

Appropriateness and quality of referral letters in gastroenterology

Development of a quality score and a referral improvement tool

Sigrun Losada Eskeland

Department of Medical Research

Bærum Hospital

Vestre Viken Hospital Trust

Institute of Clinical Medicine

Faculty of Medicine

University of Oslo

Oslo 2017

© **Sigrun Losada Eskeland, 2017**

*Series of dissertations submitted to the
Faculty of Medicine, University of Oslo*

ISBN 978-82-8377-117-6

All rights reserved. No part of this publication may be
reproduced or transmitted, in any form or by any means, without permission.

Cover: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.

Contents

Acknowledgements	5
1. List of papers	7
2. Abbreviations and definitions	8
3. Summary.....	9
4. Introduction	11
4.1 The referral.....	13
4.1.1 The referral process	13
4.1.2 Electronic referrals.....	15
4.1.3 Referral letter templates	16
4.1.4 Checklists	17
4.1.5 Quality of referrals.....	18
4.1.6 Measuring referral quality.....	19
4.2 Appropriateness.....	19
4.2.1 Appropriateness of colonoscopy	20
4.3 Vignette studies	21
5. Aims of the Thesis.....	23
6. Material and Methods.....	24
6.1 Design and study populations	24
6.2 Measurement of main outcomes.....	26
6.3 Other outcomes and assessments	28
6.4 Vignettes and IDRI web-page	28
6.5 Intervention.....	30
6.6 Statistical methods and power analyses.....	31
6.7 Ethical considerations.....	32
7. Summary of the results	33
Paper I	33
Paper II.....	34
Paper III.....	35
Paper IV	36
Overall comparison of the results from the 4 papers.....	37
8. Discussion	38
8.1 Methodological considerations	38
8.1.1 Design and study populations	38

8.1.2 Measurement of main outcomes	42
8.1.3. Other outcomes and assessments.....	44
8.1.4 Vignettes and IDRI web-page	45
8.1.5 Intervention.....	47
8.2 Discussion of the main results	47
8.2.1 Appropriateness	47
8.2.2 Referral quality	50
8.2.3 The TPS	51
8.2.4 Interactive, dynamic checklists and referral quality	52
9. Conclusions	54
10. Future perspectives.....	55
11. References.....	56

Acknowledgements

This work was funded by the South-Eastern Norway Regional Health Authority research grant, the Norwegian Medical Association grant for quality and patient safety, and the South-Eastern Norway Health Authority Research Network for Clinical Effectiveness in Gastroenterology. The practical work was completed at the Department of Medical Research at Bærum Hospital, Vestre Viken Health Trust.

First of all, I would like to express my sincere gratitude to my main supervisor, **Thomas de Lange**. Your expert knowledge in the field of gastroenterology, quality and standardization has given me a solid foundation for my thesis work. Thank you for taking your time, not only to share your knowledge, but for taking me under your wing and opening doors wherever we have gone together. Your enthusiasm is infective, and your constant positivity and support has been priceless. Every time I have felt stuck or in doubt, you have easily pulled me back up and put me back on track. I could not have asked for a better supervisor, and I look forward to exiting new projects in the years to come.

A great thank you also goes to my co-supervisor **Lars Aabakken** for sharing your great knowledge and experience in the crossing between gastroenterology and standardization work. Your sharp eye for important details and always welcome constructive feedback has been of great importance for me.

I would like to thank **Arnljot Tveit** and all of my colleagues at the Department of Medical Research at Bærum Hospital. Arnljot, thank you for always making me feel like an important part of the department even when I have been an outsider with my research. **Mona Olufsen**, thank you for your big heart and your constant presence and support, you are the glue that keeps us together. To my friends, **Anne Pernille Ofstad** and **Sara Reinvik Ulimoen**: I was so lucky to share office with you during my “formative years” as a PhD student. You taught me how to survive the hard times as well as how to enjoy the good. I could not have asked for a better welcome or better company. To my co-PhD students at the department, **Håkon Ihle-Hansen**, **Marie Helene Ursin**, **Guri Hagberg**, **Ragnhild Munthe-Kaas**, **Trygve Berge**, **Anja Wiedswang Horjen** and **Katrine Enge** : thank you for sharing ups and downs during all these years. You are the best! Thank you also to the rest of the staff and researchers at the Department of Medical Research: **Steve Enger**, **Sophia Onarheim**, **Ingrid Elisabeth Christophersen**, **Hege-Ihle Hansen**, **Vigdis Bache-Semb**, **Hilde Marie Larhammer**,

Kristine Guldvog Haider, Kristi Elisabeth Heiberg and **Vigdis Bruun-Olsen**. You have all contributed through being great colleagues.

A special thank you to **Eirin Dalén**, who took part in the initiation of the IDRI project, and participated in creating the design and content of the IDRI-webpage as well as data collection, plotting and writing for paper I. You provided me with a flying start in the project.

I would also like to thank the eHealth department in Vestre Viken and **Jan Rune Nilsen** who helped finance the IDRI web-page in silence for so many years.

In paper II, my collaborators at the participating hospitals, **Birgitte Seip, Erik Skogestad, Øistein Hovde, Tanja Owen, Kristine Wiencke, Fred-Arne Halvorsen** and **Gert Huppertz Hauss** were essential for the realization of the study. I would also like to thank the auxiliary hospital staff who helped collecting paper copies of the referral letters.

To **Cathrine Brunborg** and **Corina Sylvia Rueegg**: thank you for your skilled contribution as statisticians in the project, and for always being available and positive.

To all of the GPs who participated in the IDRI-project. You were extremely patient and positive in spite of some challenges along the way. You made this project possible, and I am forever grateful.

Thank you to Microsoft and **Murad Sæter** who programmed the IDRI web-page, to **Bjørn Anton Graff** who provided with radiology standard-phrases, and all of the people who helped test the IDRI-chat before the trial-start.

To **Michael Bretthauer** and the South-Eastern Norway Health Authority Research Network for Clinical Effectiveness in Gastroenterology: Thank you so much for your participation in the trial and for the funding that made it possible for me to finish this thesis work.

And last, but not least: All my love and gratitude goes to my always loving and supportive husband **Israel** and our daughters **Gabriela** and **Valentina**. You are the ones who make everything else unimportant. Thank you for being the centre of my world and for making coming home from work the best part of my day. You inspire me to try to always be at my best.

1. List of papers

This thesis is based on the following papers:

- I. Eskeland SL, Dalén E, Sponheim J, Lind E, Brunborg C, de Lange T. **European panel on the appropriateness of gastrointestinal endoscopy II guidelines help in selecting and prioritizing patients referred to colonoscopy-a quality control study.** Scand J Gastroenterol. 2014;49(4):492-500.
- II. Eskeland SL, Brunborg C, Seip B, Wiencke K, Hovde O, Owen T, Skogestad E, Huppertz-Hauss G, Halvorsen F-A, Garborg K, Aabakken L, de Lange T. **First quality score for referral letters in gastroenterology- a validation study.** BMJ Open. 2016;6(e012835).
- III. Eskeland SL, Brunborg C, Ruegg CS, Aabakken L, de Lange T. **An Interactive Dynamic Referral Interface (IDRI) Improves quality of referral letters- a randomized cross-over vignette trial.** Resubmitted 5th of May 2017 after revision.
- IV. Eskeland SL, Ruegg CS, Brunborg C, Aabakken L, de Lange T. **Electronic checklists improve referral letters in gastroenterology- a randomized vignette survey.** Revision due 8th of July 2017.

2. Abbreviations and definitions

ANOVA	Analysis of Variance
ASGE	American Society of Gastroenterology
CDSS	Clinical decision support system
CI	Confidence interval
CRC	Colorectal cancer
CRF	Case report form
DY	Diagnostic yield
EPAGE	European panel on the appropriateness of gastrointestinal endoscopy
EHR	Electronic health record
ELIN	Electronic information exchange
GDP	Gross domestic product
GP	General Practitioner
HSØ	South Eastern Norway Health Region
IBD	Inflammatory bowel disease
IBS	Irritable bowel syndrome
ICD	International classification of disease
ICPC	International classification of primary care
IDRI	Interactive dynamic referral interface
KITH	Competence centre for information technology in healthcare
NGF	Norwegian society of gastroenterology
NHN	Norwegian Health-Net
NPGg	National Prioritization Guideline for gastroenterology
NRS	Numeric rating scale
OR	Odds ratio
PCP	Primary care provider
REK	Regional ethics committee
SD	Standard deviation
TPS	Thirty point score
UC	Ulcerative colitis
VAS	Visual analogue scale
WHO	World Health Organization

3. Summary

The referral of patients from primary to secondary care is a critical point in the chain of care. The referral letter constitutes a key document in which the referring physician can inform the hospital specialist of all relevant information regarding a referred patient.

When information is lost in the transfer between health care levels and institutions, e.g. due to poor quality of the referral letter, it challenges patient safety. Patients with serious disease may not be perceived as such by the consultant due to a lack of clinical details in the letter. Conversely, patients with trivial disease may be scheduled for inappropriate and potentially harmful invasive procedures. Additionally, the specialists are faced with the challenge of prioritizing low-quality referrals, which is frustrating, can lead to less confident decisions and more time spent on the task.

In **paper I** of this thesis we show that 7.1% of patients referred and accepted for colonoscopies have inappropriate indications for performing the procedure, with an almost non-existent diagnostic yield. This puts a great burden on the health care system both in terms of patient safety, capacity issues and cost.

When aiming to evaluate and improve the quality of referral letters, it is imperative to measure the effect of the interventions. We therefore developed and validated a thirty point score (TPS) to improve the assessment of the quality of referral letters in gastroenterology (**paper II**). The validation study was conducted at referral centres for gastroenterology services. It showed that the average quality of the referral letters was 13.2 (95% CI 12.8-13.8) measured by the TPS, and 4.7 (95% CI 4.5-5.0) measured by a 10 cm visual analogue scale (VAS). Both means are below 50% of the maximum score, which indicates a need to improve the quality of the referral letters.

In order to improve the quality issues demonstrated in **paper II** and to facilitate the transfer of information from the GP to the specialist, we developed an electronic checklist-tool to aid the GP in creating high quality referral letters (**paper III**). In a randomized vignette cross-over trial, we demonstrated that this tool produced a significant increase in the TPS of the referral letters (mean $\Delta=6.8$, 95%CI 5.1 to 8.5, $p<0.001$), with more information items recorded in each referral letter. In **paper IV** we evaluated whether this increased referral quality also affected the gastroenterologists' opinions of the referral quality and the scheduling of the patients. The checklist-referrals were considered to be of slightly higher quality than the non-checklist referrals, but no effect on the clinical management of the patients was seen.

In conclusion, ensuring appropriate use of the health care services is imperative. Currently, both appropriateness and quality of referral letters are suboptimal. Electronic dynamic checklists seem to improve the quality of referral letters both according to a quality score and according to the specialists. The effect on the prioritization of the patients seems to be limited.

The studies in this thesis were conducted in a setting of referrals to gastroenterology services. However, the general principles and ideas are applicable to other specialties, and the findings are most likely transferable to other countries with similar referral systems.

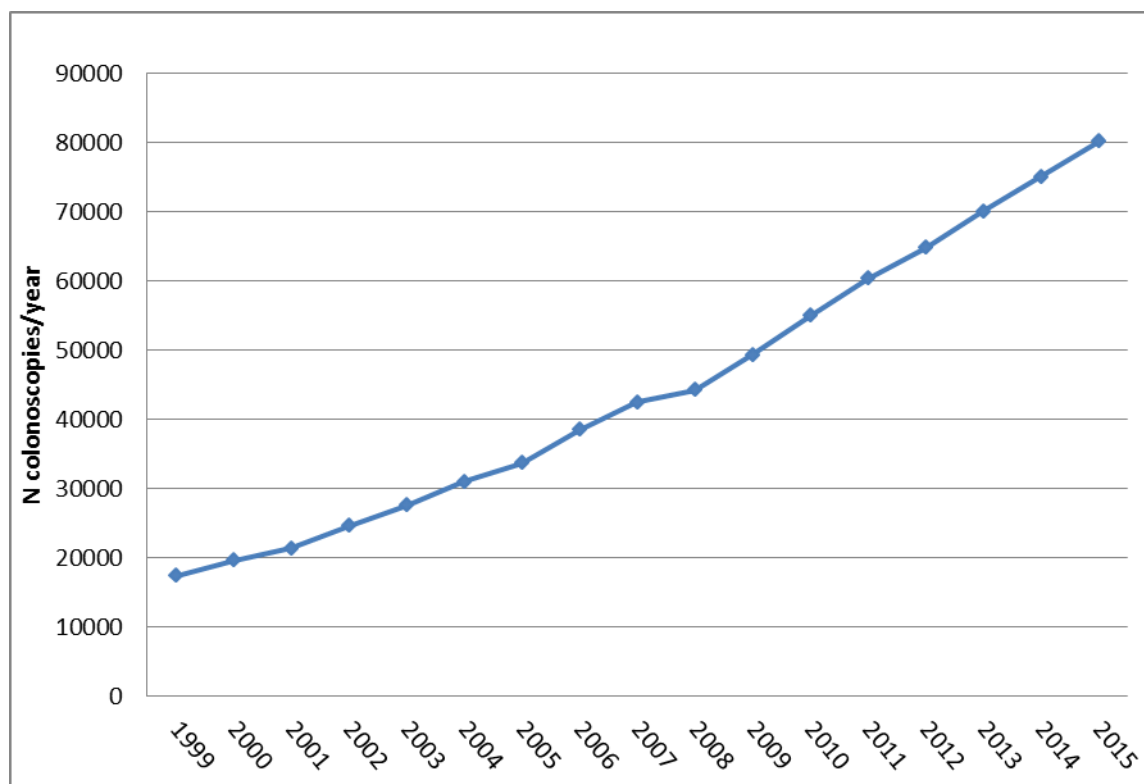
4. Introduction

The transfer of care between health care providers is a critical point where information may get lost and patient safety may be at risk. Poor communication may lead to disruptions in the continuity of care, delayed diagnoses, unnecessary testing, and iatrogenic complications (1-3) as well as increased cost for the health care system (4). The referral letter is a key document when patients are transferred from primary care to secondary care, or between secondary care institutions. It is used to convey the information needed for the consultant specialist to make a correct assessment of the patients' need for secondary care treatment or work-up, and to schedule the correct examinations/procedures with the right urgency and correct maximal waiting time.

A lack of detailed information regarding the patient's symptoms and findings makes the task of prioritizing the patients for the outpatient clinics particularly difficult, as the seriousness of the disease may be unclear to the specialist (5, 6). This may be frustrating, and can make the referral assessment more challenging (5). It may involve making decisions with reduced confidence, and also often involves extra time spent on the assessment process, including phone calls/letters to referring Primary Care Providers (PCPs) to get more information about the patient in question. Whether it can also lead to incorrect decisions with consequences for the patients' outcome has to a lesser extent been demonstrated (7, 8). However, there is some evidence indicating that referral letters that include diagnostic clues of colorectal cancer reduce the risk of delayed diagnosis (9, 10), and that inadequate diagnostic tests increase the delay (11). Only few studies report satisfactory referral quality (12), but it is unknown whether this may be due to publication bias.

The total health care expenditure has risen in Norway the last 15 years, comprising 7.7% of the Gross Domestic Product (GDP) in 2000 and 9.7% in 2015 (13). Within the field of gastroenterology, there has been an increase from approximately 31000 colonoscopies/year in 2004 to approximately 80 000 in 2015 (figure 1) (14).

Figure 1: Number of colonoscopies performed in Norway 1999-2015 (14).



In this setting of increased demands for treatments, examinations and consultations, identifying solutions that can help increase the appropriate use of the health care system is imperative. It is important for maintaining a sustainable health care system, but much more importantly it is crucial to ensure the safety of the patients, and to ultimately reduce both the burden of disease, as well as the overuse of costly examinations with potential complications (15, 16).

Quality in healthcare is a subjective feature that is not easily defined (17). It is often described by using quality indicators (18, 19) or criteria like outcome (20), but these measurements are tools to support quality improvement, not necessarily direct measures of quality (21). In a systematic review of performance measures in the referral process, Guevara et al identified factors associated with quality of the referral, including appropriateness, effectiveness, efficiency, equity, patient-centeredness, referral satisfaction, safety and timeliness of the referral (22). They do not, however, define criteria for the quality of the referral letter.

4.1 The referral

4.1.1 The referral process

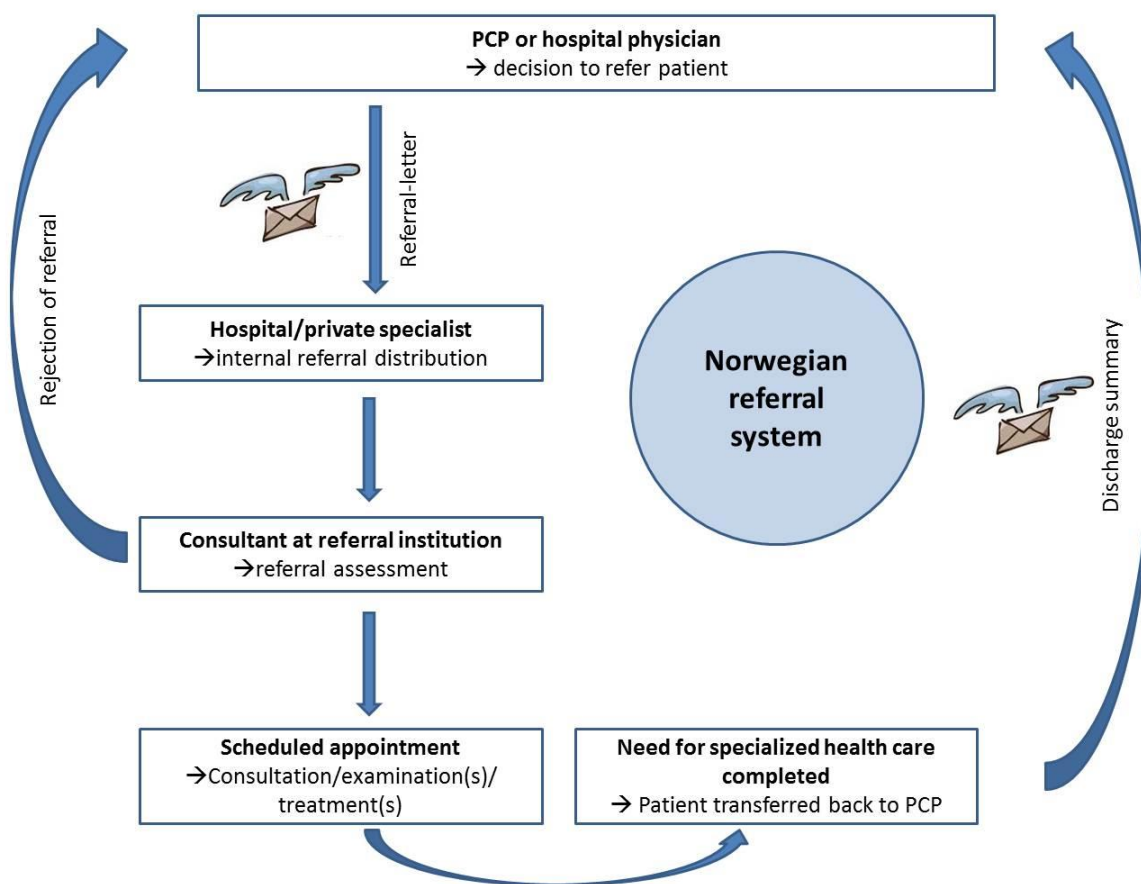
In most countries, the health care system is divided in primary- , secondary (specialist) - and tertiary (regional/national specialist services) - health care services. In Norway, as in many other European countries, access to secondary/tertiary health care is controlled by general practitioners (GPs) in a gatekeeper system (23, 24). When the need for treatment in the specialist health care services arises within a gatekeeper system, the GP can initiate the transfer of care by referring the patient. The referral can thus be defined as a formal request that another health care authority (e.g. specialist, outpatient clinic or hospital) fully or partly take over the responsibility for a patient (25, 26), either to provide specialized examinations or treatments, give advice on diagnosis or management, to request admission, or to obtain a second opinion. Referring a patient for non-emergent causes is usually done by sending a referral letter to the appropriate health care authority.

The referral letter serves several important purposes; 1) Signal the need for transfer of care from one level/institution to another, either for admission or outpatient services, 2) transfer information about- or trigger the patients' legal rights in terms of priority and acceptable waiting time, 3) transfer medical information about the patient to other health care providers, 4) form the basis for the specialist's assessment of the patients need for health care, and consequently also the scheduling of appropriate examinations, prioritization on the waiting lists and deciding the level of care (26-28). If the referral letter does not contain all necessary information to prioritize the patient it cannot be rejected or returned, and the specialist is responsible for collecting the missing information in order to make a correct referral-assessment (27).

GP referral rates are increasing both nationally (29) and internationally (30-32), and challenges the capacity of outpatient clinics. In Norway, the total number of referrals is estimated to be around 1.9 million each year, and each year Norwegian doctors spend 150 man-years on administrative work related to referrals (26, 33, 34). The referral is usually sent from the GP to a central referral reception (most common organization in the Norwegian health care system), and is distributed to the correct specialist-groups internally by hospital auxiliary staff. At this point, the specialists are presented with the referral letter and assess the need for health care, set the maximum waiting time and schedule the outpatient/inpatient visit and decide the work-up/procedures to be performed. When the patient no longer has a need for specialized health care, he/she is

transferred back to the care of the GP through a discharge letter informing the GP of the examinations/treatments received and the expected follow-up by the GP or the secondary health care system.

Figure 2: The referral and discharge system in Norway.



Both medical and non-medical reasons may be involved when a GP refers a patient (35-37). These include 1) establishing a diagnosis/providing treatment, 2) receiving advice and 3) reassuring the patient (36). The decision to refer may consist of several and complex considerations, such as expected medical benefit, costs and patient preferences (38, 39). All these aspects should be kept in mind when studying referral quality and appropriateness.

Several studies have documented differences in referral rates between GPs (29, 37, 40, 41), and some have associated this aspect with the frequency of inappropriate/unnecessary referral (42). Other studies do not find this association (43). While some patients are inappropriately referred, there may be other patients who are not referred who may have benefited from specialty care

(44). Ensuring a rational use of the referral system and the consumption of secondary health care services is an important task of the GP in the gatekeeper system. Thus, to develop and deliver tools to improve their ability to perform this task satisfactorily should be a priority.

4.1.2 Electronic referrals

The implementation of electronic referrals in Norway was initiated in 1996 when the first standard for electronic referrals was developed, and national projects like the ELectionic INformation exchange (ELIN)-project further developed these first initiatives (45). However, only 8200 referrals were sent electronically in September 2007. The same year, 125 000 electronic discharge summaries were generated (33). Now this scenario has largely changed, and in 2015 more than 1.5 million (79%) referrals from primary to secondary care were sent electronically (46), an increase of 52% from the year before. The implementation of inter-hospital referrals has been a slower process, but in 2016 a solution for electronic transfer of referrals between hospitals in the South Eastern Norway Health Region (HSØ) was piloted and will be further developed with the aim of implementation in the whole Norwegian health care system (47). The electronic referrals are transferred via the Norwegian Health-Net (NHN) for electronic messaging with an ebXML framework (48). Automatic application receipts are sent for all messages, which makes the transfer of the referral letters safe and reliable compared with paper based transfer (33). Internationally, e-Referrals are also implemented to an increasing extent, but few projects have published evaluations of their systems. Examples of local or national e-Referral systems are found in Finland (49, 50), Denmark (Medcom) (51), UK (Choose and book) (52), Scotland (SCI gateway) (53) and Netherland (Zorg Domain) (52), Australia and New Zealand (Hutt Valley). The approaches vary from country to country (52, 54). Recommendations for the design, development, implementation, improvement and monitoring of electronic outpatient referral communication were presented by Esquivel et al in 2012, defining 10 important features to consider (55). These include that 1) standardized electronic referral templates should include both structured and free-text fields, and 2) the systems should use automation to pre-populate electronic referral requests with patient-specific data (55).

Several studies show a beneficial effect of electronic referrals on referral quality (49, 56-58), as well as the safety and speed of the transfer of the referral letter (59-61). eReferrals has the potential of improving the cost-benefit of the health care system according to a Danish report (62). The electronic population of the referral letter with information from the Electronic Health

Record (EHR) system is a feature appreciated by GPs due to more efficient work processes (53). This feature is also implemented in Norwegian primary care EHR systems, where headings from the Norwegian Guideline for Referrals and previously stored patient- information like previous medical history, medications etc. are automatically added from the EHR.

The electronic format holds great potential for innovative solutions. However, the implementation of electronic tools are costly and very resource demanding. The processes from decision making through development to implementation are slow. Since implementation of Information and Communication Technology (ICT) systems in Norwegian healthcare is publically funded, decision makers have a special obligation to base such implementation on a solid evidence base, funded on research. Testing the effect of such interventions in a limited and less resource demanding environment is thus important. Parallels can be drawn to the early phases in drug development trials. Bypassing these phases in drug development would be considered highly unethical, and maybe ICT in healthcare have something to learn from clinical research.

4.1.3 Referral letter templates

In Norway, a template for the content and structure of electronic referrals was created by the Competence Centre for Information Technology in Healthcare (KITH) in 2003 (26), and in 2015 it was further modified and developed to a national guideline for referrals to the specialist health care services (28).

This national guideline provides an overall structure for the referral letter (table 1), and indicates sub-categories of information required in the referral letter, such as alcohol/tobacco consumption under personal information and information about allergies and contagious diseases under critical information. The guideline also opens for diagnose-/specialty- specific recommendations, but by 2017 only a few guidelines for more specific content of referrals have been developed, namely for referrals to mental health care, to specialist in kidney diseases, and for referrals within the standard pathways for cancer.

Table 1: Information required in the referral according to the Norwegian national referral guidelines

Recipient of referral
Referring physician
Copy recipients
Patient information
Type of referral/expected action
Urgency of referral
Referral diagnosis
Critical information
History of the current disease
Previous relevant medical history
Findings and results from examinations
Personal information
Medicines
Patient informed of referral

Other international templates for the requirements for the information in the referral letter have been difficult to identify, possibly due to lack of available publications of national official documents. However, many publications reporting on referral quality describe general features similar to the Norwegian template (63-71).

Some studies have suggested guidelines for the specific content of referral letters within different medical specialties (72-75) and some have shown a beneficial effect of referral templates/form letters on the quality of referral letters (76-80). Structured templates have also been effective for improving documentation in other medical records, particularly in reports from examinations/procedures (81-84). Thus, the evidence indicates that some degree of structure/form is warranted when the objective is to improve documentation quality.

4.1.4 Checklists

A checklist is a type of informational job aid used to reduce failure by compensating for the potential limitations of human memory or attention. They have been implemented in different areas of medicine to improve the quality and safety of the healthcare (85). The main driver for this trend has been in the surgical area, where checklists have been implemented in the operating theatre, and have dramatically improved patient safety (86).

Research in the field of referral quality has also gradually been directed towards the use of checklists. Initially, research focused mainly on structured referral letters, e.g. form letters / templates (76). At the start of the millennium referral letters were mainly paper based. Therefore, interventions to improve referrals were mainly also paper-based. The evidence regarding the effect of paper-based referral guidelines or templates indicated a beneficial effect on referral quality, but checklists were not frequently used (79, 87, 88). The interest in electronic checklists has increased the recent years following implementation of EHR. EHR providers for primary care have signalled ambitions to implement checklists in existing systems, but in Norway only few local projects have implemented electronic checklists in clinical practice (72, 89). Internationally, the evidence is also limited, but indicates that structured referral sheets could improve the quality of referral letters (88).

Many of the same quality issues seen in referral letters are also seen in other EHR documents (90). In discharge summaries the use of standardized formats to highlight the most important information has been shown to improve the perceived quality of this document (90). Endoscopy reports is another example showing a beneficial effect of a structured electronic format (81).

The similarities between a checklist and a clinical decision support system (CDSS) are many. They can both provide on-site reminders of requirements for specific clinical situations, they aim to improve clinical practice and patient outcome and they can be paper- or computer- based. There are numerous studies investigating the effect of CDSS, and a review from as early as 2005 stated that the key features associated with a beneficial effect of a CDSS were: 1) automatic provision of decision support as part of clinician workflow, 2) provision of recommendations rather than just assessments, 3) provision of decision support at the time and location of decision making, and 4) that the support is computer based (91). These requirements may also be valid for electronic checklists.

4.1.5 Quality of referrals

The quality of a referral can be understood as the quality of the indication/timing of the referral, e.g. the appropriateness of the referral, but can also be understood as the quality of the written referral letter. These aspects are closely related, but also have some major differences. A patient can be referred for an appropriate indication through a referral letter of poor quality, or for an inappropriate indication through a referral letter containing all necessary information to make a good assessment of the need for health care. Consequently, the appropriateness of the referral and the quality of the referral letter cannot be treated as the same entity.

The quality of the referral has also been related to the referral rate (8, 88, 92-94), but most studies find that there are no associations between the referral rate and the quality or appropriateness of the referral (36, 43, 95). To judge the referral quality on the basis of referral rates is thus a very simplified way of looking at referral quality, and should not be used (95).

Several studies have aimed to assess the quality of referrals from primary to secondary care, but the studies have been performed with a variable methodologies and qualities. They do, however, largely paint the same picture: both nationally (5, 63, 65, 96-98) and internationally (66-71, 74, 99-110) studies show unsatisfactory quality of referrals and referral letters. Information may be missing, both on a general level, e.g. information about current medications (63, 65, 69, 98, 109, 110), social history (67), and description of the current disease (63, 65, 74, 96, 103, 109), but also on a more detailed and disease specific level, e.g. information regarding the presence of dysphagia or hematemesis in dyspepsia referrals (74) or sensory symptoms in referrals to a spinal unit (96). Additionally, inappropriate pre-referral work-up and examinations have been reported (63, 67, 69, 70, 77, 93, 111).

Other referral-quality issues are excessive and unstructured referral information resulting in essential information getting lost in the “noise” from unnecessary text. Addressing this issue is outside the scope of this thesis, but is nonetheless an important factor to consider when looking at interventions to improve the quality of the written referrals.

I thus infer that a high quality referral consists of two main aspects: 1) An appropriate indication and timing of the referral and 2) a high quality referral letter. A high quality referral letter implicates a well-structured, concise communication of the patient’s need for health care, including all necessary content items that are relevant for the assessing the condition and scheduling treatment in the secondary or tertiary health care services.

4.1.6 Measuring referral quality

Many previous studies have documented referral quality, with different approaches ranging from the use of VAS (65), Likert scales (112), other scales (113), information item counts (57, 63, 66, 67, 96, 100, 101, 103, 105, 106, 109, 114-116), and more or less complex scores (64, 71, 75, 76, 102, 117-119). However, the lack of objective, validated tools was evident. Consequently, assessing the effect of interventions to improve referral quality, as well as comparing different studies and approaches, was challenging due to the lack of comparable endpoints. An objective, reliable and validated score would be a useful tool in this setting.

As far as we were able to discover, no score to assess referral quality in gastroenterology existed. Only one study stood out as a plausible alternative. Jiwa et al (75) aimed at creating a score to evaluate the quality of referrals to colorectal surgeons. However the main focus of this score was on the presence of alarm symptoms for colorectal cancer in the referral letters. In order to evaluate referral letters regardless of the reason for referral, there was a need for a score that also included items for assessing the presence of non-neoplastic diseases and upper GI/liver-related symptoms.

4.2 Appropriateness

Appropriateness of health care is important when working within any medical specialty. A procedure is deemed appropriate when the expected benefits of the procedure outweigh the theoretical risk or inconvenience by an acceptable margin of safety (120, 121). The appropriateness is determined by the combination of the indication and the timing, e.g. of a procedure or a referral. The most acknowledged method to define appropriate care is the RAND-

University of California at Los Angeles (UCLA) Delphi panel method (122-124). This method consists of developing multiple separate indications or case scenarios for a given procedure, reflecting any combination of scenarios that can affect the benefit and risk associated with the procedure. Following this, a multidisciplinary panel of experts rate the scenarios on a nine-point appropriateness scale, and re-rate them after panel discussions about areas of disagreement (122). Based on these ratings, the different clinical scenarios are judged appropriate, uncertain or inappropriate.

4.2.1 Appropriateness of colonoscopy

Studies have shown that a high percentage of referrals for gastrointestinal endoscopies are inappropriate (120, 125-134), meaning that the endoscopy should not be performed. Nonetheless, numerous inappropriate procedures are carried out every year. Tools to aid the clinician in prioritizing patients referred to gastroenterology services exist (27, 135), but they are less specific regarding which groups of patients may not benefit from undergoing colonoscopy.

Both the American Society of Gastroenterology (ASGE) (136) and the European Panel on the Appropriateness of Gastrointestinal Endoscopy

(EPAGE) (121) have developed guidelines to determine the appropriateness of gastrointestinal endoscopy.

The first EPAGE guidelines were developed in 1999 by a panel of international experts within family medicine, gastrointestinal surgery and gastrointestinal medicine in 1999, using the UCLA-RAND appropriateness method (137, 138). Indications for colonoscopy were rated on a scale from 1 to 9, where 7-9 were defined as appropriate indications, 4-6 were defined as uncertain and 1-3 were defined as inappropriate indications.

In 2009, the guidelines were further refined based on a comprehensive literature review and the RAND-UCLA method, and presented as the current EPAGE II guidelines (121, 139-143). This new version aims to assess the appropriateness of colonoscopy for 11 main indications (table 2) and

Table 2: Main indications for colonoscopy according to the EPAGE II guidelines.

1. Iron deficiency anaemia
2. Hematochezia
3. Discomfort or pain in the lower abdomen persisting ≥ 3 months
4. Uncomplicated chronic diarrhoea
5. Assessment of ulcerative colitis (UC)
6. Assessment of Crohns disease
7. Colorectal cancer (CRC) screening
8. CRC screening in patients with inflammatory bowel disease (IBD)
9. Surveillance colonoscopy after polypectomy
10. Surveillance colonoscopy after CRC resection
11. Miscellaneous

309 clinical scenarios.

The guidelines are available as a collection of articles presenting the rationale and conclusions for the different indications, as well as in a web-based and dynamic interface (144) where they are easily accessible for both referring physicians and for physicians prioritizing referrals for gastrointestinal endoscopies.

The EPAGE II guidelines have been evaluated in three Spanish studies for assessing appropriateness of colonoscopy (125, 128, 131), and have showed an improved sensitivity for correctly classifying significant lesions, at the expense of reduced specificity.

4.3 Vignette studies

A clinical vignette is a brief case history of a fictional patient based on a realistic clinical situation and is typically accompanied by one or more questions or tasks that explore what a physician would do if presented with the actual patient (145). They can be short and static or sequential, providing increasing amounts of information (145). Vignette studies are not to be used for testing the knowledge of individual physicians, but rather to assess clinical every-day practice for groups of physicians. Open-ended answers in the physicians' own words are preferred over closed, multiple choice replies, because closed-ended questions can lead to an over-estimation of the physicians' performance (145, 146). Clinical vignettes have been used to study quality of clinical care, documentation quality and to assess variations in clinical practice in several studies (146-149).

Assessing clinical practice and quality of care is complicated, and is not easily measured in clinical studies due to variation in patient case mix and from variations in the physicians' time, stress level and other factors that can make results difficult to interpret. The gold standard for studying quality and variation of physician practice is the use of standardized patients (150-153). A standardized patient is an actor who presents unannounced at the clinician's office and act as a patient with a strictly controlled pre-defined behaviour and answers to anamnestic questions (154). However, obtaining properly trained standardized patients has been described as expensive, intrusive and potentially unethical unless informed consent from participating physicians can be obtained (145). Vignette studies have been validated against standardized patients and medical record extraction as an appropriate way of studying quality and variation of physician practice (145, 148, 153), and have been used in various studies (146, 147, 149, 155-161). There are a few important prerequisites for designing a successful vignette study. Veloski et al

(145) name 3 important factors: 1) unambiguous instructions (information that the aim is to study clinical practice, rather than knowledge and skills), 2) realistic clinical situations and 3) a distinctive strategy for data analysis that involves adjustment for correlated data (160). Failing to adjust for correlated data can lead to underestimation of the SD and consequently increase the risk of type 2 error.

Vignette studies offer the possibility of evaluating the effect of interventions on physicians' practice at a relatively low cost compared to implementation in clinical practice. Changes in existing ICT systems are particularly costly. A first step in validating the impact of new ICT solutions should therefore include non-clinical evaluations.

5. Aims of the Thesis

The overall aim of the thesis was to use the referral letter to assess some steps of the referral process from primary to secondary care, with focus on the referral appropriateness and the quality of the referral letter. We wanted to find ways to improve these aspects of the referral, and increase the ability of the hospital consultants to select appropriate examinations for the right patients. To achieve this, we wanted to explore electronic, dynamic checklists as a tool to increase the amount of important information that is conveyed in the referral letter and whether such checklists would facilitate the task of assessing the referral letter for the hospital consultant.

Specific research objectives for each paper were:

Paper I: To use referral letters to assess the proportion of appropriate indications in patients accepted for colonoscopy according to the EPAGE and the EPAGE II guidelines in a Scandinavian setting, and whether appropriate indications resulted in higher diagnostic yield.

Paper II: To develop and validate an objective, relevant and reliable score for evaluation of the quality of referral letters. We also wanted to assess the current quality of referral letters to gastroenterologists in Norway.

Paper III: To develop interactive, electronic, dynamic, diagnose-specific checklists to help GPs improve the quality of referral letters in gastroenterology and to assess the acceptability of this tool for the GP.

Paper IV: To assess whether referral letters generated with the support of checklists were considered to be of higher quality by the gastroenterologists, and whether they could result in a better agreement in the referral assessment of similar patient

6. Material and Methods

6.1 Design and study populations

In this thesis, the study designs and populations used in the four papers differ, and were chosen based on what was considered the appropriate design to explore the different facets of referral letter appropriateness, quality and improvements in a gastroenterology-setting. While paper I explored appropriateness and paper II explored quality of clinical referrals, paper III and IV explored an intervention aiming to improve these aspects of the referrals in a standardized setting using vignettes. An overview of the material and the results from the four papers together can be found in table 6 (page 37).

In **paper I** we assessed the appropriateness of referrals (n=301) for patients (57.1 % females, mean age 59 years) accepted for colonoscopy at a primary gastroenterology centre during 2004. The referral letters had been consecutively collected during a cross-sectional quality assurance study, and the findings during the colonoscopy were recorded by the gastroenterologist after the examination. The referrals in this study were mainly paper-based.

In **Paper II** we developed a Thirty Point Score (TPS) to assess quality of referral letters in gastroenterology, and validated it in a clinical setting. The study consisted of two parts. The first part was completed in 2014, and included 25 gastroenterologists (28 % females, mean age 49 years). They completed a web-based survey where they selected the variables to be included in the TPS. The participants were recruited through personal contact or professional networks. In the second part, completed in 2015, a cross-sectional validation study was conducted at seven primary referral centres for colonoscopy. Referral letters (n=327) for patients (62.4 % females, mean age 57 years) within nine different indications (dyspepsia, dysphagia, longstanding abdominal pain, change of bowel habit, diarrhoea, constipation, gastrointestinal bleeding, weight loss and jaundice/elevated liver enzymes) for referral to gastroenterology services were collected. The referrals in this study were mainly electronic, but there were also some paper-based referrals included.

While the referral letters collected in paper I consisted of referrals for colonoscopy only, the referrals in paper II represented a wider range of indications, including upper gastrointestinal symptoms, weight loss and liver diseases. However, the lower-abdominal indications in paper II were largely colonoscopy referrals, and the study indications were therefore to some extent

comparable.

To investigate whether the quality of the letters in paper I and paper II were also in agreement , we made a random sampling of 1/3 of the referrals letters from paper I, 115 referral letters in total. After excluding referral letters regarding follow up of findings on other procedures, polyp/CRC controls and screening, we were left with 80 referral letters that could be assessed by using the TPS. We chose to eliminate referral letters not sent from primary care, and thus ended with 69 referral letters from GPs that we assessed by using the TPS. The results of this assessment can be found in table 6.

In **paper III**, completed in 2014/2015, we included 45 GPs (51% females, mean age 51 years) through already established mandatory educational groups for GPs in Norway, as well as through personal contacts. A flowchart can be seen in the appendix of the paper. The GPs participated in a randomized crossover trial where they created referral letters for 8 patient cases (vignettes) representing different clinical scenarios (table 3). The GPs were randomized to create the referrals with or without the support of electronic dynamic checklists, and after minimum 3 months they repeated the referrals with the opposite method. All GPs were contacted by email after three months for participation in the second period. The participants who did not respond to the first email invitation received repeated invitations and phone calls to encourage participation. They were invited to participate together at the hospital computer-room, or at home at their own convenience. Both options were considered feasible as they had learned to use the IDRI web page previously in the first period of the trial. Twenty five GPs completed the trial. This resulted in a total of 360 referrals (180 pairs).

In **paper IV**, completed in 2015/2016, 32 gastroenterologists (9.4 % females, mean age 50 years) participated in a web-based vignette study, to explore how the use of checklists affect the subjective quality assessment and the prioritization of the referrals. The participants were recruited through an email invitation to members of the Norwegian Association of Gastroenterology (NGF). The 16 vignettes (referral letters) used in the survey were randomly sampled from the referral letters generated in the first period of paper III to include one free text referral and one checklist referral from each of the same indications as used in paper III. They were transcribed to blind the observers for the presence of checklist-items and were presented in two rounds separated by a minimum of 3 months. Each round contained one randomly selected referral from each indication.

Both papers III and IV focus on evaluating the effect of an intervention (checklists) on the quality

of referrals, and the referral letters generated by the GPs in paper III were used in the vignette survey in paper IV. Thus, paper III and paper IV may be seen as a continuum, exploring the effect of the intervention from several angles.

6.2 Measurement of main outcomes

Objective measurements

European Panel on the Appropriateness of Gastrointestinal Endoscopy (EPAGE) guidelines:

We used the information presented in the referral letters to select the relevant EPAGE II-indication and supplementary information regarding the patient's history and symptoms on the EPAGE II web page (144). When there was more than one relevant alternative indication for the referral letter, the indication resulting in the highest appropriateness-rating was selected.

The same procedure was followed for the first EPAGE guideline (162).

Thirty point score (TPS)

In paper II we developed a score for assessing quality of referral letters in gastroenterology, the Thirty Point Score (TPS). The methods used to develop the score can be summarized in 3 main steps:

- Content validity (The ability of a core to measure all facets of a given construct): defining the score items.
 - Defining potential items for the score through available literature (27, 163, 164) and expert opinion- 2 gastroenterologists, one PhD student.
 - Selection of the 15 most important score items in a web-survey- 25 gastroenterologists.
 - Repetition of survey to test for reliability of the selection of information items - 16 gastroenterologists.
- Construct validity (The test's ability to measure what it claims to be measuring (165)): Comparison with other measure of referral quality.
 - Referral letters for gastroenterology were collected in a multi-centre study.
 - We used a VAS for assessing referral quality as a comparator for the score- gastroenterologists assessed the VAS, the PhD student calculated the TPS.
- Reliability (the test's ability to show the same result in repeated measurements and with different raters (166)): 25% of referral letters from step 2 reassessed blinded.
 - Intra- rater reliability- PhD student 6 months after the initial assessment.
 - Inter- rater reliability- Experienced gastroenterology fellow.

The TPS- items are presented in the appendix of paper II. An example (from the indication dyspepsia) of the Case Report form (CRF) used to register the VAS and the TPS- items from each referral letter is presented in the appendix of the thesis.

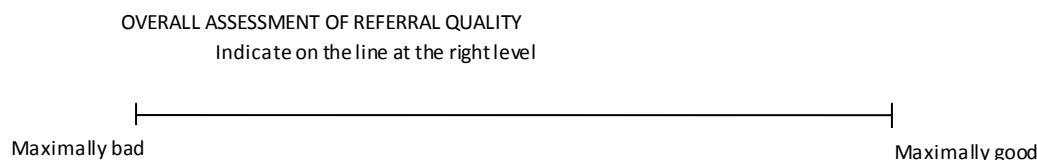
Subjective measurements

To assess gastroenterologists' subjective opinions regarding the quality of referral letters, we used two different scales, the visual analogue scale (VAS) (**paper II**) and a numeric rating scale (NRS) (**paper IV**).

Visual analogue scale (VAS)

A Visual Analogue Scale is a psychometric response scale where subjective characteristics or attitudes that are not easily assessed objectively are measured on a continuum ranging between two end-points, usually ranging from 0mm to 100mm. In **paper II** we used a VAS as shown in figure 3. The VAS was used as the comparator for referral quality in the validation trial in **paper II**, and the VAS for each referral letter was determined by the gastroenterologist who collected and assessed the referral letters at each centre.

Figure 3: The VAS scale used in paper II.



Numeric rating scale(NRS)

The primary end-point in **paper IV** was an 11-point NRS. This scale was constructed to be as similar to a VAS as possible with no numbers visible on the scale. The gastroenterologists rated the quality of referral letters by checking the appropriate button on the scale (figure 4).

Figure 4: The NRS used in paper IV.

* 12. On a scale from 0 to 10, how would you rate the quality of this referral?



6.3 Other outcomes and assessments

Diagnostic yield (DY): Assessed as the proportion of colonoscopies where a significant endoscopic diagnosis was detected. Significant diagnoses were defined as: CRC, adenomatous polyp, IBD, microscopic colitis, angiectasia and non-malignant stricture.

GP opinion of checklist: GP satisfaction, opinion regarding the user-friendliness of the checklist and willingness to implement such checklists were assessed through a web-based questionnaire after completing both periods of the vignette trial. The questionnaire was designed as a multiple choice form and consisted of a combination of Likert scales and other dichotomous and polytomous (ordinal and nominal) alternatives for the responses.

Referral assessment: The gastroenterologists assessed the referrals in paper IV according to the following variables: prioritization (maximum acceptable number of weeks before consultation/procedure), schedule appropriate consultation/procedure (colonoscopy/gastroscopy/radiology/consultation), preliminary diagnosis (free text), whether additional information regarding the patient would be warranted (yes/no/ideally, but is usually not done).

6.4 Vignettes and IDRI web-page

Vignette development

In paper III and IV, vignette studies were used. In paper III, the vignettes were dynamic and sequential simulated patient cases, while in paper IV the vignettes were static, written referral letters from the same patient cases as in paper III.

The vignettes were designed based on a potential underlying diagnose, and given symptoms matching the diagnose and different clinical scenarios from the Norwegian National Prioritization Guideline for gastroenterology (NPGg) (27) (table 3). The vignettes presented with different degrees of seriousness of the symptoms and findings.

Table 3: Patient cases with presenting symptom and underlying diagnose in paper III and IV.

	Symptom	Diagnose
Vignette 1	Upper abdominal pain/reflux	Ulcus
Vignette 2	Change of bowel habit	Diverticulosis/irritable bowel syndrome (IBS)
Vignette 3	Diarrhoea	IBD
Vignette 4	Rectal bleeding	Adenomatous polyp
Vignette 5	Long- standing abdominal pain	IBS /constipation/no findings
Vignette 6	Constipation/family history CRC	No finding, side effect of medicines
Vignette 7	Dysphagia	Oesophageal cancer
Vignette 8	Fatigue, jaundice, elevated liver enzymes	Autoimmune hepatitis

The IDRI user interface

The development of the IDRI web page was done in cooperation between the study team and Microsoft AS. It was developed on a SharePoint platform, and was designed to resemble the interfaces of common EHR for general practice as a virtual EHR simulator (figure 5). The GP could communicate with the patient case and complete a physical examination through chatting with it, and could also order laboratory tests and radiology before referring the patient.

Figure 5: Interface of IDRI web page with the different sections highlighted in colour.

The screenshot displays the IDRI web interface, which is a virtual EHR simulator. The interface is divided into several sections, each highlighted with a different color to indicate its function:

- Purple:** Chatting area on the left side, featuring a patient avatar and a text input field for communication.
- Red:** Referral support section in the center, containing a "Diagnose / Problemstilling" form with various checkboxes and text input fields for patient history and symptoms.
- Dark Blue:** Background information section at the bottom of the red area, providing detailed patient history and family background.
- Yellow:** Free-text referral section at the top right, allowing for a detailed written referral.
- Light Blue:** Section for ordering radiology and lab tests on the right side, featuring buttons for "Rekvirer prøver" and "Rekvirer røntgen".

Explanation: purple: chat-area, red: referral support, dark blue: background information, yellow: free-text referral, light blue: section for radiology and lab.

6.5 Intervention

The intervention explored in the thesis was dynamic, diagnose specific checklists as a method to improve the quality of referral letters. The checklists were developed by the study team after review of different sources of information regarding important information in referrals (27, 163, 164).

We integrated the referral support in the SharePoint platform by using an access database through a remote desktop. The referral support was designed as a checklist-like drop-down sheet where diagnose specific anamnestic questions could be answered by ticking the appropriate answers in checkboxes or answering specific questions in small free-text fields. The lists were different depending of the indication/referral diagnosis and were triggered by adding the referral diagnosis in the IDRI interface. They were optional to use and dynamic, adjusting to the replies from the GPs, e.g. the alternatives changed somewhat depending on the clicks the GPs did in the list. As an example, if the GP clicked the box indicating that the patient had lost weight, he was presented with alternatives to specify the size and duration of the weight loss.

The lists had 3 different formats: 1) checkboxes for multiple answers, 2) checkboxes with one answer, 3) small free-text fields. Every referral could have a maximum of 3 referral diagnoses, and thus potentially trigger 3 checklists. The lists were combined with a free text referral sheet as recommended in the literature (55).

6.6 Statistical methods and power analyses

Statistical analyses

Statistical analyses were performed using the Statistical Package for Social Services (SPSS) version 23.0 (SPSS Inc. Chicago IL) and STATA version 14 (StataCorp LP). In all papers, the significance level was set at 0.05. Continuous variables are presented as mean (95%CI) or mean (range). Categorical variables are described as numbers and percentages.

ICC values were interpreted as: >0.75=excellent, 0.40–0.75=fair to good and <0.40=poor(167). The κ values were interpreted as: >0.80=very good, 0.61–0.80=good, 0.41–0.60=moderate, 0.21–0.40=fair and <0.21=poor (168). The correlation coefficient was interpreted as 0.70-1.00 = high, 0.50-0.70=moderate, 0.30-0.50= low and 0.0-0.30= negligible (169).

In analyses involving multiple ratings by the same rater (**paper III and IV**), multilevel regression analyses were used for the overall analyses to avoid underestimation of the standard deviation (SD) resulting in type 2 error.

Table 4: Statistical methods and study designs used in the papers:

Paper	Design	Statistical method/measurements
Paper I	Cross-sectional study	Chi-square test Logistic regression analysis (univariate and multivariate)
Paper II	Cross-sectional study	Intraclass correlation coefficient (ICC) Weighted κ Linear regression analysis(univariate and multivariate) Pearson correlation Analysis of Variance (ANOVA)
Paper III	Randomized crossover vignette trial	Multilevel linear regression Paired t-test Linear regression analysis(multivariate and multilevel) Variance and Coefficient of Variance (CV) Logistic regression analysis(multivariate and multilevel)
Paper IV	Randomized vignette study	Linear regression analysis (univariate and multilevel) Paired t-test Chi-square test Unweighted Fleiss' κ

Power analyses

Paper I: The power estimation was done posteriorly, to ensure that the study was powered to answer the research question. It was based on the 323 collected referral letters and the diagnostic yield was defined as the primary endpoint. Power estimations related to individual groups of findings, e.g. detection of CRC was not performed.

Paper II: Sample size estimation was performed to investigate the association between VAS and the TPS, estimating a correlation coefficient of 0.6. Based on this, a minimum of 21 referral letters were collected for each indication. As some indications were more common than others, the number of referral letters per indication was in some cases higher than the minimum required.

Paper III: The power estimation in paper III was based on the detected mean TPS and SD in clinical, non-checklist referrals collected in paper II, and we considered a 30% change in the TPS clinically relevant.

Paper IV: Power estimation was based on detecting a 30% change in the NRS of the referral letters, and the VAS detected in paper II was used as reference to estimate the expected value for the non-checklist referral letters.

6.7 Ethical considerations

The work on this thesis has not implicated major ethical considerations. The completion of the IDRI trial is of potential value to ensure a more rational and ethical use of resources for implementation of ICT solutions, as it provides low-cost evidence in a setting where patients are not affected by potential system- shortcomings.

This thesis has mainly been completed as observational- or non-clinical studies. Throughout the work on the different studies, recommendations have been obtained from the local Data Protection Official for research, to ensure appropriate management of personal information, both for the participating physicians and for the patients who were referred during the project.

The Regional Ethics Committee (REK) considered the studies outside its mandate and permissions were not mandated.

7. Summary of the results

Paper I

European panel on the appropriateness of gastrointestinal endoscopy II guidelines help in selecting and prioritizing patients referred to colonoscopy – a quality control study

In this paper, we assessed the appropriateness of referrals to colonoscopy and assessed how the appropriateness related to the diagnostic yield (DY) of the examination.

The main findings in this paper were:

- 301 referrals were assessed by the EPAGE and EPAGE II criteria.
- The EPAGE II criteria were applicable in 98% (295/301) of the cases vs. 95.3% (287/301) for EPAGE.
- 81.4% (240/301) of referrals were considered appropriate by the EPAGE II criteria vs. 57.8% (166/301) for EPAGE.
- For EPAGE II, the DY was 31.3% (75/240) for appropriate referrals versus 10.9% (6/55) for uncertain/inappropriate referrals (OR = 3.5, 95% CI: 1.4-8.9, $p = 0.007$).
- For EPAGE, the DY was 34.9% (58/166) for appropriate referrals versus 17.4% (21/121) for uncertain/inappropriate referrals (odds ratio [OR] = 2.5, 95% CI: 1.8-4.4, $p = 0.003$).
- The EPAGE II criteria were more sensitive for detecting significant endoscopic lesions than the EPAGE criteria (92.6% vs. 73.4%).

Paper II

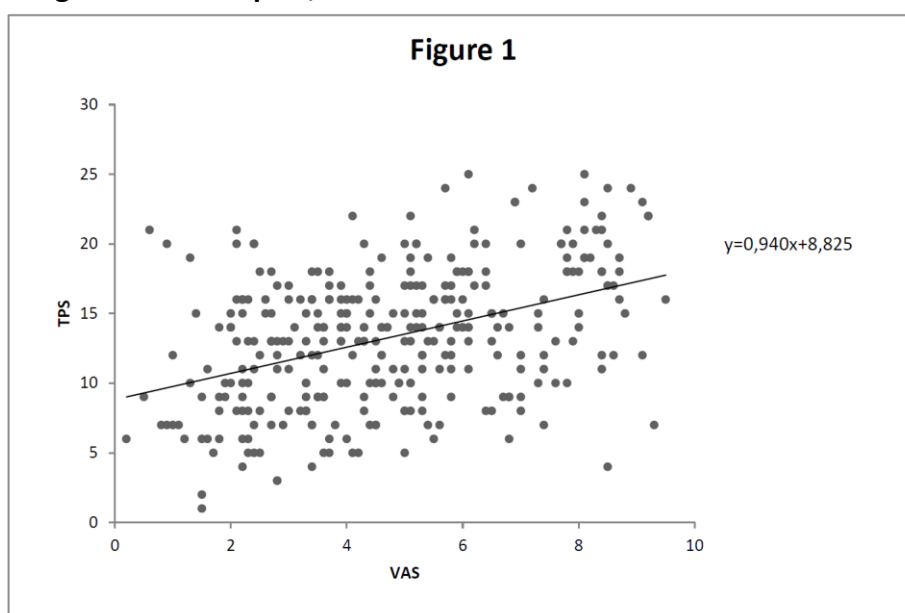
First quality score for referral letters in gastroenterology- a validation study.

In this paper we developed a score to assess the quality of referral letters, and used it to assess the quality of referral letters to Norwegian primary gastroenterology services.

The main findings in this paper were:

- The development of a novel, symptom specific 15 items score, the TPS, for nine important indications for referral to gastroenterology services, namely dyspepsia, dysphagia, chronic abdominal pain, change of bowel habit, diarrhoea, constipation, gastrointestinal bleeding, weight loss and jaundice/elevated liver enzymes. The maximum score-value is 30, indicating a high-quality referral letter. The score is presented in the appendix of the thesis.
 - Inter-rater reliability: ICC = 0.91 (95% CI 0.86-0.94)
 - Intra-rater reliability: ICC = 0.87 (95% CI 0.81-0.92)
 - Correlation between the TPS and VAS score for quality of the referral letters was $r=0.42$ (figure 6).
- Average VAS for quality of the referral letters was 4.7 (95% CI 4.5-5.0, range 0.2-9.5).
- Average TPS for quality of the referral letter was 13.2 (95% CI 12.8-13.8, range 1-25).

Figure 6: Scatterplot, correlation between TPS and VAS.



Paper III

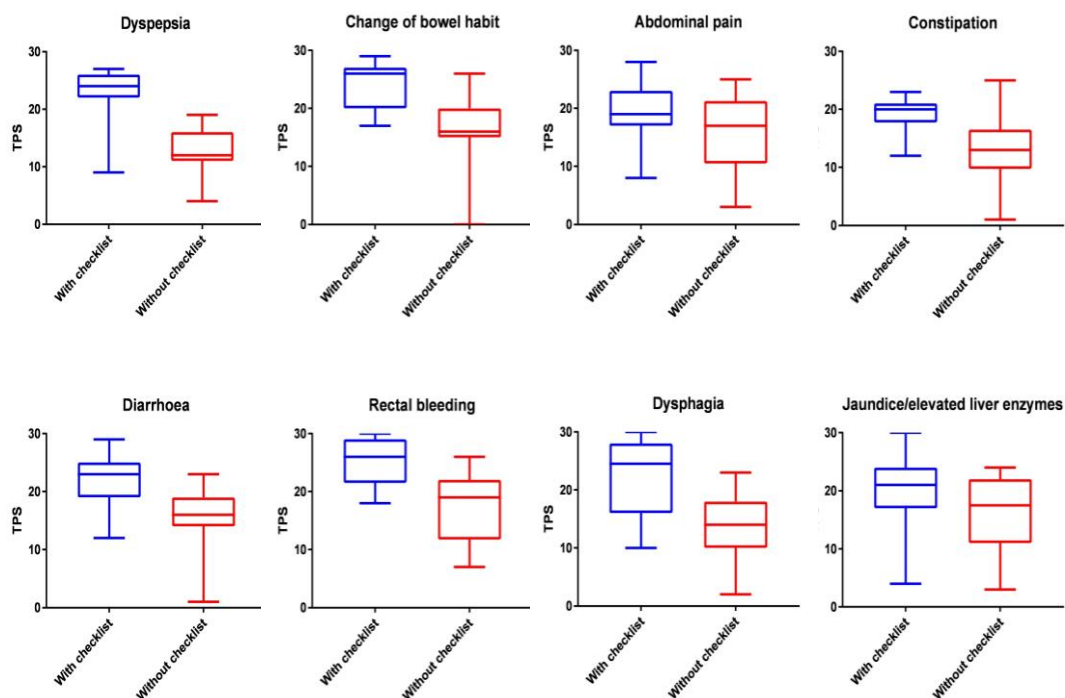
An Interactive Dynamic Referral Interface (IDRI) Improves quality of referral letters- a randomized cross-over vignette trial.

In this paper we developed symptom- specific checklists for referrals in gastroenterology, and evaluated the effect of these checklists on the quality of the referral according to the TPS.

The main findings in this paper were:

- The mean TPS was higher in referral letters with checklist than without checklist (Figure 7). Overall $\Delta=6.8$ (95% CI 5.1-8.5, $p<0.001$).
- The difference between the checklist- and the non-checklist referrals was also considerable when comparing single information items.
- The variance in the checklist-referral letters was lower in the checklist referrals (26.5) compared with the non-checklist referrals (36.2). The coefficient of variance (CV) was 23.3% for the checklist group and 39.6% for the non-checklist group.
- The GPs were generally positive to the check lists/referral support, but some 33% considered the checklists to be too extensive.

Figure 7: TPS with and without checklist for each patient case



Paper IV

Electronic checklists improve referral letters in gastroenterology- a randomized vignette survey.

In this paper we assess the gastroenterologist-assessed quality of referral letters created with-and without using electronic referral support/checklists. We also explored whether the referral support would lead to increased agreement in the assessment of referrals for similar cases.

The main findings in this paper were:

- The gastroenterologists considered the referral letters created with the checklist-support to be of a higher quality than then referral letters created without this support. The mean quality of the referral letters on the NRS was 7.0 (95% CI 6.8-7.2) for all letters combined, 6.5 (95% CI 6.2-6.8) for the free text referral letters (n=256), and 7.5 (95% CI 7.3- 7.7) for the checklist referrals (n=255) ($p < 0.001$) (table 5).
- More gastroenterologists would have required additional information regarding the patient to make a proper referral assessment for the referral letters created without the checklists, compared with the checklist- referrals.
- There was no difference in the clinical scheduling of the patients, but some more patients were correctly diagnosed in the checklist-group.

Table 5: Comparison of the rating of the two groups of referrals, and the TPS of the referrals.

Clinical case	N vignettes/ referrals	N assessment pairs	TPS for the case	Mean rating with checklist (95% CI)	TPS for the case	Mean rating without checklist (95% CI)	P-value*
Dyspepsia	2	32	27	8.5 (8.1-8.9)	12	7.8 (7.1-8.4)	0.026
Change of bowel habit	2	32	28	7.9 (7.3-8.5)	16	6.1 (5.4-6.8)	<0.001
Diarrhoea	2	32	28	8.3 (7.9-8.8)	11	5.9 (5.1-6.6)	<0.001
Rectal bleeding	2	32	28	7.4 (6.9-7.9)	22	7.8 (7.1-8.4)	0.216
Abdominal pain	2	32	20	5.7 (5.0-6.5)	23	7.1 (6.6-7.1)	<0.001
Constipation	2	32	20	7.5 (6.9-8.1)	13	5.6 (5.0-6.3)	<0.001
Dysphagia	2	32	21	7.0 (6.3-7.6)	2	4.6 (3.7-5.5)	<0.001
Jaundice/elevated liver enzymes	2	31	24	7.4 (6.9-8.0)	24	7.3 (6.6-7.9)	0.724
Total	16	255		7.5 (7.3-7.7)		6.5 (6.2-6.8)	<0.001

*Paired sample t-test for NRS scores

Overall comparison of the results from the 4 papers

Table 6: Comparison of the populations and results in the 4 papers.

	Paper I	Paper II	Paper III	Paper IV
Population	Referral letters to gastroenterology services in Asker and Bærum County.	Referral letters to gastroenterology services in the South Eastern Norway Health Region	Patient cases (vignettes) in a virtual setting	Referral letters randomly selected from paper II presented in a web-based survey as vignettes
Number of referrals	301	327	360	16
Time of study	Jan-Oct 2004	May-Sept 2014	First period completion: Apr-Dec 2014 Second period completion: Dec 2014-Jul 2015	First round completion: Jun 2015- Jan 2016 Second round completion: Sept 2015- May 2016
Study design	Cross-sectional study	Cross-sectional study	Randomized cross over vignette trial	Randomized vignette study
Physician age, years(range)	NA	47 (26-72)	51 (21-72)	50 (33-75)
Physician gender, n female (%)	NA	122 (37.3)	2 (8.0)	3 (9.4)
Patient age, years(range)	59 (19-90)	57 (6-94)	NA	NA
Patient gender, n female (%)	172 (57.1)	204 (62.4)	NA	NA
TPS, mean(95% CI)	8.4 (7.3-9.6)*	13.2 (12.8-13.8)	Checklist: 22.0 (20.6-23.4) Non-checklist: 15.2 (13.2-17.2)	Checklist: 24.5** Non-checklist: 15.4**
VAS/NRS, mean (95%CI)	NA	4.7 (4.5-5.0)	NA	Checklist: 7.5 (7.3-7.7) Non-checklist: 6.6 (6.2-6.8)
EPAGE II appropriateness	Appropriate: 81.4% Uncertain: 11.5% Inappropriate: 7.1%	Appropriate: Uncertain: Inappropriate:	NA	NA

*random selection of 1/3 of referral letter from paper I. ** means calculated from the 8 TPS- values of the vignettes(table 5).

8. Discussion

The results of this thesis show that referrals to colonoscopy are generally appropriate, and that inappropriate colonoscopies have a lower diagnostic yield than appropriate colonoscopies. A variable and low quality of referral letters to gastroenterologists has also been demonstrated. We have developed a score to assess referral quality in gastroenterology and used it to evaluate the effect of implementing dynamic, diagnose specific referral letters on the quality of referral letters. While we observed increased quality of the letters, we did not observe any impact on the prioritization of the referral letters. Below, some considerations related to the study design and methods will be discussed, before a general discussion of the main results.

8.1 Methodological considerations

There are some inherent strengths and limitations associated with different trial designs. Internal validity is the degree in which a study is appropriately designed to draw correct conclusions, e.g. between an exposure and an effect (170). The internal validity of a study is determined by the extent to which the study has succeeded in eliminating systematic (bias) and random error. Observational studies are particularly prone to systematic error, in particular selection bias, information bias and confounding (170). Cross-over studies are particularly susceptible to carry-over effects and drop-outs after the first period, and require that the studied material/condition is stable over time and does not require long-term follow-up.

Below I will discuss strengths and weaknesses of the study designs, study recruitment process and measurement tools with focus on potential sources of classical bias and other factors that may have affected the outcomes of the studies.

8.1.1 Design and study populations

Study designs

In paper I and II observational, cross-sectional studies were used to obtain data from clinical practice. Paper I presents a quality assurance study (171).

Both studies are subject to bias, in particular observer bias, as the assessors were not blinded from the outcome variables. In paper I, knowledge of the colonoscopy findings potentially influenced the researcher's classification of the appropriateness level and consequently could have led to an overestimation of the association between the appropriateness-level and the diagnostic yield. In

paper II the VAS set by the gastroenterologists was known to the TPS-assessor due to using the same form for this task. Consequently, some score items may have been classified to better match the VAS-value.

Even though we believe this did not occur, we cannot exclude these potential biases. In hindsight this risk of bias could have been avoided by using separate sheets to record the outcome variables.

In paper III we chose a randomized crossover design with a minimum of three months wash out between periods. This design is particularly useful in a setting where the variable measured at baseline is highly individual, because study subjects act as their own controls in paired statistical analysis. It also requires fewer study subjects to reach statistical significance. We expected that there would be significant inter-individual differences in the quality of the referral letters at baseline, and that any improvement caused by the checklist would be relative to this value.

The three months wash out was chosen based on clinical judgement as we were not able to identify any literature defining memory-washout time. A carry-over effect could minimize the measured effect of the intervention by increasing the score in non-checklist referrals from the second period (172). Most participating GPs completed the second period even later, and the median number of months between the periods was six. To account for any recollection of the checklist items, we adjusted statistically for the time of receiving the checklist intervention (before or after crossover) and found no significant impact on the TPS. The effect of recollection therefore seems to be limited.

Studying the clinical management of the referral (e.g. prioritization) and patient outcomes of a checklist-intervention is ideally done in a clinical setting and requires implementation of checklists in clinical practice. However, the trial in paper III gave us the opportunity to explore the effects of the checklists in a standardized setting as a pilot preceding implementation. Paper IV was therefore designed as a randomized vignette study with referral letters presented to the participants in two rounds, allowing for paired statistical analysis also in this study. Since outcomes could not be evaluated in this setting, we used surrogate end-points related to the prioritization of the referral, which was done by the gastroenterologist based on the information in the letter.

Alternative approaches to the vignette trial in paper III were considered. Both paper sheets/static interfaces presenting the patient case without an accompanying dialogue (similar to paper IV) or standardized patients (described in section 4.4) were possible alternative solutions. The advantage of using a static design in paper III would have been to avoid challenges associated with the

communication with the vignettes. However, this could to a greater extent have influenced the GPs in the same direction regarding the content of referral letters, and a “copy-paste” approach with more uniform referral letters of high quality could have been the result. Standardized patients would have been an easier approach for the GPs, but would have been more expensive, and required implementation in existing clinical EHR systems to create a realistic setting, and was discarded as method due to these issues. However, an approach using standardized patients may have increased the quality of the trial and improved the validity of the findings in an implementation-setting.

Finally, another possible alternative was to implement the checklists in existing EHR systems for general practice and complete the trial as a clinical cluster-randomized trial. This approach was considered during the planning of the trial, and the study team had some initial meetings with EHR providers to discuss the feasibility of this approach. However, it was deemed too costly and not feasible at that stage of development. Thus, a vignette- based approach was finally chosen.

Collection of referral letters and recruitment of participants

Referral letters, paper I and II

In paper I, the referral letters were collected continuously, including all patients accepted for colonoscopy during the inclusion time. The assessment according to the EPAGE- guidelines was done years later. This delay was unfortunate, and the question regarding the referral letters’ current relevance may be questioned. Referral letters for patients who were not accepted for colonoscopy were not collected, and the likelihood of finding higher proportions of inappropriate referrals in the referrals rejected for colonoscopic examination is high. Thus, the referral letters in this study may not be fully representative of the appropriateness of the total volume of referrals for colonoscopy that are sent from primary care. In this population we would expect to observe a slightly lower appropriateness, depending on the rate of rejected referrals. We were unable to identify studies assessing appropriateness in rejected referrals, or that quantify the proportion of rejected referrals.

In paper II we aimed for continuous collection of referral letters until we reached the targeted inclusion, but in some centres continuous collection was not achieved. This introduces the risk of selection bias, which could result in a non-representative sample. The generalizability of the results to the general population is therefore questionable. However, this bias would not influence the correlation between the VAS and the TPS.

Gastroenterologists, paper II and IV

We invited gastroenterologists through open invitations, with no upper limit for the number of participants. This introduced a potential selection bias if participants had different opinions of the content or quality of referral letters than non-participants. This may be a strength of the study, since gastroenterologists with an interest in the topic may also have a better understanding of what should be considered important clinical information in referral letters.

General practitioners, paper III

We recruited GPs from established educational groups and through personal relations. Some of the participants were therefore enrolled in the study by their group-leader and were variably motivated for participation. The major challenge we consequently encountered with the cross-over design was the drop-out between the first and second period of the trial despite repeated attempts to make contact with the participants. The drop-out was higher in the intervention group, and even though the estimate of the effect of the intervention may be unaffected by this due to the cross-over design, there may be issues with the user-friendliness of the checklists that has influenced the study compliance. We did not administer the post-trial questionnaire to the GPs who dropped out due to their lack of a basis for comparison, but a follow-up questionnaire to assess the reason for the non-compliance could have been a feasible solution to account for the drop-outs in the trial.

Design of questionnaires

In the four papers in this thesis we used questionnaires/CRFs that were developed by the study team due to the lack of validated questionnaires fitting the study objectives. The questionnaires were unvalidated and some flaws in the designs were discovered during the data collection.

In the CRF used for the referral assessment in paper II, we asked whether the referral letters contained too much or too little information, and whether they were unstructured or illegible. This was done to gain additional insight into other factors influencing the VAS. However, we failed to inquire about the appropriateness of the referral or to give the option of providing with a negative answer to the questions, e.g. “none of the above”. This made us unable to adjust for appropriateness as a confounder for the quality of the referral letter assessed by the VAS, and is a major limitation in the study design. Additionally, we could not quantify missing data for this question, as CRFs with intentional and unintentional omission of ticked boxes could not be

distinguished. It is unknown whether the missing cases have led to bias due to differential omission.

8.1.2 Measurement of main outcomes

The subjective nature of quality (17) is reflected by the different methods used to define quality in this thesis, and also motivated the development of the TPS early in the work with this thesis.

Below, I will discuss the different choices of methods used to describe the endpoints in this thesis.

EPAGE guidelines

The EPAGE II guidelines were chosen in favour of the ASGE- guidelines due to the European setting of the study. After revision in the Scandinavian Journal of Gastroenterology, the EPAGE guidelines were also added for comparison of the two guidelines.

However, appropriateness criteria, like the EPAGE guidelines, have been criticized for being subjective and unreliable. Indeed, studies have shown variable agreement between panels assessing the same procedures (173), largely reflecting areas where evidence from randomized controlled trials is scarce (174). Expert opinion may also differ from the opinions of other groups like GPs (175), and patient opinions are generally not considered in a RAND-UCLA process (122). Therefore, appropriateness criteria should be used with care.

Thirty point score (TPS)

The TPS was designed to provide an objective measurement of the quality of referral letters. This required score-items looking beyond the mere structural, general information items (medications, previous medical history etc.) and include specific symptom- related items mentioned in the referral letters.

We chose to focus explicitly on information quality (e.g. presence of clinical variables/information items) when developing the score, and not on structure or appropriateness of the referral or patient administrative information that is automatically added to the electronic referrals (80). Additionally, we chose to focus on items describing the patient's underlying disease and need for health care, rather than on each information item's capacity to discriminate between patients with high- or low-risk of serious underlying disease. Thus the score is useful for the majority of referral letters in gastroenterology, regardless of the seriousness of the condition or the location of the disease.

Both the targeted number of items and the grading of the points assigned were strategies chosen by the study team based on qualified opinions. A score proposed by Jiwa et al contained 15 items (75), Wåhlberg et al used 10-17 items (80), while a score proposed by Hartveit et al contained 19 items (176). Thus, the number of items chosen for the TPS is comparable with similar scores. It is possible that a less comprehensive score may be more reliable and easier to use due to time consumption and the presence of fewer interpretative items (177). Scoring instruments with many interpretive questions could lead to a lower reliability, due to variations in the way different doctors understand the criteria, or how they interpret the findings according to the criteria (177). Therefore the assessment of the reliability of the score was an important part of the validation study.

There may be a need for modification of some of the score items to make the score more comprehensible and intuitive. Finding the balance between items that are sufficiently specific, but also easy to use and generally applicable has been a challenge and may be a focus in further development of the score.

One major challenge when developing a score to assess the quality of referral letters was that there is no recognized “gold standard” to use in the validation process. In other settings, e.g. validation of a tool for prediction of colorectal cancer, the gold standard could be the result of a colonoscopy, and the presence or absence of CRC would be verifiable with a small margin of error. However, this requires availability of objective methods to quantify quality, which was not available for referral letters. Hartveit et al validated a score for evaluating quality of referrals to mental health care by using the subjective opinion of specialists as reference (176).

Visual analogue scale (VAS)

We chose VAS as a comparator for the TPS, based on the assumption that quality of referral letters rated high on the VAS was satisfactory in terms of providing relevant clinical information. VAS is a recognized tool for assessing pain (178, 179) and soft data in clinical trials (180). It has e.g. been used in a validation study for a score to assess satisfaction with an EHR system in psychiatry (181), to document construct validity in the development of an IBD disease activity score (182), to evaluate IBD disease activity (183), for assessing mucosal injury after NSAID-use (184), to assess patient satisfaction (185) as well as to evaluate the quality of referral letters (65). We did, however encounter some challenges with this approach. One is that VAS is an individual and subjective scoring method with a variable inter-rater reliability (184, 186). This may have influenced the

comparison between the TPS and the VAS if the individual gastroenterologists used the VAS differently when scoring the referral letters. Indeed, the Analysis of Variance (ANOVA) showed differences in VAS between the centres that were not explained by the objective TPS.

It is also possible that the VAS also reflected other, global features related to the referral quality like appropriateness and structure in addition to measuring the quality of the information in the letter. This may have affected the correlation between the two measurements.

Another challenge with the VAS is that no cut-off values defining high- or low quality referrals exist. It is therefore a matter of interpretation whether or not a VAS of e.g. 4.7 indicates a low- or a medium quality referral.

In spite of shortcomings of the VAS there were few good alternatives to the use of such subjective scales.

Numeric rating scale (NRS)

In paper IV the VAS was replaced by a NRS ranging from 0 to 10 due to limitations of the web interface used for quality assessment of the referral letters. NRS have been shown to compare well with VAS, and an 11-point scale has been considered the most appropriate version of the NRS for pain assessments and comparisons of the two scales (187). A similar numeric scale was also used to describe referral quality by Wåhlberg et al in a recent study (188). To increase the comparability of the two measurements, we removed the numbers below the individual points on the NRS and left only an explanatory text in each end of the scale.

8.1.3. Other outcomes and assessments

Diagnostic yield

The definition of significant lesions was determined by using the same criteria as other studies to allow for comparison of the results. This definition is most likely a simplification, in particular for adenomatous polyps, which were all considered significant. Diminutive polyps ($\leq 5\text{mm}$) have very limited malignancy potential and nearly never progress to CRC (189, 190). Additionally, the slow growth of colorectal polyp implicate that the age of the patient should be considered when determining the significance of a finding during colonoscopy (191, 192).

Another factor worth considering is that significant lesions were counted regardless of their relation to the patient's presenting symptom. This implies that when a patient was referred due to

constipation and a 2mm adenomatous polyp was detected, the finding was considered significant even if it was unrelated to the patient's symptoms.

8.1.4 Vignettes and IDRI web-page

Vignette development

A vignette study was considered relevant for studying the checklist-intervention, since implementation in clinical practice was not considered feasible. Additionally, the possibility of eliminating external confounding factors, like physician time constraints and patient case mix was considered an important feature (145).

We chose to use two different types of vignettes in this thesis because these types of vignettes represented the most realistic simulations of the situation under investigation, namely referral generation and referral assessment.

For generating referrals we chose a sequential chat functionality in an EHR simulator to mimic a real consultation and to make the anamnestic work as realistic as possible. The response of the GP was open-ended, e.g. the GPs could choose the content and design of the referral letter and used their own words to describe the patients and their actions as recommended in the literature (145, 146).

For assessing referrals we chose static vignettes, and presented the full referral letter in a web-interface. The transcription of the referral letters may have prevented detection of large differences between the letters as they may have appeared very similar at the first glimpse. The answers of the gastroenterologists were to a larger extent closed-ended with pre-defined alternatives for most of the answers. Clinical referral-assessment through the hospital EHR system also has a limited number of options, and thus the setting resembled clinical practice.

We chose a relatively low number of vignettes to compensate for the workload associated with the anamnestic work with each case without compromising the generalizability of the results. One review paper has previously reported a median number of vignettes in studies of medical choice and judgement of 25 (160). Other non-medical investigators suggest a maximum of 20 vignettes (193). However, the evidence regarding the optimal number of vignettes and attributes per vignette is scarce for medical research, and should be decided based on the total workload to avoid loss of attention associated with a high number of vignettes (160). Another non-medical paper claims that more attributes, more choice options and more vignettes decrease the response

reliability, but do not bias mean responses (194). However, in these papers, vignette numbers around 16-20 are discussed, and this is far beyond the eight we used in the IDRI-trial.

Vignette studies have some weaknesses that may have influenced the results. These include the possibility that the performance of the GPs in a vignette trial does not reflect clinical practice (195), or that participants in the trial perform better due to the knowledge of being tested or due to the lack of time pressure and distractions (196). As an example, it may be easier to write down the required examinations than to actually perform them in clinical practice, and physicians may therefore perform better in a vignette trial with no “temporal cost” of performing the required tasks(153). In the IDRI trial, this effect of this standardized setting seems to be limited, based on the overlapping confidence intervals when comparing with the mean TPS in paper II.

The IDRI user interface

Other researchers have used patient simulators (197-201) and have found it a useful tool for i.e. medical education and to assess physicians’ clinical practice and communication. In spite of efforts to optimize the functionality and interface of the simulator, the IDRI web page demonstrated some important flaws during the trial that may have influenced the results of the study. In the first group of GPs participating in the IDRI trial, we experienced an unfortunate system failure that led to the referrals of two GPs not being recorded. This resulted in exclusion of the referrals, and we do not know whether this may have influenced the results of the trial. Some of the vignettes failed to understand the questions asked by the GPs, which may also have affected the GPs’ performance, mainly through exhaustion.

Additionally, some GPs may have rushed through the trial and done a poor job referring the vignettes. This is a plausible effect of a very time consuming trial, and may have been avoided by using fewer or simpler vignettes (160).

To assess whether the vignette design may have influenced the performance of the referring physician compared with a clinical setting, we can look at the TPS detected in real life referral letters paper II of 13.7 (95% CI 12.8-13.8). The TPS recorded from the vignette- non-checklist referrals was 15.2 (95%CI 13.2-17.2). It seems like the vignette setting was comparable with a clinical setting in terms of referral quality, with possibly a small increased performance of the GPs in the IDRI trial.

8.1.5 Intervention

In this work, we chose to explore the combination of a free text field (as in a standard referral letter) with optional structured data fields. This approach has been documented in previous studies (55, 81, 202), and provides flexibility for the physician, combined with firm requirements in terms of checklist items prompting a more structured approach to describing the presence or absence of symptoms and findings. The dynamic design, with adaptation of the checklist-items to the answers of the GP, was chosen to explore the flexibility of electronic tools in contrast to the more static design of a paper-based solution.

The checklists in paper III were created mainly to demonstrate whether such checklists increased the presence of important information items. They were not subject to rigorous peer review or validation before they were used in the IDRI trial, but underwent several rounds of assessment of the contents by the study team before concluding with a final version for use. Consequently, individual checklist items are not to be considered as a final, implementable version. Modifying the contents of the checklists according to different clinical needs and local adaptations, as recommended in the literature for both endoscopic reporting (203) and implementation of referral guidelines (88), was considered to be feasible at a later stage.

The checklists in the IDRI- trial may have been too extensive. Some GPs reported this when completing the post-trial questionnaire. It is important to find a balance between obtaining enough information of value through the checklist, and creating a checklist that is so extensive that the GP fail to use it or experience an unacceptable increase in workload. In future implementation of dynamic diagnose specific checklists, this aspect should be carefully considered. Additionally, checklists should be optional and should be combined with a free-text option.

8.2 Discussion of the main results

8.2.1 Appropriateness

The high degree of appropriateness of colonoscopy detected in this thesis is supported by both previous (125, 128, 131) and recent studies (126, 204). This largely reflects the EPAGE II guidelines' lowered threshold for classifying colonoscopies as appropriate. The previous version of the guideline had been criticized for insufficient sensitivity, resulting in classifying colonoscopies with significant findings as inappropriate (205). The current EPAGE II guidelines demonstrated a

high sensitivity for endoscopic findings in accordance with other studies (131), indicating that they are relatively safe to use to select patients for colonoscopy.

None of the studies mentioned below assess appropriateness of rejected referrals, or indicate whether or not any referrals are rejected in their hospitals.

Argüello et al found an appropriateness of 70% according to EPAGE II in their material (125). This population consisted of patients referred to a university hospital by non-GP specialists.

Surveillance of neoplastic lesions was frequently considered inappropriate, most likely due to inappropriate surveillance intervals, and polyps were frequently detected in these colonoscopies. Most likely these polyps would have also been detected at appropriate surveillance intervals (125, 189, 190). Abdominal pain was also a frequent reason for inappropriate colonoscopies. In similar populations consisting of approximately 25% GP referrals, Gimeno Garcia et al found an appropriateness of 80% (131), while Carrion et al found an appropriateness of 70% (128). They observed high degree of inappropriateness associated with too short surveillance and screening intervals. In a more recent study by Marzo-Castillejo et al, an appropriateness of 73% was reported in a population where over 50% of the patients were referred from primary care physicians. This study did not include patients from the CRC screening program.

It appears from these studies that the most important factor influencing the appropriateness of referral letters is the adherence to guidelines for correct surveillance intervals for neoplastic lesions and IBD. Additionally, screening colonoscopies should be performed within screening programs defining correct age and intervals for screening. In Norway, such a screening program is currently being piloted (206). For symptomatic patients, abdominal pain/lower abdominal symptoms are among the most frequent causes of inappropriate referrals (125, 131), in line with our findings.

Studies using the ASGE guidelines have similar results with regards to the proportion of appropriate colonoscopies. Mangualde et al detected a 82.6% appropriateness in their material (207) and Suriani found 77% (208). They also report a high frequency of polyp controls and screening, similar to the EPAGE II studies. The colonoscopies most frequently classified as inappropriate in the material of Mangualde et al were abdominal pain and chronic constipation (207), while Suriani et al also found a high inappropriateness for polyp surveillance and screening (208). The EPAGE guidelines consider abdominal pain an appropriate indication in some situations, in spite of a low diagnostic yield. The ASGE guidelines have, in contrary to EPAGE, deemed this indication as generally not indicated (136).

The proportion of inappropriate colonoscopies was relatively low in our cohort (7%). With more than 80000 colonoscopies yearly in Norway (14), this amounts to 6500 potentially unnecessary procedures, and a potential cost-reduction of approximately 22.3 million Norwegian kroner (cost/diagnostic colonoscopy of NOK 3427 (209)) if these procedures could be avoided. This is a conservative estimation, as it is not unlikely that the proportion of inappropriate colonoscopies has increased the last decade in line with the increased use of colonoscopies (14). The reason for the increase is currently undocumented in Norway.

The overall diagnostic yield in the present study is also in agreement with similar previous studies, ranging between 25-41% (125, 128, 131). A more recent publication show a much higher diagnostic yield (51%) (204). However, in that study diverticular disease was included as a significant finding. Diverticular disease is a frequent finding in patients after the age of 50 years (210) and in paper I we consider this a normal finding. The diagnostic yield was found to be higher in appropriate referrals compared with inappropriate referral in several studies (125, 128, 131, 204), and advanced neoplastic lesions are rarely found in colonoscopies classified as inappropriate (125, 128, 131, 204).

The clinical effect of implementing guidelines to better select patients for secondary care is unclear. It is possible that appropriateness- guidelines can be a useful tool for the gastroenterologists working with referral assessments, mainly by ensuring that the right patients are selected for the procedure, and to avoid scheduling inappropriate surveillance intervals (211). It would benefit both the patients and the health care system if it increases the proportion of appropriate procedures. Grassini et al showed that the use of appropriateness guidelines in general practice led to increased referral appropriateness, increased the cost-effectiveness of the healthcare system, and reduced the waiting lists for colonoscopy by 15% (212). Appropriate referrals and the use of referral guidelines have also shown some evidence of a potential to decrease diagnostic delay for CRC (213, 214). However, the evidence for the effect of referral guideline is conflicting (87, 215, 216). It is likely that guideline-implementation in general practice require a high degree of collaboration between physicians in primary and secondary care to ensure common understanding and motivation for the intervention (88). Kim-Hwang et al detected increased referral quality and appropriateness after implementation of eReferral in their hospital (56). The increased appropriateness was most likely related to the interactive communication that was possible through the eReferral system (56), strengthening the impression that support for GPs in the referral process may be of value to increase appropriateness of

referrals.

Only paper based referral letters were collected in 2004 and may not be representative of the current situation regarding structure, quality and appropriateness (49, 56-58). In Norway, referrals are currently transferred electronically and to a large extent written according to national standards (26, 28) as demonstrated in paper II. When we reassessed a random sample of the referral letters collected in 2004, we found a TPS of 8.4 (95% CI 7.3-9.6). This was significantly lower than observed in 2014, where the average TPS was 13.7 (95% CI 12.8-13.8) and this strengthens the impression that these referral letters are not directly comparable.

8.2.2 Referral quality

In paper II we observed a variable referral quality, with means below 50% of the maximum score values for both the TPS and to the VAS. Unsatisfactory quality of the referral letters has been a frequent finding in the literature, although the reporting methods and the studied specialties are fairly heterogeneous (5, 6, 63, 65-71, 74, 96-110). Missing information have been reported in between 37% (63) and 98% (96) of referrals, and the most frequently missing information is reported in pre-referral work-up/examinations (63, 67, 69, 70), medications (63, 65, 69, 98, 109, 110) and details regarding the current disease (63, 65, 74, 96, 103, 109). Most studies report missing information in this general way, not elaborating on diagnose-specific clinical information items that are under-reported, even if such details may be of importance for the prioritization of the patients.

Some studies that have reported the clinical information in referral letters in more details are worth mentioning. Jiwa et al found a mean reporting of only 6/18 possible upper gastrointestinal symptoms and signs, largely matching the findings in the TPS categories dyspepsia and dysphagia (74), while Gulati et al found a mean of 6/12 required information items present in referrals to a spinal unit (96). In studies of glaucoma referral letters, a lack of clinical details have been reported (100, 102), and Cheng et al suggest referral templates to easier include these vital details in the referrals (100). DeAngelis et al observed a lack of clinical details in referrals to an Oral and Maxillofacial unit, ranging from 7% to 67% missing data for the different information items (101). In a study of rheumatology referral letters, involvement of small joints was mentioned in 14 %, and morning stiffness was mentioned in only 1% (103). In referral letters to an obesity unit, information regarding the presence of metabolic syndrome was missing in 92-97 %, while information of the patient's waist circumference was included in only 1% of referrals (109).

8.2.3 The TPS

The importance of developing objective, reliable and valid tools for assessing quality of referral letters cannot be overemphasized, as it forms the basis for evaluating interventions aiming to improve referral quality and indirectly also patient safety and quality of care. Subjective quality assessments carry a higher risk of bias and are likely to have a lower inter-assessor reliability, preventing comparison of results both within a trial, and with other trials.

A score presented by Jiwa et al included items designed to identify patients with colorectal cancer, and was validated against this endpoint. It also comprised 15 items, with a complex score value assigned to each item. Ten of the items from this score are represented in the TPS. This score is interesting, but has limited value in a gastroenterology setting to evaluate referral letters for a wide range of indications (75). Another score/variable count for upper gastrointestinal referral letters presented by Jiwa et al, largely reflect the same items as in the upper-gastrointestinal categories of the TPS (13/18 items overlapping).

More recently, a score to assess the quality of referral letters in psychiatry has been suggested by Hartveit et al (176), and exhibits many of the same features and development steps that were applied when developing the TPS in paper II. Selection of score items was done by literature reviews and group interviews with specialists in mental care and GPs, using a RAND- based approach. Convergent validity was assessed by comparing the score value with the subjective assessments of specialists on the topic. Reliability was assessed by re-scoring the referral letters. Wåhlberg et al suggested a symptom specific score for evaluating referral letters in gastroenterology, lung diseases and cardiology, which was based on the presence of items from a referral checklist (80). This score is not validated, but several of the items in the gastroenterology-scores are also items in the TPS. For referrals with suspicion of CRC 9/10 items are present in the lower-abdominal categories of the TPS, and for dyspepsia referrals 11/16 items are represented in the TPS for dyspepsia referrals.

McGoldrick et al presented a symptom-specific score for head injury referrals, consisting of 15 items for history and examination and eight items for management of the patient (117).

Some scores mainly focus on the presence of general items like previous medical history and medications, which does not provide sufficient information to evaluate the presence of specific symptoms and findings (64). Several scores have included appropriateness in the score-items (71, 105), and the lack of this is possibly a limitation of the TPS.

These works are important contributions to the effort of developing standardized tools for evaluating referral quality, and hopefully more will follow.

8.2.4 Interactive, dynamic checklists and referral quality

The results of the randomized cross over trial revealed a significant effect of the intervention, both on the referral quality measured by the TPS, and on the reporting of individual clinical information items. Others have also studied the effect of interventions aimed at improving referral quality.

A Cochrane Database review of interventions to improve the quality of referral letters concluded that referral guidelines were more likely to be effective if structured referral sheets were used (88). However, the included studies focused mainly on appropriateness (8, 77, 93) or referral rates (92-94), not on the quality of the information conveyed in the referral letter.

Some of the included studies did observe an increased reporting of pre-referral work-up (77, 93), and Emslie et al detected increased appropriateness of referrals for glue ear when a risk factor checklist and a training video was disseminated in general practice (94). Only a few of the included studies actually used structured referral sheets in the intervention (8, 77, 92-94). Additionally, none of the studies included electronic referrals or electronic structured referral sheets or checklists.

Of other relevant studies, Shaffie et al observed increased referral information after a referral pro-forma for referrals to the oral and maxillofacial department was implemented (78). Rokstad et al implemented electronic checklists for referrals in pulmonary diseases in existing EHR systems for primary care (72). They 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. Heimly et al completed a similar project in Akershus University Hospital in Norway (89, 217). They showed that the GPs were satisfied with the intervention, but no increase in referral quality was seen. Studies evaluating the effect of checklists on other medical documents have also shown to improve the documentation quality, e.g. for radiology reports (83, 84) and for reports from endoscopic procedures (81).

We were not able to demonstrate any impact on the management of the patients due to the checklist intervention. Indeed, the topic of patient outcomes associated with the quality of the referral letter has only to a limited extent been studied (7).

Jiwa et al showed that an electronic referral guideline improved referral content, but did not change the outcome for the patient (218). Another vignette study from the same author showed a

correlation between the content of the referral letters and the confidence of the physician in the assessment and prioritizing of the patient to specialist health care services (219), but this was not confirmed in a later randomized controlled trial (119). Wåhlberg et al recently showed that paper-based checklists for dyspepsia, suspected colorectal cancer, chronic obstructive lung disease, and chest pain did not influence the prioritization or quality of care for the patients (188) in spite of an 18% increase in referral quality (80). A cluster randomized trial of a checklist intervention for urology referrals by Thomas et al pointed in the direction of shorter waiting times and improved cost-effectiveness when referral guidelines were implemented (8). One observational study by Tay et al showed that insufficient referral information may cause longer waiting times to see a specialist, as well as delayed diagnosis (220), while another observational study by Sales et al failed to show any correlation between quality of the referral letter and guideline compliance (221). Some evidence suggests that failure to examine or investigate may lead to doctor's delay in the diagnosis of CRC (222). If checklists prompt the GPs to increase the pre-referral workup and examinations (77, 93) and not only increased reporting of these items in the referral letter, this may have an impact on patient outcome, but this is undocumented by clinical trials. Thus, current evidence is conflicting and of variable quality, but indicates that the effect of referral- checklists and increased referral quality has a limited effect on patient outcomes. The studies in paper III and IV were conducted in a setting of gastrointestinal patient cases, but it is likely that under the condition of availability of appropriate diagnose specific checklists, the results are transferable to other medical specialties.

9. Conclusions

- The updated EPAGE II guidelines can probably be a useful tool for both the referring physician and the consultant gastroenterologist when deciding whether or not a patient will benefit from a colonoscopy. The guideline's ability to reduce the number of patients is limited, but it seems to be safer than with the previous version of the guideline. EPAGE II can possibly help reduce the significant burden of unnecessary examinations to the patients while also reducing the cost to the public health care system.
- The Thirty Point Score is a reliable score that provides a more objective assessment of the quality of referral letters in gastroenterology, and it shows a variable agreement with the subjective VAS. The score may be of particular importance when evaluating the effect of interventions to improve referral-quality. The method used in the development of the score can serve as a model for other medical specialties. However, the score may require some additional modifications before more widespread use.
- Electronic dynamic 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 the rationale for developing and implementing corresponding checklists has been strengthened. GPs are largely positive to the idea of a checklist for referrals.
- Based on the findings in this thesis we can conclude that the gastroenterologists may benefit from the intervention, as it is likely that increased quality of the referral letter makes the letter easier to read and less time consuming to assess regardless of the outcome for the patient. However, we need clinical trials to assess the impact on the healthcare delivered to the patients and the clinical outcome. Current evidence indicates that checklists for referrals do not influence patient outcome.

10. Future perspectives

The studies in paper III and IV were done in standardized settings, and implementation in the GPs' EHR-systems and evaluation of the effect in clinical trials should be the next step. A clinical trial may also be a more appropriate setting for studying the patient-related effects, e.g. outcomes related to improved quality of the referral letter. Such a clinical trial may be used as a pilot for a more widespread implementation of checklists as standard clinical tools in the EHR-systems for GPs. The intervention could also be combined with appropriateness assessments to evaluate the effect of the checklists on the indication of patients referred for colonoscopy.

The TPS developed in paper II may require further refinement and modifications that may increase the score's ability to discriminate between high- and low quality referral letters. Validation of the TPS in new trials is also warranted for more widespread use and a high general validity. The model for developing the score may be used in other specialties to evaluate referral letters in other areas than gastroenterology.

The referral letters in paper I were collected more than 10 years ago, and a larger study on new referral letters, preferably in a multicentre prospective setting may be warranted to ensure the validity of using the EPAGE II guidelines on more recent referrals. We have no data evaluating the appropriateness of present colonoscopies in Norway, and in the light of the huge increase in the use of this procedure, a new trial should be performed to investigate whether this increase is mainly due to appropriate use. This assessment should include referral letters from both patients accepted and from those not accepted for colonoscopy, even if the diagnostic yield could not be assessed for the rejected referrals. Additionally, a trial of a more systematic dissemination of the EPAGE II guidelines in the referring physicians' clinical practices is warranted to evaluate whether it could be helpful to the GPs when selecting patients to refer for colonoscopy.

The emerging era of automated systems and generation of big data may allow for clinical systems with integrated analysis of referral information and generation of decision support to better select and prioritize patients for appropriate examinations.

11. References

1. Epstein RM. Communication between primary care physicians and consultants. *Archives of family medicine*. 1995;4(5):403-9.
2. Forster A, Murff H, Peterson J, Gandhi T, Bates D. The incidence and severity of adverse events affecting patients after discharge from the hospital. *Annals of Internal Medicine*. 2003;138.
3. Vermeir P, Vandijck D, Degroote S, Peleman R, Verhaeghe R, Mortier E, et al. Communication in healthcare: a narrative review of the literature and practical recommendations. *International journal of clinical practice*. 2015.
4. Agarwal R, Sands DZ, Schneider JD. Quantifying the economic impact of communication inefficiencies in U.S. hospitals. *Journal of healthcare management / American College of Healthcare Executives*. 2010;55(4):265-81; discussion 81-2.
5. 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.
6. Kunkle EC. COMMUNICATION BREAKDOWN IN REFERRAL OF THE PATIENT. *Jama*. 1964;187:663.
7. Jiwa M, Dadich A. Referral letter content: can it affect patient outcomes? *British Journal of Healthcare Management*. 2013;19(3):140-6.
8. Thomas RE, Grimshaw JM, Mollison J, McClinton S, McIntosh E, Deans H, et al. Cluster randomized trial of a guideline-based open access urological investigation service. *Fam Pract*. 2003;20(6):646-54.
9. Esteva M, Leiva A, Ramos M, Pita-Fernandez S, Gonzalez-Lujan L, Casamitjana M, et al. Factors related with symptom duration until diagnosis and treatment of symptomatic colorectal cancer. *BMC cancer*. 2013;13:87.
10. Singh H, Petersen LA, Daci K, Collins C, Khan M, El-Serag HB. Reducing referral delays in colorectal cancer diagnosis: is it about how you ask? *Qual Saf Health Care*. 2010;19(5):e27.
11. Vega P, Valentin F, Cubiella J. Colorectal cancer diagnosis: Pitfalls and opportunities. *World journal of gastrointestinal oncology*. 2015;7(12):422-33.
12. Kyne G, Maxwell S, Brameld K, Harrison K, Goldblatt J, O'Leary P. Compliance with professional guidelines with reference to familial cancer services. *Australian and New Zealand journal of public health*. 2011;35(3):226-30.
13. Health expenditure ratios, by country, 1995-2014, Norway [Internet]. World Health Organization(WHO) 2016.
14. Norwegian Directorate of Health. Norwegian Patient Register. Number of colonoscopies 1999-2015. Norwegian Directorate of Health; 2016.
15. Morgan DJ, Brownlee S, Leppin AL, Kressin N, Dhruva SS, Levin L, et al. Setting a research agenda for medical overuse. *Bmj*. 2015;351:h4534.
16. Crispin A, Birkner B, Munte A, Nusko G, Mansmann U. Process quality and incidence of acute complications in a series of more than 230,000 outpatient colonoscopies. *Endoscopy*. 2009;41(12):1018-25.
17. Mosadeghrad AM. Factors influencing healthcare service quality. *International journal of health policy and management*. 2014;3(2):77-89.
18. Kaminski MF, Regula J, Kraszewska E, Polkowski M, Wojciechowska U, Didkowska J, et al. Quality indicators for colonoscopy and the risk of interval cancer. *The New England journal of medicine*. 2010;362(19):1795-803.
19. World Health Organization. Quality of care: a process for making strategic choices in health systems World Health Organization 2006.
20. Donabedian A. Evaluating the Quality of Medical Care. *The Milbank Quarterly*. 2005;83(4):691-729.
21. Thomson RG, McElroy H, Kazandjian VA. Maryland Hospital Quality Indicator Project in the United Kingdom: an approach for promoting continuous quality improvement. *Quality in health care : QHC*. 1997;6(1):49-55.

22. Guevara JP, Hsu D, Forrest CB. Performance measures of the specialty referral process: a systematic review of the literature. *BMC Health Serv Res.* 2011;11:168.
23. Boerma WG, van der Zee J, Fleming DM. Service profiles of general practitioners in Europe. European GP Task Profile Study. *Br J Gen Pract.* 1997;47(421):481-6.
24. 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):209-13.
25. Directorate for e-Health. Volven www.volven.no: Directorate for e-Health; 2016 [
26. Ree AO. KITH Report 22/03 Medical content of referrals- "The good referral". In: KITH, editor. Trondheim: KITH: Kompetence center for IT in health care AS; 2003. p. 35.
27. Nygaard E, Kårikstad V. National Prioritization Guideline for gastroenterology. Oslo: The Norwegian Directory of Health; 2009.
28. Norwegian Directorate of Health. National guideline for referrals to the specialist health care services. Oslo: Norwegian Directorate of Health; 2015.
29. Ringberg U, Fleten N, Deraas TS, Hasvold T, Forde O. 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.
30. Sonnenberg A, Amorosi SL, Lacey MJ, Lieberman DA. Patterns of endoscopy in the United States: analysis of data from the Centers for Medicare and Medicaid Services and the National Endoscopic Database. *Gastrointest Endosc.* 2008;67(3):489-96.
31. Barnett ML, Song Z, Landon BE. Trends in physician referrals in the United States, 1999-2009. *Arch Intern Med.* 2012;172(2):163-70.
32. Fleming DM. The European study referrals from primary to secondary care. Maastricht: University of Limburg; 1993.
33. Heimly V. Standardization, innovation and deployment of electronic referral software in Norway. *J Telemed Telecare.* 2008;14(7):359-62.
34. National ICT(NIKT). National ICT strategy plan for 2013-2016. National ICT(NIKT) 2012.
35. Donohoe MT, Kravitz RL, Wheeler DB, Chandra R, Chen A, Humphries N. Reasons for outpatient referrals from generalists to specialists. *J Gen Intern Med.* 1999;14(5):281-6.
36. Coulter A, Noone A, Goldacre M. General practitioners' referrals to specialist outpatient clinics. I. Why general practitioners refer patients to specialist outpatient clinics. *Bmj.* 1989;299(6694):304-6.
37. Ringberg U, Fleten N, Forde OH. Examining the variation in GPs' referral practice: a cross-sectional study of GPs' reasons for referral. *Br J Gen Pract.* 2014;64(624):e426-33.
38. Bjornsson S, Sigurdsson JA, Svavarsdottir AE, Gudmundsson GH. Gatekeeping and referrals to cardiologists: General practitioners' views on interactive communications. *Scand J Prim Health Care.* 2013;31(2):79-82.
39. Foot C, Naylor C, Imison C. The quality of GP diagnosis and referral. The King's Fund. 2010.
40. Shadd J, Ryan BL, Maddocks H, Thind A. Patterns of referral in a Canadian primary care electronic health record database: retrospective cross-sectional analysis. *Informatics in primary care.* 2011;19(4):217-23.
41. Limb M. Referral rate to outpatient services in London in some practices is triple that in others, report says. *Bmj.* 2012;345:e8249.
42. Wilkin D, Smith AG. Variation in general practitioners' referral rates to consultants. *The Journal of the Royal College of General Practitioners.* 1987;37(301):350-3.
43. Knottnerus JA, Joosten J, Daams J. Comparing the quality of referrals of general practitioners with high and average referral rates: an independent panel review. *Br J Gen Pract.* 1990;40(334):178-81.
44. Roland MO, Green CA, Roberts SO. Should general practitioners refer more patients to hospital? *Journal of the Royal Society of Medicine.* 1991;84(7):403-4.
45. Christensen T, Vigen T. [ELIN-prosjektet (ELEktronisk INformasjonsutveksling)]. 2009.
46. Norwegian Health-Net. Message counter (<http://meldingsteller.nhn.no/ConfigurableReport>): Norwegian Health-Net; [
47. Directorate for e-Health. Henvisning mellom HF <https://ehelse.no/nasjonale-prosjekter/henvisning-mellom-hf>: Directorate for e-Health; 2016 [

48. Norwegian Health-Net. Norwegian Health-Net (<https://www.nhn.no/Sider/default.aspx>): Norwegian Health-Net; [
49. Reponen J, Marttila E, Paajanen H, Turula A. Extending a multimedia medical record to a regional service with electronic referral and discharge letters. *J Telemed Telecare*. 2004;10 Suppl 1:81-3.
50. Wootton R, Youngberry K, Swinfen R, Swinfen P. Referral patterns in a global store-and-forward telemedicine system. *J Telemed Telecare*. 2005;11 Suppl 2:S100-3.
51. MedCom. Henvisningshotellet REFHOST, En brugervejledning for primærleger, speciallæger, fysioterapeuter og psykologer. . OneZone2008.
52. Heimly V. Electronic referrals in healthcare: a review. *Stud Health Technol Inform*. 2009;150:327-31.
53. Bouamrane MM, Mair FS. A qualitative evaluation of general practitioners' views on protocol-driven eReferral in Scotland. *BMC medical informatics and decision making*. 2014;14:30.
54. Tian L. Improving knowledge management between primary and secondary healthcare: an e-referral project. *Health Care Inform Rev Online*. 2011;15:31-7.
55. Esquivel A, Sittig DF, Murphy DR, Singh H. Improving the Effectiveness of Electronic Health Record-Based Referral Processes. *BMC medical informatics and decision making*. 2012;12(1):107.
56. Kim-Hwang JE, Chen AH, Bell DS, Guzman D, Yee HF, Jr., Kushel MB. Evaluating electronic referrals for specialty care at a public hospital. *J Gen Intern Med*. 2010;25(10):1123-8.
57. Nash E, Hespe C, Chalkley D. A retrospective audit of referral letter quality from general practice to an inner-city emergency department. *Emergency medicine Australasia : EMA*. 2016;28(3):313-8.
58. Gandhi T, Keating N, Dittmore M, Kiernan D, Johnson R, Burdick E. Improving Referral Communication Using a Referral Tool Within an Electronic Medical Record. *Advances in Patient Safety*. 2008;3.
59. Dennison J, Eisen S, Towers M, Ingham Clark C. An effective electronic surgical referral system. *Annals of the Royal College of Surgeons of England*. 2006;88(6):554-6.
60. Warren J, White S, Day KJ, Gu Y, Pollock M. Introduction of electronic referral from community associated with more timely review by secondary services. *Appl Clin Inform*. 2011;2(4):546-64.
61. Branger PJ, van der Wouden JC, Schudel BR, Verboog E, Duisterhout JS, van der Lei J, et al. Electronic communication between providers of primary and secondary care. *Bmj*. 1992;305(6861):1068-70.
62. Cannaby S, Westcott D, Pedersen CD, Voss H, Wanscher CE. The cost benefit of electronic patient referrals in Denmark: summary report. *Stud Health Technol Inform*. 2004;100:238-45.
63. Lønning KJ. *Tidsskrift for Den norske legeforening*. 2009(Årg. 129, nr. 18):1868-9.
64. Francois J. Tool to assess the quality of consultation and referral request letters in family medicine. *Canadian family physician Medecin de famille canadien*. 2011;57(5):574-5.
65. Garasen 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.
66. Ibiyemi O, Ibiyemi T. Quality and contents of referral letters from peripheral health centers to the dental centre of a teaching hospital, southwestern Nigeria. *Acta Odontol Scand*. 2012;70(2):165-8.
67. Ong SP, Lim LT, Barnsley L, Read R. General practitioners' referral letters--Do they meet the expectations of gastroenterologists and rheumatologists? *Australian family physician*. 2006;35(11):920-2.
68. Bahous MC, Shadmi E. Health information exchange and information gaps in referrals to a pediatric emergency department. *Int J Med Inform*. 2016;87:68-74.
69. Campbell B, Vanslembroek K, Whitehead E, van de Wauwer C, Eifell R, Wyatt M, et al. Views of doctors on clinical correspondence: questionnaire survey and audit of content of letters. *Bmj*. 2004;328(7447):1060-1.
70. Durbin J, Barnsley J, Finlayson B, Jaakkimainen L, Lin E, Berta W, et al. Quality of communication between primary health care and mental health care: an examination of referral and discharge letters. *The journal of behavioral health services & research*. 2012;39(4):445-61.
71. Jacobs LG, Pringle MA. Referral letters and replies from orthopaedic departments: opportunities missed. *Bmj*. 1990;301(6750):470-3.

72. Rokstad IS, Rokstad KS, Holmen S, Lehmann S, Assmus J. 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(3):166-71.
73. 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;14:7.
74. Jiwa M, Coleman M, McKinley RK. Measuring the quality of referral letters about patients with upper gastrointestinal symptoms. *Postgraduate medical journal*. 2005;81(957):467-9.
75. Jiwa M, Mathers N, Walters S. Quality of information on referrals to colorectal surgeons: towards consensus. *Current medical research and opinion*. 2002;18(2):72-7.
76. Jenkins S, Arroll B, Hawken S, Nicholson R. Referral letters: are form letters better? *Br J Gen Pract*. 1997;47(415):107-8.
77. Emslie C, Grimshaw J, Templeton A. Do clinical guidelines improve general practice management and referral of infertile couples? *BMJ*. 1993;306(6894):1728-31.
78. Shaffie N, Cheng L. Improving the quality of oral surgery referrals. *Br Dent J*. 2012;213(8):411-3.
79. Stille CJ, Mazor KM, Meterko V, Wasserman RC. Development and validation of a tool to improve paediatric referral/consultation communication. *BMJ quality & safety*. 2011;20(8):692-7.
80. Wahlberg H, Valle PC, Malm S, Broderstad AR. 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.
81. de Lange T, Moum BA, Tholfsen JK, Larsen S, Aabakken L. Standardization and quality of endoscopy text reports in ulcerative colitis. *Endoscopy*. 2003;35(10):835-40.
82. Gouveia-Oliveira A, Raposo VD, Salgado NC, Almeida I, Nobre-Leitao C, de Melo FG. Longitudinal comparative study on the influence of computers on reporting of clinical data. *Endoscopy*. 1991;23(6):334-7.
83. Sahni VA, Silveira PC, Sainani NI, Khorasani R. Impact of a Structured Report Template on the Quality of MRI Reports for Rectal Cancer Staging. *AJR Am J Roentgenol*. 2015;205(3):584-8.
84. Barbosa F, Maciel LM, Vieira EM, Azevedo Marques PM, Elias J, Muglia VF. Radiological reports: a comparison between the transmission efficiency of information in free text and in structured reports. *Clinics (Sao Paulo, Brazil)*. 2010;65(1):15-21.
85. Project Check. [Available from: <http://www.projectcheck.org/>].
86. Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat AH, Dellinger EP, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. *The New England journal of medicine*. 2009;360(5):491-9.
87. Faulkner A, Mills N, Bainton D, Baxter K, Kinnersley P, Peters TJ, et al. A systematic review of the effect of primary care-based service innovations on quality and patterns of referral to specialist secondary care. *Br J Gen Pract*. 2003;53(496):878-84.
88. Akbari A, Mayhew A, Al-Alawi MA, Grimshaw J, Winkens R, Glidewell E, et al. Interventions to improve outpatient referrals from primary care to secondary care. *Cochrane Database Syst Rev*. 2008(4):CD005471.
89. Akershus University Hospital HF. End-report Project for Interactive Referrals. Oslo: Akershus Universitetssykehus; 2011.
90. Kripalani S, LeFevre F, Phillips CO, Williams MV, Basaviah P, Baker DW. Deficits in communication and information transfer between hospital-based and primary care physicians: implications for patient safety and continuity of care. *Jama*. 2007;297(8):831-41.
91. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *BMJ*. 2005;330(7494):765.
92. Abu-Ramadan MA. Making better use of scarce resources: the Palestinian experience, 1995-1999. *J Ambul Care Manage*. 2002;25(3):63-9.
93. Morrison J, Carroll L, Twaddle S, Cameron I, Grimshaw J, Leyland A, et al. Pragmatic randomised controlled trial to evaluate guidelines for the management of infertility across the primary care-secondary care interface. *BMJ*. 2001;322(7297):1282-4.

94. Bennett K, Haggard M, Churchill R, Wood S. Improving referrals for glue ear from primary care: are multiple interventions better than one alone? *J Health Serv Res Policy*. 2001;6(3):139-44.
95. O'Donnell CA. Variation in GP referral rates: what can we learn from the literature? *Fam Pract*. 2000;17(6):462-71.
96. Gulati S, Jakola AS, Solheim O, Gabrielsen S, Godo PA, Skogstad B, et al. Assessment of referrals to a multidisciplinary outpatient clinic for patients with back pain. *The Journal of manual & manipulative therapy*. 2012;20(1):23-7.
97. Martinussen PE. Referral quality and the cooperation between hospital physicians and general practice: the role of physician and primary care factors. *Scandinavian journal of public health*. 2013;41(8):874-82.
98. Frydenberg K, Brekke M. Communication about drug use in referrals, acute admissions and discharge letters. *Tidsskrift for den Norske Lægeforening*. 2011;131(9-10):942-5.
99. Westerman RF, Hull FM, Bezemer PD, Gort G. A study of communication between general practitioners and specialists. *Br J Gen Pract*. 1990;40(340):445-9.
100. Cheng J, Beltran-Agullo L, Trope GE, Buys YM. Assessment of the quality of glaucoma referral letters based on a survey of glaucoma specialists and a glaucoma guideline. *Ophthalmology*. 2014;121(1):126-33.
101. DeAngelis AF, Chambers IG, Hall GM. The accuracy of medical history information in referral letters. *Aust Dent J*. 2010;55(2):188-92.
102. Scully ND, Chu L, Siriwardena D, Wormald R, Kotecha A. The quality of optometrists' referral letters for glaucoma. *Ophthalmic Physiol Opt*. 2009;29(1):26-31.
103. Jack C, Hazel E, Bernatsky S. Something's missing here: a look at the quality of rheumatology referral letters. *Rheumatol Int*. 2012;32(4):1083-5.
104. Berendsen AJ, Kuiken A, Benneker WH, Meyboom-de Jong B, Voorn TB, Schuling J. How do general practitioners and specialists value their mutual communication? A survey. *BMC Health Serv Res*. 2009;9:143.
105. Jenkins RM. Quality of general practitioner referrals to outpatient departments: assessment by specialists and a general practitioner. *Br J Gen Pract*. 1993;43(368):111-3.
106. Culshaw D, Clafferty R, Brown K. Let's get physical! A study of general practitioner's referral letters to general adult psychiatry--are physical examination and investigation results included? *Scottish medical journal*. 2008;53(1):7-8.
107. Davey CJ, Green C, Elliott DB. Assessment of referrals to the hospital eye service by optometrists and GPs in Bradford and Airedale. *Ophthalmic Physiol Opt*. 2011;31(1):23-8.
108. Gandhi TK, Sittig DF, Franklin M, Sussman AJ, Fairchild DG, Bates DW. Communication breakdown in the outpatient referral process. *J Gen Intern Med*. 2000;15(9):626-31.
109. Linne Y, Rossner S. What is "obesity"--an analysis of referral letters to an obesity unit. *International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity*. 1998;22(12):1231-3.
110. McCullagh M, O'Kelly P, Gilligan P. Referral letters to the emergency department: is the medication list accurate? *Irish medical journal*. 2015;108(2):38-40.
111. Bowling A, Redfern J. The process of outpatient referral and care: the experiences and views of patients, their general practitioners, and specialists. *Br J Gen Pract*. 2000;50(451):116-20.
112. Lohr RH, West CP, Beliveau M, Daniels PR, Nyman MA, Mundell WC, et al. Comparison of the quality of patient referrals from physicians, physician assistants, and nurse practitioners. *Mayo Clinic proceedings*. 2013;88(11):1266-71.
113. Mead GE, Cunningham AL, Faulkner S, Russell KJ, Ford MJ. Can general practitioner referral letters for acute medical admissions be improved? *Health bulletin*. 1999;57(4):257-61.
114. Azzi M, Constantino M, Pont L, McGill M, Twigg S, Krass I. Medication Safety: an audit of medication discrepancies in transferring type 2 diabetes mellitus (T2DM) patients from Australian primary care to tertiary ambulatory care. *Int J Qual Health Care*. 2014;26(4):397-403.
115. Ihuhua P, Pitcher RD. Is the devil in the detail? The quality and clinical impact of information provided on requests for non-trauma emergency abdominal CT scans. *Acta radiologica (Stockholm, Sweden : 1987)*. 2016.

116. Scully P, O'Donnell B, Peters C, O'Connor M, Lyons D. Older patient hospital admissions following primary care referral: the truth is in the referring. *Irish journal of medical science*. 2016;185(2):483-91.
117. McGoldrick C, Damkat-Thomas L, Lewis H. The impact of emergency nurse practitioners on referrals to a tertiary hand trauma service: a pilot of referral quality scoring system. *The Ulster medical journal*. 2011;80(1):19-20.
118. Jiwa M, Arnet H, Bulsara M, Ee HC, Harwood A. What is the importance of the referral letter in the patient journey? A pilot survey in Western Australia. *Qual Prim Care*. 2009;17(1):31-6.
119. Jiwa M, Meng X, O'Shea C, Magin P, Dadich A, Pillai V. Impact of referral letters on scheduling of hospital appointments: a randomised control trial. *Br J Gen Pract*. 2014;64(624):e419-25.
120. Froehlich F, Pache I, Burnand B, Vader JP, Fried M, Beglinger C, et al. Performance of panel-based criteria to evaluate the appropriateness of colonoscopy: a prospective study. *Gastrointest Endosc*. 1998;48(2):128-36.
121. Juillerat P, Peytremann-Bridevaux I, Vader JP, Arditi C, Schussele Fillietaz S, Dubois RW, et al. Appropriateness of colonoscopy in Europe (EPAGE II). Presentation of methodology, general results, and analysis of complications. *Endoscopy*. 2009;41(3):240-6.
122. Naylor CD. What is appropriate care? *The New England journal of medicine*. 1998;338(26):1918-20.
123. Brook RH, Chassin MR, Fink A, Solomon DH, Kosecoff J, Park RE. A method for the detailed assessment of the appropriateness of medical technologies. *International journal of technology assessment in health care*. 1986;2(1):53-63.
124. Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR, Lazaro P, et al. *The RAND/UCLA Appropriateness Method User's Manual*. Santa Monica, CA: RAND Corporation; 2001.
125. Arguello L, Pertejo V, Ponce M, Peiro S, Garrigues V, Ponce J. The appropriateness of colonoscopies at a teaching hospital: magnitude, associated factors, and comparison of EPAGE and EPAGE-II criteria. *Gastrointest Endosc*. 2012;75(1):138-45.
126. Andujar X, Sainz E, Gali A, Loras C, Aceituno M, Espinos JC, et al. [Inappropriateness rate for colonoscopy indications in an open access unit]. *Gastroenterol Hepatol*. 2015;38(5):313-9.
127. Balaguer F, Llach J, Castells A, Bordas JM, Ppellise M, Rodriguez-Moranta F, et al. The European panel on the appropriateness of gastrointestinal endoscopy guidelines colonoscopy in an open-access endoscopy unit: a prospective study. *Aliment Pharmacol Ther*. 2005;21(5):609-13.
128. Carrion S, Marin I, Lorenzo-Zuniga V, Moreno De Vega V, Boix J. [Appropriateness of colonoscopy indications according to the new EPAGE II criteria]. *Gastroenterol Hepatol*. 2010;33(7):484-9.
129. Chan TH, Goh KL. Appropriateness of colonoscopy using the ASGE guidelines: experience in a large Asian hospital. *Chin J Dig Dis*. 2006;7(1):24-32.
130. de Bosset V, Froehlich F, Rey JP, Thorens J, Schneider C, Wietlisbach V, et al. Do explicit appropriateness criteria enhance the diagnostic yield of colonoscopy? *Endoscopy*. 2002;34(5):360-8.
131. Gimeno Garcia AZ, Gonzalez Y, Quintero E, Nicolas-Perez D, Adrian Z, Romero R, et al. Clinical validation of the European Panel on the Appropriateness of Gastrointestinal Endoscopy (EPAGE) II criteria in an open-access unit: a prospective study. *Endoscopy*. 2012;44(1):32-7.
132. Gonvers JJ, Harris JK, Wietlisbach V, Burnand B, Vader JP, Froehlich F. A European view of diagnostic yield and appropriateness of colonoscopy. *Hepatogastroenterology*. 2007;54(75):729-35.
133. Grassini M, Verna C, Niola P, Navino M, Battaglia E, Bassotti G. Appropriateness of colonoscopy: diagnostic yield and safety in guidelines. *World J Gastroenterol*. 2007;13(12):1816-9; discussion 9.
134. Harris JK, Froehlich F, Gonvers JJ, Wietlisbach V, Burnand B, Vader JP. The appropriateness of colonoscopy: a multi-center, international, observational study. *Int J Qual Health Care*. 2007;19(3):150-7.
135. Excellence NfHaC. *Referral Guidelines for Suspected Cancer*. In: NICE, editor. London 2005.
136. Early DS, Ben-Menachem T, Decker GA, Evans JA, Fanelli RD, Fisher DA, et al. Appropriate use of GI endoscopy. *Gastrointest Endosc*. 2012;75(6):1127-31.
137. Vader JP, Froehlich F, Dubois RW, Beglinger C, Wietlisbach V, Pittet V, et al. European Panel on the Appropriateness of Gastrointestinal Endoscopy (EPAGE): conclusion and WWW site. *Endoscopy*. 1999;31(8):687-94.

138. Vader JP, Burnand B, Froehlich F, Dubois RW, Bochud M, Gonvers JJ. The European Panel on Appropriateness of Gastrointestinal Endoscopy (EPAGE): project and methods. *Endoscopy*. 1999;31(8):572-8.
139. Schussele Fillietaz S, Gonvers JJ, Peytremann-Bridevaux I, Arditi C, Delvaux M, Numans ME, et al. Appropriateness of colonoscopy in Europe (EPAGE II). Functional bowel disorders: pain, constipation and bloating. *Endoscopy*. 2009;41(3):234-9.
140. Peytremann-Bridevaux I, Arditi C, Froehlich F, O'Malley J, Fairclough P, Le Moine O, et al. Appropriateness of colonoscopy in Europe (EPAGE II). Iron-deficiency anemia and hematochezia. *Endoscopy*. 2009;41(3):227-33.
141. Schussele Fillietaz S, Juillerat P, Burnand B, Arditi C, Windsor A, Beglinger C, et al. Appropriateness of colonoscopy in Europe (EPAGE II). Chronic diarrhea and known inflammatory bowel disease. *Endoscopy*. 2009;41(3):218-26.
142. Arditi C, Gonvers JJ, Burnand B, Minoli G, Oertli D, Lacaine F, et al. Appropriateness of colonoscopy in Europe (EPAGE II). Surveillance after polypectomy and after resection of colorectal cancer. *Endoscopy*. 2009;41(3):209-17.
143. Arditi C, Peytremann-Bridevaux I, Burnand B, Eckardt VF, Bytzer P, Agreus L, et al. Appropriateness of colonoscopy in Europe (EPAGE II). Screening for colorectal cancer. *Endoscopy*. 2009;41(3):200-8.
144. European Panel for Appropriateness in Gastrointestinal Endoscopy. EPAGE II guideline: European Panel of Appropriateness in Gastrointestinal Endoscopy 2008 [Available from: www.epage.ch].
145. Veloski J, Tai S, Evans AS, Nash DB. Clinical vignette-based surveys: a tool for assessing physician practice variation. *Am J Med Qual*. 2005;20(3):151-7.
146. Pham T, Roy C, Mariette X, Liote F, Durieux P, Ravaud P. Effect of response format for clinical vignettes on reporting quality of physician practice. *BMC Health Serv Res*. 2009;9:128.
147. Jiwa M, Gordon M, Arnet H, Ee H, Bulsara M, Colwell B. Referring patients to specialists: a structured vignette survey of Australian and British GPs. *BMC Fam Pract*. 2008;9:2.
148. Peabody JW, Luck J, Glassman P, Jain S, Hansen J, Spell M, et al. Measuring the quality of physician practice by using clinical vignettes: a prospective validation study. *Ann Intern Med*. 2004;141(10):771-80.
149. Landon BE, Reschovsky J, Reed M, Blumenthal D. Personal, organizational, and market level influences on physicians' practice patterns: results of a national survey of primary care physicians. *Med Care*. 2001;39(8):889-905.
150. Badger LW, deGruy F, Hartman J, Plant MA, Leeper J, Ficken R, et al. Stability of standardized patients' performance in a study of clinical decision making. *Family medicine*. 1995;27(2):126-31.
151. Colliver JA, Swartz MH. Assessing clinical performance with standardized patients. *Jama*. 1997;278(9):790-1.
152. McLeod PJ, Tamblyn RM, Gayton D, Grad R, Snell L, Berkson L, et al. Use of standardized patients to assess between-physician variations in resource utilization. *Jama*. 1997;278(14):1164-8.
153. Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *JAMA*. 2000;283(13):1715-22.
154. Glassman PA, Luck J, O'Gara EM, Peabody JW. Using standardized patients to measure quality: evidence from the literature and a prospective study. *Jt Comm J Qual Improv*. 2000;26(11):644-53.
155. Tiemeier H, de Vries WJ, van het Loo M, Kahan JP, Klazinga N, Grol R, et al. Guideline adherence rates and interprofessional variation in a vignette study of depression. *Qual Saf Health Care*. 2002;11(3):214-8.
156. Skaner Y, Bring J, Ullman B, Strender LE. Heart failure diagnosis in primary health care: clinical characteristics of problematic patients. A clinical judgement analysis study. *BMC Fam Pract*. 2003;4:12.
157. Skaner Y, Backlund L, Montgomery H, Bring J, Strender LE. General practitioners' reasoning when considering the diagnosis heart failure: a think-aloud study. *BMC Fam Pract*. 2005;6(1):4.
158. 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*. 2012.

159. Quine MA, Bell GD, McCloy RF, Devlin HB, Hopkins A. Appropriate use of upper gastrointestinal endoscopy--a prospective audit. Steering Group of the Upper Gastrointestinal Endoscopy Audit Committee. *Gut*. 1994;35(9):1209-14.
160. Bachmann LM, Muhleisen A, Bock A, ter Riet G, Held U, Kessels AG. Vignette studies of medical choice and judgement to study caregivers' medical decision behaviour: systematic review. *BMC Med Res Methodol*. 2008;8:50.
161. Englund L, Tibblin G, Svardsudd K. Variations in sick-listing practice among male and female physicians of different specialties based on case vignettes. *Scand J Prim Health Care*. 2000;18(1):48-52.
162. European Panel for Appropriateness in Gastrointestinal Endoscopy. EPAGE guideline: European Panel for Appropriateness in Gastrointestinal Endoscopy; 1999 [Available from: http://www.epage.ch/EPAGE_ADM/EPAGE1/bonjour.html].
163. Norwegian Electronic Medical Handbook for Doctors. Trondheim: Norsk Helseinformatikk AS; 2013 [Available from: <http://legehandboka.no/>].
164. UpToDate. UpToDate in Waltham, MA; 2013 [Available from: <http://www.uptodate.com/home>].
165. Cronbach LJ, Meehl PE. Construct validity in psychological tests. *Psychological Bulletin*. 1955;52(4):281-302.
166. Trochim WM. The Research Methods Knowledge Base 2006 [updated October 20, 2006 2. edition:[Available from: <http://www.socialresearchmethods.net/kb/>].
167. Fleiss JL. Design and analysis of clinical experiments: John Wiley & Sons; 2011.
168. Altman D. Practical Statistics for Medical Research. London: Chapman and Hall; 1991.
169. Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi medical journal : the journal of Medical Association of Malawi*. 2012;24(3):69-71.
170. Delgado-Rodriguez M, Llorca J. Bias. *Journal of epidemiology and community health*. 2004;58(8):635-41.
171. Stoller JK. Quality-assurance research: studying processes of care. *Respiratory care*. 2004;49(10):1175-80.
172. The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions http://handbook.cochrane.org/front_page.htm2011 [updated March 2011. Version 5.1.0:]
173. Shekelle PG, Kahan JP, Bernstein SJ, Leape LL, Kamberg CJ, Park RE. The reproducibility of a method to identify the overuse and underuse of medical procedures. *The New England journal of medicine*. 1998;338(26):1888-95.
174. Lomas J, Anderson G, Enkin M, Vayda E, Roberts R, MacKinnon B. The role of evidence in the consensus process. Results from a Canadian consensus exercise. *Jama*. 1988;259(20):3001-5.
175. Ayanian JZ, Landrum MB, Normand SL, Guadagnoli E, McNeil BJ. Rating the appropriateness of coronary angiography--do practicing physicians agree with an expert panel and with each other? *The New England journal of medicine*. 1998;338(26):1896-904.
176. Hartveit M, Aslaksen A, Vanhaecht K, Thorsen O, Hove O, Haug K, et al. Development and testing of an instrument in Western Norway to measure the quality of referral information from primary care to specialised mental health care. *International Journal of Care Coordination*. 2015;18(2-3):35-47.
177. Oremus M, Oremus C, Hall GBC, McKinnon MC, ECT, Team CSR. Inter-rater and test-retest reliability of quality assessments by novice student raters using the Jadad and Newcastle-Ottawa Scales. *BMJ Open*. 2012;2(4).
178. Huskisson EC. Measurement of pain. *Lancet*. 1974;2(7889):1127-31.
179. Olsen S, Nolan MF, Kori S. Pain measurement. An overview of two commonly used methods. *Anesthesiology review*. 1992;19(6):11-5.
180. Larsen S, Aabakken L, Lillevold PE, Osnes M. Assessing Soft Data in Clinical Trials. *Pharmaceutical Medicine*. 1991(5):29-36.
181. Boyer L, Baumstarck-Barrau K, Belzeaux R, Azorin JM, Chabannes JM, Dassa D, et al. Validation of a professionals' satisfaction questionnaire with electronic medical records (PSQ-EMR) in psychiatry. *European psychiatry : the journal of the Association of European Psychiatrists*. 2011;26(2):78-84.
182. Bodger K, Ormerod C, Shackcloth D, Harrison M. Development and validation of a rapid, generic measure of disease control from the patient's perspective: the IBD-Control questionnaire. *Gut*. 2013.

183. Bengtsson M, Ohlsson B, Ulander K. Development and psychometric testing of the Visual Analogue Scale for Irritable Bowel Syndrome (VAS-IBS). *BMC Gastroenterol*. 2007;7:16.
184. Aabakken L, Larsen S, Osnes M. Visual analogue scales for endoscopic evaluation of nonsteroidal anti-inflammatory drug-induced mucosal damage in the stomach and duodenum. *Scand J Gastroenterol*. 1990;25(5):443-8.
185. Singer AJ, Thode HC, Jr. Determination of the minimal clinically significant difference on a patient visual analog satisfaction scale. *Academic emergency medicine : official journal of the Society for Academic Emergency Medicine*. 1998;5(10):1007-11.
186. de Lange T, Larsen S, Aabakken L. Inter-observer agreement in the assessment of endoscopic findings in ulcerative colitis. *BMC Gastroenterol*. 2004;4:9.
187. Hjermstad MJ, Fayers PM, Haugen DF, Caraceni A, Hanks GW, Loge JH, et al. Studies comparing Numerical Rating Scales, Verbal Rating Scales, and Visual Analogue Scales for assessment of pain intensity in adults: a systematic literature review. *Journal of pain and symptom management*. 2011;41(6):1073-93.
188. Wahlberg H, Valle PC, Malm S, Hovde O, Broderstad AR. The effect of referral templates on out-patient quality of care in a hospital setting: a cluster randomized controlled trial. *BMC Health Serv Res*. 2017;17(1):177.
189. Eide TJ. Risk of colorectal cancer in adenoma-bearing individuals within a defined population. *International journal of cancer*. 1986;38(2):173-6.
190. Vleugels JL, Hazewinkel Y, Fockens P, Dekker E. The Natural History of Diminutive and Small Colorectal Polyps: a Systematic Literature Review. *Gastrointest Endosc*. 2016.
191. Hoff G, Foerster A, Vatn MH, Sauar J, Larsen S. Epidemiology of polyps in the rectum and colon. Recovery and evaluation of unresected polyps 2 years after detection. *Scand J Gastroenterol*. 1986;21(7):853-62.
192. Hofstad B, Vatn M, Larsen S, Osnes M. Growth of colorectal polyps: recovery and evaluation of unresected polyps of less than 10 mm, 1 year after detection. *Scand J Gastroenterol*. 1994;29(7):640-5.
193. Johnson RM, Orme BK. How Many Questions Should You Ask In Choice-Based Conjoint Studies? *ART Forum*, Beaver Creek; 1996.
194. Louviere J, Eagle T, Cohen S. Conjoint analysis: Methods, myths and much more. Center for the Study of Choice, Working Paper. 2005(05-001).
195. Morrell DC, Roland MO. Analysis of referral behaviour: responses to simulated case histories may not reflect real clinical behaviour. *Br J Gen Pract*. 1990;40(334):182-5.
196. Hartley RM, Charlton JR, Jarman B, Harris CM. Case history questionnaires in the study of doctors' use of resources. Are they measuring what we want? *Med Care*. 1985;23(10):1163-70.
197. Bediang G, Bagayoko CO, Raetzo MA, Geissbuhler A. Relevance and usability of a computerized patient simulator for continuous medical education of isolated care professionals in sub-saharan Africa. *Stud Health Technol Inform*. 2011;169:666-70.
198. Perron NJ, Perneger T, Kolly V, Dao MD, Sommer J, Hudelson P. Use of a computer-based simulated consultation tool to assess whether doctors explore sociocultural factors during patient evaluation. *J Eval Clin Pract*. 2009;15(6):1190-5.
199. Sijstermans R, Jaspers MW, Bloemendaal PM, Schoonderwaldt EM. Training inter-physician communication using the Dynamic Patient Simulator. *Int J Med Inform*. 2007;76(5-6):336-43.
200. Zary N, Johnson G, Boberg J, Fors UG. Development, implementation and pilot evaluation of a Web-based Virtual Patient Case Simulation environment-Web-SP. *BMC medical education*. 2006;6:10.
201. Weller JM. Simulation in undergraduate medical education: bridging the gap between theory and practice. *Medical education*. 2004;38(1):32-8.
202. Newton J, Eccles M, Hutchinson A. Communication between general practitioners and consultants: what should their letters contain? *Bmj*. 1992;304(6830):821-4.
203. Aabakken L, Barkun AN, Cotton PB, Fedorov E, Fujino MA, Ivanova E, et al. Standardized endoscopic reporting. *J Gastroenterol Hepatol*. 2014;29(2):234-40.
204. Marzo-Castillejo M, Almeda J, Mascort JJ, Cunillera O, Saladich R, Nieto R, et al. Appropriateness of colonoscopy requests according to EPAGE-II in the Spanish region of Catalonia. *BMC Fam Pract*. 2015;16(1):154.

205. Hassan C, Di Giulio E, Marmo R, Zullo A, Annibale B. Appropriateness of the indication for colonoscopy: systematic review and meta-analysis. *J Gastrointest Liver Dis.* 2011;20(3):279-86.
206. de Lange T, Randel K, Schult A, Knudsen M, Kirkøen B, Botteri E, et al. Sigmoidoskopi og testing for blod i avføringen - en sammenlignende screeningstudie. *Tidsskrift Norske Legeforening*- accepted for publication. 2017.
207. Mangualde J, Cremers MI, Vieira AM, Freire R, Gamito E, Lobato C, et al. Appropriateness of outpatient gastrointestinal endoscopy in a non-academic hospital. *World J Gastrointest Endosc.* 2011;3(10):195-200.
208. Suriani R, Rizzetto M, Mazzucco D, Grosso S, Gastaldi P, Marino M, et al. Appropriateness of colonoscopy in a digestive endoscopy unit: a prospective study using ASGE guidelines. *J Eval Clin Pract.* 2009;15(1):41-5.
209. Aas E. Kostnader ved screening av tykk- og endetarmskreft – basert på tall fra piloten i Østfold og Bærum https://www.kreftregisteret.no/globalassets/publikasjoner-og-rapporter/tarmkreftscreening/kostnadsanalyse_pilot_screening_tykk-endetarmskreft.pdf: Cancer Registry of Norway; 2017 [
210. Peery AF, Keku TO, Martin CF, Eluri S, Runge T, Galanko JA, et al. Distribution and Characteristics of Colonic Diverticula in a United States Screening Population. *Clin Gastroenterol Hepatol.* 2016;14(7):980-5.e1.
211. Petruzzello L, Hassan C, Alvaro D, Kohn A, Rossi Z, Zullo A, et al. Appropriateness of the indication for colonoscopy: is the endoscopist the 'gold standard'? *J Clin Gastroenterol.* 2012;46(7):590-4.
212. Grassini M, Verna C, Battaglia E, Niola P, Navino M, Bassotti G. Education improves colonoscopy appropriateness. *Gastrointest Endosc.* 2008;67(1):88-93.
213. Debnath D, Dielehner N, Gunning KA. Guidelines, compliance, and effectiveness: a 12 months' audit in an acute district general healthcare trust on the two week rule for suspected colorectal cancer. *Postgraduate medical journal.* 2002;78(926):748-51.
214. Eccersley AJ, Wilson EM, Makris A, Novell JR. Referral guidelines for colorectal cancer--do they work? *Annals of the Royal College of Surgeons of England.* 2003;85(2):107-10.
215. Cardin F, Zorzi M, Terranova O. Implementation of a guideline versus use of individual prognostic factors to prioritize waiting lists for upper gastrointestinal endoscopy. *Eur J Gastroenterol Hepatol.* 2007;19(7):549-53.
216. O'Brien K, Wright J, Conboy F, Bagley L, Lewis D, Read M, et al. The effect of orthodontic referral guidelines: a randomised controlled trial. *Br Dent J.* 2000;188(7):392-7.
217. Heimly V. Experience report from the implementation of interactive referral at AHUS(Akershus University Hospital). . Trondheim: NTNU; 2011.
218. Jiwa M, Skinner P, Coker AO, Shaw L, Campbell MJ, Thompson J. Implementing referral guidelines: lessons from a negative outcome cluster randomised factorial trial in general practice. *BMC Fam Pract.* 2006;7:65.
219. Jiwa M, Dhaliwal S. Referral Writer: preliminary evidence for the value of comprehensive referral letters. *Qual Prim Care.* 2012;20(1):39-45.
220. Tay SH, Lim AY, Lee TL, Low BP, Cheung PP. The value of referral letter information in predicting inflammatory arthritis--factors important for effective triaging. *Clinical rheumatology.* 2014;33(3):409-13.
221. Sales M, Quain D, Lasserson D, Levi C, Oldmeadow C, Jiwa M, et al. Quality of Referrals and Guideline Compliance for Time to Consultation at an Acute Neurovascular Clinic. *Journal of stroke and cerebrovascular diseases : the official journal of National Stroke Association.* 2015.
222. Mitchell E, Macdonald S, Campbell NC, Weller D, Macleod U. Influences on pre-hospital delay in the diagnosis of colorectal cancer: a systematic review. *British journal of cancer.* 2008;98(1):60-70.

List of errata

- Numbering of figures: Throughout the thesis the figures were wrongly numbered. This has been corrected.
- P 36, line 26, last line of table. 6.6 (6.2-6.8) has been changed to 6.5 (6.2-6.8).
- P 36, line 27(footnote of table): “paired sample t-test” has been replaced with “paired sample t-test for NRS scores”.

Papers I-IV

BMJ Open First quality score for referral letters in gastroenterology – a validation study

Sigrun Losada Eskeland,¹ Cathrine Brunborg,² Birgitte Seip,³ Kristine Wiencke,⁴ Øistein Hovde,^{5,6} Tanja Owen,⁷ Erik Skogestad,⁸ Gert Huppertz-Hauss,⁹ Fred-Arne Halvorsen,¹⁰ Kjetil Garborg,^{11,12} Lars Aabakken,¹³ Thomas de Lange^{4,14}

To cite: Eskeland SL, Brunborg C, Seip B, *et al.* First quality score for referral letters in gastroenterology—a validation study. *BMJ Open* 2016;**6**:e012835. doi:10.1136/bmjopen-2016-012835

► Prepublication history and additional material is available. To view please visit the journal (<http://dx.doi.org/10.1136/bmjopen-2016-012835>).

Received 26 May 2016
Revised 18 August 2016
Accepted 1 September 2016

ABSTRACT

Objective: To create and validate an objective and reliable score to assess referral quality in gastroenterology.

Design: An observational multicentre study.

Setting and participants: 25 gastroenterologists participated in selecting variables for a Thirty Point Score (TPS) for quality assessment of referrals to gastroenterology specialist healthcare for 9 common indications. From May to September 2014, 7 hospitals from the South-Eastern Norway Regional Health Authority participated in collecting and scoring 327 referrals to a gastroenterologist.

Main outcome measure: Correlation between the TPS and a visual analogue scale (VAS) for referral quality.

Results: The 327 referrals had an average TPS of 13.2 (range 1–25) and an average VAS of 4.7 (range 0.2–9.5). The reliability of the score was excellent, with an intra-rater intraclass correlation coefficient (ICC) of 0.87 and inter-rater ICC of 0.91. The overall correlation between the TPS and the VAS was moderate ($r=0.42$), and ranged from fair to substantial for the various indications. Mean agreement was good (ICC=0.47, 95% CI (0.34 to 0.57)), ranging from poor to good.

Conclusions: The TPS is reliable, objective and shows good agreement with the subjective VAS. The score may be a useful tool for assessing referral quality in gastroenterology, particularly important when evaluating the effect of interventions to improve referral quality.

Strengths and limitations of this study

- Multicentre design gives better external validity.
- Wide variety of gastrointestinal indications encompassed by score.
- Score does not assess appropriateness of the indication for the referral.
- Potential interindividual differences in the use of the visual analogue scale.

secondary healthcare services, which may result in delayed diagnosis and a poorer prognosis. It is well documented that referral letters of low quality are prevalent.^{5–21}

In Norway, referral letters are sent at the GP's discretion and generated by using free text in a standard template created by the Norwegian health authorities²² (table 1). The template includes the urgency of the referral, but does not include symptom-specific or indication-specific criteria. It is very likely that specific symptoms and clinical findings are crucial for deciding the urgency of the referral.²³ The GPs optionally add relevant information about supplementary workup, like laboratory tests and imaging results, by copying and pasting these results to the referral text. Subsequently, the referral is transferred electronically by a secure system to the hospital.²⁴ At arrival, a consultant assesses the referral letter and prioritises the urgency of the referral. If the consultant considers the referral indication inappropriate, it may be rejected. However, the legislation prohibits rejection due to poor quality.

The Norwegian Prioritization Guideline for gastroenterology (NPGg),²⁵ created by an expert group of gastroenterologists, states nine main indications for referral to gastroenterology services (open access endoscopy and consultations); dyspepsia, dysphagia, chronic abdominal pain, change of bowel habit, diarrhoea, constipation, gastrointestinal bleeding, weight loss and jaundice/elevated liver enzymes. The guidelines aim to

INTRODUCTION

General practitioners (GPs) refer numerous patients to specialised healthcare services every year, and the referral rates are increasing worldwide.^{1–2} A recent study revealed a referral rate of 13.7% with great variations between GPs.³ The referral letter is a key document for the communication between the GPs and the hospital consultants,^{4–5} and its content and quality are essential for the scheduling and prioritisation of patients. Incomplete, erroneous or extensive referral letters may delay the admission of patients to



CrossMark

For numbered affiliations see end of article.

Correspondence to

Sigrun Losada Eskeland;
sigesk@vestreviken.no

Table 1 General requirements for referral letters, according to the Norwegian national referral guidelines, including the frequency of reporting in this study

Referral information	N (%)
Administrative information (patient, GP, referral institution)	325 (99.4)
Referral diagnosis	319 (97.6)
Type of referral (examination, workup, advice)/expected action	313 (95.7)
Urgency of referral	58 (17.7)
Critical information	68 (20.8)
Allergies/infectious disease	66 (20.2)
Other ongoing critical diseases and treatments	6 (1.8)
Personal information	181 (55.4)
Family history of disease	94 (28.7)
Social history (work, school, family, living)	122 (37.3)
Alcohol and tobacco history	60 (18.3)
Previous medical history	259 (79.2)
History of the current disease	327 (100.0)
Findings	267 (81.7)
Clinical examination	164 (50.2)
Laboratory workup	173 (53.9)
Radiology/other	119 (36.4)
Current medicines	271 (82.9)
Patient informed of referral	0
Total referral letters	327

GP, general practitioner.

cover 75–80% of gastroenterology referrals. Bowel cancer screening and polyp surveillance is not a part of the indications because this is taken care of either in the hospital system or by population-based trials for bowel cancer screening.

Despite the well-documented quality issues, no validated and objective tool to assess the quality of referral letters to the specialist healthcare services exists.

An objective, standardised score would probably be beneficial to ease the assessment of referral quality and to evaluate interventions intending to improve the quality of referral letters.

The aim of the study was to develop an objective, relevant and reliable quality score for the evaluation of referral letters for the most common indications for referral to gastroenterologists, and to validate the score in a clinical setting.

MATERIALS AND METHODS

Development of the score

We used the nine core indications for referral to gastroenterologists specified in the NPGg:²⁵ dyspepsia, dysphagia, chronic abdominal pain, change of bowel habit, diarrhoea, constipation, gastrointestinal bleeding, weight loss and jaundice/elevated liver enzymes.

For each of these nine indications, we created a list of 29–36 relevant medical variables based on UpToDate,²⁶ the Norwegian Electronic Medical Handbook for

Doctors (NEL)²⁷ and the NPGg.²⁵ Administrative information including sex and date of birth of the patient is included in referral letters by default, as well as name, health personnel identification number and address of the referring physician, and therefore we did not include this information in the nine lists.

Between November 2013 and March 2014, we invited 39 gastroenterologists within the South-Eastern Norway Health Region to participate in a web-based study. They recorded demographic information such as sex, age, experience and workplace before they started the survey. For each of the nine indications, we asked the participating gastroenterologists to select the 15 most important variables for assessing and prioritising referrals. The importance of the variables was categorised from three points (most important) to one point (less important) with five variables in each category. The remaining variables were given zero points. We then summarised the points assigned to the individual variables for each indication from all the returned questionnaires, and selected the 15 variables with the highest sum of points to comprise the final variables in the score. The five highest rated variables were classified with three points, the next five with two points and the last five with one point.

After a period of a minimum 6 weeks, we repeated this process to check for reliability of the values assigned. Only the results from the first round were used in the final score. This resulted in a symptom-specific Thirty Point Score (TPS).

Validation of the score

Between May and September 2014, seven primary gastroenterology referral centres in South-Eastern Norway Health Region collected 327 referral letters, 21–50 for each of the nine indications stated in the NPGg.²⁵ Patients were mostly referred for open access endoscopy as well as consultations.

One or two gastroenterologists in each participating centre collected and assessed consecutive referrals within the nine indications. They rated the quality of the referral letter on a 10 cm visual analogue scale (VAS), where 0 cm indicated the worst possible quality and 10 cm the best possible quality. We chose VAS as a comparator to the TPS to assess the external validity in the validation process, assuming that referral letters containing all essential information for assessment and prioritisation of the referred patients would also yield a high VAS.²⁸ The gastroenterologists also recorded whether the referral letters contained too much or too little information, and whether they were unstructured or illegible. Patient age and gender were recorded before all patient data were removed from the referral letters and they were handed over to the study team.

Subsequently, one researcher from the study team assessed all the referral letters according to the TPS. Both the presence and absence of signs and symptoms were given equal points, as long as they were reported

adequately in the referral letter. We also recorded demographic data about the referring physicians (sex, age, county, size of patient list, etc) from publicly available records, and whether the referral letter complied with the national guidelines for referral letters to the specialist healthcare services.²²

Information that was not available to the consultant through the referral letter at the time of assessing the referral, for example, results from laboratory tests and radiology that were communicated later in the referral process, was not included when calculating the TPS.

To check for intra-rater and inter-rater reliability, we randomly selected 25% of the referral letters from each indication and reassessed them with the TPS. This was done after a minimum of 6 months by the same researcher from the research team and a second independent gastroenterologist. These ratings were done completely blinded from each other and from the results of the first rating.

Statistics

We present descriptive statistics as means with their 95% CI or as proportions. We use intraclass correlation coefficient (ICC) for continuous measures and weighted κ statistics for categorical measures to describe reliability of the gastroenterologists' selection of score variables. ICC for average measurements was used to calculate intra-rater and inter-rater reliability of the TPS assessment of referral letters. ICC values were interpreted as: >0.75 =excellent, 0.40 – 0.75 =fair to good and <0.40 =poor.²⁹ The κ values were interpreted as: >0.80 =very good, 0.61 – 0.80 =good, 0.41 – 0.60 =moderate, 0.21 – 0.40 =fair and <0.21 =poor.³⁰ Univariable and multivariable linear regression analysis was performed to determine whether patient-related or doctor-related factors were associated with changes in quality of the referral letter, using a manual backward elimination procedure. Any variable with a $p < 0.25$ from the univariable analysis was considered a candidate for the multivariable model. We used the Pearson correlation coefficient to assess the correlation between the VAS and the TPS and ICC for average measurements to assess agreement between the two measurements. We assessed any differences in the VAS/TPS between the different centres by using one-way analysis of variance (ANOVA). The significance level was set at 0.05. All calculations were performed using the IBM SPSS V.21 (IBM SPSS, Chicago, Illinois, USA).

Power estimation

Sample size estimation was performed to investigate the association between the VAS and the TPS. The correlation coefficient (r) was anticipated to be 0.60. Considering a probability of a type I error of 5%, 80% power and a two-sided test, at least 19 referral letters were required for each indication. Thus, during the inclusion time, a minimum of 21 referral letters were collected for each indication. Since some indications

were more common than others, the number of referral letters for each indication varies upwards.

RESULTS

Selection of TPS variables

Of the 39 invited gastroenterologists, 32 started to record their demographic data, and 26 (81.3%) also moved on to select the variables for the score. Twenty-five (64.1%) of the 39 gastroenterologists completed the whole survey and were included in the study. The excluded gastroenterologist provided answers for one single indication and then aborted the questionnaire. The reason for the dropouts cannot be determined due to the study format.

The mean age of the included gastroenterologists was 48.5 years (range 35–69). Mean experience as a licensed gastroenterologist was 9.6 years (range 1–33). Sixteen gastroenterologists (64%) repeated the questionnaire 6 weeks later to test for reliability. The characteristics of these gastroenterologists were not significantly different from the nine who did not repeat the survey. The ICC for the reliability of the sum of the scores for the variables was excellent (0.88 to 0.93) for all the indications. The κ values for the reliability of final scores showed a good to very good agreement in all of the indications. The resulting TPS for all nine indications is presented in online supplementary appendix 1. It consists of 15 items for each indication. Depending on the value of the item for the quality of the referral, it is awarded with 1, 2 or 3 points if described adequately in the referral letter. The maximum score for a referral is 30 points, indicating a high-quality referral letter.

Validation of the score

The referring physicians were on average 47.1 years old (range 26–72), 37.3% were female and 95% were GPs. The referred patients were 62.4% female, and the average age was 57.2 years (range 6–94).

Adherence to the Norwegian referral guidelines varied substantially, as shown in table 1.

In particular, information regarding allergies/critical information, family history of disease and alcohol/tobacco consumption was sparse.

The average quality of the referral letters assessed by VAS was 4.7 (95% CI 4.5 to 5.0, range 0.2–9.5). The mean TPS for all referral letters was 13.2 (95% CI 12.8 to 13.8, range 1–25). In total, 54.1% of the referral letters had a VAS below 5 (out of 10) and 59.6% had a TPS below 15 (out of 30).

Intra-rater and inter-rater reliability of the TPS for scoring referral letters was excellent (ICC=0.87 (95% CI 0.81 to 0.92) and 0.91 (95% CI 0.86 to 0.94), respectively).

The average VAS and TPS for the nine indications are shown in table 2 together with the correlation between the two scores. The VAS and the TPS showed a moderate overall correlation ($r=0.42$; figure 1),

Table 2 Referral information quality assessed by VAS and TPS, and correlation between them

Indication	N referral letters (%)	Mean TPS (95% CI)	Mean VAS (95% CI)	Correlation coefficient* (95% CI)	ICC† average measures (95% CI)
Abdominal pain	50 (15.3)	12.5 (11.0 to 14.1)	4.5 (3.9 to 5.2)	0.46	0.49 (0.09 to 0.71)
Dyspepsia	47 (14.4)	11.9 (10.6 to 13.1)	4.3 (3.7 to 4.9)	0.25	0.33 (−0.02 to 0.63)
Gastrointestinal bleeding	34 (10.4)	15.7 (14.1 to 17.3)	5.1 (4.4 to 5.9)	0.46	0.51 (0.03 to 0.76)
Change of bowel habit	48 (14.7)	14.9 (13.5 to 16.3)	5.1 (4.3 to 5.8)	0.60	0.66 (0.40 to 0.81)
Diarrhoea	38 (11.6)	11.3 (9.9 to 12.7)	4.6 (4.0 to 5.2)	0.37	0.42 (−0.11 to 0.70)
Dysphagia	36 (11.0)	11.9 (10.3 to 13.5)	5.0 (4.3 to 5.8)	0.29	0.37 (−0.24 to 0.68)
Constipation	27 (8.3)	13.5 (12.1 to 14.9)	4.5 (3.8 to 5.1)	0.24	0.30 (−0.55 to 0.68)
Weight loss	21 (6.4)	14.9 (12.5 to 17.2)	4.8 (3.9 to 5.8)	0.63	0.62 (0.06 to 0.85)
Jaundice	26 (8.0)	14.3 (12.4 to 16.2)	4.9 (3.9 to 5.8)	0.24	0.32 (−0.52 to 0.70)
Total	327	13.2 (12.8 to 13.8)	4.7 (4.5 to 5.0)	0.42	0.47 (0.34 to 0.57)

*Pearson correlation coefficient interpretation: 0–0.2=slight, 0.2–0.4=fair, 0.4–0.6=moderate, 0.6–0.8=substantial, 0.8–1.0=almost perfect.

†ICC interpretation: >0.75=excellent, 0.40–0.75=fair to good and <0.40=poor.

ICC, intraclass correlation coefficient; TPS, Thirty Point Score; VAS, visual analogue scale.

ranging from fair to substantial ($r=0.24$ to 0.63) for the different indications.

A multiple linear regression analysis with a manual backward elimination procedure showed that age and gender were the only patient-related or doctor-related variables associated with TPS (age: $\beta_{\text{adj}}=-1.156$, 95% CI -2.24 to (-0.076) , $p=0.036$; gender: $\beta_{\text{adj}}=-0.090$ (95% CI -0.136 to (-0.043) , $p<0.001$), explaining 7% of the variance of TPS ($r^2=0.07$). Further, gender was identified as the only variable associated with VAS ($\beta_{\text{adj}}=-0.513$, 95% CI -0.993 to (-0.033) , $p=0.036$), explaining 1% of the variance of VAS ($r^2=0.01$).

When the gastroenterologist had recorded that the referral letter contained too little information ($n=167$ (51.1%)), the VAS and the TPS were also significantly lower (mean difference (Δ)= 1.7 , $p<0.001$ and $\Delta=3.4$, $p<0.001$, respectively). When the gastroenterologists had recorded that the referral letter was unstructured ($n=60$ (18.3%)), the VAS was significantly lower ($\Delta=1.6$, $p<0.001$), but the TPS was unaffected ($\Delta=-0.4$, $p=0.51$).

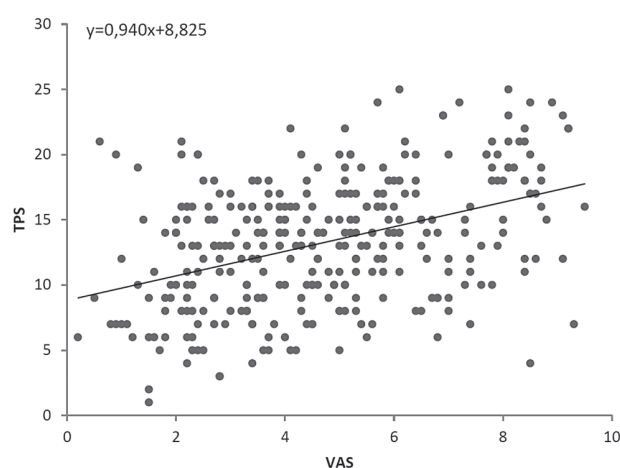
There were significant differences in the TPS and the VAS between the centres, and this difference was confirmed by the ANOVA analysis for the TPS ($p<0.001$) and the VAS ($p=0.004$). For the TPS, this significant difference disappeared by removing centre III from the calculations. The difference in the VAS disappeared by removing centre II (table 3).

Some analyses were made to identify factors that could increase the correlation between the VAS and the TPS. Eliminating the one-point items did not improve the correlation. Neither did adjusting for the number of three-point items in the referral.

DISCUSSION

Overview and principal findings

This is, to the best of our knowledge, the first study to develop an objective, reliable and validated score (TPS)

**Figure 1** Correlation between TPS and VAS

to assess the quality of gastroenterology referral letters, and it may work as a model for other medical specialties. The score is useful for the majority of referrals in gastroenterology, regardless of the seriousness of the condition or the location of the disease.

The TPS has demonstrated an excellent intra-rater and inter-rater reliability as well as a moderate correlation between the TPS and a subjective VAS score assigned by gastroenterologists. The quality of the referral letters was variable, both assessed by the TPS and the VAS.

The correlation and agreement between the TPS and the VAS was somewhat lower than expected ($r=0.42$, $\text{ICC}=0.47$). Factors not captured by the TPS may also influence the subjective assessment of the quality of referral letters, such as lack of structure or appropriateness. Such factors may have negatively influenced the correlation between the two measurements. Unstructured letters had a lower VAS despite adequate content according to the TPS.

Table 3 Relationship between TPS and VAS by centre

Centre	N (%)	Mean TPS (95% CI)	p Value*	Mean VAS (95% CI)	p Value*
Centre I	45	14.2 (12.7 to 15.7)		4.5 (4.0 to 5.1)	
Centre II	42	10.9 (9.5 to 12.4)		4.7 (4.0 to 5.4)	
Centre III	24	12.0 (10.1 to 14.2)		6.3 (5.4 to 7.2)	
Centre IV	46	13.0 (11.7 to 14.2)		5.5 (5.2 to 5.9)	
Centre V	67	13.9 (12.9 to 15.0)		4.2 (3.8 to 4.6)	
Centre VI	50	14.7 (13.2 to 16.3)		4.4 (3.6 to 5.2)	
Centre VII	53	13.1 (11.8 to 14.3)		4.6 (4.0 to 5.3)	
Total	327	13.3 (12.8 to 13.8)	0.004	4.8 (4.5 to 5.0)	<0.001

*ANOVA.

ANOVA, analysis of variance; TPS, Thirty Point Score; VAS, visual analogue scale.

There were significant differences between the centres both for the mean TPS and the mean VAS. Differences in quality of referral letters in different geographical regions have in other studies been explained by GP workload, referral culture or capacity in local nursing and care institutions, and some of these factors may also be present in this study.²³

The only patient-related or physician-related factor associated with a change in quality of the referral letters in this study was increasing age and male sex of the referring physician, both leading to small decreases in the TPS, but the changes are minor and most likely not clinically relevant.

The TPS consists of 15 items for each indication. This number may be too high, as many referral letters may contain sufficient information with fewer items.³¹ However, eliminating the one-point items did not improve the correlation between the VAS and the TPS.

Also, scoring instruments with many interpretive questions could have a lower reliability.³² Since the TPS demonstrated excellent intra-rater and inter-rater reliability, this does not seem to be an issue with the TPS.

Some score items that may seem of limited relevance for a given indication (eg, information about Faecal Occult Blood Test (FOBT) in the dyspepsia indication) may have been selected by the gastroenterologists due to the variable's ability to discriminate between serious and less serious diseases.

In conclusion, the TPS measures the quality of the information in the referral letter objectively but is not a perfect tool for assessment of overall quality of the referral letter, as this also involves consideration of indication and structure of the referral.

Comparison with the existing literature

A score to assess referral letters in colorectal surgery has previously been suggested by Jiwa *et al.*³³ This score was developed by colorectal surgeons and GPs, and was validated against the likelihood of detecting colorectal cancer, not against any 'gold standard' for referral letter quality or on a wider range of gastrointestinal conditions. Consequently, it is not possible to determine

whether high scores actually reflect a high quality of referral letters. Thus, for assessing referral letters in general, the TPS is a more useful tool.

VAS is a recognised tool for assessing pain,³⁴ and has been shown to be well suited for assessing soft data in clinical trials.³⁵ It has also been used previously in quality score validation studies.²⁸ We therefore chose VAS as a method to assess the overall quality of the referral letters. Others have also used VAS as a tool to assess quality of referral letters,¹⁶ and showed scores similar to the findings in our study, with values between 1.1 and 6.9 for the various information items. However, VAS is a subjective measurement, and cannot replace objective scoring methods for evaluating quality.

Other more general scores for content of referral letters have also been created,^{7 20} but have given little insight into the specific symptom-related items mentioned in the referral letters, and cannot be used to assess the information quality of the referral letter. Also, referral letters in Norway are mainly generated electronically within the general national referral template,³⁶ and these general scores could consequently indicate a good referral letter, regardless of the description of the patient's symptoms and signs.

Strengths and limitations of the study

An important strength of the present study is the multi-centre design, giving the results a higher external validity. This aspect is also maintained by the wide variety of indications included, covering most of the reasons for referrals to Norwegian gastroenterology units.

The inclusion of a large number and wide variety of clinical gastroenterologists in the development of the TPS also ensures that the score reflects what the specialists actually need to effectively assess the referral letters. Further, the comparison of the TPS with the gastroenterologists' subjective assessment of the quality of the referral letter (VAS) enhances the emphasis on what the actual assessors value in the referral letters.

Our study has some potential limitations. First, we have not assessed referral appropriateness and cannot distinguish appropriateness as a deciding factor for lack of correlation between the VAS and the TPS. It is

possible that appropriateness of the referral has influenced the VAS, particularly when referrals with poor indication have been well written in terms of clinical information and structure.

We also observed a difference in the VAS between the centres, and it is possible that systematic differences in the use of the VAS may have influenced the results of the study.

In this study, gastroenterologists determined the optimal content of referral letters for easy assessment and prioritisation. A score based on risk factors for gastrointestinal cancer could have been an alternative approach. However, the aim was to develop a score that reflects what makes a referral letter easier to assess and prioritise, not to identify high-risk patients.

Another weakness of the study may be a selection bias for the gastroenterologists who participated in the score development. Thirty-nine gastroenterologists were invited to participate, and the 25 who completed the study may differ from the 14 who did not. However, willingness to participate probably does not influence the validity of their opinions regarding the content of referral letters, and could also be seen as a strength, as an interest in the topic may indicate a better understanding of what should be considered important clinical information in referral letters.

Some of the items selected for the score may be somewhat too unspecific, or may be considered inappropriate for the indication by other gastroenterologists. We have, for example, chosen the unspecific term 'previous radiology' in the jaundice/elevated liver enzymes indication, while 'previous ultrasound of the liver' may be a more appropriate and specific item. We have also chosen the term 'current medical treatment' as we believe this term accounts for any relevant information regarding the patient's medication, including ingestion of anticoagulants or antiplatelet agents.

Implications of the study

We have developed and validated an objective and reliable score for assessing the quality of referral letters in gastroenterology.

The moderate quality of referral letters observed in this study suggests that a tool to facilitate creation of high-quality referral letters would be beneficial. Information technology, like checklists or clinical decision-making systems, may be a part of the solution.

Unanswered questions/future research

This study is a presentation of the TPS that resulted from the survey among the 25 consultant gastroenterologists. However, some refinement of the TPS may be warranted and may increase the ability of the TPS to discriminate between high-quality and low-quality referral letters. Validation of the TPS in other healthcare systems is also necessary to reach a TPS of a high general validity.

In this study, we found a high prevalence of referral letters with scores below the middle value of the scales for both TPS and VAS. Future research should aim to implement and evaluate interventions to improve quality of referral letters in a way acceptable to referring GPs and specialists in the hospitals. A Cochrane review of interventions to improve referrals identified active involvement of secondary care specialists, and implementation of structured referral sheets as the only interventions with effect on referral quality.³⁷ Electronic referrals are the norm in Norway, and the implementation of structured referral sheets/checklists in electronic referrals may be an interesting intervention to explore.

CONCLUSION

The TPS is reliable, objective and shows good agreement with the subjective VAS. The score may be a useful tool for assessing referral quality in gastroenterology, particularly important when evaluating the effect of interventions to improve referral quality. The method used in the development of the score can serve as a model for other medical specialties.

Author affiliations

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

²Research Support Services, Oslo Centre for Biostatistics and Epidemiology, Oslo University Hospital, Oslo, Norway

³Department of Medicine, Vestfold Hospital Trust Tønsberg, Tønsberg, Norway

⁴Department of Medicine, Vestre Viken Hospital Trust Bærum, Drammen, Norway

⁵Department of Gastroenterology, Innlandet Hospital Trust Gjøvik, Gjøvik, Norway

⁶Institute of Clinical Medicine, University of Oslo, Oslo, Norway

⁷Department of Gastroenterology, Østfold Hospital Kalnes, Kalnes, Norway

⁸Department of Medicine, Innlandet Hospital Trust Lillehammer, Lillehammer, Norway

⁹Department of Gastroenterology, Telemark Hospital Trust Skien, Skien, Norway

¹⁰Department of Medicine, Vestre Viken Hospital Trust Drammen, Drammen, Norway

¹¹Department of Transplantation Medicine, Oslo University Hospital, Oslo, Norway

¹²Department of Health Economics and Health Management, University of Oslo, Oslo, Norway

¹³Department of Gastroenterology, Oslo University Hospital, Oslo, Norway

¹⁴Cancer Registry of Norway, Oslo University Hospital, Oslo, Norway

Acknowledgements The authors specially thank the administrative staff in the local gastroenterology departments for help with collecting paper copies of all included referrals.

Contributors SLE, TdL and LA designed the study. SLE, KW, ØH, BS, F-AH, TO, ES, KG and GH-H performed the data collection. SLE and CB performed the power and data analysis. SLE drafted the paper. All authors critically reviewed and improved it. SLE is the guarantor. All authors had access to all the data and take responsibility for the integrity of the data and the analysis.

Funding The PhD student salary was funded by the South-Eastern Norway Health Authority's research grant (grant agreement number 2008040) and the Norwegian Medical Association grant for quality improvement and patient safety (grant agreement number 14/1689). Researchers were independent of the funder.

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. The presented data are anonymised and risk of identification is low.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The full data set and statistical code can be made available from the corresponding author.

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/>

REFERENCES

- Barnett ML, Song Z, Landon BE. Trends in physician referrals in the United States, 1999–2009. *Arch Intern Med* 2012;172:163–70.
- Fleming DM. *The European study referrals from primary to secondary care*. University of Limburg, 1993.
- 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.
- Long A, Atkins JB. Communications between general practitioners and consultants. *BMJ* 1974;4:456–9.
- 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.
- François J. Tool to assess the quality of consultation and referral request letters in family medicine. *Can Fam Physician* 2011;57:574–5.
- 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.
- 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.
- 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.
- Ibiyemi O, Ibiyemi T. Quality and contents of referral letters from peripheral health centers to the dental centre of a teaching hospital, southwestern Nigeria. *Acta Odontol Scand* 2012;70:165–8.
- DeAngelis AF, Chambers IG, Hall GM. The accuracy of medical history information in referral letters. *Aust Dent J* 2010;55:188–92.
- Hedelin H, Johansson N, Strömberg P. Uneven quality of referrals for lower urinary tract symptoms in men. *Läkartidningen* 2003;100:1435–7.
- Scully ND, Chu L, Siriwardena D, *et al*. The quality of optometrists' referral letters for glaucoma. *Ophthalmic Physiol Opt* 2009;29:26–31.
- 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.
- 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.
- 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.
- Jenkins RM. Quality of general practitioner referrals to outpatient departments: assessment by specialists and a general practitioner. *Br J Gen Pract* 1993;43:111–13.
- Mead GE, Cunningham AL, Faulkner S, *et al*. Can general practitioner referral letters for acute medical admissions be improved? *Health Bull (Edinb)* 1999;57:257–61.
- Lønning KJ, Kongshavn T, Husebye E. Kvaliteten på henvisninger fra fastleger til medisinsk poliklinikk. *Tidsskr Nor Lægeforen* 2009;129(Arg. 129, nr. 18):1868–9.
- 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.
- Helsedirektoratet. *Nasjonal veileder for henvisninger til spesialisthelsetjenesten*. Oslo: Helsedirektoratet, 2015.
- Martinussen PE. Referral quality and the cooperation between hospital physicians and general practice: the role of physician and primary care factors. *Scand J Public Health* 2013;41:874–82.
- Health-Net N. Norwegian Health-Net. Secondary Norwegian Health-Net. <https://www.nhn.no/Sider/default.aspx>
- Nygaard E, Kårikstad V. National Prioritization Guideline for gastroenterology. In: Health TND, ed. Oslo: The Norwegian Directory of Health, 2009.
- UpToDate. Secondary UpToDate 2013. <http://www.uptodate.com/home>
- Norwegian Electronic Medical Handbook for Doctors. Secondary Norwegian Electronic Medical Handbook for Doctors 2013. <http://legehandboka.no/>
- Boyer L, Baumstarck-Barrau K, Belzeaux R, *et al*. Validation of a professionals' satisfaction questionnaire with electronic medical records (PSQ-EMR) in psychiatry. *Eur Psychiatry* 2011;26:78–84.
- Fleiss JL. *Design and analysis of clinical experiments*. John Wiley & Sons, 2011.
- Altman D. *Practical Statistics for Medical Research*. London: Chapman and Hall, 1991.
- 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.
- Oremus M, Oremus C, Hall GBC, *et al*. ECT, Team CSR. Inter-rater and test–retest reliability of quality assessments by novice student raters using the Jadad and Newcastle–Ottawa Scales. *BMJ Open* 2012;2:pil: e001368.
- Jiwa M, Mathers N, Walters S. Quality of information on referrals to colorectal surgeons: towards consensus. *Curr Med Res Opin* 2002;18:72–7.
- Huskisson EC. Measurement of pain. *Lancet* 1974;2:1127–31.
- Larsen S, Aabakken L, Lillevold PE, *et al*. Assessing soft data in clinical trials. *Pharm Med* 1991;5:29–36.
- Ree AO. *KITH Rapport 22/03 Medisinsk- faglig innhold i henvisninger- "Den gode henvisning"*. Trondheim: KITH: Kompetansesenter for IT i helsevesenet AS, 2003:35.
- 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.

Appendix 1: Final TPS (points) and frequencies of variables in the 327 collected referrals

Dyspepsia	Points	N(%)	Dysphagia	Points	N(%)	Long-standing abdominal pain	Points	N(%)
Symptoms			Symptoms			Symptoms		
Duration of symptoms	3	27(57.4)	Duration of symptoms	3	26(72.2)	Duration of symptoms	3	32(64.0)
Abdominal pain described in detail	2	13(27.7)	Reflux symptoms	3	18(50.0)	Localisation and characterisation of pain	3	24(48.0)
Reflux symptoms	2	19(40.4)	Hematemesis	2	5(13.9)	Pain in relation to meals	1	14(28.0)
Nausea/vomiting	1	21(44.7)	Texture provoking dysphagia	3	18(50.0)	Nocturnal pain	1	10(20.0)
Hematemesis	3	7(14.9)	Subjective localisation of obstruction	1	10(27.8)	Bowel changes	1	26(52.0)
Dysphagia	3	4(8.5)	Symptoms intermittent, stable or progressive	3	10(27.1)	Quantified weight loss	2	13(26.0)
Quantified weight loss	2	14(29.8)	Regurgitation of undigested foods	2	16(44.4)	Medical history		
Medical history			Quantified weight loss	2	17(47.2)	Previous medical history	1	37(74.0)
Previous medical history	1	37(78.7)	Medical history			Treatments/exposures		
Treatments/exposures			Treatments/exposures			Current medical treatment	2	37(74.0)
Current medical treatment	3	42(89.4)	Current medical treatment	1	28(77.8)	Clinical examination		
Effect of anti-acid/PPI- treatment	1	29(61.7)	Stimulantia(alcohol/tobacco)	2	8(22.2)	Abdominal palpation	3	29(58.0)
Clinical examination			Ulcerogenic medication	1	27(75.0)	General condition	1	8(16.0)
Abdominal palpation	1	24(51.1)	Clinical examination			Endoscopies/radiology		
General condition	1	9(19.1)	General condition	1	6(16.7)	Previous endoscopies/radiology	2	18(36.0)
Endoscopies/radiology			Endoscopies/radiology			Laboratory tests		
Previous endoscopies	2	22(46.8)	Previous endoscopies	3	8(22.2)	Faecal calprotectin	2	6(12.0)
Laboratory tests			Previous radiology	2	8(22.2)	FOBT**	3	12(24.0)
FOBT**	2	5(10.6)	Laboratory tests			Lab for diagnosing anaemia	2	24(48.0)
Lab for diagnosing anaemia	3	17(36.2)	Lab for diagnosing anaemia	1	8(22.2)	Lab for liver/pancreatic function	3	17(34.0)

Diarrhoea	Points	N(%)	Change of bowel habit	Points	N(%)	Constipation	Points	N(%)
Symptoms			Symptoms			Symptoms		
Duration of symptoms	2	37(97.4)	Duration of symptoms	2	37(77.1)	Duration of symptoms	3	20(74.1)
Nocturnal diarrhoea	2	9(23.7)	Abdominal pain described in detail	1	10(20.8)	Abdominal pain described in detail	1	7(25.9)
Macroscopic rectal bleeding with details	1	16(42.1)	Type of bowel change	3	44(91.7)	Frequency and consistency of bowel movements	2	15(55.6)
Quantified weight loss	2	14(36.8)	Macroscopic rectal bleeding with details	3	24(50.0)	Main complaint constipation (defecation hard, slow, rare etc)	1	16(59.3)
Medical history			B symptoms***	1	4(8.3)	Macroscopic rectal bleeding with details	2	8(29.6)
Treatments/exposures			Quantified weight loss	2	22(45.8)	Quantified weight loss	1	8(29.6)
Current medical treatment	1	30(78.9)	Medical history			Medical history		
Recent antibiotic treatment	3	8(21.1)	Treatments/exposures			Predisposing factors constipation	1	7(25.9)
Clinical examination			Current medical treatment	2	34(70.8)	Treatments/exposures		
Digital rectal examination(DRE)	1	3(7.9)	Clinical examination			Current medical treatment	3	25(92.6)
General condition	1	11(28.9)	Abdominal palpation	1	24(50.0)	Effect of current and previous treatment attempts for constipation	2	10(37.0)
Endoscopies/radiology			Digital rectal examination(DRE)	3	13(27.1)	Clinical examination		
Previous endoscopies	1	14(36.8)	General condition	1	11(22.9)	Abdominal palpation	2	13(48.1)
Laboratory tests			Endoscopies/radiology			Digital rectal examination(DRE)	3	10(37.0)
Faecal bacteria/parasites	3	7(18.4)	Previous endoscopies	2	11(22.9)	Endoscopies/radiology		
Faecal calprotectin	3	17(44.7)	Laboratory tests			Previous endoscopies	3	6(22.2)
FOBT**	3	16(42.1)	Faecal calprotectin	2	9(18.8)	Laboratory tests		
Lab for celiac disease	3	14(36.8)	FOBT**	3	30(62.5)	FOBT**	3	8(29.6)
Lab for diagnosing anaemia	2	14(36.8)	Lab for celiac disease	1	7(14.6)	Lab for diagnosing anaemia	2	11(40.7)
Lab for infection/inflammation(CRP/LPK)	2	11(28.9)	Lab for diagnosing anaemia	3	28(58.3)	TSH/FT4	1	8(29.6)

Gastrointestinal bleeding	Points	N(%)	Weight loss	Points	N(%)	Jaundice/elevated liver enzymes	Points	N(%)
Symptoms			Symptoms			Symptoms		
Duration of symptoms	2	25(73.5)	Duration of symptoms	3	16(76.2)	Duration of symptoms	3	14(53.8)
Symptoms from upper or lower GI-tract	1	23(67.6)	Presence of abdominal symptoms	2	13(61.9)	Abdominal pain described in detail	2	9(34.6)
Bowel changes	3	17(50.0)	Abdominal pain described in detail	1	6(28.6)	Colour changes urine/faeces	2	2(7.7)
Blood colour(bright red/dark)	3	25(73.5)	B symptoms***	2	4(19.0)	Quantified weight loss	1	4(15.4)
Blood observed on paper, on faeces or in faeces	2	22(64.7)	Quantified weight loss	3	16(76.2)	Medical history		
% defecations with observed blood	1	7(20.6)	Appetite	1	9(42.9)	Previous medical history	1	24(92.3)
Hematemesis	2	6(17.6)	Food intake	1	10(47.6)	Treatments/exposures		
Quantified weight loss	1	15(44.1)	Medical history			Current medical treatment	2	25(96.2)
Medical history			Previous medical history	2	15(71.4)	Stimulantia(alcohol/tobacco)	2	8(30.8)
Treatments/exposures			Treatments/exposures			Exposure liver-toxic substances	1	3(11.5)
Current medical treatment	1	30(88.2)	Clinical examination			Clinical examination		
Clinical examination			Abdominal palpation	2	12(57.1)	Abdominal palpation	3	7(26.9)
Abdominal palpation	1	11(32.4)	General condition	2	7(33.3)	Presence of liver stigmata	2	0
Digital rectal examination(DRE)	3	17(50.0)	Endoscopies/radiology			Endoscopies/radiology		
Endoscopies/radiology			Previous radiology	3	5(23.8)	Previous radiology	3	15(57.7)
Previous endoscopies	2	12(35.3)	Laboratory tests			Laboratory tests		
Laboratory tests			FOBT**	3	11(52.4)	Hepatitis serology	3	15(57.7)
Faecal calprotectin	2	8(23.5)	Lab for celiac disease	1	1(4.8)	Lab for infection/inflammation(CRP/LPK)	1	15(57.7)
FOBT**	3	19(55.9)	Lab for diagnosing anaemia	3	13(61.9)	Lab for liver/pancreatic function	3	25(96.2)
Lab for diagnosing anaemia	3	23(67.6)	Lab tests for mal-absorption	1	3(14.3)	Lab tests for specific liver diseases	1	14(53.8)