Hunting zebra: retrieval of rare disease clinical guidelines

Authors: Manson, Jan† j.manson@nhs.net
Hilton Boon, Michele†
Ritchie, Karen†

1 Healthcare Improvement Scotland, Glasgow, Scotland.

Abstract

Introduction

RARE-Bestpractices (http://www.rarebestpractices.eu/) is a European Union 7th Framework Programme funded project aiming to improve management of rare diseases through facilitating the creation and dissemination of best practice guidelines. A key objective of the project is to create an online database for guideline documents and research recommendations. This paper presents work carried out to identify existing sources of rare disease guidelines

Aim

To investigate the recall and yield of rare disease guidelines from recognised guideline information resources.

Method

A literature search was carried out to identify any current practice in the area of rare disease guideline retrieval. A purposive sample of three rare conditions was used to test the retrieval performance of guideline databases, specialist rare disease resources, primary literature and internet searching.

Results

The guideline databases retrieved 1 of the total 15 guideline documents identified. Rare disease specialist resources provided a low number of highly specific results, the majority of which were not in English. Primary literature and internet searching retrieved the highest number of relevant documents with respective percentage yields of 0-11.8 and 0-7.

Conclusions

Primary literature and internet searching appear to be the most sensitive methods of retrieving rare disease guideline documents although percentage yield is very low. The time and resource required to identify relevant documents indicates a need for the development of a rare disease guideline database as will be delivered by the RARE-Bestpractices project.

Keywords: Rare Diseases; Guidelines as Topic; Information Storage and Retrieval; Databases as Topic; Library Collection Development
Introduction

Medical students and clinicians are encouraged to consider the most common diseases when diagnosing patient symptoms – when you hear hooves, think horses. This may hold true in the majority of cases but occasionally the horse turns out to be a zebra; a rare disease.

Rare diseases are defined in Europe as conditions which affect less than 5 in 10000 persons. There are more than 5000 recognised rare disorders which affect approximately 30 million people across Europe(1). Individual diseases may affect small numbers of the population but, collectively, rare disease contributes a significant burden to health and health care resources.

The very nature of rare conditions means that provision of and access to relevant health information for specific diseases can be severely limited. RARE Bestpractices(2) is a European Union Seventh Framework Programme funded project which aims to deliver an online platform to facilitate improved rare disease management through the exchange of high quality knowledge and information. Creation and dissemination of rare condition guidelines is a key aspect of the project; a suite of resources will be made available inform and encourage best practice and collaboration in guideline development and a new database will be created to host existing rare disease guidelines and research recommendations.

This paper describes work carried out to develop a search protocol suitable for the collection of rare disease guideline documents in the context of the RARE-Bestpractices project. The emphasis of this report is the recall and yield testing of known guideline information resources for rare disease topics.

Literature search

MEDLINE, Library Information Science and Technology Abstracts (LISTA) and Emerald databases were searched to identify any existing protocols, methods or research work specific to the retrieval of rare disease guideline documents. Searches were limited to English language material published from 2003 to August 2013. The MEDLINE search strategy used is listed in appendix I.

233 search results were retrieved and sifted for relevance. Four papers addressed aspects of rare disease information retrieval including the use of MeSH headings (3), rare disease diagnosis tools(4) and clinician’s information seeking behaviour(5).

No papers described retrieval methods for rare disease guideline documents.

Guideline search protocol

In the absence of existing methodology, guideline search protocols already in use by Healthcare Improvement Scotland (HIS) were utilised and tested for efficiency in identifying rare disease documents.
HIS incorporates the work of a variety of health care groups to provide evidence, improvement and scrutiny services which support evidence based patient centred health care in Scotland. The umbrella HIS organisation includes the Scottish Intercollegiate Guidelines Network (SIGN) and a Knowledge and Information Unit which both routinely search for published guidelines during the course of their project development.

SIGN investigates the work of other international guideline developers during the commissioning phase for new topics. To do so, a short protocol of guideline resources is used to search for key documents on the health condition under consideration. The Knowledge and Information Unit supports the development of standards and quality indicators based on best current evidence and clinical knowledge. The unit carries out comprehensive searching of potential guideline sources in order to maximise retrieval of available publications which may inform their work.

A composite guideline search protocol consisting of resources from the SIGN and the Knowledge and Information Units existing lists was taken forward for testing on rare disease topics.

This composite protocol can be divided into distinct resource groups:

**Guideline databases**

**G-I-N**

The Guidelines International Network (G-I-N) is a global network supporting the development and dissemination of evidence based clinical guidelines(6). The G-I-N website hosts a library of over 6500 international guidelines produced by G-I-N member organisations.

**National Guideline Clearinghouse**

Administered by the US Agency for Healthcare Research and Quality, the National Guideline Clearinghouse (NGC)(7) provides a number of guideline related resources including detailed summaries of more than 2500 published clinical guideline documents. Potential content of the database is identified through literature searching/guideline resource audit and through self submission by guideline developers. Guidelines documents are required to meet a minimum set of criteria in order to be considered for inclusion in the database(8).

**NICE Evidence search(9)**

The National Institute for Health and Care Excellence (NICE) hosted evidence resource search tool provides access to health care guidance documents and systematic reviews. The method of identifying documents for inclusion to this resource is unclear.

**Trip(10)**

Trip (Turning research into practice) is a database containing a wide variety of evidence resources, including guidelines, which aims to facilitate rapid access to all types of clinical
evidence materials through a single platform. Content selection is based on user and database creator recommendation and includes the G-I-N and NGC databases as well as many of the national and clinician guideline development groups considered in the HIS search protocol.

**National/regional guideline producers**

English language national or regional guideline producers representing the main governmental and health service producers for Scotland, England, Ireland, Canada, Australia and New Zealand; the work of these organisations influences project development in Healthcare Improvement Scotland.

**Clinician groups**

Organisations such as the UK Royal Colleges which produce clinical guidance documents in order to support best clinical practice in their field of expertise.

**Additional resources**

The search protocol also includes primary literature resources (MEDLINE and EMBASE) and a Google internet search limited to the first 100 results available in PDF document format.

**Modification of the protocol**

**Trip**

Although there is clear advantage to multi resource searching from a single point of access, as is offered by Trip, the objective of our work was to assess the value of specific individual information sources in the retrieval of rare disease guidelines. Trip was therefore removed from the test protocol.

**PubMed**

Healthcare Improvement Scotland accesses MEDLINE and EMBASE databases via a subscription service. RARE-Bestpractices aims to maximise accessibility and inclusivity in all aspects of the project and so PubMed was selected as an alternative resource; PubMed is a open access primary literature database.

**Rare disease specific resources**

There are many established rare disease web spaces ranging from specific disease support groups to umbrella rare disease resources(11-13)which, in most cases, provide patient information or overview descriptions of disorders.

Orphanet(12) provides expert authored disease synopses and links out to other potentially relevant information sources for the majority of recognised rare disorders. In some instances the website entry for a condition will include links to guidance documents. This appears to be
at the discretion of the synopses writer and there is no indication of guideline document retrieval methods or inclusion criteria. None the less, Orphanet provides direct access to guidelines for named rare conditions and so was added to the protocol.

EuroGentest(12) is an EU funded project to harmonise the process of genetic testing across EuropeOne output of the project is the development of clinical utility gene cards; concise guidance delivered in a standardised format on the clinical utility of genetic testing for individual conditions. EuroGentest is not specifically a rare disease project but, as around 80% of rare conditions are associated with genetic anomalies, much of the output is useful to the rare disease community. Arguably, the gene cards themselves do not adhere to standard definitions of a guideline document(14, 15) as they contain no explicit recommendations. However, they are best practice statements and, because diagnosis has been highlighted as an important aspect of rare disease management, it was decided to include this resource in the search protocol.

A list of all resources included in the test protocol is provided in appendix II.

**Rare disease topic selection**

As previously noted, there are over 6000 conditions which meet the criteria for a rare disease in Europe. These conditions vary enormously in body systems affected, strataums of the population most at risk, mode of onset, underlying cause, duration and impact on health. The resulting heterogeneity proved to be a barrier to selecting a small purposive sample of diseases. Three conditions were selected based on disease prevalence; Huntington’s Disease, Costello syndrome and Turner’s syndrome.

An ‘all disease’ list ranked by prevalence(16) was used to select a condition with a mid to high prevalence (Huntington’s Disease at 7 cases per 100 000) and an ultra-rare disorder, Costello syndrome: 300 recorded cases.

Huntington’s disease is a hereditary, progressive disorder of the central nervous system which usually develops in adulthood and can cause a wide range of neurological symptoms (Huntington’s Disease Association(17)

Costello syndrome is caused by a single gene mutation and results in multiple congenital abnormalities. Characteristics of the syndrome include short stature, distinctive facial appearance, intellectual disability, cardiac abnormalities and an increased risk of developing specific malignant tumours(18)

There are a number of genetically distinct rare conditions which have similar or overlapping phenotypical manifestations to Costello syndrome(19). This includes Turner syndrome, which a common disorder in rare disease terms (20 cases per 100 000). As such, Turner’s syndrome was added to our test disease selection as a high prevalence disease.
Search and document selection

Each resource in the search protocol provided different search functionality. The common approach was to use each disease name synonym in turn as a simple keyword search.

The PubMed primary literature search was limited using a filter for guideline documents. The Google search combined disease keywords with guideline document keywords and was limited to PDF documents.

All searches and document selection were carried out by an information scientist.

Document inclusion and exclusion criteria

A document was considered to be a guideline where it had been developed by a stakeholder group and contained recommendations for practice, reflecting key elements of the WHO and IOM guideline definitions (14, 15). Full inclusion and exclusion criteria are listed in table 1.

<table>
<thead>
<tr>
<th>Document type</th>
<th>Inclusion</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type</strong></td>
<td>Any document produced by a stakeholder group which is described as a guideline, consensus statement, or best practice statement <strong>AND</strong> contains recommendations* for practice.</td>
<td>Patient information documents. Local (e.g. hospital) policy documents. Single author publications. Systematic reviews and Health technology assessments.</td>
</tr>
<tr>
<td><strong>Year</strong></td>
<td>Published within 10 years</td>
<td></td>
</tr>
<tr>
<td><strong>Language</strong></td>
<td>English</td>
<td></td>
</tr>
<tr>
<td><strong>Topic</strong></td>
<td>Directly relating to the named condition</td>
<td>Generic symptom management e.g. dementia management. Guidance on single interventions e.g. review of a particular drug.</td>
</tr>
<tr>
<td><strong>Format</strong></td>
<td>PDF, web document, print document, journal article</td>
<td>Text book or Ebook</td>
</tr>
</tbody>
</table>

*Does not necessarily apply to diagnostic testing protocols

Table 1. Inclusion and exclusion criteria for guideline documents.

Results and discussion

Both the national/regional and clinical group guideline developer organisations included in the search protocol did not recall any potentially relevant documents for the three test conditions. Tables 2 to 4 show the recall and percentage yield of the remaining resources. These tables are reproduced from Hilton-Boon et al 2014(20)
<table>
<thead>
<tr>
<th>Source</th>
<th>Recall</th>
<th>Reasons for exclusion</th>
<th>Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G-I-N database</td>
<td>1</td>
<td>--</td>
<td>1 (100)</td>
</tr>
<tr>
<td>National Guidelines Clearinghouse</td>
<td>12</td>
<td>8 Not specific to Huntington’s</td>
<td>1 (8.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 Not relevant</td>
<td></td>
</tr>
<tr>
<td>Orphanet</td>
<td>5</td>
<td>4 Not English language</td>
<td>1 (20)</td>
</tr>
<tr>
<td>Eurogentest</td>
<td>65</td>
<td>64 Not relevant</td>
<td>1 (1.5)</td>
</tr>
<tr>
<td>PubMed</td>
<td>34</td>
<td>1 Not English language</td>
<td>4 (11.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>29 Not relevant</td>
<td></td>
</tr>
<tr>
<td>Evidence Search (NICE)</td>
<td>52</td>
<td>Not relevant</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Google pdf first 100</td>
<td>--</td>
<td>93 Not relevant</td>
<td>7 (7)</td>
</tr>
<tr>
<td>Pearl growing</td>
<td>--</td>
<td>--</td>
<td>6 (0.3%)</td>
</tr>
<tr>
<td>Total</td>
<td>269</td>
<td>De-duplication</td>
<td>11 (6.5)</td>
</tr>
</tbody>
</table>

*Table 2. Huntington’s disease guidelines retrieved, by source.*

<table>
<thead>
<tr>
<th>Source</th>
<th>Recall</th>
<th>Reasons for exclusion</th>
<th>Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G-I-N database</td>
<td>1</td>
<td>Not English language</td>
<td>0 (0)</td>
</tr>
<tr>
<td>National Guidelines Clearinghouse</td>
<td>6</td>
<td>Not relevant</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Orphanet</td>
<td>2</td>
<td>Not English language</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Eurogentest</td>
<td>1</td>
<td>Not relevant</td>
<td>0 (0)</td>
</tr>
<tr>
<td>PubMed</td>
<td>1</td>
<td>Not relevant</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Evidence Search (NICE)</td>
<td>53</td>
<td>Not relevant</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Google pdf first 100</td>
<td>--</td>
<td>99 Not relevant</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 Not a guideline</td>
<td></td>
</tr>
<tr>
<td>Pearl growing</td>
<td>--</td>
<td>--</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Total</td>
<td>64</td>
<td>De-duplication</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

*Table 3. Costello syndrome guidelines retrieved, by source.*
Table 4. Turner syndrome guidelines retrieved, by source.

| Total | 596 | De-duplication | 4 (0.7) |

Guideline databases

Of the fifteen guidelines documents identified across the three test conditions, the established guideline databases identified four (27%). Reasons for this low return might be explained by the characteristics of the databases and features typical of rare disease guideline documents.

The G-I-N database publishes guidelines produced by subscribed member organisations. Much of G-I-N’s activities involve development, streamlining and collaboration on areas of guideline methodology; a highly useful and appropriate objective for organisations working in the field of clinical guideline development. Rare disease guidelines appear to be more often produced by patient organisations or small networks of stakeholders where guideline development per se is not an area of extended interest. For these types of groups, the benefits of G-I-N membership might not justify the costs.

Clinical guidelines must meet a series of criteria in order to be included in the National Guideline Clearinghouse database(8). Guidelines must demonstrate or supply supporting documentation to show that specific methodological processes have been followed including systematic review of the literature and an explicit demonstration of translation of evidence into recommendations.

It can be argued that rare disease guidelines should be judged using the same standards of methodology demanded from those on common conditions. However rare disease guideline developers can face additional barriers to complying with conventional guideline quality standards, resulting in their omission from resources such as NGC. There may be a lack of reliable evidence on the condition of interest. Individual uncommon disorders may not be a priority for research and, even when studies are commissioned, there are often problems with recruitment and ethics. Recommendations are therefore often expert consensus based rather than evidence driven. Although it is still possible to provide records of the consensus process, it is more challenging to meet the inclusion criteria of NGC when the empirical evidence base is weak or non-existent.

Also, adhering to recognised guideline standards requires a guideline development group to have some knowledge of best practice in methodology. It is common practice for established guideline development organisations to have at least one methodologist advising on each project(21) but small rare disease networks might not easily have access to this kind of resource.

National/clinical group guideline developers

No potentially relevant documents for the three test conditions were identified from the national/regional or clinical group guideline developer organisations included in the search protocol. These organisations are usually required to prioritise resources into conditions
where there is the potential for significant health improvement at a population level. This is unlikely to include diseases affecting very small numbers of the population.

Rare disease specific resources

The structure of the Orphanet database provides a dedicated page entry for each rare condition meaning that, in the instances where guidelines are signposted, they are highly relevant to the disease under consideration. Orphanet returned at least one guideline for each condition searched. The majority of these were excluded only because they were in languages other than English.

Although EuroGentest yielded only one guideline across the three test conditions it should be noted that a Gene Utility Card for Costello syndrome is currently ‘in progress’; had this been complete at the time of searching it would have provided the only English language guidance document for this condition.

Primary literature – PubMed

PubMed performed comparatively well in our test searches both in terms of yield and number of relevant documents retrieved. Journal submission appears to be a commonly used method of publishing and disseminating rare disease guideline documents produced by clinical groups.

Internet search – Google

The percentage yield for Google remained very low (0-7%) but the proportion of the total number of relevant guidelines retrieved was higher than in any other resource (73%). In addition to all of the guidelines retrieved by Pubmed, Google was able to identify further documents made available through methods other than journal publishing. This may be an important consideration in the rare disease field where direct web publishing can be a preferred option; established patient groups which develop guidelines are likely to use their own web presence as the primary tool for guideline publishing and dissemination, small specialised networks may not have the resource to engage in the academic publishing arena.

Additional guidelines – pearlgrowing

In the Huntington’s disease example, one guideline identified via the search protocol acknowledged five further guidelines developed by the same patient organisation. These additional guidelines were not retrieved through our search protocol; they were not indexed for PubMed and, due is seems to an apparent meta-data issue, were not returned in the Google search. Although this may be an anomaly, it does suggest the usefulness of pearlgrowing methods should not be discounted when searching for rare disease material.

Limitations
This study considered a test sample of three conditions and it is possible that the resources in the protocol may perform differently with alternative conditions or according to specific aspects common to different conditions. Our disease choice was based on prevalence but the sample size is not large enough to draw any conclusions on resource performance relating to size of disease population.

Included guidelines were limited to English language. This resulted in several potentially relevant documents being excluded and an underestimation of the performance of some resources.

Searching and sifting were carried out by one person and may have introduced bias in to the document selection process.

**Conclusions**

To our knowledge, this is the first work carried out to assess the performance of guideline information resources on retrieval of rare disease guideline documents. The results of our testing demonstrate that the large established international databases do not cater particularly well for rare disease topics. Rare condition specific resources provide highly relevant but sporadic access to guidance documents. Google and PubMed appear to be the most successful resources in terms of returning the highest number of relevant guidelines although yield is extremely low.

These findings indicate that the most common methods of disseminating rare disease guidelines are through primary medical literature and the World Wide Web. There is no single source for guideline documents and searches return a large proportion of irrelevant results. This may present a barrier to clinicians and patients attempting to access guideline information and so the findings support the need for the RARE-Bestpractices guideline and research recommendation databases.
Appendix I

MEDLINE search strategy for rare disease guideline retrieval.

1. Rare Diseases/
2. (rare disease* or rare condition* or rare disorder* or rare syndrome*).tw.
3. (orphan disease* or orphan condition* or orphan disorder*).tw.
4. (uncommon disease* or uncommon condition* or uncommon disorder* or uncommon syndrome*).tw.
5. 1 or 2 or 3 or 4
6. "Review Literature as Topic"/
7. Evidence-Based Medicine/
8. exp "Information Storage and Retrieval"/
9. exp subject headings/
10. Knowledge Management/
11. exp terminology as topic/ or semantics/ or vocabulary/
12. exp medlars/ or exp online systems/
14. search term*.tw.
15. subject heading*.tw.
16. (Database* or online).tw.
17. information retrieval.tw.
18. (search* or identify or retriev* or locate).tw.
19. (search* or identify or retriev* or locate or protocol* or systematic* or process* or method*).tw.
20. or/6-19
21. 5 and 20
22. limit 21 to guideline
23. exp Guidelines as Topic/
24. (guideline or guidance).tw.
25. 23 or 24
26. 21 and 25
27. 22 or 26
28. find*.tw.
29. 21 and 25 and 28
Appendix II

List of included information sources

**Guideline databases**

Guideline International Network  
AHRQ National Guideline Clearing House  
NICE Evidence Search

**National guideline producers**

SIGN  
NICE  
GAIN: Guidelines and Audit Implementation Network (Northern Ireland)  
Canadian Medical Association Infobase  
Australian NHMRC  
New Zealand Guidelines Group  
MonashHealth Centre for Clinical Effectiveness

**Clinical group guideline producers**

RCN Clinical Guidelines  
Australian Safety and Efficiency Register of New Interventionsal Procedures (ASERNIP)  
Oncoline  
Royal College of Physicians (London and Edinburgh)  
Royal College of Radiologists  
Royal Society of medicine  
American Cancer Society  
British association of surgical oncology  
Royal college of surgeons

**Rare disease specific resources**

Orphanet  
EuroGentest

**Primary literature**

PubMed

**Internet searching**

Google PDF top 100
References
