experiences Questionnaire (GS-PEQ): identification of core items from a survey in Norway. *BMC Health Serv Res* 2011; 11(1):88.

17. Skudal KE, Bjernæs ØA, Holmboe O et al. Commonwealth Fund-undersøkelsen 2010: Resultater fra en komparativ befolkningsundersøkelse i 12 land. Oslo: Nasjonalt kunnskapssenter for helsetjenesten 2010.

18. Lee KJ, Carlin JB. Multiple Imputation for Missing Data: Fully Conditional Specification Versus Multivariate Normal Imputation. *Am J Epidemiol* 2010; 171(5):624-632.

19. van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res* 2007; 16(3):219-242.

20. Xiao H, Barber JP. The effect of perceived health status on patient satisfaction. *Value Health* 2008; 11(4):719-725.

21. Fitzpatrick R. Surveys of patient satisfaction: II--Designing a questionnaire and conducting a survey. *BMJ* 1991; 302(6785):1129-1132.

22. Da Costa D, Clarke AE, Dobkin PL et al. The relationship between health status, social support and satisfaction with medical care among patients with systemic lupus erythematosus. *Int J Qual Health Care* 1999; 11(3):201-207.

23. Twisk J. Applied Multilevel Analysis. Cambridge: Cambridge University Press 2006.

24. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2011; 30(4):377-399.

25. Eldridge S, Kerry S. A Practical Guide to Cluster Randomised Trials in Health Services Research. 1 ed. Chichester: John Wiley & Sons, Ltd 2012.

26. Finch W. Imputation Methods for Missing Categorical Questionnaire Data: A Comparison of Approaches. *J Data Sci* 2010; 8:361-378.

27. Leite W, Beretvas S. The Performance of Multiple Imputation for Likert-type Items with Missing Data. *J Mod Appl Stat Methods* 2010;(9):64-74.

28. Sitzia J, Wood N. Response rate in patient satisfaction research: an analysis of 210 published studies. *Int J Qual Health Care* 1998; 10(4):311-317.

29. Lasek RJ, Barkley W, Harper DL et al. An Evaluation of the Impact of Nonresponse Bias on Patient Satisfaction Surveys. *Med Care* 1997; 35(6).

30. Perneger TV, Chamot E, Bovier PA. Nonresponse Bias in a Survey of Patient Perceptions of Hospital Care. *Med Care* 2005; 43(4).

31. Guldvog B, Hofoss D, Pettersen KI et al. PS-RESKVA -pasienttilfredshet i sykehus [PS-RESKVA -patient satisfaction in hospitals]. *Tidsskr Nor Laegeforen* 1998; 118:386-391.

32. Dahle KA, Iversen HH, Sjetne IS et al. Do non-respondents' characteristics and experiences differ from those of respondents? Poster - 25th International Conference by the Society for Quality in Health Care (ISQua) 2008.

33. Iversen, HH, Dahle KA, Homlboe O. Using telephone interviews to explore potential non-response bias in a national postal survey. Poster - 26th International Conference by the Society for Quality in Health Care (ISQua) 2009.

34. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; 30(6):473-483.

35. Eaker S, Bergstrøm R, Bergstrøm A et al. Response Rate to Mailed Epidemiologic Questionnaires: A Population-based Randomized Trial of Variations in Design and Mailing Routines. *Am J Epidemiol* 1998; 147(1):74-82.

36. Sitzia J, Wood N. Patient satisfaction: A review of issues and concepts. *Soc Sci Med* 1997; 45(12):1829-1843.

37. Pettersen KI, Veenstra M, Guldvog B et al. The Patient Experiences Questionnaire: development, validity and reliability. *Int J Qual Health Care* 2004; 16(6):453-463.

38. Haugum M, Bjertnæs ØA, Iversen HH et al. Commonwealth Fund's health policy survey in 11 countries: Norwegian results in 2013 and development since 2010. Oslo: Norwegian Knowledge Centre for the Health Services 2013.

39. Skudal KE, Holmboe O, Iversen HH et al. Inpatients' experience with Norwegian hospitals: National results in 2011 and changes from 2006. Oslo: Norwegian Knowledge Centre for the Health Services 2012.

40. Canadian Medical Association. National Survey: experiences with referrals from primary to specialty care. Canadian Medical Association 2013.

41. Bodek S, Ghori K, Edelstein M et al. Contemporary referral of patients from community care to cardiology lack diagnostic and clinical detail. *Int J Clin Pract* 2006; 60(5):595-601.

42. Staniszewska SH, Henderson L. Patients evaluations of the quality of care: influencing factors and the importance of engagement. *J Adv Nurs* 2005; 49(5):530-537.

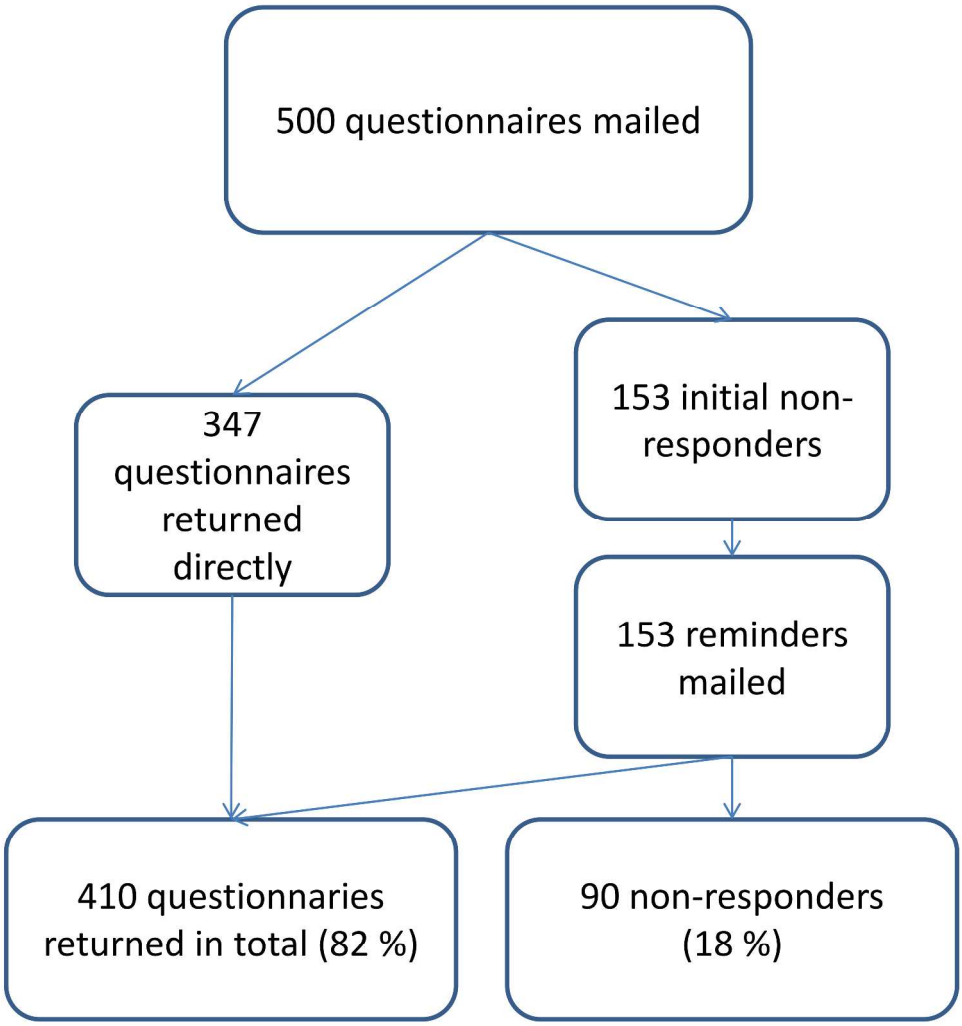
43. Williams B, Coyle J, Healy D. The meaning of patient satisfaction: An explanation of high reported levels. *Soc Sci Med* 1998; 47(9):1351-1359.

44. Jenkinson C, Coulter A, Bruster S et al. Patients' experiences and satisfaction with health care: results of a questionnaire study of specific aspects of care. *Qual Saf Health Care* 2002; 11(4):335-339.

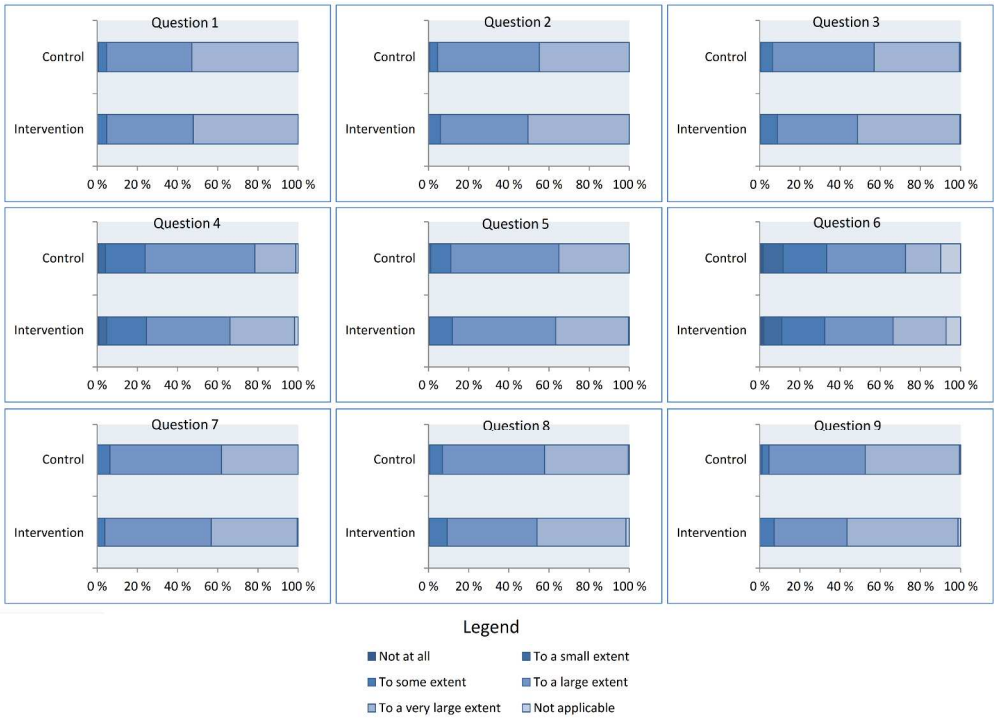
45. Hiidenhovi H, Nojonen K, Laippala P. Measurement of outpatients' views of service quality in a Finnish university hospital. *J Adv Nurs* 2002; 38(1):59-67.

46. Levinson W, Roter DL, Mullooly JP et al. Physician-patient communication: The relationship with malpractice claims among primary care physicians and surgeons. *JAMA* 1997; 277(7):553-559.

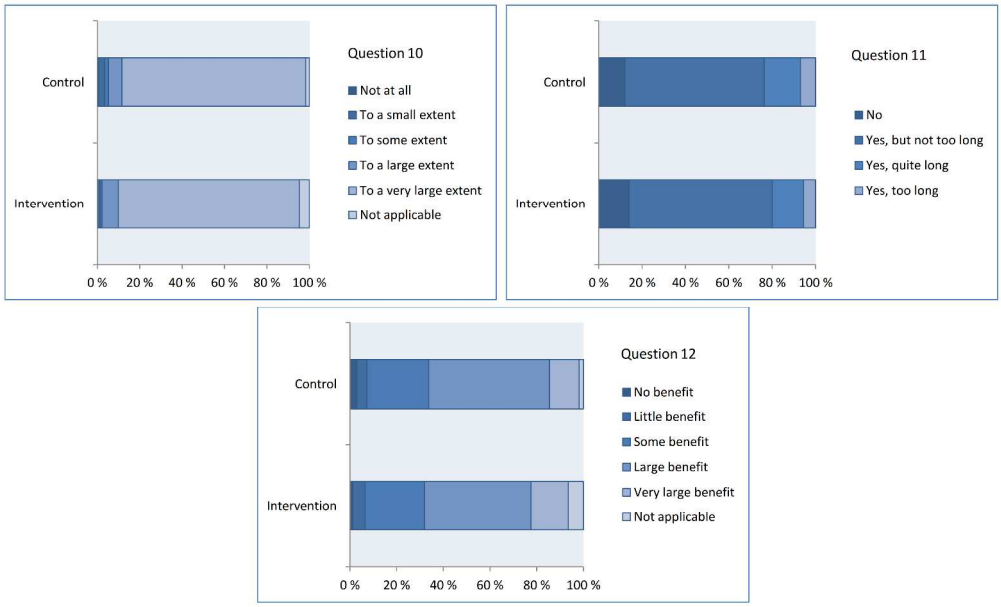
47. Beckman HB, Markakis KM, Suchman AL, Frankel RM. The doctor-patient relationship and malpractice: Lessons from plaintiff depositions. *Arch Intern Med* 1994; 154(12):1365-1370.



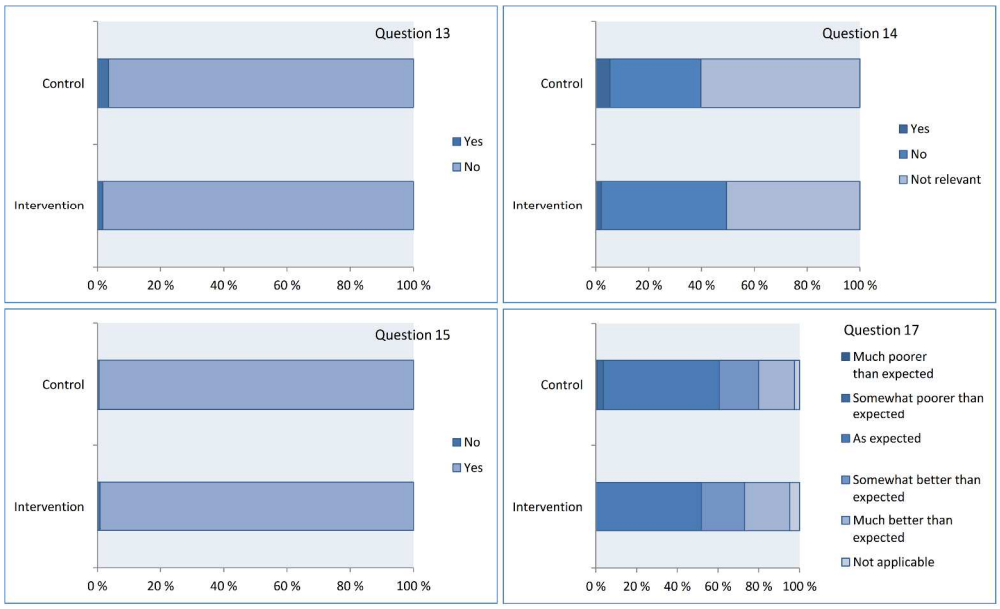
Questionnaire response
Figure 1
641x679mm (150 x 150 DPI)



Graphical depiction of response to individual questions for intervention vs control group
Figure 2a-c
1016x762mm (150 x 150 DPI)



Graphical depiction of response to individual questions for intervention vs control group
Figure 2a-c
1014x624mm (150 x 150 DPI)



Graphical depiction of response to individual questions for intervention vs control group
Figure 2a-c
1012x616mm (150 x 150 DPI)

Table 1: CONSORT 2010 checklist of information to include when reporting a cluster randomised trial

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Title and abstract				
	1a	Identification as a randomised trial in the title	Identification as a cluster randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)[1,2]	See table 2	2
Introduction				
Background and objectives	2a	Scientific background and explanation of rationale	Rationale for using a cluster design	6
	2b	Specific objectives or hypotheses	Whether objectives pertain to the the cluster level, the individual participant level or both	7-8
Methods				
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	Definition of cluster and description of how the design features apply to the clusters	6
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons		None
Participants	4a	Eligibility criteria for participants	Eligibility criteria for clusters	6
	4b	Settings and locations where the data were collected		8
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	Whether interventions pertain to the cluster level, the individual participant level or both	7
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and	Whether outcome measures pertain to the cluster level, the individual participant level or both	7

		when they were assessed		
	6b	Any changes to trial outcomes after the trial commenced, with reasons		None
Sample size	7a	How sample size was determined	Method of calculation, number of clusters(s) (and whether equal or unequal cluster sizes are assumed), cluster size, a coefficient of intracluster correlation (ICC or *k*), and an indication of its uncertainty	7
	7b	When applicable, explanation of any interim analyses and stopping guidelines		
Randomisation:				
Sequence generation	8a	Method used to generate the random allocation sequence		6
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	Details of stratification or matching if used	6
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	Specification that allocation was based on clusters rather than individuals and whether allocation concealment (if any) was at the cluster level, the individual participant level or both	7
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Replace by 10a, 10b and 10c	
	10a		Who generated the random allocation sequence, who enrolled clusters, and who assigned clusters to interventions	8
	10b		Mechanism by which individual participants were included in clusters for the purposes of the trial (such as complete	8

			enumeration, random sampling)	
	10c		From whom consent was sought (representatives of the cluster, or individual cluster members, or both), and whether consent was sought before or after randomisation	8
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how		7
	11b	If relevant, description of the similarity of interventions		Not relevant
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	How clustering was taken into account	13
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses		13/14
Results				
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	For each group, the numbers of clusters that were randomly assigned, received intended treatment, and were analysed for the primary outcome	14
	13b	For each group, losses and exclusions after randomisation, together with reasons	For each group, losses and exclusions for both clusters and individual cluster members	14
Recruitment	14a	Dates defining the periods of recruitment and follow-up		8
	14b	Why the trial ended or was stopped		Not stopped early
Baseline data	15	A table showing baseline demographic and clinical	Baseline characteristics for the individual and cluster levels as	16

		characteristics for each group	applicable for each group	
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	For each group, number of clusters included in each analysis	16
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	Results at the individual or cluster level as applicable and a coefficient of intracluster correlation (ICC or k) for each primary outcome	16-20
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended		
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory		
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms[3])		None
Discussion				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses		21-24
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	Generalisability to clusters and/or individual participants (as relevant)	
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		21-24
Other information				
Registration	23	Registration number and		2

		name of trial registry	
Protocol	24	Where the full trial protocol can be accessed, if available	Provided & methods paper referenced
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	25

** Note: page numbers optional depending on journal requirements*

Table 2: Extension of CONSORT for abstracts[1,2] to reports of cluster randomised trials

Item	Standard Checklist item	Extension for cluster trials
Title	Identification of study as randomised	**Identification of study as cluster randomised**
Trial design	Description of the trial design (e.g. parallel, cluster, non-inferiority)	
Methods		
Participants	Eligibility criteria for participants and the settings where the data were collected	**Eligibility criteria for clusters**
Interventions	Interventions intended for each group	
Objective	Specific objective or hypothesis	**Whether objective or hypothesis pertains to the cluster level, the individual participant level or both**
Outcome	Clearly defined primary outcome for this report	**Whether the primary outcome pertains to the cluster level, the individual participant level or both**
Randomization	How participants were allocated to interventions	**How clusters were allocated to interventions**
Blinding (masking)	Whether or not participants, care givers, and those assessing the outcomes were blinded to group assignment	
Results		
Numbers randomized	Number of participants randomized to each group	**Number of clusters randomized to each group**
Recruitment	Trial status[1]	
Numbers analysed	Number of participants analysed in each group	**Number of clusters analysed in each group**
Outcome	For the primary outcome, a result for each group and the estimated effect size and its precision	**Results at the cluster or individual participant level as applicable for each primary outcome**
Harms	Important adverse events or side effects	
Conclusions	General interpretation of the results	
Trial registration	Registration number and name of trial register	
Funding	Source of funding	

[1] Relevant to Conference Abstracts

REFERENCES

1 Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet* 2008, 371:281-283

2 Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG at al (2008) CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. *PLoS Med* 5(1): e20

3 Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, Moher D. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004; 141(10):781-788.

BMJ Open

The impact of referral templates on patient experience of the referral and care process. A cluster randomized trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-011651.R1
Article Type:	Research
Date Submitted by the Author:	13-Jun-2016
Complete List of Authors:	Wåhlberg, Henrik; Universitetet i Tromso Institutt for Samfunnsmedisin ISM, ; Sykehuset Østfold Kalnes, Medical department Braaten, Tonje; University of Tromsø, Department of Community Medicine Broderstad, Ann Ragnhild; Uit the Arctic University of Norway, Faculty of Health; University Hospital of North Norway, medical department
Primary Subject Heading:	Health services research
Secondary Subject Heading:	General practice / Family practice, Patient-centred medicine
Keywords:	Patient experience, Referral, General practitioner

Title: The impact of referral templates on patient experience of the referral and care process. A cluster randomized trial

Henrik Wåhlberg – corresponding author

Department of Community Medicine, UiT The Arctic University of Norway

Faculty of Health Sciences, UiT The Arctic University of Norway

9037 Tromsø, Norway

University Hospital of North Norway Harstad

St. Olavsgate 70, 9480 Harstad, Norway

henrik.wahlberg@uit.no/henrik_waahlberg@hotmail.com

Tonje Braaten

Department of Community Medicine, UiT The Arctic University of Norway

Faculty of Health Sciences, UiT The Arctic University of Norway

9037 Tromsø, Norway

tonje.braaten@uit.no

Ann Ragnhild Broderstad

Centre for Sami Health Research, UiT The Arctic University of Norway

UiT The Arctic University of Norway, 9037 Tromsø, Norway

University Hospital of North Norway Harstad

St. Olavsgate 70, 9480 Harstad, Norway

ann.ragnhild.broderstad@uit.no

Keywords: Patient experience. Referral. General practitioner.

Wordcount: 3370 words (excluding abstract, references, tables and author contributions)

ABSTRACT

Objectives: To evaluate if a referral intervention improves the patient experience of the referral and treatment process.

Setting: Interface between fourteen primary care surgeries and a district general hospital.

Participants: The fourteen GP surgeries (seven intervention, seven control) in the area around the University Hospital of North Norway Harstad were randomised and all completed the study. Consecutive individual patients were recruited at their hospital appointment. A total of 500 patients were recruited with 281 in the intervention and 219 in the control arm.

Interventions: Dissemination of referral templates for four diagnostic groups (dyspepsia, suspected colorectal cancer, chest pain and COPD) coupled with intermittent surgery visits by study personnel. Control arm continued standard referral practice. The intervention was in use for 2.5 years.

Outcome: The main outcome was a quality indicator score. This paper reports a secondary outcome, the patient experience, as measured by self-report questionnaires. General practitioners in the intervention group could not be blinded. Patients were blinded to intervention status. Analysis was based on single question comparison with a questionnaire subscore used to assess the effect of clustering.

Results: On the individual questions overall satisfaction was very high with minor differences between the intervention and control group. Interestingly the most negative responses, in both groups, were concerned questions relating to patient interaction and information. Very little evidence of clustering was found with an estimated intracluster correlations coefficient at $1.21e^{-11}$.

Conclusion: In total this indicates no clear effect of the implementation of referral templates on the patient experience, in a setting of generally high patient satisfaction.

Trial registration: This trial has been registered at ClinicalTrials.gov. The trial registration number is NCT01470963.

ARTICLE SUMMARY

Strengths and limitations of the study

- Clinically relevant research in a regular district hospital setting
- High response rate
- Newly developed questionnaire hampers wider generalisation

INTRODUCTION

Evaluation of patient experience and satisfaction is widespread with a wealth of literature concerning the development and use of questionnaires[1-5]. The evaluation of patient experience can help drive quality improvement[6] and improved patient experience is associated with safety and clinical effectiveness[7].

Care coordination is an important aspect of a well-functioning high quality health service. It has been defined as "the deliberate integration of patient care activities between two or more participants involved in a patient's care to facilitate the appropriate delivery of health care services."[8] In the US the National Quality Forum (NQF) has published preferred practices for care coordination, including transitions of care[9]. This report includes clear recommendations for participation of the patient, or his/her designee, in the decision, planning and execution of a care transition. This is important, as exemplified by a recent Australian article, where patients with colorectal cancer perceived that poor information exchange led to suboptimal care[10]. Hence assessing patient experience of the referral process may be beneficial in assessing the effect of a referral intervention.

This article presents the patient experience aspect of a cluster randomized study evaluating the effect of the implementation of referral templates for four diagnostic groups – dyspepsia, suspected colorectal cancer, chest pain and COPD - in the patient referral pathway[11]. Previously we have shown that the referral templates led to increased referral quality[12], and the effect on the main outcome, quality of care at the

hospital, is in publication. This publication aims to assess whether the implementation of a referral template in the transition of care from the general practitioner to the hospital has affected the patient experience of the care process.

METHODS

Study setting

In Norway the health care system is quite uniformly organised throughout the country. GPs act as gatekeepers to secondary care[13] with specialist health services delivered by governmentally owned regional health authorities, mainly through public hospitals. Some specialist outpatient care is delivered by private specialists, but this is mainly purchased by the regional health authorities. The access to private specialists in the geographical area of the current study is very limited.

Study design

The study was designed as a cluster randomized study with the general practitioner (GP) surgery as the clustering unit. A total of 14 surgeries were randomized, seven to the intervention and seven to the control group. The clustered design was chosen to avoid possible spill-over effect from the intervention to control GPs. Randomization was done by simple drawing by a person not connected to the research team, stratified by town vs countryside location of surgery.

As the intervention was to be actively used by the GPs the referring GP could not be blinded. Patients, hospital doctors and outcome evaluators were blinded to the intervention status of the patients. Due to the design of the intervention the referral letter would sometimes reveal the intervention status, if the electronic template was used. No separate sample size calculation was performed for the patient experience outcome. The full study details are published in the methods paper[11].

Intervention

The intervention consisted of the distribution of four separate referral templates to the intervention surgeries. These templates covered four clinical areas (dyspepsia, suspected colorectal malignancy, chest pain and chronic obstructive pulmonary disease). The templates were to be use when initiating a new referral to the medical outpatient clinic at the University Hospital of North Norway, Harstad (UNN Harstad). The templates were distributed by the corresponding author (HW) during educational and/or lunch meetings and were provided as laminated reference sheets or in electronic form. In addition follow up visits were conducted regularly during the study period and intermittent mail leaflets and reminders were distributed to the intervention offices. Control offices continued standard referral practice.

Outcomes

The main outcome in the project was the quality of care delivered to each individual patient. In addition health process indicators such as correct prioritization

were recorded and referral quality was also compared between the intervention and control group. The current paper presents the patient experience aspect of the study, as measured by self-report questionnaires.

Participants

The fourteen GP surgeries primarily served by UNN Harstad were included in the randomization process. In 2013 these surgeries had a total list size of 39,523 patients. The individual patients were recruited from consecutive new patients within one of the four clinical areas referred to the medical outpatient clinics at UNN Harstad. Study information and a consent form were sent to each individual patient together with their appointment letter. Further information, including a new consent form if appropriate, was provided at their hospital appointment. The individual patients were analyzed as part of the intervention or control group depending on the intervention status of the GP surgery they were referred from. Children (<18 years of age) and patients with reduced capacity to consent were excluded from the study.

Recruitment

The study recruited patients for approximately 2.5 years and a total of 538 patients were included with 290 in the intervention arm and 227 in the control arm. The remaining 21 patients were referred from GP surgeries outside the regular area of UNN Harstad, and as such neither in the intervention nor the control group. These 21 were

excluded, together with 17 patients that did not fill the inclusion criteria. In total this left 281 patients in the intervention arm and 219 patients in the control arm (Figure 1).

Questionnaire development

Multiple tools exist for measuring different aspects of care coordination[14] and patient experience, however no complete questionnaire was located that covered the area in the current study completely. Therefore a questionnaire was developed by combining validated questionnaires regarding patient experiences and care coordination. The questions used were the full version of the Generic Short Patient Experiences Questionnaire (GS-PEQ)[15], together with 2 further questions used in patient experience questionnaires in the Norwegian health care system (Question 11 and 12)[16] and the two questions about health interaction from the Commonwealth Fund Survey 2010[17]. Three further questions were added to assess (1) who referred the patient, (2) if the referral was seen as appropriate and (3) an overall evaluation of the institution. Table 1 presents the questions in the questionnaire. GS-PEQ and Question 11-12 use Likert style response categories. The health interaction questions had a yes/no response. The full questionnaire, including the demographic questions, is available upon request.

Table 1 – Questionnaire details

Question No	Wording of question
1	Did the clinicians[a] talk to you in a way that was easy to understand?
2	Do you have confidence in the clinicians' professional skills?
3	Did you get sufficient information about how examinations and tests were to be performed?
4	Did you get sufficient information about your diagnosis/conditions?

5	Did you perceive the treatment to be adapted to you situation?
6	Were you involved in decisions regarding your treatment?
7	Did you perceive the institution work practices to be well organized?
8	Did you perceive the equipment at the institution to be in good working order?
9	Overall, was the help and treatment you received at the institution satisfactory?
10	Do you believe that you were in any way given incorrect treatment (according to your own judgment)?
11	Did you have to wait before you were given an appointment at the institution?
12	Overall, what benefit have you had from the care at the institution?
13	Did the hospital specialist miss basic medical information from your GP about the reason for your visit or test results?
14	After your saw the hospital specialist did your GP miss importation information about the care you got from the specialist?
15	Was the referral to the outpatient department necessary (according to your own judgment)?
16a	Were you referred by your GP for the outpatient appointment?
16b	If no in question 16a; who referred you?
17	If you take an overview of your entire treatment process, how would you evaluate the institution?

[a] With 'clinicians' we mean those who had the main treatment responsibility. This is linguistically clearer in the Norwegian wording.

The questionnaire was piloted for content validity with four local health professionals; these felt that it covered the important aspects of patient experience and care coordination. It was then piloted with five outpatients with a median age of 72 years (average 68.8 years) to ensure face validity and acceptability. Two patients needed clarification on one of the questions before they felt they could answer, and the wording of this question was adjusted accordingly. The patients felt the questionnaire was acceptable, with logical response categories and that the questions covered their clinical path during the referral process well. These patients did not take part when the project was later initiated. No further formal evaluation of the questionnaire was carried out.

Questionnaire distribution

The questionnaire was mailed to patients who had consented to take part in the referral project presented above. To increase response rates a pre-paid response envelope was included, addresses were handwritten, the questionnaire was kept short and association with research bodies was indicated[3]. For non-responders one reminder was sent approximately one month after the first questionnaire, with a new questionnaire and pre-paid response envelope.

Ethics

The study followed the principles of the Helsinki Declaration. Before recruitment started it was presented to the Regional Committee for Medical and Health Research Ethics Northern Norway,, who determined it not to be within the scope of the Health Research Act (REK NORD 2010/2259). The project has been approved by the Data Protection Officer for Research. The study is registered at www.clinicaltrials.gov with trial registration number NCT01470963. All patients provided written informed consent.

Imputation

Missing data hampered the analysis beyond single item comparison. To further aid the assessment of clustering missing data was therefore imputed. For the imputation answers set as 'not applicable' were counted as missing. Missing data was seen to be random and multiple imputation using chained equations was employed. This has been shown to perform well for a variety of variable scaling types[18]. Every variable used in

further statistical analysis was entered into the imputation model, as failure to do so may bias estimates towards the null[19]. The ordinal response scales for each single question were to be combined into a continuous score, and as such it was determined that imputation with predictive mean matching was appropriate. As shown by van Buuren et al the number of iterations can usually be quite low, between 5 and 20[19]. In this study the Stata standard of 10 iterations as burn-in period was used.

Statistical analysis

Results are presented on single question basis with comparison between the two groups using the Mann-Whitney U test for ordinal data and Chi square test for nominal data. No correction for clustering was made as the estimated ICC was very low (shown below). Aggregation of scores was postulated in the methods paper[11], but discarded as a main outcome as properties of the questionnaire, with a "not applicable" answering category are not easily suitable for such an approach. However, to assess the effect of clustering a sum of scores from the GS-PEQ part of the questionnaire was calculated and a multilevel regression model was built with the GS-PEQ score from the questionnaire as the dependent variable. Intervention status was included in the model as this is the main point of interest. Gender, age, education level and self-perceived health were included in the model, as these tend to influence patient experience[20-22]. Age was centered to ease interpretation in a mixed model analysis[23]. Self-perceived health was reported on a five level Likert style scale and education level in four categories. Both were included as dummy variables in the analysis. Other confounding variables were assessed in the model and included if their inclusion led to a >10%

change in the main outcome when added to the base model (main outcome + intervention status). Relevant interactions were checked for relevant variables, where $p<0.10$ was set as the significance level. As imputation was used, Monte Carlo error estimates were employed to assess the level of simulation error, as suggested by White and Royston[24]. Normal evaluation of multilevel models with log likelihood ratio tests were not carried out, as this is not well defined for multilevel models with imputed data. The analysis employed restricted maximum likelihood techniques throughout, as suggested when the number of clusters is small[25]. As described multiple imputation was used to account for missing data in all analysis beyond single question analysis.

RESULTS

Response rate

Response rate was 69.4% before reminders were sent out, rising to 82.0% after reminders (Figure 1). Mean age for responders were 61 years and for non-responders 47 years (t-test <0.0001). There was no significant gender difference between the responders and non-responders, and the response rate did not differ significantly between the intervention and control group (χ^2-test).

Missing data

Missing data for most questions was low, ranging from 0 to 11 out of 410 answered questionnaires. Statistically these were considered missing completely at

random (MCAR) with no clear relation to either age, gender, self-perceived health, disease severity or other variables[26]. However, Questions 6, 10 and 12 in the general part of the questionnaire had higher amounts of not applicable ranging from 14 to 34 representing 3.4% to 8.3% of returned questionnaires. In these questions the word 'treatment' was used. This was intended to cover the medical examination and interventions during the outpatient visit. However, it seems that this has been misunderstood by several patients. It seems reasonable to assume that patients who underwent 'only' diagnostic evaluation felt that they had received no 'treatment', and hence felt unable to answer the question. This was also highlighted by one patient in a free-text response in the questionnaire. "Not applicable" to Question 6, 10 and 12 did not vary significantly with age (t-test), intervention group status (χ^2-test), gender (χ^2-test) or self-reported health (χ^2-test). This was treated as missing at random for imputation purposes (MAR)[26]. Question 14 had a missing rate of 15.9% but also yielded a high level of not applicable responses, at 46.0% of returned questionnaires. This was expected, as many people will not have had a new appointment with their GP following the hospital outpatient evaluation. It is also reasonable to assume that the high amount of missing was related to the same concept. The response "not applicable" did not significantly vary with age (t-test p=0.868), intervention group status (χ^2-test p=0.064) or self-reported health (χ^2-test p=0.459).

A histogram of responses to questions with five categories showed all response sets to be skewed to the left. However, earlier work has indicated that multiple imputation can perform well, even when the categorical variable is non-normally distributed, as long as MAR does not exceed 10%[27]. In a 2010 article Finch argues that multiple imputation performs well for imputation of missing categorical questionnaire

data[26]. There was no association between levels of missing data and the multilevel structure of the data.

Baseline characteristics

Baseline characteristics are presented in Table 2. There was no major difference between the intervention and control group with regard to gender, age, urban or rural residency or questionnaire response. The effect of the referral intervention on referral quality has previously been shown to be clinically significant with an effect of 18% (95% CI 11, 25 p<0.001)[12]. However this was for the full dataset of 500 patients. To ensure that this was also representative of the subpopulation who answered the questionnaire, the multi-level regression model was employed using data from only the 410 patients who answered it. This showed an intervention effect of 19% (95% CI 12, 27 p<0.001) on referral quality, well within the 95% CI of the full analysis.

Table 2: Selected patient baseline characteristics by intervention status

	Intervention group	Control group	p-value
Female/male, N(%)	140 (59.3)/96 (40.7)	102 (58.6)/72 (41.4)	0.89
Age (year), mean (±SD)	60.9 ±12.5	60.3 ± 13.5	0.63
Urban/rural, N(%)	145 (61.4)/91 (38.6)	95 (54.6)/79 (45.4)	0.17
Clinical group, N(%)			
- dyspepsia	117 (49.6)	96 (55.2)	0.29
- suspected colorectal malignancy	75 (31.8)	57 (32.8)	
- chest pain	40 (17.0)	18 (10.3)	
- chronic obstructive	4 (1.7)	3 (1.7)	

pulmonary disease			
Hospital appointment with senior house officer/ specialist, N(%)	107 (45.3)/129 (54.7)	78 (44.8)/96 (55.2)	0.92
Questionnaire returned promptly/after mailed reminder, N(%)	202 (85.6)/34 (14.4)	145 (83.3)/29 (16.7)	0.53

Questionnaire results

Overall satisfaction with services was high and as presented in Table 3 there was little difference between the intervention and control group for the individual questions. Using the Mann-Whitney U test and χ^2-test only two questions had significant p-values (Q14 and Q17), however in both these questions the absolute difference in numbers was very small.. All response sets were skewed to the left, that is, towards more positive responses.

Interestingly the highest numbers of scores indicating dissatisfaction were for Questions 4 and 6, both for the intervention and control group patients. These questions concern patient interaction and information in the treatment process.

Table 3: Questionnaire results

Question	Answering categories[a]	Intervention	Control	p-value
Question 1 [b]		5 (4, 5)	5 (4, 5)	0.92
Question 2 [b]		5 (4, 5)	4 (4, 5)	0.39
Question 3 [b]		5 (4, 5)	4 (4, 5)	0.23
Question 4 [b]		4 (3, 5)	4 (4, 4)	0.12
Question 5 [b]		4 (4, 5)	4 (4, 5)	0.88
Question 6 [b]		4 (3, 5)	4 (3, 4)	0.19

Question 7 [b]		4 (4, 5)	4 (4, 5)	0.22
Question 8 [b]		4 (4, 5)	4 (4, 5)	0.81
Question 9 [b]		5 (4, 5)	4 (4, 5)	0.15
Question10 [b]		5 (5, 5)	5 (5, 5)	0.60
Question 11 [c]	No	33 (14.0)	21 (12.1)	0.33
	Yes, but not too long	155 (66.0)	111 (64.2)	
	Yes, quite long	34 (14.5)	29 (16.8)	
	Yes, too long	13 (5.5)	12 (6.9)	
Question 12 [c]	No benefit	3 (1.4)	5 (3.1)	0.56
	Little benefit	12 (5.5)	7 (4.3)	
	Some benefit	59 (27.2)	44 (27.0)	
	Large benefit	106 (48.9)	86 (52.8)	
	Very large benefit	37 (17.1)	21 (12.9)	
Question 13 [c]	Yes	4 (1.7)	6 (3.5)	0.25
	No	229 (98.3)	165 (96.5)	
Question 14 [c]	Yes	4 (4.2)	8 (13.1)	0.04
	No	92 (95.8)	53 (86.9)	
Question 15 [c]	Yes	232 (99.2)	170 (99.4)	0.75
	No	2 (0.8)	1 (0.6)	
Question 17 [c]	Much poorer than expected	0 (0)	1 (0.6)	0.05
	Somewhat poorer than expected	0 (0)	5 (3.1)	
	As expected	119 (54.1)	94 (58.4)	
	Somewhat better than expected	50 (22.7)	32 (19.9)	
	Much better than expected	51 (23.2)	29 (18.0)	

[a] for question 1-10 the following scoring was used: 1 = Not at all, 2 = To a small extent, 3 = To some extent, 4 = To a large extent and 5 = To a very large extent
[b] data presented as median(25th percentile, 75th percentile)
[c] data presented as number(%)

The Cronbach alpha for Question 1-15 was 0.83 and for Question 1-10 0.88.

Assessment of clustering effect

In the regression model, no significant difference was seen in the GSPEQ score between the intervention and control group with crude regression coefficient 0.55 (95% CI -0.37, 1.47 p=0.24) and adjusted 0.57 (95% CI -0.31, 1.46 p=0.20). No significant interaction was found and the result was not confounded by GP specialist status, GP gender, specialist status of hospital doctor or seriousness of final diagnosis. The Monte Carlo error estimates were within the limits recommended[24]. Initial multilevel analysis of the data revealed virtually no variance of the intercepts. The ICC was estimated at $1.21e^{-11}$. Hence very little of the variation in the data was related to the clustered design.

DISCUSSION

In the presentation of the data from each question in Table 3 it is quite clear that, for the most part, patients in this project report positive experiences, with no differences between the intervention and control group. It hence seems that although the intervention has increased the referral quality significantly[12], this has not translated into a more positive patient experience with the referral process and treatment, as measured by self-report questionnaires. In the current study in depth data analysis with imputation and multilevel regression modeling was employed to further explore the effect of clustering. No clear effect of clustering was found.

A strength of the current study is the fairly high response rate (82.0%) compared to other mail response studies[28]. However, the potential for non-response bias is always

present. Others have previously shown the effect of this to be small[29;30]. Earlier Norwegian studies have suggested only minor differences between answers provided by responders and non-responders, when the latter have been obtained through telephone follow-up interviews[31-33]. A clear limitation is the use of short form questionnaires with single items, which may be less valid than longer forms[34]. However, shorter forms will increase the response rate[4;35]. The current project aimed to assess the effect of a health system intervention and the patient experiences with care after this intervention. We hence decided to keep the questionnaire short to enable a high response rate and keep the patient and staff workload manageable.

The current project used a newly developed questionnaire to assess patient experience by combining previously validated questions. The general nature of the final questionnaire may be seen as a weakness, as small changes in the patient experience induced by the intervention may have been missed. Further piloting might have revealed more clearly if the questionnaire did indeed assess the patient experience with the referral and care process in an adequate way. However in this clinically oriented project the authors hoped that a more general questionnaire would highlight whether the intervention would cause a more overall positive, or even a negative, change. It is probable that for each individual patient it is the experience with the entire process that matters, as opposed to the experience of a subpart of the process. If large scale implementation of referral guidance is contemplated a more specific questionnaire may need to be validated.

An additional weakness was the lack of a sound analytical plan proposed in the methods paper[11]. To ensure transparency the analysis presented in this paper is therefore simple and based on single question assessment. Given the clustered nature of the study an assessment of clustering is given for a subsection of the questionnaire, but very little effect was seen.

Comparison with other studies was difficult as no clearly comparable analysis was found, except for the two health interaction questions. In the current study 1.7% in the intervention group and 3.5% in the control group felt the hospital specialist lacked information from the GP. 4.2% in the intervention and 13.1% in the control group felt the GP lacked information from the specialist. In the Norwegian part of the 2010 Commonwealth Fund Survey, the same questions gave much higher negative ratings, with 12.1% indicating that the specialist lacked information from the GP and 38.3% indicating that the GP lacked information from the hospital[17]. Data from the 2013 Commonwealth Fund Survey suggest similar ratings as in 2010, although the wording of the questions is slightly different[36]. A Norwegian report concerning patient experience as inpatients also suggests higher dissatisfaction with co-operation between the hospital and the GP[37] than in the current study. In total this clearly suggests that the patient experience of the GP/specialist communication is better in a small district hospital than the country average suggests. It is therefore possible that the effect of the intervention on patient experience could have been higher if the level of dissatisfaction with the health care cooperation had been higher in the local population. However, this also may suggest that although the hospital consultants often feel information is lacking in the referrals[38;39], this is not necessarily experienced as a problem by patients.

In the current study, two questions were answered more negatively than others. These questions therefore probably provide the most interesting points for further quality improvement at the local facility. These two questions represent areas where communication is the main concept, namely patient involvement in the treatment process and information from doctors to patients. Others have previously shown communication and information errors as a cause for dissatisfaction[40], and in other jurisdictions even malpractice claims[41;42].

CONCLUSION

In this project, patient satisfaction, as measured by patient experience questionnaires, was generally high, with no major differences between the intervention and control group. No clear effect of the implementation of referral templates on patient satisfaction was evident.

Interestingly, the most negative feedback, from both intervention and control group, was concerning patient interaction, involvement and information. Effecting communication and involving patients in decision making may help to increase patient satisfaction to an even higher level.

AUTHOR CONTRIBUTIONS:

The administration and daily running of the study was performed by HW, who was also the grant holder. ARB and HW developed the questionnaire. All authors participated in the analysis and interpretation of the data. All authors revised drafts of the manuscript and approved the final version.

FUNDING STATEMENT:

This work was funded by a research grant from the Northern Norway Regional Health Authority (Helse Nord RHF) with grant number HST1026-11.

COMPETING INTERESTS:

The authors declare that they have no competing interests.

DATA SHARING:

No additional data available. The dataset supporting the conclusions in this article may be available on request to the main author (HW).

FIGURE LEGENDS:

Figure 1: Patient inclusion and questionnaire response

Reference List

(1) Crow R, Gage H, Hampson S, Hart J, Kimber A, Storey L et al. The measurement of satisfaction with healthcare: implications for practice from a systematic review of the literature. *Health Technol Assess* 2002; 6(32):1-244.

(2) Garrat A, Solheim E, Danielsen K. National and cross-national surveys of patient experiences: a structured review. 7-2008. 2008. Oslo, Norwegian Knowledge Centre for the Health Services.

(3) Edwards PJ, Roberts I, Clarke MJ, DiGuiseppi C, Wentz R, Kwan I et al. Methods to increase response to postal and electronic questionnaires. *Cochrane Database Syst Rev* 2009;(3):MR000008.

(4) Edwards P. Questionnaires in clinical trials: guidelines for optimal design and administration. *Trials* 2010; 11(1):2.

(5) Fitzpatrick R, Davey C, Buxton M, Jones D. Evaluating patient-based outcome measures for use in clinical trials: a review. *Health Technology Assessment* 1998; 2(14):74.

(6) Ellins J. Great expectations? Reflections on the future of patient and public involvement in the NHS. *Clinical Medicine* 2011; 11(6):544-547.

(7) Doyle C, Lennox L, Bell D. A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. *BMJ Open* 2013; 3(1).

(8) McDonald K, Sundaram V, Bravata D, et al. Closing the Quality Gap: A Critical Analysis of Quality Improvement Strategies (Vol. 7: Care Coordination). Rockville (MD): Agency for Healthcare Research and Quality (US); 2007.

(9) National Quality Forum. Preferred Practices and Performance Measures for Measuring and Reporting Care Coordination: A Consensus Report. National Quality Forum, editor. 2010. Washington, DC, National Quality Forum.

(10) Pascoe S, Veitch C, Crossland L, Beilby J, Spigelman A, Stubbs J et al. Patients' experiences of referral for colorectal cancer. *BMC Family Practice* 2013; 14(1):124.

(11) Wahlberg H, Valle PC, Malm S, Broderstad AR. Practical health co-operation - the impact of a referral template on quality of care and health care co-operation: study protocol for a cluster randomized controlled trial. *Trials* 2013; 14(1):7.

(12) Wahlberg H, Valle P, Malm S, Broderstad A. Impact of referral templates on the quality of referrals from primary to secondary care: a cluster randomised trial. *BMC Health Services Research* 2015; 15(1):353.

(13) Carlsen B, Norheim OF. Introduction of the patient-list system in general practice. Changes in Norwegian physicians' perception of their gatekeeper role. *Scand J Prim Health Care* 2003; 21(4):209-213.

(14) Schultz E, Pineda N, Lonhart J, Davies S, McDonald K. A systematic review of the care coordination measurement landscape. *BMC Health Services Research* 2013; 13(1):119.

(15) Sjetne I, Bjertnaes O, Olsen R, Iversen H, Bukholm G. The Generic Short Patient Experiences Questionnaire (GS-PEQ): identification of core items from a survey in Norway. *BMC Health Services Research* 2011; 11(1):88.

(16) Sjetne IS, Bjertnes OA, Iversen HH, Olsen RV. Pasienterfaringer i spesialisthelsetjenesten. Et generisk, kort spørreskjema. 2009. Oslo, Nasjonalt kunnskapssenter for helsetjenesten.

(17) Skudal KE, Bjertnes OA, Holmboe O, Bukholm G, Rottingen JA. The 2010 Commonwealth Fund survery: Results from a comparative population survery in 11 countries. 21. 2010. Oslo, The Norwegian Knowledge Centre for the Health Services.

(18) Lee KJ, Carlin JB. Multiple Imputation for Missing Data: Fully Conditional Specification Versus Multivariate Normal Imputation. *American Journal of Epidemiology* 2010; 171(5):624-632.

(19) van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research* 2007; 16(3):219-242.

(20) Xiao H, Barber JP. The effect of perceived health status on patient satisfaction. *Value Health* 2008; 11(4):719-725.

(21) Fitzpatrick R. Surveys of patient satisfaction: II--Designing a questionnaire and conducting a survey. *BMJ* 1991; 302(6785):1129-1132.

(22) Da Costa D, Clarke AE, Dobkin PL, Senecal JL, Fortin PR, Danoff DS et al. The relationship between health status, social support and satisfaction with medical care among patients with systemic lupus erythematosus. *Int J Qual Health Care* 1999; 11(3):201-207.

(23) Twisk J. Applied Multilevel Analysis. Cambridge: Cambridge University Press; 2006.

(24) White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Statist Med* 2011; 30(4):377-399.

(25) Eldridge S, Kerry S. A Practical Guide to Cluster Randomised Trials in Health Services Research. 1 ed. Chichester: John Wiley & Sons, Ltd; 2012.

(26) Finch W. Imputation Methods for Missing Categorical Questionnaire Data: A Comparison of Approaches. *Journal of Data Science* 2010; 8:361-378.

(27) Leite W, Beretvas S. The Performance of Multiple Imputation for Likert-type Items with Missing Data. *Journal of Modern Applied Statistical Methods* 2010;(9):64-74.

(28) Sitzia J, Wood N. Response rate in patient satisfaction research: an analysis of 210 published studies. *Int J Qual Health Care* 1998; 10(4):311-317.

(29) Lasek RJ, Barkley W, Harper DL, Rosenthal GE. An Evaluation of the Impact of Nonresponse Bias on Patient Satisfaction Surveys. *Medical Care* 1997; 35(6).

(30) Perneger TV, Chamot E, Bovier PA. Nonresponse Bias in a Survey of Patient Perceptions of Hospital Care. *Medical Care* 2005; 43(4).

(31) Guldvog B, Hofoss D, Pettersen KI, Ebbesen J, Ronning OM. PS-RESKVA - pasienttilfredshet i sykehus [PS-RESKVA -patient satisfaction in hospitals]. *Tidsskr Nor Laegeforen* 1998; 118:386-391.

(32) Do non-respondents' characteristics and experiences differ from those of respondents? Poster at the the 25th International Conference by the Society for Quality in Health Care (ISQua). 25th International Conference by the Society for Quality in Healthcare (ISQua); 08; 2008.

(33) Using telephone interviews to explore potential non-response bias in a national postal survey. Poster at the the 26th International Conference by the Society for Quality in Health Care (ISQua). 26th International Conference by the Society for Quality in Health Care (ISQua); 09; 2009.

(34) Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; 30(6):473-483.

(35) Eaker S, Bergstrøm R, Bergstrøm A, Adami HO, Nyren O. Response Rate to Mailed Epidemiologic Questionnaires: A Population-based Randomized Trial of Variations in Design and Mailing Routines. *American Journal of Epidemiology* 1998; 147(1):74-82.

(36) Haugum M, Bjertnes OA, Iversen HH, Lindahl AK, Nylenna M. Commonwealth Fund's health policy survey in 11 countries: Norwegian results in 2013 and development since 2010. 27-11-2014. Oslo, The Norwegian Knowledge Centre for the Health Services.

(37) Skudal KE, Holmboe O, Iversen HH, Sjetne IS, Haugum M, Bjertnaes O. Inpatients' experience with Norwegian hospitals: National results in 2011 and changes from 2006. 3-2012. 2012. Oslo, Norwegian Knowledge Centre for the Health Services.

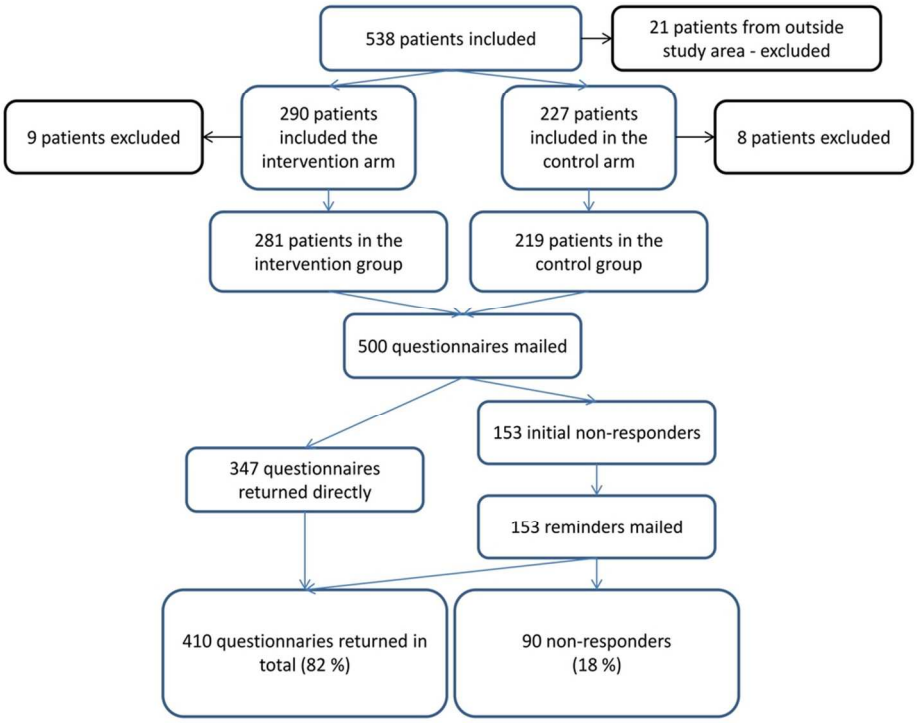
(38) Canadian Medical Association. National Survey: experiences with referrals from primary to specialty care. 2013. Canadian Medical Association.

(39) Bodek S, Ghori K, Edelstein M, Reed A, MacFadyen RJ. Contemporary referral of patients from community care to cardiology lack diagnostic and clinical detail. *Int J Clin Pract* 2006; 60(5):595-601.

(40) Hiidenhovi H, Nojonen K, Laippala P. Measurement of outpatients' views of service quality in a Finnish university hospital. *Journal of Advanced Nursing* 2002; 38(1):59-67.

(41) Levinson W, Roter DL, Mullooly JP, Dull VT, Frankel RM. Physician-patient communication: The relationship with malpractice claims among primary care physicians and surgeons. *JAMA* 1997; 277(7):553-559.

(42) Beckman HB, Markakis KM, Suchman AL, Frankel RM. The doctor-patient relationship and malpractice: Lessons from plaintiff depositions. *Archives of Internal Medicine* 1994; 154(12):1365-1370.



Patient inclusion and questionnaire response

95x71mm (300 x 300 DPI)

Table 1: CONSORT 2010 checklist of information to include when reporting a cluster randomised trial

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Title and abstract				
	1a	Identification as a randomised trial in the title	Identification as a cluster randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)[1,2]	See table 2	2
Introduction				
Background and objectives	2a	Scientific background and explanation of rationale	Rationale for using a cluster design	6
	2b	Specific objectives or hypotheses	Whether objectives pertain to the the cluster level, the individual participant level or both	7-8
Methods				
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	Definition of cluster and description of how the design features apply to the clusters	6
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons		None
Participants	4a	Eligibility criteria for participants	Eligibility criteria for clusters	6
	4b	Settings and locations where the data were collected		8
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	Whether interventions pertain to the cluster level, the individual participant level or both	7
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and	Whether outcome measures pertain to the cluster level, the individual participant level or both	7

		when they were assessed		
	6b	Any changes to trial outcomes after the trial commenced, with reasons		None
Sample size	7a	How sample size was determined	Method of calculation, number of clusters(s) (and whether equal or unequal cluster sizes are assumed), cluster size, a coefficient of intracluster correlation (ICC or *k*), and an indication of its uncertainty	7
	7b	When applicable, explanation of any interim analyses and stopping guidelines		
Randomisation:				
Sequence generation	8a	Method used to generate the random allocation sequence		6
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	Details of stratification or matching if used	6
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	Specification that allocation was based on clusters rather than individuals and whether allocation concealment (if any) was at the cluster level, the individual participant level or both	7
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Replace by 10a, 10b and 10c	
	10a		Who generated the random allocation sequence, who enrolled clusters, and who assigned clusters to interventions	8
	10b		Mechanism by which individual participants were included in clusters for the purposes of the trial (such as complete	8

			enumeration, random sampling)	
	10c		From whom consent was sought (representatives of the cluster, or individual cluster members, or both), and whether consent was sought before or after randomisation	8
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how		7
	11b	If relevant, description of the similarity of interventions		Not relevant
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	How clustering was taken into account	13
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses		13/14
Results				
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	For each group, the numbers of clusters that were randomly assigned, received intended treatment, and were analysed for the primary outcome	14
	13b	For each group, losses and exclusions after randomisation, together with reasons	For each group, losses and exclusions for both clusters and individual cluster members	14
Recruitment	14a	Dates defining the periods of recruitment and follow-up		8
	14b	Why the trial ended or was stopped		Not stopped early
Baseline data	15	A table showing baseline demographic and clinical	Baseline characteristics for the individual and cluster levels as	16

		characteristics for each group	applicable for each group	
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	For each group, number of clusters included in each analysis	16
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	Results at the individual or cluster level as applicable and a coefficient of intracluster correlation (ICC or k) for each primary outcome	16-20
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended		
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory		
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms[3])		None
Discussion				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses		21-24
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	Generalisability to clusters and/or individual participants (as relevant)	
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		21-24
Other information				
Registration	23	Registration number and		2

		name of trial registry	
Protocol	24	Where the full trial protocol can be accessed, if available	Provided & methods paper referenced
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	25

** Note: page numbers optional depending on journal requirements*

Table 2: Extension of CONSORT for abstracts[1,2] to reports of cluster randomised trials

Item	Standard Checklist item	Extension for cluster trials
Title	Identification of study as randomised	**Identification of study as cluster randomised**
Trial design	Description of the trial design (e.g. parallel, cluster, non-inferiority)	
Methods		
Participants	Eligibility criteria for participants and the settings where the data were collected	**Eligibility criteria for clusters**
Interventions	Interventions intended for each group	
Objective	Specific objective or hypothesis	**Whether objective or hypothesis pertains to the cluster level, the individual participant level or both**
Outcome	Clearly defined primary outcome for this report	**Whether the primary outcome pertains to the cluster level, the individual participant level or both**
Randomization	How participants were allocated to interventions	**How clusters were allocated to interventions**
Blinding (masking)	Whether or not participants, care givers, and those assessing the outcomes were blinded to group assignment	
Results		
Numbers randomized	Number of participants randomized to each group	**Number of clusters randomized to each group**
Recruitment	Trial status[1]	
Numbers analysed	Number of participants analysed in each group	**Number of clusters analysed in each group**
Outcome	For the primary outcome, a result for each group and the estimated effect size and its precision	**Results at the cluster or individual participant level as applicable for each primary outcome**
Harms	Important adverse events or side effects	
Conclusions	General interpretation of the results	
Trial registration	Registration number and name of trial register	
Funding	Source of funding	

[1] Relevant to Conference Abstracts

REFERENCES

1 Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet* 2008, 371:281-283

2 Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG at al (2008) CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. *PLoS Med* 5(1): e20

3 Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, Moher D. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004; 141(10):781-788.

BMJ Open

The impact of referral templates on patient experience of the referral and care process. A cluster randomized trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-011651.R2
Article Type:	Research
Date Submitted by the Author:	13-Sep-2016
Complete List of Authors:	Wåhlberg, Henrik; Universitetet i Tromso Institutt for Samfunnsmedisin ISM, ; Sykehuset Østfold Kalnes, Medical department Braaten, Tonje; University of Tromsø, Department of Community Medicine Broderstad, Ann Ragnhild; Uit the Arctic University of Norway, Faculty of Health; University Hospital of North Norway, medical department
Primary Subject Heading:	Health services research
Secondary Subject Heading:	General practice / Family practice, Patient-centred medicine
Keywords:	Patient experience, Referral, General practitioner

Title: The impact of referral templates on patient experience of the referral and care process. A cluster randomized trial

Henrik Wåhlberg – corresponding author

Department of Community Medicine, UiT The Arctic University of Norway

Faculty of Health Sciences, UiT The Arctic University of Norway

9037 Tromsø, Norway

University Hospital of North Norway Harstad

St. Olavsgate 70, 9480 Harstad, Norway

henrik.wahlberg@uit.no/henrik_waahlberg@hotmail.com

Tonje Braaten

Department of Community Medicine, UiT The Arctic University of Norway

Faculty of Health Sciences, UiT The Arctic University of Norway

9037 Tromsø, Norway

tonje.braaten@uit.no

Ann Ragnhild Broderstad

Centre for Sami Health Research, UiT The Arctic University of Norway

UiT The Arctic University of Norway, 9037 Tromsø, Norway

University Hospital of North Norway Harstad

St. Olavsgate 70, 9480 Harstad, Norway

ann.ragnhild.broderstad@uit.no

Keywords: Patient experience. Referral. General practitioner.

Wordcount: 3370 words (excluding abstract, references, tables and author contributions)

ABSTRACT

Objectives: To evaluate if a referral intervention improves the patient experience of the referral and treatment process.

Setting: Interface between fourteen primary care surgeries and a district general hospital.

Participants: The fourteen GP surgeries (seven intervention, seven control) in the area around the University Hospital of North Norway Harstad were randomised and all completed the study. Consecutive individual patients were recruited at their hospital appointment. A total of 500 patients were recruited with 281 in the intervention and 219 in the control arm.

Interventions: Dissemination of referral templates for four diagnostic groups (dyspepsia, suspected colorectal cancer, chest pain and COPD) coupled with intermittent surgery visits by study personnel. Control arm continued standard referral practice. The intervention was in use for 2.5 years.

Outcome: The main outcome was a quality indicator score. This paper reports a secondary outcome, the patient experience, as measured by self-report questionnaires. General practitioners in the intervention group could not be blinded. Patients were blinded to intervention status. Analysis was based on single question comparison with a questionnaire subscore used to assess the effect of clustering.

Results: On the individual questions overall satisfaction was very high with minor differences between the intervention and control group. Interestingly the most negative responses, in both groups concerned questions relating to patient interaction and information. Very little evidence of clustering was found with an estimated intracluster correlations coefficient at $1.21e^{-11}$.

Conclusion: In total this indicates no clear effect of the implementation of referral templates on the patient experience, in a setting of generally high patient satisfaction.

Trial registration: This trial has been registered at ClinicalTrials.gov. The trial registration number is NCT01470963.

ARTICLE SUMMARY

Strengths and limitations of the study

- Clinically relevant research in a regular district hospital setting
- High response rate
- Newly developed questionnaire hampers wider generalisation

INTRODUCTION

Evaluation of patient experience and satisfaction is widespread with a wealth of literature concerning the development and use of questionnaires[1-5]. The evaluation of patient experience can help drive quality improvement[6] and improved patient experience is associated with safety and clinical effectiveness[7].

Care coordination is an important aspect of a well-functioning high quality health service. It has been defined as "the deliberate integration of patient care activities between two or more participants involved in a patient's care to facilitate the appropriate delivery of health care services."[8] In the US the National Quality Forum (NQF) has published preferred practices for care coordination, including transitions of care[9]. This report includes clear recommendations for participation of the patient, or his/her designee, in the decision, planning and execution of a care transition. This is important, as exemplified by a recent Australian article, where patients with colorectal cancer perceived that poor information exchange led to suboptimal care[10]. Hence assessing patient experience of the referral process may be beneficial in assessing the effect of a referral intervention.

This article presents the patient experience aspect of a cluster randomized study evaluating the effect of the implementation of referral templates for four diagnostic groups – dyspepsia, suspected colorectal cancer, chest pain and COPD - in the patient referral pathway[11]. Previously we have shown that the referral templates led to increased referral quality[12], and the effect on the main outcome, quality of care at the

hospital, is in publication. This publication aims to assess whether the implementation of a referral template in the transition of care from the general practitioner to the hospital has affected the patient experience of the care process.

METHODS

Study setting

In Norway the health care system is quite uniformly organised throughout the country. GPs act as gatekeepers to secondary care[13] with specialist health services delivered by governmentally owned regional health authorities, mainly through public hospitals. Some specialist outpatient care is delivered by private specialists, but this is mainly purchased by the regional health authorities. The access to private specialists in the geographical area of the current study is very limited.

Study design

The study was designed as a cluster randomized study with the general practitioner (GP) surgery as the clustering unit. A total of 14 surgeries were randomized, seven to the intervention and seven to the control group. The clustered design was chosen to avoid possible spill-over effect from the intervention to control GPs. Randomization was done by simple drawing by a person not connected to the research team, stratified by town vs countryside location of surgery.

As the intervention was to be actively used by the GPs the referring GP could not be blinded. Patients, hospital doctors and outcome evaluators were blinded to the intervention status of the patients. Due to the design of the intervention the referral letter would sometimes reveal the intervention status, if the electronic template was used. No separate sample size calculation was performed for the patient experience outcome. The full study details are published in the methods paper[11].

Intervention

The intervention consisted of the distribution of four separate referral templates to the intervention surgeries. These templates covered four clinical areas (dyspepsia, suspected colorectal malignancy, chest pain and chronic obstructive pulmonary disease). The templates were to be use when initiating a new referral to the medical outpatient clinic at the University Hospital of North Norway, Harstad (UNN Harstad). The templates were distributed by the corresponding author (HW) during educational and/or lunch meetings and were provided as laminated reference sheets or in electronic form. In addition follow up visits were conducted regularly during the study period and intermittent mail leaflets and reminders were distributed to the intervention offices. Control offices continued standard referral practice.

Outcomes

The main outcome in the project was the quality of care delivered to each individual patient. In addition health process indicators such as correct prioritization

were recorded and referral quality was also compared between the intervention and control group. The current paper presents the patient experience aspect of the study, as measured by self-report questionnaires.

Participants

The fourteen GP surgeries primarily served by UNN Harstad were included in the randomization process. In 2013 these surgeries had a total list size of 39,523 patients. The individual patients were recruited from consecutive new patients within one of the four clinical areas referred to, at the medical outpatient clinics at UNN Harstad. Study information and a consent form were sent to each individual patient together with their appointment letter. Further information, including a new consent form if appropriate, was provided at their hospital appointment. The individual patients were analyzed as part of the intervention or control group depending on the intervention status of the GP surgery they were referred from. Children (<18 years of age) and patients with reduced capacity to consent were excluded from the study.

Recruitment

The study recruited patients for approximately 2.5 years and a total of 538 patients were included with 290 in the intervention arm and 227 in the control arm. The remaining 21 patients were referred from GP surgeries outside the regular area of UNN Harstad, and as such neither in the intervention nor the control group. These 21 were

excluded, together with 17 patients that did not fill the inclusion criteria. In total this left 281 patients in the intervention arm and 219 patients in the control arm (Figure 1).

Questionnaire development

Multiple tools exist for measuring different aspects of care coordination[14] and patient experience, however no complete questionnaire was located that covered the area in the current study completely. Therefore a questionnaire was developed by combining validated questionnaires regarding patient experiences and care coordination. The questions used were the full version of the Generic Short Patient Experiences Questionnaire (GS-PEQ)[15], together with 2 further questions used in patient experience questionnaires in the Norwegian health care system (Question 11 and 12)[16] and the two questions about health interaction from the Commonwealth Fund Survey 2010[17]. Three further questions were added to assess (1) who referred the patient, (2) if the referral was seen as appropriate and (3) an overall evaluation of the institution. Table 1 presents the questions in the questionnaire. GS-PEQ and Question 11-12 use Likert style response categories. The health interaction questions had a yes/no response. The full questionnaire, including the demographic questions, is available upon request.

Table 1 – Questionnaire details

Question No	Wording of question
1	Did the clinicians[a] talk to you in a way that was easy to understand?
2	Do you have confidence in the clinicians' professional skills?
3	Did you get sufficient information about how examinations and tests were to be performed?
4	Did you get sufficient information about your diagnosis/conditions?

5	Did you perceive the treatment to be adapted to your situation?
6	Were you involved in decisions regarding your treatment?
7	Did you perceive the institution work practices to be well organized?
8	Did you perceive the equipment at the institution to be in good working order?
9	Overall, was the help and treatment you received at the institution satisfactory?
10	Do you believe that you were in any way given incorrect treatment (according to your own judgment)?
11	Did you have to wait before you were given an appointment at the institution?
12	Overall, what benefit have you had from the care at the institution?
13	Did the hospital specialist lack basic medical information from your GP about the reason for your visit or test results?
14	After your saw the hospital specialist did your GP lack important information about the care you got from the specialist?
15	Was the referral to the outpatient department necessary (according to your own judgment)?
16a	Were you referred by your GP for the outpatient appointment?
16b	If no in question 16a; who referred you?
17	If you take an overview of your entire treatment process, how would you evaluate the institution?

[a] With 'clinicians' we mean those who had the main treatment responsibility. This is linguistically clearer in the Norwegian wording.

The questionnaire was piloted for content validity with four local health professionals; these felt that it covered the important aspects of patient experience and care coordination. It was then piloted with five outpatients with a median age of 72 years (average 68.8 years) to ensure face validity and acceptability. Two patients needed clarification on one of the questions before they felt they could answer, and the wording of this question was adjusted accordingly. The patients felt the questionnaire was acceptable, with logical response categories and that the questions covered their clinical path during the referral process well. These patients did not take part when the project was later initiated. No further formal evaluation of the questionnaire was carried out.

Questionnaire distribution

The questionnaire was mailed to patients who had consented to take part in the referral project presented above. To increase response rates a pre-paid response envelope was included, addresses were handwritten, the questionnaire was kept short and association with research bodies was indicated[3]. For non-responders one reminder was sent approximately one month after the first questionnaire, with a new questionnaire and pre-paid response envelope.

Ethics

The study followed the principles of the Helsinki Declaration. Before recruitment started it was presented to the Regional Committee for Medical and Health Research Ethics Northern Norway,, who determined it not to be within the scope of the Health Research Act (REK NORD 2010/2259). The project has been approved by the Data Protection Officer for Research. The study is registered at www.clinicaltrials.gov with trial registration number NCT01470963. All patients provided written informed consent.

Imputation

To further aid the assessment of clustering missing data was imputed. For the imputation answers set as 'not applicable' were counted as missing. Missing data was seen to be random and multiple imputation using chained equations was employed. This has been shown to perform well for a variety of variable scaling types[18]. Every variable used in further statistical analysis was entered into the imputation model, as failure to do

so may bias estimates towards the null[19]. The ordinal response scales for each single question were to be combined into a continuous score, and as such it was determined that imputation with predictive mean matching was appropriate. As shown by van Buuren et al the number of iterations can usually be quite low, between 5 and 20[19]. In this study the Stata standard of 10 iterations as burn-in period was used.

Statistical analysis

Results are presented on single question basis with comparison between the two groups using the Mann-Whitney U test for ordinal data and Chi square test for nominal data. No correction for clustering was made as the estimated ICC was very low (shown below). Aggregation of scores was postulated in the methods paper[11], but discarded as a main outcome as properties of the questionnaire, with a "not applicable" answering category are not easily suitable for such an approach. However, to assess the effect of clustering a sum of scores from the GS-PEQ part of the questionnaire was calculated and a multilevel regression model was built with the GS-PEQ score from the questionnaire as the dependent variable. Intervention status was included in the model as this is the main point of interest. Gender, age, education level and self-perceived health were included in the model, as these tend to influence patient experience[20-22]. Age was centered to ease interpretation in a mixed model analysis[23]. Self-perceived health was reported on a five level Likert style scale and education level in four categories. Both were included as dummy variables in the analysis. Other confounding variables were assessed in the model and included if their inclusion led to a >10% change in the main outcome when added to the base model (main outcome +

intervention status). Relevant interactions were checked for relevant variables, where $p<0.10$ was set as the significance level. As imputation was used, Monte Carlo error estimates were employed to assess the level of simulation error, as suggested by White and Royston[24]. Normal evaluation of multilevel models with log likelihood ratio tests were not carried out, as this is not well defined for multilevel models with imputed data. The analysis employed restricted maximum likelihood techniques throughout, as suggested when the number of clusters is small[25]. As described multiple imputation was used to account for missing data in the multilevel regression model assessing the effect of clustering. Stata version 13.1 (StataCorp 2013, TX) were used for all analysis.

RESULTS

Response rate

Response rate was 69.4% before reminders were sent out, rising to 82.0% after reminders (Figure 1). Mean age for responders were 61 years and for non-responders 47 years (t-test <0.0001). There was no significant gender difference between the responders and non-responders, and the response rate did not differ significantly between the intervention and control group (χ^2-test).

Missing data

Missing data for most questions was low, ranging from 0 to 11 out of 410 answered questionnaires. Statistically these were considered missing completely at

random (MCAR) with no clear relation to either age, gender, self-perceived health, disease severity or other variables[26]. However, Questions 6, 10 and 12 in the general part of the questionnaire had higher amounts of not applicable ranging from 14 to 34 representing 3.4% to 8.3% of returned questionnaires. In these questions the word 'treatment' was used. This was intended to cover the medical examination and interventions during the outpatient visit. However, it seems that this has been misunderstood by several patients. It seems reasonable to assume that patients who underwent 'only' diagnostic evaluation felt that they had received no 'treatment', and hence felt unable to answer the question. This was also highlighted by one patient in a free-text response in the questionnaire. "Not applicable" to Question 6, 10 and 12 did not vary significantly with age (t-test), intervention group status (χ^2-test), gender (χ^2-test) or self-reported health (χ^2-test). This was treated as missing at random for imputation purposes (MAR)[26]. Question 14 had a missing rate of 15.9% but also yielded a high level of not applicable responses, at 46.0% of returned questionnaires. This was expected, as many people will not have had a new appointment with their GP following the hospital outpatient evaluation. It is also reasonable to assume that the high amount of missing was related to the same concept. The response "not applicable" did not significantly vary with age (t-test p=0.868), intervention group status (χ^2-test p=0.064) or self-reported health (χ^2-test p=0.459).

A histogram of responses to questions with five categories showed all response sets to be skewed to the left. However, earlier work has indicated that multiple imputation can perform well, even when the categorical variable is non-normally distributed, as long as MAR does not exceed 10%[27]. In a 2010 article Finch argues that multiple imputation performs well for imputation of missing categorical questionnaire

data[26]. There was no association between levels of missing data and the multilevel structure of the data.

Baseline characteristics

Baseline characteristics are presented in Table 2. There was no major difference between the intervention and control group with regard to gender, age, urban or rural residency or questionnaire response. The effect of the referral intervention on referral quality has previously been shown to be clinically significant with an effect of 18% (95% CI 11, 25 p<0.001)[12]. However this was for the full dataset of 500 patients. To ensure that this was also representative of the subpopulation who answered the questionnaire, the multi-level regression model was employed using data from only the 410 patients who answered it. This showed an intervention effect of 19% (95% CI 12, 27 p<0.001) on referral quality, well within the 95% CI of the full analysis.

Table 2: Selected patient baseline characteristics by intervention status

	Intervention group	Control group	p-value
Female/male, N(%)	140 (59.3)/96 (40.7)	102 (58.6)/72 (41.4)	0.89
Age (year), mean (±SD)	60.9 ±12.5	60.3 ± 13.5	0.63
Urban/rural, N(%)	145 (61.4)/91 (38.6)	95 (54.6)/79 (45.4)	0.17
Clinical group, N(%)			
- dyspepsia	117 (49.6)	96 (55.2)	0.29
- suspected colorectal malignancy	75 (31.8)	57 (32.8)	
- chest pain	40 (17.0)	18 (10.3)	
- chronic obstructive	4 (1.7)	3 (1.7)	

pulmonary disease			
Hospital appointment with senior house officer/ specialist, N(%)	107 (45.3)/129 (54.7)	78 (44.8)/96 (55.2)	0.92
Questionnaire returned promptly/after mailed reminder, N(%)	202 (85.6)/34 (14.4)	145 (83.3)/29 (16.7)	0.53

Questionnaire results

Overall satisfaction with services was high and as presented in Table 3 there was little difference between the intervention and control group for the individual questions. Using the Mann-Whitney U test, χ^2-test and Fisher exact test only two questions had significant p-values (Q14 and Q17), however in both these questions the absolute difference in numbers was very small. All response sets were skewed to the left, that is, towards more positive responses.

Interestingly the highest numbers of scores indicating dissatisfaction were for Questions 4 and 6, both for the intervention and control group patients. These questions concern patient interaction and information on the treatment process.

Table 3: Questionnaire results

Question	Answering categories[a]	Intervention	Control	p-value
Question 1 [b]		5 (4, 5)	5 (4, 5)	0.92
Question 2 [b]		5 (4, 5)	4 (4, 5)	0.39
Question 3 [b]		5 (4, 5)	4 (4, 5)	0.23
Question 4 [b]		4 (3, 5)	4 (4, 4)	0.12
Question 5 [b]		4 (4, 5)	4 (4, 5)	0.88
Question 6 [b]		4 (3, 5)	4 (3, 4)	0.19

Question 7 [b]		4 (4, 5)	4 (4, 5)	0.22
Question 8 [b]		4 (4, 5)	4 (4, 5)	0.81
Question 9 [b]		5 (4, 5)	4 (4, 5)	0.15
Question10 [b]		5 (5, 5)	5 (5, 5)	0.60
Question 11 [c]	No	33 (14.0)	21 (12.1)	0.33
	Yes, but not too long	155 (66.0)	111 (64.2)	
	Yes, quite long	34 (14.5)	29 (16.8)	
	Yes, too long	13 (5.5)	12 (6.9)	
Question 12 [c]	No benefit	3 (1.4)	5 (3.1)	0.56
	Little benefit	12 (5.5)	7 (4.3)	
	Some benefit	59 (27.2)	44 (27.0)	
	Large benefit	106 (48.9)	86 (52.8)	
	Very large benefit	37 (17.1)	21 (12.9)	
Question 13 [c]	Yes	4 (1.7)	6 (3.5)	0.25
	No	229 (98.3)	165 (96.5)	
Question 14 [c]	Yes	4 (4.2)	8 (13.1)	0.04
	No	92 (95.8)	53 (86.9)	
Question 15 [c]	Yes	232 (99.2)	170 (99.4)	0.75
	No	2 (0.8)	1 (0.6)	
Question 17 [c]	Much poorer than expected	0 (0)	1 (0.6)	0.03
	Somewhat poorer than expected	0 (0)	5 (3.1)	
	As expected	119 (54.1)	94 (58.4)	
	Somewhat better than expected	50 (22.7)	32 (19.9)	
	Much better than expected	51 (23.2)	29 (18.0)	

[a] for question 1-10 the following scoring was used: 1 = Not at all, 2 = To a small extent, 3 = To some extent, 4 = To a large extent and 5 = To a very large extent
[b] data presented as median(25th percentile, 75th percentile)
[c] data presented as number(%)

The Cronbach alpha for Question 1-15 was 0.83 and for Question 1-10 0.88.

Assessment of clustering effect

In the regression model, no significant difference was seen in the GSPEQ score between the intervention and control group with theregression coefficient 0.55 (95% CI -0.37, 1.47 p=0.24) when taking clustering into account and adjusted for confounding variables 0.57 (95% CI -0.31, 1.46 p=0.20). No significant interaction was found and the result was not confounded by GP specialist status, GP gender, specialist status of hospital doctor or seriousness of final diagnosis. The Monte Carlo error estimates were within the limits recommended[24]. Initial multilevel analysis of the data revealed virtually no variance of the intercepts. The ICC was estimated at $1.21e^{-11}$. Hence very little of the variation in the data was related to the clustered design.

DISCUSSION

In the presentation of the data from each question in Table 3 it is quite clear that, for the most part, patients in this project report positive experiences, with no differences between the intervention and control group. It hence seems that although the intervention has increased the referral quality significantly[12], this has not translated into a more positive patient experience with the referral process and treatment, as measured by self-report questionnaires. In the current study in depth data analysis with imputation and multilevel regression modeling was employed to further explore the effect of clustering. No clear effect of clustering was found.

A strength of the current study is the fairly high response rate (82.0%) compared to other mail response studies[28]. However, the potential for non-response bias is always present. Others have previously shown the effect of this to be small[29;30]. Earlier Norwegian studies have suggested only minor differences between answers provided by responders and non-responders, when the latter have been obtained through telephone follow-up interviews[31-33]. A clear limitation is the use of short form questionnaires with single items, which may be less valid than longer forms[34]. However, shorter forms will increase the response rate[4;35]. The current project aimed to assess the effect of a health system intervention and the patient experiences with care after this intervention. We hence decided to keep the questionnaire short to enable a high response rate and keep the patient and staff workload manageable.

The current project used a newly developed questionnaire to assess patient experience by combining previously validated questions. The general nature of the final questionnaire may be seen as a weakness, as small changes in the patient experience induced by the intervention may have been missed. Further piloting might have revealed more clearly if the questionnaire did indeed assess the patient experience with the referral and care process in an adequate way. However in this clinically oriented project the authors hoped that a more general questionnaire would highlight whether the intervention would cause a more overall positive, or even a negative, change. It is probable that for each individual patient it is the experience with the entire process that matters, as opposed to the experience of a subpart of the process. If large scale implementation of referral guidance is contemplated a more specific questionnaire may need to be validated.

An additional weakness was the lack of a sound analytical plan proposed in the methods paper[11]. To ensure transparency the analysis presented in this paper is therefore simple and based on single question assessment. Given the clustered nature of the study an assessment of clustering is given for a subsection of the questionnaire, but very little effect was seen.

Comparison with other studies was difficult as no clearly comparable analysis was found, except for the two health interaction questions. In the current study 1.7% in the intervention group and 3.5% in the control group felt the hospital specialist lacked information from the GP. 4.2% in the intervention and 13.1% in the control group felt the GP lacked information from the specialist. In the Norwegian part of the 2010 Commonwealth Fund Survey, the same questions gave much higher negative ratings, with 12.1% indicating that the specialist lacked information from the GP and 38.3% indicating that the GP lacked information from the hospital[17]. Data from the 2013 Commonwealth Fund Survey suggest similar ratings as in 2010, although the wording of the questions is slightly different[36]. A Norwegian report concerning patient experience as inpatients also suggests higher dissatisfaction with co-operation between the hospital and the GP[37] than in the current study. In total this clearly suggests that the patient experience of the GP/specialist communication is better in a small district hospital than the country average suggests. It is therefore possible that the effect of the intervention on patient experience could have been higher if the level of dissatisfaction with the health care cooperation had been higher in the local population. However, this

also may suggest that although the hospital consultants often feel information is lacking in the referrals[38;39], this is not necessarily experienced as a problem by patients.

In the current study, two questions were answered more negatively than others. These questions therefore probably provide the most interesting points for further quality improvement at the local facility. These two questions represent areas where communication is the main concept, namely patient involvement in the treatment process and information from doctors to patients. Others have previously shown communication and information errors as a cause for dissatisfaction[40], and in other jurisdictions even malpractice claims[41;42].

CONCLUSION

In this project, patient satisfaction, as measured by patient experience questionnaires, was generally high, with no major differences between the intervention and control group. No clear effect of the implementation of referral templates on patient satisfaction was evident.

Interestingly, the most negative feedback, from both intervention and control group, was concerning patient interaction, involvement and information. Effecting communication and involving patients in decision making may help to increase patient satisfaction to an even higher level.

AUTHOR CONTRIBUTIONS:

The administration and daily running of the study was performed by HW, who was also the grant holder. ARB and HW developed the questionnaire. All authors participated in the analysis and interpretation of the data. All authors revised drafts of the manuscript and approved the final version.

FUNDING STATEMENT:

This work was funded by a research grant from the Northern Norway Regional Health Authority (Helse Nord RHF) with grant number HST1026-11.

COMPETING INTERESTS:

The authors declare that they have no competing interests.

DATA SHARING:

No additional data available. The dataset supporting the conclusions in this article may be available on request to the main author (HW).

FIGURE LEGENDS:

Figure 1: Patient inclusion and questionnaire response

Reference List

(1) Crow R, Gage H, Hampson S, Hart J, Kimber A, Storey L et al. The measurement of satisfaction with healthcare: implications for practice from a systematic review of the literature. *Health Technol Assess* 2002; 6(32):1-244.

(2) Garrat A, Solheim E, Danielsen K. National and cross-national surveys of patient experiences: a structured review. 7-2008. 2008. Oslo, Norwegian Knowledge Centre for the Health Services.

(3) Edwards PJ, Roberts I, Clarke MJ, DiGuiseppi C, Wentz R, Kwan I et al. Methods to increase response to postal and electronic questionnaires. *Cochrane Database Syst Rev* 2009;(3):MR000008.

(4) Edwards P. Questionnaires in clinical trials: guidelines for optimal design and administration. *Trials* 2010; 11(1):2.

(5) Fitzpatrick R, Davey C, Buxton M, Jones D. Evaluating patient-based outcome measures for use in clinical trials: a review. *Health Technology Assessment* 1998; 2(14):74.

(6) Ellins J. Great expectations? Reflections on the future of patient and public involvement in the NHS. *Clinical Medicine* 2011; 11(6):544-547.

(7) Doyle C, Lennox L, Bell D. A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. *BMJ Open* 2013; 3(1).

(8) McDonald K, Sundaram V, Bravata D, et al. Closing the Quality Gap: A Critical Analysis of Quality Improvement Strategies (Vol. 7: Care Coordination). Rockville (MD): Agency for Healthcare Research and Quality (US); 2007.

(9) National Quality Forum. Preferred Practices and Performance Measures for Measuring and Reporting Care Coordination: A Consensus Report. National Quality Forum, editor. 2010. Washington, DC, National Quality Forum.

(10) Pascoe S, Veitch C, Crossland L, Beilby J, Spigelman A, Stubbs J et al. Patients' experiences of referral for colorectal cancer. *BMC Family Practice* 2013; 14(1):124.

(11) Wahlberg H, Valle PC, Malm S, Broderstad AR. Practical health co-operation - the impact of a referral template on quality of care and health care co-operation: study protocol for a cluster randomized controlled trial. *Trials* 2013; 14(1):7.

(12) Wahlberg H, Valle P, Malm S, Broderstad A. Impact of referral templates on the quality of referrals from primary to secondary care: a cluster randomised trial. *BMC Health Services Research* 2015; 15(1):353.

(13) Carlsen B, Norheim OF. Introduction of the patient-list system in general practice. Changes in Norwegian physicians' perception of their gatekeeper role. *Scand J Prim Health Care* 2003; 21(4):209-213.

(14) Schultz E, Pineda N, Lonhart J, Davies S, McDonald K. A systematic review of the care coordination measurement landscape. *BMC Health Services Research* 2013; 13(1):119.

(15) Sjetne I, Bjertnaes O, Olsen R, Iversen H, Bukholm G. The Generic Short Patient Experiences Questionnaire (GS-PEQ): identification of core items from a survey in Norway. *BMC Health Services Research* 2011; 11(1):88.

(16) Sjetne IS, Bjertnes OA, Iversen HH, Olsen RV. Pasienterfaringer i spesialisthelsetjenesten. Et generisk, kort spørreskjema. 2009. Oslo, Nasjonalt kunnskapssenter for helsetjenesten.

(17) Skudal KE, Bjertnes OA, Holmboe O, Bukholm G, Rottingen JA. The 2010 Commonwealth Fund survery: Results from a comparative population survery in 11 countries. 21. 2010. Oslo, The Norwegian Knowledge Centre for the Health Services.

(18) Lee KJ, Carlin JB. Multiple Imputation for Missing Data: Fully Conditional Specification Versus Multivariate Normal Imputation. *American Journal of Epidemiology* 2010; 171(5):624-632.

(19) van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research* 2007; 16(3):219-242.

(20) Xiao H, Barber JP. The effect of perceived health status on patient satisfaction. *Value Health* 2008; 11(4):719-725.

(21) Fitzpatrick R. Surveys of patient satisfaction: II--Designing a questionnaire and conducting a survey. *BMJ* 1991; 302(6785):1129-1132.

(22) Da Costa D, Clarke AE, Dobkin PL, Senecal JL, Fortin PR, Danoff DS et al. The relationship between health status, social support and satisfaction with medical care among patients with systemic lupus erythematosus. *Int J Qual Health Care* 1999; 11(3):201-207.

(23) Twisk J. Applied Multilevel Analysis. Cambridge: Cambridge University Press; 2006.

(24) White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Statist Med* 2011; 30(4):377-399.

(25) Eldridge S, Kerry S. A Practical Guide to Cluster Randomised Trials in Health Services Research. 1 ed. Chichester: John Wiley & Sons, Ltd; 2012.

(26) Finch W. Imputation Methods for Missing Categorical Questionnaire Data: A Comparison of Approaches. *Journal of Data Science* 2010; 8:361-378.

(27) Leite W, Beretvas S. The Performance of Multiple Imputation for Likert-type Items with Missing Data. *Journal of Modern Applied Statistical Methods* 2010;(9):64-74.

(28) Sitzia J, Wood N. Response rate in patient satisfaction research: an analysis of 210 published studies. *Int J Qual Health Care* 1998; 10(4):311-317.

(29) Lasek RJ, Barkley W, Harper DL, Rosenthal GE. An Evaluation of the Impact of Nonresponse Bias on Patient Satisfaction Surveys. *Medical Care* 1997; 35(6).

(30) Perneger TV, Chamot E, Bovier PA. Nonresponse Bias in a Survey of Patient Perceptions of Hospital Care. *Medical Care* 2005; 43(4).

(31) Guldvog B, Hofoss D, Pettersen KI, Ebbesen J, Ronning OM. PS-RESKVA - pasienttilfredshet i sykehus [PS-RESKVA -patient satisfaction in hospitals]. *Tidsskr Nor Laegeforen* 1998; 118:386-391.

(32) Do non-respondents' characteristics and experiences differ from those of respondents? Poster at the the 25th International Conference by the Society for Quality in Health Care (ISQua). 25th International Conference by the Society for Quality in Healthcare (ISQua); 08; 2008.

(33) Using telephone interviews to explore potential non-response bias in a national postal survey. Poster at the the 26th International Conference by the Society for Quality in Health Care (ISQua). 26th International Conference by the Society for Quality in Health Care (ISQua); 09; 2009.

(34) Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; 30(6):473-483.

(35) Eaker S, Bergstrøm R, Bergstrøm A, Adami HO, Nyren O. Response Rate to Mailed Epidemiologic Questionnaires: A Population-based Randomized Trial of Variations in Design and Mailing Routines. *American Journal of Epidemiology* 1998; 147(1):74-82.

(36) Haugum M, Bjertnes OA, Iversen HH, Lindahl AK, Nylenna M. Commonwealth Fund's health policy survey in 11 countries: Norwegian results in 2013 and development since 2010. 27-11-2014. Oslo, The Norwegian Knowledge Centre for the Health Services.

(37) Skudal KE, Holmboe O, Iversen HH, Sjetne IS, Haugum M, Bjertnaes O. Inpatients' experience with Norwegian hospitals: National results in 2011 and changes from 2006. 3-2012. 2012. Oslo, Norwegian Knowledge Centre for the Health Services.

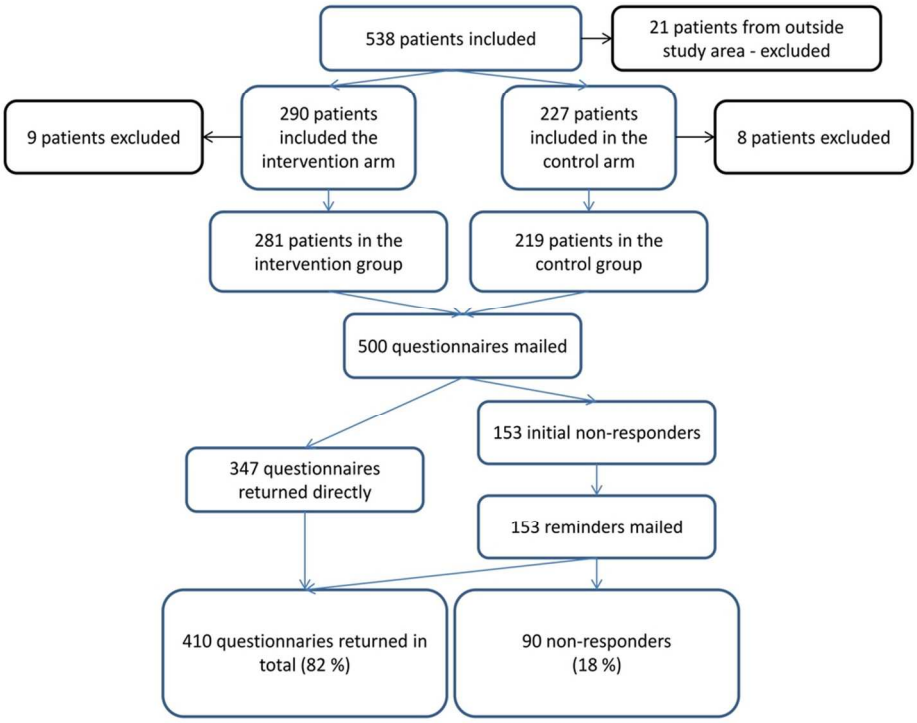
(38) Canadian Medical Association. National Survey: experiences with referrals from primary to specialty care. 2013. Canadian Medical Association.

(39) Bodek S, Ghori K, Edelstein M, Reed A, MacFadyen RJ. Contemporary referral of patients from community care to cardiology lack diagnostic and clinical detail. *Int J Clin Pract* 2006; 60(5):595-601.

(40) Hiidenhovi H, Nojonen K, Laippala P. Measurement of outpatients' views of service quality in a Finnish university hospital. *Journal of Advanced Nursing* 2002; 38(1):59-67.

(41) Levinson W, Roter DL, Mullooly JP, Dull VT, Frankel RM. Physician-patient communication: The relationship with malpractice claims among primary care physicians and surgeons. *JAMA* 1997; 277(7):553-559.

(42) Beckman HB, Markakis KM, Suchman AL, Frankel RM. The doctor-patient relationship and malpractice: Lessons from plaintiff depositions. *Archives of Internal Medicine* 1994; 154(12):1365-1370.



Patient inclusion and questionnaire response

95x71mm (300 x 300 DPI)

Table 1: CONSORT 2010 checklist of information to include when reporting a cluster randomised trial

Section/Topic	Item No	Standard Checklist item	Extension for cluster designs	Page No *
Title and abstract				
	1a	Identification as a randomised trial in the title	Identification as a cluster randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)[1,2]	See table 2	2
Introduction				
Background and objectives	2a	Scientific background and explanation of rationale	Rationale for using a cluster design	6
	2b	Specific objectives or hypotheses	Whether objectives pertain to the the cluster level, the individual participant level or both	7-8
Methods				
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	Definition of cluster and description of how the design features apply to the clusters	6
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons		None
Participants	4a	Eligibility criteria for participants	Eligibility criteria for clusters	6
	4b	Settings and locations where the data were collected		8
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	Whether interventions pertain to the cluster level, the individual participant level or both	7
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and	Whether outcome measures pertain to the cluster level, the individual participant level or both	7

		when they were assessed		
	6b	Any changes to trial outcomes after the trial commenced, with reasons		None
Sample size	7a	How sample size was determined	Method of calculation, number of clusters(s) (and whether equal or unequal cluster sizes are assumed), cluster size, a coefficient of intracluster correlation (ICC or *k*), and an indication of its uncertainty	7
	7b	When applicable, explanation of any interim analyses and stopping guidelines		
Randomisation:				
Sequence generation	8a	Method used to generate the random allocation sequence		6
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	Details of stratification or matching if used	6
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	Specification that allocation was based on clusters rather than individuals and whether allocation concealment (if any) was at the cluster level, the individual participant level or both	7
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	Replace by 10a, 10b and 10c	
	10a		Who generated the random allocation sequence, who enrolled clusters, and who assigned clusters to interventions	8
	10b		Mechanism by which individual participants were included in clusters for the purposes of the trial (such as complete	8

			enumeration, random sampling)	
	10c		From whom consent was sought (representatives of the cluster, or individual cluster members, or both), and whether consent was sought before or after randomisation	8
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how		7
	11b	If relevant, description of the similarity of interventions		Not relevant
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	How clustering was taken into account	13
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses		13/14
Results				
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	For each group, the numbers of clusters that were randomly assigned, received intended treatment, and were analysed for the primary outcome	14
	13b	For each group, losses and exclusions after randomisation, together with reasons	For each group, losses and exclusions for both clusters and individual cluster members	14
Recruitment	14a	Dates defining the periods of recruitment and follow-up		8
	14b	Why the trial ended or was stopped		Not stopped early
Baseline data	15	A table showing baseline demographic and clinical	Baseline characteristics for the individual and cluster levels as	16

		characteristics for each group	applicable for each group	
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	For each group, number of clusters included in each analysis	16
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	Results at the individual or cluster level as applicable and a coefficient of intracluster correlation (ICC or k) for each primary outcome	16-20
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended		
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory		
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms[3])		None
Discussion				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses		21-24
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	Generalisability to clusters and/or individual participants (as relevant)	
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		21-24
Other information				
Registration	23	Registration number and		2

		name of trial registry	
Protocol	24	Where the full trial protocol can be accessed, if available	Provided & methods paper referenced
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	25

** Note: page numbers optional depending on journal requirements*

Table 2: Extension of CONSORT for abstracts[1,2] to reports of cluster randomised trials

Item	Standard Checklist item	Extension for cluster trials
Title	Identification of study as randomised	**Identification of study as cluster randomised**
Trial design	Description of the trial design (e.g. parallel, cluster, non-inferiority)	
Methods		
Participants	Eligibility criteria for participants and the settings where the data were collected	**Eligibility criteria for clusters**
Interventions	Interventions intended for each group	
Objective	Specific objective or hypothesis	**Whether objective or hypothesis pertains to the cluster level, the individual participant level or both**
Outcome	Clearly defined primary outcome for this report	**Whether the primary outcome pertains to the cluster level, the individual participant level or both**
Randomization	How participants were allocated to interventions	**How clusters were allocated to interventions**
Blinding (masking)	Whether or not participants, care givers, and those assessing the outcomes were blinded to group assignment	
Results		
Numbers randomized	Number of participants randomized to each group	**Number of clusters randomized to each group**
Recruitment	Trial status[1]	
Numbers analysed	Number of participants analysed in each group	**Number of clusters analysed in each group**
Outcome	For the primary outcome, a result for each group and the estimated effect size and its precision	**Results at the cluster or individual participant level as applicable for each primary outcome**
Harms	Important adverse events or side effects	
Conclusions	General interpretation of the results	
Trial registration	Registration number and name of trial register	
Funding	Source of funding	

[1] Relevant to Conference Abstracts

REFERENCES

1 Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet* 2008, 371:281-283

2 Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG at al (2008) CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. *PLoS Med* 5(1): e20

3 Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, Moher D. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004; 141(10):781-788.