UNIVERSITA' DEGLI STUDI
DI MODENA E REGGIO EMILIA

DOTTORATO DI RICERCA IN CLINICAL AND EXPERIMENTAL MEDICINE

CICLO XXVIII

TELEPHONE MEDICAL CONSULTATION QUALITY AND IMPROVEMENT

Candidato: Alberto Vaona

Relatore: Prof. Roberto D'Amico

Direttore: Prof.Giuseppe Biagini

Telephone medical consultation quality and improvement

Cover design and thesis layout: King's College London

http://www.kcl.ac.uk/library/researchsupport/disseminate/theses/index.aspx

Abstract

Background

The ability to consult by telephone has become an integral part of modern healthcare. Safety and legal concerns come from studies showing the clinical quality of medical telephone consultation (MTC) is low. Little is know about the quality of MTC in Italy.

No systematic review is currently available to address specific training of health professionals on MTC. Computerised decision support systems (CDSS) could play a role in improving MTC quality but their effectiveness is not clear, especially when CDSS are applied to telephone triage (clinical cases urgency estimate).

E-learning, the integration between educational programmes and interactive electronic systems, could be useful to deliver these skills but little is known about its effectiveness in comparison with traditional training.

Objectives

1) To assess the clinical quality of MTC in a high MTC rate service.

2) To systematically review the literature about the efficacy of CDSS on mortality and morbidity prevention, of e-learning in comparison with traditional learning, of training intervention to improve clinicians telephone skills

3) To test the efficacy and safety of a CDSS for telephone triage in a high MTC rate service

4) To study barriers and facilitators toward the uptake of CDSS

Methods

A cross sectional study detected the MTC quality in a high MTC rate service; an off-topic Cochrane review was performed to learn Cochrane methodology; three systematic reviews (two Cochrane and one non-Cochrane) found and pooled the results of existing studies (estimating their risk of bias) about CDSS efficacy on mortality and morbidity prevention, about e-learning efficacy in comparison with traditional learning and about training intervention for telephone skills improvement; a randomised control trial assessed the efficacy and safety of a CDSS for telephone triage and one more cross sectional study addressed barriers and facilitators towards CDSS uptake.

Main results

The clinical quality of MTC in the study setting was low: doctors asked only about 30% of the obligatory questions.

CDSS use did not affect mortality but a significant effect was detected in the prevention of morbidity, even if selective outcome reporting and publication bias are not excluded.

The off-topic Cochrane systematic review showed the main two drugs for multiple sclerosis are similar in terms of relapse rate, progression and imaging outcomes.

E-learning when compared to traditional learning does not improve patient outcomes and health professionals behaviours; it possibly leads to slightly less improvement in health professionals skills but no difference on health professionals knowledge.

We found no evidence about training interventions for improving clinicians telephone communication skills on patient or clinicians outcomes.

Due to technical problems the RCT assessing the efficacy of a CDSS for telephone triage failed to collect reliable results: the CDSS could not be used to a sufficient extent and this resulted in a loss of power and a irreversible imbalance between the study arms.

The protocol about barriers and facilitators of CDSS uptake was performed in a study by the other researchers and they could design a model to guide the adoption of CDSS.

Authors' conclusions

Despite the phone is a very frequently used tool, the clinical quality of the MTC in Italy seems low. The currently available evidence on how to address the improvement of clinicians telephone skills is poor but CDSS proved to be effective even if specific CDSS for the MTC remain to be tested. Once evidence is available on how to improve, e-learning will be a potential cost-effective training way: it seems as effective as traditional training in terms of transmission of knowledge even if potentially little less effective on skills transmission.

Introduzione

La consultazione telefonica medica (MTC) è diventata una parte integrante della moderna assistenza sanitaria. Alcuni studi mostrano che la qualità clinica della consultazione telefonica medica (MTC) è bassa e è dato sapere circa la qualità della MTC in Italia. Nessuna revisione sistematica è attualmente disponibile per indirizzare la formazione dei sanitari sulla MTC. I moderni sistemi informatici di supporto decisionale (CDSS) potrebbero svolgere un ruolo importante nel migliorare la qualità della MTC ma la loro efficacia non è chiara. L'e-learning, l'integrazione tra programmi educativi e sistemi elettronici interattivi, potrebbe essere utile per formare i sanitari ma poco si sa circa la sua efficacia a confronto con la formazione tradizionale.

Obiettivi

1) Valutare la qualità clinica della MTC.
2) Revisionare la letteratura sull'efficacia di
- CDSS su mortalità e prevenzione della morbilità,
- e-learning in confronto con la formazione tradizionale,
- interventi formativi utili a migliorare le competenze telefoniche dei sanitari.
3) Testare l'efficacia e la sicurezza di un CDSS per il triage telefonico
4) Studiare le barriere e i facilitatori all'utilizzo dei CDSS.

Metodi

E' stato condotto uno studio trasversale per rilevare la qualità della MTC; una revisione sistematica Cochrane off-topic è stato realizzata al fine di apprendere la metodologia Cochrane; tre revisioni sistematiche (due Cochrane) hanno reperito e combinato i risultati degli studi esistenti sull'efficacia dei CDSS in termini di mortalità e prevenzione della morbilità, di efficacia dell'e-learning a confronto con la formazione tradizionale e circa gli interventi formativi per il miglioramento delle abilità telefoniche dei sanitari; un trial controllato e randomizzato ha valutato l'efficacia e la sicurezza di un CDSS per triage telefonico e un altro studio trasversale ha valutato le barriere e i facilitatori verso l'utilizzo dei CDSS.

Risultati

La qualità clinica della MTC che è bassa perché i medici pongono solo circa il 30% delle domande obbligatorie. L'uso del CDSS non influenza la mortalità ma presenta un effetto significativo sulla prevenzione della morbilità. La revisione sistematica Cochrane off-topic ha mostrato che i due principali farmaci per la sclerosi multipla sono simili in termini di risultati tasso di recidiva, progressione e di risultati radiologici. L'e-learning se confrontato con l'apprendimento tradizionale non migliora i risultati che il paziente trae dall'assistenza né i comportamenti dei sanitari; è possibile che porti ad un miglioramento leggermente minore delle competenze dei sanitari e non aumenti la loro conoscenza. Non abbiamo trovato alcuno studio sugli interventi formativi per migliorare le abilità telefoniche dei sanitari.
A causa di problemi tecnici l'RCT che si proponeva di valutare l'efficacia di un CDSS per il triage telefonico non è riuscito a raccogliere risultati affidabili. Sulla base del protocollo di studio su

barriere e facilitatori all'uso dei CDSS, è stato condotto da altri ricercatori uno studio utile a progettare un modello per guidare l'adozione dei CDSS.

Conclusioni

Nonostante il telefono sia uno strumento molto utilizzato, la qualità clinica della MTC in Italia sembra essere bassa. Le prove attualmente disponibili sul modo di affrontare il miglioramento delle abilità telefoniche dei sanitari sono scarse ma l'utilizzo dei CDSS ha dimostrato di essere efficace anche se l'efficacia e la sicurezza di CDSS specifici per la MTC rimangono da testare. Una volta disponibili prove di efficacia su come migliorare, l'e-learning sarà un canale formativo potenzialmente efficace come la formazione tradizionale, anche se può essere leggermente meno efficace nella trasmissione delle competenze.

Table of Contents

INTRODUCTION 16

1.1 THE MEDICAL TELEPHONE CONSULTATION (MTC) 16

The role of medical telephone consultations 16

The MTC dimensions and quality 17

The computerized decision support systems (CDSS) 19

OOH service in Italy and telephone triage 20

e-Learning for health professionals 20

1.2 COCHRANE AND NON-COCHRANE SYSTEMATIC REVIEWS 22

1.3 THE PURPOSES OF THIS RESEARCH 22

The “process” research purpose; getting the method 22

Why it is important to do this research 23

The “outcome” research purposes 23

Why it is important to do this research 24

METHODS 26

2.1 EFFICACY AND SAFETY OF INFS-BETA VS GS IN RELAPSING REMITTING MULTIPLE SCLEROSIS 26

Criteria for considering studies for this review 26

Search methods for identification of studies 28

Searching other resources 28

Selection of studies 29

Data extraction and management 29

Assessment of risk of bias in included studies 29

Measures of treatment effect 29

Unit of analysis issues 30

Dealing with missing data 30

Assessment of heterogeneity 30

Assessment of reporting biases 30

Data synthesis 30

Subgroup analysis and investigation of heterogeneity 30

Sensitivity analysis 31

Summary of findings table 31

2.2 THE QUALITY OF TELEPHONE TRIAGE 31

2.3 THE EFFICACY OF CDSS 34

Eligibility Criteria 34

Data Sources 35

Study Selection and Data Extraction 35

Risk of Bias Assessment 36

Data Synthesis 36

2.4 THE EFFICACY OF E-LEARNING 37

Criteria for considering studies for this review 37

Search methods for identification of studies 39
Data collection and analysis 40
Selection of studies 40
Data extraction and management 40
Assessment of risk of bias in included studies 41
Measures of treatment effect 41
Unit of analysis issues 42
Dealing with missing data 43
Assessment of heterogeneity 43
Assessment of reporting biases 44
Data synthesis 44
Subgroup analysis and investigation of heterogeneity 45
Sensitivity analysis 45
2.5 TRAINING INTERVENTION FOR IMPROVING TELEPHONE SKILLS 45
Criteria for considering studies for this review 45
Search methods for identification of studies 47
Data collection and analysis 48
Data extraction and management 48
Assessment of risk of bias in included studies 49
Measures of treatment effect 49
Unit of analysis issues 50
Dealing with missing data 50
Assessment of heterogeneity 50
Assessment of reporting biases 50
Data synthesis 50
Subgroup analysis and investigation of heterogeneity 51
Sensitivity analysis 51
2.6 EFFICACY AND SAFETY OF A CDSS TO SUPPORT TELEPHONE TRIAGE 52
Study Design and research hypotheses 52
The setting 52
The intervention, Odyssey Teleassess 53
The outcomes 53
The study period 54
Sample size, randomization, allocation and blindness 54
Data analysis 55
Duality of interest 55
2.7 BARRIERS AND FACILITATORS TO UPTAKE CDSS 55
Setting and participants 56
Data collection 56
Data analysis 57
Ethics and funding 57

RESULTS 58

3.1 EFFICACY AND SAFETY OF INFS-BETA VS GA IN RRMS 58
3.2 THE QUALITY OF TELEPHONE TRIAGE 73
3.3 THE EFFICACY OF CDSS 76
3.4 THE EFFICACY OF E-LEARNING 79
3.5 TRAINING INTERVENTION FOR IMPROVING TELEPHONE SKILLS 89
3.6 EFFICACY AND SAFETY OF A CDSS TO SUPPORT TELEPHONE TRIAGE 94
3.7 BARRIERS AND FACILITATORS TO UPTAKE CDSS 95

DISCUSSION 96

4.1 EFFICACY AND SAFETY OF INFS-BETA VS GA IN RRMS 96
4.2 THE QUALITY OF TELEPHONE TRIAGE 99
4.3 THE EFFICACY OF CDSS 100
4.4 THE EFFICACY OF E-LEARNING 102
4.5 TRAINING INTERVENTION FOR IMPROVING TELEPHONE SKILLS 105

CONCLUSIONS 106

5.1 EFFICACY AND SAFETY OF INFS-BETA VS GA IN RRMS 106
5.2 THE QUALITY OF TELEPHONE TRIAGE 106
5.3 THE EFFICACY OF CDSS 107
5.4 THE EFFICACY OF E-LEARNING 107
5.5 TRAINING INTERVENTION FOR IMPROVING TELEPHONE SKILLS 109

FINAL STATEMENT 110

REFERENCES 111

APPENDICES 140

Table of Figures

Figure 1. Study flow diagram INFs vs GA in RRMS 58

Figure 2. Judgements about each risk of bias item for each included study 63

Figure 3. Analysis 1.1. Number of participants who experienced at least one relapse 66

Figure 4. Analysis 1.2. Number of participants who experienced at least one relapse (likely scenario)..... 66

Figure 5. Analysis 2.1. Number of participants with confirmed worsening 67

Figure 6. Analysis 2.2. Number of participants with confirmed worsening (likely scenario) 67

Figure 7. Analysis 3.1. Number of participants who withdrew from or dropped out 68

Figure 8. Analysis 3.2. Number of participants who withdrew from or dropped out 68

Figure 9. Analysis 4.1-2. Frequency of relapse 69

Figure 10. Analysis 5.1. Time to first relapse 69

Figure 11. Analysis 7.1. Participants treated with steroids for relapse of MS 70

Figure 12. Analysis 9.1-3. Mean number of active T2-hyperintense lesions per participant 70

Figure 13. Analysis 10.1-3. Mean number of new T1-hypointense lesions per participant 71

Figure 14. Analysis 11.1-3. Mean change in total T2-hyperintense lesion volume 72

Figure 15. Analysis 12.1, 12.2. Mean change in total T1-hypointense lesion volume 72

Figure 16. Analysis 13.1-2. Mean change in total brain volume 73

Figure 17. Study flow diagram CDSS Efficacy 76

Figure 18. Judgements about each risk of bias item for each included study 77

Figure 19. Forest plots CDSS effictiveness on mortality and morbidity prevention. 78

Figure 20. Study Flow Diagram e-Learning Efficacy 79

Figure 21. Judgements about each risk of bias item for each included study 83

Figure 22. Analysis 1.1. Patients appropriately screened (dyslipidaemia) 86

Figure 23. Analysis 1.2. Patients appropriately treated (dyslipidaemia) 86

Figure 24. Analysis 1.3. Patients appropriately screened (dyslipidaemia and diabetes) 86

Figure 25. Analysis 1.4. Patients appropriately treated (dyslipidaemia and hypertension) 86

Figure 26. Analysis 1.5. Patients appropriately screened (dyslipidaemia) and treated (hypertension) 87

Figure 27. Analysis 3.1 Knowledge at any time (fixed effect) 88

Figure 28. Analysis 3.2. Knowledge at any time (random effects) 88

Figure 29. Analysis 3.3 Knowledge at immediately after the training 88

Figure 30. Analysis 3.4 Knowledge at after 3 or more months 88
Figure 31. Study Flow Diagram Training Intervention for Telephone Skills Improvement 91
Figure 32. Judgements about each risk of bias item for each included study 91
Figure 33. RCT Study results (non frail population) 94
Figure 34. RCT Study results (frail population) 95

Table of Tables

Table 1. Obligatory questions to be asked in each case 33
Table 2. INFs vs GA for RRMS Summary of Findings Table 64
Table 3. Adult with nosebleed. Number of doctors who asked the obligatory questions 74
Table 4. Adult with fever. Number of doctors who asked the obligatory questions 74
Table 5. Child with vomiting. Number of doctors who asked the obligatory questions 74
Table 6. Child with fever. Number of doctors who asked the obligatory questions 75
Table 7. Management decisions concerning the 40 simulated cases 75
Table 8. e-Learning Efficacy Summary of Finding Table 84
Table 9. Telephone Skills Summary of Findings Table 93

Acknowledgements

I would like to dedicate this PhD thesis to Hay Derkx, a good friend and the best medical telephone consultation master.

I would like to thank Giulio Rigon and Anna Longobardi for their support: without them I would have never begun this PhD.

Thanks to Prof. Giuseppe Biagini and Prof Roberto D'Amico for the time and the encouragement they provided me.

Thanks to all the researchers who worked with me during these years: I appreciated them very much; they let an Italian GP to taste the world of research and the pleasure to give his own very little contribution to what is known.

Abbreviations

Abbreviation	Meaning
CDSS	Computerized Decision Support Systems
CME	Continuing Medical Education
DMT	Disease Modifing Therapies
EBM	Evidence Based Medicine
EHR	Electronic Health Records
GA	Glatiramer Acetate
GP	General Practitioner
GT	Grand Therapy
INFb	Interferon Beta
ISP	Incognito Standardised Patient
LHT	Local Health Trust
MS	Multiple Sclerosis
MTC	Medical Telephone Consultation
OOH	Out Of Hours
RRMS	Relapsing Remitting Multiple Sclerosis
SPMS	Secondary Progressive Multiple Sclerosis

INTRODUCTION

1.1 The medical telephone consultation (MTC)

Since 1879, the year of the first documented medical telephone consultation, the ability to consult by telephone has become an integral part of any modern patient-centred healthcare system (Anonymous 1879, Evans 2003); for many years it has been reported that in the United States one eighth of professional time is spent by doctors assessing clinical cases on the telephone (Bergman 1966) and up to more than a quarter of all care consultations are conducted by this method (Mendenhall 1981; Patel 1997). More recently, the British Medical Association (BMA) has provided guidance for general practitioners (GPs) entitled *Consulting in the modern world: guidance for GPs* (BMA 2001); this guidance advises "telephone consultations when correctly conducted can be considered to be a safe and acceptable practice". Reisman 2005 described how telephone communication is the primary mode of communication between physicians and patients outside of an office visit. Car 2004 and Patel 2005 argued that telephone consulting is both a feasible and effective form of clinical intervention.

Bunn 2004 described telephone consulting as a process whereby patients receive medical advice by one or more qualified healthcare professionals via the telephone. The authors concluded that telephone consultations appear to be safe and that healthcare users were just as satisfied with them as with face-to-face consultations. They also suggested that telephone consultations appear to decrease the number of immediate visits to doctors without increasing attendance to emergency departments. This conclusion seems to be denied by a recent large study (Campbell 2014) which found that introducing telephone triage results merely in a redistribution of GP workload away from face-to-face consultations and towards more telephone consultations or nurse-led care.

Italian National Health Service 2013 Data Report[1] say more than 11.500 general practitioners work in the Out of Hours (OOH) Services and many studies show they deal by phone up tp 60% of their clinical cases (Colombo 2016, Buja 2015) as it happens in other countries (UK National Audit Office 2014).

The role of medical telephone consultations

Telephones are used to provide a range of healthcare services including delivery of routine and emergency care, obtaining repeat prescriptions, gathering results of laboratory investigations and facilitating health promotion (Car 2003, Giesen 2011). Examples of telephone consultations include the management of conditions such as heart failure (Clark 2007; Riegel 2002), asthma (Gruffydd-Jones 2005; Patel 2009; Pinnock 2003) and palliative care (Pimentel 2015; Zhou 2012).

[1] http://www.salute.gov.it/imgs/C_17_pubblicazioni_2536_allegato.pdf

Telephone consultations may reduce doctors' face-to face workloads and enhance access to care without the inconvenience and cost associated with physically attending a consultation, thus increasing the flexibility and availability of service (Hallam 1992; Patel 2005, Campbell 2014).

Katz 2008, Car 2008, Huibers 2011 highlighted some of the safety concerns that exist in relation to telephone consultations as the vulnerability of patients to errors in management and of clinicians to malpractice claims. The authors suggested that the most effective risk-management strategy is to improve the quality of telephone care and service to patients. They also suggested that prevention should include a more disciplined approach to documentation, improved workload systems, increased skills training and an upfront commitment to evaluation.

Bunn 2004 claimed that there are still questions about the effect of telephone consultations upon service use. Since telephone consultations play a role in patient management, it is essential that when consulting over the telephone, healthcare professionals feel confident with their skills to conduct and document the interview with accuracy and clinical competency. It is therefore important that they receive adequate training to enable them to carry out their clinical roles effectively.

Purc-Stephenson 2012 found that overall patient compliance with triage advice provided by tele-nurses was 62% and it was influenced by the interactive role of patient perceptions and the quality of provider communication. The authors highlight the need for communication-skills training in a telephone-consultation context that is patient centred and specifically addressed, building active listening and advising skills and providing guidance on how to structure the call.

The MTC dimensions and quality

Telephone triage is a specific role of MTC and it can be described as the care process by which the degree of urgency of a patient clinical problem presented by telephone and the care needed are determined.

This care process can be divided into the first phase of information gathering followed by the second phase of determining the degree of urgency and the care needed. The call handler, also called "triagist", actually handles the request for medical advice. In most cases this person is a specially trained nurse or a physician. The quality of telephone triage depends on the clinical knowledge and the communication skills of the triagist and his or her expertise in evaluating the information gathered (Derkx 2008, Derkx 2009).

Most of the studies about the MTC dimensions and quality were performed in the primary care out of hours service setting and total quality management of MTC should consider three dimensions: the clinical content, the communication and the reporting of the consultation.

1.1.3.1 MTC clinical content

Derkx 2009 used incognito standardised patients to present to 17 out of hours Dutch centres seven clinical cases three times each over a period of 12 months, making a total of 357 calls. The mean percentage of obligatory questions asked compared with the standard was 54% and only 21% of the questions asked were obligatory. Answers to questions about the clinical condition were not always correctly evaluated from a clinical viewpoint. The quality of

information on home management and safety net advice varied, but it was consistently poor for all cases and for all out of hours centres. Triagists achieved the appropriate triage outcome in 58% of calls.

Similar findings were reported many years before by Brown 1974 and Yanovski 1992; these two authors noted that nether the quantity nor the quality of questions asked was associated with any of the measures of experience and many physicians in private practice did not demonstrate the skills needed to perform adequate telephone triage despite their great number of patient contacts.

All these researches conclude physicians must not only ask good questions, but must carefully listen to and evaluate the answers if they are to manage patients effectively on the telephone.

1.1.3.2 MTC relational content

Assessment of the medical problem and the quality of 'care-by-phone' depend on the medical and communication skills of the call handlers.

The scale, known as the RICE rating scale, has 17 items divided over four different phases of the telephone consultation: Reason for calling, Information gathering, Conclusion and Evaluation (RICE). This instrument can be used to give feedback to call handlers (Derkx 2007).

In Derkx 2008 telephone incognito standardised patients (TISPs) called 17 OOH centres presenting different clinical cases and the assessment of communication skills was carried out using the RICE-communication rating list.

The mean overall score for communication skills was 35% of the maximum feasible. Triagists usually asked questions about the clinical situation correctly and little about the patient personal situation, perception of the problem or expectation. Advice about the outcome of triage and self-care advice was usually given without checking for patient understanding and acceptance of the advice. Calls were often handled in an unstructured way, without summarizing or clarifying the different steps within the consultation.

Authors conclude training in telephone consultation should focus more on patient-centred communication with active listening, active advising and structuring the call.

1.1.3.3 MTC reporting

After a telephone consultation at an out of hours centre, the call handler writes a medical report to record what has been discussed with the caller: the report of a telephone consultation serves different purposes. First, it is important to secure continuity of care. The report is the only documented and therefore rapidly accessible source of information about the content of a call (Derkx 2010).

Reports of telephone consultations of out of hours centres contained little information on patients' clinical and personal condition. This could potentially endanger patient continuity of care and might pose legal consequences for the triagist.

In Derkx 2010 the out of hours centres returned a report for 78% of the 357 calls. For the remaining 22% of the calls, no report was written. Reports contained almost always information about the medical reason for calling but little information about details of the clinical history. Patient expectation, personal situation or perception of the care advice was seldom

documented. In all but one out of hours centre, answers to obligatory questions were reported by triagists, although they had not been asked, varying between 1% and 54% of all questions entered. Triagists entered a subjective evaluation of a patients' condition in 12% of the reports.

The computerized decision support systems (CDSS)

A CDSS can be defined as an information system aimed to support clinical decision-making, which links patient specific information in electronic health records (EHRs) with evidence-based knowledge to generate case-specific guidance messages through a rule- or algorithm-based software (Chaudhry 2006).

Current research demonstrates the potential of computerized decision support systems (CDSS) to assist with problems raised in clinical practice, increase clinician adherence to guideline- or protocol-based care, and, ultimately, improve the overall efficiency and quality of health care delivery systems (Chaudhry 2006, Ash 2012, Roshanov 2013). CDSS have been additionally shown to increase the use of preventive care in hospitalized patients, facilitate communication between providers and patients, enable faster and more accurate access to medical record data, improve the quality and safety of medication prescribing, and decrease the rate of prescription errors (Kaushal 2003, Bonnabry 2008, de Lusignan 2008, Romano 2011 Bright 2012). A recent study estimated that the adoption of Computerized Physician Order Entry and Clinical Decision Support could prevent 100 000 inpatient adverse drug events (ADEs) per year, resulting in increased inpatient bed availability by more than 700 000 bed-days and opportunity savings approaching €300 million in the studied European Union member states (i.e., theCzech Republic, France, the Netherlands, Sweden, Spain, and the United Kingdom) (Swedish Presidency of the EU 2014). Electronic Health Records (EHRs) represent another innovation that is gaining momentum in health care systems. In the United States, the use of EHRs is encouraged by the $27 billion allocated in reimbursement incentives by the 2009 Health Information Technology for Economic and Clinical Health (HITECH) Act. Under the Act, clinicians and hospitals must demonstrate "meaningful use" of EHRs by adhering to a set of criteria, which includes the implementation of clinical decision support rules relevant to a specialty or high priority hospital condition such as diagnostic test ordering (Miriovsky 2012). The integration of CDSSs with EHRs through the delivery of guidance messages to health care professionals at the point of care may maximize the impact of both innovations.

Information that is accessible at the point-of-care is seen by clinicians as a key factor in the successful incorporation of evidence in routine care (Bright 2012). Despite consistent findings demonstrating the potential of CDSSs to improve patient outcomes and health professional behavior, the mere provision of the technology does not guarantee its uptake. Even if a CDSS is readily available within a hospital, clinicians often fail to adopt its recommendations, ignoring up to 96% of its alerts (Bright 2012).

So far no CDSS has been studied in terms of efficiency and safety in order to support telephone triage even though several are available.

OOH service in Italy and telephone triage

Italian Out Of Hours (OOH) Service is provided by doctors only all over the country. Telephone consultations account for 50% of the whole amount of consultations. Telephone triage is currently performed without computerised support. Several studies show that not supported telephone triage is partially unsafe (Derkx 2008, Huibers 2011) while evidence from not randomized controlled trial highlights that supported telephone triage could be safer (Meer 2012). Results from observational studies suggest telephone triage could reduce costs by reducing hospital referrals (Fry 2009, Dunt 2001, Dale 2001, Raftery 2000). In the Azienda ULSS 20 Verona setting, the local health trust (LHT) of Verona, MTC referral rate has been estimated as 13% of all the calls while in other countries it is lower: 4% in the Netherlands, 4,7% in the UK and 5,1% in the USA (Giesen 2011, Forrest 2003).

No study has evaluated so far the impact of a CDSS on doctors telephone triage outcomes.

e-Learning for health professionals

E-learning is a broad concept that involves the transfer and usage of knowledge through educational programmes integrated with interactive electronic systems. Currently, there is no standardised or recognised definition of e-learning for research purposes. The Medical Subjects Headings Vocabulary, for example, does not provide a specific item different from 'distance education' that includes, in addition to computer networks, correspondence, radio and television as media tools.

Many different terms refer to e-learning in biomedical literature: web-based learning or training, online learning or education, computer-assisted or -aided instructions (CAI) or computer-based instructions (CBI), and Internet-based learning (Ruiz 2006, Cook 2008a), multimedia learning, technology-enhanced learning or virtual learning. These terms however differ from e-learning because they address only a specific part of the concept such as the medium (e.g. computer-assisted instruction) or the delivery system (e.g. online learning). Although the term e-learning has sometimes been used to define conjunctions of electronic systems with face-to-face teaching (blended interventions), it is generally seen as a particular evolution of distance education: it could be defined as the use of information technologies in order to deliver education to remote learners. When these learners are computer-assisted, interconnected through computer networks and access online packages for learning, distance education can unequivocally be referred to as e-learning (Ward 2001, Ruiz 2006).

Although e-learning shares many features with traditional learning systems (Zimitat 2001), several aspects are distinctive. Besides adequate quality assurance and careful educational design, the quality assessment of an e-learning programme should also involve an analysis of navigability, multimedia approach, degree of interactivity, and other key factors in the development of an optimal e-learning framework (Straus 2004). The traditional role of trainers is evolving from a 'distributor of content' to a 'facilitator', enhancing the learner-centred characteristics of the educational programme (Wentling 2000).

Applying the latest information technologies to education takes advantage of the increasing availability of Internet access (using optical fibres, Wi-Fi and 3G/4G mobile phone technology),

allowing a broad use of contents across diverse settings (home, workplaces, and public places such as libraries, parks, and Internet points).

The delivery advantages of an e-learning programme are easily recognised: lower costs, widespread distribution, increased accessibility to information, frequent content updates, and personalised instruction in terms of content and pace of learning are some of the most cited benefits of e-learning (Wentling 2000). Moreover, its interactivity and ability to link educational programmes with past experiences and specific needs fit the adult learning paradigm (Gibbons 2000).

As a result of these advantages, online learning is growing in popularity. E-learning systems are rapidly increasing in number in several countries and offer many speciality modules in their portfolios (Coppus 2007, Moja 2007, Ruiz 2007), even though there may be potential disadvantages such as technology-related costs, cost involved in developing programmes, possible technical problems, limited direct interaction, lack of exchanges and relations with other learners, absence of the physical presence of the teacher, decrease in motivation to learn, need for greater self-discipline and attenuation of the desire to compete with other learners (Poon 2015, Cook 2007, Welsh 2003). Moreover, equity should be considered carefully: poor access and lack of computer and Internet literacy could exclude some health professionals, especially in low- or middle-income countries, from participation.

Previous systematic reviews (Cook 2008a, Lam-Antoniades 2009, Cook 2010a) on the efficacy and efficiency of e-learning focused on Kirkpatrick's outcomes set (Kirkpatrick 1996): satisfaction, knowledge/attitudes, skills (in a test setting), behaviours (in a practice setting), and effects on patients. Knowledge measurement by standardised tests is the most direct outcome for both traditional and e-learning systems. The progression from cognitive to behavioural steps, from acquiring knowledge to performing a task in practice, is neither linear nor simple: many other factors influence health professionals' behaviours, including system-related factors, (e.g. government incentives, guidelines, laws) and individual-related factors (e.g. patients' expectations, relationship with peers) (Rethans 2002).

These reviews found:

- e-learning is associated with large positive effects when compared with no intervention (Cook 2008);
- e-learning is associated with small positive effects when compared with non e-learning educational interventions, suggesting effectiveness similar to traditional methods (Cook 2008;
- e-learning and non-e-learning educational interventions take similar time to participate in or complete (Cook 2010);
- interactivity, practice exercises, repetition, and feedback play pivotal roles in e-learning and seem to be associated with improved learning outcomes (Cook 2010);

A further relevant finding was the large heterogeneity among study designs, participants, instructional designs and outcomes. The authors conclude that e-learning is not a single entity, although educators and researchers frequently view it as a single activity, or a cluster of single activities, with relatively homogeneous effects (Cook 2010).

1.2 Cochrane and non-Cochrane Systematic Reviews

In 1972, Archie Cochrane expressed the need for higher quality empirical evidence around the development of health services (McKenzie 2013). Cochrane believed that randomised controlled trials play a major role in the development of this evidence, but realized that there was no systematic way to disseminate results from randomised trials to the professional medical field (Starr 2009). As a result, in 1993, The Cochrane Collaboration was established to conduct meta-analytical reviews on health care related topics, specifically randomised trials, enabling physicians and other key decision-makers to access high-quality information on evidence-based results. Because of its rigorous and analytic methodology, standardization of approaches, and transparency, the Cochrane Collaboration is often considered to be the gold standard for meta-analytic reviews, is deemed robust against bias (Jorgensen 2006), and is highly trusted by clinicians (Rosenbaum 2008).

Not infrequently, two or more meta-analyses are independently published on the same topic, though such studies often fail to reference each other's findings and may yield conflicting results (Siontis 2013, Helfer 2015). Reviews conducted within the Cochrane Collaboration follow a standardized set of methods that non-Cochrane reviews are not bound to. In theory, this might introduce systematic differences between the two. Several studies provide empirical evidence that Cochrane reviews tend to be of higher quality, were less vulnerable to bias, acknowledged more limitations, and were generally more conservative in how the results were endorsed than non-Cochrane reviews (Moseley 2009, Tricco 2009).

Cochrane Collaboration's methodology has many advantages: standardization of methodology, transparency, and the breadth of analyses assessed in one report. Nonetheless, this approach limits the numbers of individuals or organizations that can commit the time and labor to adhering to the Cochrane Collaboration's standards. One consequence of that is to limit the overall number of analyses that are conducted by the Cochrane Collaboration. Given that meta-analyses are essential tools in clinical research, the need for meta-analyses conducted outside of the Cochrane Collaboration is not in dispute (Useem 2015).

1.3 The purposes of this research

The "process" research purpose; getting the method

This PhD Thesis has had a "process" purpose concerning the learning of Cochrane methodology for systematic reviews. In order to get this methodology an "off-topic" systematic review was performed with Multiple Sclerosis (MS) Cochrane Review Group about a comparison between the main two drugs to treat this disease: interferon beta (INFs-beta) and glatiramer acetate (GA).

The main objective of the review was to assess whether IFNs-beta and GA differ in terms of safety and efficacy in the treatment of patients with relapsing-remitting MS (RRMS).

Why it is important to do this research

The choice of specific drug for MS remains a relevant issue and the final decision should be based on a thorough evaluation of the risk/benefit profile, impact on quality of life and potential neuroprotective and long-term effects of a given drug (Compston 2008). At present, no clear evidence can be found on the relative efficacy of different disease modifying therapies (DMT) in the treatment of patients with relapsing-remitting MS (RRMS). Comparative trials are difficult to run for several reasons, including the difficulty of achieving proper blinding and the need for large sample sizes; however, they remain the best tool for acquiring objective information (Goodin 2008).

Direct comparative data evaluating IFNs-beta and GA in RRMS are now available (Goodin 2008a), but no systematic reviews of head-to-head trials of IFNs versus GA have been performed. Specific details on their relative efficacy and tolerability may help physicians make a more precise and unbiased therapeutic choice for their patients.

The "outcome" research purposes

1.3.3.1 The quality of telephone triage

The aim of this part of the research was to evaluate the validity of telephone triage by comparing the questions asked by call-handlers working at OOH centres during telephone triage with those that should have been asked and examining their decisions about therapy. In order to address this issue we conducted a cross sectional study.

1.3.3.2 The efficacy of CDSS

The aim of this part of the research was to rigorously evaluate the impact of CDSSs linked to EHRs on critical outcomes—mortality morbidity, and costs—and adopted a narrow definition of the intervention to facilitate its coherent and accurate evaluation. In order to address this issue we conducted a non-Cochrane systematic review.

1.3.3.3 The efficacy of e-learning

The aim of this part of the research was to assess the effectiveness of e-learning programmes in comparison with traditional learning (educational interventions without access to e-learning; e.g. paper-based books, residential courses) in improving patient outcomes and behaviours, skills or knowledge of licensed health professionals. In order to address this issue we conducted a Cochrane systematic review.

1.3.3.4 The efficacy of training intervention for improving telephone skills

The aim of this part of the research was to assess the effectiveness of training interventions on clinicians' telephone skills and their effects on patient outcomes. In order to address this issue we conducted a Cochrane systematic review.

1.3.3.5 The efficacy and safety of a CDSS to support telephone triage

We designed and performed a superiority parallel group randomized controlled trial in order to test two hypotheses.

The first hypothesis was whether the availability of a CDSS reduces the calls closed by doctors at the phone as “hospital referrals”.
The second hypothesis was whether the availability of a CDSS increases the proportion of obligatory questions asked by doctors providing telephone triage in a primary care out of hours service.

1.3.3.6 Barriers ad facilitators to uptake CDSS

This part of the research seeks to identify potential barriers and facilitators to the adoption of EBM-focused CDSS linked to EHRs in a hospital setting. We plan to address the limitations of previous studies by considering: a) multiple health professionals, including physicians, nurses, and hospital managers and b) hospitals with different levels of CDSS infrastructure and use as well as experience in implementing EBM practice.

Why it is important to do this research

1.3.4.1 The quality of telephone triage

Results from previous studies show triagists seem to carry out a rapid clinical scan before they came to a conclusion, without considering in sufficient detail different causes for a symptom or its possible consequences. Telephone triage should aim at minimising risks to a patient health and the safety of telephone triage might be enhanced by using computer based decision support systems. No study has been performed in Italy so far about this issue.

1.3.4.2 The efficacy of CDSS

Several studies and reviews are available about efficacy of CDSS but the inclusion of studies with variable interventions across diverse health care settings precluded systematic reviews from reaching a decisive understanding of the impact of CDSSs.

1.3.4.3 The efficacy of e-learning

E-learning is gaining in popularity and e-learning programmes are rapidly increasing in number. Their relatively low costs, high flexibility, and reduced dependence on geographical or site boundaries are attracting the investments of stakeholders (countries, networks, and universities) and increasing the demands of learners. Previous systematic reviews have weakness that can limit their results and more precise data about the effectiveness of e-learning programmes have the potential to influence future investments regarding continuing medical education (CME) programmes. If e-learning is more or as effective as traditional learning, it will be possible to use it as a mean of provision of education about telephone skills and CDSS use to health professionals.

1.3.4.4 The efficacy of training intervention for improving telephone skills

There is an important role for telephone consultations within healthcare, so it is important to know which is the best way to provide the adequate skills to the relevant healthcare professionals.

1.3.4.5 The efficacy of CDSS to support telephone triage

Several studies show that not supported telephone triage could be unsafe while evidence from not randomized controlled trial highlights that supported telephone triage is safe. Results from observational studies suggest telephone triage could reduce costs by reducing unnecessary hospital referrals.

No randomised controlled trial has evaluated the impact of CDSS on doctors telephone triage outcomes so far.

1.3.4.6. Barriers and facilitators to uptake CDSS

Various factors must be considered when planning CDSS introduction in healthcare settings. The findings of this study has been planned to guide the development of strategies to facilitate their successful integration into the regular clinical workflow

METHODS

2.1 Efficacy and safety of INFs-beta vs GS in Relapsing Remitting Multiple Sclerosis

Criteria for considering studies for this review

Types of studies

We included randomised, double-blind or single-blind and open-label active control trials, as well as cross-over trials, comparing all types of IFNs versus GA in participants with RRMS. Studies including participants with secondary progressive multiple sclerosis (SPMS) were excluded. Quasi-randomised and cross-over trials were accepted but were not found. Randomised controlled trials (RCTs) designed with multiple groups were included only with regard to data provided by groups given GA and IFNs. Uncontrolled non-randomised trials, add-on trials (i.e. trials with drug associations) and observational studies were excluded. Trials with a follow-up period shorter than three months were excluded.

Trials comparing head-to-head different types and dosages of IFNs without a GA arm were excluded.

Types of participants

Patients of any age, gender and race affected by RRMS according to Poser's (Poser 1983) or McDonald's (McDonald 2001; Polman 2005; Polman 2011) criteria were included.

A relapsing-remitting course is characterised by relapses and remissions, with or without complete recovery between relapses.

Study participants had to have an Expanded Disability Status Scale (EDSS) score (Kurtzke 1983) of 0 to 6.0.

Types of interventions

We included trials in which participants received recombinant IFN-beta 1a (Rebif, Avonex) or IFN-beta 1b (Betaferon, Betaseron, Extavia) at any dose and by any route of administration in any setting.

Comparison: GA at any dose, route of administration and setting.

For trials comparing multiple groups of participants, only the following designs were considered.

IFN-beta 1a versus IFN-beta 1b versus GA.GA (dose 1) versus GA (dose 2) versus IFNs-beta.GA versus IFNs-beta (dose 1) versus IFNs-beta (dose 2).

Treatment duration had to be at least three months.

Types of outcome measures

Primary outcomes

CLINICAL EFFICACY OUTCOMES

1. Number of participants who experienced at least one relapse at 12 - 24 months and at the end of follow-up.

2. Number of participants whose condition worsened during the study; we defined worsening as a 0.5-point increase from starting EDSS score (Kurtzke 1983) ≥ 5.5 or a 1-point increase from starting EDSS score ≤ 5.0. Worsening must have been confirmed during two subsequent neurological examinations separated by at least six months of time free of relapses (Rudick 2010). We attempted to ascertain confirmed worsening at 12 and at 24 months and at the end of follow-up. Less stringent criteria (such as an increase in EDSS score sustained for three months) were considered.

CLINICAL SAFETY OUTCOMES

3. Number of participants who withdrew from or dropped out of the study because of adverse events (AEs), was defined according to the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2011) as unfavourable outcomes that occur during or after use of a drug but not necessarily caused by it, and/or as 'side effects' (any unintended effects, adverse or beneficial, of a drug that occur at doses normally used for treatment). Serious adverse events (SAEs) were included and were defined according to the US Food and Drug Administration (FDA) as participant outcomes such as death, life-threatening events, hospitalisation, disability, permanent damage or congenital anomaly/birth defect (FDA 2013).

Secondary outcomes

CLINICAL OUTCOMES

4. Frequency of relapse (number of relapses/patient-year: Annual relapse rate (ARR) was defined as the total number of relapses across all participants divided by total time on the study across all participants). Relapse was defined as newly developed or recently worsened symptoms of neurological dysfunction, with objective confirmation, lasting longer than 24 hours. Less stringent criteria (i.e. without objective confirmation) were considered.

5. Time to first relapse after the start of the study.

6. Percentage of participants free of disease activity: no relapses, no change in EDSS and no MRI changes (T1-T2).

7. Number of participants treated with steroids for relapse of MS.

8. Mean changes in quality of life (QOL) measured by validated questionnaires such as the Multiple Sclerosis Quality of Life-54 instrument (MSQOL-54) (Vickrey 1995).

MRI OUTCOMES

9. Mean number of active (new or enlarged) T2-hyperintense lesions per participant at 6-12-24 months from the start of the study and at the end of the scheduled follow-up period.

10. Mean number of new contrast-enhancing T1 lesions per participant at 6-12-24 months from the start of the study and at the end of the scheduled follow-up period.

11. Mean change in total T2-hyperintense lesion volume at 12-24 months from the start of the study and at the end of the scheduled follow-up period.

12. Mean change in total T1-hypointense lesion volume at 12-24 months from the start of the study and at the end of the scheduled follow-up period.
13. Mean change in total brain volume (as a measure of atrophy) at 12-24 months from the start of the study and at the end of the scheduled follow-up period.

Search methods for identification of studies

A systematic search with no restrictions was conducted to identify all relevant published and unpublished RCTs.

The Trials Search Co-ordinator searched the Trials Specialised Register of the Cochrane Multiple Sclerosis and Rare Diseases of the Central Nervous System Group (last search on 08 August 2016), which contains the following.

The Cochrane Central Register of Controlled Trials (CENTRAL) (08 August 2016).MEDLINE (PubMed) (1966 to 08 August 2016).EMBASE (EMBASE.com) (1974 to 08 August 2016).Cumulative Index to Nursing and Allied Health Literature (CINAHL) (EBSCO host) (1981 to 08 August 2016). Latin American and Caribbean Health Science Information Database (LILACS) (Bireme) (1982 to 08 August 2016).Clinical trial registries (http://clinicaltrials.gov, www.clinicaltrialsregister.eu). World Health Organization (WHO) International Clinical Trials Registry Portal (http://apps.who.int/trialsearch/).

Information on the Trial Register of the Review Group and details of search strategies used to identify trials can be found in the 'Specialised Register' section within the Cochrane Multiple Sclerosis and Rare Diseases of the Central Nervous System Group module.

The keywords used to search for studies for this review are listed in Appendix A.

Searching other resources

The search was extended to other resources, including:

1. screening of reference lists of review articles and primary studies found;
2. screening of abstract books of the main MS meetings (European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS), European Neurological Society (ENS), American Academy of Neurology (AAN)) from 2000 to 08 August 2016; and
3. contact with drug manufacturers.

Data collection and analysis

Two review authors independently determined the eligibility of the intervention by examining the study report and the description of the intervention. If necessary, we referred to other related papers or reports (e.g. protocol or register records) and sent requests to the study authors for additional information, especially if e-learning programmes were unclear or the measures to monitor outcomes changes were not clearly identified.

Multiple reports of the same studies were collated so that each study, rather than each report, was the unit of interest in the review.

Where means and standard deviations (SDs) were not reported in the original article, we sent request to the study authors for additional information.
We examined any relevant retraction statements and errata for information, and searched for any key unpublished information that is missing from the reports of included studies.
We used RevMan software to manage the included studies data (RevMan 2014).

Selection of studies

Titles and abstracts were independently reviewed by three review authors. If the study did not meet eligibility criteria, it was excluded. If the title and the abstract did not provide sufficient information, the full paper was obtained and was evaluated independently by the three previously mentioned review authors. If no consensus was reached on inclusion/exclusion criteria of an individual study, the final decision was made by all review authors. Study authors were contacted in cases of ambiguity or missing data.

Data extraction and management

Two couples of review authors quarried independently the data for each trial and summarised the information on a predefined collection form; this information was further confirmed by a third review author. Disagreements about extracted data were resolved by consensus with involvement of all review authors.

Assessment of risk of bias in included studies

Two review authors independently graded the selected trials according to the domain-based evaluation described in the Cochrane Handbook for Systematic Reviews of Interventions (Sterne 2011). The review authors compared evaluations and discussed and resolved disagreements.
Review authors assessed the following domains as 'yes' (i.e. low risk of bias), 'unclear' (uncertain risk of bias) or 'no' (i.e. high risk of bias): sequence generation, allocation concealment, blinding (of participants, personnel and outcome assessors), incomplete outcome data (when rate of dropout or loss to follow-up was greater than 20%, the risk was judged as high), s elective outcome reporting.
Overall quality of the studies was considered good if all domains of selection, attrition and detection were at low risk of bias, moderate if one was at high risk and poor in the other cases.

Measures of treatment effect

Data were processed according to a modified intention-to-treat (ITT) principle, using the number of randomly assigned participants who took at least the first dose of the drug. For dichotomous data, study results were summarised as risk ratios (RRs) with 95% confidence intervals (CIs) according to Mantel-Haenszel methods. When possible, we calculated the number needed to treat or to harm. For continuous outcomes, weighted mean difference (MD) and standardised mean difference (95% CI) methods were used.

When event rates were reported as the occurrence of events in the overall population over a specific interval in time, we used log-RR, and when the events were reported as hazard ratio (HR), we used log-HR.

Unit of analysis issues

Studies with parallel-group design were included: Participants randomly assigned to intervention or control were analysed at the individual allocation level. We decided to include cross-over studies by considering only data from the first half of the cross-over trial, but no cross-over studies were found. We performed a separate analysis at various periods (time frames: short term ≤18 months, medium term 24 and 36 months, and long term >36 months) of different outcomes based on different periods of follow-up.

Dealing with missing data

We addressed the effects of withdrawal and loss to follow-up by performing a sensitivity analysis (see Sensitivity analysis section). When we discovered that some data, with focus on predefined outcomes, were missing, we contacted the corresponding authors. Missing data are provided in the studies tables. Additional data from the BEYOND trial were provided by Bayer (O'Connor 2009; Pleimes 2013), and requests for additional data from the other study authors (Lublin, Mikol and Calabrese) or from drug companies (Merck Serono) were not answered.

Assessment of heterogeneity

Clinical diversity and methodological diversity have been considered as heterogeneity. Heterogeneity among trial results has been examined by using the I^2 test (Higgins 2003). This provides an estimate of the percentage of variability due to heterogeneity rather than to chance alone. We interpreted an I^2 estimate ≥ 50% as indicating the presence of heterogeneity, and random-effects models were applied (DerSimonian 1986). We investigated diversity in clinical and methodological aspects of included trials.

Assessment of reporting biases

The small number of trials included in this review did not permit an assessment of publication bias. In future updates, we will assess publication bias by following recommendations provided by the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2011).

Data synthesis

The meta-analysis was conducted using RevMan software (Review Manager 2013). If significant heterogeneity was not found among the included trials, we aggregated all included data in the final analyses by using a fixed-effect model (Yusuf 1985). If substantial clinical diversity was noted between included studies, we used the random-effects model with studies grouped by intervention. However, we decided to present results using a random-effects model.

Subgroup analysis and investigation of heterogeneity

Small numbers of trials and small quantities of data did not permit a subgroup analysis according to IFN type and baseline EDSS. In future updates, and if further data become

available, we plan to carry out subgroup analyses for primary outcomes according to IFN types beta 1a and 1b and baseline EDSS higher or lower than 3.0 points.

Sensitivity analysis

Sensitivity analyses were performed to explore missing data by likely scenario, attributing the outcome of interest to both treatment groups, as described in Section 16.2.2 (Higgins 2011).

Summary of findings table

These outcomes are included in the Summary of findings (SoF) table.

1. Number of participants who experienced at least one relapse.
2. Number of participants whose condition worsened during the study.
3. Number of participants who withdrew from or dropped out of the study because of adverse events.
4. Mean number of active (new or enlarging) T2 lesions.
5. Mean number of new contrast-enhancing T1 lesions.
6. Mean change in total T2-hyperintense lesion volume.
7. Mean change in total T1-hypointense lesion volume.

We assessed the quality of evidence as it relates to the studies that contributed data to the meta-analyses for prespecified outcomes using GRADEpro software (GRADEpro 2008). We justified all decisions to downgrade or upgrade the quality of studies by using footnotes, and we provided comments to aid readers' understanding of the review when necessary, as recommended by The Cochrane Collaboration (Schünemann 2011). The SoF table includes overall grading of the quality of evidence related to each of the outcomes, using the GRADEpro approach (GRADE Working Group 2004). Quality of evidence was graded as high, moderate, low or very low, upon consideration of within-study risk of bias, directness of evidence, heterogeneity, precision of effect estimates and risk of publication bias. Control event rates used in the calculation of absolute risks were based on the numbers of events reported in the included studies.

2.2 The quality of telephone triage

We evaluated the quality of telephone triage and the appropriateness of the decisions resulting from it, at two call centres in Verona. In this service, the call-handlers were doctors. A total of 22 OOH doctors were involved in the study, working in two different locations. Before the study began, the doctors were given an information letter about the study. All the doctors gave their consent to participate. The study was approved by the Ethics Committee and by the public company providing OOH care services (ULSS 20).

Four standardized clinical cases were used: an adult with nosebleed, an adult with fever, a child with fever and a child with vomiting. There was a set of obligatory questions for each case, translated from those used in a previous study (Derkx 2008). In the Italian translation, it was necessary to change the number of questions in order to express the same content (Table 1).

We also used the expected decisions in terms of how the clinical cases should have been managed following the MTC conducted by the primary care OOH centre, proposed in a previous study (Derkx 2008). The previous study used a focus group of General Practitioners (GP) with experience in telephone triage to define the lists of questions for each case study, based on national protocols. Our clinical cases were chosen so that only advice via telephone was required and the patient (or parent) was advised to call back only if necessary.

Different advice given by telephone triage call-handlers, such as advice to see their GP the next day, to contact the local health care centre to see a doctor or require a doctor's visit at home, or obtain a hospital referral, was considered incorrect. The four cases were handled by primary care OOH telephone triage staff in a four month study period. The cases were randomly proposed using the Incognito Standardized Patient (ISP) method (Derkx 2009, Rethans 2007, Moriarty 2003). Researchers phoned as ISP to one of the OOH services in the Verona city area and gave information concerning an illness without revealing that the case was simulated and not real. The clinical cases did not require the call-handler to advise a home visit or to call the emergency ambulance service. If the doctor answering the calls had given such advice, callers would have revealed their true identity immediately and the nature of the study. Over the 4-month period each of the four simulated clinical cases was used five times in calls to the two centres involved, i.e. there was a total of 40 calls. No sample size calculation was conducted before the study began.

Because of their monthly shifts, each telephone triage call-handler received about 1 call per month. The calls were made between 9:30 pm ans 11:30 pm for night shifts and from 11:00 am to 7:00 pm for day shifts, thus avoiding calling at times particularly inconvenient for the telephone triage staff (deep night and at the times of shift change). The researchers involved in the role of the Incognito Standardized Patient were three men and three women, aged 24–42 years. They received structured instructions and were supervised during their performance in order to reduce heterogeneity in the presentation of the cases.

To reduce the chance of repeating a clinical case to the same doctor, we took into account the shifts of the telephone triage participants. When making the calls we switched between simulated patients of both sexes, changing the patient names given to the doctor for every telephone call. We also used as many different telephones as possible, because the number was visible to the doctor and was stored in an archive of previous calls. The standardized clinical cases were administered with the support of a written framework where the ISP could take notes and flag the questions asked. All the calls were recorded and listened again if necessary.

For each call made we noted: (1) the proportion of obbligatory questions the call-handler asked compared to the questions on the standard reference list; (2) the proportion of correct decisions made by the call-handler; and (3) the length of the call itself. The duration of the simulated call was measured from the moment the conversation began until the parties hung up.

We could not link the OOH doctors demographic to performance data because ISPs were blind with regard to doctors' identity. Also, out-of-hours doctors do not usually introduce themselves at the beginning of telephone consultations and asking for the doctor's identity could suggest to the call-handler that the caller was not a real patient.

Table 1. Obligatory questions to be asked in each case

Adult with nosebleed

1. Duration of nosebleed?
2. How many nosebleeds have you had in the past 48 hours?
3. How much blood has been lost?
4. Have you vomited any blood?
5. Has there been a blow to the head?
6. Are you feeling faint?
7. Might you have a foreign object in the nose?
8. Do you have bruises anywhere?
9. Do you have a bleeding disorder?
10. Do you take anticoagulant treatment?

Adult with fever

1. Did the fever start with shaking/rigour?
2. How high is the temperature?
3. Do you have a stiff neck?
4. Do you have any difficulties with breathing?
5. Is there any pain when passing urine?
6. How much fluid has been taken in the past 12 hours?
7. When did you last pass urine?
8. Any major long term health problem?
9. Did you recently travel abroad?

Child with vomiting

1. Can you describe the child's behaviour now?
2. How often did the child vomit in the past six hours?
3. How much did the child vomit?
4. Did the child vomit blood?
5. Did the child drink in the past six hours?
6. When did the child pass urine the last time?
7. Did the child complain about pain when passing urine?
8. Does the child have a rash anywhere?
9. Does the child have fever?
10. Does the child complain about a headache?
11. Can the child touch forehead on knees(or kiss knees)?
12. Does the child complain about abdominal pain?
13. Does the child complain about photophobia?
14. Did the child have a head injury recently?
15. Did the child eat the wrong food?

Child with fever

1. Can you describe the child's behaviour now?
2. How high is the temperature of the child?

3. Has the child had a fit?
4. Does the child have pain anywhere?
5. Has the child got, or has he or she had, a headache?
6. Can the child touch forehead on knees (or kiss knees)?
7. Does the child seem breathless or is there indrawing of the chest/tummy?
8. How much fluid has been taken in the past 12 hours?
9. When did the child last pass urine?
10. Does the child have a rash?
11. Did the child recently travel abroad?

2.3 The efficacy of CDSS

Eligibility Criteria

Population

Postgraduate health professionals (medical, nursing, and allied health) in primary, secondary, and tertiary care settings. Only interventions that were implemented in real, non-simulated, clinical settings were considered.

Types of interventions

We adapted the definition of a CDSS by Haynes 2010 and Eberhardt 2012. We defined a CDSS as an information system aimed to support clinical decision-making, linking patient-specific information in EHRs with evidence-based knowledge to generate case-specific guidance messages through a rule- or algorithm-based software.

Our inclusion criteria emphasize the implementation of evidence based medicine, meaning that computer-generated guidance messages had to be based on literature or a priori evidence (e.g.,guidelines or point-of-care services) and not on expert opinions.

This knowledge had to then be delivered to medical doctors or allied health care professionals through electronic media (e.g.,computer, smartphone, or tablet).

We did not exclude a CDSS, however, based on the degree of literature it covered in the literature surveillance system. In other words, we included a CDSS if it integrated a single evidence-based guideline or incorporated multiple evidence-based guidelines. We also included CDSSs irrespective of the level of patient information archived in the EHR.

Systems that alter the guidance based on previous experience or average behaviors were excluded.

We included software guidance messages, irrespective of the form (e.g., recommendations, alerts, prompts, or reminders), as well as guidance messages, regardless of the target assistance (e.g., diagnostic test ordering and interpretation, treatment planning, therapy recommendations, primary preventive care, therapeutic drug monitoring and dosing, drug prescribing, or chronic disease management).

Patient-specific information had to derive from EHRs.

Our operational definitions for considering a study "compliant" with the EHR were inclusive: from clinical data repository and health data repository (CDHR), to electronic medical---patient record (EMR and EPR) (Gunter 2005)
Our inclusion criteria match the "6S" Haynes' model for evidence based literature products (Dicenso 2009) and the evolution of online point of care services (Moja 2011).

Types of comparison groups.

To address our objectives, we considered standard care with no access to CDSS, CDSS that do not generate advice or CDSSs that are not based on evidence. Trials comparing arms accessing the same CDSS at different intensities (e.g., one arm having guidance messages pushed to the health professional vs another arm having guidance message statically available in a folder) were not pooled together with the other trials in the quantitative analyses.

Types of outcomes and assessment measures.

We identified a priori the following (primary) outcome measures for included studies:
1. Mortality: we selected mortality as it is the most relevant and objective outcome, although there may exist variability across studies with regards to the time frame during which mortality is captured.
2. Morbidity: we selected and grouped objective patient of illness (e.g., pneumonia, myocardial infarction, stroke), progression of diseases and hospitalizations.
3. Economic outcomes: Information about health care utilization (e.g., length of stay, emergency department visits, and primary care consultations) and costs. We did not consider the following outcomes: patient satisfaction, measures of process, and health care professional activity or performance (e.g., adherence to guidelines, rates of screening and other preventive measures, provision of counseling, rates of appropriate drug administration, and identification of at-risk behaviors).

Types of studies.

To be eligible, studies had to be randomised controlled trials (RCTs). Randomization was allowed to be either at the individual- or at the cluster-level.

Data Sources

We systematically searched the English-language literature indexed in MEDLINE, EMBASE, Cochrane Central Register of Controlled Trials and Cochrane Database of Abstracts of Reviews of Effects. Studies found in the bibliographies of Systematic Reviews on CDSS, as well as those identified by experts, were also considered.

Study Selection and Data Extraction

We identified RCTs of the CDSS fulfilling the aforementioned eligibility criteria. We combined the results into a reference management software program (EndNote X5 for Windows, Thomson Reuters, Philadelphia, PA). The database was filtered for duplications to derive a unique set of records.

We independently examined the search results and screened the titles and abstracts; the full text reports of all potentially relevant trials were subsequently screened. We independently abstracted information on CDSS characteristics and effect estimates from all included trials using a modified version of The Cochrane Effective Practice and Organisation of Care Review Group (EPOC) data collection checklist: study setting and methods (design), comparators, computerized CDSS characteristics, patient or provider characteristics, and outcomes. We performed all steps in the study selection and data extraction processes in duplicate.
When necessary, we attempted to contact the study authors to clarify uncertainties in the study design or results.

Risk of Bias Assessment

We assessed the potential risk for bias in included studies using the criteria outlined in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2014). The assessment involved the following key domains: sequence generation, allocation concealment, blinding of outcome assessors, incomplete outcome data, selective outcome reporting, and other sources of bias (e.g., extreme baseline imbalance or failure to disclose source of funding for the study). We did not assess the blinding of personnel and participants given the nature of the intervention. In fact, the use of masking procedures to prevent personnel and participants from knowing the allocation to the intervention or control arms was impractical. Furthermore, blinding does not affect mortality, an outcome of this review. Our assessment referred only to studies reporting mortality or morbidity outcomes. Any disagreement was resolved by discussion or by the involvement of a third investigator.

Data Synthesis

Risk ratios and 95% confidence intervals (CIs) were calculated for each trial by reconstructing contingency tables based on the number of patients randomly assigned and the number of patients with the outcome of interest (analysis in accordance with the intention to treat principle). For the cluster randomized trials, to calculate adjusted (inflated) CIs that account for the clustering, we performed an approximate analysis as recommended in the Cochrane Handbook (Higgins 2014). Our approach was to multiply the standard error of the effect estimate (from the analysis ignoring the clustering) by the square root of the design effect. For this, we used an intracluster correlation coefficient (ICC = 0.027) borrowed from an external source (Health Services Research Unit 2014). Then, each meta-analysis was performed twice, assuming either a fixedeffects or a random-effects model (DerSimonian 1986, Mantel 1959). In the absence of heterogeneity, the fixed-effects and the random-effects models provide similar results. When heterogeneity is found, the random effects model is considered to be more appropriate, although both models may be biased (Pettiti 1999). For all statistical analyses we used the R software environment (R Core Team 2013), version 3.0.1, and the “meta” package for R (Schwarzer 2013), version 2.3-0. Selective outcome reporting or publication bias was assessed using the Begg and Mazumdar adjusted rank correlation test (Begg 1994) and the Egger regression asymmetry test (Egger 1997). To evaluate whether the results of the studies were homogeneous, we used the Cochran Q test with a 0.10 level of significance (Cochran

1964). We also calculated the I^2 statistic (Higgins 2003) that describes the percentage variation across studies that is attributed to heterogeneity rather than chance. We regarded an I^2 value less than 40% as indicative of "not important heterogeneity" and a value higher than 75% as indicative of "considerable heterogeneity" (Higgins 2014). To evaluate the stability of the results, we also performed a "leave-one-out" sensitivity analysis.

The scope of this approach was to evaluate the influence of individual studies by estimating the summary relative risk in the absence of each study (Tobias 1999). All P values are 2-tailed. For all tests (except for heterogeneity), a probability level less than .05 was considered statistically significant.

2.4 The efficacy of e-learning

Criteria for considering studies for this review

Types of studies

We included randomised controlled trials (RCTs) and cluster RCTs. We used the Cochrane definitions and criteria for RCTs (Higgins 2011). We excluded non-randomised trials (e.g. controlled before-after studies or interrupted time series) as they are prone to a wider range of potential risks of bias and add little to what is known when sufficient evidence is available from RCTs (EPOC 2013). Non-randomised quality-improvement intervention trials often overstate the strength of causal inference between intervention and outcomes compared to RCTs (Li 2009). Conclusions from meta-analyses exploring the causality of e-learning might be undermined if largely based on studies that adopt intrinsically weaker research designs (Banzi 2009).

We included studies published in all languages and providing data about any follow-up periods.

Types of participants

We included studies assessing e-learning programmes aimed to improve patient outcomes or behaviours, skills or knowledge of licensed health professionals (doctors, nurses and allied health professionals). We focused on post-graduate education, considering the license to practice without supervision as a discriminating factor, that is, health professionals who can fully practice a specific health-related profession versus those who cannot. Only those licensed to practice were included in this review. If the description was not sufficient, we sent requests to the study authors for additional information before excluding the studies.

We excluded studies recruiting undergraduate students, trainees and residents or a mix of licensed and unlicensed participants if data on the eligible participants were not separately provided by the authors after a formal request by email.

Types of interventions

Definition of e-learning programme

Any intervention in which clinical content is distributed and facilitated primarily by Internet, extranet or Intranet: web-based tutorials, virtual clinical vignettes and patients, discussion groups, Internet-mediated videoconferencing, web seminars, emails, pod casts, and virtual

social networks was included. We excluded CD-ROMs and applications not distributed through the media mentioned above. The learners may have had access to interventions through a variety of technologies [e.g. computers, personal digital assistant (PDA), smart phones, etc]. We applied no restrictions on the basis of the programme length: we included short programmes such as single lectures, workshops, and modules, as well as more extended educational programmes. We included an intervention if the description is sufficient to allow us to establish whether it could potentially improve knowledge or behaviours by any kind of intervention mentioned above and when the description was not sufficient, we sent a request to the study authors for additional information before excluding the studies.

We excluded e-learning programmes focusing on non-clinical medical topics (e.g. bio-terrorism) defined as subjects different from the seven roles that all physicians need to have to be better doctors: medical expertise, communication, collaboration, leadership, health advocacy, scholarship, and professionalism (The CanMEDS Framework).

We only included interventions in which e-learning is considered a core or essential element. However, in multifaceted educational interventions (e.g. those applying two or more interventions to change health professionals' practice), the e-learning component may have different degrees of centrality. Thus, we categorised studies into three groups: e-learning alone; e-learning as a core, essential component of a multifaceted intervention; e-learning as a component of a multifaceted intervention but not considered core and essential.

A study was classified 'core' when e-learning was described to provide the foundation for the entire educational intervention (e.g. e-learning together with papery guideline dissemination). When the components other than e-learning could be used in the absence of e-learning or e-learning was merely added to a multifaceted intervention that could easily be offered in its absence (e.g. audit and feedback interventions), the study was classified as 'not core'.

We included RCTs where the eligible comparators were educational interventions on the same topic without access to e-learning (e.g. paper-based material, residential face-to-face courses) or multifaceted educational interventions without e-learning on the same topic.

Types of outcome measures

This review included studies reporting the following outcomes: patient outcomes or behaviours, skills or knowledge of health professionals (Kirkpatrick 1996, Straus 2004).

Several conceptual models have been developed for assessing knowledge and competence. Miller 1990 identifies four stages of development: 'knows, knows how, shows how, and does' that are the essential facets of clinical competence. An individual progresses through cognitive and behavioural steps, from acquiring knowledge to performing a task in practice. Miller's theory assumes that knowledge and competence predict behaviours. We slightly modified the stages of development to better reflect the different components targeted by educational interventions in clinical practice (Moja 2008):

1. knowledge defined as factual knowledge or basic learning, for example knowing the benefits and risks of different interventions (e.g. in patients with unstable angina, aspirin is beneficial);

2. skills defined as deep learning or competence, for example posing structured clinical questions considering patients, treatment, comparison, and outcomes and understanding quantitative aspects (e.g. relative or absolute risk reduction, number needed to treat);
3. behaviours defined as ability or performance: the incorporation of knowledge into practice, with the adoption of proven treatments and interventions that can potentially improve patients' health;
4. patient outcomes.

Non-objectively assessed outcomes were excluded (e.g. learner self-reported knowledge or satisfaction).

Primary outcomes

Patient outcomes and health professionals' behaviours: any objective measure of patient outcomes (e.g. blood pressure, number of caesarean sections, medical errors) or professional performance (e.g. number of tests ordered, prescriptions for a particular drug) was considered as a primary outcome measure.

An intervention was included if the description is sufficient to allow us to establish that it was aimed at improving clinical practice, such as starting effective treatment, or dismissing ineffective or harmful treatment; if the description was not clear, we sent a request to the study authors for additional information before excluding the studies.

Primary outcome results were computed at two major time points: immediately after the e-learning intervention at the longest duration of follow-up available.

Secondary outcomes

Health professionals skills and knowledge: skills and knowledge are clinical competence dimensions related to the concept of 'knows' (knowledge) and 'knows how' (skills) (Miller 1990). We included studies reporting any objective measure of skills such as the assessment of learners' ability to demonstrate a procedure or technique (e.g. problem solving, objective structured clinical examination scores) or any objective measure of learners' knowledge such as assessment of factual or conceptual understanding (e.g. multiple-choice test of knowledge).

Search methods for identification of studies

Electronic searches

The EPOC Trials Search Co-ordinator (TSC) wrote the search strategies in consultation with the authors. The TSC searched the Cochrane Database of Systematic Reviews (CDSR) and the Database of Abstracts of Reviews of Effects (DARE) (The Cochrane Library) for related systematic reviews, and the following databases for primary studies.

MEDLINE, 1946 to 7 July 2016, Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE Daily and Ovid MEDLINE, OvidSPEMBASE, 1980 to 7 July 2016 week 6, OvidSPCochrane Central Register of Controlled Trials, Issue 6, 2016, WileyHealth Technology Assessment, Issue 2, 2016, WileyNHS Economic Evaluation Database, Issue 2, 2016, WileyDatabase of Abstracts of Reviews of Effects, Issue 2, 2016, Wiley

The MEDLINE strategy is provided in Appendix B and was translated using appropriate syntax and vocabulary for other databases. Results were limited by two methodological filters: the Cochrane Highly Sensitive Search Strategy for identifying randomised trials in MEDLINE (2008 revision; Lefebvre 2011) to identify randomised trials, and an EPOC methodology filter. We applied no language or date limit.

Trial Registries

We searched the following trial registries for ongoing and completed trials;
International Clinical Trials Registry Platform (ICTRP), Word Health Organization (WHO) http://www.who.int/ictrp/en/ClinicalTrials.gov, US National Institutes of Health (NIH) http://clinicaltrials.gov/. We also reviewed reference lists of all included studies, relevant systematic reviews and primary studies.

Searching other resources

We examined the reference lists of the included trials and relevant reviews published in the field of e-learning (e.g. Wentling 2000, Chumley-Jones 2002, Wutoh 2004, Ruiz 2006, Cook 2008a, Lam-Antoniades 2009).

Data collection and analysis

Two review authors independently determined the eligibility of the intervention by examining the study report and the description of the intervention. If necessary, we referred to other related papers or reports (e.g. protocol or register records) and sent requests to the study authors for additional information, especially if e-learning programmes were unclear or the measures to monitor outcomes changes were not clearly identified.
Multiple reports of the same studies were collated so that each study, rather than each report, was the unit of interest in the review.
Where means and standard deviations (SDs) were not reported in the original article, we sent request to the study authors for additional information.
We examined any relevant retraction statements and errata for information and searched for any key unpublished information that is missing from the reports of included studies.
We used RevMan software to manage the included studies data (RevMan 2014).

Selection of studies

Two review authors independently screened the titles and abstracts and applied inclusion and exclusion criteria. We searched for complete manuscripts in the cases of uncertainty, and resolved disagreements through discussion and consensus.
We documented the studies selection process in a Prisma flow diagram (Liberati 2009).

Data extraction and management

Two authors independently extracted data from the included studies, using a data sheet developed as a modified version of the EPOC data collection checklist (EPOC 2015).
We extracted the following information.

Characteristics of participants: total number at baseline, total number at completion of the study, and type of target health professionals. Interventions and controls: number of groups, interventions applied, frequency, duration and main components. Methods: study design, duration of the study, setting and provider. Outcomes: type of outcome measures, scales of measure, values for means and standard deviations. Results: measures at follow-up (including means and SD/standard errors (SEs)/CIs for continuous data and summary table for dichotomous data), withdrawals and loss to follow-up.

We resolved any disagreement by discussion and consensus among the authors.

Assessment of risk of bias in included studies

Two reviewers independently assessed the quality of all eligible studies using the EPOC risk of bias criteria (EPOC 2013). Any discrepancies in quality ratings was resolved by discussion and consensus between the authors. We collected the sources of information for each risk of bias assessment (e.g. quotation, summary of information from trial reports, correspondence with investigators). For each study, we assessed the following components for risk of bias:

1. Adequate generation of allocation sequence
2. Adequate concealment of allocation sequence (selection)
3. Baseline outcome measurement similarity
4. Baseline participants' characteristics similarity
5. Incomplete outcome data reporting (attrition)
6. Adequate prevention of knowledge by outcome assessors about allocation to the interventions (blinding)Adequate protection against contamination
7. Freedom from selective outcome reporting
8. Freedom from other risk of bias (e.g. duality of interest)

We summarised the overall risk of bias for the single studies, considering the risk of bias in the concealment of allocation sequence, incomplete outcome data reporting and prevention of knowledge of outcome assessors about allocation to the interventions (Wood 2008, Savovic 2012, Chan 2004, Dwan 2008, Kirkham 2010) as key domains. The overall risk of bias of the single study was judged as high if one of these items was rated at high risk of bias and as low if all of the items were low. We used the risk of bias of the single studies in the sensitivity analysis as detailed below.

Measures of treatment effect

We separately analysed patient outcomes, health professionals' behaviours, skills and knowledge.

When possible, we calculated the outcome measures in accordance with the intention-to-treat principle (i.e. all data were analysed according to randomised group assignment, regardless of whether some of the participants violated the protocol, were not compliant or were lost to follow-up). Accordingly, we contacted study authors to obtain additional primary trial data, when necessary.

We based analyses upon the consideration of dichotomous (e.g. proportion of patients managed according to e-learning programme) or continuous process measures (e.g. change in

learners' knowledge scores). Where studies reported more than one measure for each endpoint, we planned to abstract the primary measure (as defined by the study authors) or the median measure identified. For example, if the comparison reported five continuous knowledge test variables in which none of them were denoted as the primary variable, we ranked the effect sizes for the five variables and took the median value.

We extracted the outcomes from each study in natural units. We planned to combine final values if the same scale was used across studies, or convert the effect size back into the natural units of the outcome measure most familiar to the target audience or provide standardized effect size.

We only included continuous data from a trial in the analyses if means and SD were available or can be calculated and there was no clear evidence of a skewed distribution (e.g. as indicated by the ratio between the difference between the minimum or maximum value of the scale and the SD (Deeks 2011)).

In cases where some trials reported changes in scores from baseline to final values as the measure of treatment effect for continuous outcomes, while other trials reported only the final values, we combined both the final values and the changes in scores into the same meta-analysis (Higgins 2011). For studies providing both measures of treatment effect, we used the final scores as, due to randomisation, we did not expect differences between experimental and control group baseline scores.

We planned to use results from both periods of cross-over trials, unless there was a reason to believe a carryover of effects from one period to another, which presents a serious flaw. For cross-over trials, we planned to use paired estimates of the effect (e.g. means and its SE), or calculated them from the exact statistical test results (e.g. paired t-test for continuous data or McNemar's test for binary outcomes) (Elbourne 2002, Cook 2008).

Unit of analysis issues

Studies with more than two arms

If more than one comparison from a study with more than two arms was eligible for the same comparison, we planned to adjust the number of health professionals to avoid double counting. We sought to make the adjustment by dividing the number of health professionals in the shared arm approximately evenly among the comparisons.

Cluster RCTs

Due to the focus on an educational intervention, we expected that trials were randomised by cluster. In cluster-RCTs or “clusters”, groups of individuals are randomly allocated to study arms and outcomes are then measured based on the individual cluster members. Under such circumstances, it is necessary to adjust the results from primary trials for clustering before they are included in the meta-analysis in order to avoid spurious precision in 95% CIs. We included cluster-RCTs with adequate definition of participants and clusters, as suggested by the Ottawa Statement for cluster-RCTs (Weijer 2012).

For the cluster-RCTs, in order to calculate adjusted (inflated) CIs that account for the clustering, we planned to proceed to an approximate analysis. Our approach was to multiply the SE of the effect estimate (from the analysis ignoring the clustering) by the square root of the design effect. For this, we used intra-cluster correlation coefficients borrowed from an external source (University of Aberdeen 2015).

Performing meta-analysis using studies with unit of analysis errors required us to make a number of assumptions about the magnitude of unreported parameters, such as the intra-class correlation coefficients and the distributions of patients across clusters. We planned to re-analyse studies with potential unit of analysis errors where possible. If a study was re-analysed, we reported the re-analysed results (observed SEs, P values, or CIs) along with the original results. If this was not possible, we reported only the original results and the study was excluded from the meta-analyses.

Dealing with missing data

For all outcomes across all studies, we carried out analyses as far as possible on an intention-to-treat basis (i.e. we attempted to include all participants randomised to each group in the analyses, and analysed all participants in the group to which they were allocated, regardless of whether or not they received the allocated intervention). If intention-to-treat data were not available, or for dichotomous and continuous data that were missing, we made no assumptions about loss to follow-up, but we based analyses on participants completing the trial. If there was a discrepancy between the number randomised and the number analysed in each treatment group, we calculated and reported the percentage of loss to follow-up in each group.

Where standard deviations were not specified, we calculated them using a test of significance (e.g. P value related to t or F statistic) wherever possible. If P values and variance were not forthcoming, because studies used similar measurement scales, featured similar degrees of measurement error, or had similar time periods, we planned to use the average SD from the other similar studies (Cook 2008).

The impact of missing data was considered separately for each primary and secondary outcome reported in each study.

Assessment of heterogeneity

To assess the contextual heterogeneity of the included trials (the differences in populations, context, interventions, comparators, follow-up), we planned to conduct subgroup analyses according to important clinical and methodological characteristics, such as settings, interventions, comparators, etc. Between-study heterogeneity was planned to be assessed overall, and within the subgroups.

We included all the pre-specified outcomes available from the individual studies in the meta-analysis, with heterogeneity reported by the Q (Chi^2) and the I^2 statistics (Deeks 2011). The I^2 describes the percentage of the variability in effect estimates that is due to heterogeneity rather than chance (sampling error). The Cochrane Handbook (Higgins 2011) gives the following guidance on this decision based on I^2 values to classify the inconsistency of the effect measures across studies:

0% to 40%: might not be important;
30% to 60%: may represent moderate heterogeneity;
50% to 90%: may represent substantial heterogeneity;
75% to 100%: considerable heterogeneity.

In cases of moderate/substantial heterogeneity, we performed the analysis using both the fixed and the random-effects model. Where considerable heterogeneity existed, we explored the magnitude and direction of the effects: if I^2 was more than 75%, but the large majority of studies were in the direction of benefit and a random-effects meta-analysis yielded highly statistically significant benefit, we accepted the results. In this scenario, we could be uncertain about the amount of benefit, but not about its existence; it is safe to conclude that the intervention is beneficial (Virgili 2009). If substantial heterogeneity existed, studies were sparse, or directions were discordant, we did not pool data from the trials and we did not conclude in favour of or against the intervention.

Assessment of reporting biases

We planned to use funnel plots to assess the reporting biases. We planned to evaluate the funnel plot asymmetry, not only visually, but also with the use of tests for funnel plot asymmetry if we found more than 10 studies to include in meta-analysis. We planned to use the test proposed by Egger 1997 and by Harbord 2006 for continuous and dichotomous outcomes, respectively. If asymmetry was detected, we discussed possible explanations (e.g. publication bias or poor methodological quality of the studies) on the basis of available information (Higgins 2011) and, subsequently, performed a sensitivity analysis. Funnel plots were interpreted cautiously as they may be misleading.

Data synthesis

The studies were grouped according to important clinical and methodological (conceptual) characteristics, such as settings, interventions, comparators, etc. Accordingly, similar studies reporting homogeneous (similar) outcomes and outcome measures were synthesised.

We entered outcomes into Review Manager (RevMan) 5.3 (RevMan 2014) as effect sizes and their SEs.

We assessed the certainty of evidence as it relates to the studies that contributed data to the meta-analyses for pre-specified outcomes using GRADEpro software (GRADEpro 2008). We justified all decisions to downgrade or upgrade the rating using footnotes, and we provided comments to aid readers' understanding of the review when necessary, as recommended by The Cochrane Collaboration (Schünemann 2011). The certainty of evidence was graded as high, moderate, low or very low, downgrading the initial level of confidence considering the risk of bias, inconsistency and indirectness of evidence, imprecision of effect estimates and risk of publication bias.

Subgroup analysis and investigation of heterogeneity

We performed the following subgroup analyses:

- Change in content: e-learning programmes sub-grouped by medical, surgical or rehabilitation topics with the hypothesis that e-learning about medical topics (more likely to be centred on knowledge than skills or behaviours) are more effective than other topics.
- Change in targeted health professionals: doctors, nurses or physiotherapists with the hypothesis that e-learning for doctors are more effective than for other health professionals.
- Change in regulation: formally accredited versus non-accredited e-learning programmes with the hypothesis that accredited e-learning programmes are more effective than not-accredited.
- Change in format: high-interactive programmes (combination of at least three components, e.g. web module, chat, emails) or low-interactive programmes (fewer than three components); short (i.e. attendance less than one week) or long programmes (more than one week)
- with the hypothesis that high-interactive and short programmes are more effective than low-interactive and long programmes.

Some of these factors have been identified by others as those which may influence the effect of educational e-learning programmes (Ruiz 2006, Cook 2008). We undertook the standard test for heterogeneity across subgroup results to investigate the differences between two subgroups (Borenstein 2009). We used these analyses to investigate potential sources of heterogeneity and reported them as post-hoc exploratory data analyses only.

Sensitivity analysis

We planned to perform the following sensitivity analyses:

- excluding studies assessed as at high risk of bias.
- excluding cross-over trials.

We decided to aggregate studies at unclear risk of bias to those at high risk of bias. This decision was not specified in the protocol. We assumed a conservative approach, assuming that the quality was inadequate when the information was not provided at all ('guilty until proven innocent') (Moja 2014).

2.5 Training intervention for improving telephone skills

Criteria for considering studies for this review

Types of studies

- We considered the following types of studies meeting the minimum criteria used by the Cochrane Effective Practice and Organisation of Care Group (EPOC 2013).
- Randomised controlled trials (RCTs)
- Non-randomised controlled trials (NRCTs)
- Controlled before-after (CBA) studies with a minimum of two intervention and two control sites

- Interrupted time series studies (ITS) of interventions with a clearly defined point in time when the intervention occurred and at least three data points before and three after the intervention

This decision was based on our initial evidence search that identified few randomised controlled trials in this area.

Types of participants

We included clinicians (a broad term that encompasses all doctors, nurses and other health professionals) who have undergone educational interventions for developing and improving telephone consultation skills with patients. We included studies from all settings including primary care, outpatient, inpatient and public health. We excluded studies regarding communication between clinicians.

Types of interventions

We considered any kind of intervention aiming at improving the clinicians' telephone consultation skills regardless the mean and the way they were delivered (computerised, written, face-to-face training programmes or decision support software).

The eligible comparators were any control intervention with a possible effect on the same outcomes set or the absence of intervention.

Consequently the comparison could be, for example, an interactive e-learning programme on telephone consultation structure versus a classroom intervention on the same topic or no intervention; or a computerised decision support software use versus written management algorithms or no intervention.

Types of outcome measures

We considered the following types of outcomes:

Primary outcomes

- Patient outcomes
- health outcomes (e.g. validated tools, biomedical markers and patient behaviour)
- effect upon morbidity/mortality
- patient satisfaction
- urgency assessment accuracy
- adverse events
- Clinicians' telephone consulting skills as measured/assessed by a validated tool (e.g. RICE tool, Derkx 2007);

We use the definition of "validated tool" provided by the Joint Commission: a tool is validated when it is "*an instrument that has been psychometrically tested for reliability (the ability of the instrument to produce consistent results), validity (the ability of the instrument to produce true results), sensitivity (the probability of correctly identifying a patient with the condition)*" (https://manual.jointcommission.org/Manual/Questions/UserQuestionId03Sub0015)

Secondary outcomes

- Clinician knowledge gain
- Attitudes to telephone consultation (e.g. confidence, satisfaction)
- Time effectiveness (length and frequency of consultations, avoidance of face-to-face contact, effect on further clinical contact)
- Referral patterns
- Economic evaluation (litigation issues, resource issues, time effectiveness)

We only included studies if they assessed primary outcomes (e.g. not those with just secondary outcomes).

Search methods for identification of studies

The EPOC Information Specialist (IS), developed the search strategies in consultation with the review authors and ran the searches of The Cochrane Database of Systematic Reviews and the Database of Abstracts of Reviews of Effects (DARE) for related systematic reviews and the databases listed below for primary studies. The most recent search was conducted on May 19, 2016.

We searched the following databases:
Cochrane Central Register of Controlled Trials via OVID (May 2016)
Cochrane Library via Wiley (May 2016) including
Database of Reviews of Effects (DARE)
Economic Evaluation Database (EED) and
Health Technology Assessment Database (HTA)
MEDLINE via OVID (1946 to May 2016)
EMBASE via OVID (1947 to May 2016)
CINAHL via EbscoHost (1980-May 2016)
EPOC specialized register via Reference Manager (to 2012)
We searched the Cochrane EPOC Group Specialised Register only to 2012 as it has not been updated since that time.

We did not apply language nor date restrictions to the searches. We used two methodological search filters to limit retrieval to appropriate study designs: the Cochrane Highly Sensitive Search Strategy (2015 revision) to identify randomised trials (cf. Cochrane Handbook for Systematic Reviews of Interventions 6.4d) (Lefebvre 2011). To retrieve non-randomised controlled trials, controlled before/after studies (CBAs) and interrupted time series (ITSs) the Effective Practice and Organization of Care (EPOC) Group Methods Filter 2.6 (January 2013 version) was applied. For other databases, where no filter exists, study designs were identified at the screening stage (see Types of studies). Detailed search strategies used for searches are provided in Appendix C.

The search strategy was devised for the OVID Medline interface and then adapted for the other databases. Relevant individuals and organisations were consulted for information about unpublished or ongoing studies.

Searching other resources

We searched the following trial registries and additional thesis resources; selected grey literature and Google Scholar (we screened the first 500 items retrieved).

Trial Registries

WHO International Clinical Trials Registry Platform (ICTRP) http://www.who.int/ictrp/en/

ClinicalTrials.gov http://clinicaltrials.gov/

TrialsCentralTM (www.trialscentral.org)

Current Controlled Trials (www.controlled-trials.com)

Theses Portals

Australasian Digital Theses Program (http://adt.caul.edu.au/)

EThOS, Electronic Thesis Online Service, British Library (http://ethos.bl.uk)

Networked Digital Library of Theses and Dissertations (http://www.ndltd.org)

Index to Theses (http://www.theses.com/) (Great Britain and Ireland)

We also:

a) Screened individual journals and conference proceedings (e.g. hand searching).

b) Reviewed reference lists of relevant systematic reviews or other publications.

c) Contacted authors of relevant studies or reviews to clarify reported published information or seek unpublished results/data (when necessary).

d) Contacted researchers with expertise relevant to the review topic or EPOC interventions.

e) Conducted cited reference searches in ISI Web of Science/Web of Knowledge.

Data collection and analysis

Selection of studies

Two pairs of review authors independently assessed the eligibility of all titles and abstracts identified from electronic searches. We retrieved full text copies of all articles judged to be potentially eligible. The same review authors independently assessed the retrieved articles to determine whether they met the inclusion criteria.

The final list of included and excluded studies was agreed between the authors. Where there was insufficient detail about the study to decide whether it met the inclusion criteria, we contacted the study authors to enable a more informed decision. If necessary, another review author was asked to resolve any potential differences of opinion.

Data extraction and management

Two review authors independently extracted data from all included studies using a standard data recording form derived from the data extraction template provided by the Cochrane EPOC Group EPOC 2013a. We compared results and resolved disagreements by discussion and,

when necessary, through the involvement of an third review author. We contacted study authors to obtain or clarify data from included studies. We planned to use RevMan 2014 to manage the study data.

Assessment of risk of bias in included studies

Two review authors independently assessed the risk of bias of the included study using the nine standard criteria for RCTs and the seven standard criteria for ITS as outlined by the Cochrane EPOC Group (EPOC 2015). We planned to use a template to guide our assessment of risk of bias, judging each item as having a low risk of bias, a high risk of bias or 'unclear' risk of bias and providing a description to explain the decision using the guidance outlined by the Cochrane EPOC Group (EPOC 2015) and in the section 8.3 of Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2008). We compared judgments and resolved any disagreements by discussion and consensus, and by consulting a third review author where necessary.

Low risk of bias within selected studies was decided if all the above mentioned elements were deemed to be low risk. Conversely, if one or more of these key elements were found to have a high risk of bias, the selected study was planned to be classified as at high risk.

When necessary we contacted study authors for additional information about the included studies.

Measures of treatment effect

We reported the findings of the included study in narrative form as described by the study authors. When further studies are identified and included in this review, we plan to analyse effect measures in relation to the primary outcome measures to assess whether there are definable and significant changes in a variety of outcomes after the training intervention. We anticipate that with additional studies, the primary outcomes will reveal data that can be assessed by measures such as mean difference (MD), standardised mean difference (SMD) and proportions where appropriate.

For dichotomous data: where feasible, we plan to analyse outcomes with dichotomous data (such as confidence rating scales) with relative effect.

For continuous data: we plan to report the mean difference (MD) or standardised mean difference (SMD) (if there was a difference in measurement of scales across trials), using 95% confidence intervals (CI) as measures of the amount of random errors influencing the outcome estimates.

Future included studies, could use standardised assessment tools of consulting (such as Pendleton's Consultation Rating Scale). These again, could be measured using MD and SD or SMD if different tools were used. If medians are used, then interquartile ranges (IR) will be measured. Where total numbers and effect sizes are not recorded then we will describe results narratively.

Unit of analysis issues

We did not evaluate unit of analysis issues as only one study was included. When future studies are included, and cluster trials identified, we plan to analyse the data according to recommendations in the Cochrane Collaboration Open Learning Module on issues related to the unit of analysis (Alderson 2002).

Dealing with missing data

We contacted the authors of the included study for missing data.

Assessment of heterogeneity

We did not assess heterogeneity as only one study was included. When future studies are included, we will evaluate heterogeneity using tables and box plots to compare effect sizes of studies grouped according to potential effect modifiers. These include:

1. Type of health professional.
2. Type of intervention.
3. Duration of education/intervention.
4. Outcomes of intervention.
5. Setting and contextual factors: primary/secondary care, face-to-face/eLearning
6. Study design (e.g. RCT, CCT, CBA, ITS).
7. Methodological quality of studies.

With additional included studies, we expect to find substantial variation in the study results due to differences in types of interventions, the type of healthcare professional (targeted population), the design of the intervention, duration of the intervention and the context in which the intervention was implemented. We will conduct subgroup analyses based on type of intervention, type of health professional and study setting when two or more studies considering the same outcomes or using the same intervention in a similar population.

Assessment of reporting biases

We did not assess reporting bias as only one study was included. When future studies are included, we plan to use funnel plots to assess for the potential existence of small study bias. As there are a number of explanations for asymmetry in a funnel plot (Sterne 2001), we plan to carefully interpret results (Lau 2006).

Data synthesis

We did not perform quantitative analysis as only one study was included.

We summarized the findings for each primary outcome in a 'Summary of findings' table to draw conclusions about the certainty of the evidence for the main comparison within the text of the review. Two review authors independently assessed the certainty of the evidence (high, moderate, low, and very low) using the five GRADE considerations (study limitations, consistency of effect, imprecision, indirectness and publication bias) reported in the specific guidance developed by EPOC (EPOC 2013). We resolved disagreements on certainty ratings

by discussion and provide justification for decisions to down- or up-grade the ratings using footnotes in the table(s).

When future studies are included, we plan to begin the data synthesis with a narrative overview of the findings and a table systematically summarising the extracted results. We will assess the participants, interventions and outcomes for comparability, which is necessary for statistical pooling. We will look for studies sufficiently similar in terms of study design, setting, intervention, follow-up and outcome measures in order to combine the study data in a meta-analysis. We plan to review the appropriateness to carry out a meta-analysis collectively as a review team.

The choice of model will depend on the heterogeneity of the studies included in the meta-analysis. We plan to conduct the analysis according to the Cochrane Handbook for Systematic Reviews of Interventions guidance (Higgins 2011). If meta-analysis is feasible, we plan to use a random-effects model, which provides a more conservative estimate of effect and can be used where there is moderate heterogeneity.

With additional included studies, will measure median effect sizes across groups as originally described by Grimshaw 2004, a method which has been used by several subsequent authors (Jamtvedt 2006, Shojania 2004, Shojania 2009, Steinman 2006, Walsh 2006).This method is considered to help measure the median effect of each outcome within a study and subsequently measure the overall single effect size for that study. It is from these single effect sizes for each study that the median effect size and interquartile range across all studies can be calculated. This type of analysis is still subject to limitations, e.g. studies would be assumed to have equal weight. However Grimshaw 2004 argued the process of using median as opposed to the mean results means the summary estimate is less likely to be resulting from a few out-lying results.

We will synthesise data through specific analysis of outcome measures previously described. Where possible, we plan to present separately results of studies comparing:

- the intervention to no intervention (e.g. telephone training programmes alone);
- the intervention to other forms of intervention (e.g. telephone consulting training versus face-to-face consulting training).

Subgroup analysis and investigation of heterogeneity

We did not perform subgroup analysis. When future studies are included and where data are sufficient and it is appropriate in the context of the study, we plan to conduct subgroup analysis according to several factors (type of participants, patient characteristics, location of the study, year of publication, type of intervention, disease specific training interventions and development of protocols). This will allow the examination of the effect of certain studies on the pooled effects of the intervention.

Sensitivity analysis

We did not perform sensitivity analysis. When future studies are included, we planned to remove studies from the analysis deemed to be at high risk of bias after examination of individual study characteristics, to examine the effect on the pooled effects of the intervention.

2.6 Efficacy and safety of a CDSS to support telephone triage

Study Design and research hypotheses

We design a superiority parallel group randomized controlled trial. Telephone calls were randomized to be managed as usual (without CDSS availability) or by the availability of a CDSS. This kind of randomisation could be affected theoretically by a risk of contamination because doctors could learn from the CDSS the question to be asked and use them even when the CDSS is not available. Three alternatives were taken into account: doctor randomisation, periods randomisation, centres randomisation but:

- the first solution was judged to be at risk of contamination too (because doctors speak to each other sharing the same work place and can even share the access to the CDSS if they gamble) and it would be not acceptable by doctors in term of consequent financial incentives disparity (and not allowed by regional law);
- the second solution would have increased the difficulties of statistical analysis and interpretation of the data collected.
- the third solution would not have been achievable because the calls would have been randomized to be answered by a centre that could not be the centre expected to provide a subsequent face-to-face evaluation (with particular regard to home visits) and that would have been a significant complication for the service.

Considering the wide variability of clinical cases presented by callers (the first reason for encounter accounts for 14,5% of all the reasons, the second for 6,5% and the following three for 5% each), the low probability to be referred of the majority of clinical cases and the low mean number of cases every single doctor was be required to assess by CDSS (on average about 20-30 per month), we assumed the contamination risk deriving from randomisation at single call level is not relevant.

The setting

Seven primary care out of hours (OOH) service centres were led by 58 doctors (20 concurrently in shift during weekend day-time and 15 concurrently in shift during night-time – 7 doctors concurrently at the phone at any time – 1 per centre) providing service to about 470.000 inhabitants

All patients ages, all clinical problems, first and third party callers (ie calling on behalf of someone else) were included.

The calls related to a problem that did not require clinical assessment (e.g. seeking advice about medications or where a pharmacy is etc) were excluded.

All the centres and doctors were eligible to use CDSS when randomly offered. All the telephone lasting more than 90 seconds were considered eligible to the study, assuming that this duration is the minimum duration for medical telephone clinical assessment.

Doctors answering to the phone received financial incentives to use the support at least in a basic way, answering to 100% of critical questions (the CDSS has critical, urgent and ordinary questions for every questions set) in the 75% of the times the CDSS was offered. So up to 25% of the times the CDSS will be offered, the triage data could be lost.

The intervention, Odyssey Teleassess

Odyssey TeleAssess is a clinical decision support software using more than 460 clinical scenarios to prompt health professionals to ask all relevant questions during the telephone triage procedure and suggesting an urgency level depending from the answers given. It has been used for over 20 million calls without clinical incident and its development is accredited in the UK by the National Institute of Health and Clinical Excellence (NICE).

Doctors were exposed to a dedicated training program in order to learn more about how to use Odyssey TeleAssess as a CDSS.

Odyssey TeleAssess was integrated into the usual software recording calls data (Em.Ma.Web by Beta80 Group) and its availability was randomly offered to doctors during the calls. The patient personal data were stored into Em.Ma. Web database (ULSS 20 Service) and just data about call progressive code (as ID), patient's age and patient's gender were passed from Em.Ma Web to Odyssey TeleAssess in order to let it provide the required clinical assessment.

The whole study dataset was produced integrating data provided by EmMa and by Odyssey.

The outcomes

Primary Outcome

1) N° calls from not frail patients closed as "referred to hospital emergency room"/ total number of calls from not frail patients – EFFICIENCY OUTCOME

Secondary outcomes

2) N° calls from frail patients closed as referred to hospital emergency room/ total number of calls from frail patients – EFFICIENCY OUTCOME

3) Mean (SD) proportion of obligatory questions asked by doctors at the phone among the obligatory questions previously defined as obligatory - SAFETY OUTCOME assessed by ISP[2]

4) Mean (SD) proportion of obligatory questions asked by doctors at the phone among all the questions asked by them- - EFFICIENCY OUTCOME assessed by ISP

5) Mean (SD) proportion of appropriate care advice given by doctors at the phone at the end of telephone consultation versus advice previously defined as "required advice" - SAFETY OUTCOME assessed by ISP

6) Mean call duration – EFFICIENCY OUTCOME

[2] In order to assess SAFETY outcomes 6 standardized clinical cases commonly presented in OOH clinical setting and taken from the literature (Derkx 2008) were used with their questions sets as obligatory questions. The evaluators - 6 family medicine residents – called and presented the cases using the Incognito Standardized Patient (ISP) technic. ISP is a recognized method to study medical telephone consultation quality, using standardised clinical cases, administered by sham patients (incognito evaluators).

The study period

The study was planned to start on 1 July 2014 and stop on 28 February 2015 with a run-in phase (for doctor training) form 1 July 2014 to 31 August 2015 and a possible extension up to 31 December 2015.

Sample size, randomization, allocation and blindness

Previous performance data from the service showed the mean hospital referral rate was about 13%. The hypothesis we used to estimate the sample size was that CDSS would have determined a reduction of hospital referrals from 13% to 10%. To detect such a difference, considering a 5% alfa and 95% statistical power, we calculated to need 3003 calls per arm with about 6.000 calls as a whole amount of calls to include in the study.

The 7 OOH centres were divided into 4 major centres and 3 minor centres; the 4 major centres receive a mean of 600 calls on a monthly base while the minor centres receive a mean of 50 calls on a monthly base (with a total of about 2550 calls potentially included in the study per month), so we plan to reach the necessary sample size in a 3 months period. As doctors receive financial incentives to accept the support in the 75% times it is offered, we plan a loss of 25% of data in the CDSS availability arm. Not all the calls needed a clinical evaluation and not all the patients accepted to participate (we had no preliminary data about that). For this reasons we decided to extend the study length to 6 month. In the case this period was not enough to reach the sample size because of unplanned reasons (for example a lot of denied consents from callers), we planned to extent the study for 12 more months (up to 18 months as a maximum period).

Each evaluator performed 50 incognito calls during the data collecting study phase. This size was decided in a pragmatic way.

Randomisation to support availability arm or to support unavailability arm was automatically computer-generated on a single service centre basis (in every centre the computer performed the randomisation; this means the randomisation was not performed on a whole service basis: in this way for every centre about 50% of the calls were allocated to support availability and 50% to support unavailability.

The allocation sequence was computer generated and the offer to the doctors at the phone to use CDSS was made at the moment the call arrives and the doctors have to open the data recording window, so the potential loss of a maximum of 25% selectively in the CDSS arm was not related to clinical severity of the case (assessed later) and to the consequent probability to be referred to hospital care.

There was no need to actively enrol participants because they are the patients normally accessing the OOH service by phone.

In this study the phone callers were blind to assignment while for doctors directly providing care it was impossible to be blind. There was no need of external or subsequent outcome assessment or confirmation because doctors were obliged to record the call final result in terms of referred to hospital or not by the Italian law and by the software structure (there was a block on this software item). Data analysis was planned to be performed by a statistician blind to assignment to intervention.

The evaluators were blind with regards to study design so they did not know if doctors used CDSS or not. There was no need for external or subsequent outcome assessment or confirmation because evaluators recorded the call outcomes in a predefined checklist.

Data analysis

For binary outcomes differences between proportions and their relative 95% confidence intervals were planned to be estimated.

For continuous outcome differences between mean and medians and their relative 95% confidence intervals were planned to be estimated.

We planned to use non parametric statistics for variables not-normally distributed.

We planned to perform a subgroup analysis on the basis of the following variables:

- patient age (<14; 15-64; >65; >80)
- patient level of frailty
- centre
- doctor age (<36, 36-45, 46-55, >55)
- time of day/day of week

No interim analysis was performed but the study was planned to be stopped as soon as the planned sample size was reached and a check of sample size reached was planned to be performed every 2 months.

The local Ethics Committee approved the protocol about two months before the study start-up.

Duality of interest

The principal nvestigator declared no financial or other competing interest.

Doctors taking part to the study received financial incentives as part of a contract, regulating several parts of their activity, their unions signed with the LHT.

2.7 Barriers and facilitators to uptake CDSS

We conducted a qualitative, cross-sectional study based on semi-structured interviews to examine individual and contextual barriers and facilitators to CDSS uptake. The interview was designed and analyzed accordingly to the constructivist Grounded Theory (GT) approach, the inductive development of theory from data (Glazer 2004). Instead of initiating the study with a hypothesis or research question, the GT method begins with empirical observations on the field of interest (Glazer 2004). Data on the main features, conditions, outcomes, and contextual factors of the object of study is used to ground and systematically generate a theory surrounding the social or socio-psychological process (Charmaz 2006, Glazer 2004). By adopting this approach, we aim to capture the complex and multi-dimensional processes and relationships involved in the use and adoption of CDSS across different stages of implementation.

One of the main elements of the GT method is the collection of data before its analysis. The former is driven by theoretical sampling, the sequential selection of individuals in a study

sample according to the state of theory generation (Charmaz 2006). Following the criteria of theoretical saturation (Morse 2009), we plan to progressively recruit participants until no new information is raised from the interviews.

Setting and participants

We adopted a purposive sampling strategy and select the first participants using a maximum variability logic (Patton 2002). This strategy was chosen to explore contexts with different levels of familiarity with CDSS as well as participants with diverse organizational roles.

To determine the impact of particular clinical setting characteristics on CDSS uptake, we selected three specialty research hospitals located in northern Italy based on the following criteria: the hospital reported the use of evidence during practice, an EHR, and a level of CDSS implementation and familiarity. Specifically, setting 'A' is an oncology hospital that abandoned paper-based clinical documents and fully adopted an EHR beginning in 2008. The hospital's EHR is linked to a variety of CDSS, including evidence-based messages on treatment and diagnosis using care management algorithms. All participants from setting 'A' were considered current users of CDSS and compliant with EBM. Setting 'B' is an orthopaedic research hospital that has been using an HER since 2011, but does not have a CDSS. This hospital's EHR is not sufficiently developed to link to a CDSS or to be utilized by health professionals who have adopted standard international codes, a prerequisite to the activation of CDSS guide messages. Health professionals in setting 'B' may or may not be compliant with EBM depending on their own capacity and willingness. Setting 'C' is an orthopedic research hospital that does not have an EHR or a CDSS. This setting was considered an environment reluctant to innovation and the use of evidence in practice.

In order to determine the impact of particular professional or organizational roles and characteristics on CDSS perception, we interviewed frontline physicians, nurses, information technology staff, and members of the hospital board of directors. We used purposive sampling to capture the perspectives of a diverse and representative sample of professionals.

The demographics or seniority of participants were not be considered in the selection criteria; nonetheless, this information was collected during the interview. Consistent with the theoretical sampling strategy, we added to our sample throughout the data collection process based on the provisional results from the analyses. For example, we increased the number of participants in one cluster of stakeholders or include health professionals that were not expected to participate in the sample if their positions and experiences require further consideration or elaboration.

Data collection

Data were collected through interviews in hospital consultation rooms. The interview was address the following topics: a) participants' beliefs and experiences with information technology, in general (e.g., the use of personal computers, tablets, and smartphones to obtain information relevant to their clinical practice); b) beliefs and experiences with CDSSs, specifically; c) willingness to adopt EBM and clinical guidelines in their routine practice; and d) perceptions regarding the potential of CDSSs to integrate evidence and guidelines in clinical practice. We asked participants to discuss both their own experiences and those of their

colleagues in their workplace. We followed the GT principle of progressively refining the interview framework according to participants' answers around the object of interest. In other words, interviews did not follow a prescribed structure; rather, questions were developed continuously for each interview. Participants were contacted by the three Unit Coordinators (LM, MM, ON) and invited to participate in the study. Each interview was conducted by two individuals: a trained investigator who conducted the interview, and a medical doctor with competence in the participating hospital's field of specialty who supported the investigator in clinical topics. The expected duration of the interview is 30 to 90 minutes. The interviews were taped and transcribed verbatim.

Data analysis

Investigators analyzed all interview transcripts according to the procedure outlined in GT content analysis, which involves three sequential phases of coding (Morse 2009, Charmaz 2006).

In the first step, open coding, investigators identified and label preliminary concepts found in the data (e.g., 'CDSS is less reliable than colleagues'). Investigators analyzed the interview transcripts line-by-line to detect 'in vivo' codes that directly use the participant's wording. In axial coding, the second analytical step, investigators reassembled particular sets of data based on central concepts that emerge from the ongoing analysis; in other words, codes were progressively aggregated into broader categories (e.g., 'resilience of paper-based culture' or 'power and hierarchy issues'). This step involved the recurrent identification and comparison of themes both within and across sub-categories and broader categories. In selective coding, the final step of data analysis, investigators further defined, developed and refined discrete concepts and categories. The core categories, pivotal concepts encapsulating the whole phenomenon under investigation were selected and systematically related to the other categories. The combined categories and their interrelationships ultimately formed a larger storyline surrounding the process of CDSS uptake (Morse 2009, Charmaz 2006).

The coding process was conducted by three investigators. The NVivo software (version 10) (QSR International Pty) was used to support the analysis.

Ethics and funding

The study was approved by the Research Ethics Committees of IRCCS Istituto Ortopedico Rizzoli (approved January 31, 2014; file number 0003938/2014), IRCCS Ospedale San Raffaele (approved March 6, 2014), and IRCCS Istituto per la Cura dei Tumori della Romagna (approved February 21, 2014; file number 1155/2014).

This work was supported by the Italian Ministry of Health (GR-2009-1606736) and by Regione Lombardia (D.R.G. IX/4340 26/10/2012). Funding sources had no role in the writing of this manuscript or the decision to submit it for publication.

Before beginning the interview, participants were given an informed consent form as well as information outlining the purpose of the study and participant rights.

Participants were notified that their involvement was voluntary and could be withdrawn at any time and that confidentiality was protected through the anonymization of all collected data.

RESULTS

3.1 Efficacy and safety of INFs-beta vs GA in RRMS

Results of the search

We identified overall 636 reports through the search strategy (MEDLINE 356, Embase 225, CINAHL 14, CENTRAL 14, clinical trials registries 11, CDR database 7, other databases 2, additional articles 7). We excluded 592 articles on the basis of abstracts considered not pertinent. (See the corresponding references section).

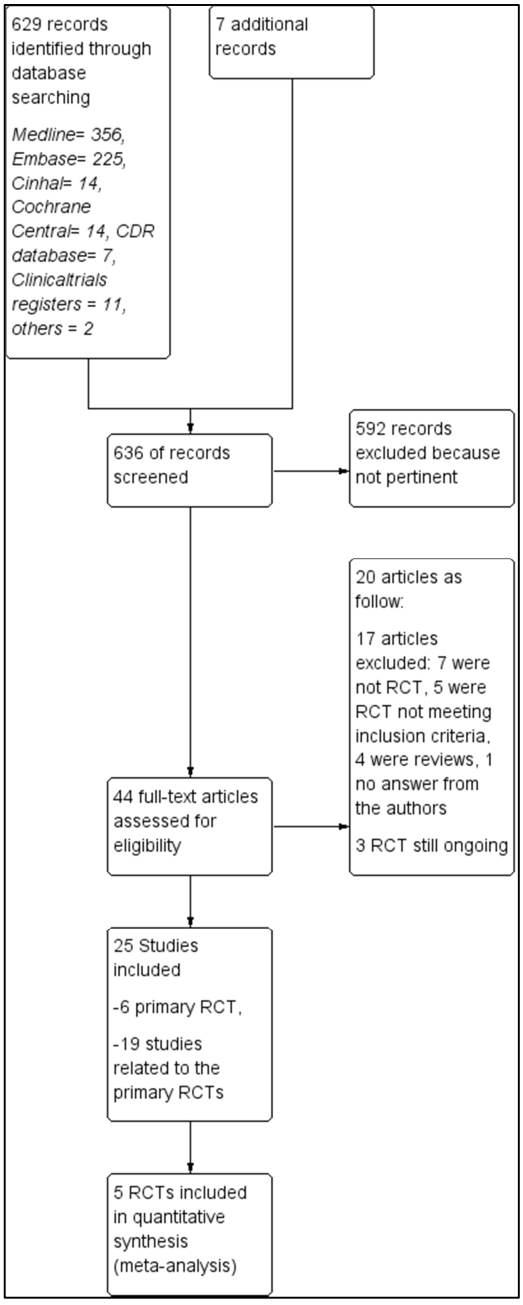

Figure 1. Study flow diagram INFs vs GA in RRMS

Overall, a total of 44 articles were provisionally selected as potentially fulfilling the inclusion criteria (Figure 1).

Seventeen studies were excluded: seven because they were not RCTs; five RCTs because the design included no active comparator treatment, one because it was a comparative trial evaluating two dosages of GA and two because evaluated drugs different from those considered in this review; four publications were review articles. Finally one study was excluded because authors had never answered to our request for additional information.

Three trials are still ongoing. One ongoing trial (NCT00176592) is a phase IV RCT, with head-to-head comparisons of IFN-beta 1b (250 mg of Betaseron) and GA (Copaxone) for the treatment of participants with CIS and RR forms of MS by using acute changes on MRI as the primary outcome. The second one (EUCTR2012-003735-32-GR) is a phase IIb/III trial that compares masitinib with interferon beta-1a, interferon beta-1b, peginterferon beta-1a or glatiramer acetate in patients with RRMS who do not respond to these first line treatments. The third (NCT01623596) is a 12 month study where patients with RRMS are randomized 1:1 to fingolimod or approved disease modifying therapy. Patients were treatment naive or have only been treated with one class of DMT (Interferon beta preparation or glatiramer acetate). Primary objective was to evaluate efficacy of fingolimod by assessing patients retention on treatment.

Twenty-five studies were considered for inclusion: five published RCTs (Cadavid 2009; Calabrese 2012; Lublin 2013; Mikol 2008; O'Connor 2009a), and 1 unpublished study (NCT01058005) met the selection criteria. Nineteen articles were related to primary studies, excluding the aforementioned ongoing trial (NCT00176592), which was related to the Cadavid 2009 study.

1. Three articles were pertinent to the Cadavid 2009 trial. Investigators analysed the impact of therapies on MRI measures as follows: one was a post hoc analysis of MRI/clinical activity (Cadavid 2011), another was an analysis of the development of focal lesions suggestive of brain injury (Cadavid 2009b) and the last described changes in brain volume (Cheriyan 2012).
2. Six articles were pertinent to the Lublin 2013 trial. One study reported the protocol (Lindsey 2012) and five articles were related to meeting reports of the same trial (Lublin 2012; Lublin 2013b; Lublin 2013; Wolinski 2012; Wolinsky 2013).
3. Two articles were pertinent to the Mikol 2008 trial. One was a post hoc analysis of tolerability (Coyle 2010), and the other commented on the same trial (Sorensen 2008).
4. Eight articles were pertinent to the O'Connor 2009a trial. One study was a meeting report of preliminary results (O'Connor 2008), another described the immunological effects of treatments (Goodin 2012), another was a post hoc analysis evaluating MRI measures of degeneration among participants included in the BEYOND trial (Filippi 2011), three were Errata reports (O'Connor 2009b; O'Connor 2011; O'Connor 2012) and one was a post hoc analysis (Lampi 2013). Moreover, we analysed additional data from the BEYOND trial as provided by Bayer (Pleimes 2013).

Included studies

Six RCTs met our predefined selection criteria: two studies compared the effects of GA versus IFN-beta 1b (Cadavid 2009a; O'Connor 2009a), and four compared GA versus IFN-beta 1a

(Calabrese 2012; Lublin 2013; Mikol 2008; NCT01058005), with two comparing GA versus IFN-beta 1a 44 mcg SC (Mikol 2008; NCT01058005), one GA versus IFN-beta 1a 30 mcg IM (Lublin 2013) and one GA versus both IFN-beta 1a 44 mcg SC and IFN-beta 1a 30 mcg IM (Calabrese 2012) and one comparing GA versus IFN-beta 1a 44 mcg SC and natalizumab (NCT01058005):

1. Cadavid 2009 (BECOME) evaluated the efficacy of IFN-beta 1b and GA in 79 participants with RRMS or CIS (36 treated with IFN and 39 with GA). The primary outcome was MRI measures of activity (combined active lesion counts).
2. Calabrese 2012 evaluated the efficacy of IFN-beta 1a (44 mcg SC three times weekly), IM IFN-beta 1a (30 mcg weekly) or GA in 165 participants with RRMS (55 participants in each group). The primary outcome was the development of new cortical lesions and cortical atrophy progression among participants with RRMS.
3. Lublin 2013 (CombiRx) evaluated the efficacy of combined use of interferon-beta 1a 30 mcg IM weekly and GA 20 mg daily versus each single agent with matching placebo in 1008 participants with RRMS (250 treated with IFN, 259 with GA and 499 with IFN + GA). The primary outcome was annualised relapse rate (ARR).
4. Mikol 2008 (REGARD) evaluated the efficacy of SC IFN-beta 1a (44 mcg three times weekly) versus GA in 764 participants with RRMS (386 in IFN group and 378 in GA group). The primary outcome was time to first relapse.
5. O'Connor 2009a (BEYOND) evaluated the efficacy of IFN-beta 1b at two different doses (250 mcg and 500 mcg every other day) and of GA in 2244 participants with RRMS (897 treated with IFN 250 mcg, 899 with IFN 500 mcg and 448 with GA). The primary outcome was risk of relapse. Missing data were provided by Bayer (Pleimes 2013).
6. NCT01058005 (SURPASS) evaluated the safety of natalizumab (300 mg intravenous injection every 4 weeks), IFN-beta 1a (44 mcg SC injection 3 times weekly), or GA (20 mg SC injection once daily) in 84 participants. The primary outcome was the Incidence of treatment-emergent Serious Adverse Events (SAEs).

Four of the five trials (Cadavid 2009a; Lindsey 2012; Mikol 2008; O'Connor 2009a) defined relapse as new or recurrent neurological abnormalities associated with an increase in Scripps Neurological Rating Scale (SNRS) score (Cadavid 2009a) or in EDSS score (Cadavid 2009a; Lublin 2013; Mikol 2008; O'Connor 2009a), lasted at least 24 hours (Cadavid 2009a; Lublin 2013; O'Connor 2009a) or at least 48 hours (Mikol 2008) and occurred without fever or infection. NCT01058005 (SURPASS) underwent an early termination due to significantly slower than expected enrolment: all clinical efficacy and magnetic resonance imaging (MRI) procedures were removed from the protocol and safety assessments were to be managed through standard of care activities.The results were reported on www.clinicaltrials.gov. Based on these data this study was selected because fitting with our predefined inclusion criteria, although it was excluded in the previous version of this review. Actually no published data have been retrieved and this study did not contribute to any analyses for the lack of available data.

Cadavid 2009a and Lublin 2013 specified that the increase in EDSS score identified a relapse as follows.

1. Increase in total EDSS of 0.5 point.

2. Increase in Kurtzke Functional System Score (FSS) of 2.0 points.

3. Increase in two or more Kurtzke FSS scores of 1.0 point, or decrease in SNRS score of 7.0 points (Cadavid 2009a).

Disability progression was analysed by three RCTs and was defined as:

1. six months' sustained increase in EDSS > ≥1.0 point from baseline or at least 0.2 standard deviations (SDs) on

baseline Multiple Sclerosis Functional Composite (MSFC) score (post hoc analysis) (Cadavid 2011);

2. six months' sustained increase in EDSS > ≥ 1.0 point (0.5 for baseline EDSS ≥ 5.5) (secondary outcome) (Lublin 2013); or

3. three months' sustained increase in EDSS > ≥ 1.0 point (secondary outcome) (O'Connor 2009a).

All RCTs included participants with RRMS with low disability and active disease (frequency of relapse ≥ 1/y). No significant differences were found between studied populations in terms of mean age (range 34.8 to 39.0 years), mean EDSS (1.9 to 2.35 points) and mean relapse frequency (0.97 to 1.9). Mean disease duration was different, ranging from 0.9 (Cadavid 2009a) to 6.55 (Mikol 2008).

Only three studies reported the numbers of participants with enhancing baseline lesions at MRI (Cadavid 2009; Lublin 2013; Mikol 2008).

The overall number of participants included in the five RCTs was 4256. However, we decided to exclude from the analysis high dosages of IFN-beta 1b (500 mcg to 899 patients), as used in the O'Connor 2009 study, and combined IFN + GA (509 participants), as used in Lublin 2013 study arms, because these schedules are not used in clinical practice. The overall population considered in our analysis was 2858 (1679 participants treated with IFN and 1179 with GA). The drugs analysed in comparison with GA were IFN-beta 1b (two trials, 933 participants), SC IFN-beta 1a 44 mcg (two trials, 441 participants) and IM IFN-beta 1a 30 mcg (two trials, 305 participants).

Duration of treatment and follow-up was three years for the Lublin 2013 study and two years for the other four RCTs.

Excluded studies

We excluded 17 articles: 7 were not RCT, 5 were RCT non meeting inclusion criteria, 4 were review, 1 no answer from the authors.

Risk of bias in included studies

The risk of bias was variable across studies (Figure 2): incomplete outcome data was the main biased dimension (high risk of bias in all studies) because of high levels of dropout and missing data, followed by blinding of participants and investigators and by selective outcome reporting (high risk of bias in three and two studies, respectively).

Allocation (selection bias)

All studies—with the exception of Cadavid 2009 (in which the item was not mentioned)—used computer systems to generate the allocation sequence, but none clearly explained how the sequence was concealed (with the exception of Lublin 2013, in which the sequence was masked by the computer system).

Blinding (performance bias and detection bias)

Participants treated in Mikol 2008, O'Connor 2009 Cadavid 2009 and NCT01058005 were not blinded. In Calabrese 2012, the item is not mentioned and in Lublin 2013, the computer system ensured blindness. Treating physicians were aware of the treatments in Mikol 2008 and O'Connor 2009 and were unaware in Lublin 2013. The other studies did not mention the item. In all studies, outcome assessors were blinded. Two studies were at low risk for detection bias for MRI measures (Cadavid 2009; Calabrese 2012), which were the primary outcomes for these studies, and were unclear for clinical outcomes.

Incomplete outcome data (attrition bias)

Incomplete outcome data was the main biased dimension because of the high level of loss to follow-up. Participants who dropped out accounted for 13% to 30%, with higher values for IFN groups in four RCTs. An ITT analysis was performed in Cadavid 2009, Mikol 2008 and Lublin 2013, but in the other two studies, it was not performed. Reasons for loss to follow-up were clearly reported only by Cadavid 2009 and Mikol 2008.

Selective reporting (reporting bias)

Cadavid 2009 and O'Connor 2009 failed in matching outcomes planned with outcomes reported: O'Connor 2009 reported five tertiary outcomes not planned, and Cadavid 2009 reported one outcome more than those planned.

Other potential sources of bias

All studies were sponsored by the drug industry (Bayer sponsored Cadavid 2009 and O'Connor 2009; Merck Serono sponsored Mikol 2008 and Calabrese 2012), with the exception of Lublin 2013, which was funded by the National Institutes of Health and the National Institute of Neurological Disorders and Stroke.

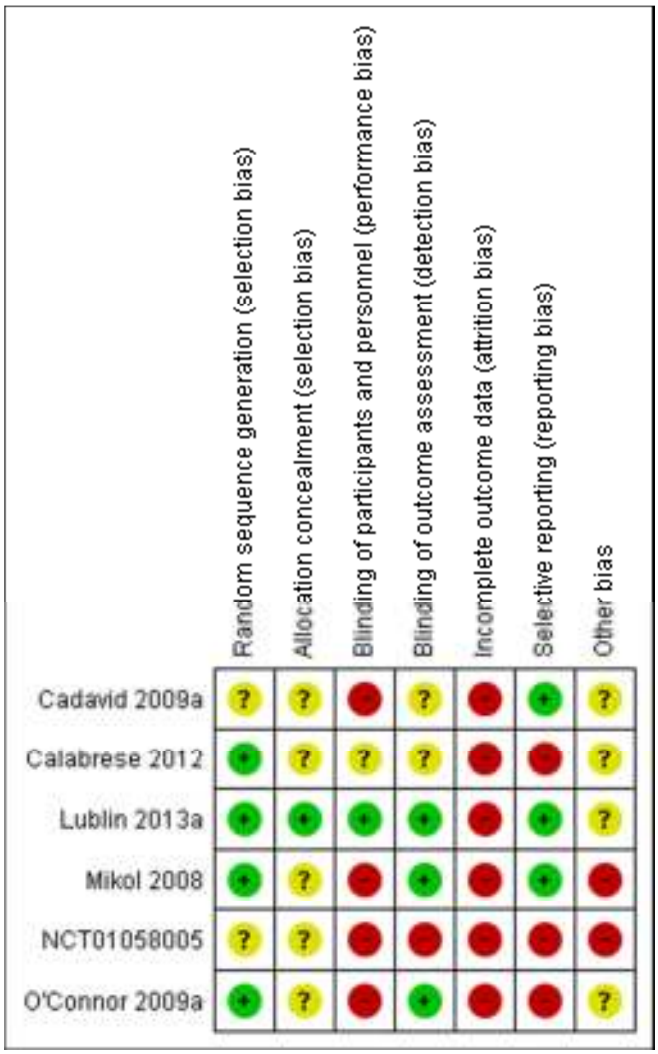

Figure 2. Judgements about each risk of bias item for each included study

Effects of interventions

Main findings are described in Summary of Findings Table (Table 2).

Table 2. INFs vs GA for RRMS Summary of Findings Table

Interferons compared with glatiramer acetate for participants with relapsing-remitting multiple sclerosis						
Patient or population: patients with relapsing-remitting multiple sclerosis **Settings:** secondary care **Intervention:** interferons **Comparison:** glatiramer acetate						
Outcomes	**Illustrative comparative risks* (95% CI)**		**Relative effect (95% CI)**	**No. of participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
	Assumed risk (control)	**Corresponding risk (intervention)**				
	Glatiramer acetate	**Interferons**				
Number of participants with relapse Risk ratio (M-H, random, 95% CI) Follow-up: 24 months	**Study population**		**RR 1.04** (0.87 to 1.24)	2184 (3 studies)	⊕⊕⊕⊖ **moderate**[a]	Detection bias risk for clinical outcomes was judged as high for 1 study and low for the other 2 RCTs
	36 per 100	**38 per 100** (31 to 45)				
	Moderate					
	35 per 100	**36 per 100** (30 to 43)				
Number of participants with confirmed progression Risk ratio (M-H, random, 95% CI) Follow-up: 24 months	**Study population**		**RR 1.11** (0.91 to 1.35)	2169 (3 studies)	⊕⊕⊕⊖ **moderate**[a]	Detection bias risk for clinical outcomes was judged as high for 1 study and low for the other 2 RCTs
	15 per 100	**16 per 100** (13 to 20)				
	Moderate					
	15 per 100	**17 per 100** (14 to 21)				
Number of participants who dropped out for AEs Risk ratio (M-H, random, 95% CI) Follow-up: 24 months	**Study population**		**RR 0.95** (0.64 to 1.4)	2685 (4 studies)	⊕⊕⊖⊖ **low**[a,b]	
	4 per 100	**4 per 100** (3 to 6)				
	Moderate					
	5 per 100	**5 per 100** (3 to 7)				
Mean number of		**0.15 lower in**		1790 (3 studies)	⊕⊕⊖⊖ **low**[b,c]	Detection bias risk for MRI outcomes was judged as low for all

active T2 lesions Mean difference (IV, random, 95% CI) Follow-up: 24 months		**IFN versus GA groups** (0.68 lower to 0.39 higher)				studies
Mean number of new enhancing lesions Mean difference (IV, random, 95% CI) Follow-up: 24 months		**0.14 lower in IFN versus GA groups** (0.3 lower to 0.02 higher)		1734 (3 studies)	⊕⊕⊕⊖ **moderate**[d]	Detection bias risk for MRI outcomes was judged as low for all studies
Mean change in total T2-hyperintense lesion load Mean difference (IV, random, 95% CI) Follow-up: 24 months		**0.58 lower in IFN versus GA groups** (0.99 to 0.18 lower)		1608 (2 studies)	⊕⊕⊕⊖ **moderate**[d]	Detection bias risk for MRI outcomes was judged as low for both studies
Mean change in total T1-hypointense lesion load Follow-up: 24 months		**-0.20 lower in IFN versus GA groups** (-0.33 to -0.07)		1602 (2 studies)	⊕⊕⊕⊖ **moderate**[d]	Detection bias risk for MRI outcomes was judged as low for both studies

*The basis for the **assumed risk** (e.g. median control group risk (GA) across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group (IFNs) and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **RR:** Risk ratio.

GRADE Working Group grades of evidence.
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

Primary outcomes

CLINICAL EFFICACY OUTCOMES

1. Number of participants who experienced at least one relapse at 12 - 24 months and at the end of follow-up

This outcome was assessed by three trials at 24 months (Cadavid 2009; Mikol 2008; O'Connor 2009) (2184 participants; 76%) and by one trial at 36 months (Lublin 2013) (509 participants; 18%) and was not available from the Calabrese 2012 trial. From these data, we found no significant differences in effect at 24 months (RR 1.04, 95% CI 0.87 to 1.24) and at 36 months (RR 1.27, 95% CI 0.92 to 1.75). No significant heterogeneity was found among studies (Figure 3. Analysis 1.1). The results were unchanged when missing data were explored at sensitivity analysis by a likely scenario (Figure 4. Analysis 1.2).

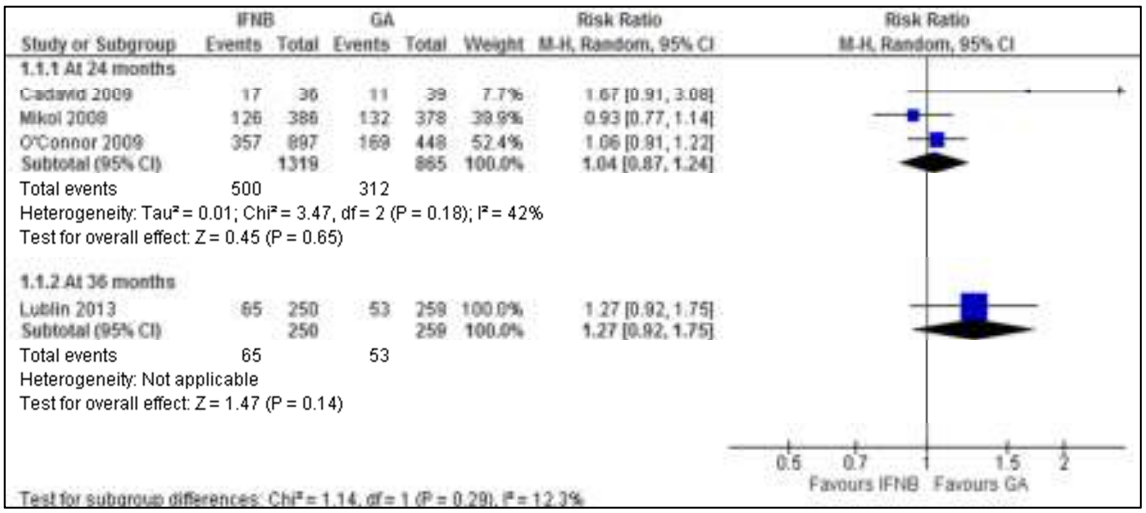

Figure 3. Analysis 1.1. Number of participants who experienced at least one relapse

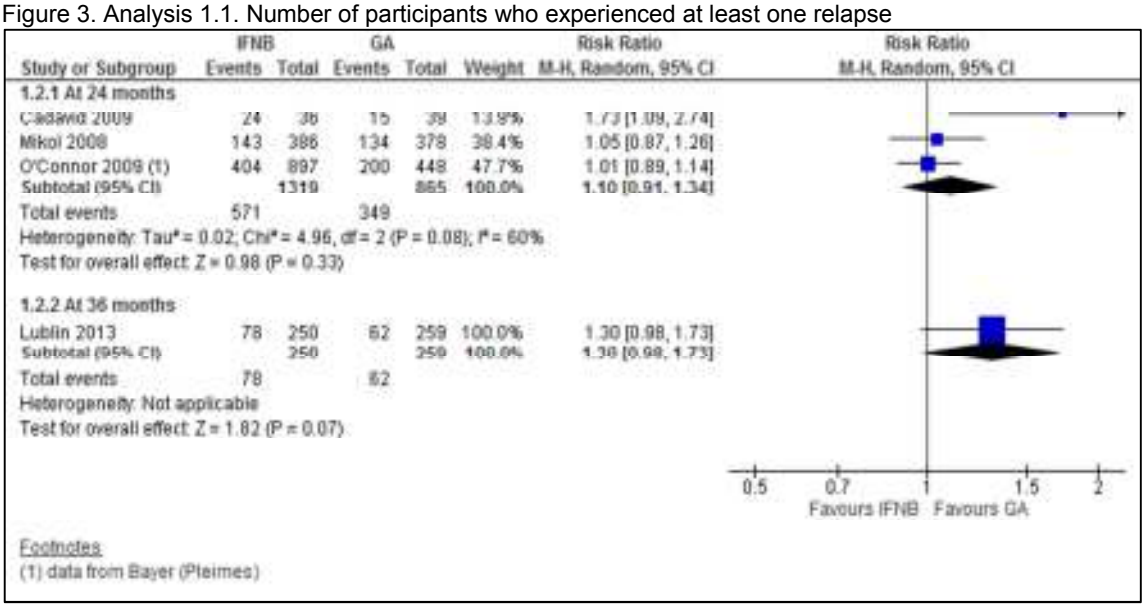

Figure 4. Analysis 1.2. Number of participants who experienced at least one relapse (likely scenario)

2. Number of participants with confirmed worsening at 12 - 24 months and at the end of follow-up

This outcome was assessed by three trials at 24 months (2169 participants; 76%) and by one trial at 36 months. Three trials adopted the predefined criteria of worsening (confirmed during

two subsequent neurological examinations separated by at least six months) at 24 months (Cadavid 2009; Mikol 2008) for 839 participants (29%) and at 36 months (Lublin 2013) for 758 participants (26%). One study used less stringent criteria (i.e. an increase in EDSS sustained for three months (O'Connor 2009) for 1345 participants (47%)).

No differences were found when confirmed progression was analysed at 24 months (RR 1.11, 95% CI 0.91 to 1.35) or at 36 months (RR 0.87, 95% CI 0.63 to 1.20) (Figure 5. Analysis 2.1).

Results were unchanged when missing data were explored by a likely scenario (Figure 6. Analysis 2.2)

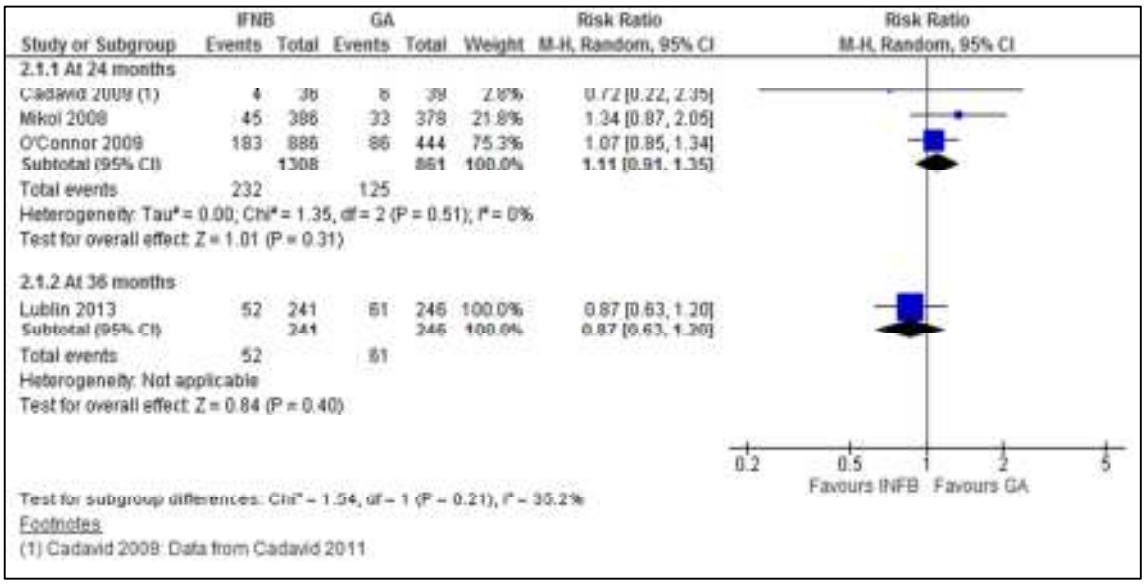

Figure 5. Analysis 2.1. Number of participants with confirmed worsening

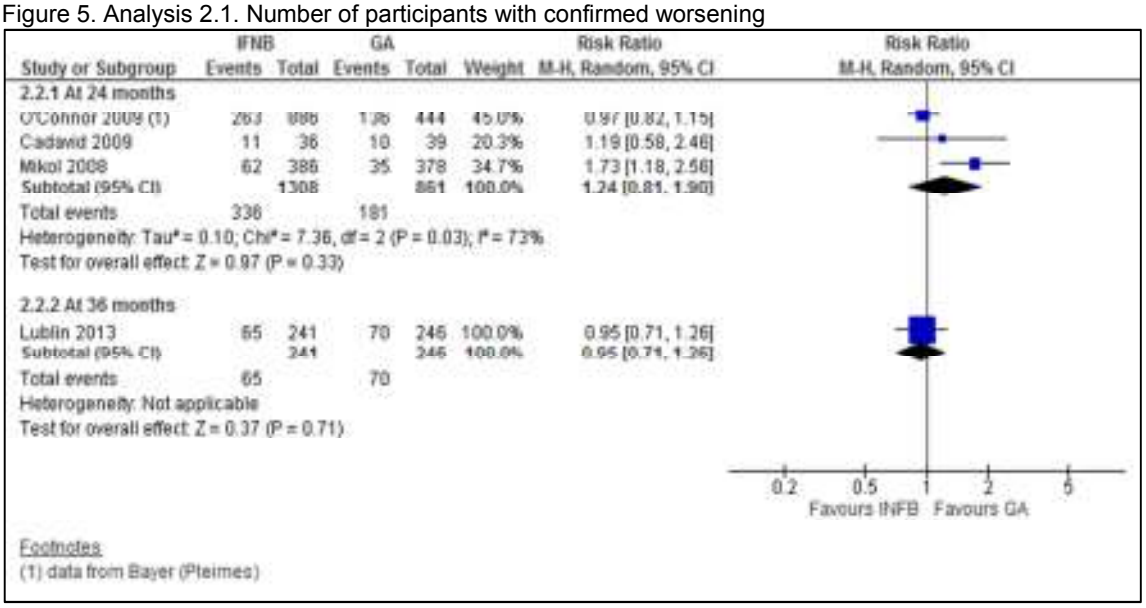

Figure 6. Analysis 2.2. Number of participants with confirmed worsening (likely scenario)

CLINICAL SAFETY OUTCOMES

3. Number of participants who withdrew from or dropped out of the study because of adverse events

The number of participants who withdrew from or dropped out of the study because of adverse events was available for four studies (2685 participants; 93%) and was not reported in the

Calabrese 2012 study. No differences was found between the two treatment groups (RR 0.95, 95% CI 0.64 to 1.39). No heterogeneity was found (Figure 7. Analysis 3.1).

Similar results were found when SAEs were considered (RR 0.99, 95% CI 0.63 to 1.56) (Lublin 2013; Mikol 2008; O'Connor 2009). Seven deaths were reported: five in the IFN group (three in the 500-mcg arm) and two in the GA group: Among participants treated with IFN, one died (suicide) about three months after taking the last dose of study drug (IFN-beta 1a) (Mikol 2008), one died as the result of pulmonary embolism (Lublin 2013) and three died for unexplained reasons (O'Connor 2009). No reason was specified for one death reported in the GA group (O'Connor 2009), and one participant died of a large cell lymphoma of the CNS (Lublin 2013) (Figure 8. Analysis 3.2).

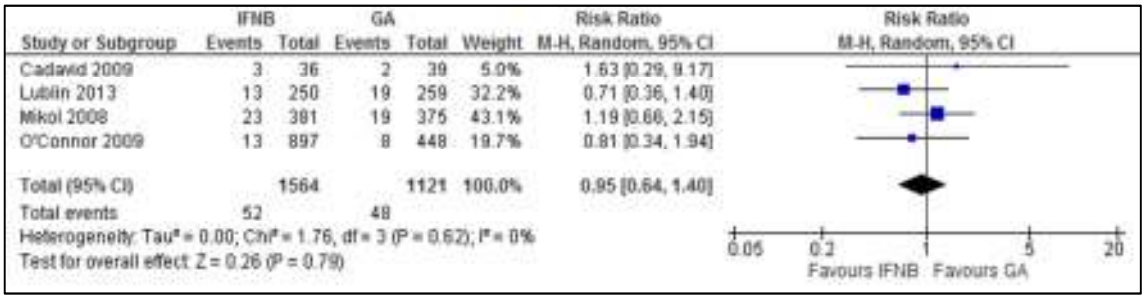

Figure 7. Analysis 3.1. Number of participants who withdrew from or dropped out

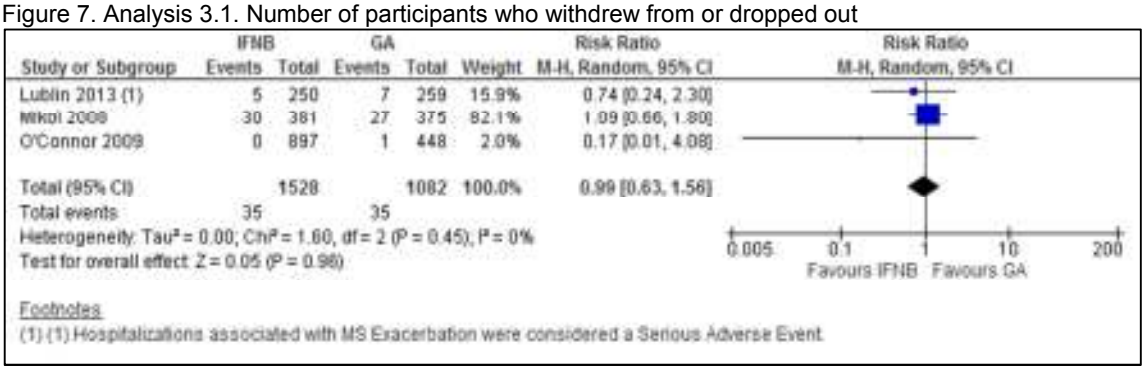

Figure 8. Analysis 3.2. Number of participants who withdrew from or dropped out

Lublin 2013 is the only trial that provided details on numbers of SAEs by organ system as experienced by participants. In this study, SAEs occurred more often in the IFN group than in the GA group: nervous system disorders, including relapses: 4.4% versus 1.9%; neoplasms benign and malignant: 2% versus 0.4%; surgical and medical procedures: 2% versus 1.9%; infections and infestations: 1.6% versus 0.4%; gastrointestinal disorders: 1.2% versus 0.8% and cardiac disorders: 1.2% versus 0.8%.

By contrast, the proportions of participants treated with GA were higher for experiencing hepatobiliary disorders (0.8% vs 0.4%) and musculoskeletal, connective tissue and bone disorders (0.8% vs 0.4%).

Surpass study (NCT01058005) reported 2 SAE, one Meningitis Herpes and one Cerebral Venous Thrombosis in 2/22 participants treated with IFN1a.

Analysis of the number needed to treat or to harm was not provided because reliable data were insufficient.

Secondary outcomes

CLINICAL OUTCOMES

4. Frequency of relapse

The frequency of relapse was analysed in terms of log rate ratio at 24 months for four studies (Cadavid 2009; Calabrese 2012; Mikol 2008; O'Connor 2009). The rate ratio (1.06, 95% CI 0.95 to 1.18) showed no difference between the two groups. At 36 months, data were provided by one study (Lublin 2013); the rate ratio (1.40, 95% CI 1.13 to 1.74) was significantly higher in the IFN group (P value 0.002), favouring GA (Figure 9. Analysis 4.1).

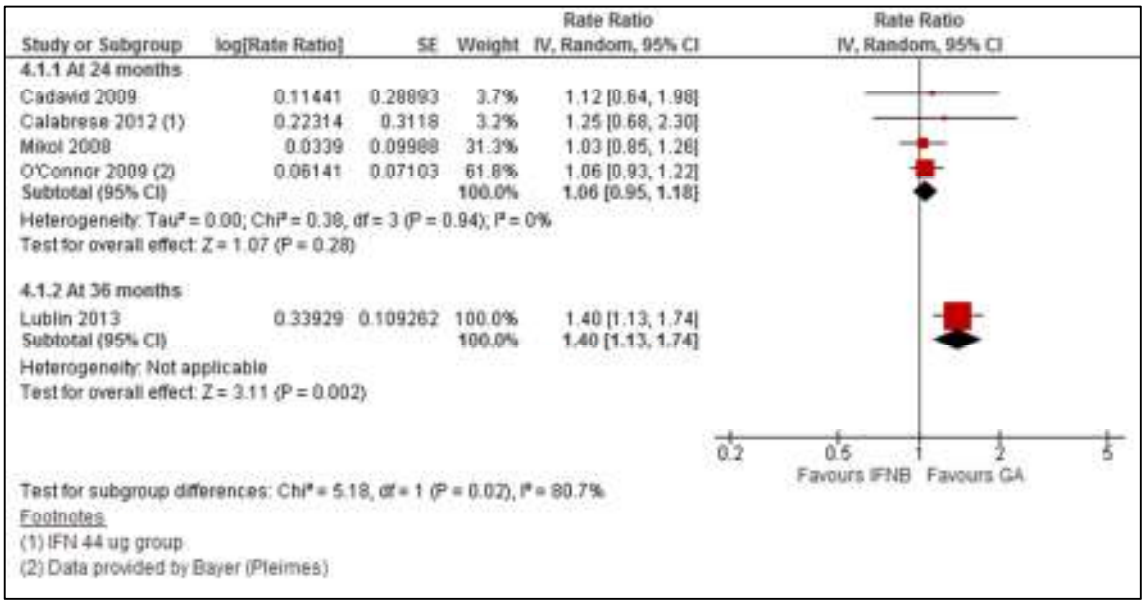

Figure 9. Analysis 4.1-2. Frequency of relapse

5. Time to first relapse

No differences were found (HR 1.01, 95% CI 0.87 to 1.16) without heterogeneity among studies (Figure 10. Analysis 5.1).

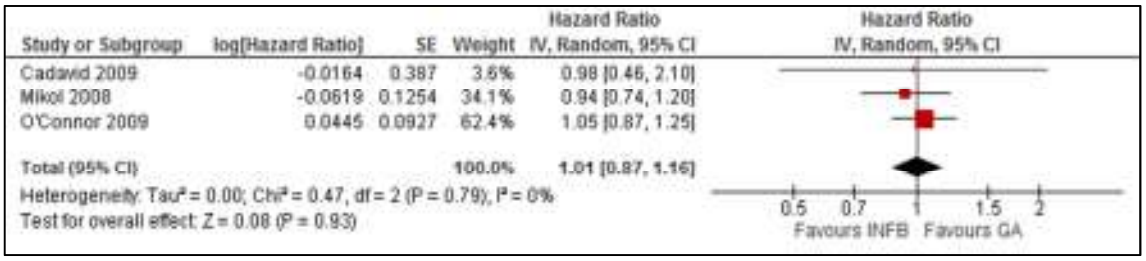

Figure 10. Analysis 5.1. Time to first relapse

6. Percentage of participants free of disease activity: no relapses, no changes in EDSS and no MRI changes

No analysis was possible because data were insufficient. Data were reported by only one study (Lublin 2013): a preplanned assessment of percentages of participants with no clinical or MRI activity was performed; no significant differences were found between IFNs and GA (21.2% and 19.4%).

7. Participants treated with steroids for relapse of MS

This outcome was available for two studies (Cadavid 2009; O'Connor 2009) (1420 participants; 50%). Results did not show a statistically significant difference between the two therapies (RR 1.30, 95% CI 0.76 to 2.24) (Figure 11. Analysis 7.1). Significant heterogeneity was found between studies (I^2 = 63%).

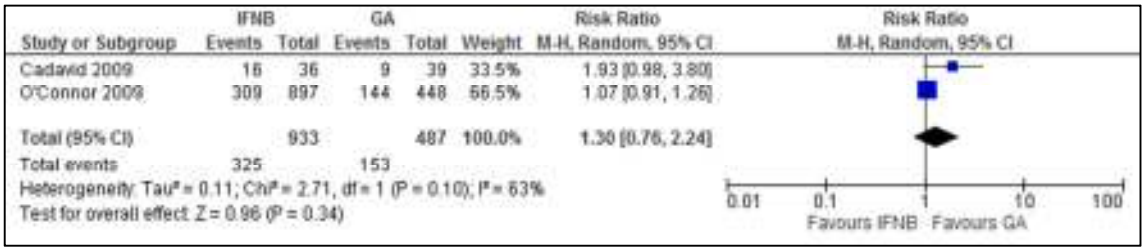

Figure 11. Analysis 7.1. Participants treated with steroids for relapse of MS

8. Mean changes in quality of life

No data were available.

MRI OUTCOMES

9. Mean number of active (new or enlarged) T2-hyperintense lesions per participant at 6 - 12 - 24 months

This outcome was available for one study at six months (396 participants; 14%), for two studies at 12 months (1722 participants; 60%) and for three studies at 24 months (1790 participants; 62%). For the latter time point, data from Calabrese 2012 referred to the IFN 44 mcg group. At six months, the number of active (new or enlarging) T2 lesions was significantly lower in IFN-treated than in GA-treated participants (MD -0.86, 95% CI -1.32 to -0.40; P value 0.0003). No significant differences were found between IFN and GA for this outcome at 12 months or at 24 months (MD -0.52, 95% CI -1.12 to 0.09, and MD -0.15, 95% CI -0.68 to 0.39, respectively). At 24 months, heterogeneity was significant (I^2 = 68%) (Figure12. Analysis 9.1-3). At 24 months, the results did not change when data from Calabrese 2012 evaluating the effects of IM IFN beta-1a 30 mcg versus GA were analysed (MD 0.11, 95% CI -0.67 to 0.44).

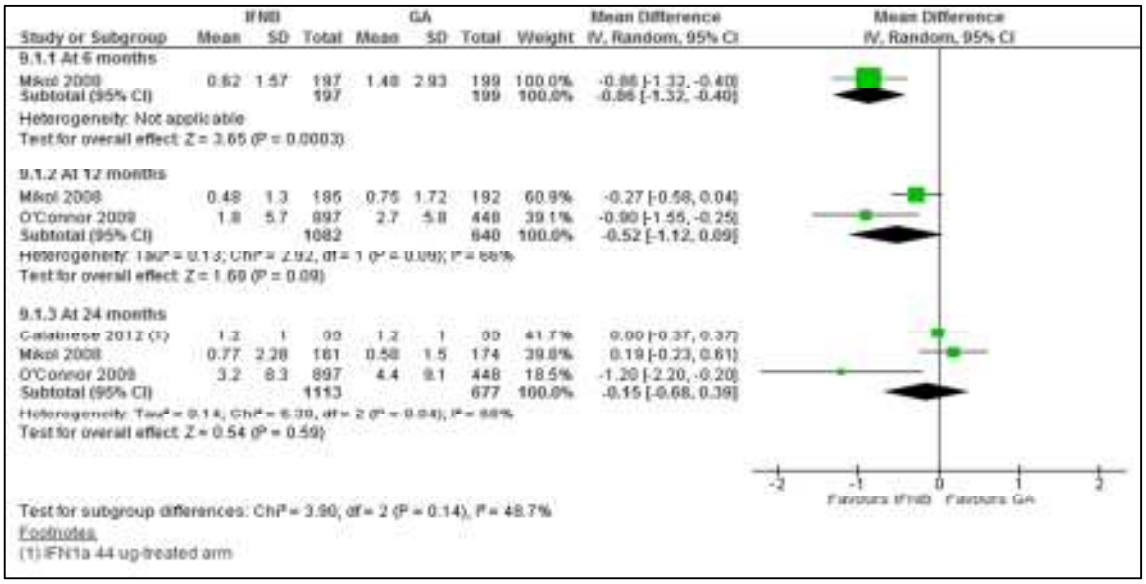

Figure 12. Analysis 9.1-3. Mean number of active T2-hyperintense lesions per participant

10. Mean number of new contrast-enhancing T1 lesions per participant at 6 - 12 - 24 months

This outcome was available at 12 months for one study (1233 participants; 43%) and at 24 months for three studies (1734 participants; 61%). For the latter time point, data from Calabrese 2012 referred to the IFN 44 mcg group. No significant differences between SC IFN 44 mcg and GA were found for this outcome at 12 months (MD -0.10, 95% CI -0.26 to 0.06) and at 24 months (MD -0.14, 95% CI -0.30 to 0.02) (Figure 13. Analysis 10.1). At 24 months, the results did not change when data from Calabrese 2012 evaluating the effects of IM IFN-beta 1a 30 mcg versus GA were analysed (MD -0.16, 95% CI -0.31 to 0.00).

No data at six months were available.

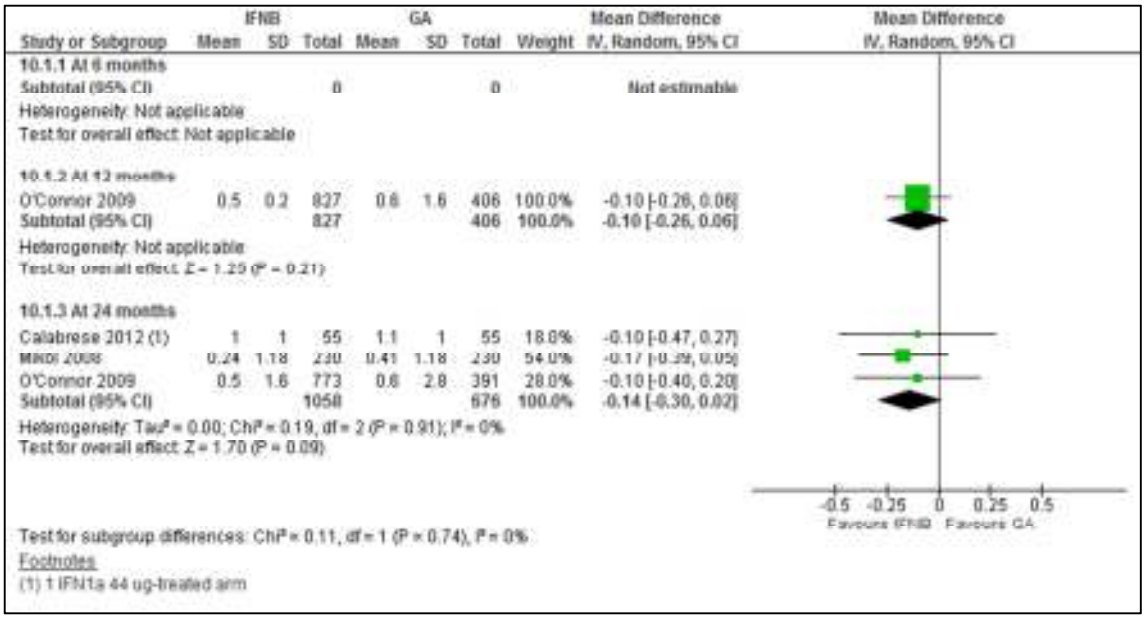

Figure 13. Analysis 10.1-3. Mean number of new T1-hypointense lesions per participant

11. Mean change in total T2-hyperintense lesion volume at 12 - 24 months

This outcome was available for one study at 12 months (1221 participants; 43%), for two studies at 24 months (1608 participants; 56%) and for another study at 36 months (509 participants; 19%). The mean increase in T2 lesion volume was significantly less in IFN-treated than in GA-treated participants at month 12 and at month 24 (MD -0.40, 95% CI -0.59 to -0.21, and MD -0.58, 95% CI -0.99 to -0.18; P values < 0.0001 and 0.004, respectively). No heterogeneity was found. The difference favouring IFN (MD -0.26, 95% CI -1.04 to 0.52) was also present in 36-month data from a single study (Lublin 2013) but it did not reach statistical significance (Figure 14. Analysis 11.1-3).

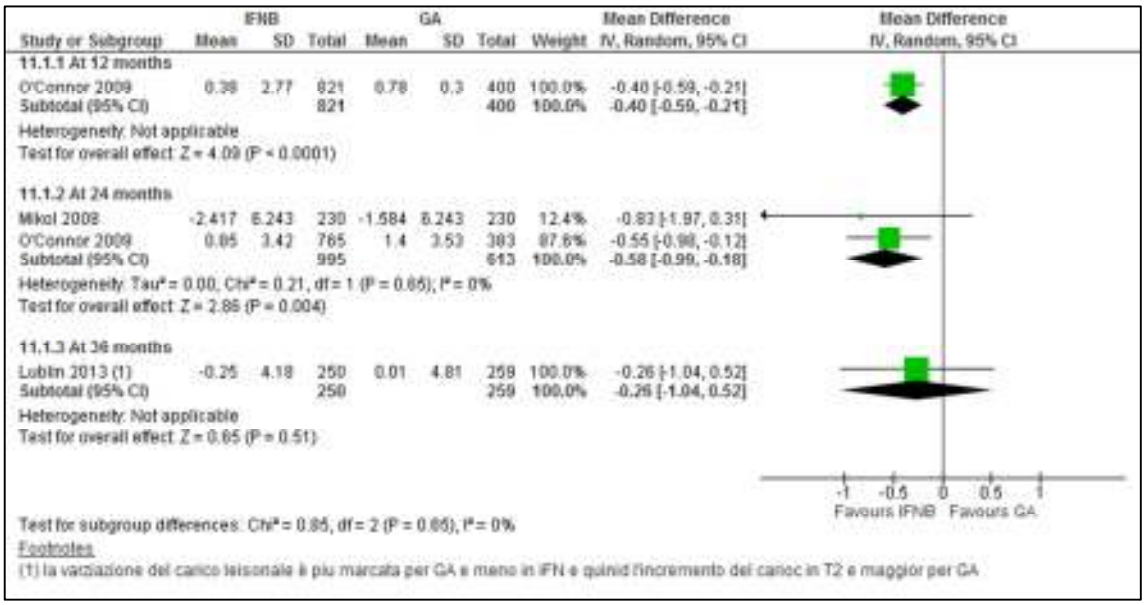

Figure 14. Analysis 11.1-3. Mean change in total T2-hyperintense lesion volume

12. Mean change in total T1-hypointense lesion volume at 12 - 24 months

This outcome was available at 12 months for one study (1207 participants; 42%) and at 24 months for two studies (1602 participants; 56%). The mean increase in T1 lesion volume was significantly less in IFN-treated than in GA-treated participants at month 24 (MD -0.20, 95% CI -0.33 to -0.07; P value 0.003). The difference favouring IFN (MD -0.06, 95% CI -0.18 to 0.07) was also seen in 12-month data from a single study (O'Connor 2009) but did not reach statistical significance (Figure 15. Analysis 12.1-2).

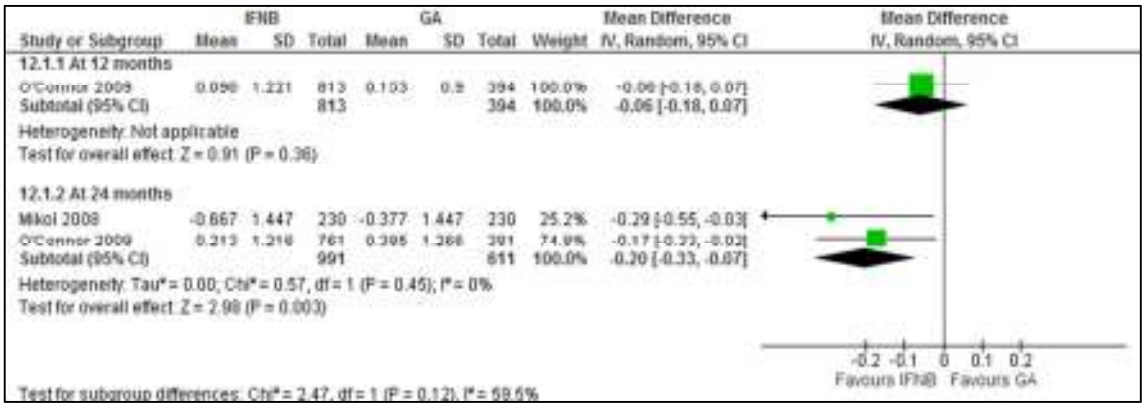

Figure 15. Analysis 12.1, 12.2. Mean change in total T1-hypointense lesion volume

13. Mean change in total brain volume (as a measure of atrophy) at 12 - 24 months

This outcome was available at 12 months for one study (1137 participants; 40%) and at 24 months for two studies (1552 participants; 54%). Data at 12 months did not show a significant difference between IFN and GA with regard to brain volume changes (MD -0.10, 95% CI -0.22 to 0.02). At 24 months, mean brain volume reduction was significantly greater in IFN-treated than in GA-treated participants (MD -0.12, 95% CI -0.23 to -0.01; P value 0.04). At 24 months, the heterogeneity was significant (I^2 = 83%) (Figure 16. Analysis 13.1-2).

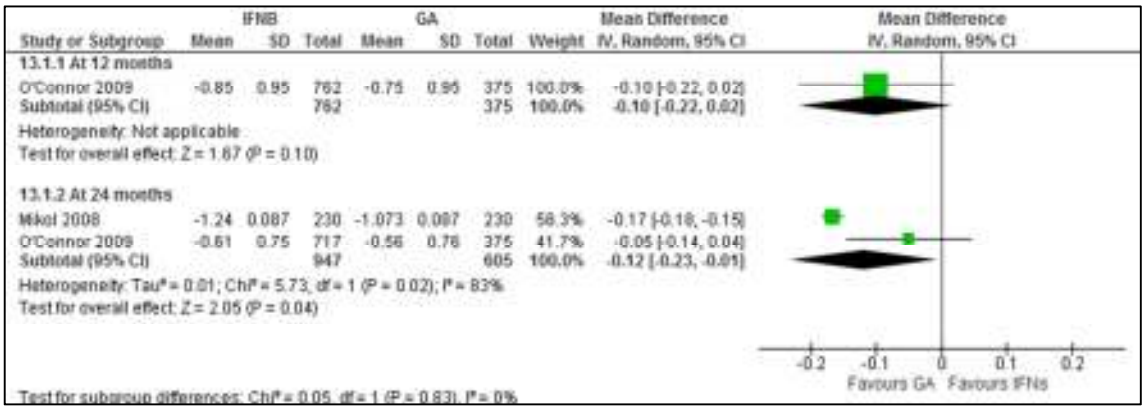

Figure 16. Analysis 13.1-2. Mean change in total brain volume

3.2 The quality of telephone triage

Obligatory questions asked compared to those expected to be asked was consistently below 100% for each of the four clinical cases: 36% in the clinical case of a vomiting child, 32% in the case of a child with fever, 28% in the case of the adult with fever and 27% of the adult with nosebleed (Tables 3–6).

Few questions were extremely relevant in determining the clinical severity of the patients symptoms but no doctor asked for them. Doctors were used to ask a few questions, probably because the symptoms described in the simulated cases were very common and in most situations they were unlikely to mask serious diseases. Both in the adult (Tables 3 and 4) and the child patients (Tables 5 and 6), no attention was given to symptoms or conditions which might signal a more severe disease, such as asking about travel abroad to exclude any tropical disease.

In analysing the management decisions we found, as expected, that no doctor advised calling the emergency service or requested a home visit to the patient, in any of the simulations. In only three clinical cases out of the 40 simulated (2 adults with fever and 1 adult with epistaxis), all of which were considered to be manageable by telephone, the patient was advised to go to an OOH centre for a face-to-face consultation. In the remaining 37 cases, the call handler opted for telephone advice aimed at self-management of the patients (Table 7). The average duration of the 40 simulated calls was 3 min 47 s.

At the end of the study we asked the doctors involved if they had suspected that the calls received were fake calls; no doctor gave a positive response.

Table 3. Adult with nosebleed. Number of doctors who asked the obligatory questions

How many episodes in the last few hours?	6
How long ago did nosebleed end?	5
Amount of blood lost?	5
Any recent concussion?	4
Did or do you feel any object in your nose or nasal passages?	4
Medicine recently taken or currently being taken?	2
General feeling of weakness?	1
Any bruises on the skin?	0
Any problems concerning coagulation?	0
Any blood in vomit?	0

Table 4. Adult with fever. Number of doctors who asked the obligatory questions

Any flu symptoms?	8
Duration of febrile state?	7
Highest temperature measured?	6
Any nausea, vomiting or diarrhoea?	4
Chills?	1
Any recent trips abroad?	1
Any problems urinating?	1
Any blotches on skin?	0
Any recent illnesses?	0
Any neck stiffness or neck pain?	0

Table 5. Child with vomiting. Number of doctors who asked the obligatory questions

General conditions of the child?	9
Any stomach ache?	9
How many times has the child vomited?	8
Running a temperature?	5
Headache?	5
How much did the child vomit?	4
Any possibility of having eaten foods gone bad?	4
Has the child ingested liquids in the last 6 hrs?	4
Any recent head injuries?	3
Any back stiffness?	2
Is the child sensitive to light?	0
Any blotches on the skin?	0
Is urination painful?	0
Any blood visible in vomit?	0
Last time child urinated?	0

Table 6. Child with fever. Number of doctors who asked the obligatory questions

Any flu symptoms?	8
How high is the temperature?	8
General conditions of child?	7
How long has fever lasted?	7
Any particular pains?	5
Headache?	4
Any back stiffness or pain?	2
How much liquid consumed in past few hours?	2
Any nausea/vomiting/diarrhoea?	2
Any blotches on skin?	1
Any recent trips abroad?	1
Stomach ache?	0
Has the child had a recent injury?	0
Last time child urinated?	0
How is child's breathing?	0

Table 7. Management decisions concerning the 40 simulated cases

Patients instructed to treat themselves 21	21
Patients instructed to treat themselves and call back if necessary	16
Patients advised to go to an outpatient clinic 3	3
Patient advised to call a doctor for a home visit 0	0
Patient advised to go to hospital 0	0
Total	40

3.3 The efficacy of CDSS

The results of our search and selection process are presented in Figure 17. We identified 28 RCTs, which met the predefined inclusion criteria. Eighteen studies reported mortality or morbidity data and were included in the meta-analyses, while 10 more studies reported only economic outcomes (see the corresponding references section).

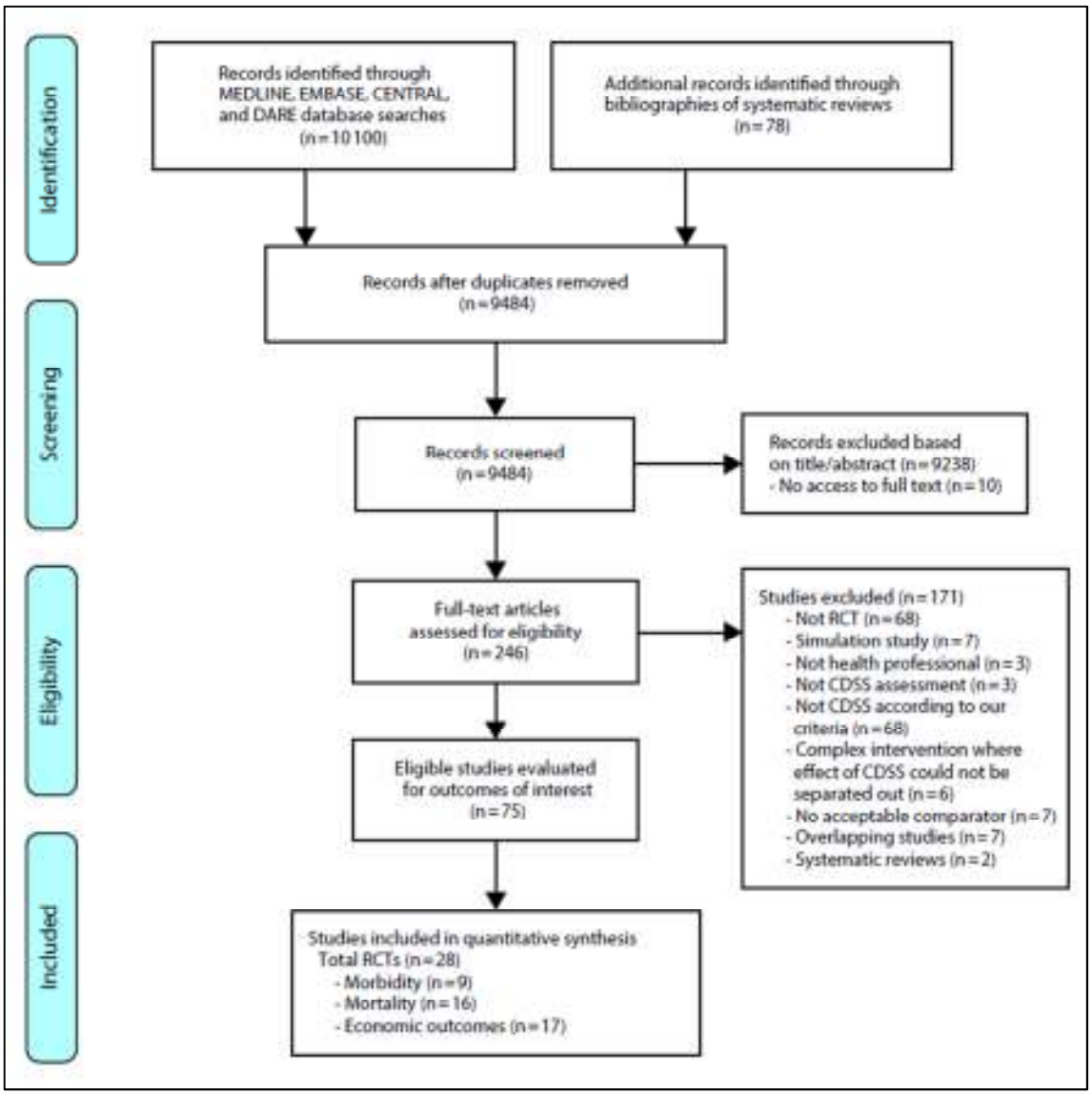

Figure 17. Study flow diagram CDSS Efficacy

Risk of Bias in Studies

Overall, the assessment of the 18 studies incorporated in the meta-analyses indicated high risk of bias across 7 (39%) and unclear risk for 10 studies (56%). Only 1 study (5%) was judged to be at low risk for bias (Figure 18). We noticed that the majority of trials did not measure mortality as an outcome, but reported it as additional information, often as a reason for loss to follow-up. Readers should be aware that our risk of bias assessment did not evaluate studies based on their intended outcomes but according to 2 outcomes of our systematic review: mortality and morbidity.

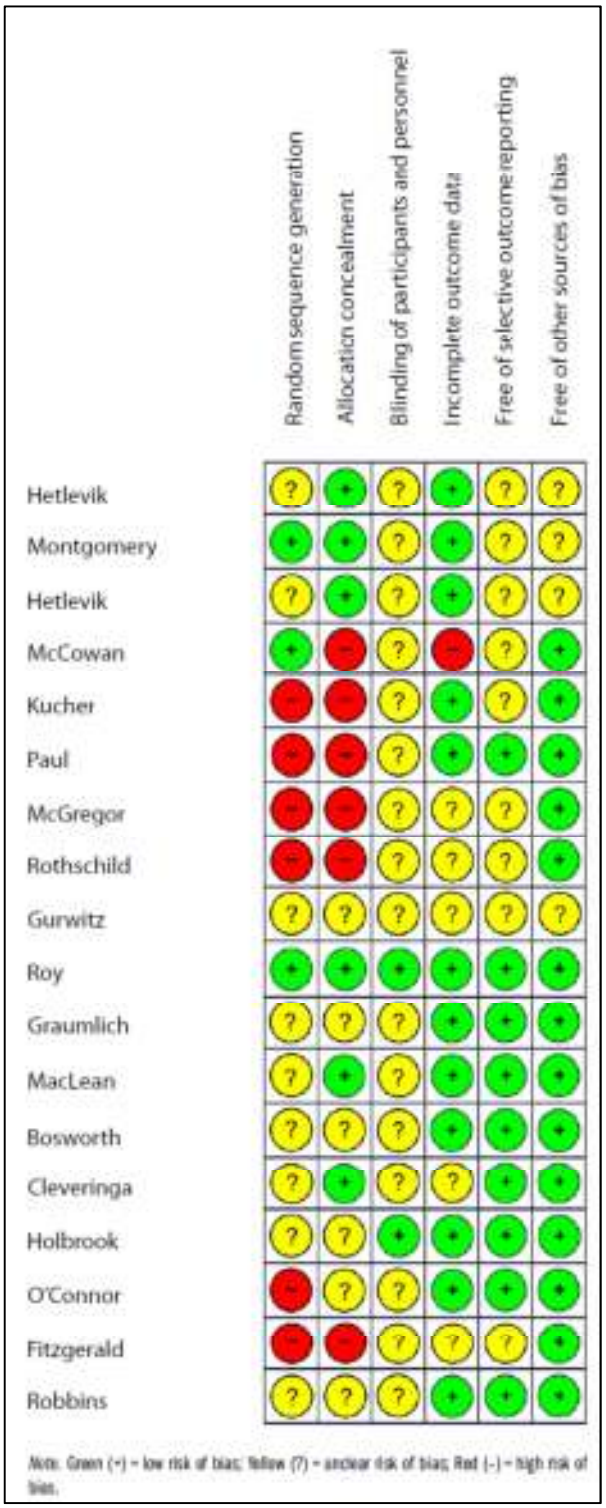

Figure 18. Judgements about each risk of bias item for each included study

Meta-Analysis of Mortality Outcomes

Sixteen RCTs contributed to this analysis. A total of 37395 individuals participated in these trials: 18848 in the intervention groups and 18547 in the control groups. Seven trials reported a lower mortality in the intervention group, while 8 trials reported a higher mortality. Only 3 were statistically significant. The overall mortality rate on all 16 RCTs was 6.2% in the intervention groups (1171 deaths) and 6.0% in the control groups (1111 deaths). The pooled effect estimate was not statistically significant assuming either a fixed effects model (RR =1.00; 95% CI = 0.92,

1.08), or a random effects model (RR = 0.96; 95% CI = 0.85, 1.08). Figure 19a shows the forest plot of the RR estimates and 95% CIs from the individual trials and the pooled results.

Meta-Analysis of Morbidity Outcomes

Nine RCTs contributed to this analysis. A total of 13868 individuals participated in these trials. The analysis revealed a weak inverse association between CDSS use and morbidity from any disease. The difference between the CDSS and control groups in the occurrence of morbidity outcomes was marginally significant assuming a random effects model (RR = 0.82; 95% CI = 0.68, 0.99), but not significant assuming a fixed-effects model (RR = 0.91; 95% CI = 0.83, 1.00). Figure 19b shows the forest plot of the RR estimates and 95% CIs from the individual trials and the pooled results.

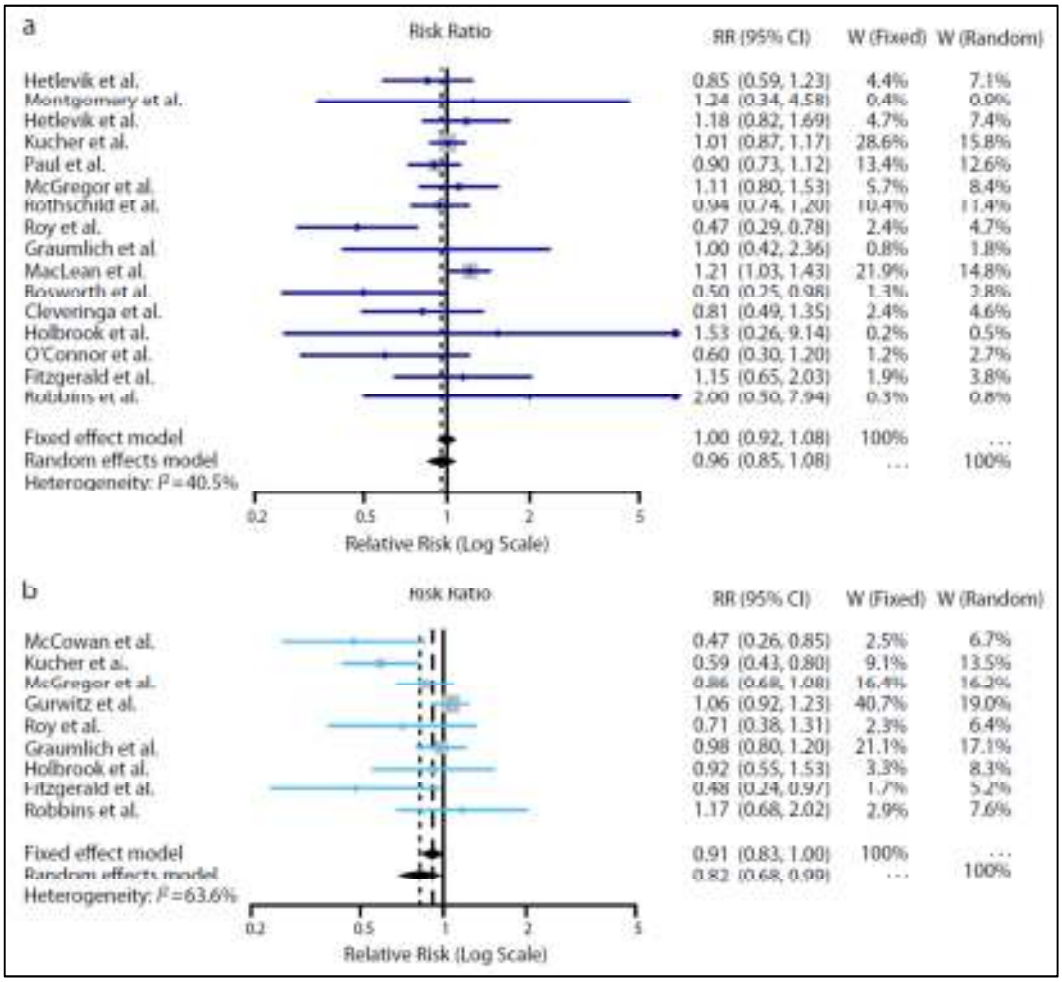

Figure 19. Forest plots CDSS effictiveness on mortality and morbidity prevention.

Qualitative Assessment of Economic Outcomes

Seventeen RCTs reported economic outcomes. Three of these presented the economic data in separate publications. Differences were seen for costs and health service utilization (e.g., drug or test orders), but these were often small in magnitude. Across economic outcomes, interventions equipped with CDSSs did not consistently perform better than non equipped ones.

3.4 The efficacy of e-learning

Results of the search

We identified 3465 articles through the search strategy (MEDLINE 2398, EMBASE 608, CENTRAL 417, CDSR 6, DARE 7, CMR 17, HTA 9, NHSEED 2) and one additional article from the other reviews we took into account. We excluded 3328 articles based on the abstracts. A total of 142 articles were provisionally selected as potentially fulfilling the inclusion criteria (Figure 20). (See the corresponding references section).

We searched for the full text of 137 articles to determine their eligibility for inclusion; we could not retrieve 3 articles and we assessed the remaining 134: 120 studies were excluded and 14 were included.

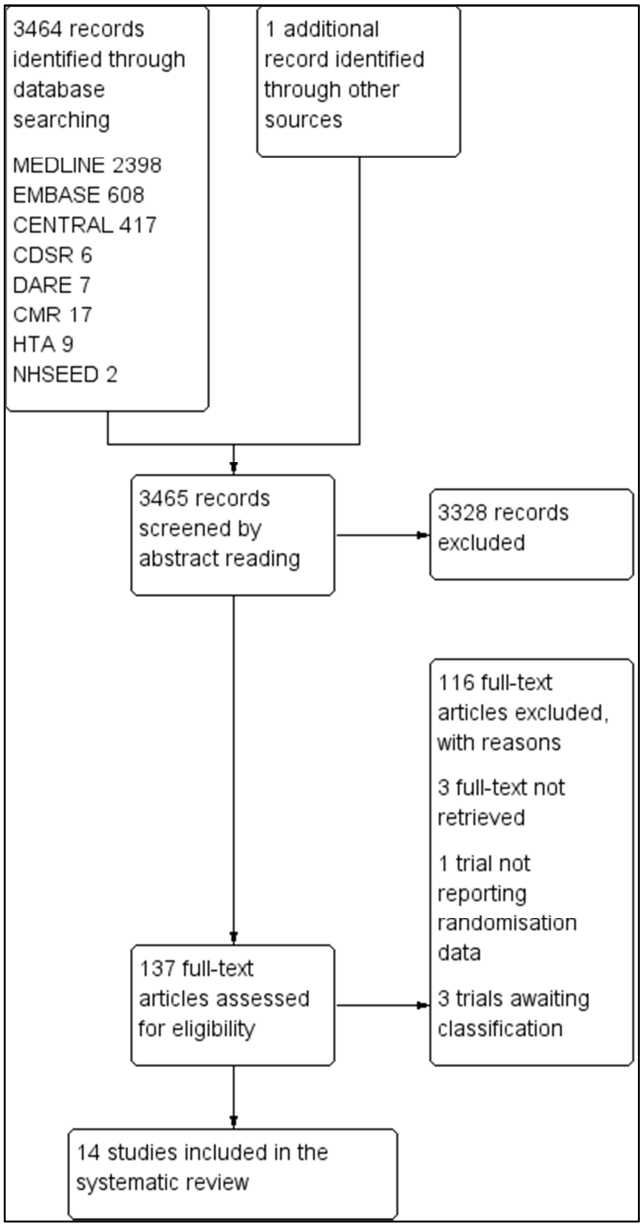

Figure 20. Study Flow Diagram e-Learning Efficacy

Included studies

Finally, 14 RCTs providing data on 5590 participants met our predefined selection criteria. All trials were published between 2005 and 2014 and had a parallel design. The mean sample size is 400 participants but only 3 trials have more than 150 participants. Five trials were performed in the USA (Benjamin 2008, Fordis 2005, Harris 2008, Le 2010, Levine 2011), while the remaining nine studies were performed in Japan (Horiuchi 2009), Netherlands (Hugenholtz 2008), Finland (Mäkinen 2006), Australia (Maloney 2011, Perkins 2012), Brasil (Paladino 2007), UK (Perkins 2012), Taiwan (Sheen 2008), Norway (Simonsen 2014), Iran (Khatony 2009); only one study was performed in two countries (Perkins 2012). All the included studies but two (Paladino 2007, Sheen 2008) addressed medical expertise as educational topic according to The CanMEDS Framework. Two studies (Paladino 2007, Sheen 2008) assessed interventions aimed at improving leadership or mixed competences, respectively.

Characteristics of participants and settings

Three trials randomised 4714 mixed health professionals (Levine 2011, Maloney 2011, Perkins 2012), six trials randomised 543 nurses (Horiuchi 2009, Khatony 2009, Mäkinen 2006, Paladino 2007, Sheen 2008, Simonsen 2014), four trials randomised 300 doctors (Fordis 2005, Harris 2008, Hugenholtz 2008, Le 2010) and one trial randomised 33 children-care health consultants (Benjamin 2008). Four trials were performed in a primary care setting (Fordis 2005, Harris 2008, Le 2010, Levine 2011), five trials in a secondary care hospital setting (Horiuchi 2009, Khatony 2009, Mäkinen 2006, Paladino 2007, Sheen 2008), two in a mixed setting (Perkins 2012, Simonsen 2014), one in a rehabilitation setting (Maloney 2011) and two in other settings (Benjamin 2008, Hugenholtz 2008).

Characteristics of educational interventions used in the trials

Among the 14 included trials, 12 compared an e-learning intervention with face-to-face residential learning, while two compared e-learning with guideline dissemination or availability (Le 2010, Levine 2011). In five trials, the educational intervention was accredited for CME purposes (Fordis 2005, Harris 2008, Hugenholtz 2008, Le 2010, Levine 2011). In six trials the duration of the e-learning intervention, in terms of time needed to be spent on learning, was the same as the control intervention (Harris 2008, Hugenholtz 2008, Levine 2011, Maloney 2011, Perkins 2012, Simonsen 2014); in three trials, the duration of the educational session was longer in the control groups than in the e-learning groups (Horiuchi 2009, Mäkinen 2006, Paladino 2007); in the remaining cases, this information was not reported or confused with the time the intervention was available to the participants (online or in terms of the time the guidelines could be studied before the outcome assessment). We considered the amount of time needed to be spent on learning to be short (less than one week) in all trials except two (Le 2010, Levine 2011). In nine trials, e-learning was administered alone, not in combination with other interventions; in the five remaining trials (Fordis 2005, Le 2010, Levine 2011, Maloney 2011, Perkins 2012) we considered e-learning as being a core and essential element of a multifaceted educational intervention. The e-learning tools interactivity was high (combination of

at least three components) in nine trials and low in the others (Harris 2008, Horiuchi 2009, Hugenholtz 2008, Paladino 2007, Sheen 2008).

Outcome assessment

The patient outcomes were detected by administrative data analysis, health professionals' behaviours by patients chart audit and by administrative data analysis, and skills by written skills tests or by simulations and objective structured clinical examinations. The knowledge outcome was detected by questionnaires: in four trials, the authors reported that the questionnaire was previously validated (Fordis 2005, Harris 2008, Khatony 2009, Perkins 2012), while the other studies did not include information about this item.

Duration of follow-up and outcome assessment times

The median follow-up time from the conclusion of the educational intervention to the last outcome assessment was 1.5 weeks ranging from 0 to 52 weeks. Only three trials had more than one outcome assessment follow-up time during the study (Fordis 2005, Harris 2008, Le 2010)

Excluded studies

We excluded 120 studies for the following reasons: 49 articles because of the control group (no intervention at all, intervention on a different topic, or different types of e-learning in the control group), 30 articles because of the type of included participants (students or trainees); 21 articles because of the study design (non randomised controlled trials); 10 articles because of the type of intervention used (not e-learning, not delivered by Internet, not core and essential or not compliant with CANMEDS criteria); 6 articles because of type of outcome assessed (no outcome of interest or self reported outcomes); 3 because the full text was not retrievable and 1 because it lacked data on the number of participants randomised per group and the authors were unable to answer our request for clarification.

Ongoing trials

No ongoing trial was found.

Risk of bias in included studies

We could not adequately assess the risk of bias of studies due to poor reporting in the majority of studies. The overall risk of bias was high for 7 studies and unclear for 6 studies. Perkins 2012 was judged at a low overall risk of bias as we were able to exclude a high risk of bias for all individual dimensions. Incomplete data was the main source of bias (evaluated at high risk of bias in nearly 50% of the studies, seven out of 14), followed by other risks of bias (duality of interest; three studies at high risk of bias) and imbalance in the baseline characteristics (four studies at high risk of bias).

Allocation (selection bias)

Seven studies used acceptable methods to generate the allocation sequence, including computerised random number generators (Fordis 2005, Horiuchi 2009, Maloney 2011, Perkins 2012, Simonsen 2014), a blind name draw (Harris 2008) and a coin flip (Sheen 2008); one study was at high risk of selection bias as participants from the same practice were matched into pairs before randomisation (Le 2010). Eight studies clearly explained how the sequence was concealed (Benjamin 2008, Fordis 2005, Harris 2008, Horiuchi 2009, Le 2010, Levine 2011, Perkins 2012, Sheen 2008), while the remaining did not mention the methods used by the investigators.

Blinding (performance bias and detection bias)

Participant blinding is not achievable in educational studies, so performance bias might be unavoidable in this setting. We considered the blinding of assessors: the risk of detection bias was rated as high in one study (Sheen 2008) because the authors clearly stated that the assessors were not blind. The study was so small that assessors could possibly know and remember the allocation of the participants. Also in Perkins 2012 authors were unable to ensure blinding of outcome assessors. However this study was so large that we assumed some degrees of separation between participants and assessors, and the process of measurement was well structured, limiting the risk of bias. Only three studies reported that the knowledge of the allocated interventions was adequately prevented (Fordis 2005, Mäkinen 2006, Maloney 2011). The remaining studies did not report information on blinding of outcome assessors.

Incomplete outcome data (attrition bias)

Seven studies were judged at high risk of attrition bias (Fordis 2005, Harris 2008, Horiuchi 2009, Le 2010, Levine 2011, Maloney 2011, Sheen 2008): one study (Sheen 2008) used a per protocol analysis as the remaining six studies at high risk of bias reported a loss ranging from 15% (Fordis 2005) to 47% for (Levine 2011). In four out of six aforementioned studies (Fordis 2005, Harris 2008, Le 2010, Maloney 2011), the attrition was bigger in the e-learning group than in the control group. Three studies were judged at low risk of attrition bias (Hugenholtz 2008, Perkins 2012, Simonsen 2014).

Selective reporting (reporting bias)

With the exception of one (Horiuchi 2009), no study presented inconsistencies between outcomes declared in the methods section and outcomes reported in the results section.

Other potential sources of bias

We considered the duality of interest as a potential source of bias. Two studies (Fordis 2005, Harris 2008) were supported by private sponsor grants and one (Le 2010) received support in terms of evaluation tool or e-learning modules development.

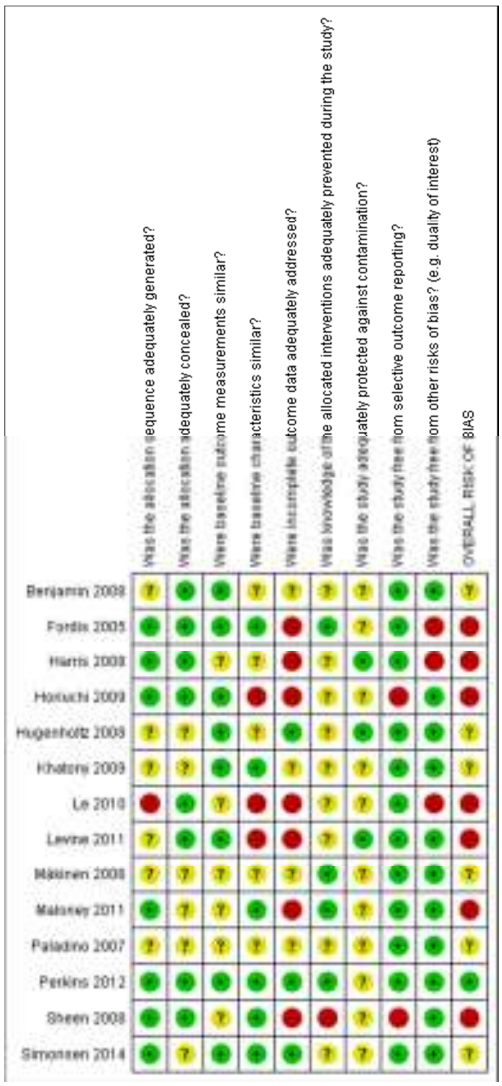

Figure 21. Judgements about each risk of bias item for each included study

Effects of interventions

The Summary of Findings Table (Table 8) reports the effectiveness of e-learning compared to traditional learning in terms of patient outcomes, health professionals' behaviours, skills and knowledge.

PRIMARY OUTCOMES

Patient outcomes

One study addressed patient outcomes (Levine 2011). This study randomised 168 primary care clinics (847 health professionals) to a highly interactive e-learning versus face-to-face traditional learning. After at least 12 months of exposure to the interventions, a patient administrative data review was performed and the groups were compared for two primary patient outcomes indicators. No difference was found in terms of proportion of patients with on-target low-density lipoprotein cholesterol level [6399 patients; adjusted difference in improvement between the groups 4.0% (95% CI -0.3 to 7.9)] and in terms of proportion of patients with on-target glycated haemoglobin level [3114 participants patients; adjusted difference in improvement between the groups 4.6% (95% CI -1.5 to 9.8)]. The data about SD or SE were unavailable and no further elaboration of these data was possible.

Table 8. e-Learning Efficacy Summary of Finding Table

E-learning compared to traditional learning for health professionals							
Patient or population: health professionals **Settings: postgraduate education** **Intervention: e-learning** **Comparison: traditional learning**							
Outcomes	**Impact**				**No of Participants (studies)**	**Certainty of the evidence (GRADE)**	**Comments**
Patient outcomes Follow up 12 months	One study found a no difference in: - Last measured LDL level <100mg/dL [6399 participants patients; adjusted difference 4.0% (95% CI -0.3 to 7.9%). - Last measured glycated haemoglobin level <8% [3114 participants patients; adjusted difference 4.6% (95% CI -1.5 to 9.8%)				168 primary care clinics (847 health professionals) (1 study)	⊕⊕⊖⊖ **low**[1]	
Health professionals' behaviours Follow up 3 to 12 months	**Illustrative comparative risks* (95% CI)**			**Relative effect (95% CI)**			
	Behaviour	**Assumed risk** **Traditional learning**	**Corresponding risk** **E-learning**				
	Screening for dyslipidaemia	891 screened per 1000	880 screened per 1000	**OR 0.90** (95% CI 0.77 to 1.06)	950 health professionals (2 studies)	⊕⊕⊖⊖ **low**[2]	Studies reported multiple outcomes without specifying the primary outcome. To assess consistency, we explored three other possible combinations between the two studies indicators.
	Treatment for dyslipidaemia	952 treated per 1000	958 treated per 1000	**OR 1.15** (95% CI 0.89 to 1.48)			
Health professionals' skills (cardiac arrest simulation test - CASTest) measured as passes/fails) Follow up 0-2 weeks	**Illustrative comparative risks* (95% CI)**		**Relative effect (95% CI)**				
	Assumed risk **Traditional learning**	**Corresponding risk** **E-learning**					
	799 passes per 1000	730 passes per 1000	**OR 1.46** (1.22 to 1.76)	2823 health professionals (4 studies)	⊕⊕⊕⊖ **moderate**[3]		Four trials considered this outcome. Data were not pooled

						because studies provided incomplete reporting of data or could not be aggregated. According to the authors' conclusions, three studies (2640 participants) but one favoured traditional learning. One large study (2562 participants) at low risk of bias provided complete data which were used for the skill benefit estimate.
	Average effect (fixed effect) *	**Average effect (random effects) ***				
Health professionals' knowledge (measured as correct answers) Any follow up 0-12 weeks	**0.04 SDs** (-0.03 lower to 0.11 higher) favouring traditional learning	**-0.09 SDs** (-0.27 lower to 0.09 higher) favouring e-learning		3236 (11 studies - 8 included into meta-analysis)	⊕⊕⊖⊖ **low**[4]	Heterogeneity among studies ($I^2$47%). Predictive interval was -0.55 to 0.37. Irrespective of the preferred meta-analytic model, the magnitude of differences between groups is limited, negligible to minor in extent.

* SD: standard deviation

We interpreted SMDs using the following rules suggested by the Cochrane Handbook:

<0.40 represents a small effect size

0.40 to 0.70 represents a moderate effect size

>0.70 represents a large effect size

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

Health professionals behaviours

Two studies addressed this outcome on 950 health professionals (Fordis 2005, Levine 2011). Fordis 2005 randomised 103 primary care physicians to a highly interactive and multifaceted e-learning versus face-to-face traditional learning. After 12 weeks, a patient chart review was performed for 20 randomly selected doctors per group and the groups were compared in terms of appropriate screening and treatment of dyslipidaemia. Levine 2011 reported data from three performance quality indicators, which we considered as behaviours outcomes: beta-blocker prescription, statin prescription, ACE Inihibitor or angiotensin receptor antagonist prescription. To assess consistency we explored all the possible combinations between the indicators reported by the two studies, as we did not have any rational to prefer one indicator over the others. No difference was found in terms of the proportion of patients appropriately screened or treated in any combination. The results are compatible with either a beneficial or a detrimental effect of e-learning, with OR ranging from 0.90 to 1.15 and the 95% CI always crossing the line of no effect (Figure 22-26. Analysis 1.1-1.5). The 95% CIs, where the actual effect may be, show that e-learning may lead to better behaviours but may also worsen the outcome. The random-effects method and the fixed-effect method gave identical results as there was no heterogeneity among the two studies.

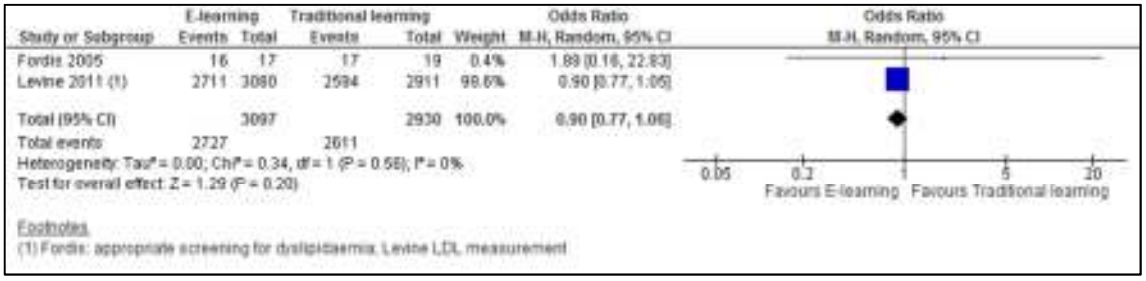

Figure 22. Analysis 1.1. Patients appropriately screened (dyslipidaemia)

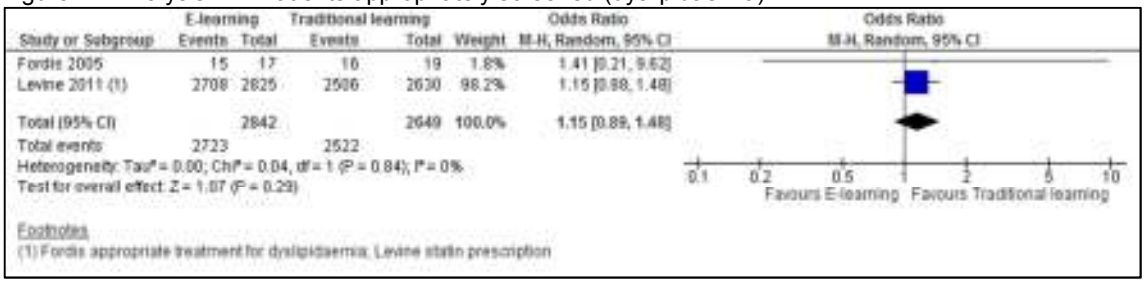

Figure 23. Analysis 1.2. Patients appropriately treated (dyslipidaemia)

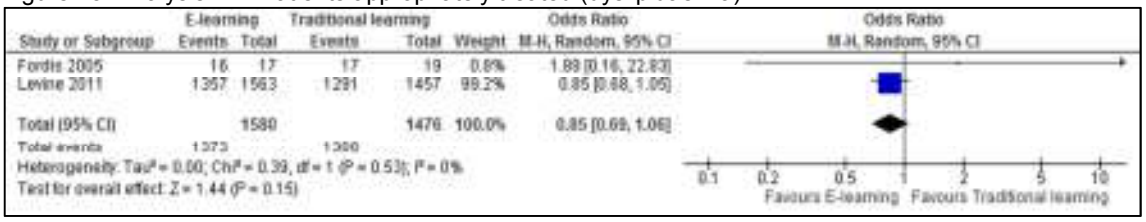

Figure 24. Analysis 1.3. Patients appropriately screened (dyslipidaemia and diabetes)

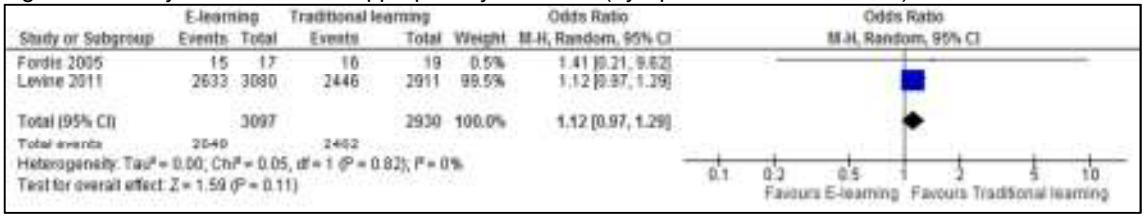

Figure 25. Analysis 1.4. Patients appropriately treated (dyslipidaemia and hypertension)

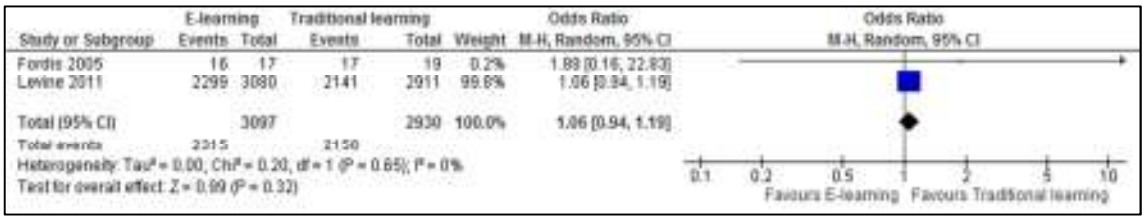

Figure 26. Analysis 1.5. Patients appropriately screened (dyslipidaemia) and treated (hypertension)

SECONDARY OUTCOMES

Health professionals skills

Four studies provided data about skills on 2823 health professionals (three week median follow-up): according to the authors' conclusions Mäkinen 2006, Perkins 2012, Sheen 2008 (2640 participants) favoured traditional learning, while Simonsen 2014 (183 participants) reported no difference. These results could not be pooled in a meta-analysis: Mäkinen 2006, gave no data about standard deviation, Sheen 2008 reported the results without providing quantitative estimates, referring to non significant difference between the groups and Simonsen 2014 was the only study reporting the data by mean and standard deviation (e-learning score 11.6 versus traditional 11.9).

Perkins 2012 requires special consideration. The study tested individual candidate performance in a cardiac arrest simulation test (CASTest) assessing resuscitation skills during a simulated cardiac arrest. The full analysis on the mixed population showed no difference between e-learning and traditional learning but the authors provided unpublished data excluding students and participants with missing professional status (2562 participants, 91% of all the participants for skills outcome). More health professionals passed the test in the traditional learning group compared to the e-learning group (OR 1.46, 95% CI 1.22 to 1.76).

Health professionals knowledge

Eleven trials (3236 participants) assessed this outcome but three trials with 154 participants (Le 2010, Maloney 2011, Sheen 2008) poorly reported their data and they cannot be pooled. The remaining eight trials (3082 participants) contributed to our meta-analyses; seven of these studies (3012 participants) (Benjamin 2008, Fordis 2005, Harris 2008, Hugenholtz 2008, Khatony 2009, Paladino 2007, Perkins 2012) assessed the results immediately after the training and three studies (225 participants) (Fordis 2005, Harris 2008, Horiuchi 2009) after a follow-up (4 to 12 weeks from the end of the intervention).

We report the data for the overall evaluation of the knowledge outcome at any time, considering the longest available follow-up data for each study. A moderate heterogeneity was found among the eight studies ($I^2$47%), so we report both fixed- and random-effects analyses: we found no differences in the effect between the e-learning and traditional learning groups, both in the fixed-effect model (SMD 0.04, 95% CI -0.03 to 0.11; Figure 27) and in the random-effects model (SMD -0.09, 95% CI -0.27 to 0.09; Figure 28). The corresponding predictive interval was -0.55 to 0.37.

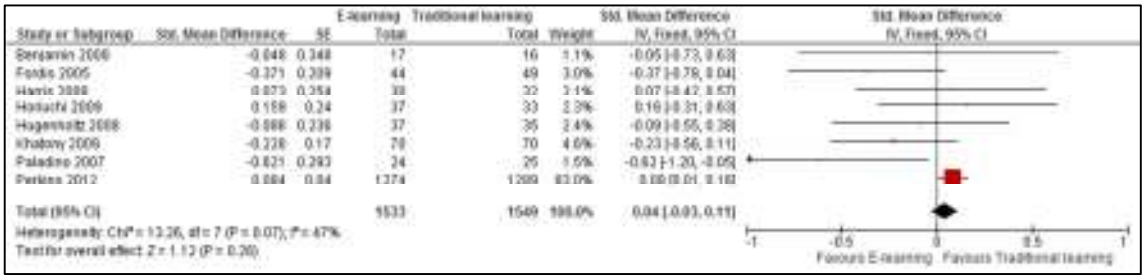

Figure 27. Analysis 3.1 Knowledge at any time (fixed effect)

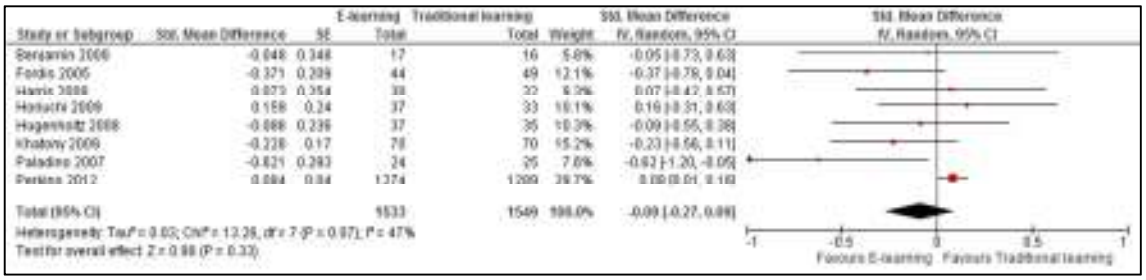

Figure 28. Analysis 3.2. Knowledge at any time (random effects)

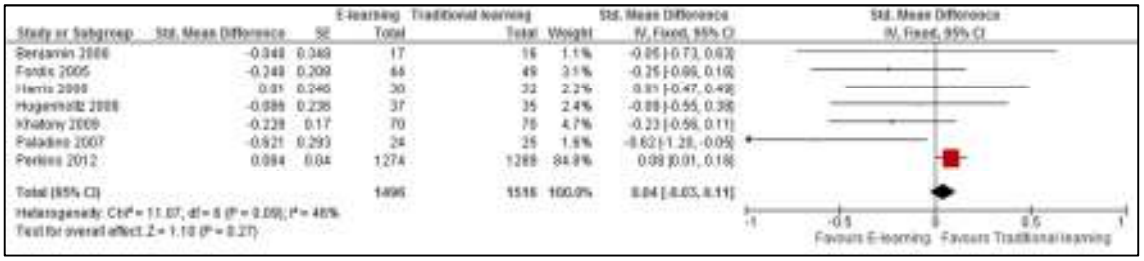

Figure 29. Analysis 3.3 Knowledge at immediately after the training

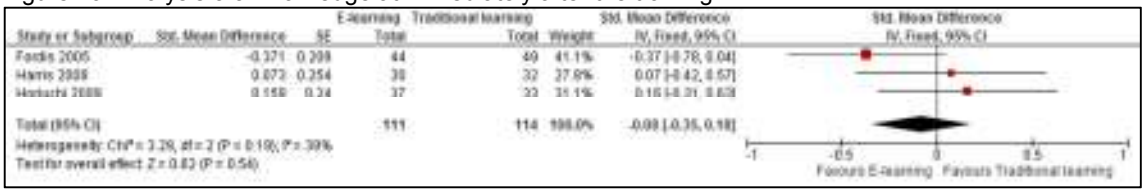

Figure 30. Analysis 3.4 Knowledge at after 3 or more months

Perkins 2012 needs special consideration: this is the larger study and it is at low risk of bias. In the fixed effect main analysis (Figure 27), its weight (2563 participants) accounts for 83% of the cumulative information and favours traditional learning (SMD 0.08, 95% CI 0.01 to 0.16), while the others studies, each contributing 1.1% to 4.6% of the overall information, are often in the opposite direction. In the random effects meta-analysis the weight of Perkins 2012 is diluted, so the point estimate of the meta-analysis, dragged by smaller studies, goes in the opposite direction favouring e-learning. Irrespective of the preferred meta-analytic model, the magnitude of differences between groups is limited, negligible to minor in extent.

The analysis immediately post training (Figure 29. Analysis 3.3) and after three or more months of follow-up period (Figure 30. Analysis 3.4) provided similar results.

Assessment of reporting bias

We did not have enough data to perform reporting bias analyses.

Subgroup analysis and investigation of heterogeneity

All of the subgroup analyses reported were performed with regards to knowledge outcome at any time. We did not undertake the standard test for heterogeneity across subgroup results to investigate the differences between two subgroups due to paucity of data.

Change in content: all included studies assessed medical expertise with the exception of Paladino 2007, which assessed managerial tools (i.e. leadership): we found no difference [SMD 0.05, 95% CI -0.02 to 0.12, I^2 26%] between e-learning versus traditional learning for medical expertise.

Change in targeted health professionals: three studies involved doctors (Fordis 2005, Harris 2008, Hugenholtz 2008), three studies nurses (Horiuchi 2009, Khatony 2009 Paladino 2007) and two studies (Benjamin 2008, Perkins 2012) other health professionals. The comparison of e-learning versus traditional learning for doctors and nurses largely overlapped [SMD -0.16, 95% CI -0.42 to 0.10, I^2 38%; SMD -0.19, 95% CI -0.44 to 0.05, I^2 54%, respectively].

Change in regulation: three studies evaluated formally accredited interventions (Fordis 2005, Harris 2008, Hugenholtz 2008): the comparison of accredited e-learning intervention versus traditional learning found no difference [SMD -0.16, 95% CI -0.42 to 0.10, I^2 0%,].

Change in format: high-interactive programmes: four studies assessed highly interactive e-learning interventions (Benjamin 2008, Fordis 2005, Khatony 2009, Perkins 2012): no difference was detected between e-learning intervention with high interaction and traditional learning [SMD -0.05, 95% CI -0.02 to 0.13, I^2 44%]; all except one study (Khatony 2009) assessed a short e-learning intervention: no difference was found between short e-learning interventions and traditional learning.

Sensitivity analysis

Excluding studies assessed at an overall high or unclear risk of bias was not applicable because all of the studies were rated at high or unclear risk of bias except Perkins 2012; no cross-over trial was found.

3.5 Training intervention for improving telephone skills

Results of the search

We identified 12209 articles through the search strategy and 1 additional article from experts suggestion. We excluded 12168 articles on the basis of abstracts considered not pertinent. A total of 42 articles were provisionally selected as potentially fulfilling the inclusion criteria. We considered these as potentially eligible and retrievable. We evaluated the full text and excluded 41 with reasons: 18 were excluded because of the absence of an experimental design; 7 were before-after studies without control; 2 were excluded because of use of a non validated tool to measure the clinicians' telephone skills change; 14 were excluded for other reasons. Only one study fulfilled our inclusion criteria. The study selection process is summarised in Figure 31. (See the corresponding references section).

Included studies

One controlled before and after study met the inclusion criteria and was included in this review (Wood 1989). This study assessed the effect of a role play telephone management curriculum for paediatric residents on history taking and case management skills. Six residents participated in three one-half-hour group sessions with a role play curriculum stressing a structured approach to telephone management of two paediatric problems; each resident had the opportunity to participate in several mock telephone conversations in which one resident played the parent of a sick child and the other resident played the clinic doctor.

A seven residents control group received no formal instruction in telephone case management. Pre-test and three months later post-test calls were made to a particular resident during times when he or she was assigned to function as “telephone doctor” in the residents continuity clinic. Using standardized scripts, a simulated mother played the role of a mother calling the clinic for advice concerning her sick child. The standardized rating form was composed of three scales, each consisting of weighted items that were summed and then expressed as a percentage of possible perfect score, but only one scale [Specific History taking Scale (SHS) and its subset Specific History Triage (SHT)] was validated and could be taken into account.

Unfortunately no quantitative data were provided by the paper on specific history scale (and specific history triage scale) and authors could not provide additional data: the only results they stated are that “*There were no differences between intervention and control groups on the Specific History taking Scale. Even when specific history taking scores were calculated using a subset of heavily weighted items (Specific History Triage), there were no differences between groups*”.

Excluded studies

We contacted the authors to be sure that there were no additional unreported information available whenever the characteristics of the study were not completely clear. We excluded Greenberg 1999 and Ottolini 1998 because, despite their agreement with all the other inclusion criteria, we found no evidence of previous validation for the tools they used in order to detect clinicians telephone consulting skills changes.

Risk of bias in included studies

We assessed the risk of bias of the included study using the nine standard criteria suggested by the Cochrane EPOC Group (EPOC 2015). We assessed the overall risk of bias for this study to be high.

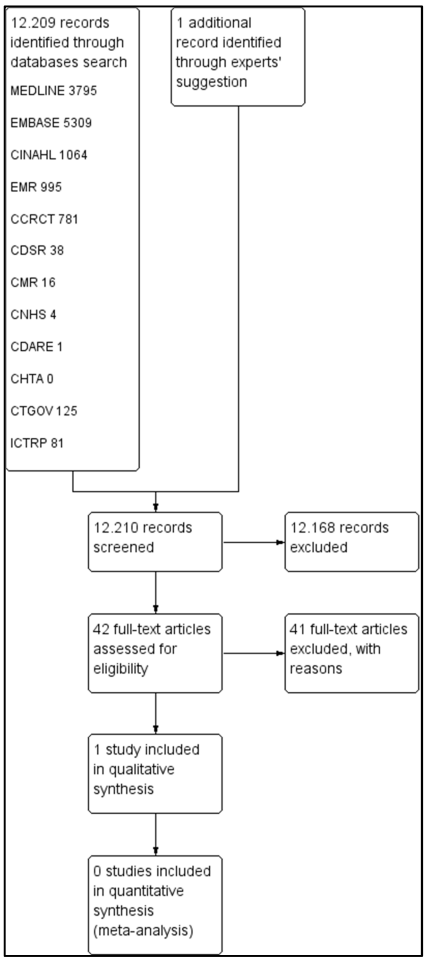

Figure 31. Study Flow Diagram Training Intervention for Telephone Skills Improvement

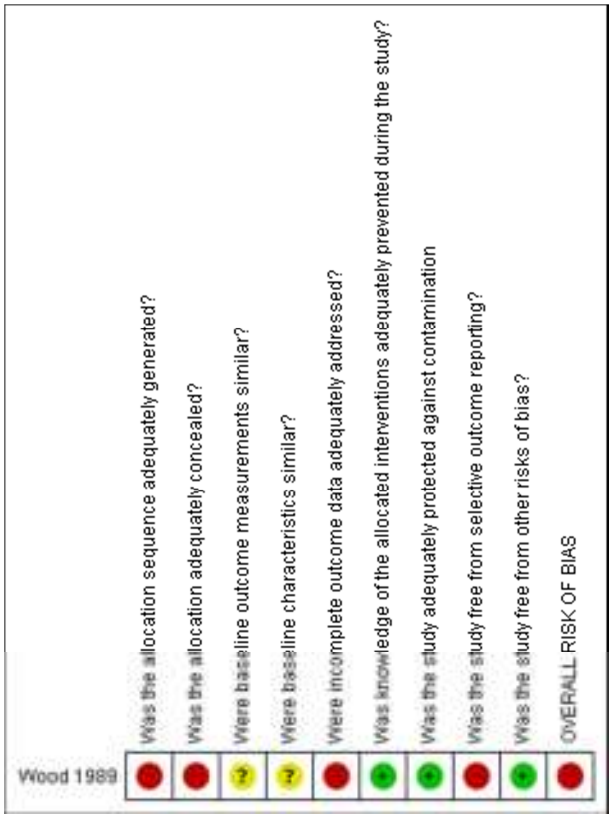

Figure 32. Judgements about each risk of bias item for each included study

Allocation (selection bias)

For generation and concealment of allocation sequence a controlled before-after study - as the study we found - should be scored “high risk” by default according to suggested risk of bias criteria for EPOC reviews (EPOC 2015). About baseline outcome measurements and characteristics no data were provided by authors and the risk of selection bias was unclear.

Blinding (performance bias and detection bias)

Audiotapes were subsequently transcribed and rated by a single trained rater who was unaware of the resident group so the detection bias was low. Nothing about contamination is mentioned but the design of the study allowed us to state against it.

Incomplete outcome data (attrition bias)

The authors stated the unpaired audiotapes were excluded from the analysis determining a high risk of bias for incomplete outcome data management.

Selective reporting (reporting bias)

Many outcomes are provided in the results section and not in the methods one: we rated the risk of reporting bias as high risk .

Other potential sources of bias

No other source of bias was identified. The overall risk of bias was rated high because of a high risk of bias in at least a key domain.

Effects of interventions

We present the evidence found in the Summary of findings table 9 for the main comparisons.

No eligible study was found that reported effects of intervention on primary patient outcomes (health outcomes, mortality, morbidity, satisfaction, urgency assessment accuracy or adverse events). The only study identified to report effects of training intervention on clinicians telephone consulting skills found no difference between intervention and control among 11 paediatrics residents but authors did not provided quantitative data.

We only included studies that assessed primary outcomes so no secondary outcomes were evaluated.

Certainty of the evidence

We judged the overall quality/certainty of the evidence for the outcomes reported to be very low (Summary of findings table 9) because the initial level of confidence about the only included study was low (non randomised evidence) and we downgraded the confidence because the risk of bias was rated as high (incomplete outcome data management and selective outcome reporting) and because of the impossibility to assess consistency of effect, imprecision, indirectness and publication bias due to lack of other studies. Based on our GRADE assessment of the certainty evidence it is uncertain whether the intervention improves clinicians' history taking and management skills (very low certainty).

Table 9. Telephone Skills Summary of Findings Table

Training intervention compared with no intervention for improving telephone consultation skills in clinicians

Patient or population: health professionals
Settings: all
Intervention: Traininng intervention
Comparison: no intervention

Outcomes	Impact	No of studies (participants)	Certaintyh of the evidence (GRADE)
Patient health outcomes	not assessed	0	no evidence available
Patients mortality	not assessed	0	no evidence available
Patients morbidity	not assessed	0	no evidence available
Patient satisfaction	not assessed	0	no evidence available
Urgency assessment accuracy	not assessed	0	no evidence available
Adverse events	not assessed	0	no evidence available
Clinicians' telephone consulting skills	It is uncertain whether the intervention improves clinicians' history taking and management skills	1 (11)	⊕⊖⊖⊖ **very low**

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

3.6 Efficacy and safety of a CDSS to support telephone triage

The study was started on 28 June 2014 and kept on until 29 May 2016.

Unfortunately since the beginning and during all the study period the internet lines power showed major problems making unusable the CDSS for the most of times in the most of centres. Figure 33 and 34 show the results of the study in terms of referrals from the non-frail and respectively frail population in the study arms when, in the intervention arm, the CDSS was accepted (fully used and not fully used) or refused and in the control arm.

Considering that the CDSS could be used just in 10,3% in the non frail population and in 16,5% of the frail population of the times it was offered, the study was judged to be failed.

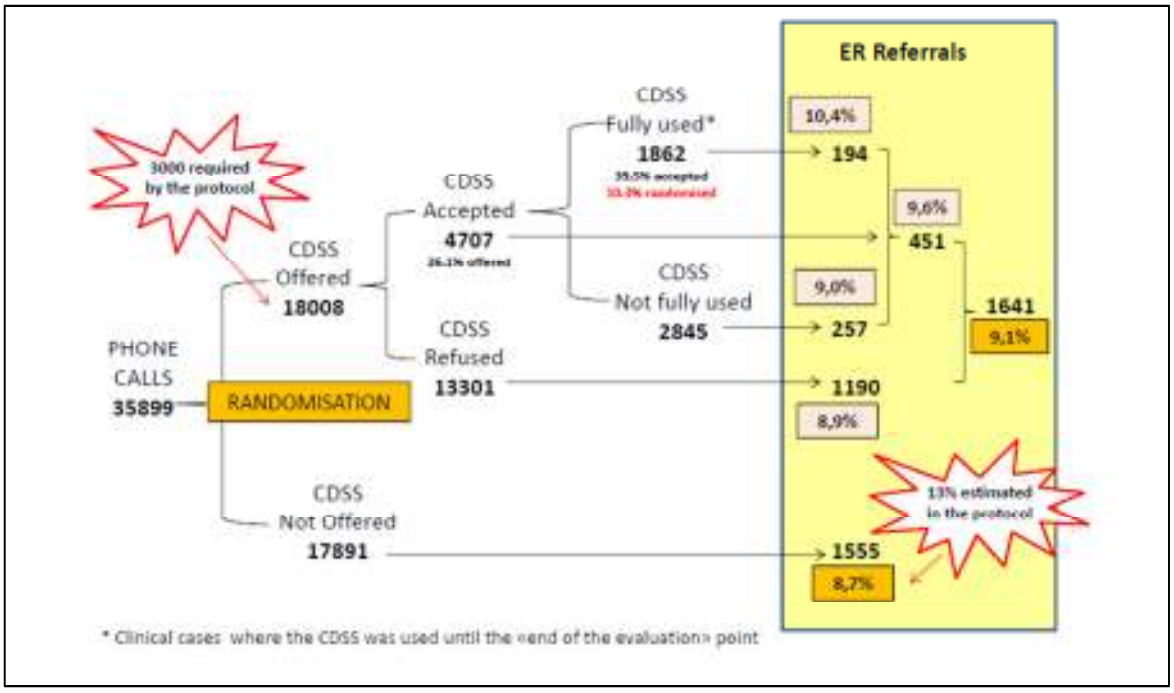

Figure 33. RCT Study results (non frail population)

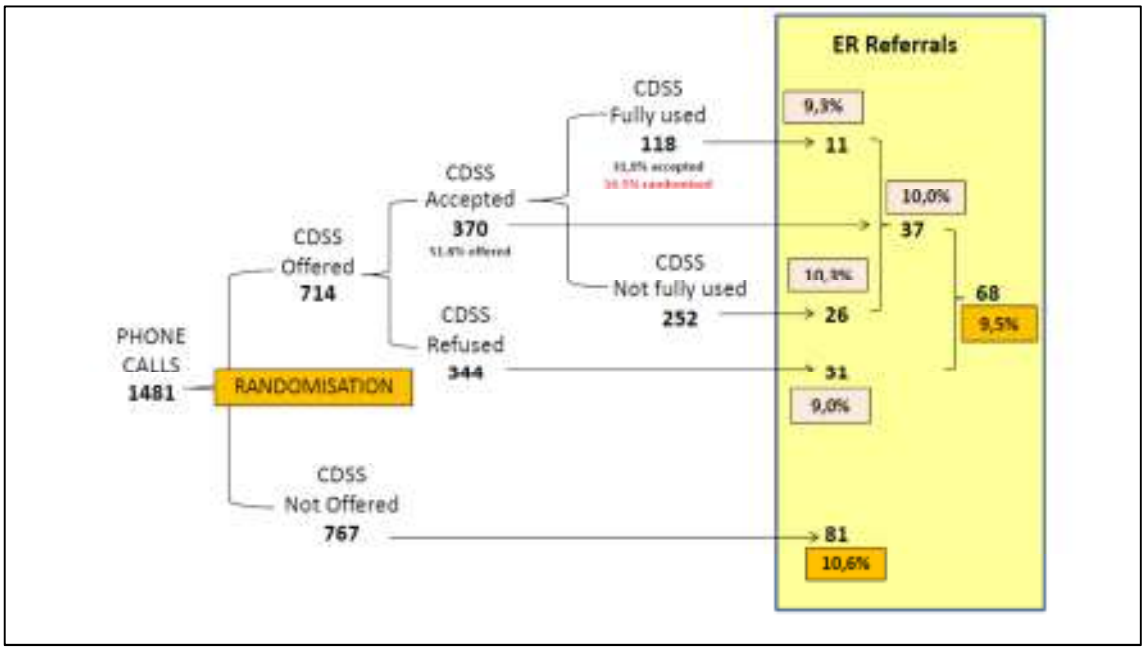

Figure 34. RCT Study results (frail population)

3.7 Barriers and facilitators to uptake CDSS

The conduction of the study was performed by the other researchers.

They reach the following results.

The adoption of CDSSs by health care professionals can be represented as a process that consists of six “positionings,” each corresponding to an individual’s use and perceived mastery of the technology. In conditions of low mastery, the CDSS is perceived as an object of threat, an unfamiliar tool that is difficult to control. On the other hand, individuals in conditions of high mastery view the CDSS as a helpful tool that can be locally adapted and integrated with clinicians’ competences to fulfil their needs.

In the first positionings, the uptake of CDSSs is hindered by representational obstacles. The last positionings, alternatively, featured technical obstacles to CDSS uptake.

DISCUSSION

4.1 Efficacy and safety of INFs-beta vs GA in RRMS

Summary of main results

The purpose of this study was to evaluate whether treatment with IFN-beta and GA in participants with RRMS was different in terms of efficacy and safety, by performing a systematic review of head-to-head RCTs.

Five RCTs contributed to the analysis . The overall population included 2858 participants (1679 participants treated with IFN and 1179 with GA). The drugs analysed in comparison with GA were SC IFN-beta 1b (two trials, 933 participants), SC IFN-beta 1a 44 mcg (two trials, 441 participants) and IM IFN-beta 1a 30 mcg (two trials, 305 participants). Duration of treatment and follow-up was three years for Lublin 2013 and two years for the other four RCTs. All studies were at high risk for attrition bias.

The main conclusion was that the two therapies seemed to have similar effects in terms of clinical efficacy, when primary outcomes (number of participants with relapse or risk of confirmed progression) and other supportive outcomes (time to first relapse, number of participants treated with steroids) were considered. However, we found some evidence, limited by the small number of participants derived from only one study, that the relapse ratio was significantly lower at 36 months in GA as compared with IFN.

The number of participants who withdrew from or dropped out of the studies because of adverse events (safety end point) was similar in the two groups. The percentage of discontinuation due to adverse effects was similar (3% for IFNs and 4% for GA). The percentage of discontinuation due to serious adverse events was also similar between groups.

In light of MRI outcomes, results showed that the effects on new/enlarging T2 or gadolinium (Gd)+ lesions at 24 months were similar. However, IFNs have a more pronounced effect on T2 and T1 lesion volume accrual at 12 and 24 months. Conversely, a lower rate of brain volume (BV) loss at 24 months was observed in GA-treated participants. The values for new T2 lesions and for brain atrophy data were limited by the heterogeneity of the results. Methodological issues were highlighted (Rudick 2009).

The quality of evidence for primary outcomes was judged as moderate for clinical end points, but for safety and some MRI outcomes (number of active T2 lesions) as low.

Overall completeness and applicability of evidence

The RCTs included in the review were homogeneous in terms of included populations, treatment schedules and outcomes measures.

All studies included only participants with RRMS with active disease (pre-study relapse frequency ranging from 0.97 to 1.9) and low disability (EDSS 1.9 to 2.35). Participants had similar baseline characteristics in terms of gender and mean age (34.8 to 39 years), although

mean disease duration was different, ranging from 0.9 (Cadavid 2009) to 6.55 years (Mikol 2008).

Studies compared the effects of different IFN products used in clinical practice versus GA. However, evaluation of the comparative effects of IFN-beta 1a IM and GA was limited because of the different time points provided (24 months and 36 months, respectively). For this reason, the main comparison in this review was that between GA and IFNs at high frequencies and dosages.

The objectives of this review were reasonably well achieved, allowing comparison of the two treatments in terms of predefined clinical and MRI measures of efficacy. The main outcome measure not available for analysis was quality of life, which impaired the possibility of adequate evaluation and comparison of the relative tolerability of different DMTs in terms of patient-reported outcomes.

We selected outcome measures that evaluated changes in clinical activity (relapse rate, attack-free status or time to first relapse), MRI activity (new or enlarging T2 lesions or Gd-enhancing lesions), clinical severity (confirmed disability progression) or MRI severity (T1 lesion load accrual or BV decrease).

Overall, results showed no differences in clinical effect between the two treatments. Different effects on some measures of MRI activity suggest that IFN might have a stronger and faster anti-inflammatory action as compared with GA. Such an anti-inflammatory effect might also account for observed differences between the two drugs in terms of brain volume reduction at 24 months, because they could be related to non–tissue-related BV loss (“pseudoatrophy” effect), although pseudoatrophy is seen primarily only during the first year (Zivadinov 2008). Furthermore, the clinical impact of these data is currently being studied (Fisniku 2008: Popescu 2013).

The results of this systematic review have several limitations.

1. The main limitation was related to high risk of attrition bias. The percentage of dropouts was higher in the IFN groups than in the GA group for different reasons (adverse effects, lack of consent, tolerability). This aspect might have a role in the final assessment of effectiveness of these drugs, although sensitivity analysis using a likely scenario did not capture different results.

2. The definition of relapse was similar, with objective confirmation required, in all studies. However, Calabrese 2012 did not define the outcome and Mikol 2008 counted as relapses both qualifying and non-qualifying attacks. We could not exclude the possibility that an inhomogeneous relapse definition might constitute a bias.

3. Pivotal trials comparing DMTs versus placebo have suggested different efficacy for disease activity (Hillert 2012), although the results of this review show similar effects for IFN and GA. The reasons for this are uncertain, but the primary source of discrepancy is probably related to differences in the enrolled populations of studies, especially in the behaviour of placebo groups (Goodin 2008).

4. Data did not allow verification of the effects of IFNs and GA as change therapy in participants who were not responders to previous therapies. A recent comparative study evaluating DMTs in the real world clinical practice using international MSBase registry data has been conducted using propensity score matching. Slightly lower relapse incidence was found among patients

treated with glatiramer acetate or subcutaneous interferon beta-1a relative to intramuscular interferon beta-1a and interferon beta-1b. No differences in 12-month confirmed progression of disability were observed (Kalincik 2015) The ongoing RCTs might possibly contribute to this issue comparing the new approved therapies.

5. Analysis of the safety profile was restricted to the number of participants who withdrew from or dropped out of the study; drug-related adverse effects (tolerability) were not considered. Data derived from one study (reported but not planned) suggest a greater incidence of SAEs such as tumour or infection in IFN groups than in GA groups, although hepatobiliary and musculoskeletal disorders were reported more often among GA-treated participants. Other studies showed that tolerability was generally good. It is well known that up to 75% of IFN-treated patients experience flu-like illness, headache, myalgia. Increased alanine aminotransferase and leukopenia occur significantly more often in the IFN group, whereas pruritus, swelling, induration at the injection site (56% to 71%), dyspnoea and postinjection systemic reactions (15% to 38%) occur significantly more often in the GA group (Carter 2010; Parkenov 2013).

6. Significant heterogeneity of results on MRI assessment of new or enlarging T2 and brain volume changes was probably related to differences in the MRI protocol and measures of image acquisition protocol (Cadavid 2009; O'Connor 2009; Filippi 2011)

Quality of the evidence

Six RCTs were included, but only five contribute to the analyses involving large numbers of participants with similar disease characteristics. No heterogeneity of results was found when clinical end points were considered, supporting no difference in the protective effects of IFNs and GA on risk of relapse and progression.

The body of evidence was judged to be of moderate quality for most clinical and radiological outcomes, although the quality of evidence for the safety profile was found to be low.

Potential biases in the review process

The trials search strategy and contacts initiated with the main investigators suggest the likelihood that all relevant studies were identified and all relevant data were obtained.

Agreements and disagreements with other studies or reviews

This is the first systematic review comparing IFNs and GA by using head-to-head RCTs in participants with RRMS.

A Cochrane systematic review by Filippini (Filippini 2013) reported a meta-analysis of all available therapies in MS, including comparisons between active and placebo groups and head-to-head studies. Head-to-head trials on IFN and GA were also considered. However, in terms of risk of relapse and worsening at 24 months, the quality of the evidence was judged too low to allow meaningful comparisons, given the small sample size and the wide 95% confidence intervals. In this review, only two studies were considered (Mikol 2008; O'Connor 2009).

Furthermore, some published reviews have evaluated the efficacy and safety of DMTs and other approved therapies through direct or indirect comparison in participants with RRMS, but some have used different outcome measures and designs. Among these reviews, Qizilbash 2012 considered both RCTs and comparative observational cohort studies, focused on participants treated with GA for RRMS or CIS versus placebo, the comparators in a subanalysis were also IFNs and standard treatment. A total of 11 studies were included. Similar results were found in terms of relapse-free outcomes and safety. An 18% reduction in clinical progression (RR 0.82, 95% CI 0.68 to 0.98) was seen in GA-treated participants compared with those given IFNs, but only two studies were included (Mikol 2008; O'Connor 2009). Similar to our findings, study discontinuations due to adverse events did not differ among participants treated with GA and those given IFNs (RR 0.89, 95% CI 0.57 to 1.41). Hadjigeorgiou 2013 evaluated through direct and indirect comparisons the approved existing treatment for RRMS, but no clear conclusions were provided. Zagmutt 2013 conducted a network meta-analysis to assess the rates of adverse events and dropout from treatment for RRMS. Placebo-controlled RCTs evaluating immunosuppressive drugs and DMTs were included. The GA group had a significantly lower incidence rate of adverse events when compared with groups given other products. Dropouts were not significantly different across treatments, except for a higher incidence in the SC IFN-beta 1a 22 mcg and 44 mcg groups compared with the fingolimod group. Recent comparative study evaluating DMTs in the real world clinical practice using international MSBase registry data have been conducted using propensity score matching, slightly lower relapse incidence was found among patients treated with glatiramer acetate or subcutaneous interferon beta-1a relative to intramuscular interferon beta-1a and interferon beta-1b No differences in 12-month confirmed progression of disability were observed (Kalincik 2015).

A systematic non-Cohrane review of observational studies and in extension of randomized controlled trial (not included in our review) has been conducted to evaluate the long-term impact of IFNs or GA on disability progression in MS. The quantitative estimate of the treatment effect in reducing progression to EDSS 6 was HR pooled=0.49 (95% CI: 0.34-0.69), $p<0.001$, showing that treatment with immunomodulators seems to reduce long-term probability of disability progression (Signori 2016).

4.2 The quality of telephone triage

The results of the study showed that there were significant deficiencies in the telephone triage provided in the out of hours primary care service and, therefore, wide room for improvement. Indeed the study showed that the quality of telephone triage, in terms of mandatory questions asked during the call, was far below the desired standard, although the ability to assess the case, the clinical management decisions made and the treatment advised was almost always appropriate (it should be kept in mind, however, that the cases were not especially severe). The deficiencies found during triage might have been due to the call-handlers' inexperience or lack of training, but were probably the result of fatigue resulting from night shifts, repetitiveness and

the difficulty of dealing with the callers' emotions and expectations together with the problem of making a remote evaluation of the seriousness and urgency of the case.

In comparing the results of the previous study conducted in the Netherlands and those of our study, involving the same simulated clinical cases, we found 70% versus 36% of mandatory questions asked in the case of a child with vomiting, 47% versus 32% in the case of the child with fever, 36% versus 28% in the case of the adult with fever, and 54% compared to 27% of the adult with nosebleed, respectively. A service in which only one-third of the questions necessary to make a correct decision is actually asked represents a high medical risk and could become a medicolegal burden on patients and doctors.

Telephone triage is a delicate phase of a telephone consultation during which there is the risk that insufficient information is given by the caller resulting in an underestimation of the case, with potentially serious consequences for the patient, or an overestimation of the case with advice for inappropriate treatment that could also be economically disadvantageous. In other words, the quality of the telephone triage service and the risks that teleconsultations and telediagnoses entail are closely related.

The present study had several limitations: (1) the small size due to limited resources; (2) the use of clinical cases of mild severity, because the call-handler might have decided that a face-to-face consultation was required; (3) the complexity of the ISP method; (4) the lack of a comprehensive statistical analysis; (5) the limited generalizability; (6) the use of standardized clinical cases built by expert consensus, but not directly supported by evidence-based guidelines (because no evidence based clinical guideline provides questions to be administered in a telephone triage setting). Nonetheless, the study demonstrated that the model is feasible, and enabled us to assess to what extent and in what areas the quality of telephone triage concerning out-of-hours primary care requires improvement.

4.3 The efficacy of CDSS

This systematic review of 28 RCTs revealed little evidence for a difference in mortality when pooling results from comparisons of adoption of a CDSS integrated with an EHR versus health care settings without a CDSS. Our review indicates that differences in mortality outcomes, if they exist, appear small across studies and health care services, and may exist only in particular settings with specific diseases and circumstances. However, most of the studies were underpowered and too short to prove or exclude an effect on mortality, and effects as large as a 25% increase or reduction could still be possible. We found weak evidence that an active CDSS is associated with a lower risk for morbidity. All morbidity outcomes selected were relevant from a clinical and health services perspective.

Again, results on morbidity outcomes were very diverse, limiting quantitative inferences; however, the summary RR morbidity decrease of 10% to 18% places CDSS linked to EHRs at the top of the spectrum of quality improvement interventions for their potential impact on health outcomes.

The beneficial effects of CDSS might still be greater than that suggested by the current analysis given the limited number of actual studies providing results on hard outcomes.

Finally, we observed differences for costs and health service utilization, but these were often small in magnitude.

Several other systematic reviews provided pooled estimates of the RRs for CDSSs. All reviews observed large between-study heterogeneity.

This is expected given the variability in intervention, settings, diseases, and study designs. Despite this limitation, they concluded in favour of CDSS.

Our review exhibits several differences. We adopted stricter inclusion criteria, selecting only CDSS featuring a rule- or algorithm based software integrated with EHRs and evidence-based knowledge.

The CDSS we included can be viewed as a second generation in terms of their technology, information management and linkage to EHRs. Furthermore, we did not include process and laboratory outcomes such as adherence to guideline recommendations or change in blood values. Analyzing estimates from process outcomes is problematic. Their relevance is questionable and the quality of the data may have been less than optimal, particularly when the data sources were administrative rather than clinical. The overlap between our review and others is limited: approximately 50% in terms of the studies and less in terms of the rough data.

The results of our review complement previous analyses showing that CDSS are best oriented to directly affect process outcomes (recommendation adherence) and with decreasing impact morbidity and mortality.

Several included studies were cluster-RCTs that did not report if they accounted for clustering effects.

Trials randomizing at the group level should not be analysed at the individual participant level. If the clustering is ignored, P values will be artificially small: this problem might result in false positive conclusions that the CDSS has an effect when it does not.

Thus, we adjusted estimates of the RRs for our data synthesis using a method that inflates variances. However, such adjusted results should be interpreted cautiously; if the clustering effect is limited across studies, the analysis may be too conservative.

Our meta-analysis has additional limitations. We did not evaluate the quality of the evidence-based information supporting the CDSS recommendations. We accepted study authors description of a CDSS as evidence-based at face value, even if the authors did not explain the source of evidence or knowledge in detail. Furthermore, the limited number of trials, especially regarding the meta-analysis for the morbidity outcomes, increases the uncertainty of the findings and conclusions. The trials included were conceptually heterogeneous in terms of their design, setting, participants, and interventions, as well as the definition and measurement of outcomes.

In addition, although our literature search was as inclusive as possible without the exclusion of studies based on methodological characteristics, the search was restricted to studies published in indexed journals. We did not search for unpublished studies or for source data. Moreover, the trials included in this meta-analysis were not designed to specifically analyze the relationship between mortality and CDSS use. In fact, mortality was additional information provided often as a reason behind loss to follow-up. Additionally, the follow-up was too short to detect a sufficient number of deaths to show potentially relevant differences. Finally, we cannot exclude that

pooling the mortality outcome across different settings (e.g., intensive care units versus primary care) could have influenced the overall result toward a null effect with primary care studies bearing larger weight in the meta-analysis.

The results of this review may provide sufficient evidence to fuel the debate on the prospects of CDSS linked to EHRs. For those perceiving CDSS as an autocratic command to doctors, our systematic review may be interpreted as evidence that they do not affect patient mortality, on average and should be abandoned. For those interested in CDSS dissemination, our results, which show a decrease in morbidity across all settings by one fifth, may be used as an argument to increase CDSS adoption within health care services.

Both interpretations might be exaggerated as the evidence is still in its infancy along with the technology and implementation. Many of the trials adopted locally developed CDSS interventions, which may have compromised their level of integration into clinicians workflow. The next generation of CDSS trials should focus on systems with a more global outlook featuring authoritative point-of care services (Banzi 2010) and full integration with EHRs. The conclusion of a landmark article by Sim 2001 published almost 15 years ago, still reflects the current scenario: "*Although the promise of clinical decision support system-facilitated evidence-based medicine is strong, substantial work remains to be done to realize the potential benefits* (Sim 2001).

4.4 The efficacy of e-learning

Summary of main results

This systematic review of 14 randomised controlled studies revealed e-learning makes little or no difference when compared to traditional learning on patient outcomes or health professionals behaviours (low certainty of evidence); e-learning possibly leads to slightly less improvement in health professionals skills when compared to traditional learning (moderate certainty of evidence) even if important differences are unlikely. E-learning when compared to traditional learning may make little or no difference even on health professionals knowledge but our meta-analytic scenario was peculiar: a large trial at low risk of bias (Perkins 2012) detected a small difference in favour of traditional training, while smaller studies all reported no difference with the exception of one favouring e-learning (Paladino 2007). In such a case, it is not possible to completely exclude that the small advantage detected by the larger trial in favour of traditional learning is dragged into the non-significant difference area by the presence of other small studies. In this scenario high heterogeneity was expected but, we observed only moderate heterogeneity, which reinforces the confidence that we can have on our results. Moderate heterogeneity can be explained by the consistent small effects associated with the interventions. Another reason why we found moderate heterogeneity is that we limited the inclusion to randomised studies. The randomised design partially solved the problems associated with differences in case-mix, baseline skills or knowledge, and other setting characteristics associated with observational studies which can lead to spurius, significant results. Considering the predictive intervals for knowledge outcomes, the probability that new

studies will change our results is low, as we expect new studies to cluster around the no difference area.

Study-level variables such as change in content type, targeted health professionals, accreditation and format (interactivity or length) did not influence the relative efficacy of e-learning and traditional learning.

Overall completeness and applicability of evidence

The randomised controlled trials included in the review were sufficiently homogeneous in terms of included populations, comparison between e-learning versus traditional learning and outcome measures. With the exception of one study including child health consultants, all studies included doctors or nurses, the two health professionals groups that are possibly most represented in the world (WHO 2006). Twelve trials compared e-learning intervention with face-to-face learning as two trials evaluated e-learning against guideline dissemination or availability. These comparisons are representative of different options at stake when one educational approach has to be prioritised over another. Outcome measures were similar and, in most cases, a meta-analytic approach was deterred due to the unavailability of data rather than because of high variability of outcome measures. Statistical heterogeneity was never substantial, supporting the confidence we can have in the precision of effect sizes. However, it is possible that the included studies were different in the details of the settings or of the intervention.

Quality of the evidence

Our results indicate that differences in knowledge outcomes appear small. If any gain in knowledge exists, it might be easily lost with time; too little data are available to say how long it can last. The implication is that changes in patient outcomes or health professionals behaviours may be difficult to obtain or might not be attainable at all, at least when a training program is implemented for short time period (i.e. few months). In our review, the inconsistent effects of e-learning across knowledge outcomes may not be reflective of the intervention genuine ineffectiveness, but derive from the lack of compliance by health professionals or insufficient replication over time. Given the variability in individual knowledge level outcomes and high attrition, the sample size may have been increased to consider the scarce compliance of health professionals. The number of participants who withdrew from or dropped out of the studies was more than 30% in several studies. The loss to follow-up, when reported, could have introduced imbalances between the groups: incomplete outcome data was the most biased dimension in terms of the number of studies rated at high risk for this item. The overall quality of evidence was often rated as low because of some methodological limitations (imbalance at baseline and incomplete data) of primary studies and imprecision surrounding soft outcomes. The results of Perkins 2012, a large trial at low risk of bias, assessing skills during a simulated cardiac arrest, contribute in defining the role of e-learning in post-graduate education, but its research findings are open-ended. Only through repeated research can certainties emerge.

Potential biases in the review process

The trials search strategy suggests the likelihood that all relevant studies were identified and all relevant data were collected.

Agreements and disagreements with other studies or reviews

Previous systematic reviews have found e-learning to be associated with small positive effects when compared with non e-learning educational interventions, suggesting effectiveness similar or slightly better than traditional methods. In Cook 2008 is quantitative meta-analysis including 201 published studies on Internet-based learning (Cook 2008). This review considered three relevant outcomes: knowledge, skills, learner behaviours and patient outcomes. The first comparison focused on e-learning and no intervention; the second on e-learning and other types of educational activities (e.g. meetings or residential learning in class). In the first comparison, significant differences favouring e-learning were observed for all outcomes. These differences were also relevant in terms of the magnitude of the effect size. In the second comparison, the significance was formally maintained for knowledge, although the effect size was reduced. There was a direction effect for skills and patient outcomes but no difference was detected. Our review did not find a significant difference between e-learning and traditional leaning interventions with regards to patient outcomes and health professionals behaviours. On the contrary, although we were unable to pool the data on skills, we found studies favouring traditional training on skills and knowledge outcomes. This diversity of findings may be attributed to the type of studies included in Cook's reviews: while our review considered only randomised controlled trials and licensed health professionals, Cook additionally included non-randomised trials and undergraduate participants. Only two out of 76 studies included in Cook had the same PICO framework (Fordis 2005, Mäkinen 2006) of our review and were randomised trials and only 14% of participants were practicing health professionals (the other participants were different types of students). Although Cook and colleagues conducted a comprehensive and rigorous systematic review, neither methodological accuracy nor the broad inclusion criteria can overcome the weakness of the primary research included in their analysis.

A recent document from the US Department of Education reported the results of a review and meta-analysis of online learning studies on undergraduate students, finding, on average, that students in online learning environments performed modestly better than those receiving face-to-face instructions. The comparison with our review suggests that different populations may learn with different effectiveness using the same online tools; this may further pertain to different generations within the same type of health professionals, in which case, future studies should evaluate whether or not younger health professionals may benefit more from new online tools compared to previous generations. This phenomenon is well known in social sciences research where it is known as a "cohort effect", defined as "the effect that having been born in a certain time, region, period or having experienced the same life experience (in the same time period) has on the development of a particular group" (Glen 2005).

4.5 Training intervention for improving telephone skills

Summary of main results

We did not identify any studies that assessed the effect of training intervention for clinicians on patient primary outcomes (health outcomes measured by validated tools or biomedical markers or patient behaviours; patient morbidity or mortality; patient satisfaction; urgency assessment accuracy; adverse events). We identified one controlled before-after study evaluating the effect of a training intervention on clinicians telephone consulting skills by a validated tool. The study found no difference between intervention and control on history taking and case management skills but no quantitative data were provided in the article and the study authors could not supply additional data. We rated this study as being at high risk of bias. Based on our GRADE assessment of the certainty evidence it is uncertain whether the intervention improves clinicians history taking and management skills (very low certainty).

Overall completeness and applicability of evidence

Several studies in the telephone medicine literature were excluded because they did not have an experimental design and two studies were excluded, despite fulfilling all other inclusion criteria, because they used non-previously validated evaluation tools: no reliable evidence was available for this research topic.

Quality of the evidence

The overall quality/certainty of the evidence was judged to be very low because the initial level of confidence about the only included study was low (non randomised evidence) and we downgraded the confidence because the risk of bias was rated as serious (incomplete outcome data management and selective outcome reporting) and because of the impossibility to assess consistency of effect, imprecision, indirectness and publication bias due to lack of other studies.

Potential biases in the review process

This review did not formally explore publication bias because only one study was included.

Agreements and disagreements with other studies or reviews

To our knowledge this is the only systematic review on this topic.

CONCLUSIONS

5.1 Efficacy and safety of INFs-beta vs GA in RRMS

Implications for practice

The results of this review fit into current clinical practice. IFNs and GA have played historic roles in transforming MS into a treatable disease and remain the cornerstone of treatment for RRMS as first-line drugs or as switching therapies, even after 15 years. IFNs and GA are generally perceived as very safe drugs, in contrast with the awareness of uncommon but severe AEs associated with the new available MS treatments (Hillert 2012; Hutchinson 2012 Dubey 2016).

The results of this review support clinicians in using either of these therapies for RRMS, because of their similar safety and efficacy in the prevention of disease activity.

However, some differences in clinical and MRI measures should be underlined.

Evidence from a single study suggests better efficacy of GA over IFN in terms of relapse rate at three years of follow-up. In addition, greater and faster reduction in MRI lesion load accrual was observed in IFN-treated compared with GA-treated participants with MS. These results need to be considered with caution because they are derived from a limited number of studies and participants, and/or because significant heterogeneity of the results is apparent.

Furthermore, the clinical impact of these data is uncertain.

We were not able to draw conclusions on quality of life and tolerability of these two drugs, as these parameters were not fully assessed in this review.

Implications for research

Two primary considerations should be applied to future research efforts.

Authors of RCTs should be required to include a clear data presentation, such as reasons for loss to follow-up; this may allow more accurate interpretation and less problematic comparisons of trial results. This is true also for MRI outcome measures that need comparable protocols for image acquisition.

Comparison of IFN and GA using patient-related outcomes such as fatigue and other specific MS-related quality of life measures is needed, as is comparison of the effects of DMTs on cognitive function.

Researchers should be required to apply standardised sets of outcomes to produce manageable data that could be further merged in meta-analysis.

5.2 The quality of telephone triage

Implications for practice

The results of the present study show that the quality of telephone triage in the OOH centres studied was low and that there was room for improvement. The doctors answering the telephone asked a very low proportion of obligatory questions and this could affect patient

safety and risk management for the service. Doctor on call should consider to be very cautious in the telephone decision making because despite the attention they are paying, they are probably asking less questions their should.

Implications for research

Several suggestions for future research emerged from the study. We chose not to call at particularly inconvenient times for the out-of-hours staff (e.g. late night and shift change). However, these times could be particularly important, as these are times when mistakes might be more likely to occur. Another question that emerges from the study regards the efficacy of the strategies that could be adopted in order to improve the quality of telephone triage. These could include the effect of specific training, in particular training on management techniques and structuring telephone consultations and the effectiveness, efficiency and security that computerized support could offer the telephone triage doctors in decision making.

Previous studies have shown that decision-making tools are safe and can reduce the clinical risks associated with unassisted telephone triage call handlers, as well as being cost-effective and possibly reducing the number of unnecessary visits to emergency room services.

5.3 The efficacy of CDSS

Our results on health care services equipped with versus health care not equipped with CDSS suggest, in broad terms, that this technology does not result in substantial benefits or risks for patients in terms of mortality.

This effect, when it occurs, is largely dependent on the disease and setting characteristics. Focusing on subgroup analyses, however, can lead to misleading claims when the overall data are limited and unavoidably weak because of inherent design problems. Effects on morbidity might exist and the magnitude of the effect, in the order of 10% to 20%, could be large enough to impact mortality if appropriate follow-up is ensured. The results of this study may provide enough evidence to advance the debate on the prospects of CDSS.

5.4 The efficacy of e-learning

Implications for practice

Our results suggest in broad terms that e-learning does not itself result in major benefits for patient and health professionals outcomes. For members of a university, primary or secondary care institution or royal college making the decision between the offering of traditional or e-learning courses, the judgement is complex, based not only on the relative efficacy of the methods, but on alternative dimensions such as accessibility and usability by participants, higher compliance or satisfaction, cost or economic feasibility, as well as its alignment with an existing portfolio of learning activities. Medical education providers should carefully consider the purposes and the specific setting of their programmes, favouring, for instance, traditional learning to improve knowledge or skills in small groups of health professionals when meeting

attendance is sustainable, or e-learning programmes when the aim is to reach a large number of health professionals with limited costs. Blended courses potentially balance the benefits of the two learning strategies. Assumptions about the potential superiority of one strategy over the other within the context of a specific dimension should be explored. One may argue that differences in skills or knowledge outcomes with e-learning versus traditional learning are expected to exist for particular diseases and settings. Focusing on subgroups, however, can lead to misleading claims when the overall data suggest similar effectiveness of interventions.

The results of this review fit into the current health professionals continuing medical education provider's policy with special regards to CME and its implications towards government agencies, scientific societies, healthcare organisations, private companies. Although our results do not completely support the effectiveness of e-learning for skills and procedural outcomes, the results do not override some benefits of e-learning with regards to accessibility and flexibility. There is insufficient evidence to provide recommendations about accreditation, interactivity, length of e-learning programmes or about targeting the courses on specific types of participants or contents. We have limited understanding of the characteristics of the targeted knowledge, professionals and settings that may influence the effectiveness of different e-learning programs. Thus, for those working in the e-learning setting, the findings from this meta-analysis provide limited information to guide the choice or optimise the components of such complex educational interventions. The effectiveness of e-learning is likely to be modified by characteristics such as the attitudes of the health professionals or their perceived ability to transform passive information into tangible actions.

Implications for research

Future research is unlikely to change the main results of our review. The probability of finding a large effect in a future study either favouring e-learning or traditional learning is likely to be small, unless the learning patterns of health professionals change in the long term as a result of subsequent generations greater exposure and familiarity with technology (e.g. millennials).

Next studies should privilege a randomised design with an appropriate sample size. They should additionally privilege patient outcomes and health professionals' behaviours rather than skills or knowledge and should focus on which components of e-learning could eventually change not only knowledge and skills but also change at least behaviour,

Multiple outcomes measure points in the time should be planned in order to detect effects persistence and attention should be given to studies duration and follow-up. Considering the difference between e-learning and traditional learning is likely to be small, cost-effectiveness becomes a relevant outcome for future studies. All studies, irrespective of the outcomes considered, should use predefined data scales and reporting rules in order to improve their reporting.

More data are needed to evaluate the relative efficacy of e-learning in specific medical areas or rare conditions (i.e. e-learning programs assisting in surgical teaching) and the importance of accreditation, interactivity, and length of e-learning programmes.

The feasibility of these studies is challenged by the large number of participants that need to be enrolled and by longer follow-up but existing educational settings providing training interventions could be taken into account to override this problem.

5.5 Training intervention for improving telephone skills

Implications for practice

The lack of studies assessing the effect of interventions aiming to improve clinicians' telephone skills, the limited size and the low quality of the only study found, do not allow robust evidence-based conclusions. Essentially, this review cannot provide any guidance on effective methods to train health care professionals in telephone skills. Given the established use of telephone consultations within the medical field, this apparent lack of evidence seems surprising. We have described in the background section some of the important roles that telephone communication has within healthcare: this lack of knowledge is even more severe because telephone consultation is nowadays an important means of initial assessment of clinical cases and management of everyday practice in all clinical specialities.

Our review seems to describe a paradox: a rapidly increasing use of the telephone as a method of medical consultations, a patients' compliance to recommendations influenced by the quality of clinicians' communication and, as a counterpart, the widespread low quality of clinicians communication skills, the lack of specific training during the undergraduate and postgraduate education and the complete absence of specific evidence needed to inform this training.

At the moment the training of clinicians on telephone consultation has to be guided by studies and models set on face to face communication that do not consider the differences between these two communicative dimensions.

The very limited evidence that this review has identified would suggest that telephone consultation should be employed cautiously and carefully in healthcare provision until proven methods are developed around the training and practice of remote consultation skills.

Implications for research

High quality randomised controlled trials assessing by validated tools the effect of training interventions on clinicians telephone skills on both patient-oriented and clinician-orientated outcomes should be undertaken to ensure telephone skills can be taught based on reliable evidence. In a similar way to face-to-face consultation skills, we need to ensure telephone consultation skills are taught in accordance with rigorous and robust methods. A substantial evidence base is lacking and future studies could focus on a variety of different aspects including researching the type of teaching methods to be employed, developing telephone consultation models and validated assessment tools.

FINAL STATEMENT

Despite the phone is a very frequently used tool, the clinical quality of the MTC seems low: doctors answering the phone ask all the obligatory questions in a low proportion of the clinical cases they manage. This raises concerns on the service patient safety and about the potential legal consequences for the doctors and the LHT.

We showed CDSS could be effective tools and they can help doctors, maybe even to manage the risk of remote consulting in the 100% of the calls, but unfortunately we cannot test the efficacy and safety of a specific CDSS for medical telephone triage.

Future research should focus not only on clinical side of medical telephone consultation improvement but also on relational side, because patient compliance depends on this.

No evidence is available so far to address clinicians education towards telephone skills improvement but, once evidence on how to improve is available, e-learning will be a potential cost-effective training way for CDSS uptake and telephone skills education: it seems as effective as traditional training in terms of transmission of knowledge and, even if potentially little less effective on skills transmission, its advantages could make it the best way to quickly fil the gap.

REFERENCES

NOTE: given the high number of citations, the references section is divided in several parts:

1. General References
2. References for the review “INFs vs GA for RRMS Interferons-beta versus glatiramer acetate for relapsing-remitting multiple sclerosis”
3. References for the review “Effectiveness of Computerized Decision Support Systems Linked to Electronic Health Records: A Systematic Review and Meta-Analysis
4. References for the review “e-Learning for licensed health professionals”
5. References for the review “Training interventions for improving telephone consultation skills in clinicians”

General References

Alderson P, Green S. Issues related to the unit of analysis. Cochrane Collaboration Open Learning Material 2002.

Anonymous. Notes, Short Comments, and Answers to Correspondents. Lancet 1879;114(2935):819

Ash JS, McCormack JL, Sittig DF, Wright A, McMullen C, Bates DW. Standard practices for computerized clinical decision support in community hospitals: a national survey. J Am Med Inform Assoc. 2012;19(6):980-987.

Banzi R, Moschetti I, Moja L. Internet-based education for health professionals. JAMA 2009;301(6):599; author reply 599-600.

Banzi R, Liberati A, Moschetti I, Tagliabue L, Moja L. A review of online evidence-based practice point-of-care information summary providers. J Med Internet Res. 2010;12(3):e26.

Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics. 1994; 50:1088-1101

Bergman AB, Dassel SW, Wedgwood RJ. Time-motion study of practicing paediatricians. Pediatrics 1966;38(2):254-263.

Bonnabry P, Despont-Gros C, Grauser D, Casez P, Despond M, Pugin D. A risk analysis method to evaluate the impact of a computerized provider order entry system on patient safety. J Am Med Inform Assoc. 2008;15(4):453---460.

Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to Meta-Analysis. First edition. Chichester, West Sussex, UK: John Wiley & Sons Ltd., April 27, 2009.

Bright TJ, Wong A, Dhurjati R, Bristow E, Bastian L, Coeytaux RR. Effect of clinical decision-support systems: a systematic review. Ann Intern Med. 2012;157(1):29---43.

British Medical Association. BMA, General Practitioners Committee. London: British Medical Association, 2001. [Other: http://www.bma.org.uk/images/consultingmarch2001_tcm41-19941.pdf]

Brown SB, Eberle BJ. Use of the telephone by pediatric house staff: a technique for pediatric care not taught. J Pediatr. 1974 Jan;84(1):117-9.

Buja A, Toffanin R, Rigon S, Sandonà P, Carraro D, Damiani G. Out-of-hours primary care services: demands and patient referral patterns in a Veneto region (Italy) Local Health Authority. Health Policy. 2015 Apr;119(4):437-46

Bunn F, Byrne G, Kendall S. Telephone consultation and triage: effects on health care use and patient satisfaction. Cochrane Database of Systematic Reviews 2004, Issue 4. Art. No.: CD004180 DOI: 10.1002/14651858.CD004180.pub2

Campbell JL, Fletcher E, Britten N, Green C, Holt TA, Lattimer V et al. Telephone triage for management of same day consultation requests in general practice (the ESTEEMtrial): a cluster randomised controlled trial and cost consequence analysis. Lancet 2014;384(9957):1859-688.

Car J, Freeman GK, Partridge MR, Sheikh A. Improving quality and safety of telephone based delivery of care: teaching telephone consultation skills. Quality and Safety in Health Care 2004 Feb;13:2-3.

Car J, Koshy E, Derek B et al. Telephone triage in out of hours call centres. BMJ 2008;337:a1167.

Car J, Sheikh A. Telephone consultations. BMJ 2003 May 3;326:966-9.

Carter NJ, Keating GM. Glatiramer acetate: a review of its use in relapsing-remitting multiple sclerosis and in delaying the onset of clinically definite multiple sclerosis. Drugs 2010;70(12):1545-77.

Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomised trials: comparison of protocols to published articles. JAMA 2004;291:2457-2465.

Charmaz K: Constructing Grounded Theory: A Practical Guide through Qualitative Analysis. Thousand Oaks, CA: SAGE Publications Ltd.; 2006.

Chaudhry B, Wang J, Wu S, Maglione M, Mojica W, Roth E et al. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. Ann Intern Med. 2006;144(10):742---752.

Chumley-Jones HS, Dobbie A, Alford CL. Web-based learning: sound educational method or hype? A review of the evaluation literature. Academic Medicine 2002;77(10) Suppl:S86-S93.

Clark RA, Inglis SC, McAlister FA, Cleland JGF, Stewart S. Telemonitoring or structured telephone support programs for patients with chronic heart failure: systematic review and meta-analysis. BMJ 2007 April 10;334(942):1-9.

Cochran WG. The combination of estimates from different experiments. Biometrics. 1954;10:101---129.

Colombo A, Bert F, Camussi E, Gualano MR, Arensi D, Castaldo D et al. Analysis of users and reasons for use of a physician-on-call service in an Italian local health unit. Fam Pract. 2016 Aug 24. pii: cmw089

Compston A, Coles A. Multiple sclerosis. Lancet 2008;372(9648):1502-17.

Cook DA, Garside S, Levinson AJ, Dupras DM, Montori VM. What do we mean by web-based learning? A systematic review of the variability of interventions. Medical education 2010;44(8):765-74.

Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Instructional design variations in Internet-based learning for health professions education: a systematic review and meta-analysis. Academic medicine 2010;85:909-22

Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM. Internet-based learning in the health professions: a meta-analysis. JAMA 2008;300(10):1181-96.

Cook DA, Levinson AJ, Garside S. Time and learning efficiency in Internet-based learning: a systematic review and meta-analysis. Advances in Health Sciences Education Theory and Practice 2010;15(5):755-70.

Cook DA. Web-based learning: pros, cons and controversies. Clinical Medicine 2007;7(1):37-42.

Coppus SF, Emparanza JI, Hadley J, Kulier R, Weinbrenner S, Arvanitis TN, et al. A clinically integrated curriculum in evidence-based medicine for just-in-time learning through on-the-job training: the EU-EBM project. BMC Medical Education 2007;7:46.

Dale J, Crouch R, Lloyd D. Computer assisted assessment and advice for 'non-serious' 999 ambulance service callers: the potential impact on ambulance despatch. Emergency Medicine Journal 2001; 20:178-183

de Lusignan S, Chan T. The development of primary care information technology in the United Kingdom. J Ambul Care Manage. 2008;31(3):201---210.

Deeks JJ, Higgins JPT, Altman DG (editors). Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available at www.cochrane-handbook.org.

Derkx H, Rethans JJ, Muijtjens A, Maiburg B, Winkens R, van Rooij H, Knottnerus A. 'Quod scripsi, scripsi.' The quality of the report of telephone consultations at Dutch out-of-hours centres. Qual Saf Health Care. 2010 Dec;19(6):e1

Derkx HP, Rethans JJ, Knottnerus JA, Ram PM. Assessing communication skills of clinical call handlers working at an out-of-hours centre: development of the RICE rating scale. Br J Gen Pract. 2007 May;57(538):383-7.

Derkx HP, Rethans JJ, Maiburg BH, Winkens RA, Muijtjens AM, van Rooij HG, Knottnerus JA. Quality of communication during telephone triage at Dutch out-of-hours centres. Patient Educ Couns. 2009 Feb;74(2):174-8.

Derkx HP, Rethans JJ, Muijtjens AM, Maiburg BH, Winkens R, van Rooij HG, Knottnerus JA. Quality of clinical aspects of call handling at Dutch out of hours centres: cross sectional national study. BMJ. 2008 Sep 12;337:a1264.

DerSimonian R, Laird N. Meta-analysis in clinical trials. Controlled Clinical Trials 1986;7(3):177-88.

Dicenso A, Bayley L, Haynes RB.Accessing pre-appraised evidence: finetuning the 5S model into a 6S model. Evid Based Nurs. 2009;12(4):99-101.

Dubey D, Cano CA, Stüve O. Update on monitoring and adverse effects of approved second-generation disease-modifying therapies in relapsing forms of multiple sclerosis. Current Opinion in Neurology 2016;29(3):278-85.

Dunt DW. Impact of telephone triage on emergency after hours GP Medicare usage: A time-series analysis. Australian and New Zealand Health Policy. 2001;4:21

Dwan K, Altman DG, Arnaiz JA, Bloom J, Chan AW, Cronin E et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias.. PLoS One 2008;3::e3081.

Eberhardt J, Bilchik A, Stojadinovic A. Clinical decision support systems: potential with pitfalls. J Surg Oncol. 2012;105(5):502---510.

Effective Practice and Organisation of Care (EPOC). Data extraction and management. EPOC Resources for review authors. Oslo: Norwegian Knowledge Centre for the Health Services 2015;Available at: http://epoc.cochrane.org/epoc-specific-resources-review-authors.

Effective Practice and Organisation of Care (EPOC). What study designs should be included in an EPOC review and what should they be called? EPOC Resources for review authors. Oslo: Norwegian Knowledge Centre for the Health Services; 2013. Available at: http://epocoslo.cochrane.org/epoc-specific-resources-review-authors.

Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ.1997;315:629---634.

Elbourne DR, Altman DA, Higgins JP, Curtin F, Worthington HV, Vail A. Meta-analyses involving cross-over trials: methodological issues. International Journal of Epidemiology 2002;31(1):140-9.

Evans R, Edwards A, Elwyn G. The future for primary care: increased choice for patients. Quality and Safety in Health Care 2003;12:83-4.

FDA. Safety - What is a Serious Adverse Event? http://www.fda.gov/safety/medwatch/howtoreport/ucm053087.htm (accessed November 2013)

Filippini G, Del Giovane C, Vacchi L, D'Amico R, Di Pietrantonj C, Beecher D, et al. Immunomodulators and immunosuppressants for multiple sclerosis: a network meta-analysis. Cochrane Database of Systematic Reviews 2013, Issue 6. Art. No.: CD008933. DOI: 10.1002/14651858.CD00893

Fisniku LK, Chard DT, Jackson JS, Anderson VM, Altmann DR, Miszkiel KA, et al. Gray matter atrophy is related to long-term disability in multiple sclerosis. Annals of Neurology 2008;64(3):247-54.

Forrest CB. Primary care in the United States: primary care gatekeeping and referrals: effective filter or failed experiment? BMJ. 2003;326:692-5.

Fry MM. Barriers and facilitators or successful after hours care model implementation: reducing ED utilisation. Australasian Emergency Nursing Journal. 2009;12:137-144

Gibbons A, Fairweather P. Computer-based instruction. In: Tobias S, Fletcher J (eds).Training & Retraining: A Handbook for Business, Industry, Government, and the Military. New York: Macmillan Reference USA, 2000.

Giesen P, Smits M, Huibers L, Grol R, Wensing M. Quality of after-hours primary care in the Netherlands: a narrative review. Ann Intern Med 2011;155(2):108-113.

Glaser BG, Holton J. Remodeling Grounded Theory. In Forum Qualitative Sozialforschung/Forum: Qualitative Social Research; Vol 5, No 2 (2004): Qualitative Market, Media and Opinion Research

Goodin D. Comparative studies of glatiramer acetate and interferon beta. International Multiple Sclerosis Journal 2008;15(2):39-41.

Goodin DS. Disease-modifying therapy in multiple sclerosis. Update and clinical implications. Neurology 2008;71 Suppl 3:8–13

GRADE Working Group. Grading quality of evidence and strength of recommendations. British Medical Journal 2004;328:1490-4.

GRADEpro [Computer program]. Brozek J, Oxman A, Schunemann H. Version 3.2 for Windows. GRADE Working Group, 2008

Grimshaw JM, Thomas RE, MacLennan G, Fraser C,Ramsay CR, Vale L, et al.. Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technology Assessment 2004;8:iii-iv:1-72.4

Gruffydd-Jones K, Hollinghurst S, Ward S, Taylor G. Targeted routine asthma care in general practice using telephone triage. British Journal of General Practice 2005 Dec;55:918-23.

Gunter TD, Terry NP. The emergence of national electronic health record architectures in the United States and Australia: models, costs, and questions. J Med Internet Res. 2005;7(1):e3

Hadjigeorgiou G, Doxani C, Miligkos M, Ziakas P, Bakalos G, Papadimitriou D, et al. A network meta-analysis of randomized controlled trials for comparing the effectiveness and safety profile of treatments with marketing authorization for relapsing multiple sclerosis. Journal of Clinical Pharmacy and Therapeutics 2013;38(6):43

Hallam L. Patient access to general practitioners by telephone: the doctor's view. British Journal of General Practice 1992 May;42:186-9.

Harbord RM, Egger M and Sterne JA. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. Statistics in medicine 2006;25:344-57.

Haynes RB, Wilczynski NL. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: methods of a decisionmaker-researcher partnership systematic review. Implement Sci. 2010;5:12.

Health Services Research Unit. Database of ICCs: Spreadsheet (empirical estimates of ICCs from changing professional practice studies). Available at:http://www.abdn.ac.uk/hsru/research/delivery/behaviour/methodologicalresearch.Accessed July 1, 2014.

Helfer B, Prosser A, Samara MT, Geddes JR, Cipriani A, Davis JM, et al. Recent meta-analyses neglect previous systematic reviews and meta-analyses about the same topic: a systematic examination. BMC medicine 2015;13: 82.

Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. www.cochrane-handbook.org.

Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327:557---560

Hillert J. In the coming year we should abandon interferons and glatiramer acetate as first line therapy for MS: No. Multiple Sclerosis (Houndmills, Basingstoke, England) 2012;19(1):26-8.

Huibers L, Smits M, Renaud V, Giesen P, Wensing M.Safety of telephone triage in out-of-hours care: a systematic review. Scand J Prim Health Care. 2011 Dec;29(4):198-209

Hutchinson M. In the coming year we should abandon interferons and glatiramer acetate as first line therapy for MS: commentary. Multiple Sclerosis (Houndmills, Basingstoke, England) 2012;19(1):29-30.

Italian National Health Service 2012 Data Report www.salute.gov.it/imgs/C_17_pubblicazioni_2370_allegato.pdf. Last Visit 01 Nov 2016

Jamtvedt G, Young JM, Kristoffersen DT, O'Brien MA,Oxman AD. Does telling people what they have been doing change what they do? A systematic review of the effects of audit and feedback. Quality and Safety in Health Care 2006;15:433-6.

Jorgensen AW, Hilden J, Gotzsche PC. Cochrane reviews compared with industry supported meta-analyses and other meta-analyses of the same drugs: systematic review. BMJ 2006;333: 782.

Katz HP, Kaltsounis D, Halloran L, Mondor M. Patient Safety and Telephone Medicine. Some Lessons from Closed Claim Case Review. Journal of General Internal Medicine 2008 Jan 29;23:517-22.

Kaushal R, Shojania KG, Bates DW. Effects of computerized physician order entry and clinical decision support systems on medication safety: a systematic review. Arch InternMed. 2003;163(12):1409---1416.

Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews.. BMJ 2010;340:c365.

Kirkpatrick DL. Revisiting Kirkpatrick's four-level-model. Training & Development 1996;1:54-7

Kurtzke JF. Rating neurological impairment in multiple sclerosis: an Expanded Disability Status Scale (EDSS). Neurology 1983;33(11):1444-52.

Lam-Antoniades M, Ratnapalan S, Tait G. Electronic continuing education in the health professions: an update on evidence from RCTs. The Journal of Continuing Education in the Health Professions 2009;29(1):44-51.

Lau J, Ioannidis JP, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. BMJ 2006;333(7568):597-600.

Lefebvre C, Manheimer E, Glanville J. Chapter 6: Searching for studies. In: Higgins JPT, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Li LC, Moja L, Romero A, Sayre EC, Grimshaw JM. Nonrandomized quality improvement intervention trials might overstate the strength of causal inference of their findings. Journal of Clinical Epidemiology 2009;62(9):959-66.

Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Medicine 2009;6:e1000100.

Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst. 1959;22:719---748.

Moja L, Pecoraro V, Ciccolallo L, Dall'Olmo L, Virgili G, Garattini S. Flaws in animal studies exploring statins and impact on meta-analysis. Eur J Clin Invest 2014;44(6):597-612.

McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the Diagnosis of Multiple Sclerosis. Annals of Neurology 2001;50(1):121-7.

McKenzie JE, Salanti G, Lewis SC, Altman DG (2013) Meta-analysis and The Cochrane Collaboration: 20 years of the Cochrane Statistical Methods Group. Systematic reviews 2: 80. doi: 10.1186/2046-4053-2-80

Meer A, Gwerder T, Duembergen L et al. Is a computer-assisted telephone triage safe? A prospective surveillance study in walk-in patients with non-life threatening medical condition. Emerg Med J 2012; 29:124-128

Mendenhall R. Medical practice in the United States. Princeton, New Jersey: Robert Wood Johnson Foundation, 1981

Miller GE. The assessment of clinical skills/competence/performance. Academic Medicine 1990;65:S63-67.

Miriovsky BJ, Shulman LN, Abernethy AP. Importance of health information technology, electronic health records, and continuously aggregating data to comparative effectiveness research and learning health care. J Clin Oncol. 2012;30(34):4243-4248.

Moja L, Banzi R. Navigators for medicine: evolution of online point-ofcare evidence-based services. Int J Clin Pract. 2011;65(1):6-11.

Moja L, Moschetti I, Cinquini M, Sala V, Compagnoni A, Duca P, et al. Clinical evidence continuous medical education: a randomised educational trial of an open access e-learning program for transferring evidence-based information - ICEKUBE (Italian Clinical Evidence Knowledge Utilization Behaviour Evaluation) - study protocol. Implement Sci. 2008;3(37):1-11

Moja L, Moschetti I, Liberati A, Manfrini R, Deligant C, Satolli R, et al. Using Clinical Evidence in a national continuing medical education program in Italy. PLos Medicine 2007;4(5):e113

Moja L, Pecoraro V, Ciccolallo L, Dall'Olmo L, Virgili G, Garattini S. Flaws in animal studies exploring statins and impact on meta-analysis. Eur J Clin Invest 2014;44(6):597-612.

Moriarty H, McLeod D, Dowell A. Mystery shopping in health service evaluation. Br J Gen Pract 2003;53:942–6.

Morse JM, Stern PN, Corbin J, Bowers B, Clarke AE, Charmaz K: Developing Grounded Theory: The Second Generation. 1st edition. Walnut Creek, CA: Left Coast Press; 2009.

Moseley AM, Elkins MR, Herbert RD, Maher CG, Sherrington C. Cochrane reviews used more rigorous methods than non-Cochrane reviews: survey of systematic reviews in physiotherapy. Journal of clinical epidemiology 2009;62: 1021–1030

National Audit Office 2014. Out-of-hours GP services in England. https://www.nao.org.uk/wp-content/uploads/2014/09/Out-of-hours-GP-services-in-England1.pdf. Last visit 01 Nov 2016

O'Connor P, Arnason B, Comi G, Filippi M, Cook S, Goodin D, et al. Interferon beta-1b 500 mcg, interferon beta-1b 250 mcg and glatiramer acetate: primary outcomes of the Betaferon® efficacy yielding outcomes of a new dose study. Neurology 2008;71:153.

Parkenov V, Schluep M, Du Pasquier R. Assessing risk of multiple sclerosis therapies. Journal of Neurosurgical Sciences 2013;332(1-2):59-65.

Patel A, Dale J, Crouch R. Satisfaction with telephone advice from an accident and emergency department: identifying areas for service improvement. Quality in Health Care 1997;6:140-5.

Patel H, Patel M, Car J. Telephone consultations in general practice: Areas for improvement. Journal of Telemedicine and Telecare 2005;11:265-6.

Patel R, Gray W, Saltoun C, Grammer L. Improving asthma care for the elderly: a randomised, controlled trial using telephone intervention conducted by non-medically trained personnel. Journal of Asthma 2009 Feb 1;46:30-5

Patton MQ: Qualitative Research & Evaluation Methods. 3rd edition. Thousand Oaks, CA: SAGE Publications, Inc; 2002.

Petitti DB. Statistical methods in meta-analysis. In: Petitti DB, ed. Meta-Analysis, Decision Analysis, and Cost-Effectiveness Analysis. New York: Oxford University Press; 1999.

Pimentel LE, Yennurajalingam S, Chisholm G, Edwards T, Guerra-Sanchez M, De La Cruz M et al. The frequency and factors associated with the use of a dedicated Supportive Care Center Telephone Triaging Program in patients with advanced cancer at a comprehensive cancer centre.. J Pain Symptom Manage 2015;49(5):939-944.

Pinnock H, Bawden R, Proctor S, Wolfe S, Scullion J, Price D. Accessibility, acceptability, and effectiveness in primary care of routine telephone review of asthma: pragmatic, randomised controlled trial. BMJ 2003;326:477-9.

Pleimes D, Pohl C, Beckmann K, Stolz C. BEYOND Study – Data for Cochrane Analyses Protocol No. 306440. Bayer HealthCare Pharmaceuticals Affairs Specialized Therapeutics 2013.

Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Annals of Neurology 2011;69(2):292-302.

Polman CH, Reingold SC, Edan G, Filippi M, Hartung HP, Kappos L, et al. Diagnostic criteria for multiple sclerosis: 2005 revision to the "McDonald Criteria". Annals of Neurology 2005;58(6):840-6.

Poon WB, Tagamolila V, Toh YP, Cheng ZR. Integrated approach to e-learning enhanced both subjective and objective knowledge of aEEG in a neonatal intensive care unit. Singapore Med J 2015;56(3):150-156.

Popescu V, Agosta F, Hulst HE, Sluimer IC, Knol DL, Sormani MP, et al. Brain atrophy and lesion load predict long term disability in multiple sclerosis. Journal of Neurology, Neurosurgery and Psychiatry 2013;84(10):1082-91.

Poser CM, Paty DW, Scheinberg L, McDonald WI, Davis FA, Ebers GC, et al. New diagnostic criteria for multiple sclerosis:guidelines for research protocols. Annals of Neurology 1983;13(3):227-31.

Purc-Stephenson RJ, Thrasher C. Patient compliance with telephone triage recommendations: a meta-analytic review. Patient Educ Couns 2012;87(2):135-142.

QSR International Pty Ltd: NVivo Qualitative Data Analysis Software. 10th edition. 2012.

Qizilbash N, Mendez I, Sanchez-de la Rosa R. Benefit-risk analysis of glatiramer acetate for relapsing-remitting and clinically isolated syndrome. Multiple Sclerosis Clinical Therapy 2012;34(1):159-76

R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2013. Available at: http://www.R-project.org. Accessed July 1, 2014.

Raftery J. Nurse telephone consultation for out of hours primary care can save money through reduced ER admission, surgery attendance and GP home visits. Evidence-based Health care 2000; 4:61-161

Reisman AB, Brown KE. Preventing communication errors in telephone medicine. Journal of General Internal Medicine 2005;20(10):959-963

Rethans JJ, Gorter S, Bokken L, Morrison L. Unannounced standardised patients in real practice: a systematic literature review. Med Educ 2007;41:537–49.

Rethans JJ, Norcini JJ, Baron-Maldonado M, Blackmore D, Jolly BC, LaDuca T, et al. The relationship between competence and performance: implications for assessing practice performance. Medical Education 2002;36:901-9.

Review Manager (RevMan) [Computer program]. Version 5.2. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2013.

Review Manager (RevMan) [Computer program]. Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Riegel B, Carlson B, Kopp Z, LePetri B, Glaser D, Unger A. Effect of a standardized nurse case-management telephone intervention on resource use in patients with chronic heart failure. Archives of Internal Medicine 2002 Mar 25;162:705-12.

Romano MJ, Stafford RS. Electronic health records and clinical decision support systems: impact on national ambulatory care quality. Arch Intern Med.2011;171(10):897---903.

Rosenbaum SE, Glenton C, Cracknell J. User experiences of evidence-based online resources for health professionals: user testing of The Cochrane Library. Bmc Medical Informatics and Decision Making 2008;8:34

Roshanov PS, Fernandes N, Wilczynski JM, Hemens BJ, You JJ, Handler SM et al. Features of effective computerised clinical decision support systems: meta-regression of 162 randomised trials. BMJ. 2013;346:f657

Rudick RA, Lee LC, Cutter GR, Miller DM, Bourdette D, Weinstock-Guttman B, et al. Disability progression in a clinical trial of relapsing-remitting multiple sclerosis: eight year follow-up. Archives of Neurology 2010 12 Jul

Ruiz JG, Mintzer MJ, Leipzing RM. The impact of e-learning in medical education. Academic Medicine 2006;81(3):207-12.

Ruiz JG, Teasdale TA, Hajjar I, Shaughnessy M, Mintzer MJ. The consortium of e-learning in geriatrics instruction. Journal of the American Geriatrics Society 2007;55(3):458-63.

Savović J, Jones H, Altman D, Harris R, Jűni P, Pildal J et al. Influence of reported study design characteristics on intervention effect estimates from randomised controlled trials: combined analysis of meta-epidemiological studies. Health Technol Assess 2012;16(35):1-82.

Schünemann HJ, Oxman AD, Higgins JPT, Vist GE, Glasziou P, Guyatt GH. Chapter 11: Presenting results and ‘Summary of findings' tables. In: Higgins JPT, Green S (editors), Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. www.cochrane-handbook.org.

Schwarzer G. Meta: an R package for meta-analysis. 2013. Available at: http://cran.r-project.org/package=meta. Accessed July 1, 2014.

Shojania KG, McDonald KM, Wachter RM, OwensDK. Closing the quality gap: a critical analysis of quality improvement strategies. In: Series Overview and Methodology. Vol. 1. Rockville MD: Agency for Healthcare Research and Quality, August 2004.

Shojania KG, Jennings A, Mayhew A, Ramsay CR, Eccles MP, Grimshaw J. The effects of on-screen, point of care computer reminders on processes and outcomes of care. Cochrane Database of Systematic Reviews 2009, Issue 3. Art. No.: CD001096. DOI: 10.1002/14651858.CD001096.pub2.

Sim I, Gorman P, Greenes RA, et al. Clinical decision support systems for the practice of evidence-based medicine. J Am Med Inform Assoc. 2001;8(6):527---534.

Signori A, Gallo F, Bovis F, Di Tullio N, Maietta I, Sormani MP. Long-term impact of interferon or Glatiramer acetate in multiple sclerosis: A systematic review and meta-analysis. Multiple Sclerosis and Related Disorders 2016;6:57-63.

Siontis KC, Hernandez-Boussard T, Ioannidis JP. Overlapping meta-analyses on the same topic: survey of published studies. BMJ 2013;347: f4501.

Starr M, Chalmers I, Clarke M, Oxman AD. The origins, evolution, and future of The Cochrane Database of Systematic Reviews. International journal of technology assessment in health care 2009(25) Suppl 1: 182–195.

Steinman MA, Ranji SR, Shojania KG, Gonzales R. Improving antibiotic selection: a systematic review and quantitative analysis of quality improvement strategies. Medical Care 2006;44:617-28.

Sterne JA, Egger M, Smith GD. Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. BMJ 2001;323(7304):101-5.

Sterne JAC, Egger M, Moher D (editors). Chapter 10: Addressing reporting biases. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Intervention. Version 5.1.0

Straus SE, Green ML, Bell DS, Badgett R, Davis D, Gerrity M, et al. Evaluating the teaching of evidence based medicine: conceptual framework. BMJ 2004;329:1029-32

Swedish Presidency of the EU, e-Health for a Healthier Europe! Opportunities for a better use of healthcare resources, 2009. Available at: http://www.government.se/content/1/c6/12/98/15/5b63bacb.pdf. Accessed July 1, 2014.

Tobias A. Assessing the influence of a single study in the meta-analysis estimate. Stata Tech Bull. 1999;8:15---17.

Tricco AC, Tetzlaff J, Pham B, Brehaut J, Moher D. Non-Cochrane vs. Cochrane reviews were twice as likely to have positive conclusion statements: cross-sectional study. Journal of clinical epidemiology 2009;62: 380–386 e381.

University of Aberdeen, Health Services Research Unit. Database of intra-correlation coefficients (ICCs). Available at: http://www.abdn.ac.uk/hsru/research/delivery/behaviour/methodological-research/ 2015.

Useem J, Brennan A, LaValley M, Vickery M, Ameli O,Reinen N et al. Systematic Differences between Cochrane and Non-Cochrane Meta-Analyses on the Same Topic: A Matched Pair Analysis. PLoS One 2015;10(12): e0144980]

Virgili G, Conti AA, Moja L, Gensini GF, Gusinu R. Heterogeneity and meta-analyses: do study results truly differ? Internal and Emergency Medicine 2009;4(5):423-7.

Walsh JM, McDonald KM, Shojania KG, Sundaram V, Nayak S, Lewis R, et al. Quality improvement strategies for hypertension management: a systematic review. Medical Care 2006;44:646-57.

Ward JP, Gordon J, Field MJ, Lehmann HP. Communication and information technology in medical education. Lancet 2001;357:792-6

Weijer C, Grimshaw JM, Eccles MP, McRae AD, White A, Brehaut JC, Taljaard M. The Ottawa Statement on the Ethical Design and Conduct of Cluster Randomized Trials. PLoS Med 2012;9(11):e1001346.

Welsh ET, Wanberg CR, Brown KG, Simmering MJ. E-learning: Emerging uses, empirical results and future directions. International Journal of Training and Development 2003;7(4):245-58.

Wentling TL, Waight C, Gallagher J, La Fleur J, Wang C, Kanfer A. E-learning - A Review of Literature. University of Illinois at Urbana-Champaign 2000;http://learning.ncsa.uiuc.edu/papers/elearnlit.pdf.

Wood L, Egger M, Gluud LL, Schulz KF, Jüni P, Altman DG et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. BMJ 2008;336(7644):601-5.

Wutoh R, Boren SA, Balas EA. eLearning: a review of Internet-based continuing medical education. Journal of Continuing Education in the Health Professions 2004;24:20-30.

Yusuf S, Peto R, Lewis J, Collins R, Sleight P. Beta-blockade during and after myocardial infarction: an overview of the randomised trials. Progress in Cardiovascular Diseases 1985;27(5):335-71.

Zagmutt F, Carroll C. A network meta-analysis assessing the rate of adverse events and drop outs of alternative treatments for relapsing forms of multiple sclerosis. In: Neurology. Vol. 80. 2013:1.

Zhou M, Holden L, Bedard G, Zeng L, Lam H, Chu D et al. The utilization of telephone follow-up in the advanced cancer population: a review of the literature. J Comp Eff Res 2012;1(6):509-517.

Zimitat C. Designing effective on-line continuing medical education. Medical Teacher 2001;23(2):117-22

Zivadinov R, Reder AT, Filippi M, Minagar A, Stüve O, Lassmann H, et al. Mechanisms of action of disease-modifying agents and brain volume changes in multiple sclerosis. Neurology 2008;71(2):136-44.

References for the review “INFs vs GA for RRMS Interferons-beta versus glatiramer acetate for relapsing-remitting multiple sclerosis”

<u>Included studies</u>

Cadavid 2009a

Cadavid D, Cheriyan J, SkurnickJ, LincolnJ A, Wolansky LJ, Cook SD. New acute and chronic black holes in patients with multiple sclerosis randomised to interferon beta-1b or glatiramer acetate. Journal Neurology Neurosurgery and Psychiatry 2009;80(12):1337-43.

Cadavid D, Kim S, Peng B, Skurnick J, Younes M, Hill J, et al. Clinical consequences of MRI activity in treated multiple sclerosis. Multiple Sclerosis Journal 2011;17(9):1113–21.

* Cadavid D, Wolansky LJ, Skurnick J, Lincoln J, Cheriyan J, Szczepanowski K, et al. Efficacy of treatment of MS with IFNbeta-1b or glatiramer acetate by monthly brain MRI in the BECOME study. Neurology 2009;72(23):1976-83.

Cheriyan J, Kim S, Wolansky LJ, Cook SD, Cadavid D. Impact of inflammation on brain volume in multiple sclerosis. Archives of Neurology 2012;69(1):82-8.

Calabrese 2012

Calabrese M, Bernardi V, Atzori M, Mattisi I, Favaretto A, Rinaldi F, et al. Effect of disease-modifying drugs on cortical lesions and atrophy in relapsing–remitting multiple sclerosis. Multiple Sclerosis Journal 2012;18(4):418–24.

Lublin 2013

Lindsey JW, Scott TF, Lynch SG, Cofield SS, Nelson F, Conwit R, et al. The CombiRx trial of combined therapy with interferon and glatiramer acetate in relapsing remitting MS: design and baseline characteristics. Multiple Sclerosis and Related Disorders 2012;1(2):81-6.

Lublin F, Cofield S, Cutter G, Conwit R, Narayana P, Nelson F, et al. The CombiRx trial: a multi-center, double-blind, randomized study comparing the combined use of interferon beta-1a and glatiramer acetate to either agent alone in participants with relapsing remitting multiple sclerosis - clinical outcomes. Neurology 2012;78(1Suppl1):PL02.003.

Lublin F, Cofield S, Cutter G, Salter A, Wang J, Conwit R, et al. EDSS changes in CombiRx: blinded, 7-year extension results for progression and improvement. Neurology 2013;80(7Suppl1):P04.121.

Lublin F, Cofield S, Cutter G, Salter A, Wang J, Conwit R, et al. Relapse activity in the CombiRx trial: blinded, 7-year extension results. Neurology 2013;80(7 Suppl 1):S01.002.

* Lublin FD, Cofield SS, Cutter GR, Conwit R, Narayana PA, Nelson F, et al. Randomized study combining interferon and glatiramer acetate in multiple sclerosis. Annals of Neurology 2013;73(3):327–40.

Wolinsky J, Narayana P, Nelson F, Datta S, Cofield S, Cutter G, et al. The CombiRx trial: a multi-center, double-blind, randomized study comparing the combined use of interferon beta-1a and glatiramer acetate to either agent alone in participants with relapsing remitting multiple sclerosis - MRI outcomes. Neurology 2012;78(Suppl 1):S11.002.

Wolinsky J, Salter A, Narayana P, Datta S, Nelson F, Cofield S, et al. MRI outcomes in CombiRx: blinded, 7-year extension results. Neurology 2013;80(Suppl 7):S01.003.

Mikol 2008

Coyle PK, Cornelisse P, Lehr L, Stubinski B. Time course of injection-site reactions to subcutaneous interferon beta-1a or glatiramer acetate in the REGARD Study. In: Proceedings of the 24th Annual Meeting of the Consortium of Multiple Sclerosis Centers, June 2–5, San Antonio, Texas, USA. 2010.

* Mikol DD, Barkhof F, Chang P, Coyle PK, Jeffery DR, Schwid SR, et al. Comparison of subcutaneous interferon beta-1a with glatiramer acetate in patients with relapsing multiple sclerosis (the REbif vs Glatiramer Acetate in Relapsing MS Disease [REGARD] study): a multicentre, randomised, parallel, open-label trial. Lancet Neurology 2008;7(10):903-14.

* Sørensen S. REGARD: what can we learn from randomised, open-label, head-to-head studies? The Lancet 2008;7(10):864-6.

* NCT01058005. Study Evaluating Rebif, Copaxone, and Tysabri for Active Multiple Sclerosis [A Multicenter, Randomized, Open-Label, Parallel-Group, Active-Controlled Study to Evaluate the Benefits of Switching Therapy (Glatiramer Acetate or Interferon Beta-1a) to Natalizumab in Subjects With Relapsing Remitting Multiple Sclerosis]. http://clinicaltrials.gov/show/NCT01058005 (first received 26 January 2010).

O'Connor 2009a

Filippi M, Rocca MA, Camesasca F, Cook S, O'Connor P, Arnason BG, et al. Interferon-1b and glatiramer acetate effects on permanent black hole evolution. Neurology 2011;76(14):1222–8.

Goodin DS, Hartung HP, O'Connor P, Filippi M, Arnason B, Comi G, et al. Neutralizing antibodies to interferon beta-1b multiple sclerosis: a clinico-radiographic paradox in the beyond trial. Multiple Sclerosis 2012;18(2):181-95.

Lampl C, Nagl S, Arnason B, Comi G, O Connor P, Cook S, et al. Efficacy and safety of interferon beta-1b SC in older RRMS patients-a post hoc analysis of the BEYOND study. Journal of Neurology 2013;260(7):1838-45.

O'Connor P, Arnason B, Comi GC, Filippi M, Cook S, Newark NJ, et al. Interferon beta-1b 500 mcg, Interferon beta-1b 250 mcg and glatiramer acetate: primary outcomes of the betaferon® efficacy yielding outcomes of a new dose study. Neurology 2008;71:LBS.004.

O'Connor P, Filippi M, Arnason B, Comi G, Cook S, Goodin D et al. 250 µg or 500 µg interferon beta-1b versus 20 mg glatiramer acetate in relapsing-remitting multiple sclerosis: a prospective, randomised, multicentre study [Errata]. Lancet Neurology 2009;8(11):981.

O'Connor P, Filippi M, Arnason B, Comi G, Cook S, Goodin D, et al. 250 µg or 500 µg interferon beta-1b versus 20 mg glatiramer acetate in relapsing-remitting multiple sclerosis: a prospective, randomised, multicentre study. Errata. Lancet Neurology 2012;11(1):27.

O'Connor P, Filippi M, Arnason B, Comi G, Cook S, Goodin D. 250 µg or 500 µg interferon beta-1b versus 20 mg glatiramer acetate in relapsing-remitting multiple sclerosis: a prospective, randomised, multicentre study. Errata. Lancet Neurology 2011;10(2):115.

* O'Connor P, Filippi M, Arnason B, Comi G, Cook S, Goodin D, et al. 250 microg or 500 microg interferon beta-1b versus 20 mg glatiramer acetate in relapsing-remitting multiple sclerosis: a prospective, randomised, multicentre study. Lancet Neurology 2009;8(10):889-97.

Pleimes D, Pohl C, Beckmann K, Stolz C. BEYOND Study – Data for Cochrane Analyses Protocol No. 306440. Bayer HealthCare Pharmaceuticals Affairs Specialized Therapeutics 2013.

Excluded studies

Barbato LM, Schofield L, McCague K, Pestreich L, Tobias K, Malhotra M, et al. Randomized, open-label study to evaluate patient-reported outcomes (PRO) with fingolimod after changing from prior disease-modifying therapy (DMT) for relapsing multiple sclerosis. In: American Neurological Association (ANA) 136th Annual Meeting, September 25 - 27, 2011; San Diego, California. 2011.

Beer K, Müller M, Hew-Winzeler AM, Bont A, Maire P, You X, et al. The prevalence of injection-site reactions with diseasemodifying therapies and their effect on adherence in patients with multiple sclerosis: an observational study. Bacteriologia, Virusologia, Parazitologia, Epidemiologia 2011;11(144):2-7.

Carra A, Onaha P, Luetic G, Burgos M, Crespo E, Deri N, et al. Therapeutic outcome 3 years after switching of immunomodulatory therapies in patients with relapsing-remitting multiple sclerosis in Argentina. European Journal Neurology 2008;15(4):386-93.

Carter NJ, Keating GM. Glatiramer acetate: a review of its use in relapsing-remitting multiple sclerosis and in delaying the onset of clinically definite multiple sclerosis. Drugs 2010;70(12):1545-77.

Comi G, Martinelli V, Rodegher M, Moiola L, Bajenaru O, Carra A, et al. Effect of glatiramer acetate on conversion to clinically definite multiple sclerosis in patients with clinically isolated syndrome (PreCISe study): a randomised, double-blind, placebo-controlled trial. Lancet 2009;374:1503–11.

Comi G, Cohen JA, Arnold DL, Wynn D, Filippi M, FORTE Study Group. Phase III dose-comparison study of glatiramer acetate for multiple sclerosis. Annals of Neurology 2011;69(1):75-82.

Del Santo F, Maratea D, Fadda V, Trippoli S, Messori A. Treatments for relapsing–remitting multiple sclerosis: summarising current information by network meta-analysis. European of Journal Clinical Pharmacology 2011;68(4):441-8.]

Ghezzi A, Amato MP, Capobianco M, Gallo P, Marrosu G, Martinelli V, et al. Disease-modifying drugs in childhood-juvenile multiple sclerosis: results of an Italian co-operative study. Multiple Sclerosis 2005;11(4):420-4.

Kalincik T, Jokubaitis V, Izquierdo G, Duquette P, Girard M, Grammond P, et al. Comparative effectiveness of glatiramer acetate and interferon beta formulations in relapsing-remitting multiple sclerosis. Multiple Sclerosis (Houndmills, Basingstoke, England) 2015;21:1159-71.

Khan OA, Tselis AC, Kamholz JA, Garbern JY, Lewis RA, Lisak RP. A prospective, open-label treatment trial to compare the effect of IFNbeta-1a (Avonex), IFNbeta-1b (Betaseron), and glatiramer acetate (Copaxone) on the relapse rate in relapsingremitting multiple sclerosis: results after 18 months of therapy. Multiple Sclerosis (Houndmills, Basingstoke, England) 2001;

Khan O, Bao F, Shah M, Caon C, Alexandros T, Selis A, et al. Effect of disease-modifying therapies on brain volume in relapsing–remitting multiple sclerosis: results of a five-year brain MRI study. Journal of the Neurological Sciences 2012;

Khan O, Rieckmann P, Boyko A, Selmaj K, Zivadinov R, GALA Study Group. Three times weekly glatiramer acetate in relapsing-remitting multiple sclerosis. Annals Neurology 2013;73(6):705-13.

Khoury SJ, Healy BC, Kivisäkk P, Viglietta V, Egorova S, Guttmann CR, et al. A randomised controlled double-masked trial of albuterol add-on therapy in patients with multiple sclerosis. Archives of Neurology 2010;67(9):1055-61.

Ouallet JC. Immunomodulatory treatments for multiple sclerosis: lessons from direct comparative studies [Traitements de fond de la sclérose en plaques: enseignements des études randomisées comparatives directes]. Reveu Neurologique 2010; 166(1):21-31.

Qizilbash N, Mendez I, Sanchez-de la Rosa R. Benefit-risk analysis of glatiramer acetate for relapsing-remitting and clinically isolated syndrome. Multiple Sclerosis Clinical Therapy 2012;34:159-76.

Salama HH, Abu-Hashim EM, El Bakry MA, Zhang J, El Mongui A. Twelve-month comparative study of the impacts of IFNb-1a (Avonex), IFNb-1b (Betaseron) and glatiramer acetate (Copaxone) on the clinical, MRI and immunological responses in relapsing-remitting multiple sclerosis. Neurosciences (Official Journal of the Pan Arab Union of Neurological Sciences) 2003;8:93-4.

Spelman T, Kalincik T, Zhang A, Pellegrini F, Wiendl H, Kappos L, et al. Comparative efficacy of switching to natalizumab in active multiple sclerosis. Annals of Clinical and Translational Neurology 2015;2(4):373–87.

Ongoing studies

EUCTR2012-003735-32-GR. Study to compare the efficacy and/or safety of masitinib to interferon beta-1a, interferon beta-1b, peginterferon beta-1a or glatiramer acetate in patients with relapsing remitting multiple sclerosis with unsatisfactory response to these first line treatments [A 96-weeks, prospective, multicentre, randomised, open label, active-controlled, parallel groups, phase 2b/3 study to compare efficacy and safety of masitinib to first line treatment, in patients with relapsing remitting multiple sclerosis with unsatisfactory response to first line treatment]. www.clinicaltrialsregister.eu/ctrsearch/trial/2012-003735-32/GR (first received 19 November 2015).

NCT00176592 Interferons-beta versus glatiramer acetate for relapsing-remitting multiple sclerosis. Phase IV Study, Betaseron Versus Copaxone for Relapsing Remitting or CIS Forms of MS Using Triple Dose Gad 3 T MRI [Phase IV, rater-blinded, randomized study, comparing 250 mg of betaseron with 20 mg of copaxone in patients with the relapsing-remitting(RR) or CIS forms of ms using 3 tesla (3T) magnetic resonance imaging (MRI) with triple-dose gadolinium]. ClinicalTrials.gov/show/NCT00176592 (first received 13 September 2005

NCT01623596. Evaluation of Patient Retention of Fingolimod vs. Currently Approved Disease Modifying Therapy in Patients With Relapsing Remitting Multiple Sclerosis [A 12-month, Prospective, Randomized, Active-controlled, Open-label Study to Evaluate the Patient Retention of Fingolimod vs. Approved First-line Disease Modifying Therapies in Adults With Relapsing Remitting Multiple Sclerosis (PREFERMS)]. https://clinicaltrials.gov/ct2/show/study/NCT01623596 (first received 18 June 2012).

References for the review “Effectiveness of Computerized Decision Support Systems Linked to Electronic Health Records: A Systematic Review and Meta-Analysis”

Included studies

Abdel-Kader K, Fischer GS, Li J, Moore CG, Hess R, Unruh ML. Automated clinical reminders for primary care providers in the care of CKD: a small cluster-randomized controlled trial. Am J Kidney Dis. 2011;58(6):894---902.

Apkon M, Mattera JA, Lin Z, et al. A randomized outpatient trial of a decision support information technology tool. Arch Intern Med. 2005;165(20):2388---2394.

Bell LM, Grundmeier R, Localio R, et al. Electronic health record-based decision support to improve asthma care: a cluster-randomized trial. Pediatrics. 2010;125(4):e770---e777.

Bosworth HB, Olsen MK, Dudley T, et al. Patient education and provider decision support to control blood pressure in primary care: a cluster randomized trial. Am Heart J. 2009;157(3):450---456.

Cleveringa FG, Welsing PM, van den Donk M, et al. Cost-effectiveness of the diabetes care protocol, a multifaceted computerized decision support diabetes management intervention that reduces cardiovascular risk. Diabetes Care. 2010; 33:258---263.

Cobos A, Vilaseca J, Asenjo C, et al. Cost effectiveness of a clinical decision support system based on the recommendations of the European Society of Cardiology and other societies for the management of hypercholesterolemia: report of a cluster-randomized trial. Dis Manag Health Outcomes. 2005;13(6):421---432.

Fitzgerald M, Cameron P, Mackenzie C, et al. Trauma resuscitation errors and computer-assisted decision support. Arch Surg. 2011;146(2):218---225.

Gonzales R, Anderer T, McCulloch CE, et al. A cluster randomized trial of decision support strategies for reducing antibiotic use in acute bronchitis. JAMA Intern Med. 2013;173(4):267---273.

Graumlich JF, Novotny NL, Nace GS, Aldag JC. Patient and physician perceptions after software-assisted hospital discharge: cluster randomized trial. J Hosp Med. 2009;4(6):356---363.

Gurwitz JH, Field TS, Rochon P, et al. Effect of computerized provider order entry with clinical decision support on adverse drug events in the long-term care setting. J Am Geriatr Soc. 2008;56(12):2225---2233.

Hetlevik I, Holmen J, Krüger O, Kristensen P, Iversen H, Furuseth K. Implementing clinical guidelines in the treatment of diabetes mellitus in general practice. Evaluation of effort, process, and patient outcome related to implementation of a computer-based decision support system. Int J Technol Assess Health Care. 2000;16(1):210---227.

Hetlevik I, Holmen J, Krüger O. Implementing clinical guidelines in the treatment of hypertension in general practice. Evaluation of patient outcome related to implementation of a computerbased clinical decision support system. Scand J Prim Health Care. 1999;17(1): 35-40.

Holbrook A, Pullenayegum E, Thabane L, et al. Shared electronic vascular risk decision support in primary care: Computerization of Medical Practices for the Enhancement of Therapeutic Effectiveness (COMPETE III) randomized trial. Arch Intern Med. 2011;171(19):1736---1744.

Holbrook A, Thabane L, Keshavjee K, et al. Individualized electronic decision support and reminders to improve diabetes care in the community: COMPETE II randomized trial. CMAJ. 2009;181(1-2):37---44.

Javitt JC, Rebitzer JB, Reisman L. Information technology and medical missteps: evidence from a randomized trial. J Health Econ. 2008;27(3):585---602.

Kucher N, Koo S, Quiroz R, et al. Electronic alerts to prevent venous thromboembolism among hospitalized patients. N Engl J Med. 2005;352(10):969---977.

Maclean CD, Gagnon M, Callas P, Littenberg B. The Vermont diabetes information system: a cluster randomized trial of a population based decision support system. J Gen InternMed. 2009;24(12):1303---1310.

McCowan C, Neville RG, Ricketts IW, et al. Lessons from a randomized controlled trial designed to evaluate computer decision support software to improve the management of asthma. Med Inform Internet Med. 2001;26(3):191---201.

McGregor JC, Weekes E, Forrest GN, et al. Impact of a computerized clinical decision support system on reducing inappropriate antimicrobial use: a randomized controlled trial. J Am Med Inform Assoc. 2006;13(4):378---384.

Montgomery AA, Fahey T, Peters TJ, MacIntosh C, Sharp DJ. Evaluation of computer based clinical decision support system and risk chart for management of hypertension in primary care: randomized controlled trial. BMJ. 2000; 320(7236):686---690.

O'Connor PJ, Sperl-Hillen JM, Rush WA, et al. Impact of electronic health record clinical decision support on diabetes care: a randomized trial. Ann Fam Med. 2011;9(1):12---21.

Paul M, Andreassen S, Tacconelli E, et al. Improving empirical antibiotic treatment using TREAT, a computerized decision support system: cluster randomized trial. J Antimicrob Chemother. 2006; 58(6):1238---1245.

Robbins GK, Lester W, Johnson KL, et al. Efficacy of a clinical decision-support system in an HIV practice: a randomized trial. Ann Intern Med. 2012;157(11): 757---766.

Rothschild JM, McGurk S, Honour M, et al. Assessment of education and computerized decision support interventions for improving transfusion practice. Transfusion. 2007;47(2):228---239.

Roy PM, Durieux P, Gillaizeau F, et al. A computerized handheld decision support system to improve pulmonary embolism diagnosis: a randomized trial. Ann Intern Med. 2009;151(10):677---686.

Subramanian S, Hoover S, Wagner JL, et al. Immediate financial impact of computerized clinical decision support for long-term care residents with renal insufficiency: a case study. J Am Med Inform Assoc. 2012;19(3):439---442.

Tamblyn R, Eguale T, Buckeridge DL, et al. The effectiveness of a new generation of computerized drug alerts in reducing the risk of injury from drug side effects: a cluster randomized trial. J Am Med Inform Assoc. 2012;19(4):635---643.

Tierney WM, Overhage JM, Murray MD, et al. Effects of computerized guidelines for managing heart disease in primary care. J Gen Intern Med. 2003; 18(12):967---976.

References for the review "e-Learning for licensed health professionals"

Included studies

Benjamin SE, Tate DF, Bangdiwala SI, Neelon BH, Ammerman AS, Dodds JM, et al. Preparing Child Care Health Consultants to address childhood overweight: a randomised controlled trial comparing web to in-person training. Matern Child Health J. 2008;12(5):662-9.

Fordis M, King JE, Ballantyne CM, Jones PH, Schneider KH, Spann SJ, et al. Comparison of the instructional efficacy of Internet-based CME with live interactive CME workshops: a randomised controlled trial. JAMA 2005;294(9):1043-51.

Harris JM Jr, Elliott TE, Davis BE, Chabal C, Fulginiti JV, Fine PG. Educating generalist physicians about chronic pain: live experts and online education can provide durable benefits. Pain Med. 2008;9(5):555-63.

Horiuchi S, Yaju Y, Koyo M, Sakyo Y, Nakayama K. Evaluation of a web-based graduate continuing nursing education program in Japan: A randomised controlled trial. Nurse Educ Today 2009;29(2):140-9.

Hugenholtz NI, de Croon EM, Smits PB, van Dijk FJ, Nieuwenhuijsen K. Effectiveness of e-learning in continuing medical education for occupational physicians. Occup Med (Lond) 2008;58(5):370-2.

Khatony A, Nayery ND, Ahmadi F, Haghani H, Vehvilainen-Julkunen K. The effectiveness of web-based and face-to-face continuing education methods on nurses' knowledge about AIDS: a comparative study. BMC Med Educ. 2009;Jul(10):9:41.

Le TT, Rait MA, Jarlsberg LG, Eid NS, Cabana MD. A randomised controlled trial to evaluate the effectiveness of a distance asthma learning program for paediatricians. J Asthma 2010;47(3):245-50.

Levine DA, Funkhouser EM, Houston TK, Gerald JK, Johnson-Roe N, Allison JJ, et al. Improving care after myocardial infarction using a 2-year Internet-delivered intervention: the Department of Veterans Affairs myocardial infarction-plus cluster-randomised trial. Arch Intern Med. 2011;171(21):1910-7.

Mäkinen M, Castrèn M, Tolska T, Nurmi J, Niemi-Murola L. Teaching basic life support to nurses. Eur J Anaesthesiol. 2006;Apr;23(4):327-31.

Maloney S, Haas R, Keating JL, Molloy E, Jolly B, Sims J, et al. Effectiveness of Web-based versus face-to-face delivery of education in prescription of falls-prevention exercise to health professionals: randomised trial. J Med Internet Res. 2011;Dec 22(13):e116.

Padalino Y, Peres HH. E-learning: a comparative study for knowledge apprehension among nurses. Rev Lat Am Enfermagem. 2007;May-Jun(15):397-403.

Perkins GD, Kimani PK, Bullock I, Clutton-Brock T, Davies RP, Gale M, et al. Improving the efficiency of advanced life support training: a randomised, controlled trial. Ann Intern Med. 2012;3(157):19-28.

Sheen ST, Chang WY, Chen HL, Chao HL, Tseng CP. E-Learning education program for registered nurses: the experience of a teaching medical centre. J Nurs Res. 2008;Sep;16(3):195-201.

Simonsen BO, Daehlin GK, Johansson I, Farup PG. Improvement of drug dose calculations by classroom teaching or e-learning: a randomised controlled trial in nurses. BMJ Open. 2014;Oct 24(4):e006025.

Excluded studies

Alfieri J, Portelance L, Souhami L, Steinert Y, McLeod P, Gallant F, et al. Development and impact evaluation of an e-learning radiation oncology module. Int J Radiat Oncol Biol Phys. 2012;82(3):e573-80.

Allison JJ, Kiefe CI, Wall T, Casebeer L, Ray MN, Spettell CM, et al. Multicomponent Internet continuing medical education to promote chlamydia screening. Am J Prev Med. 2005;28(3):285-90.

Anderson C. Training efforts to reduce reports of workplace violence in a community health care facility. J Prof Nurs. 2006;22(5):289-95.

Andolsek K, Rosenberg MT, Abdolrasulnia M, Stowell SA, Gardner AJ. Complex cases in primary care: report of a CME-certified series addressing patients with multiple co-morbidities. Int J Clin Pract. 2013;67(9):911-7.

Bayar MR, Poyraz BC, Aksoy-Poyraz C, Arikan MK. Reducing mental illness stigma in mental health professionals using a web-based approach. Isr J Psychiatry Relat Sci. 2009;46(3):226-30.

Beckley S, Stenhouse E, Greene K. The development and evaluation of a computer-assisted teaching programme for intrapartum fetal monitoring. BJOG. 2000;107(9):1138-44.

Beeckman D, Schoonhoven L, Boucqué H, Van Maele G, Defloor T. Pressure ulcers: e-learning to improve classification by nurses and nursing students. J Clin Nurs. 2008;17(13):1697-707.

Bello G, Pennisi MA, Maviglia R, Maggiore SM, Bocci MG, Montini L, et al. Online vs live methods for teaching difficult airway management to anesthesiology residents. Intensive Care Med. 2005;31(4):547-52.

Benedict N, Schonder K, McGee J. Promotion of self-directed learning using virtual patient cases. Am J Pharm Educ. 2013;77(7):151.

Beyea JA, Wong E, Bromwich M, Weston WW, Fung K. Evaluation of a particle repositioning maneuver Web-based teaching module. Laryngoscope 2008;118(1):175-80.

Bode R et al. E-learning software to enhance paediatric medical education. In: Conference: Annual Meeting of the Association of Pediatric Program Directors. San Antonio, TX: Academic pediatrics, 2012.

Boespflug AG, Dalle S, Thomas L. Enhancement of Customary Dermoscopy Education With Spaced Education e-Learning: A Prospective Controlled Trial. JAMA Dermatology 2015;151(8):847-853.

Bonevski B, Sanson-Fisher RW, Campbell E, Carruthers A, Reid AL, Ireland M. Randomized controlled trial of a computer strategy to increase general practitioner preventive care. Prev Med. 1999;29(6):478-86.

Browne L, Mehra S, Rattan R, Thomas G. Comparing lecture and e-learning as pedagogies for new and experienced professionals in dentistry. British Dental Journal 2004;197:95-97.

Buijze GA, Guitton TG, van Dijk CN, Ring D. Training improves interobserver reliability for the diagnosis of scaphoid fracture displacement. Clin Orthop Relat Res. 2012;470(7):2029-34.

Butler CC, Simpson SA, Dunstan F, Rollnick S, Cohen D, et al. Effectiveness of multifaceted educational programme to reduce antibiotic dispensing in primary care: practice based randomised controlled trial. BMJ 2012;2(344):d8173.

Butzlaff M, Vollmar HC, Floer B, Koneczny N, Isfort J, Lange S. Learning with computerized guidelines in general practice?: A randomised controlled trial. Fam Pract. 2004;21(2):183-8.

Carney PA, Bowles EJ, Sickles EA, Geller BM, Feig SA, Jackson S, et al. Using a tailored web-based intervention to set goals to reduce unnecessary recall. Acad Radiol. 2011;18(4):495-503.

Carney PA, Abraham L, Cook A, Feig SA, Sickles EA, Miglioretti DL, et al. Impact of an educational intervention designed to reduce unnecessary recall during screening mammography. Acad Radiol. 2012;19(9):1114-20.

Casap N, Nadel S, Tarazi E, Weiss EI. Evaluation of a navigation system for dental implantation as a tool to train novice dental practitioners. J Oral Maxillofac Surg. 2011;69(19):2548-56.

Chan DH, Leclair K, Kaczorowski J. Problem-based small-group learning via the Internet among community family physicians: a randomised controlled trial. MD Comput. 1999;16(3):54-8.

Chenkin J, Lee S, Huynh T, Bandiera G. Procedures can be learned on the Web: a randomised study of ultrasound-guided vascular access training. Acad Emerg Med. 2008;15(10):949-54.

Chung S, Mandl KD, Shannon M, Fleisher GR. Efficacy of an educational Web site for educating physicians about bio-terrorism. Acad Emerg Med. 2004;11(2):143-8.

Cook DA, Beckman TJ, Thomas KG, Thompson WG. Adapting web-based instruction to residents' knowledge improves learning efficiency: a randomised controlled trial. J Gen Intern Med. 2008;23(7):985-90.

Crenshaw K, Curry W, Salanitro AH, Safford MM, Houston TK, Allison JJ, et al. Is physician engagement with Web-based CME associated with patients' baseline haemoglobin A1c levels? The Rural Diabetes Online Care study. Acad Med. 2010;85(9):1511-7.

Curtis JR, Westfall AO, Allison J, Becker A, Melton ME, Freeman A, et al. Challenges in improving the quality of osteoporosis care for long-term glucocorticoid users: a prospective randomised trial. Arch Intern Med. 2007;167(6):591-6.

de Beurs DP, de Groot MH, de Keijser J, Mokkenstorm J, van Duijn E, de Winter RF, et al. The effect of an e-learning supported Train-the-Trainer programme on implementation of suicide guidelines in mental health care. J Affect Disord. 2015;175:446-53.

de Beurs DPdG, de Keijser J. van Duijn E, de Winter RF, Kerkhof AJ. Evaluation of benefit to patients of training mental health professionals in suicide guidelines: cluster randomised trial. British Journal of Psychiatry 2016;208(5):477-483.

Dimeff LA, Woodcock EA, Harned MS, Beadnell B. Can dialectical behavior therapy be learned in highly structured learning environments? Results from a randomised controlled dissemination trial. Behav Ther. 2011;42(2):263-75.

Esche CA, Warren JI, Woods AB, Jesada EC, Iliuta R.. Traditional classroom education versus computer-based learning: how nurses learn about pressure ulcers.. J Nurses Prof Dev 2015;31(1):21-7.

Estrada CA, Safford MM, Salanitro AH, Houston TK, Curry W, Williams JH, et al. A web-based diabetes intervention for physician: a cluster-randomised effectiveness trial. Int J Qual Health Care 2011;23(6):682-9.

Fary RES, Chua J, Ranelli S, Chan M, Briggs AM. Policy-into-practice for rheumatoid arthritis: randomized controlled trial and cohort study of e-learning targeting improved physiotherapy management. Arthritis Care and Research 2015;67(7):913-922.

Fisher WW, Luczynski KC, Hood SA, Lesser DA , Machado MA, Piazza CC. Preliminary findings of a randomised clinical trial of a virtual training program for applied behavior analysis technicians. Res Autism Spectr Disord. 2014;8(9):1044–1054.

Foroudi F, Pham D, Bressel M, Tongs D, Rolfo A, Styles C, et al. The utility of e-Learning to support training for a multicentre bladder online adaptive radiotherapy trial (TROG 10.01-BOLART). Radiother Oncol. 2013;109(1):165-9.

Fox N, O'Rourke A, Roberts C, Walker J. Change management in primary care: Design and evaluation of an Internet-delivered course. Med Educ. 2001;35(8):803-5.

Franchi CT, Djade CD, Pasina L, Mannucci PM, Onder G, Gussoni et al. E-learning in order to improve drug prescription for hospitalised older patients: a cluster-randomized controlled study. British Journal of Clinical Pharmacology 2016;82(1):53-63.

Funk M, Rose L, Fennie K. Challenges of an Internet-based education intervention in a randomised clinical trial in critical care. AACN Adv Crit Care 2010;21(4):376-9.

Gerbert B, Bronstone A, Maurer T, Berger T, McPhee SJ, Caspers N. The effectiveness of an Internet-based tutorial in improving primary care physicians' skin cancer triage skills. J Cancer Educ. 2002;17(1):7-11.

Ghoncheh R, Kerkhof AJ, Koot HM. Effectiveness of adolescent suicide prevention e-learning modules that aim to improve knowledge and self-confidence of gatekeepers: study protocol for a randomised controlled trial. Trials 2014 Feb 8 2014;15(52):1-7.

Gordon M1, Chandratilake M, Baker P. Improved junior paediatric prescribing skills after a short e-learning intervention: a randomised controlled trial. Arch Dis Child 2011;96(12):1191-1194.

Gordon MA, Baker P, Chandratilake M. Is a short e-learning course effective at improving paediatric prescribing skills among foundation doctors? An open label randomised controlled trial. Archives of Disease in Childhood 2011;96:A22-A23.

Gordon JS, Mahabee-Gittens EM, Andrews JA, Christiansen SM, Byron DJ. A randomised clinical trial of a web-based tobacco cessation education program. Pediatrics 2013;131(2):e455-62.

Gordon M, Chandratilake M, Baker P. Low fidelity, high quality: a model for e-learning. The clinical teacher 2013;10:258-263.

Granpeesheh D, Tarbox J, Dixon RD, Peters CA, Thompson K, Kenzer A. Evaluation of an eLearning tool for training behavioral therapists in academic knowledge of applied behavior analysis. Res Autism Spectr Disord. 2010;4(1):11-17.

Gyorki DE, Shaw T, Nicholson J, Baker C, Pitcher M, Skandarajah A, et al. Improving the impact of didactic resident training with online spaced education. ANZ J Surg. 2013;83(6):477-80.

Hansen KE, Rosenblatt ER, Gjerde CL, Crowe ME. Can an online osteoporosis lecture increase physician knowledge and improve patient care? J Clin Densitom. 2007;10(1):10-20.

Harris JM Jr, Sun H. A randomised trial of two e-learning strategies for teaching substance abuse management skills to physicians. Acad Med. 2013;88(9):1357-62.

Hearty T, Maizels M, Pring M, Mazur J, Liu R, Sarwark J, et al. Orthopaedic resident preparedness for closed reduction and pinning of paediatric supracondylar fractures is improved by e-learning: a multisite randomised controlled study. J Bone Joint Surg Am. 2013;4(95):e1261-7.

Houwink EJ, van Teeffelen SR, Muijtjens AM, Henneman L, Jacobi F, van Luijk SJ, et al. Sustained effects of online genetics education: a randomised controlled trial on oncogenetics. Eur J Hum Genet. 2014;22(3):310-6.

Jensen ML, Mondrup F, Lippert F, Ringsted C. Using e-learning for maintenance of ALS competence. Resuscitation 2009;80(8):903-8.

Kemper KJ, Amata-Kynvi A, Sanghavi D, Whelan JS, Dvorkin L, Woolf A, et al. Randomized trial of an Internet curriculum on herbs and other dietary supplements for health care professionals. Acad Med. 2002;77(9):882-9.

Kerfoot BP. Adaptive spaced education improves learning efficiency: a randomised controlled trial. J Urol. 2010;183(2):678-81.

Kerfoot BP, Baker H. An online spaced-education game for global continuing medical education: a randomised trial. Ann Surg. 2012;256(1):33-8.

Kim JH, Shin JS. Effects of an online problem-based learning program on sexual health care competencies among oncology nurses: a pilot study. J Contin Educ Nurs. 2014;45(9):393-401.

Kobak KA, Engelhardt N, Lipsitz JD. Enriched rater training using Internet based technologies: a comparison to traditional rater training in a multi-site depression trial. J Psychiatr Res. 2006;40(3):192-99.

Kontio R, Lahti M, Pitkänen A, Joffe G, Putkonen H, Hätönen H, et al. Impact of eLearning course on nurses' professional competence in seclusion and restraint practices: a randomised controlled study. J Psychiatr Ment Health Nurs. 2011;18(9):813-21.

Kontio R, Hätönen H, Joffe G, Pitkänen A, Lahti M, Välimäki M. Impact of eLearning course on nurses' professional competence in seclusion and restraint practices: 9-month follow-up results of a randomised controlled study. J Psychiatr Ment Health Nurs. 2013;20(5):411-8.

Kontio R, Pitkänen A, Joffe G, Katajisto J, Välimäki M. eLearning course may shorten the duration of mechanical restraint among psychiatric inpatients: a cluster-randomized trial. Nord J Psychiatry 2014;68(7):443-9.

Legris MÈ, Séguin NC, Desforges K, Sauvé P, Lord A, Bell R, et al. Pharmacist Web-based training program on medication use in chronic kidney disease patients: impact on knowledge, skills, and satisfaction. J Contin Educ Health Prof. 2011;31(3):140-50.

Liaw SY, Wong LF, Chan, SWC, Ho JTY, Mordiffi SZ, Ang SB et al. Designing and evaluating an interactive multimedia Web-based simulation for developing nurses' competencies in acute nursing care: randomized controlled trial. Journal of Medical Internet Research 2015;17(1):e5.

Liaw SYW, Chan SW, Ho JT, Mordiffi SZ, Ang SB, Goh PS et al. Designing and evaluating an interactive multimedia Web-based simulation for developing nurses' competencies in acute nursing care: randomized controlled trial. Journal of Medical Internet Research 2015;17(1):e5.

Little P, Stuart B, Francis N, Douglas E, Tonkin-Crine S, Anthierens S, et al. Effects of internet-based training on antibiotic prescribing rates for acute respiratory-tract infections: a multinational, cluster, randomised, factorial, controlled trial. Lancet 2013;382(9899):1175-82.

Liu WI, Rong JR, Liu CY. Using evidence-integrated e-learning to enhance case management continuing education for psychiatric nurses: A randomised controlled trial with follow-up. Nurse Education Today 2014;34:1361-1367.

Liu WI, Chu RC, Chen SC. The development and preliminary effectiveness of a nursing case management e-learning program. Computers, Informatics, Nursing 2014;32(7):343-352.

Lu DF, Lin ZC, Li YJ. Effects of a Web-based course on nursing skills and knowledge learning. J Nurs Educ. 2009;48(2):70-7.

Maloney S, Haas R, Keating JL, Molloy E, Jolly B, Sims J, et al. Breakeven, cost benefit, cost effectiveness, and willingness to pay for web-based versus face-to-face education delivery for health professionals. J Med Internet Res. 2012;14(2):e47.

Markova A, Weinstock MA, Risica P, Kirtania U, Shaikh W, Ombao H, et al. Effect of a web-based curriculum on primary care practice: basic skin cancer triage trial. Fam Med. 2013;45(8):558-68.

Marshall E, York J, Magruder K, Yeager D, Knapp R, De Santis ML, et al. Implementation of online suicide-specific training for VA providers. Acad Psychiatry 2014;38(5):66-74.

MC Cormack H, Bipasha C. Using a web-based electronic resource to study cardiac anatomy. Clinical Anatomy 2012;25(6):798-812.

McCrow JS, Beattie ER. Delirium knowledge and recognition: a randomized controlled trial of a web-based educational intervention for acute care nurses. Nurse Education Today 2014;34(6):912-917.

Meckfessel S, Stühmer C, Bormann KH, Kupka T, Behrends M, Matthies H, et al. Introduction of e-learning in dental radiology reveals significantly improved results in final examination. J Craniomaxillofac Surg. 2011;39(1):40-8.

Midmer D, Kahan M, Marlow B. Effects of a distance learning program on physicians' opioid- and benzodiazepine-prescribing skills. J Contin Educ Health Prof. 2006;26(4):294-301.

Moja L, Moschetti I, Cinquini M, Sala V, Compagnoni A, Duca P, et al. Clinical evidence continuous medical education: a randomised educational trial of an open access e-learning program for transferring evidence-based information - ICEKUBE (Italian Clinical Evidence Knowledge Utilization Behaviour Evaluation) - study protocol. Implement Sci. 2008;3(37):1-11.

Moorthy K, Jiwanji M, Shah J, Bello F, Munoz Y, Darzi A. Validation of a web-based training tool for lumbar puncture. Medicine Meets Virtual Reality 2003;94:219-225.

Moreira IC, Ventura SR, Ramos I, Rodrigues PP. Development and assessment of an e-learning course on breast imaging for radiographers: a stratified randomized controlled trial. J Med Internet Res. 2015;5(17):e3.

NCT00394017 Waldorff FB. The use of reminders in implementing an e-learning program in general practice. https://clinicaltrials.gov/ct2/show/NCT00394017?term=NCT00394017&rank=1 (accessed 04 October 2015).

NCT00815724 Cabana MD, Le T. Evaluating a distance learning asthma education program for paediatricians (The DALI Study). https://clinicaltrials.gov/ct2/show/NCT00815724?term=NCT00815724&rank=1 (accessed 04 October 2015).

NCT00934141 Gustafson DH. Evaluating improvement strategies in addiction treatment (NIATx 200). https://clinicaltrials.gov/ct2/show/NCT00934141?term=NCT00934141&rank=1 (accessed 04 October 2015).

NCT00962455 Schermer TR, Crockett AJ, Pieters W. Feedback reports and e-learning in primary care spirometry (FRESCO). https://clinicaltrials.gov/ct2/show/NCT00962455?term=NCT00962455&rank=1 (accessed 04 October 2015).

NCT01326936 Harris JM. Online primary care physician (PCP) training in screening, brief Intervention, referral, and treatment. https://clinicaltrials.gov/ct2/show/NCT01326936?term=NCT01326936&rank=1 (accessed 04 October 2015).

NCT01427660 Heisler M. iDecide.Decido: diabetes medication decision support study. https://clinicaltrials.gov/ct2/show/NCT01427660?term=NCT01427660&rank=1 (accessed 04 October 2015).

NCT01834521 Reyners AKL. Web-based screening and tailored support (ENCOURAGE). https://clinicaltrials.gov/ct2/show/NCT01834521?term=NCT01834521&rank=1 (accessed 04 October 2015).

NCT01955005 Turvey CL. Use of the Internet for health information. https://clinicaltrials.gov/ct2/show/NCT01955005?term=NCT01955005&rank=1 (accessed 04 October 2015).

Nesterowicz K, Fereshtehnejad SM, Edelbring S. e-learning in continuing pharmacy education is effective and just as accepted as on-site learning. Pharmacy Education 2015;15(1):22-26.

Paul CL, Piterman L, Shaw J, Kirby C, Sanson-Fisher RW, Carey ML, et al. Diabetes in rural towns: effectiveness of continuing education and feedback for healthcare providers in altering diabetes outcomes at a population level: protocol for a cluster randomised controlled trial. Implement Sci. 2013;13(8):30.

Pearce-Smith N. Issues and problems for librarians conducting research - an example of a randomised controlled trial comparing the effect of e-learning with a taught workshop on the knowledge and search skills of health professionals. In: Third International Evidence-Based Librarianship Conference. Brisbane, Australia, 2005.

Pelayo M, Cebrián D, Areosa A, Agra Y, Izquierdo JV, Buendía F. Effects of online palliative care training on knowledge, attitude and satisfaction of primary care physicians. BMC Fam Pract 2011;12(37):1-11.

Pelayo-Alvarez M, Perez-Hoyos S, Agra-Varela Y. Clinical effectiveness of online training in palliative care of primary care physicians. J Palliat Med 2013;16(10):1188-1196.

Perkins GD, Kimani PK, Bullock I, Clutton-Brock T, Davies RP, Gale M, et al; Electronic Advanced Life Support Collaborators. Improving the efficiency of advanced life support training: a randomised, controlled trial. Ann Intern Med. 2012;157(1):19-28.

Pham DH, Foroudi F, Kron T, Bressel M, Hilder B, Chesson B. A Multidisciplinary Evaluation of a Web-based eLearning Training Programme for SAFRON II (TROG 13.01): a Multicentre Randomised Study of Stereotactic Radiotherapy for Lung Metastases. Clinical Oncology (Royal College of Radiologists) 2016;(in press):1-8.

Platz E, Goldflam K, Mennicke M, Parisini E, Christ M, Hohenstein C. Comparison of web- versus classroom-based basic ultrasonographic and EFAST training in 2 European hospitals. Ann Emerg Med. 2010;56(6):660-7.

Rafalski V, Andreeva I. Promoting access to Cochrane Collaboration outputs in developing countries: the possible role of post-graduate distance education. In: Publisher Name:, editor(s). 12th Cochrane Colloquium: Bridging the Gaps. 177p. 2004 Oct 2-6, :-6 edition. Ottawa, Ontario, Canada, 2004:177.

Rankin JA, Then KL, Atack L. Can emergency nurses' triage skills be improved by online learning? Results of an experiment. J Emerg Nurs. 2013;39(1):20-6.

Ruzek JI, Rosen RC, Marceau L, Larson MJ, Garvert DW, Smith L, et al. Online self-administered training for post-traumatic stress disorder treatment providers: design and methods for a randomised, prospective intervention study. Implement Sci. 2012 May 14;7:43. 2012;7(43):1-14.

Schermer TR, Akkermans RP, Crockett AJ, van Montfort M, Grootens-Stekelenburg J, et al. Effect of e-learning and repeated performance feedback on spirometry test quality in family practice: a cluster trial. Ann Fam Med. 2011;9(4):330-6.

Schopf T, Flytkjaer V. Impact of interactive web-based education with mobile and email-based support of general practitioners on treatment and referral patterns of patients with atopic dermatitis: randomised controlled trial. J Med Internet Res. 2012;14(6):e171.

Sharma V, Chamos C, Valencia O, Meineri M, Fletcher SN. The impact of Internet and simulation-based training on transoesophageal echocardiography learning in anaesthetic trainees: a prospective randomised study. Anaesthesia 2013;68(6):621-7.

Shaw T, Long A, Chopra S, Kerfoot BP. Impact on clinical behavior of face-to-face continuing medical education blended with online spaced education: a randomised controlled trial. J Contin Educ Health Prof. 2011;31(2):103-8.

Simpson SA, Butler CC, Hood K, Cohen D, Dunstan F, Evans MR, et al. Stemming the Tide of Antibiotic Resistance (STAR): a protocol for a trial of a complex intervention addressing the 'why' and 'how' of appropriate antibiotic prescribing in general practice. BMC Fam Pract. 2009;10(20):1-10.

Smeekens AE, Broekhuijsen-van Henten DM, Sittig JS, Russel IM, ten Cate OT, Turner NM, et al. Successful e-learning programme on the detection of child abuse in emergency departments: a randomised controlled trial. Arch Dis Child. 2011;96(4):330-4.

Stein BDC, Swartz HA, DeRosier ME, Sorbero MJ, Brindley RA, Burns RM et al. Implementing a Web-based intervention to train community clinicians in an evidence-based psychotherapy: A pilot study. Psychiatric services 2015;66(9):988-991.

Stewart M, Marshall JN, Østbye T, Feightner JW, Brown JB, Harris S, et al. Effectiveness of case-based on-line learning of evidence-based practice guidelines. Fam Med. 2005;37(2):131-8.

Sung YH, Kwon IG, Ryu E. Blended learning on medication administration for new nurses: integration of e-learning and face-to-face instruction in the classroom. Nurse Educ Today 2008;28(8):943-52.

Thompson JS, Lebwohl B, Syngal S, Kastrinos F. Knowledge of quality performance measures associated with endoscopy among gastroenterology trainees and the impact of a web-based intervention. Gastrointest Endosc. 2012;76(1):e1-4.

Tung CY, Chang CC, Ming JL, Chao KP. Occupational hazards education for nursing staff through web-based learning. Int J Environ Res Public Health 2014;11(12):13035-46.

Valish AU, Boyd NJ. The role of computer assisted instruction in continuing education of registered nurses: an experimental study. J Contin Educ Nurs. 1975;6(1):13-32.

Van de Steeg L, Langelaan M, Ijkema R, Wagner C.. The effect of a complementary e-learning course on implementation of a quality improvement project regarding care for elderly patients: a stepped wedge trial. Implement Sci. 2012;7(13):1-7.

van Stiphout FZ-vR; Aarts JE, Koffijberg H, Klarenbeek-deJonge E, Krulder, M.; Roes KC et al. MEDUCATE trial: effectiveness of an intensive EDUCATional intervention for IT-mediated MEDication management in the outpatient clinic - study protocol for a cluster randomized controlled trial. Trials [Electronic Resource] 2015;16:223.

Veredas FJ, Ruiz-Bandera E, Villa-Estrada F, Rufino-González JF, Morente L. A web-based e-learning application for wound diagnosis and treatment. Comput Methods Programs Biomed. 2014;116(3):236-48.

Vidal-Pardo JI, Pérez-Castro TR, López-Álvarez XL, Santiago-Pérez MI, García-Soidán FJ, Muñiz J. Effect of an educational intervention in primary care physicians on the compliance of

indicators of good clinical practice in the treatment of type 2 diabetes mellitus [OBTEDIGA project]. Int J Clin Pract. 2013;67(8):750-8.

Viguier MR, Aubin F, Leccia MT, Richard MA, Esposito-Farese M, Gaudin P et al. Online training on skin cancer diagnosis in rheumatologists: results from a nationwide randomized web-based survey. PLoS ONE [Electronic Resource] 2015;10(5):e0127564.

Wakefield PL, Wilson MA. Enhancing nurses' knowledge regarding the complex care of hospitalised patients on insulin. J Nurses Prof Dev. 2014;30(4):174-80.

Ward L. E-learning versus workshops to teach critical appraisal to health professionals: a randomised controlled equivalence study. In: Third International Evidence-Based Librarianship Conference; Brisbane, Australia. 2005.

Weaver MR, Crozier I, Eleku S, Makanga G, Mpanga Sebuyira L, Nyakake J, et al. Capacity-building and clinical competence in infectious disease in Uganda: a mixed-design study with pre/post and cluster-randomised trial components. PLoS One 2012;7(12):e51319.

Wehrs WH, Pfafflin M, May TW. E-learning courses in epilepsy—concept, evaluation, and experience with the E-learning course “genetics of epilepsies”. Epilepsia 2007;48(5):872-879.

Weston CM, Sciamanna CN, Nash DB. Evaluating online continuing medical education seminars: evidence for improving clinical practices. Am J Med Qual. 2008;23(6):475-83.

Worm BS, Buch SV. Does competition work as a motivating factor in e-learning? A randomised controlled trial. PLoS One 2014;17(9):e85434.

Yao KU, Muto M, Ishikawa H, Cardona HJ, Castro Filho EC, Pittayanon, R et al. Development of an E-learning system for the endoscopic diagnosis of early gastric cancer: An international multicenter randomized controlled trial. Gastrointestinal Endoscopy 2015;81(5 Supp 1):AB 327.

Studies awaiting classification

Bredesen IM, Bjøro K, Gunningberg L, Hofoss D. Effect of e-learning program on risk assessment and pressure ulcer classification - A randomized study. Nurse Educ Today 2016;40:191-197.

Khanal P, Vankipuram A, Ashby A, Vankipuram M, Gupta A, Drumm-Gurnee D et al. Collaborative virtual reality based advanced cardiac life support training simulator using virtual reality principles. J Biomed Inform 2014;51:49-59.

Soh BP, Reed W, Poulos A, Brennan PC. Visual search strategy of naive readers interpreting mammograms before and after an e-learning tutorial. Proceedings of Singapore Healthcare 2010;19:S214.

Wilson-Sands C, Brahn P, Graves K. The Effect of Instructional Methodon Cardiopulmonary Resuscitation Skill Performance. Journal for Nurses in Professional Development 2015;31(5):E1-7.

References for the review “Training interventions for improving telephone consultation skills in clinicians”

Included studies

Wood PR, Littlefield JH, Foulds DM. Telephone management curriculum for paediatric interns: a controlled trial. Pediatrics 1989;83(6):925-930.

Excluded studies

Afzali M, Lippert A, Ostergaard D. Health professional operators at dispatch centres shall ensure identification of cardiac arrest and initiate counselling by telephone [Sundhedsfaglige operatører på vagtcentralenskal erkende hjertestop og starte telefonisk rådgivning]. Ugeskrift for Laeger 2012;174(46):2852-2854.

Anonymous. Training centres gear up to teach new standards. Aids alert 1997;12(8):89-90.

Banach G, Livermore L. Does the training of 911 center personnel impact calls to a poison center from 911 centers? Clinical Toxicology 2012;50(7):630-631.

Bediang G, Franck C, Raetzo MA, Doell J, Ba M, Kamga Y et al. Developing clinical skills using a virtual patient simulator in a resource-limited setting. Stud Health Technol Inform 2013;192:102-106.

Bell KR, Esselman P, Garner MD, Doctor J, Bombardier C, Johnson K et al.. The Use of a world wide web-based consultation site to provide support to telephone staff in a traumatic brain injury demonstration project. Journal of Head Trauma Rehabilitation 2003;18(6):504-511.

Berlin C, Kogan J, Rusk M. A brief telephone medicine curriculum incorporating a telephone mini-cex to teach residents telephone medicine. Journal of General Internal Medicine 2010;25:S434-S435.

Billington J, Coster S, Murrells T, Norman I. Evaluation of a Nurse-Led Educational Telephone Intervention to Support Self-Management of Patients With Chronic Obstructive Pulmonary Disease: A RandomizedFeasibility Study. COPD: Journal of Chronic Obstructive Pulmonary Disease 2015;12:395-403.

Bishop A, Gamlin J, Hall J, Hopper C, Foster NE. PhysioDirect: supporting physiotherapists to deliver telephone assessment and advice services within the context of a randomised trial. Physiotherapy 2013;99(2):113-118.

Boutin H, Robichaud P, Valois P, Labrecque M. Impact of a continuing education activity on the quality of telephone interventions by nurses in an adult asthma client base. Journal of Nursing Care Quality 2006;21(4):335-343.

Brown SB, Eberle BJ. Use of the telephone by paediatric house staff: a technique for paediatric care not taught. Journal of Pediatrics 1974;84(1):117-119.

Campbell JL, Fletcher E, Britten N, Green C, Holt T, Lattimer V et al. The clinical effectiveness and cost-effectiveness of telephone triage for managing same-day consultation requests in general practice: a cluster randomised controlled trial comparing general practitioner-led and nurse-led management systems with usual care (the ESTEEM trial). Health Technol Assess 2015;19(13):1-212.

* Campbell JL, Fletcher E, Britten N, Green C, Holt TA, Lattimer V. Telephone triage for management of same-day consultation requests in general practice (the ESTEEM trial): a

cluster-randomised controlled trial and cost-consequence analysis. Lancet 2014;384(9957):1859-68.

Caralis P. Teaching residents to communicate: the use of a telephone triage system in an academic ambulatory clinic. Patient Education & Counseling 2010;80(3):351-353.

Chandrasekaran PN, Shlutter L. A survey and intervention to improve effective communication between Critical care team and Neurology team in a busy neurointensive care unit-A portal to improve patient care. Neurology 2015;84:Supplement P7.331.

Cunningham NJ, Weiland TJ, Van DJ, Paddle P, Shilkofski N, Cunningham NY. Telephone referrals by junior doctors: a randomised controlled trial assessing the impact of SBAR in a simulated setting. Postgraduate Medical Journal 2012;88(1045):619-626.

Curry TA, Schwartz MW. Telephone assessment of illness: what is being taught and learned? Pediatrics 1978;62(4):603-605.

Delvin M. The pitfalls of telephone advice: getting it right. Journal of Intensive Care Society 2014;15(4):325-327.

Duherty D, King Helm K, Hickey K, Hysen O, Kaud Faisal, Puckett R et al. Telephone etiquette. Hospital Food & Nutrition Focus 1988;4(6):8.

Evens S, Curtis P. Using patient-simulators to teach telephone communication skills to health professionals. Journal of Medical Education 1983;58(11):894-898.

Gleason K, O'Neill EB, Goldschmitt J, Horigan J, Moriarty L. Ambulatory oncology nurses making the right call: assessment and education in telephone triage practices. Clinical Journal of Oncology Nursing 2013;17(3):335-336.

Greenberg L, Pedreira F, Getson P. Teaching telephone management skills to paediatric residents in a pilot program using a standardised patient. Southern Medical Journal 1999;92(4):394-399.

Hamilton G, Ortega R, Hochstetler V, Pierson K, Lin P, Susan L. Teaching communication skills to hospice teams: comparing the effectiveness of a communication skills laboratory withIn-person, second life, and phone role-playing. American Journal of Hospice and Palliative Medicine 2014;31:611-618.

Kaakinen P, Kyngäs H, Tarkiainen K, Kääriäinen M. The effects of intervention on quality of telephone triage at an emergency unit in Finland: Nurses' perspective. International Emergency Nursing 2016;26:26-31.

King EB, Gregory RP, Schlundt DG, Garton R, Fielder MT, Snow DC et al. The effect of problem-solving training on the counselling skills of telephonic nurse care managers. Journal for Nurses in Staff Development 2007;23(5):229-237.

Kosower E, Inkelis SH, Seidel JS. Telephone T.A.L.K.: a telephone communication program. Pediatric Emergency Care 1991;7(2):76-79.

Lamb MP. Telephone precepting: the development of a curriculum. Teaching & Learning in Medicine 2004;16(3):276-278.

Lane AB. Combining telephone peer counselling and professional services for clients in intensive psychiatric rehabilitation. Psychiatric Services 1998;49(3):312-314.

Larsen JH, Risor O. Telephone consultations at the emergency service, Copenhagen County: analysis of doctor-patient communication patterns. Family Practice 1997;14(5):387-393.

Letizia A, Whitman T, Zapor M, Byers D, Martin G, Ressner R, Wortman G, Hartzell J. Answering the mail: Using a case-based model to teach teleconsultation skills to infectious disease fellows. American Journal of Tropical Medicine and Hygiene 2012;87(5 Suppl 1):289. [EMBASE: 71041579]

MacFarlane A, Harrison R, Murray E, Berlin A, Wallace P. A qualitative study of the educational potential of joint teleconsultations at the primary-secondary care interface. Journal of Telemedicine & Telecare 2006;12 Suppl 1:22-24.

NCT00799461 Syrjala K. Internet-Based Program With or Without Telephone-Based Problem-Solving Training in Helping Long-Term Survivors of Hematopoietic Stem Cell Transplant Cope With Late Complications. https://clinicaltrials.gov/ct2/show/study/NCT00799461?term=NCT00799461&rank=1.

Nestel D, Sains P, Wetzel CM, Nolan C, Tay A, Kneebone RL et al. Communication skills for mobile remote presence technology in clinical interactions. Journal of Telemedicine & Telecare 2007;13(2):100-104.

Orlov OI, Levanov VM, Merrell RC, Lavrentyev VA, Doarn CR. A Pilot Telemedicine Project in the Privolzhsky District, Russia. Telemedicine Journal and e-Health 2003;9(3):291-295.

Ottolini MC, Greenberg L. Development and evaluation of a CD-ROM computer program to teach residents telephone management. Pediatrics 1998;101(3):e2.

Record J, Niranjan-Azadi A, Christmas C, Hanyok LA, Rand CS, Hellmann DB et al. Telephone calls to patients after discharge from thehospital: an important part of transitions of care. Med Educ Online 2015;20:doi.org/10.3402/meo.v20.26701.

Reisman AB, Brown KE. Preventing communication errors in telephone medicine. Journal of General Internal Medicine 2005;20(10):959-963.

Rothwell EW, Ellington L, Planalp S, Crouch BI. Tele-health: Lessons and strategies from specialists in poison information. Patient Education and Counseling 2011;85(3):440-445.

Rowe N. Telephone advice and triage within paediatric oncology.. http://clok.uclan.ac.uk/11089/2/Rowe%20Nicola%20Final%20e-Thesis%20(Master%20Copy).pdf.

Rudd M, Rodgers H, Curless R, Sudlow M, Huntley S, Madhava B et al. Remote specialist assessment for intravenous thrombolysis of acute ischaemic stroke by telephone. Emergency Medicine Journal 2012;29(9):704-708.

Santos MV, Oliveira DC,Novaes MA. A tele-health strategy for increasing adherence in the treatment of hypertension in primary care. Telemedicine Journal & E-Health 2013;19(4):241-247.

Strasser PH, Levy JC, Lamb GA, Rosekrans J. Controlled clinical trial of paediatric telephone protocols. Pediatrics 1979;64(5 II Suppl):553-557.

Whitson HE, Hastings SN, Lekan DA, Sloane R, White HK, McConnell ES. A quality improvement program to enhance after-hours telephone communication between nurses and physicians in a long-term care facility. Journal of the American Geriatrics Society 2008;56(6):1080-1086.

APPENDICES

Appendix A. Search strategy for the review “INFs vs GA for RRMS Interferons-beta versus glatiramer acetate for relapsing-remitting multiple sclerosis

{interferon*} OR {interferon beta} OR {beta-1 interferon} OR {beta 1 interferon} OR {interferon beta-1*} OR {rebif} OR {avonex} OR {betaseron} OR {beta-seron} OR {betaferon} OR {beta-IFN-1*} OR {interferon beta-1*} OR {Interferon-beta*} OR {interferon beta*} OR {recombinant interferon beta-1*}

AND

{copolymer-1} OR {cop-1} OR {copaxone} OR {glatiramer acetate} OR {cpx} OR {cop1} OR {copolymer} OR {glatiramer} OR {immunomodulation*} OR {immunomodulator*} OR {immunosuppression}

AND

{relapsing remitting} OR {relapsing-remitting } OR {remitting-relapsing} OR {remitting relapsing}

Appendix B. Search strategy for the review “e-Learning for licensed health professionals”

Medline (OVID)

No.	Search terms	Results
1	("e-learning" or elearning).ti.	857
2	("e-learning" or elearning).ab.	1376
3	or/1-2	1662
4	*internet/ and *education/	55
5	((electronic or internet or internet-based or online or "on line" or remote or distance or mobile or web or "web 2*" or web-based or web deliver*) adj2 (class or classes or classroom? or class-room? or course or courses or course-work or education* or inservice or in-service or instruction* or learning or seminar? or teaching or workshop? or work-shop?)).ti,ab.	7437
6	((computeri?ed or computer-assisted or computer-mediated* or computer-based) adj2 (class or classes or classroom? or class-room? or course or courses or coursework or course-work or education or inservice or in-service or instruction* or learning or seminar? or teaching or workshop?)).ti,ab.	1743
7	((e-mail* or email* or e-mail-based or email-based) adj2 (class or classes or classroom? or class-room? or course or courses or course-work or education* or inservice or in-service or instruction* or learning or seminar? or teaching or workshop? or work-shop?)).ti,ab.	83
8	(e-education or e-instruction or elearning or "e learning" or "e train*" or "e curricul*" or "e program*" or m-learn*).ti,ab.	1792
9	(virtual adj2 (class or classes or classroom? or course? or education* or inservice or in-service or instruction* or instructor? or learning or seminar? or teacher? or teaching or training or trainer? or workshop*)).ti,ab.	1243
10	((3g or 4g or ipad or iphone or handheld or (tablet adj5 computer?) or android or cell phone or mobile phone) adj4 (educational or class)).ti,ab.	27
11	(distributed adj3 (curricul* or education or learning)).ti,ab.	298
12	spaced learning.ti,ab.	35
13	("remote course*" or "remote education" or "remote seminar?" or "remote learning" or "remote workshop*" or (remote participation adj4 (education? or workshop or course or learning))).ti,ab.	40
14	(virtual or online or web or internet).ti.	51312
15	or/4-14	59766
16	*postgraduate education/ or *continuing education/ or *in service training/ or *professional development/	3449
17	(post-graduate or graduate education or graduate degree? or ((master? or doctoral) adj2 degree?) or doctorate or doctoral or post-professional).ti,ab.	8089
18	(continuing adj2 (medical or nursing or pharmacist? or physician? or doctor? or allied health) adj3 education?).ti,ab.	5321
19	(inservice training or professional development or cme).ti,ab.	11093
20	or/16-19	26273
21	(15 and 20) not 3	913
22	*nurse/ or exp *paramedical personnel/ or exp *physician/ or *medical personnel/	132064
23	(continuing adj2 education?).ti,ab,hw.	62702
24	(and/15,22-23) not (or/3,21)	77
25	*dental education/ or *medical education/ or *nursing education/	68626
26	25 not (undergraduate? or first year or second year or third year or preclinical or pre-clinical).ti,ab,hw.	63971
27	(26 and 15) not (or/3,21,24)	1166
28	controlled clinical trial/ or controlled study/ or randomized controlled trial/	510348
29	randomi?ed.ti. or ((random* or control) adj3 (group? or cohort? or patient? or hospital* or	641737

	department?)).ab. or (controlled adj2 (study or trial)).ti.	
30	(multicenter and (study or trial)).ti.	20362
31	(random sampl* or random digit* or random effect* or random survey or random regression).ti,ab. not randomized controlled trial/	62344
32	(exp animals/ or exp invertebrate/ or animal experiment/ or animal model/ or animal tissue/ or animal cell/ or nonhuman/) and (human/ or normal human/ or human cell/)	16144262
33	(exp animals/ or exp invertebrate/ or animal experiment/ or animal model/ or animal tissue/ or animal cell/ or nonhuman/) not 32	4275233
34	(or/28-30) not (or/31,33)	841718
35	3 and 34	176
36	21 and 34	58
37	24 and 34	9
38	27 and 34	54
39	or/35-38	297

Embase

No.	Search terms	Results
1	("e-learning" or elearning).ti.	1157
2	("e-learning" or elearning).ab.	2220
3	or/1-2	2597
4	computer-assisted instruction/	62027
5	((electronic or internet or internet-based or online or "on line" or remote or distance or mobile or web or "web 2*" or web-based or web deliver*) adj2 (class or classes or classroom? or class-room? or course or courses or course-work or education* or inservice or in-service or instruction* or learning or seminar? or teaching or workshop? or work-shop?)).ti,ab.	9126
6	((computeri?ed or computer-assisted or computer-mediated* or computer-based) adj2 (class or classes or classroom? or class-room? or course or courses or coursework or course-work or education or inservice or in-service or instruction* or learning or seminar? or teaching or workshop?)).ti,ab.	2086
7	((e-mail* or email* or e-mail-based or email-based) adj2 (class or classes or classroom? or class-room? or course or courses or course-work or education* or inservice or in-service or instruction* or learning or seminar? or teaching or workshop? or work-shop?)).ti,ab.	156
8	(e-education or e-instruction or elearning or "e learning" or "e train*" or "e curricul*" or "e program*" or m-learn*).ti,ab.	2778
9	(virtual adj2 (class or classes or classroom? or course? or education* or inservice or in-service or instruction* or instructor? or learning or seminar? or teacher? or teaching or training or trainer? or workshop*)).ti,ab.	1632
10	((3g or 4g or ipad or iphone or handheld or (tablet adj5 computer?) or android or cell phone or mobile phone) adj4 (educational or class)).ti,ab.	45
11	(distributed adj3 (curricul* or education or learning)).ti,ab.	352
12	spaced learning.ti,ab.	46
13	("remote course*" or "remote education" or "remote seminar?" or "remote learning" or "remote workshop*" or (remote participation adj4 (education? or workshop or course or learning))).ti,ab.	55
14	(virtual or online or web or internet).ti.	59771
15	or/4-14	128433
16	education, medical, continuing/ or education, medical, graduate/ or exp "internship and residency"/ or education, nursing, continuing/ or education, nursing, graduate/ or education, pharmacy, continuing/ or education, pharmacy, graduate/ or pharmacy residencies/ or inservice training/ or staff development/	660488
17	(post-graduate or graduate education or graduate degree? or ((master? or doctoral) adj2 degree?) or doctorate or doctoral or post-professional).ti,ab.	10031
18	(continuing adj2 (medical or nursing or pharmacist? or physician? or doctor? or allied health) adj3 education?).ti,ab.	6614

19	(inservice training or professional development or cme).ti,ab.	15275
20	or/16-19	674033
21	(15 and 20) not 3	49387
22	exp allied health personnel/ or exp *dentists/ or exp medical staff/ or exp nurses/ or pharmacists/ or exp physicians/	907485
23	(continuing adj2 education?).ti,ab,hw.	43200
24	(and/15,22-23) not (or/3,21)	176
25	education, dental/ or education, medical/ or education, nursing/ or education, pharmacy/	537908
26	25 not (undergraduate? or first year or second year or third year or preclinical or pre-clinical).ti,ab,hw.	514219
27	(26 and 15) not (or/3,21,24)	27
28	(randomized controlled trial or controlled clinical trial).pt. or randomized.ab. or placebo.ab. or clinical trials as topic.sh. or randomly.ab. or trial.ti.	981031
29	exp animals/ not humans.sh.	21860327
30	28 not 29	92471
31	(3 or 21 or 24 or 27) and 30	232

The Cochrane Library (Wiley)

No.	Search terms	Results
#1	("e-learning" or elearning):ti	117
#2	("e-learning" or elearning):ab	188
#3	{or #1-#2}	216
#4	[mh "computer-assisted instruction"]	1039
#5	((electronic or internet or internet-based or online or "on line" or remote or distance or mobile or web or "web 2*" or web-based or web deliver*) near/2 (class or classes or classroom? or class-room? or course or courses or course-work or education* or inservice or in-service or instruction* or learning or seminar? or teaching or workshop? or work-shop?)):ti,ab	656
#6	((computeri?ed or computer-assisted or computer-mediated* or computer-based) near/2 (class or classes or classroom? or class-room? or course or courses or coursework or course-work or education or inservice or in-service or instruction* or learning or seminar? or teaching or workshop?)):ti,ab	276
#7	((e-mail* or email* or e-mail-based or email-based) near/2 (class or classes or classroom? or class-room? or course or courses or course-work or education* or inservice or in-service or instruction* or learning or seminar? or teaching or workshop? or work-shop?)):ti,ab	25
#8	(e-education or e-instruction or elearning or "e learning" or "e train*" or "e curricul*" or "e program*" or m-learn*):ti,ab	275
#9	(virtual near/2 (class or classes or classroom? or course? or education* or inservice or in-service or instruction* or instructor? or learning or seminar? or teacher? or teaching or training or trainer? or workshop*)):ti,ab	174
#10	((3g or 4g or ipad or iphone or handheld or (tablet near/5 computer?) or android or cell phone or mobile phone) near/4 (educational or class)):ti,ab	4
#11	(distributed near/3 (curricul* or education or learning)):ti,ab	15
#12	spaced learning:ti,ab	52
#13	("remote course*" or "remote education" or "remote seminar?" or "remote learning" or "remote workshop*" or (remote participation near/4 (education? or workshop or course or learning))):ti,ab	3
#14	(virtual or online or web or internet):ti	5035
#15	{or #4-#14}	6458
#16	[mh "education, medical, continuing"] or [mh "education, medical, graduate"] or [mh "internship and residency"] or [mh "education, nursing, continuing"] or [mh "education, nursing, graduate"] or [mh "education, pharmacy, continuing"] or [mh "education, pharmacy, graduate"] or [mh "pharmacy residencies"] or [mh "inservice training"] or [mh "staff development"]	2528
#17	(post-graduate or graduate education or graduate degree? or ((master? or doctoral) near/2	225

	degree?) or doctorate or doctoral or post-professional):ti,ab	
#18	(continuing near/2 (medical or nursing or pharmacist? or physician? or doctor? or allied health) near/3 education?):ti,ab	2
#19	(inservice training or professional development or cme):ti,ab	730
#20	{or #16-#19}	3340
#21	(#15 and #20)	339
#22	[mh "allied health personnel"] or [mh *dentists] or [mh "medical staff"] or [mh nurses] or [mh pharmacists] or [mh physicians]	4047
#23	(continuing near/2 education?):ti,ab,kw	2
#24	#15 and #22 and #23	0
#25	[mh "education, dental"] or [mh "education, medical"] or [mh "education, nursing"] or [mh "education, pharmacy"]	3454
#26	#25 not (undergraduate? or first year or second year or third year or preclinical or pre-clinical):ti,ab,kw	2873
#27	#26 and #15	456
#28	#3 or #21 or #24 or #27	720

Appendix C. Search strategy for the review "Training interventions for improving telephone consultation skills in clinicians"

	CINAHL (Ebsco)	
#	**Query**	**Results**
S97	S95 AND S85 [EPOC Results]	655
S96	S94 AND S95 [RCT Results]	316
S95	(S54 OR S55 OR S56 OR S57 OR S58 OR S59) OR (S12 AND S50) [Results before filters]	1,368
S94	S86 OR S87 OR S88 OR S89 OR S90 OR S91 OR S92 OR S93 [RCT Filter]	137,865
S93	TI controlled AND TI (trial or trials or study or experiment* or intervention)	15,931
S92	AB ((multicent* n2 design*) or (multicent* n2 study) or (multicent* n2 studies) or (multicent* n2 trial*)) or AB ((multi-cent* n2 design*) or (multi-cent* n2 study) or (multi-cent* n2 studies) or (multi-cent* n2 trial*))	5,903
S91	TI multicentre or multicenter or multi-centre or multi-center	3,900
S90	TI (cluster N2 trial* or cluster N2 study or cluster N2 group or cluster N2 groups or cluster N2 cohort or cluster N2 design or cluster N2 experiment*) OR AB (cluster N2 trial* or cluster N2 study or cluster N2 group or cluster N2 groups or cluster N2 cohort or cluster N2 design or cluster N2 experiment*)	1,454
S89	TI (control group or control groups OR control* experiment* or control* design or controlled study) OR AB (control group OR control groups or control* cohort* or controlled experiment* controlled design or controlled study)	44,895
S88	TI random* or AB random*	98,033
S87	TI ("clinical study" or "clinical studies") or AB ("clinical study" or "clinical studies")	6,327
S86	(MM "Clinical Trials+")	7,551
S85	S61 or S62 or S63 or S64 or S65 or S66 or S67 or S68 or S69 or S70 or S71 or S72 or S73 or S74 or S75 or S76 or S77 or S78 or S79 or S80 or S81 or S82 or S83 or S84 [EPOC Filter]	382,111
S84	TI ((time points n3 over) or (time points n3 multiple) or (time points n3 three) or (time points n3 four) or (time points n3 five) or (time points n3 six) or (time points n3 seven) or (time points n3 eight) or (time points n3 nine) or (time points n3 ten) or (time points n3 eleven) or (time points n3 twelve) or (time points n3 month*) or (time points n3 hour*) or (time points n3 day*) or (time points n3 "more than")) or AB ((time points n3 over) or (time points n3 multiple) or (time points n3 three) or (time points n3 four) or (time points n3 five) or (time points n3 six) or (time points n3 seven) or (time points n3 eight) or (time points n3 nine) or (time points n3 ten) or (time points n3 eleven) or (time points n3 twelve) or (time points n3 month*) or (time points n3 hour*) or (time points n3 day*) or (time points n3 "more than"))	1,357
S83	TI ((control w3 area) or (control w3 cohort*) or (control w3 compar*) or (control w3 condition) or (control w3 group*) or (control w3 intervention*) or (control w3 participant*) or (control w3 study)) or AB ((control w3 area) or (control w3 cohort*) or (control w3 compar*) or (control w3 condition) or (control w3 group*) or (control w3 intervention*) or (control w3 participant*) or (control w3 study))	41,546
S82	TI (multicentre or multicenter or multi-centre or multi-center) or AB random*	88,922
S81	TI random* OR controlled	30,266
S80	TI (trial or (study n3 aim) or "our study") or AB ((study n3 aim) or "our study")	74,120

S79	TI (pre-workshop or preworkshop or post-workshop or postworkshop or (before n3 workshop) or (after n3 workshop)) or AB (pre-workshop or preworkshop or post-workshop or postworkshop or (before n3 workshop) or (after n3 workshop))	286
S78	TI (demonstration project OR demonstration projects OR preimplement* or pre-implement* or post-implement* or postimplement*) or AB (demonstration project OR demonstration projects OR preimplement* or pre-implement* or post-implement* or postimplement*)	1,198
S77	(intervention n6 clinician*) or (intervention n6 community) or (intervention n6 complex) or (intervention n6 design*) or (intervention n6 doctor*) or (intervention n6 educational) or (intervention n6 family doctor*) or (intervention n6 family physician*) or (intervention n6 family practitioner*) or (intervention n6 financial) or (intervention n6 GP) or (intervention n6 general practice*) Or (intervention n6 hospital*) or (intervention n6 impact*) Or (intervention n6 improv*) or (intervention n6 individualize*) Or (intervention n6 individualise*) or (intervention n6 individualizing) or (intervention n6 individualising) or (intervention n6 interdisciplin*) or (intervention n6 multicomponent) or (intervention n6 multi-component) or (intervention n6 multidisciplin*) or (intervention n6 multi-disciplin*) or (intervention n6 multifacet*) or (intervention n6 multi-facet*) or (intervention n6 multimodal*) or (intervention n6 multi-modal*) or (intervention n6 personalize*) or(intervention n6 personalise*) or (intervention n6 personalizing) or (intervention n6 personalising) or (intervention n6 pharmaci*) or (intervention n6 pharmacist*) or (intervention n6 pharmacy) or (intervention n6 physician*) or (intervention n6 practitioner*) Or (intervention n6 prescrib*) or (intervention n6 prescription*) or (intervention n6 primary care) or (intervention n6 professional*) or (intervention* n6 provider*) or (intervention* n6 regulatory) or (intervention n6 regulatory) or (intervention n6 tailor*) or (intervention n6 target*) or (intervention n6 team*) or (intervention n6 usual care)	36,970
S76	TI (collaborativ* or collaboration* or tailored or personalised or personalized) or AB (collaborativ* or collaboration* or tailored or personalised or personalized)	33,986
S75	TI pilot	10,372
S74	(MH "Pilot Studies")	26,778
S73	AB "before-and-after"	15,409
S72	AB time series	1,576
S71	TI time series	219
S70	AB (before* n10 during or before n10 after) or AU (before* n10 during or before n10 after)	29,229
S69	TI ((time point*) or (period* n4 interrupted) or (period* n4 multiple) or (period* n4 time) or (period* n4 various) or (period* n4 varying) or (period* n4 week*) or (period* n4 month*) or (period* n4 year*)) or AB ((time point*) or (period* n4 interrupted) or (period* n4 multiple) or (period* n4 time) or (period* n4 various) or (period* n4 varying) or (period* n4 week*) or (period* n4 month*) or (period* n4 year*))	44,497
S68	TI ((quasi-experiment* or quasiexperiment* or quasi-random* or quasirandom* or quasi control* or quasicontrol* or quasi* W3 method* or quasi* W3 study or quasi* W3 studies or quasi* W3 trial or quasi* W3 design* or experimental W3 method* or experimental W3 study or experimental W3 studies or experimental W3 trial or experimental W3 design*)) or AB ((quasi-experiment* or quasiexperiment* or quasi-random* or quasirandom* or quasi control* or quasicontrol* or quasi* W3 method* or quasi* W3 study or quasi* W3 studies or quasi* W3 trial or quasi* W3 design* or experimental W3 method* or experimental W3 study or experimental W3 studies or experimental W3 trial or experimental W3 design*))	10,969
S67	TI pre w7 post or AB pre w7 post	8,092
S66	MH "Multiple Time Series" or MH "Time Series"	1,210
S65	TI ((comparative N2 study) or (comparative N2 studies) or evaluation study or evaluation studies) or AB ((comparative N2 study) or (comparative N2 studies) or evaluation study or evaluation studies)	9,471
S64	MH Experimental Studies or Community Trials or Community Trials or Pretest-Posttest Design + or Quasi-Experimental Studies + Pilot Studies or Policy Studies + Multicenter	30,981

	Studies	
S63	TI (pre-test* or pretest* or posttest* or post-test*) or AB (pre-test* or pretest* or posttest* or "post test*) OR TI (preimplement*" or pre-implement*) or AB (pre-implement* or preimplement*)	6,241
S62	TI (intervention* or multiintervention* or multi-intervention* or postintervention* or post-intervention* or preintervention* or pre-intervention*) or AB (intervention* or multiintervention* or multi-intervention* or postintervention* or post-intervention* or preintervention* or pre-intervention*)	132,670
S61	(MH "Quasi-Experimental Studies")	5,300
S60	(S12 AND S50) NOT (S54 or S55 or S56 or S57 or S58 or S59)	232
S59	(S17 AND S50) NOT (S54 or S55 or S56 or S57 or S58)	290
S58	(S17 AND S53 AND S50) NOT (S54 or S55 or S56 or S57)	17
S57	(S17 AND S38 AND S50) NOT (S54 or S55 or S56)	128
S56	(S17 AND S34 AND S50) NOT (S54 or S55)	649
S55	(S17 AND S23 AND S50) NOT S54	37
S54	S8 AND S50	279
S53	S51 OR S52	73,431
S52	TI (professional patient or physician patient or nurse patient) OR AB (professional patient or physician patient or nurse patient)	32,618
S51	MH Professional-Patient Relations OR MH Physician-Patient Relations OR MH Nurse-Patient Relations	46,421
S50	S39 OR S40 OR S41 OR S42 OR S43 OR S44 OR S45 OR S46 OR S47 OR S48 OR S49	83,710
S49	AB (telephone# or phone) N3 skill#	42
S48	(MM "Communication+" or "health communication") AND (TI skill# OR AB skill# N2 develop*)	1,502
S47	TI communication# skill# OR AB communication# skill#	3,569
S46	(TI CME OR AB CME) AND (TI education OR AB education OR MW education)	1,496
S45	TI continuing N2 education* OR AB continuing N2 education*	8,879
S44	TI skill# N2 develop* OR AB skill# N2 develop*	3,924
S43	TI (education* N2 (intervention* or program* or hospital# or office# or practitioner# or GP or doctor#)) OR AB (education* N2 (intervention* or program* or hospital# or office# or practitioner# or GP or doctor#))	20,254
S42	TI rounds OR AB rounds	3,812
S41	TI (inservice or ((staff or physician# or nurse or nurses or doctor# or resident# or residency or intern or interns or practitioner#) N2 (educational* or train* or development#))) OR AB (inservice or ((staff or physician# or nurse or nurses or doctor# or resident# or residency or intern or interns or practitioner#) N2 (educational* or train* or development#)))	14,872
S40	MH Staff Development OR "Inservice Training"	16,936
S39	(MH "Education, Continuing+" or MH "Internship and Residency" or MH "Preceptorship") OR ("Clinical Clerkship" or "Teaching Rounds")	25,093
S38	S35 OR S36 OR S37	291,252
S37	TI (doctor# or nurse or nurses or physician# or practitioner#) OR (TI ((medical or health* or nursing or allied health*) N2 (personnel or staff*)) OR AB ((medical or health* or nursing or allied health*) N2 (personnel or staff*)))	150,524
S36	MH "allied health personnel" or MH "nursing assistants+" or MH "physician assistants+"	9,211
S35	MH "health personnel" or MH "infection control practitioners" or MH "medical staff" or MH "nurses+" or MH "pharmacists" or MH "physicians+" or "nursing staff"	199,788
S34	S24 OR S25 OR S26 OR S27 OR S28 OR S29 OR S30 OR S31 OR S32 OR S33	1,088,777

S33	(MH "Public Health" or MH "Preventive Health Care" or "Community Health Nursing") OR ("Preventive Psychiatry" or "Public Health Practice")	42,665
S32	TI (patient# N2 (assess* or care or diagnos* or evaluat* or screen*)) OR AB (patient# N2 (assess* or care or diagnos* or evaluat* or screen*))	80,244
S31	TI history N2 taking OR AB history N2 taking	922
S30	(MH "Patient History Taking+")	10,251
S29	MH "health services+" OR MH community health services OR MH emergency medical services OR MH triage OR MH nursing care OR MH nursing service	504,609
S28	("patient care management" NOT (MH "telemedicine+" or MH "telenursing")) OR (MH health care delivery or MH disease management or MH multidisciplinary care team or MH patient-centered care or "comprehensive health care" or "nurse's practice patterns" or "physician's practice patterns")	44,331
S27	((MH "Diagnostic services+" or MH "neonatal assessment")) OR ("mass screening" or "anonymous testing" or "mass chest x-ray" or "multiphasic screening")	38,908
S26	(MH "Diagnosis+")	610,376
S25	MH Family Practice	9,663
S24	MH "Nursing Care+" or MH "Patient Care" or MH "After Care" or MH "Ambulatory Care" or MH "Postoperative Care" or MH "Preoperative Care" or MH "Palliative Care" or MH "Perinatal Care" or MH "Postnatal Care" or MH "Prenatal Care" or MH "Prepregnancy Care"	244,491
S23	S18 OR S19 OR S20 OR S21 OR S22	25,892
S22	MH Remote Consultation	499
S21	TI consultation# OR AB consultation#	11,152
S20	MH "referral and consultation"	14,943
S19	TI (e-care or ecare or e-consult* or econsult* or e-diagnos* or ediagnosis* or e-health* or ehealth* or e-medicine or emedicine or e-nurse# or enurse# or e-nursing or enursing or e-physician# or ephysician# or e-psych* or epsych* or e-therapy or etherapy) OR AB (e-care or ecare or e-consult* or econsult* or e-diagnos* or ediagnosis* or e-health* or ehealth* or e-medicine or emedicine or e-nurse# or enurse# or e-nursing or enursing or e-physician# or ephysician# or e-psych* or epsych* or e-therapy or etherapy)	942
S18	TI (remote N2 (care or consult* or diagnos* or evaluat* or monitor* or treat* or therap*)) OR AB (remote N2 (care or consult* or diagnos* or evaluat* or monitor* or treat* or therap*))	465
S17	S13 OR S14 OR S15 OR S16	24,701
S16	AB telephone# or phone or phones	13,198
S15	AB telephone based or phone based	692
S14	TI telephone# or phone or phones or transtelephon*	3,496
S13	MH Telephone or Wireless Communications	14,862
S12	S9 OR S10 OR S11	6,793
S11	AB tele* N2 (care or counselling or diagnos* or health* or intervention# or manag* or therap* or treat* or medicine or medical or nursing or nurse# or physician# or doctor# or practitioner#)	2,096
S10	TI (teleassist* or tele-assist* or teleaudiolog* or tele-audiolog* or telebased or tele-based or telecancer or tele-cardiolo* or telecardiolog* or telecounselling or tele-counselling or teledental or tele-dental or telederm* or tele-derm* or telediagnos* or tele-diagnos* or teledialysis or tele-dialysis or teleecho* or tele-echo* or teleemerg* or tele-emerg* or teleepileps* or tele-epileps* or telefollow* or tele-follow* or teleguidance or tele-guidance or telehealth* or tele-health* or telehome* or tele-home* or teleICU or tele-ICU or teleintervention* or tele-intervention* or telemanag* or tele-manag* or telemedicine or tele-medicine or telemental* or tele-mental* or telemonitor* or tele-monitor* or telenurs* or tele-nurs* or teleoncolo* or tele-oncolo* or teleopthalm* or tele-opthalm* or telepalliat* or tele-palliat* or tele-patholog* or tele-patholog* or teleprocedu* or tele-procedu* or telepsych* or	2,782

	tele-psych* or teleradiol* or tele-radiol* or telerefer* or tele-refer* or telerehab* or tele-rehab* or telesurger* or tele-surger* or telesurgic* or tele-surgic* or teletherap* or tele-therap* or teletreat* or tele-treat* or teletriage or tele-triage) OR AB (teleassist* or tele-assist* or teleaudiolog* or tele-audiolog* or telebased or tele-based or telecancer or tele-cardiolo* or telecardiolog* or telecounselling or tele-counselling or teledental or tele-dental or telederm* or tele-derm* or telediagnos* or tele-diagnos* or teledialysis or tele-dialysis or teleecho* or tele-echo* or teleemerg* or tele-emerg* or teleepileps* or tele-epileps* or telefollow* or tele-follow* or teleguidance or tele-guidance or telehealth* or tele-health* or telehome* or tele-home* or teleICU or tele-ICU or teleintervention* or tele-intervention* or telemanag* or tele-manag* or telemedicine or tele-medicine or telemental* or tele-mental* or telemonitor* or tele-monitor* or telenurs* or tele-nurs* or teleoncolo* or tele-oncolo* or teleopthalm* or tele-opthalm* or telepalliat* or tele-palliat* or tele-patholog* or tele-patholog* or teleprocedu* or tele-procedu* or telepsych* or tele-psych* or teleradiol* or tele-radiol* or telerefer* or tele-refer* or telerehab* or tele-rehab* or telesurger* or tele-surger* or telesurgic* or tele-surgic* or teletherap* or tele-therap* or teletreat* or tele-treat* or teletriage or tele-triage)	
S9	MH telemedicine or telepathology or teleradiology or Telenursing	4,348
S8	S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7	2,831
S7	AB (telephone# or phone) N3 skill#	42
S6	MH Remote Consultation AND ((TI telephon* OR AB telephon* OR MW telephon*) or (TI (phone or phones) OR AB (phone or phones)))	109
S5	TI (telephone or telephones or phone or phones) and (care or counselling or diagnos* or health* or intervention# or manag* or therap* or treat* or medicine or medical or nursing or nurse# or physician# or doctor# or practitioner#)	1,242
S4	TI (telephone management or telephone communication or telephone medicine or telephone intervention* or telephone skill*) OR AB (telephone management or telephone communication or telephone medicine or telephone intervention* or telephone skill*)	971
S3	TI ((telephon* or phone# or phoning) N3 (physician# or GP or nurse or nurses or doctor# or general practitioner# or family doctor# or family practitioner# or consultant#)) OR AB ((telephon* or phone# or phoning) N3 (physician# or GP or nurse or nurses or doctor# or general practitioner# or family doctor# or family practitioner# or consultant#))	672
S2	TI (((telephon* or phone or phoning or phones or phoned) N3 (advice or advise# or advising or consult* or diagnos* or evaluat*))) OR AB (((telephon* or phone or phoning or phones or phoned) N3 (advice or advise# or advising or consult* or diagnos* or evaluat*)))	704
S1	TI (teleconsult* or tele-consult*) OR AB (teleconsult* or tele-consult*)	78

Publications related to this PhD Thesis (in order of presentation within this thesis)	**Publication type**	**Year**	**Publication state**
Pasini A, Rigon G, Vaona A. A cross-sectional study of the quality of telephone triage in a primary care out-of-hours service. J Telemed Telecare. 2015 Mar;21(2):68-72.	Cross sectional study	2015	Published
La Mantia L, Di Pietrantonj C, Rovaris M, Rigon G, Frau S, Berardo F, Gandini A, Longobardi A, Weinstock-Guttman B, Vaona A.. Interferons-beta versus glatiramer acetate for relapsing-remitting multiple sclerosis. Cochrane Database Syst Rev. 2014 Jul 26;(7):CD009333	Cochrane systematic review	2014	Published
La Mantia L, Di Pietrantonj C, Rovaris M, Rigon G, Frau S, Berardo F, Gandini A, Longobardi A, Weinstock-Guttman B, Vaona A. Comparative efficacy of interferon β versus glatiramer acetate for relapsing-remitting multiple sclerosis. J Neurol Neurosurg Psychiatry. 2015 Sep;86(9):1016-20.		2015	Co-published
La Mantia L, Di Pietrantonj C, Rovaris M, Rigon G, Frau S, Berardo F, Gandini A, Longobardi A, Weinstock-Guttman B, Vaona A.Interferons-beta versus glatiramer acetate for relapsing-remitting multiple sclerosis.Cochrane Database of Systematic Reviews 2016, Issue 11. Art. No.: CD009333.DOI: 10.1002/14651858.CD009333.pub3		2016	Up-date
Moja L, Kwag KH, Lytras T, Bertizzolo L, Brandt L, Pecoraro V, Rigon G, Vaona A, Ruggiero F, Mangia M, Iorio A, Kunnamo I, Bonovas S. Effectiveness of computerized decision support systems linked to electronic health records: a systematic review and meta-analysis. Am J Public Health. 2014 Dec;104(12):e12-22.	Non-Cochrane Systematic review	2014	Published
Vaona A, Banzi R, Kwag KH, Rigon G, Cereda D, Pecoraro V, Tramacere I, Moja L. E-learning for licensed health professionals. Cochrane Database Syst Rev. pending citation	Cochrane systematic review	2016	In press
Vaona A, Pappas Y, Grewal RS, Ajaz M, Majeed A, Car J. Training interventions for improving telephone consultation skills in clinicians. Cochrane Database Syst Rev. pending citation	Cochrane systematic review	2016	In press
Efficiency and safety of a clinical decision support software for medical telephone triage in a primary care out of hours service: a randomized controlled trial	Randomized Controlled Study	2016 failed	Unpublished
Moja L, Liberati EG, Galuppo L, Gorli M, Maraldi M, Nanni O, Rigon G, Ruggieri P, Ruggiero F, Scaratti G, Vaona A, Kwag KH.Barriers and facilitators to the uptake of computerized clinical decision support systems in specialty hospitals: protocol for a qualitative cross-sectional study. Implement Sci. 2014 Aug 28;9:105	Qualitative, cross sectional study (protocol)	2014	Published