Citation: Brice, A., Price, A. & Burls, A. (2015). Creating a database of internet-based clinical trials to support a public-led research programme: A descriptive analysis. Digital Health, 1, doi: 10.1177/2055207615617854

Original Article

DIGITAL HEALTH

Digital Health
2015, Vol. 0(0) 1–13
© The Author(s) 2015
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/2055207615617854
dhj.sagepub.com
SAGE

Creating a database of internet-based clinical trials to support a public-led research programme: A descriptive analysis

Anne Brice[1], **Amy Price**[1] **and Amanda Burls**[2]

Abstract

Background: Online trials are rapidly growing in number, offering potential benefits but also methodological, ethical and social challenges. The International Network for Knowledge on Well-being (ThinkWell^{TM}) aims to increase public and patient participation in the prioritisation, design and conduct of research through the use of technologies.

Objective: We aim to provide a baseline understanding of the online trial environment, determining how many trials have used internet-based technologies; how they have been used; and how use has developed over time.

Methods: We searched a range of bibliographic databases to March 2015, with no date limits, supplemented by citation searching and references provided by experts in the field. Results were screened against inclusion and exclusion criteria, and included studies mapped against a number of key dimensions, with key themes developed iteratively throughout the process.

Results: We identified 1992 internet-based trials to March 2015. The number of reported studies increased substantially over the study timeframe. The largest number of trials were conducted in the USA (49.7%), followed by The Netherlands (10.2%); Australia (8.5%); the United Kingdom (5.8%); Sweden (4.6%); Canada (4%); and Germany (2.6%). South Korea (1.5%) has the highest number of reported trials for other continents. There is a predominance of interventions addressing core public health challenges including obesity (8.6%), smoking cessation (5.9%), alcohol abuse (7.7%) and physical activity (10.2%); in mental health issues such as depression (10.9%) and anxiety (5.6%); and conditions where self-management (16.6%) or monitoring (8.1%) is a major feature of care.

Conclusions: The results confirm an increase in the use of the internet in trials. Key themes have emerged from the analysis and further research will be undertaken in order to investigate how the data can be used to improve trial design and recruitment, and to build an open access resource to support the public-led research agenda.

Keywords

Internet-based clinical trials, randomised controlled trial as topic, patient and public involvement, information science, information retrieval

Submission date: 12 July 2015; Acceptance date: 27 October 2015

Introduction

Online clinical trials are rapidly growing in number. This novel method for assessing the effectiveness of health interventions offers tantalising potential benefits to researchers, patients and society, such as: reducing the cost of studies; reducing the marginal cost of recruiting additional participants, enabling larger sample sizes and reducing random error; improving

---

[1] Evidence-Based Health Care Programme, Department of Primary Health Care Sciences; Department of Continuing Education, Oxford, UK
[2] Health Services Research and Management Division, School of Health Sciences, City University, London, UK

Corresponding author:
Anne Brice, Evidence-Based Health Care Programme, Department of Primary Health Care Sciences; Department of Continuing Education, New Radcliffe House, Radcliffe Observatory Quarter, Woodstock Road, OX2 6GG, UK.
Email: anne.brice@kellogg.ox.ac.uk
Twitter: @annebriceuk

the representativeness of trials by enabling the involvement of participants who tend to be neglected by current research methods; improving recruitment to trials; improving stakeholder involvement in trials; cheaper treatments; permitting easier long-term follow-up of participants; allowing novel questions to be addressed that would not have been logistically possible to test in traditional trials; and improved dissemination of results. However, the online method also brings potential threats such as: the ease of setting up trials leading to poor quality assurance; ease of recruitment leading to large attrition rates and uninterpretable results; contamination between arms due to social media or communication between participants; the technological challenges of doing research online leading to failed trials; and exposure to hacking and potential breaches of participant confidentiality. It is important, therefore, to evaluate this new technology to identify the actual advantages and disadvantages of online clinical trials in practice and to learn what factors can help maximise benefits and minimise risks.

ThinkWell

The International Network for Knowledge on Wellbeing (ThinkWell^{TM})[1] is a not-for-profit organisation set up by a group of consumers, researchers, health service users and clinicians in 2006 with the mission of using modern communication technologies to enable the public and patients (a) to make health decisions informed by the best available research evidence and (b) to participate in health research on the issues that affect the day-to-day decisions of individuals, involving the public in the prioritisation, design and conduct of research. ThinkWell's focus is on interventions that people can choose to do for themselves, without the involvement of a health care professional (such as exclusion diets, online psychological therapies, lifestyle choices, dietary supplements, etc.), and the organisation is committed to an evidence-based approach and stakeholder partnership throughout the research process. The type of intervention ThinkWell is interested in (self-management) and the fact that all research is designed and undertaken in partnership with the public and patients means that online trials are a particularly appropriate methodology to address proven uncertainties in this area. Therefore, it was important for ThinkWell to learn from best practice in online health trials regarding how to achieve meaningful stakeholder collaboration to inform the methodological, safety, and ethical considerations specific to the online environment. However, when ThinkWell started work in this area there was little published evidence about these issues. We decided to begin by looking at existing online trials to learn about their advantages and disadvantages; how stakeholders had been involved; and identify good and bad practice in conduct and reporting. A first step was to systematically identify as many relevant online clinical trials as efficiently as we could. This led to the creation of the Online Randomised Controlled Trials of Health Information Database (ORCHID). This paper describes the rationale and methods of creating ORCHID.

Rationale for public-led trials

Why involve the public and patients as partners in health research? Measures to improve the health of citizens by enabling them to make better-informed decisions about their health and well-being have been supported by national public health initiatives, media campaigns, the development of patient education resources, and tools to support shared decision making. An increase in the use of internet and mobile technologies as a channel for health-related information and advice has led to the development of computer-supported decision aids[2–4] and other self-management initiatives,[5] in tandem with the informal explosion of social media applications,[6] through forums and discussion groups,[7] and communication activities designed to promote behaviour change.[8] Pew internet research states that 35% of adults have searched online to find health information,[9] and in the UK government policies to tackle issues such as smoking cessation and obesity, self-management of long-term conditions, and improvements in mental health have been implemented through national digital initiatives such as Change4Life[10] and NHS Choices.[11]

Internet technologies may provide an efficient and cost-effective mechanism for the delivery of good-quality health information, public health-focused interventions, and health and social care transactions. The rapid rise in the use of smart phones, tablets and an increase in the accessibility of mobile internet technologies has led to the promotion of a 'digital first' strategy for the Department of Health in England,[12,13] which has the stated aim of making all public-facing transactions in health and social care digital by default. However, there are concerns about a lack of research into the impact of digital technologies in terms of harms;[14,15] on issues relating to digital skills;[16] self-diagnosis;[17] and on health inequalities.[18–20] Mobile technologies may provide an opportunity to support community engagement,[21] and although systematic reviews report that community engagement initiatives can impact positively on policy and practice, on health, and on public health outcomes, there is a lack of strong evidence on whether there is a beneficial effect on health inequalities, comparative effectiveness between models, or on cost-effectiveness.[22] More high-quality syntheses or studies are needed to inform reliable policy and practice about whether the increasing use of internet-based technologies has the potential to improve health equity.

In addition to the increased awareness of the need to promote better-informed and shared decision making, there has also been an acknowledgement of the need for a corresponding increase in patient and public involvement (PPI) in policy making, planning and service delivery,[23–26] health and social care guidance,[27] patient safety,[28] and importantly, in the research process itself.[29] Lack of meaningful stakeholder involvement has the potential to add to the waste identified in the research process by,[30] and represents a lost opportunity to harness new technologies for, socially led research.[31]

It has been suggested that without full involvement at all stages of the research process, much of the research activity that is conducted will not answer the questions that are most important to patients and the public,[32–35] and that a more participative research process will lead to better quality and more relevant research output. A range of policies and programmes have been implemented with the aim of improving PPI in health-related research, although there are concerns that this has been largely driven by a researcher-led agenda.[36] Chalmers[37] suggests that most health-related research is commissioned by policy makers and designed mainly by researchers, and often only involves patients and future participants after its initiation. Failure to involve potential participants in the prioritisation and design of research may partially explain why trials fail to recruit sufficient numbers or encounter a lack of cooperation from those who enrol patients at the point of care. The use of social networking sites to increase public involvement in health research and practice has been increasing, and guidance has been published to aid researchers;[38] however, further exploration of the full potential is required. Studies are often observational in design and lack the capacity to fully explore the interactional potential of the media.

Randomised controlled clinical trials are the current methodology of choice for testing the effectiveness of interventions because they minimise potential biases (for example, pre-existing differences in the groups being compared). However, there are known limitations and problems associated with their conduct, particularly in the area of public health and well-being. Problems include failures to recruit sufficient sample sizes, the costs associated with running large trials, problems in follow-up, and relevance.[39,40] This may impact on the ability to provide answers to the most important questions for patients, the public and clinicians, but also on the successful involvement of patients and the public in trial design and conduct,[41] with a lack of engagement contributing to concerns about poor reporting, publication bias and the suppression of information about adverse events and harms.[42] This lack of transparency undermines the confidence placed in reported findings[43–45] as well as providing another barrier to increased public involvement. These concerns have led to a number of initiatives to improve the quality of research evidence through synthesis in systematic reviews, better reporting standards[46] and a commitment to make patients aware of research in the NHS Constitution in England,[47] although the failure to register clinical trials remains a matter of serious concern.[48]

Consideration of these issues has led to specific initiatives that aim to harness the development of new technologies as a mechanism to both improve the useful outcomes of clinical research, built on the principles of involvement and through the development of public-led research.

While a growing interest in the use of internet technologies in general within clinical trials has been identified from an early stage in the development of the medium,[49,50] studies using internet technologies vary in the aspect and degree to which these tools are employed at different stages of the trial process, including recruitment, gaining consent, data collection, and the dissemination of results. As the adoption of internet-based technologies in clinical trials has increased rapidly, it has led to unstructured approaches to reporting and indexing, which in turn leads to problems in the retrieval of relevant studies, and in the assessment of quality and reliability. Murray[51] identified a set of issues concerning the use of the internet in clinical trials, and has suggested that online trials share common issues with traditional trials, such as recruitment and retention, randomisation, data quality and fidelity, but face additional challenges related to their technological characteristics, including spamming, data theft, and ‘cybersquatting’.

In a rapidly changing technical environment, the problems associated with the ability to be able to replicate interventions in reports of clinical trials[52] may be an even greater challenge.[53] Internet technologies have the potential to be used in all stages of trial conduct; however, scoping searches reveal that the most common application in the reports of clinical trials is in the delivery of the intervention or comparison, such as in studies investigating substance misuse;[54,55] smoking cessation;[56] weight loss;[57] and in the use of cognitive behavioural therapies.[58] The growth of this type of trial is reflected through an increasing number of published systematic reviews of these types of studies,[59–62] and also in the appearance of national guidance.[63] However, there are currently no agreed standards for the description or indexing of these concepts in individual studies, or in systematic reviews.

Given the lack of a robust evidence base, it was decided that a more systematic approach to the identification of uncertainties in relation to internet-based clinical trials would be useful for the emerging ThinkWell programme and for other researchers in the field of health research. UK DUETS[64,65] was established in 2007 in order to identify, collate and publish uncertainties about the effects of treatments. Information in UK DUETs not only helps health professionals identify treatment uncertainties quickly, but can also help those responsible for promoting and supporting research focus on important gaps in knowledge and the unmet information needs of patients and clinicians. A project to adapt the UK DUETs methodology to investigate the most important questions in relation to the use of internet-based technologies in health and well-being research was piloted, and analysis of the initial results showed that a number of important questions could benefit from further research.[66] These included areas such as data security, data collection, self-report issues, and technological solutions, but also raised broader questions concerning ethical, methodological and qualitative perspectives. It was unclear how many clinical trials had already been conducted using internet-based approaches, and to what extent they had been used in specific stages or elements of the trial process. We set out to consider the epidemiology of internet-based clinical trials, with the aim of establishing some core baselines and a better understanding of current activities, through a project to identify, classify, record and analyse reports of trials.

The ORCHID project

The ORCHID project was proposed to create a repository of internet-based clinical trials in health and well-being topics, to provide a baseline understanding of the nature of the internet-based clinical trial environment. We propose to use ORCHID to mine for trends in trial topics and types of interventions, to identify methods used in online trials, to explore how and when PPI is used, as well as any other emerging trends in the development of online trials.

The initial project aimed to address the following questions:

- How many randomised controlled trials have used internet-based technologies as part of their trial processes?
- In what stages of the trial process have internet-based technologies been used and how might these be characterised?
- How has the use of internet-based technologies in randomised controlled trials developed over time?

The mapping process aimed to establish the relevant concepts and the size of the literature, and will feed into the process of establishing priorities for research in this area alongside further development of the ThinkWell uncertainties harvesting project.

Methods

Definitions and scope

One of the major impediments in identifying internet-based clinical trials is related to the lack of a consistent taxonomy, and the wide range of terms used in both the reporting of the trial and the quality of indexing of trials in bibliographical databases. For the purposes of this project, the following definitions were used for key concepts.

Internet-based. The ‘internet’ concept is defined in the Oxford English Dictionary as “a global computer network providing a variety of information and communication facilities, consisting of interconnected networks using standardized communication protocols”. For the purposes of this research, the term ‘internet-based’ was used predominantly to refer to applications which use the world wide web as a primary means of communication or data entry, including the use of mobile technologies where these are used in combination with the internet.

Randomised controlled trials. The first phase of the database development focused on identifying a core set of reports of internet-based randomised controlled trials. To be considered for inclusion in the initial development phase, reports of studies needed to include reference to a randomisation process in the abstract, to be indexed with a relevant publication type, or to have ‘randomised controlled trial’ included in a declarative title. No assessment of trial quality was undertaken at this stage.

Health and well-being. The potential range and scope of this topic is extremely large, and the concept of ‘health’ is subject to a range of different interpretations and definitions. The focus for the initial database development was therefore limited to reports of trials relating to interventions in diagnosable illnesses or disorders, including non-communicable diseases and conditions. ICD10 categories were used to provide additional guidance for the inclusion or exclusion of borderline topics, for example in the area of addiction. For the term ‘well-being’, there is even less consensus concerning how this should be defined, particularly in the boundaries between the concepts of ‘mental health’ and ‘well-being’. The recent government strategy for mental health *No health without mental health; a cross-government mental health outcomes strategy for people of all ages, 2011*[67] provides the following definitions:

> Good or positive mental health is more than the absence or management of mental health problems; it is the foundations for well-being and effective functioning both for individuals and for their communities.
>
> Well-being is a positive state of mind and body, feeling safe and able to cope, with a sense of connection with people, communities and the wider environment

Factors affecting adult health, such as sexual and reproductive health and drug misuse, were included, but only where there was a defined health care, or public health, intervention. In order to manage the scope of the initial phase of development, other social determining factors such as housing, transport and planning, or broader environmental health; issues around ageing; domestic violence; or accidental injury were excluded.

Trial participants. Trials would be included if the participants were patients or members of the public. Trials dealing exclusively with health care professionals would be identified but not included for analysis at this stage.

Identifying and selecting studies

Search methods. A broad and comprehensive search strategy was developed, and a wide range of sources used in order to retrieve relevant studies. The following sources were searched to identify relevant studies: MEDLINE; EMBASE; PsycINFO; CINAHL; The Cochrane Controlled Trials Register (CCTR); Physiotherapy Evidence Database (PEDro); OT Seeker; ERIC; and LILACS. The search strategies are listed in Appendix 1.

Given the wide range of databases searched, their different functionalities and tools, the search strategy was adapted for each source. For health-related databases the 'internet' concepts were combined with a methodological filter for randomised controlled trials. For non-health-related databases the 'internet' concepts were combined with health or well-being terms, as well as any relevant randomised controlled trial concepts or methodological filters if available.

It was determined that although the preference was for conducting a comprehensive search in a wide range of sources, i.e. not just MEDLINE and EMBASE, in some smaller databases it would be more pragmatic and cost effective to use a single term or less sensitive approach, if a more precise or tested filter were not available. This would be supplemented by hand searching of journals, personal contacts, and citation and pearl searching of key papers. No specific date limits were applied to the search, due to a lack of clarity as to when trials first started to use the internet as a tool, and also due to the variable rate of development of different internet technologies.

Health and well-being research and randomised controlled trials are international in scope and relevance, and therefore no language limits were applied. However, reports of research in languages other than English were only included if an English abstract was available.

Search terms were derived from MeSH, EMBASE, and other relevant taxonomies; from core articles identified from scoping searches; and by experts in the field. A process of text word analysis was undertaken, and the identified relevant thesaurus and free-text terms, and combinations of terms, tested where possible. As previously reported, there is a very wide variation both in terminology, definition and application of internet-based technologies within the research environment, resulting in a complex set of activities which are not well controlled in terms of taxonomy or indexing.

Internet-based technologies. The term 'internet' itself represents a poorly controlled concept, and as use of the technology itself has increased over time, the term has become almost ubiquitously used in study reports. For example, the terms "web" and "online" are frequently retrieved as text words due to statements such as "we searched the web for xx". Some of the terms that were tested retrieved very large numbers of irrelevant results, for example, using "e-" as a prefix to terms such as "e-research" retrieves examples such as "Vitamin E". Where possible, controlled vocabularies were used; for instance, the MeSH term "internet" was created in 1999, with the previous indexing term of "computer communication networks" becoming a broader term. All subheadings of "internet/" were included to increase sensitivity.

Randomised controlled trials. Relevant search filters for randomised controlled trials were identified for MEDLINE, EMBASE, PsycINFO and CINAHL. In order to ensure that a systematic approach was taken to the choice of filter, a literature search was undertaken to identify what was known about specific filter characteristics and performance.[68–71] Poor and inconsistent indexing within and between databases made it difficult to compare the performance of specific terms and filters. This meant that the choice of methodological filter to identify reports of randomised controlled trials would have a high impact on the number of retrieved records. Test searches were run to estimate likely outcomes which resulted in poor performance in terms of the precision of the tested filters, and the very broad and sensitive nature of the other concept terms of the search led to the potential retrieval of a very large number of results. Therefore although the preference was to use a highly sensitive search strategy in order to increase sensitivity, a pragmatic approach had to be taken to achieve maximum balance between sensitivity and precision.

The initial database searches were run in January 2011, with further update searches completed in September 2013, and January 2015. The combined searches identified 21,813 results, which were downloaded to EndNote, and then transferred to the SENTE software package for screening and indexing. The results were then screened using the inclusion and exclusion criteria. Results were initially ordered by year of publication, to identify how the number had increased over time.

Selection of studies

Stages of the trial process. The database was developed to identify at which stage of the trial process internet technologies have been used. This would include use of the internet to gain informed consent, self-enrolment, the delivery of the intervention, data collection and analysis, and outcome measurement. Given that the use of internet tools for trial registration, randomisation, publicity and promotion, or dissemination has become standard practice in most clinical trials, reports of studies where the only reference to the use of the internet is related to these concepts were excluded.

Textbox 1. Inclusion and exclusion criteria.

All studies that meet the following criteria were included:

- randomised controlled trials using internet-based technologies in the trial process
- studies using mobile technologies where there is also internet-based activity
- studies investigating health research, public health research topics and settings
- studies that include educational or behavioural interventions of health and well-being topics
- Studies involving patients or members of the public

The following studies were excluded:

- Studies using mobile telecommunications technologies exclusively, with no internet-based content
- Studies investigating interventions in social care or educational settings, where the main topic of investigation is not health related, or where clinical diagnoses are not included in ICD10
- Studies where participants were health professionals or students

Iterative classification and labelling. The retrieved studies were then indexed to record their core characteristics, including the country the trial took place in, the conditions being investigated, and the interventions being tested. Initially studies were indexed using the terms applied by the source database, such as MeSH. Given that the source databases use a range of different controlled vocabularies, and the lack of consistency of indexing within databases, a pragmatic, hybrid system was developed. This entailed a two-stage process, where the condition and intervention indexing terms were extracted from the relevant field in each study report, using the classification label provided by the source database. This was then mapped against the existing set of codes, and an iterative process. If a relevant code was already in the structure, this was applied. If no relevant code was available, a new code was created using the terminology from the study report in question. When the indexing was completed, the results were scanned for consistency and re-labelled where appropriate. As it was envisaged that detailed analysis within specific condition or intervention areas would be undertaken by topic experts as part of secondary projects, no further work on the classification was undertaken at this stage.

Results

The current number of identified internet-based randomised controlled trials in the database as of January 2015 is 1992.

Identified studies

Date of publication. The number of reports of internet-based randomised controlled trials retrieved and added to the database by year of publication increased as expected over time, with the majority published between 2007 and 2015 (Figure 1).

Geography. The largest number of trials reported as using internet technologies were conducted in the USA (49.7%), followed by The Netherlands (10.2%), Australia (8.5%), the United Kingdom (5.8%), Sweden (4.6%), Canada (4%), and Germany (2.6%). South Korea (1.5%) has the highest number of reported trials for other continents.

Topic. The topics most often studied were those reporting interventions relating to non-communicable conditions such as obesity, smoking and substance abuse, depression, anxiety and phobias. The second largest group contained conditions where self-management is already a major feature of clinical care, for example in diabetes and cardiovascular conditions. The fastest

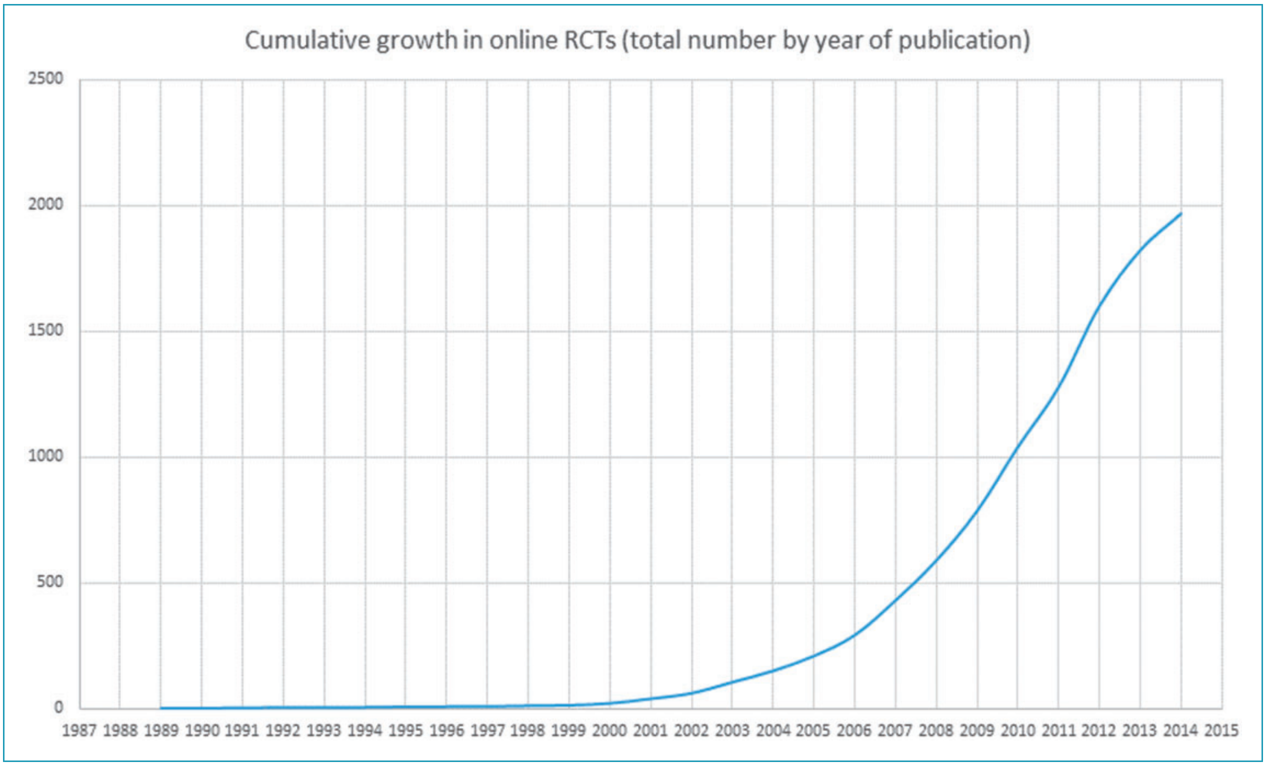

Figure 1. Cumulative growth in internet-based RCTs.

growing fields of interest in the period between 2010 and 2014 were in treatments for depressive disorders and obesity (Table 1).

Intervention. The highest numbers of concepts indexed for study interventions related closely to the nature of the conditions, as reported previously; for example, behaviour change therapies, such as cognitive behavioural therapy for anxiety and depression; smoking cessation interventions, and interventions to increase physical activity (Table 1). The ranked concepts reflect the most specific terms as indexed by the source database, and are therefore an indication of the intervention and not grouped by permuted index (Table 2).

Table 1. Highest ranked topics or condition concepts.

	No of trials	% of included trials
Topic		
depression	218	10.9
obesity	171	8.6
cancer	162	8.1
alcohol abuse	154	7.7
diabetes	133	6.7
cardiovascular diseases	131	6.6
anxiety	111	5.6
stress	79	4
phobias	49	2.5

Discussion

A number of key themes have emerged from the initial analysis of the study reports that have been included in ORCHID. These themes reflect the rapid growth of the use of internet technologies in health and well-being research, but also raise issues which require further investigation.

Location of trials

It is clear from the analysis of the trial location mapping that a high percentage of internet-based trials were conducted in a small number of countries. This may reflect existing publication bias in the scientific literature, and in the selection of journals indexed in bibliographic databases, previously found in studies of country and language of publication, such as in MEDLINE, where in a 60-year period the USA

Table 2. Highest ranked intervention concepts as indexed.

	No of trials	% of included trials
Intervention		
Psychotherapy	674	33.8
Behaviour change	613	30.8
Education	398	20
Cognitive behavioural therapy	376	18.9
Health promotion	335	16.8
Self-management	331	16.6
Communication	214	10.7
Physical activity	204	10.2
Health information/literacy	167	8.4
Monitoring	162	8.1
Social support	157	7.9
Feedback	156	7.8
Weight loss	155	7.8
Web-based care	135	
Risk assessment	130	
Smoking cessation	117	
Counselling	109	
Nutrition	105	

(22.2%) and United Kingdom (12.5%) produced most publications reporting randomised controlled trials.[72] Of 468,191 journal articles added to MEDLINE in 2000, 67.9% were initiated in the 'Anglo' countries.[73] This is confirmed in some respects from the ORCHID data, although we have observed a trend where a small number of key centres of excellence have developed, with expertise in the conduct of specific types of internet-based clinical trials. This may reflect the interest and publications of one research group or researcher, who publish multiple studies on either the same set of trials or sequences of trials on similar interventions, and may skew the position of that country in the overall table.

Expertise in the conduct and evaluation of internet-based clinical trials has developed in small pockets, and this learning could be shared to benefit others in terms of the methodological and economic advantages. Improvements in the ability to recruit and retain trial participants could help in a number of environments, including large-scale public health trials, and in particular in the availability and conduct of research in underdeveloped countries. At this moment in time research using these technologies tends to favour the diseases and conditions of high-income countries.

Type of topic and intervention

There is clearly a predominance of results in two main topic areas. These are those relating to modifiable risk factors for preventable causes of death and illness that impact on public health priorities, such as tobacco, obesity and lack of physical activity, and alcohol and substance abuse; and mental health concerns such as depressive disorders, anxiety and phobias.

It is not surprising, then, that the most common interventions in the database are those related to existing treatment options for these conditions, and where the delivery of the intervention transfers more easily to an internet-based mode of delivery, such as those for weight loss and smoking cessation programmes, talking therapies and physical activity interventions. It is not known how much the nature of the technology may pre-dispose to a particular choice of intervention.

To improve the relevance of treatment options being tested, future research in partnership with patients and the public could be based on uncertainties and the most important questions for patients, carers and practitioners. This could ensure that the technological options under consideration are driven by research need rather than by researcher preference. It is not clear whether the predominant topic areas reflect the most relevant use for internet-based technologies, or whether those particular researchers bring their interest in the technology to bear on their subject through a preference or interest in that technology.

Stage of trial

The majority of internet applications represented in ORCHID are used in the delivery of the intervention. Some 669 studies were indexed as reporting use of the internet for data collection, and 387 indexed as self-reporting. A large number of the most popular interventions use the internet to enable participants to self-report data, but there is not yet a body of trials using this in non-behavioural trials. Sixty-four studies were indexed as reporting use of the internet for data analysis or follow-up. Very few of the studies identified use internet technologies in other stages of the trial process. Some 176 studies were retrieved that reported on the methodological issues related to internet-based trials. These were not entered into the database, but were incorporated into work to conduct a methodological systematic review. Further work will be undertaken by ThinkWell researchers to investigate specific internet applications within these trials, and to develop more specific taxonomies.

Limitations

The large number of results from the US, UK and Australia, and therefore the predominance in English-speaking, high-income countries, may be due to publication bias in terms of reported studies, journal indexing and choice of databases searched. The issue of resource allocation must be considered as other countries may experience bandwidth limitations, limited access to the technology required for online trials or reduced access to the internet or internet-wired devices. Other considerations such as research capacity, governance, online and data security, and technological capacity will also need to be in place to fully exploit internet-based trial potential. ThinkWell is actively seeking partnerships with researchers in low-income countries in order to collaborate and promote knowledge transfer.

A major challenge of the project was to develop and pilot a set of search terms, phrases, and combinations in order to retrieve references to, or about, internet-based clinical controlled trials. Indexing in some non-medical databases is poor and uncontrolled, leading to retrieval problems and the need for excessive amounts of screening and filtering due to poor reporting of studies and abstracts. Although comprehensive search strategies and filters were developed as part of this project, the changing nature of the terminology, which is in line with the pace of socio-technical change, requires an exponential list of free-text search terms in the absence of recognised controlled vocabularies and ontologies.

Given the need to retrieve as many relevant studies as possible, the search strategy needed to be highly sensitive, but when combined with the poorly controlled terms for the 'internet' concept, this led to extremely large numbers of results with very poor precision. Because of this, the range and scope of databases used was restricted, and made it difficult to determine the impact of not using validated filters.

A number of different approaches to the description of terms were identified, including the development of taxonomies on behaviour change,[74] regular updates of MeSH,[75] the work of individual review groups or researchers, and organisations such as WHO[76] and the National Institute of Health and Care Excellence (NICE).[77] Work is underway to improve the reporting of interventions,[78] and this research on the identification of internet-based trials could feed into the better co-ordination of approaches to taxonomies and labelling. Standardisation in indexing methodology could complement transparency in research and lead to more effective searching and reporting of internet-based clinical trials research.

Definition of concept of 'health and well-being'

Health and well-being is a very wide concept. Using the ICD categories was a pragmatic way of limiting the results, for instance in relation to the concept of addiction, which was restricted to drug/alcohol but excluded internet addiction, as the primary focus of the project was on health care and public health settings. This may have resulted in missed trials.

Indexing and coding

Similar limitations were experienced in the terms used for coding and indexing included studies. A search of the literature revealed a number of potential alternative systems for the coding of studies, particularly in the area of behaviour change interventions. Grimshaw et al.[79] draw attention to four Cochrane Reviews that cover interactive health communication applications; interventions to enhance medication adherence; contracts; and new methods of communication. Within the original cited reviews, specific elements of existing controlled terms, consistent with MeSH and other major taxonomies, are used in different combinations, including social support, decision support, and behaviour change support. Haynes et al.[80] identify a range of interventions used within trials for medication adherence, including instruction, counselling, family interventions, mentoring, etc. This approach, although ensuring consistency with the original decisions made by indexers, does not reflect consistency across all the studies, and therefore the categories tabled may include a range of intervention headings and classifications that are not comparable.

The mapping relied on the terms used by authors to report their studies, and on the indexing terms chosen by the host database. No quality appraisal of individual studies was made, other than that of determining whether the study was a randomised controlled trial by looking for reports of randomisation.

Conclusions

There is clearly a need to develop communities of practice to share learning and expertise in internet-based clinical trials, and to ensure that the public-led aspects embedded in the nature of the technologies is harnessed. This will be one of the major aims of the ThinkWell and PLOT-IT initiatives. The original purpose for creating ORCHID was as an internal resource for ThinkWell; however, as interest from other researchers was high, we decided to continue to add results after the initial searches in 2011 and to publish the details of ORCHID's methods and findings to date. Interest from external research teams has led to collaboration on other projects, including a text-mining project.

To increase public engagement and to inform future internet-based clinical trials, the ThinkWell community will embed participatory processes in the future development of the database, the PLOT-IT platform[81] and the harvesting and prioritisation of research priorities. A core group of citizens is being recruited to help plan this process, and carry out a social network analysis. A two-stage systematic review is being undertaken to map what is known about internet-based clinical trials, and to synthesise specific aspects of the known uncertainties, including questions about ethics, methodology, and qualitative perspectives.[82] Qualitative work is also underway to explore both researcher and public experience in internet-based clinical trials, which will help to identify additional learning and support tools for all those involved in this area.

ORCHID has the potential to be a useful resource for researchers to mine data from the included studies, and will also be expanded to include a repository of what is known about the conduct of internet-based trials. Future development is planned to enable open access to this resource, to support the development of a community of practice which any researcher or member of the public can use. It is currently used as a resource to identify key terms for search strategies, and to further determine trends in this rapidly developing area. Regular updates are projected, and additional sources will be explored to increase comprehensive coverage, and to enable ThinkWell to expand the progression of public-led research.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of interest: None declared.

Contributorship: ABu conceived the study, and advised on the protocol. ABr developed the full protocol, conducted the searches and developed the database structure. ABr developed and applied the coding system and carried out the preliminary analysis. ABu and AP contributed to the data analysis and discussion. ABr wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Guarantor: AB

Ethical approval: Not applicable

Peer-review: This manuscript was reviewed by Sarah Lim Choi Keung, University of Warwick and one other reviewer who wishes to remain anonymous.

Acknowledgements

We would like to thank Jose Emperanza and Juan Cabello for their assistance and guidance in this research; the information scientists at the Donostia Unibertsitate Ospitaleko Liburutegia, Eukene Ansuategi and Marimar Ubeda, for their help in searching the LILACS database, and for testing search filters; Nia Roberts, Information Scientist at the Department of Primary Health Care Sciences, University of Oxford, and Nicola Pearce-Smith, Senior Information Scientist for the NHS National Knowledge Services, for advice and review of the original search strategies; Ally Naughten, CTU, Department of Primary Health Care Sciences, University of Oxford, for help in determining a pragmatic approach to categorising the stages of a clinical trial.

References

1. ThinkWell: The international network for knowledge about wellbeing, http://www.ithinkwell.org/ (2015, accessed 17 April 2015).
2. Elwyn G, Frosch D, Thomson R, et al. Shared decision making: a model for clinical practice. *J Gen Intern Med* 2012; 27: 1361–1367.
3. Edwards AGK, Naik G, Ahmed H, et al. Personalised risk communication for informed decision making about taking screening tests. *Cochrane Database Syst Rev* 2013; 2: CD001865.
4. Stacey D, Légaré F, Col NF, et al. Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev* 2014; 1: CD001431.
5. Coulter A. Making shared decision-making a reality: No decision about me, without me, http://www.kingsfund.org.uk/publications/making-shared-decision-making-reality (2011, accessed 17 April 2015).
6. Corrigan K. Number of consumer health apps to reach 13,000 by next summer, http://karencorrigan.blogspot.co.uk/2011/09/number-of-consumer-health-apps-to-reach.html (2011, accessed 5 April 2015).
7. Eysenbach G, Powell J, Englesakis M, et al. Health related virtual communities and electronic support groups: Systematic review of the effects of online peer to peer interactions. *BMJ* 2004; 328: 1166.
8. Webb T, Joseph J, Yardley L, et al. Using the internet to promote health behavior change: A systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J Med Internet Res* 2010; 12: e4.
9. Fox S and Duggan M. Health Online 2013: One in three American adults have gone online to figure out a medical condition, http://www.pewinternet.org/2013/01/15/health-online-2013/ (2013, accessed 19 April 2015).
10. Department of Health. Change for Life, http://www.nhs.uk/change4life/Pages/change-for-life.aspx (2015, accessed 4 June 2015).
11. Department of Health. NHS Choices, http://www.nhs.uk/Pages/HomePage.aspx (2015 accessed 4 June 2015).
12. Department of Health. http://digital.innovation.nhs.uk/pg/dashboard (2015, accessed 17 April 2015).
13. Wyatt JC. Digital healthcare: When, for whom, and at what cost? http://www.openforumevents.co.uk/digital-healthcare-cost/ (2013, accessed 5 June 2015).
14. Eysenbach G and Kohler C. Does the internet harm health? Database of adverse events related to the internet has been set up. *BMJ* 2002; 324: 239.

15. Kiley R. Does the internet harm health? Some evidence exists that the internet does harm health. *BMJ* 2002; 324: 238–239.
16. Van Deursen AJ. Internet skill-related problems in accessing online health information. *Int J Med Inform* 2012; 81: 61–72.
17. Farmer SEJ, Bernardotto M and Singh V. How good is Internet self-diagnosis of ENT symptoms using Boots WebMD symptom checker? *Clin Otolaryngol* 2011; 36: 517–518.
18. Neter E and Brainin E. eHealth Literacy: Extending the digital divide to the realm of health information. *J Med Internet Res* 2012; 14: e19.
19. Viswanath K and Ackerson LK. Race, Ethnicity, language, social class, and health communication inequalities: A nationally-representative cross-sectional study. *PLoS One* 2011; 6: e14550.
20. Pak R, Price M and Thatcher J. Age-sensitive design of online health information: Comparative usability study. *J Med Internet Res* 2009; 11: e45.
21. Han K, Shih PC, Rosson MB, et al. Understanding local community attachment, engagement and social support networks mediated by mobile technology. *Interact Comput* 2014; Nov 17, DOI:10.1093/iwc/iwu040 (online).
22. O'Mara-Eves A, Brunton G, McDaid D, et al. Community engagement to reduce inequalities in health: A systematic review, meta-analysis and economic analysis. *Public Health Res* 2013; 1: i–525.
23. Ziebland S. Why listening to health care users really matters. *J Health Serv Res Policy* 2012; 17: 68–69.
24. Coleman A, Checkland K, Mcdermott I, et al. Patient and public involvement in the restructured NHS. *J Integr Care* 2011; 19: 30–36.
25. Mockford C, Staniszewski S, Griffiths F, et al. The impact of patient and public involvement on UK NHS health care: A systematic review. *Int J Qual Health* 2012; 24: 28–38.
26. Department of Health. Patient and public involvement in health: The evidence for policy implementation, http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4082334.pdf (2004, accessed 17 April 2015).
27. Barham L. Public and patient involvement at the UK National Institute for Health and Clinical Excellence. *Patient* 2011; 4: 1–10.
28. Ocloo JE and Fulop NJ. Developing a 'critical' approach to patient and public involvement in patient safety in the NHS: Learning lessons from other parts of the public sector? *Health Expect* 2012; 15: 424–432.
29. Oliver S. Public involvement in research: Making sense of the diversity. *J Health Serv Res Policy* 2015; 20: 45.
30. Chalmers I and Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet* 2009; 374: 86–89.
31. Hibbard JH and Greene J. What the evidence shows about patient activation: Better health outcomes and care experiences; fewer data on costs. *Health Aff* 2013; 32: 207–214.
32. Stewart R and Oliver S. A systematic map of studies of patients' and clinicians' research priorities, http://www.lindalliance.org/pdfs/JLA%20Internal%20Reports/090712_JLAreport_RS_map_studies_PPI%20&%20Clinicians_research%20priorities_with%20appendices.pdf (2008, accessed 6 June 2015).
33. Moran R and Davidson P. An uneven spread: A review of public involvement in the National Institute of Health research's health technology assessment program. *Int J Technol Assess Health Care* 2011; 27: 343–347.
34. Partridge N and Scadding J. The James Lind Alliance: Patients and clinicians should jointly identify their priorities for clinical trials. *Lancet* 2004; 364: 1923–1924.
35. Testing Treatments Interactive, http://www.testingtreatments.org/ (2013, accessed 17 April 2015).
36. Tallon D, Chard J and Dieppe P. Relation between agendas of the research community and the research consumer. *Lancet* 2000; 355: 2037–2040.
37. Chalmers I. What do I want from health research and researchers when I am a patient? *BMJ* 1995; 310: 1315–1318.
38. INVOLVE. Guidance on the use of social media to actively involve people in research, http://www.invo.org.uk/wp-content/uploads/2014/11/9982-Social-Media-Guide-WEB.pdf (2014, accessed 19 April 2015).
39. Campbell MK, Snowdon C, Francis D, et al. Recruitment to randomised trials: Strategies for trial enrolment and participation study: The STEPS study. *Health Technol Assess* 2007; 11: 1–126.
40. Freiman JA, Chalmers TC, Smith H, et al. The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial. Survey of 71 "negative" trials. *N Engl J Med* 1978; 299: 690–694.
41. Rose D and Williamson T. Developing the evidence base of patient and public involvement in health and social care research: The case for measuring impact. *Int J Consum Stud* 2011; 35: 628–632.
42. Hazell L and Shakir SA. Under-reporting of adverse drug reactions: A systematic review. *Drug Saf* 2006; 29: 385–396.
43. Mello MM, Clarridge BR and Studdert DM. Academic medical centers' standards for clinical-trial agreements with industry. *N Engl J Med* 2005; 352: 2202–2210.
44. Trouiller P, Torreele E, Olliaro P, et al. Drugs for neglected diseases: A failure of the market and a public health failure? *Trop Med Int Health* 2001; 6: 945–951.
45. World Health Organization. Public-private partnerships for neglected diseases: Opportunities to address pharmaceutical gaps for neglected diseases. In: *Priority medicines for Europe and the world project "a public health approach to innovation"*. Chapter 8.1, http://mednet3.who.int/prioritymeds/report/index.htm (2005, accessed 18th December 2006).
46. The EQUATOR network: enhancing the quality and transparency of health research, http://www.equator-network.org/ (2015, accessed 19 April 2015).
47. Department of Health. The NHS Constitution, https://www.gov.uk/government/uploads/system/uploads/

attachment_data/file/170656/NHS_Constitution.pdf (2013, accessed 17 April 2015).

48. Jones CW, Handler L, Crowell KE, et al. Non-publication of large randomized clinical trials: Cross sectional analysis. *BMJ* 2013; 347: f6104.
49. Kelly MA and Oldham J. The Internet and randomised controlled trials. *Int J Med Inform* 1997; 47: 91–99.
50. Paul J, Seib R and Prescott T. The internet and clinical trials: Background, online resources, examples and issues. *J Med Internet Res* 2005; 7: e5.
51. Murray E. Methodological challenges in online trials. *J Med Internet Res* 2009; 1192: e9.
52. Hoffmann T, English T and Glasziou P. Reporting of interventions in randomised trials: An audit of journal Instructions to Authors. *Trials* 2014; 15: 20.
53. Lorencatto F, West R, Stavri Z, et al. How well is the intervention content described in published reports of smoking cessation interventions? *Nicotine Tob Res* 2013; 15: 1274–1282.
54. Newton NC, Andrews G, Teesson M, et al. Delivering prevention for alcohol and cannabis using the internet: A cluster randomised controlled trial. *Prev Med* 2009; 48: 579–584.
55. Kypri K, Hallett J, Howat P, et al. Randomized controlled trial of proactive web-based alcohol screening and brief intervention for university students. *Arch Intern Med* 2009; 169: 1508–1511.
56. Graham AL, Cobb NK, Papandonatos GD, et al. A randomized trial of internet and telephone treatment for smoking cessation. *Arch Int Med* 2011; 171: 46–53.
57. Almeida FA, You W, Harden SM, et al. Effectiveness of a worksite-based weight loss randomized controlled trial: The worksite study. *Obesity* 2015; 23: 737–745.
58. Powell J, Hamborg T, Stallard N, et al. Effectiveness of a web-based cognitive-behavioral tool to improve mental well-being in the general population: Randomized controlled trial. *J Med Internet Res* 2013; 15: e2.
59. Broekhuizen K, Kroeze W, van Poppel MN, et al. A systematic review of randomized controlled trials on the effectiveness of computer-tailored physical activity and dietary behavior promotion programs: An update. *Ann Behav Med* 2012; 44: 259–286.
60. Civljak M, Stead LF, Hartmann-Boyce J, et al. Internet-based interventions for smoking cessation. *Cochrane Database Syst Rev* 2013; 7: CD007078.
61. Murray E, Burns J, See Tai S, et al. Interactive health communication applications for people with chronic disease. *Cochrane Database Syst Rev* 2005; 4: CD004274.
62. Shahab L and McEwen A. Online support for smoking cessation: A systematic review of the literature. *Addiction* 2009; 104: 1792–1804.
63. NICE. Computerized cognitive behaviour therapy for depression and anxiety. Technology Appraisal 97, https://www.nice.org.uk/guidance/ta97/chapter/about-this-guidance (2006, accessed 20 April 2015).
64. UK DUETS, http://www.library.nhs.uk/duets/ (2015, accessed 6 June 2015).
65. Fenton M, Brice A and Chalmers I. Harvesting and publishing patients' unanswered questions about the effects of treatments, http://www.lindalliance.org/pdfs/Publications/Fenton,Brice,Chalmers%20UK%20DUETs%20Harvesting%20uncertainties.pdf (2009, accessed 6 June 2015).
66. Brice A and Burls A. Working with uncertainty: The role of DUETs in harvesting what we don't know, https://www.researchgate.net/publication/275339762_Working_with_uncertainty_the_role_of_DUETs_in_harvesting_what_we_dont_know (2008, accessed 6 June 2015).
67. Department of Health. No health without mental health: A cross-government mental health outcomes strategy for people of all ages, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213761/dh_124058.pdf (2011, accessed 17 April 2015).
68. The InterTASC Information Specialists' Sub-Group Search Filter Resource, https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home (Accessed 17 April 2015; archived by WebCite® at http://www.webcitation.org/6Xrl1UOHA).
69. Wong SS, Wilczynski NL and Haynes RB. Optimal CINAHL search strategies for identifying therapy studies and review articles. *J Nurs Scholarsh* 2006; 38: 194–199.
70. Eady AM, Wilczynski NL and Haynes RB. PsycINFO search strategies identified methodologically sound therapy studies and review articles for use by clinicians and researchers. *J Clin Epidemiol* 2008; 61: 34–40.
71. Sharon L, Wong SS, Wilczynski NL, et al. Developing optimal search strategies for detecting clinically sound treatment studies in EMBASE. *J Med Libr Assoc* 2006; 94: 41–47.
72. Gluud C and Nikolova D. Likely country of origin in publications on randomised controlled trials and controlled clinical trials during the last 60 years. *Trials* 2007; 8: 7.
73. Loria A and Arroyo P. Language and country preponderance trends in MEDLINE and its causes. *J Med Libr Assoc* 2005; 93: 381–385.
74. Michie S, Ashford S, Falko F, et al. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: The CALO-RE taxonomy. *Psychol Health* 2011; 26: 1479–1498.
75. US National Library of Medicine. MeSH: Medical Subject Headings, http://www.nlm.nih.gov/mesh/MBrowser.html (2015, accessed 20 April 2015).
76. World Health Organisation. A Taxonomy for Patient Safety, http://www.who.int/patientsafety/implementation/taxonomy/en/ (2015, accessed 20 April 2015).
77. NICE taxonomy, http://www.nice.org.uk/proxy/?sourceUrl=http%3a%2f%2fwww.nice.org.uk%2faboutnice%2fnicewebsitedevelopment%2fstandards%2fmetadata%2fniceencodingschemes%2fnicetaxonomysubjectencodingscheme%2fnice_taxonomy_subject_encoding_scheme.jsp%3fdomedia%3d1%26mid%3d0C2 (2011, accessed 2015).
78. Hoffmann T, English T and Glasziou P. Reporting of interventions in randomised trials: An audit of journal Instructions to Authors. *Trials* 2014; 15: 20.
79. Grimshaw JM, Eccles MP, Lavis JN, et al. Knowledge translation of research findings. *Implement Sci* 2012; 7: 50.

80. Haynes RB, Ackloo E, Sahota N, et al. Interventions for enhancing medication adherence. *Cochrane Database Syst Rev* 2008; 16: CD000011.
81. Brice A, Price A and Burls A. PLOT-IT: Public empowerment and engagement in clinical trials, http://www.medicine20congress.com/ocs/index.php/med/med2014b/paper/view/2534 (2014, Accessed 18 April 2015).
82. Brice A, Price A, Burls A, et al. A protocol for a systematic review of methodological and qualitative factors that might influence the effectiveness of internet-based controlled clinical trials: Evidence mapping and syntheses, http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014013308#.VTKvoVxEgy5 (2014, accessed: 26 June 2015).

Appendix 1

Medline – Ovid Medline 1948 – January Week 1 2011 Update 13/01/13

1. exp Computer Communication Networks/
2. (internet$ or web$ or online or on-line).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept, rare disease supplementary concept, unique identifier]
3. 1 or 2
4. randomized controlled trial.pt.
5. controlled clinical trial.pt.
6. randomized.ab.
7. placebo.ab.
8. clinical trials as topic.sh.
9. randomly.ab.
10. trial.ti.
11. 4 or 5 or 6 or 7 or 8 or 9 or 10
12. exp animals/not humans.sh.
13. 11 not 12
14. 3 and 13

Embase: Ovid Embase 1980 – 2011 Week 02: 18/1/11 Update 13/01/13

1. (internet$ or online or on-line).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]
2. random$.tw. or placebo$.mp. or double-blind$.tw.
3. 1 and 2

PyscINFO: 1987 to January Week 1 2011

1. (internet$ or web$ or online or on-line).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
2. (random$ adj assigned).tw.
3. double-blind.tw.
4. control.tw.
5. 2 or 3 or 4
6. 1 and 5

Cinahl: EBSCO 01 Feb 2011

1. (MH "Internet+")
2. web* OR internet* OR online
3. s1 OR s2
4. PT clinical trial
5. MH "Treatment Outcomes+"
6. TX randomized
7. s4 OR s5 OR s6
8. s3 AND s7

ERIC: ProQuest 03 Feb 2011

tbc

Pedro

1. Internet OR web OR online in Abstract or Title AND Method: clinical trial

OT Seeker

1. internet OR web OR online AND Method: Randomised Controlled Trial