

BMJ Open Assessment of the effect of an Interactive Dynamic Referral Interface (IDRI) on the quality of referral letters from general practitioners to gastroenterologists: a randomised cross-over vignette trial

Sigrun Losada Eskeland,^{1,2} Cathrine Brunborg,³ Corina Silvia Rueegg,³ Lars Aabakken,⁴ Thomas de Lange^{5,6}

To cite: Eskeland SL, Brunborg C, Rueegg CS, *et al.* Assessment of the effect of an Interactive Dynamic Referral Interface (IDRI) on the quality of referral letters from general practitioners to gastroenterologists: a randomised cross-over vignette trial. *BMJ Open* 2017;7:e014636. doi:10.1136/bmjopen-2016-014636

► Prepublication history and additional material are available. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2016-014636>).

Received 7 October 2016

Revised 6 May 2017

Accepted 30 May 2017



CrossMark

For numbered affiliations see end of article.

Correspondence to

Sigrun Losada Eskeland;
sigesk@vestreviken.no

ABSTRACT

Objectives We evaluated whether interactive, electronic, dynamic, diagnose-specific checklists improve the quality of referral letters in gastroenterology and assessed the general practitioners' (GPs') acceptance of the checklists.

Design Randomised cross-over vignette trial.

Setting Primary care in Norway.

Participants 25 GPs.

Intervention The GPs participated in the trial and were asked to refer eight clinical vignettes in an internet-based electronic health record simulator. A referral support, consisting of dynamic diagnose-specific checklists, was created for the generation of referral letters to gastroenterologists. The GPs were randomised to refer the eight vignettes with or without the checklists. After a minimum of 3 months, they repeated the referral process with the alternative method.

Main outcome measures Difference in quality of the referral letters between referrals with and without checklists, measured with an objective Thirty Point Score (TPS). Difference in variance in the quality of the referral letters and GPs' acceptance of the electronic dynamic user interface.

Results The mean TPS was 15.2 (95% CI 13.2 to 16.3) and 22.0 (95% CI 20.6 to 22.8) comparing referrals without and with checklist assistance ($p<0.001$), respectively. The coefficient of variance was 23.3% for the checklist group and 39.6% for the non-checklist group. Two-thirds (16/24) of the GPs thought they had included more relevant information in the referrals with checklists, and considered implementing this type of checklists in their clinical practices, if available.

Conclusions Dynamic, diagnose-specific checklists improved the quality of referral letters significantly and reduced the variance of the TPS, indicating a more uniform quality when checklists were used. The GPs were generally positive to the checklists.

INTRODUCTION

High-quality, written communication is essential for adequate management of patients

Strengths and limitations of this study

- We used vignettes to standardise the setting, making the results objective, quantifiable and comparable.
- The randomised cross-over design makes comparison of individual general practitioners' (GPs') changes in referral letter quality possible, regardless of the GPs' initial quality level for referral letters.
- Score for objective measurements of the quality of referral letters developed by our research team
- Vignette design may have appeared unrealistic to the participating GPs, and may have resulted in frustration due to challenges in the virtual communication.
- Appropriateness and structure of the referral letters were not measured.

referred from primary to secondary care. The referral letter is frequently the only information available to the specialist when deciding the patient's priority and selecting appropriate work-up or examinations before the first consultation at the outpatient clinic.

Referral rates from general practitioners (GPs) to secondary healthcare services are increasing^{1 2} and challenge the capacity of the secondary care outpatient clinics. A lowered threshold for referral may also potentially cause medical overuse³ and reduced effectiveness of the healthcare system. A considerable proportion of referral letters are of low quality or inappropriate.^{4–20} Such letters are a challenge for the consultants when assessing the relevance and the priority of the referrals.¹⁵ A discrepancy has also been observed between the GPs' and the specialists' considerations of referral letters in terms of quality and content.²¹ The lack of essential

information may reduce the quality, safety and cost-effectiveness of the healthcare system due to the scheduling of potentially erroneous work-up or waiting times, or even unnecessary/redundant procedures.^{3 22}

The use of electronic health record (EHR) systems and electronic referrals has increased substantially in the last decade,²³ and electronic referrals have improved referral quality.^{24 25} This shift to increased digitalisation of the healthcare services has also opened for new solutions to facilitate the referral process. Nevertheless, menu-driven structured report generation is neither available in the GPs' EHR nor in the hospital EHR. Initiatives to increase appropriateness of referrals have been implemented and tested with varying success.^{26–30} According to a Cochrane review, the implementation of structured referral paper templates has been one of few interventions with a documented effect on referral quality,²⁶ and electronic checklists have been shown to decrease the time spent evaluating referral letters.²⁷ However, the effect of interactive, electronic, dynamic checklists on the quality of referral letters has to our knowledge never been evaluated, neither in a clinical setting nor in a completely standardised trial, and it would be important to test the solution in a virtual setting before launching expensive development and implementation in the GPs' EHR system.

The aim of the present trial was to assess whether interactive, electronic, dynamic and diagnose-specific checklists improved referral quality and reduced the variation in the quality for referral letters in gastroenterology, one of the major specialities in internal medicine and with the majority of patients followed in the outpatient clinic. To avoid bias from clinical and organisational variation, we wanted to perform the trial in an entirely standardised setting by using vignettes.^{31 32}

We further wanted to assess whether the electronic dynamic user interface was well accepted by the GPs.

We hypothesised that referral letters generated with the use of interactive checklists contain more relevant information with less variation than free-text referral letters.

The primary end point of the trial was the quality of the referrals measured by a Thirty Point Score (TPS).³³ The secondary end points were the variance in the quality of the referrals, and the user satisfaction of the GPs.

METHODS

Study design

Between 30 April 2014 and 6 October 2014, we recruited GPs to participate in the trial, mainly in groups through already established mandatory educational groups for GPs in Norway. All MDs working in general practice in Norway were eligible for participation in the trial. GPs in the Asker and Bærum region (n=135), as well as some GPs in Oslo (n=9) and Bergen (n=4) were contacted directly and offered to participate in the trial. Additionally an email with information about the trial and invitation to participate was distributed through a national email-based

debate forum for GPs in Norway (Eyr). The GPs did not receive any material compensation for participating in the trial.

We designed the study as a two-armed cross-over trial, where we block-randomised participating GPs to refer eight virtual patients either with an electronic free-text referral or with a combination of the free-text referral together with electronic checklists as referral support. Dropouts after randomisation, before starting the trial (eg, those who did not show up on the agreed date for participation), were replaced by new GPs by continuing downwards on the randomisation key. The randomisation was done before the day of the participation by using permuted block randomisation with different block sizes, generated through the website www.randomization.com.

After a minimum of 3 months, the GPs referred the same eight virtual patients again with the alternative referral letter interface, for example, those who had referred with standard free text in the first round now referred with the checklist support. We chose the 3-month interval to avoid recollection of the checklist items by the GPs who had used these in the first round.

We instructed the GPs to create the referral letters the same way they normally do with a similar real patient, using the same structure, contents and time.

In the first round of the trial, we gathered groups of GPs (n=1–7) for participation together in the hospital computer room/other venue with computer/internet access. One investigator from the study team was present to give information technology support on how to get started, and also to facilitate the communication with the vignettes when necessary, for example, by suggesting alternative phrasing of questions to the vignettes when the simulator failed to give appropriate responses. This was necessary in the beginning as the simulator was sensitive to spelling and did not have a complete natural language. The GPs were quickly accustomed to the form of communication, and mostly required minimal support. In the second round, the GPs could choose whether they wanted to complete the trial in the same way, or if they wanted to do it from their offices or homes at a time of their convenience.

Intervention

Interactive Dynamic Referral Interface programme

We created an EHR simulator for generating referral letters combined with a virtual patient simulator, in cooperation with Microsoft Norway AS (figure 1). The user interface resembled common EHR systems for primary care, with a section displaying the patient's medical history, current medication, allergies and family history. It was possible to transfer this information directly to the referral letter by clicking an interactive button. The randomisation procedure determined whether a section for generating standard free-text referral letters or the semistructured referral Interactive Dynamic Referral Interface (IDRI) was activated in the user interface.

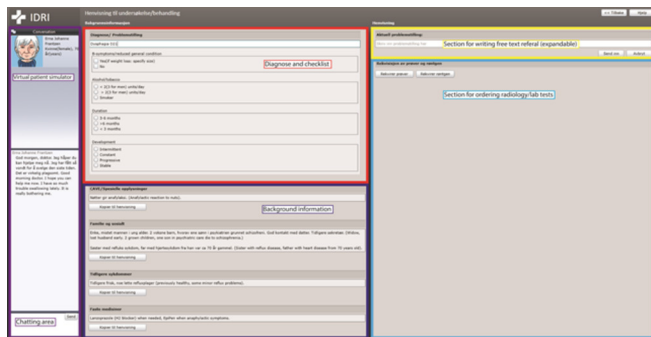


Figure 1 Interface of the virtual electronic health record/patient simulator.

The patient simulator, based on a chat functionality, was displayed on one side of the interface. An initial statement indicated the patient's reason for seeking medical care, for example, "Hi, I am really troubled by loose bowels lately, and it is getting worse. What do you think it could be?" Necessary information regarding relevant symptoms and findings could be obtained by chatting with the virtual patient. The GP could write questions to the patient in the dialogue box, either in whole sentences or using keywords, and the simulator would provide the patient's answer. In addition, it was possible to order relevant laboratory and radiological examinations. The results of the tests were displayed immediately on the screen and could be transferred to the referral letter. The information provided by the virtual patient during the 'consultation' could be written in the referral section using free text or registered by using the checklist function when activated.

The EHR simulator was set up according to the randomisation allocation before the day of the participation by the trial investigator who was also present at the first round of the trial.

The programme is available on the IDRI webpage: www.idri.no (username: IDRIopen, password open123).

Vignettes

We created eight vignettes (virtual patient cases) presenting symptoms and findings within the gastrointestinal (GI) field. The symptoms were chosen according to the main clinical situations in the Norwegian Prioritisation Guidelines for Gastroenterology (NPGg),³⁴ namely dyspepsia, change of bowel habit, diarrhoea, rectal bleeding, long-standing abdominal pain, constipation dysphagia and jaundice/elevated liver enzymes.

Subsequently, we integrated the eight vignettes in the virtual patient simulator with a unique set of answers to anamnestic questions and to laboratory and radiology tests that were made available through using the chat function or ordering the tests in the EHR simulator.

Interactive dynamic referral checklists/support

Sending a referral letter from primary care in Norway requires that at least one International Classification of Primary Care (ICPC)-2³⁵ diagnosis is stated in the

referral letter. We made a selection of relevant ICPC-2 (2005 version)³⁵ diagnoses for digestive diseases (selected from D01 to D99), omitting, for example, diagnoses for acute and paediatric diseases as well as non-GI-specific diseases and anal/oropharyngeal diseases. The selection was made based on which diagnoses could be relevant for the eight vignettes. We also added T08, weight loss, as a potential diagnosis.

When the IDRI functionality was turned on, the GPs' selection of ICPC-2 diagnosis activated the corresponding checklist. In total, we generated 10 checklists. The list of ICPC-2 codes used and the corresponding checklists can be found in the online supplementary appendix 1.

The checklists were made based on criteria for referral letters stated respectively in the NPGg,³⁴ the Norwegian Handbook for doctors³⁶ and UpToDate.³⁷ After creating an initial draft for the checklists, we adjusted and reduced the content of the checklists based on feedback from experienced gastroenterologists in the study team as well as clinical gastroenterologists.³³ The selection of the checklist items was done based on which clinical information items were considered most valuable for assessing and prioritising referral letters in gastroenterology. The checklists can be accessed through the IDRI web page, and a paper-based example can be seen in the online supplementary appendix 2.

The checklists consisted of drop-down menus with check boxes or free-text fields with symptom-specific or finding-specific questions where the GP had the possibility to choose the appropriate variable. Depending on the answers, new checklist items were activated if relevant. The use of checklists was not mandated to refer the virtual patient.

Primary outcome: quality of the referral letters

We assessed all referral letters generated in the trial by using a predeveloped score, the TPS, for objectively measuring quality of referral letters.³³ The TPS is a symptom-specific score that consists of the 15 most important variables for assessing and prioritising referrals for nine important GI symptoms (dyspepsia, change of bowel habit, diarrhoea, rectal bleeding, long-standing abdominal pain, constipation dysphagia, jaundice/elevated liver enzymes and weight loss). Points are assigned to the referrals depending on whether the variable/item is described in the referral or not. Both positive (eg, the patient has seen red blood in the stool) and negative (eg, the patient has not seen any blood in the stool) findings are assigned points if adequately described. The five most important variables are classified with 3 points, the next five with 2 points and the last five with 1 point, resulting in a maximum score value of 30 points.

One investigator from the study team scored all the referral letters in the present trial. The investigator was not blinded to the intervention, as the checklists with the GPs' answers were displayed together with the referral letters as a supplement to the information in the letters.

Secondary outcomes

As secondary outcomes, we also assessed the difference in variance in the quality of the referral letters and compared the frequencies of which important variables were included in referral letters generated with and without checklists.

To give an example of the frequency of individual variables in the referral letters, we did an a priori selection of some variables that we considered relevant for the referrals. We analysed six variables that were relevant for all the referral letters, four variables that were relevant for only the lower abdomen cases (diarrhoea, change of bowel habit, constipation, long-standing abdominal pain and rectal bleeding) and two variables that were relevant for only the upper abdomen cases (dyspepsia, dysphagia). For one variable, we only used the dyspepsia case as the dysphagia case would have included information of the presence of dysphagia by default in the referral letter.

After completing both rounds of the trial, we asked the GPs to complete a questionnaire where they provided information about age, size of their medical practice, years of experience as a doctor and their impression of the checklists in terms of usefulness and format.

STATISTICS

Power estimation

In our previous paper, we reported a mean TPS of 13.3 (SD 4.9) for standard referral letters in gastroenterology without the use of a computer-based checklist.³³ We expected an increase of 30% in the score when using the checklist, as well as a smaller variance of the score. The sample size calculation, comparing a mean TPS of 13.3 in the standard referrals with an expected mean of 17.3 in the referrals with checklist, with a two-sided type 1 error probability of 0.05 and a power of 0.80, yielded a minimal sample size of 21 referral pairs needed. In order to secure this quantity of referrals from each indication, we included as many GPs as possible from the local community. This resulted in 25 GPs completing the cross-over study, producing between 21 and 24 referrals per indication.

Statistical analysis

Descriptive variables are reported as means or proportions with 95% CI. We compared the mean overall TPS between referral letters with and without checklist using a multilevel linear regression model, adjusting for clinical case and the cluster GP. Paired t-test was used to compare the mean TPS stratified by clinical case. In a sensitivity analysis, we tested if the differences in the mean TPS of referral letters with or without checklist were influenced by whether the checklist was used in the first or second round (check for any learning effect). To do this we fitted a mixed linear regression model, adjusted for clinical case and added an interaction term between use of checklist and time of use of checklist (first or second round). We also performed a multivariable linear regression analysis

to assess whether the scores differed for age, gender and clinical case. To assess the variance of the TPS between referral letters with and without checklist, we calculated the variance and coefficient of variance (CV) ($= [SD/\text{mean}] \times 100\%$) for all referral letters with and without checklist and displayed box plots showing the median, IQR and minimum/maximum values by type of clinical case. We performed multivariable logistic regression models to compare the proportion of single variables in referral letters generated with or without checklist, adjusting for the clinical case and accounting for the cluster GP. A p value <0.05 was considered statistically significant. The statistical analyses were conducted using SPSS V.23.0 and STATA V.14 (StataCorp).

RESULTS

Study population

Between 30 April 2014 and 6 October 2014, 55 GPs were randomised and 45 GPs were included in the first period of the trial. Ten randomised GPs did not show up at the assigned date for participation and were therefore not included in the study. Of the 45 who attended, 25 (55.6%) participated in both rounds of the cross-over trial. The second round was completed between 3 December 2014 and 5 July 2015. The inclusion was ended because the targeted N for paired referrals was reached. One GP was excluded after the first round because he did not activate the ICPC-2 code in the checklist and consequently did not receive the intervention. Some GPs omitted the diagnosis in the individual referral letters or did not complete all of the eight referral letters in each round, resulting in 21–24 pairs of referral letters per clinical case (flow chart, online supplementary appendix 3). The participating GPs were, on average, 53 years old and more women (58%) than men participated (table 1). Sixty-four per cent (16/25) of the GPs who completed the trial had 20 years or more of experience as a doctor. Of the 25 GPs completing both rounds of the trial, 24 GPs answered the final questionnaire about the user friendliness of the system. In the counties Asker and Bærum in Norway, the average age of GPs is 50.8 years, with 47% female GPs.³⁸

Table 1 Characteristics of participating general practitioners

Characteristics	All participants (n=45)	Participants who completed both rounds (n=25)
Age (years), mean (range)	51.0 (31–72)	52.3 (33–63)
Female, % (n)	51.1 (23)	60.0 (15)
Checklist first round, % (n)	48.9 (22)	44.0 (11)
Time between rounds (days), mean (range)	–	181.8 (96–371)

Table 2 TPS for each patient vignette, comparing referral with or without checklist

Clinical case	Referral pairs (n)	With checklist (95% CI)	Without checklist (95% CI)	p Value*	Mean TPS difference
Dyspepsia	23	22.8 (20.9 to 24.7)	12.9 (11.2 to 14.5)	<0.001	10.0 (7.8–12.1)
Change of bowel habit	23	24.1 (22.6 to 25.6)	16.2 (13.4 to 18.9)	<0.001	7.9 (5.6–10.2)
Diarrhoea	24	21.9 (20.0 to 23.8)	15.4 (12.9 to 17.9)	<0.001	6.5 (3.8–9.1)
Rectal bleeding	24	25.3 (23.7 to 27.0)	17.7 (15.3 to 20.1)	<0.001	7.6 (5.5–10.7)
Abdominal pain	21	19.5 (17.1 to 21.9)	15.7 (12.5 to 18.8)	0.016	3.8 (0.8–6.8)
Constipation	21	18.8 (17.3 to 20.2)	13.3 (10.4 to 16.2)	<0.001	5.4 (2.7–8.1)
Dysphagia	22	22.5 (19.7 to 25.4)	13.8 (11.4 to 16.1)	<0.001	8.8 (5.8–11.8)
Jaundice/elevated liver enzymes	22	20.3 (17.6 to 22.9)	16.2 (13.3 to 19.2)	0.009	4.0 (1.1–7.8)
Total†	180	22.0 (20.6 to 23.4)	15.2 (13.2 to 17.2)	<0.001	6.8 (5.1–8.5)

*p Value is calculated using paired sample t-test.

†Results predicted from multilevel linear regression model adjusting for clinical case and general practitioner cluster.

TPS, Thirty Point Score.

Primary outcome: quality of referral letter by TPS

The mean TPS was higher in referral letters with checklist than without checklist overall (mean $\Delta=6.8$, 95% CI 5.1 to 8.5, $p<0.001$) and across all clinical cases (range of mean $\Delta=3.8$ to 10.0), but differed significantly for the different cases (global p value for clinical case <0.001 , from multivariable linear regression adjusting for time of checklist and GP cluster) (table 2 and figure 2). The smallest difference was observed in the abdominal

pain referral letters ($\Delta=3.8$ (0.8 to 6.8) and the biggest difference in the dyspepsia referral letters ($\Delta=10.0$ (7.8 to 12.1). Multivariable regression analysis did not show any influence of gender and age of the GP on the quality of the referral letters. In a sensitivity analysis, we tested whether the increase in score differed between GPs who had the checklist in the first or second round and found no difference ($p=0.303$ for interaction between use of checklist and time of checklist in a multivariable mixed linear regression model) (figure 3)

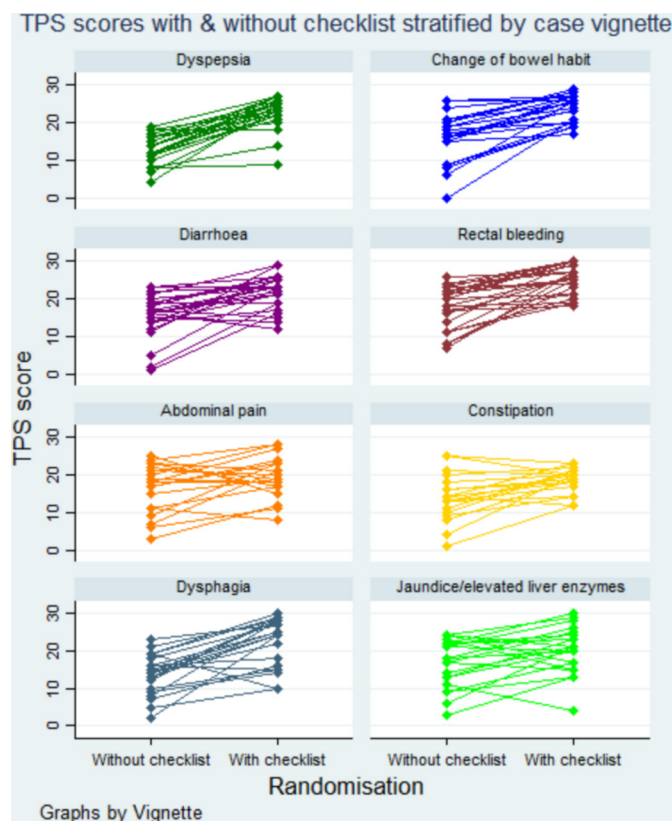


Figure 2 Thirty Point Score (TPS) for each case with and without checklist, stratified by clinical case.

Secondary outcome: variance in the quality of referral letters and frequency of specific variables

The variance in the checklist referral letters was 26.5, with a range from 4 to 30 points. The variance in the non-checklist referral letters was 36.2, with a range from 0 to 26 points. The CV was 23.3% for the checklist group and 39.6% for the non-checklist group. The box plots in figure 4 and line graphs in figure 2 graphically display the

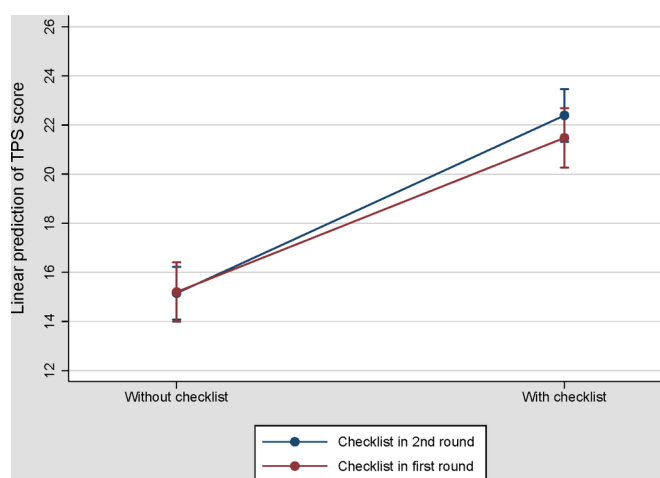


Figure 3 Predicted mean Thirty Point Score (TPS) (and 95% CI) of referral letters with and without checklist, stratified by time of the checklist.

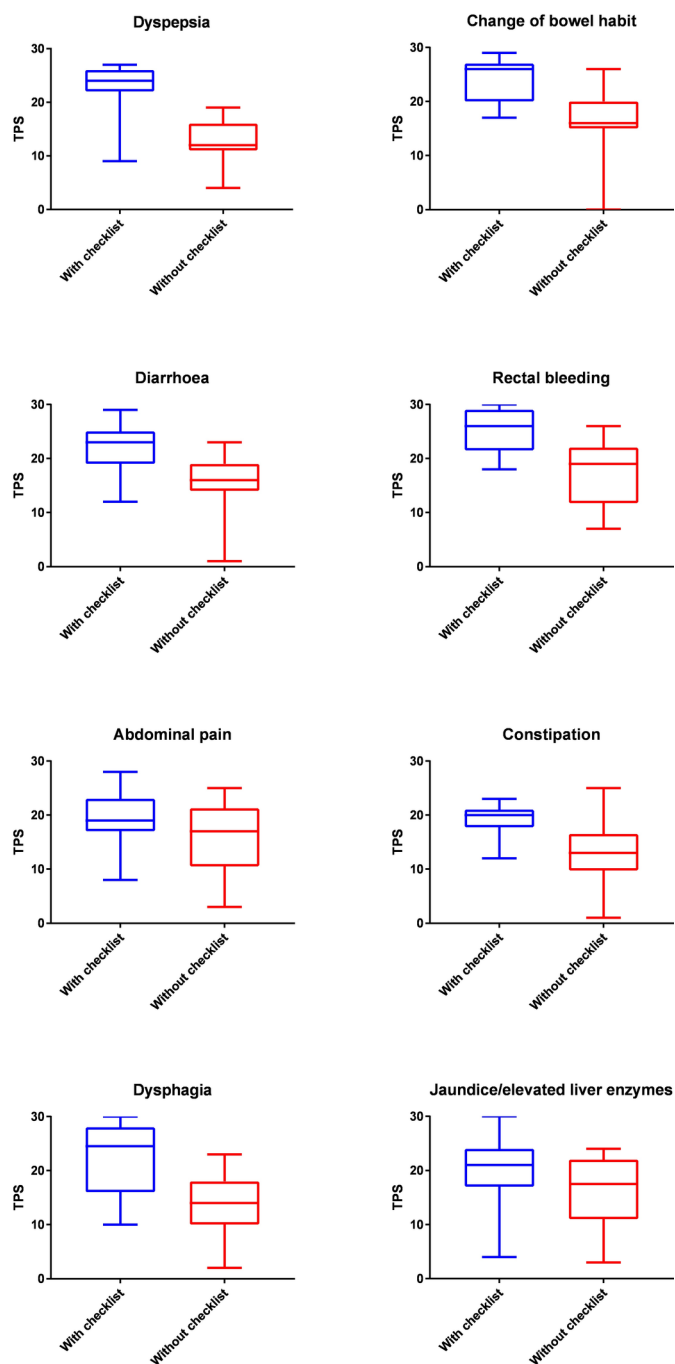


Figure 4 Box plot of the Thirty Point Score (TPS) stratified by intervention and clinical case.

larger TPS variance in referral letters without checklist; however, the pictures differ by type of clinical case.

When looking at essential clinical variables used to differentiate between serious and less serious conditions, the difference between the checklist referrals and the non-checklist referrals was also considerable. These variables are shown in [table 3](#).

Secondary outcome: acceptance of checklist among GPs

The GPs were generally positive on the checklists/referral support ([figure 5](#)). Two-thirds, 66.7% (16/24, 95% CI 44.7% to 84.4%), thought they had included

more relevant information in the referrals when using the checklists, and that they had included information in the referrals that they would not have included if they had not had access to the checklists. Seventy-six per cent (76.2%) (16/21, three missing, CI 52.8% to 91.8%) reported that they would consider implementing the checklists in their clinical practices if available and 95.5% (21/22, two missing, CI 77.2% to 99.9%) reported that the checklists had potential for improving the collaboration between the primary and the secondary healthcare services. The GPs largely found the size of the checklists to be appropriate, but 33.3% (8/24, CI 15.6% to 55.3%) found them to contain too many questions. A combination between checkboxes and drop-down menus were preferred by 47.8% (11/23, one missing, CI 26.8 to 69.4).

DISCUSSION

We have evaluated the effect of dynamic diagnose-specific checklists on the quality of referral letters in a standardised setting using vignettes.

Statement of principal findings

We found a significantly higher quality of referral letters with more important clinical variables included when they were written with access to checklists, compared with the ones written using standard free-text templates. The variance in the quality of the referral letters was smaller in the checklist referral letters than in the non-checklist referral letters. The majority of GPs found the checklists useful and would consider using them in clinical practice. The trial was set in Norway, using vignettes from the field of gastroenterology. However, we believe that the results are likely to be transferable to other medical specialties and applicable to other countries with similar referral systems.

Comparison with existing literature

To our knowledge, no other trials have rigorously evaluated the effect of electronic interactive checklists on referral letter quality in a standardised virtual setting.

When exploring relevant literature, the general tendency is that studies aiming to improve referral quality and appropriateness have been largely ineffective. The Cochrane review on such interventions published in 2008²⁶ identified only a few successful studies, but with the main focus on referral rates³⁹ and referral appropriateness.⁴⁰ They concluded that structured referral templates may improve referral practice, but none of the included studies had used electronic referral forms or checklists. Thus, these conclusions may not be valid for comparison with the present trial.

For clinical decision support systems, a review from 2005 stated four important features associated with a beneficial effect, including (1) Automatic provision of the support as a part of clinician workflow, (2) provision of recommendations rather than just assessments, (3) provision of decision support at the time and localisation of the decision-making and (4) being computer based.⁴¹

Table 3 Selected specific clinical variables in the referrals

Variable	Number of vignettes (number of referral pairs)	With checklist N (%)	Without checklist N (%)	OR (95% CI)*	p Value*
Duration	8 (180)	156 (86.7)	121 (67.2)	3.4 (1.7 to 6.8)	<0.001
Weight loss	8 (180)	156 (86.7)	93 (51.7)	6.6 (3.4 to 12.9)	<0.001
General condition	8 (180)	143 (79.4)	35 (19.4)	17.7 (6.5 to 48.1)	<0.001
Current medication	8 (180)	150 (83.3)	124 (69.3)	2.5 (1.2 to 5.2)	0.019
Findings on clinical examination	8 (180)	148 (82.2)	116 (64.4)	2.6 (1.4 to 4.8)	0.002
Medical history	8 (180)	168 (93.3)	159 (88.3)	1.9 (0.6 to 5.5)	0.262
DRE	5 (113)	73 (64.6)	44 (38.9)	3.1 (1.8 to 5.1)	<0.001
Rectal bleeding	5 (113)	78 (69.0)	73 (64.6)	1.3 (0.7 to 2.3)	0.422
FOBT	5 (113)	103 (91.2)	86 (76.1)	3.3 (1.3 to 8.4)	0.015
Hb, ferritin/MCV	5 (113)	105 (92.9)	79 (69.9)	5.7 (2.0 to 16.6)	0.001
Haematemesis	2 (45)	35 (77.8)	14 (31.1)	7.8 (3.3 to 18.3)	<0.001
Reflux details	2 (45)	33 (73.3)	17 (37.8)	10.6 (2.3 to 48.5)	0.001
Dysphagia	1 (23)	19 (82.6)	3 (13.0)	31.7 (5.3 to 189.5)	<0.001

*OR and p values calculated from multilevel logistic regression model, adjusting for clinical case and the cluster GP. DRE, digital rectal examination; FOBT, faecal occult blood test; Hb, haemoglobin; MCV, mean corpuscular volume.

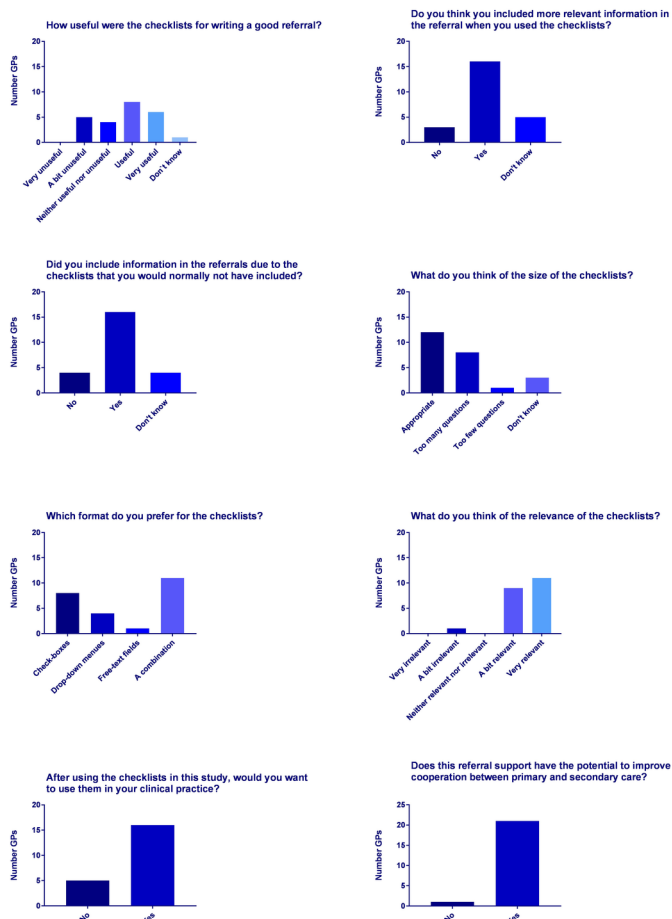


Figure 5 General practitioner (GP) opinions after completing the Interactive Dynamic Referral Interface (IDRI) trial.

These requirements are probably also valid for our intervention and care should be taken to integrate checklist solutions in the EHR referral systems to minimise any extra workload on the GPs receiving the support.

A few important later studies supporting the use of checklists/referral templates should be mentioned. Rokstad *et al*²⁷ made a direct implementation of checklists in existing EHR systems of GPs and found that this electronic optional guideline tool for referrals resulted in higher quality of referral letters and 34% less time spent by the specialist on evaluating each referral letter. Wahlberg *et al*²⁹ assessed the effect of paper referral templates distributed to local GP clinics, and observed an 18% improved quality of referral letters compared with a control group. These clinical studies support the findings of the present trial, and the evidence seems to indicate that checklists are in fact useful for both the referring and the receiving clinician.

In a study of GI endoscopy reports, structured electronic checklists increased the quality of the documentation, and concluded that this was likely due to the reminder effect of the checklist.⁴² It also concluded that a combination of free text and checklists seems to be the best way to document the procedure.⁴² These conclusions form the basis for the choice of referral design in the present study, also allowing for a combination of free text and the electronic checklist.

To measure the effect of a checklist on the quality of the referral letter can be challenging because implementation in clinical practice is subject to bias both from

patient case mix and from variations in the physicians' time, stress level, etc. Vignette studies are validated against standardised patients (gold standard) and medical record extraction as an appropriate way of studying quality and variation of physician practice,^{31 32 43 44} and have been used in various studies.^{45–53} A virtual solution using clinical vignettes was therefore chosen as the most appropriate way to standardise the setting.³¹

Strengths and limitations of the trial

The current trial has several strengths. First, we used vignettes to standardise the clinical setting, making the results objective, quantifiable and comparable in a way which would not have been possible using real patients. Additionally, the randomised cross-over design made comparison of individual GPs' changes in referral letter quality possible, regardless of the GPs' initial quality level for referral letters.

However, the study has some weaknesses that we would like to address. First, we used a score (TPS) for objective measurements of the quality of referral letters developed by our research team, which could potentially have influenced the results of the current study. This score was validated in another study,³³ and the results from this study show that the mean TPS for real referrals in gastroenterology is 13.7, thus somewhat lower than the TPS for free-text referrals in the present study. This most likely reflects a volunteer effect, where the participating GPs may have written referrals of higher quality than what they would have done under normal conditions in a clinical setting. However, the cross-over design makes us able to see past this potential bias.

Second, we did not achieve blinding of the TPS assessor due to the very apparent presence of checklist items in the referrals.

Third, the design of the vignettes may not have appeared realistic to the participating GPs, and also may have resulted in frustration due to challenges in the virtual communication. The alternative would have been to use standardised patients⁴³ or to evaluate the effect of checklists in a clinical setting. However, standardised patients would have been expensive and required much more resources, and real patients would have required implementing our system in existing EHR systems before actually knowing the effect of the intervention.

Lastly, the TPS does not measure all aspects of referral quality but rather quantifies the amount of relevant information covered in each referral letter. Thus, appropriateness and structure of the referral letter are not measured by the score.³³ However, the checklists did not aim to improve these aspects but rather the amount of relevant information in the referral letters, and the TPS is well suited for this purpose.

Implications of the study

The findings of the present study support the findings of Rokstad *et al*²⁷ and Wählberg *et al*,²⁹ and the evidence supporting the effect checklists on referral letter quality

now seems to be well documented. EHR providers should be encouraged to cooperate with both primary and secondary care specialists in developing and implementing appropriate electronic checklists and conduct randomised controlled trials to assess whether it also has an impact on patient outcome or healthcare costs.

The results from the GP survey indicate that a somewhat shorter checklist may be preferred to the one used in the present study.

Unanswered questions/future research

The present study have demonstrated an effect of electronic dynamic checklists on the quality of the referral letters, but the effect on the clinical management of the referred patients is still unknown.

CONCLUSION

Dynamic diagnose-specific checklists have a positive effect on the quality of referral letters in gastroenterology. The effect is most likely present also for other medical specialties, and corresponding checklists could be developed. GPs are largely positive to the idea of a checklist for referrals. The impact on the clinical management of the referrals remains unanswered.

Author affiliations

¹Department of Medical Research, Bærum Hospital, Vestre Viken Hospital Trust, Drammen, Norway

²Faculty of Medicine, University of Oslo, Oslo, Norway

³Department of Biostatistics, Oslo Centre for Biostatistics and Epidemiology, University of Oslo, Oslo, Norway

⁴Department of Transplantation Medicine, Section of GI Endoscopy, Division of Surgery, Inflammatory Medicine and Transplantation, Oslo University Hospital Rikshospitalet, Oslo, Norway

⁵Department of Medicine, Bærum Hospital, Vestre Viken Hospital Trust, Drammen, Norway

⁶Departement of Bowel Cancer Screening, Cancer Registry of Norway, Majorstuen, Oslo, Norway

Acknowledgements A special thanks to the participating GPs for their contribution to the project. We could not have done it without you. We would also like to thank the eHealth department in the Vestre Viken Hospital Trust for the continued support throughout the project, as well as Murad Sæther in Microsoft AS for the help in developing the EHR simulator.

Contributors SLE, TdL and LA designed the study. SLE performed the data collection. SLE, CB and CSR performed the power and data analysis. SLE drafted the paper. All authors critically reviewed and improved it. SLE is guarantor. All authors had access to all the data and take responsibility for the integrity of the data and the analysis.

Competing interests None declared.

Ethics approval The study was reported to and approved by the Data Protection Official for research. The Regional Ethics Committee considered the study outside its mandate, and its approval was not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Full data set and statistical code can be made available from the corresponding author. Participants' consent was not obtained in accordance with Norwegian Data Protection legislation, but the presented data are anonymised and risk of identification is low.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is

properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

- Ringberg U, Fleten N, Deraas TS, et al. High referral rates to secondary care by general practitioners in Norway are associated with GPs' gender and specialist qualifications in family medicine, a study of 4350 consultations. *BMC Health Serv Res* 2013;13:147.
- Barnett ML, Song Z, Landon BE. Trends in physician referrals in the United States, 1999-2009. *Arch Intern Med* 2012;172:163-70.
- Morgan DJ, Brownlee S, Leppin AL, et al. Setting a research agenda for medical overuse. *BMJ* 2015;351:h4534.
- Ong SP, Lim LT, Barnsley L, et al. General practitioners' referral letters--do they meet the expectations of gastroenterologists and rheumatologists? *Aust Fam Physician* 2006;35:920-2.
- Lønning KJ. *Tidsskrift for Den norske legeforening*. 2009;1868-9. Årg. 129, nr. 18.
- Mead GE, Cunningham AL, Faulkner S, et al. Can general practitioner referral letters for acute medical admissions be improved? *Health Bull* 1999;57:257-61.
- Jenkins RM. Quality of general practitioner referrals to outpatient departments: assessment by specialists and a general practitioner. *Br J Gen Pract* 1993;43:111-3.
- Berendsen AJ, Kuiken A, Benneker WH, et al. How do general practitioners and specialists value their mutual communication? A survey. *BMC Health Serv Res* 2009;9:143.
- Garåsén H, Johnsen R. The quality of communication about older patients between hospital physicians and general practitioners: a panel study assessment. *BMC Health Serv Res* 2007;7:133.
- Jack C, Hazel E, Bernatsky S. Something's missing here: a look at the quality of rheumatology referral letters. *Rheumatol Int* 2012;32:1083-5.
- Scully ND, Chu L, Siriwardena D, et al. The quality of optometrists' referral letters for glaucoma. *Ophthalmic Physiol Opt* 2009;29:26-31.
- Hedelin H, Johansson N, Strömberg P. (Uneven quality of referrals for lower urinary tract symptoms in men). *Lakartidningen* 2003;100:1435-7.
- DeAngelis AF, Chambers IG, Hall GM. The accuracy of medical history information in referral letters. *Aust Dent J* 2010;55:188-92.
- Ibiyemi O, Ibiyemi T. Quality and contents of referral letters from peripheral health centres to the dental centre of a teaching hospital, southwestern Nigeria. *Acta Odontol Scand* 2012;70:165-8.
- Thorsen O, Hartveit M, Baerheim A. The consultants' role in the referring process with general practitioners: partners or adjudicators? a qualitative study. *BMC Fam Pract* 2013;14:153.
- Jiwa M, Coleman M, McKinley RK. Measuring the quality of referral letters about patients with upper gastrointestinal symptoms. *Postgrad Med J* 2005;81:467-9.
- Cheng J, Beltran-Agullo L, Trope GE, et al. Assessment of the quality of glaucoma referral letters based on a survey of glaucoma specialists and a glaucoma guideline. *Ophthalmology* 2014;121:126-33.
- François J. Tool to assess the quality of consultation and referral request letters in family medicine. *Can Fam Physician* 2011;57:574-5.
- Westerman RF, Hull FM, Bezemer PD, et al. A study of communication between general practitioners and specialists. *Br J Gen Pract* 1990;40:445-9.
- Gulati S, Jakola AS, Solheim O, et al. Assessment of referrals to a multidisciplinary outpatient clinic for patients with back pain. *J Man Manip Ther* 2012;20:23-7.
- O'Malley AS, Reschovsky JD. Referral and consultation communication between primary care and specialist physicians: finding common ground. *Arch Intern Med* 2011;171:56-65.
- Vermeir P, Vandijck D, Degroote S, et al. Communication in healthcare: a narrative review of the literature and practical recommendations. *Int J Clin Pract* 2015;69:1257-67.
- Norwegian Health-Net. Message counter. Norwegian Health-Net. <http://meldingsteller.nhn.no/ConfigurableReport>.
- Walsh C, Siegler EL, Cheston E, et al. Provider-to-provider electronic communication in the era of meaningful use: a review of the evidence. *J Hosp Med* 2013;8:589-97.
- Nash E, Hespe C, Chalkley D. A retrospective audit of referral letter quality from general practice to an inner-city emergency department. *Emerg Med Australas* 2016;28:313-8.
- Akbari A, Mayhew A, Al-Alawi MA, et al. Interventions to improve outpatient referrals from primary care to secondary care. *Cochrane Database Syst Rev* 2008;4:CD005471.
- Rokstad IS, Rokstad KS, Holmen S, et al. Electronic optional guidelines as a tool to improve the process of referring patients to specialized care: an intervention study. *Scand J Prim Health Care* 2013;31:166-71.
- Akershus University Hospital HF. *End-report Project for Interactive Referrals*. Oslo: Akershus Universitetssykehus, 2011.
- Wählberg H, Valle PC, 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;15:353.
- Kim Y, Chen AH, Keith E, et al. Not perfect, but better: primary care providers' experiences with electronic referrals in a safety net health system. *J Gen Intern Med* 2009;24:614-9.
- Veloski J, Tai S, Evans AS, et al. Clinical vignette-based surveys: a tool for assessing physician practice variation. *Am J Med Qual* 2005;20:151-7.
- Peabody JW, Luck J, Glassman P, et al. Measuring the quality of physician practice by using clinical vignettes: a prospective validation study. *Ann Intern Med* 2004;141:771-80.
- Eskeland SL, Brunborg C, Seip B, et al. First quality score for referral letters in gastroenterology-a validation study. *BMJ Open* 2016;6:e012835.
- Nygaard E, Kårikstad V. *National Prioritization Guideline for gastroenterology*. Oslo: The Norwegian Directory of Health, 2009.
- World Organization of National Colleges A, Physicians AAoGPF. *ICPC-2-R: International Classification of Primary Care*. Oxford University Press, 2005.
- Norwegian Electronic Medical Handbook for doctors. Trondheim: Norsk Helseinformatikk AS. 2013. <http://legehandboka.no/>.
- UpToDate. UpToDate in Waltham, MA. 2013. <http://www.uptodate.com/home>.
- Helsenorge.no. Oversikt over fastleger. 2017. <https://helsenorge.no/min-helse/bytte-fastlege?fylke=02&kommuner=0220,0219>.
- Abu-Ramadan MA. Making better use of scarce resources: the palestinian experience, 1995-1999. *J Ambul Care Manage* 2002;25:63-9.
- Bennett K, Haggard M, Churchill R, et al. Improving referrals for glue ear from primary care: are multiple interventions better than one alone? *J Health Serv Res Policy* 2001;6:139-44.
- Kawamoto K, Houlihan CA, Balas EA, et al. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *BMJ* 2005;330:765.
- de Lange T, Moum BA, Tholfsen JK, et al. Standardization and quality of endoscopy text reports in ulcerative colitis. *Endoscopy* 2003;35:835-40.
- Peabody JW, Luck J, Glassman P, et al. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *JAMA* 2000;283:1715-22.
- Glassman PA, Luck J, O'Gara EM, et al. Using standardized patients to measure quality: evidence from the literature and a prospective study. *Jt Comm J Qual Improv* 2000;26:644-53.
- Tiemeier H, de Vries WJ, van het Loo M, et al. Guideline adherence rates and interprofessional variation in a vignette study of depression. *Qual Saf Health Care* 2002;11:214-8.
- Skånér Y, Bring J, Ullman B, et al. Heart failure diagnosis in primary health care: clinical characteristics of problematic patients. A clinical judgement analysis study. *BMC Fam Pract* 2003;4:12.
- Skånér Y, Backlund L, Montgomery H, et al. General practitioners' reasoning when considering the diagnosis heart failure: a think-aloud study. *BMC Fam Pract* 2005;6:4.
- Schwappach DL, Frank O, Davis RE. A vignette study to examine health care professionals' attitudes towards patient involvement in error prevention. *J Eval Clin Pract* 2013;19:n0.
- Quine MA, Bell GD, McCloy RF, et al. Appropriate use of upper gastrointestinal endoscopy--a prospective audit. Steering Group of the Upper Gastrointestinal Endoscopy Audit Committee. *Gut* 1994;35:1209-14.
- Pham T, Roy C, Mariette X, et al. Effect of response format for clinical vignettes on reporting quality of physician practice. *BMC Health Serv Res* 2009;9:128.
- Landon BE, Reschovsky J, Reed M, et al. Personal, organizational, and market level influences on physicians' practice patterns: results of a national survey of primary care physicians. *Med Care* 2001;39:889-905.
- Jiwa M, Gordon M, Arnet H, et al. Referring patients to specialists: a structured vignette survey of Australian and British GPs. *BMC Fam Pract* 2008;9:2.

53. Bachmann LM, Mühleisen A, Bock A, *et al.* Vignette studies of medical choice and judgement to study caregivers' medical

decision behaviour: systematic review. *BMC Med Res Methodol* 2008;8:50.