HOW LOW CAN YOU GO? : TOWARD A HIERARCHY OF GREY LITERATURE.

Jess Tyndall
Flinders University
Bedford Park, South Australia, 5042
jessica.tyndall@flinders.edu.au

ABSTRACT
In recent times health sciences librarians have become increasingly involved in working closely with researchers who are undertaking literature reviews (often, but not always, systematic literature reviews). Due to the evidence-based push throughout the medical and health practice area, strong database searching skills have become essential to locate and identify the research necessary to inform practice. These are skills that librarians already have. In hospitals and universities, librarians have been heartened to find their expertise openly acknowledged as their role dovetails neatly into a position of strength. Working with researchers we find our opinions are sought and valued, and this extends to questioning our knowledge of the grey literature. While searching the published (black) literature is a given, there is a growing recognition that grey literature should be included to fully reflect the existing evidential base. Herein lies the rub. Without a clearly established hierarchy, what grey literature is admissible: particularly if we consider grey resources other than the unpublished equivalents of traditional studies and trials? Literature reviews need to include the most significant research available. And that might mean conference papers, reports, legislation or working papers - all grey. This paper examines the variety, relevance, advantages and disadvantages of grey literature and where it might ‘sit’ in the overall research picture.

BRIEF BIOGRAPHY OF PRESENTER
Jess is the Liaison Librarian for the School of Medicine, Flinders University, in South Australia. She gives lectures, runs workshops, small group, and individual sessions in research areas and in literature searching, in EBP and in identifying and accessing grey literature. She is increasingly involved in an advisory and consultative capacity with research projects

INTRODUCTION
In this paper I’m not attempting anything as ambitious or rigorous as proposing an evidential hierarchy for all types of grey literature. The very nature of the material makes nonsense of such a proposal. Rather I’m attempting to shed some light on the importance of the role that grey literature can play, and put forward for discussion a general evaluation framework which allows the researcher to decide whether a certain resource is worth including or not.
I’m the Liaison Librarian for the School of Medicine at Flinders University, in South Australia and am located at Flinders Medical Centre. As a University staff member, being located in a teaching hospital means that I have many opportunities to build up and develop a network of strong contacts with the wider health sector. Working closely with University students and academic and research staff in the areas of medicine, nutrition and dietetics, medical science, speech pathology, audiology, public health, palliative care, medical biotechnology, environmental health etc is the core work.

Running workshops for hospital-based staff including physiotherapists, radiographers, nurses, occupational therapists, social workers, medical imaging and clinical educators means further developing an understanding of what practice, and the research that informs practice, requires.

I give lectures and presentations and run workshops and tutorials in a wide range of health topics, and have a diary that is heavily booked with individual appointments with staff and students, particularly higher degrees students. Apart from helping a student plan their research approach for their dissertation, I am involved in supporting research and grant proposals, and work with various groups, often in an advisory capacity, who have high level funding for systematic reviews.

My role here is often to strengthen their literature search strategy and to advise on relevant databases and resources. Rarely would I do the actual searching myself, rather guiding those doing the research and trouble-shooting when questions or queries arise. An exception to this would be the Cochrane systematic review where I did all the literature searching and where the reviewers were generous enough to make me a co-author (Johnson, Sandford & Tyndall, 2003). I found this a very valuable experience.

Over the last 3 or 4 years my involvement in, and awareness of, research in the Australian primary health care area has made it abundantly clear that a lot of the relevant literature is grey. If I had to make a guesstimate I’d probably say more than 50%. By grey literature here, I mean it in the broad sense of material found outside the recognised databases which document the black literature.

Definitions of grey literature

The Fourth International Conference on Grey Literature held in Washington, DC, in October 1999 defined grey literature as:

"That which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers."

There are a number of other useful definitions (American Psychological Association, 2006; National Library of Australia, 2006) including that of Turner, Liddy & Bradley (2002):

“Documents that are fugitive, ephemeral, produced along non-traditional commercial pathways such as meeting-notes, think tank reports, legislative documents, pamphlets, pre-prints and annual summaries”

McKimmie & Szurmak (2002) include in their definition of grey literature “clippings… personal files, listserv queries, consultations and personal contacts…” Bichteler (1991)
reminds us that specific disciplines have different grey literature. She writes of the uniqueness of geological grey literature which includes:

“aerial photographs, guidebooks, rock and mineral specimens, field reports, well logs, surveys and cores”

These definitions give us cause to think about different types of grey literature, or whether it even has to be strictly “literature”. Perhaps we should be talking about grey “resources” as the line between formats blurs.

### Types of grey literature

<table>
<thead>
<tr>
<th>Conference papers</th>
<th>Blogs</th>
<th>Government documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standards</td>
<td>Lectures</td>
<td>Technical reports</td>
</tr>
<tr>
<td>Patents</td>
<td>Announcement</td>
<td>Scientific reports</td>
</tr>
<tr>
<td>Newsletters</td>
<td>Course materials</td>
<td>Bibliographies</td>
</tr>
<tr>
<td>Brochures/pamphlets</td>
<td>Bulletins</td>
<td>Non-commercial translations</td>
</tr>
<tr>
<td>Research reports</td>
<td>Bulletin boards</td>
<td>Trade magazines</td>
</tr>
<tr>
<td>Policy statements</td>
<td>Catalogues</td>
<td>Repair manuals</td>
</tr>
<tr>
<td>Annual reports</td>
<td>Posters</td>
<td>Preprints</td>
</tr>
<tr>
<td>Business documents</td>
<td>Press releases</td>
<td>E-prints</td>
</tr>
<tr>
<td>Dissertations</td>
<td>Essays</td>
<td>Memoranda</td>
</tr>
<tr>
<td>Patient handouts</td>
<td>Speeches</td>
<td>Legislation</td>
</tr>
<tr>
<td>Fact sheets</td>
<td>Statistics</td>
<td>Symposia</td>
</tr>
<tr>
<td>Leaflets</td>
<td>Tenders</td>
<td>White papers</td>
</tr>
<tr>
<td>Working papers</td>
<td>Interviews</td>
<td>Surveys</td>
</tr>
<tr>
<td>Abstracts</td>
<td>Directories</td>
<td>Guidelines</td>
</tr>
<tr>
<td>Book chapters</td>
<td>Programs/Projects</td>
<td>Toolkits</td>
</tr>
<tr>
<td>Bulletins</td>
<td>Questionnaires</td>
<td></td>
</tr>
</tbody>
</table>

And the list can go on ………diaries, letters, opinion polls, speeches, transcripts, power-point presentations, position papers, podcasts, emails, vodcasts, webinars. Siegel (2004) suggests even tax forms and bus timetables might be considered as pertinent to some, thus reminding us of the need at some point to define what constitutes ephemera.

There are already a mind-boggling number of types of grey literature, and with the developments coming via Web 3.0 and beyond, it will begin to take forms that are currently unimaginable.

### Accessing grey literature

Grey literature is easier to find than ever before, and much of it is being collected and archived (digitally or otherwise) with libraries cataloguing and indexing collections,
using metadata, and universities digitising theses and developing their own academic commons as well as collaborating in shared repositories like Arrow. There are country-based initiatives like Pandora (Australia’s Web Archive) and the Collect Britain initiative, in which the British Library has used state-of-the-art technology to bring to life the material it preserves [by selecting] “maps, manuscripts, topographical drawings, photographs, rare sound recordings and even long-forgotten advertisements and music-hall songs that chart the changing face of Britain and her people”.

There are specialised resources like PsycEXTRA, a gray literature database, which is a companion to the scholarly PsycINFO database. There is Oaister, a union catalogue of digital resources which mines the deep web “often hidden from search engine users behind web scripts”, with 17 million records; Scirus which indexes preprints, eprints and trawls through over 450 million science-specific web pages, enabling the user to identify and access scientific, scholarly, technical and medical data like preprints, patents and reports. And of course Google, and everything it does.

Never before have organisations and individuals directed so much effort and so much money to record and organise and classify and publicise and disseminate those efforts to the world. Which in turn means our knowledge of what grey literature is out there has never, ever been greater.

Now accessibility is not so much the problem. Now it’s about the sheer volume and what we do with it all, how we work out what’s relevant to us and our particular area of study or research.

What grey literature is acceptable in systematic reviews?

Conference papers
The Cochrane Handbook for Systematic Reviews of Interventions (Higgins, 2006), states that conference proceedings are important to hand-search because individual conference papers are rarely indexed. They are not included in MEDLINE, or many other databases, though they are covered by Sociological Abstracts, CINAHL and Web of Science. Abstracts and other grey literature have been shown to be sources of approximately 10% of the studies referenced in Cochrane Reviews (Mallett 2002). Over one-half of trials reported in conference abstract never reach full publication (Scherer 2007). In addition, grey literature in general has been found to be more likely than health care journals to contain ‘negative’ reports (McAuley 2000). Thus, failure to identify trials reported in conference proceedings might affect the results or threaten the validity of a systematic review.

Hopewell (2006) in the Cochrane systematic review ‘Grey literature in meta-analyses of randomized trials of health care interventions’ concluded “published trials tend to be larger and show an overall greater treatment effect than grey trials … this means that those carrying out systematic reviews need to search for trials in both the published, and grey literature, in order to help minimise the effects of publication bias in their review.”

Dissertations
And on dissertations, Vickers (2000) in “Incorporating data from dissertations in systematic reviews” states that although the authors identified a number of dissertations
in the subject area of their systematic review which were eventually excluded “we would argue against any conclusion that searching for theses is wasteful”.

The Cochrane Handbook for Systematic Reviews of Interventions (Higgins, 2006) 6.2.1.7 states that dissertations, theses and conference abstracts must be considered to “identify relevant studies” so the importance of these forms of grey literature, particularly in the health sciences, is definitely acknowledged. However, apart from this “means to a possible end” approach, conference proceedings and dissertations are not seen as having any intrinsic value.

Different types of studies, established evidence hierarchies and “expert opinion”

NHMRC
If we look at the types of studies that health researchers are seeking, we find that they are defined, explained and ranked by strength, in a number of well-established and well-documented models depicting “levels of evidence”. The NHMRC is currently reviewing their evidence hierarchy due to feedback from guideline developers who have indicated that it is too restrictive. “Guidelines can have different purposes, dealing with clinical questions such as intervention, diagnosis, prognosis, aetiology and screening. To address these clinical questions adequately, guideline developers need to include different research designs. This consequently requires different evidence hierarchies that recognise the importance of research designs relevant to the purpose of the guideline.” (NHMRC, 2008) Proposed new hierarchies currently have a June 2009 end date for the deliberation process.

This consultative document includes detailed analysis and description of different study types, and in part B goes on to describe “How to assess the body of evidence and formulate recommendations”. There is also growing recognition of the work needed to be done in the area of appraisal of qualitative research. Well over one hundred tools and frameworks currently exist, and the debate as to how this appraisal can best be achieved is continuing. The main aim is to produce greater consistency of judgment in relation to inclusion of qualitative research in systematic reviews. Trish Greenhalgh, who is much cited for her excellent “How to read a paper” series has written a very readable article in BMJ “Papers that go beyond numbers (qualitative research)” (Greenhalgh, 1997), and has previously written about narrative-based medicine (Greenhalgh, 2005) and even the relationship between intuition and evidence (Greenhalgh, 2002).

Clearly grey, unpublished equivalents of traditional studies and trials that fit the bill are acceptable, but it seems that no other grey literature is admissible. Not everyone involved in research is caught up in the lengthy and costly process of producing this type of systematic review. Surely many literature reviews need to include the most significant research available? For many others there is a need to look at whatever work has been done, not just the literature that meets the currently defined hierarchy of evidence. There is absolutely no question that research (health research in particular) needs the standards like those set by the NHMRC, but there should also be no question that for many researchers, it makes very good sense to look further and wider and deeper, and in different places.
Joanna Briggs Institute

Consider the NHMRC model with this one of evidence-based healthcare from the Joanna Briggs Institute:

“The JBI model of evidence-based healthcare adopts a pluralistic approach to the notion of evidence whereby the findings of qualitative research studies are regarded as rigorously generated evidence, and other text derived from opinion, experience and expertise is acknowledged as forms of evidence when the results of research are unavailable” (Pearson et al. 2005).

“Expert opinion”

This notion of expertise, experience and opinion is echoed by the editors of the popular evidence-based clinical information resource UpToDate, who derive their evidence from the usual resources, but state in their editorial policy that they also include proceedings of major national meetings and “the clinical experience and observations of our authors, editors and peer reviewers” (UpToDate, 2008). Many other examples can be found where a pyramid or table designed to illustrate levels of evidence, gives the bottom rung to “background information/expert opinion” and the like (Wormald, 2004, Glover et al., 2007)

The broader role of grey literature in research

The subject of this paper came about as a direct result of the growing interest in grey literature from student and staff researchers, both from the hospital and the university. For the last three years I have been running an “Accessing the Grey literature” workshop for the Staff Development Unit at Flinders University, for postgraduate students and academic staff. One of the best things to come out of these workshops has been that they are always required to multidisciplinary, and so although my strength is firmly in the medical and health sciences I had to design and develop the workshops to draw examples from a wide range of subject areas. We worked with resources from the sciences, social sciences, arts and humanities. This expanded my view of not only what grey literature could encompass but also how the definition of what was ‘eligible’ as valid for research purposes, differed from discipline to discipline.

Sifting through the results of various searches in the grey literature often led to discussion about what material was “strong enough” to be included. Without the equivalent of a clearly established evidence hierarchy, how could we decide what grey literature “made the grade”, particularly if we considered grey resources other than the unpublished equivalents of traditional studies and trials? Was there a way of evaluating the resources to decide what to keep and what to discard, and if so, how could this be done?

If we cast our minds back to our long list of grey resources the dilemma is clear. Community health workers need to read field reports, students of politics need to read speeches and press releases. Legislation, standards, patents and statistics are relevant to many areas of research ……and so on..

This brings me back to the core question “how low can you go”. Researchers want to know. “can I use blogs in my literature review”, or, is a newspaper article “strong enough”? My suggestion is that we open our minds as widely as we are able, and look at the core purpose of any research.
Picture this. You are researching Australian university life in the early 70’s, with an emphasis on the role of student politics. Essential primary material includes student newspapers which give contemporary voice to mood, thoughts, philosophies and attitudes of the protagonists. Flinders University student body produced a student newspaper at the time called “The persistent goose” which (I think) ran to only 2 issues. It is held in Special Collections, has not yet been catalogued and could be seen as ephemera. Should you use this in your research or your literature review (systematic or otherwise)?

Here’s a second scenario. You’re at a dinner party and seated next to Stephen Hawking, who in conversation tells you that he has been working on a new theory regarding black holes. Encouraged by your great interest, and in a mood to share enthusiasms he asks his assistant to sketch out the basic principles onto a dinner napkin which is then given to you. As the science writer for a major newspaper you can’t believe your luck but you know your editor will not accept such “anecdotal information’ even from one as eminent as Stephen Hawking. You take a chance and email the great man, reminding him of your conversation, and he duly replies, reiterating the points he made to you earlier, at the dinner party.

Does it matter that the information was verbal, was sketched onto a serviette, or came via email?

Referring back to the Joanna Briggs model, and information “derived from opinion, experience and expertise” we can infer that the validity of such information might not be so far-fetched. They have outlined in their model a way of approaching this evaluation process: examining the opinion; identifying the credibility of the source of the opinion; establishing the motives that underlie the opinion and locating alternative opinions that give credence to it (or conversely, question it).

Rather than the format, which we often seem to get rather caught up in, doesn’t it matter more who is behind the information, who is responsible for the intellectual content? Stephen Hawking must at the very least count as “expert opinion”. In fact, in both our scenarios the resources speak with authority. The student union president writing in the Flinders newspaper is arguably also an expert voice in representing the political climate of that time and that place.

Student newspapers, emails, conversations, annotated serviettes – all grey literature (or resources) and all significant to the right researcher: significant being defined as (Geddes, 2005): “full of meaning, esp a special or hidden one; momentous, important; highly expressive, indicative”.

I would like to propose, in the absence of any other neat template or identifiable hierarchy, that grey literature (other than studies and systematic reviews or meta-analyses that equate with the evidence levels in the black literature) be critically appraised for strength and validity using a simple approach. And that is, marrying the concept of “expert opinion/insider knowledge” with the general principles used to evaluate web resources. I believe this is somewhat in the spirit of Joanna Briggs.
If we agree that the “intellectual content” principle is paramount, then perhaps we could amalgamate this with the “expert opinion” concept and say that the first principle for selection of grey literature/resources admissible as value to research, is Authority.

When I became sufficiently involved with grey Literature that I felt I needed a way to analyse and organise types, I searched high and low for a nice hierarchy or checklist that somebody else had done. I didn’t find one, presumably because most grey literature formats can’t be approached in a quantitative way and is therefore difficult to evaluate and measure. I did find some interesting things along the way though. I found that tools were being developed that met specific or niche needs and were directed at evaluating particular types of grey resources. For example, CHERRIES (The Checklist for Reporting the Results of Internet eSurveys) (Eysenbach, 2004) is designed to appraise the quality of web-based surveys, in same way that the CONSORT statement is used to evaluate RCTs. The SEVAL Standards (from Switzerland) are used to evaluate public administration reports and incorporate four quality dimensions: Propriety, Accuracy, Utility and Feasibility.

What I initially hoped to find was something akin to the checklists for black literature that said: conference papers – yes; dissertations – yes; standards – yes; reports (broken down into types – some yes, some no); newspapers – maybe; book chapters – probably …and so on.

If someone else had put one together, and justified it, I would’ve been happy to use it. But examination makes it clear that it’s just not possible to take purely the format approach to the majority of grey literature. Well-respected academics write books and extremists post blogs – but the reverse is also true.

**A proposed checklist**

We’ve already looked at the concept of “expert opinion” and there are many who feel it should have greater recognition (Ferrie, 2005). I propose a checklist that has the flexibility to be applied to the widest range of resources: from models of primary healthcare to dissertations, maps, diaries, podcasts, blogs and so on. A checklist which brings together the guidance of the following sources…

Joanna Briggs Institute: “Expertise, experience and opinion”

UpToDate: “the clinical experience and observations of our authors, editors and peer reviewers”

SEVAL standards: Propriety, Accuracy, Utility and Feasibility

Glover: “background information/expert opinion”

Wormald: “opinions of respected authorities, based on clinical experience; descriptive studies or reports of expert committees”

… and uses common sense to meld them together with the generally accepted web evaluation criteria of: Authority, Accuracy, Objectivity, Coverage and Date, which can be applied to each and every item of grey literature we come across, and which I’ll expand upon now.

**Authority**
Incorporates expertise, propriety, experience, credibility, reliability: it asks the question who is responsible for the intellectual content? Does the author have qualifications, association with responsible recognised organisations? What else have they written/published/produced? (black or grey). Do they work in the field? Are they a recognised expert, identified in a variety of sources? Are they cited often by others (use Google Scholar as a quick guide). Are they a Masters or Ph.D. student from a known university and under “expert” supervision? Do they know what they’re talking about? Do we know who they are? Of course, usually the simple rule No author = No use applies.

Accuracy
Does it seem to be right? Has it used a clearly stated methodology? Has it been peer-reviewed? Is it supported by authoritative references, credible sources or documented references? Can it be validated as representative? How does it “sit” with other work on the same topic? Try cross-checking and counterbalancing with other resources on a similar topic and accessing background information to gain context. Be creative in approach, particularly when making an assessment of secondary literature. For example, a policy brief of a technical report needs to be interpreted in light of the original for its accuracy to be fairly judged.

Coverage
Try to be aware of any coverage limitations, stated or otherwise. The content of the work could be limited geographically or culturally for example, or its coverage could be very specific in any number of ways. This is quite acceptable but like objectivity it needs to be clearly stated so that any “sphere of influence” or context is understood and taken into consideration during any selection process for eligibility.

Objectivity
This incorporates opinion, expert or otherwise. Is it balanced or does it have an unstated bias? A position paper from an organisation has a particular point of view. A check against background context of the work is always advised. For a web-based document use Google Advanced search to see who is linking to the document in question. Any links (or no links) will give an indication of peer or general acceptance or expert recognition. It may be a strong document to cite but bias need to be acknowledged.

Date
Does the resource have a clear date related to content? If the date is very recent but doesn’t seem to take into account important material it should, then this should ring alarm bells. In fact any resource should be checked against contemporary material. If necessary, considerable creativity should be employed to obtain at least a probable date for undated grey resources. Like the no author rule, no date = no use.

This selection/assessment criteria is generally applicable and relates to the item or resource itself. One more important category needs to be addressed in the light of eligibility for inclusion. Is the resource significant?
Significance
This incorporates SEVAL’s feasibility and utility, as well as relevance. Is it meaningful? Does it add context? Does it enrich and add to the current research, particularly add something unique? Or does it strengthen a current position? Would the research be lesser without it? Is it integral, representative, typical? Does it have ‘impact’; for example, in the sense of influencing the work or behaviour of others?

AACODS
AACODS (Authority, Accuracy, Coverage, Objectivity, Date, Significance) is a simplified “in principle” way of dealing with a complicated picture. Developed in the spirit of “expert opinion”, this checklist attempts to guide researchers through a broad appraisal process. One that is not restricted to a particular discipline or format, but is broad enough, and flexible enough, to address most examples of grey literature or grey resources.

Conclusion
Specific subject areas sometimes have particular forms of grey literature appraisal and are developing eligibility criteria and assessment of their own. For example, there has been some discussion in the public health arena regarding ways of ‘ranking’ various types of reports. While these tools are still in the process of being developed, this discussion paper would like to suggest that AACODS could go some way towards bridging the gap, and answering the question “how low can you go”, by helping to clarify “how low is too low”.

BIBLIOGRAPHY


Glover, J, Izzo, D, Odato, K & Wang, L, 2006 et al Hierarchy of Evidence-based Medicine (EBM) resources , viewed May 13 2008

< http://www.mclibrary.duke.edu/subject/ebm/ebmpyramid.html >


Greenhalgh, T 1997, ‘How to read a paper: Papers that go beyond numbers (qualitative research)’ BMJ; vol. 315: pp. 740-743


Johnson A, Sandford J & Tyndall J. 2003, ‘Written and verbal information versus verbal information only for patients being discharged from acute hospital settings to home’ (Cochrane Review), viewed 10 June 2007, (online Wiley InterScience/ Cochrane Database of Systematic Reviews).


National Health and Medical Research Council. NHMRC additional levels of evidence and grades for recommendations for developers of guidelines STAGE 2 CONSULTATION Early 2008 – end June 2009, viewed July 17 2008,

Scherer, RW, Langenberg, P & von Elm, E 2007, *Full publication of results initially presented in abstracts. Cochrane Database of Systematic Reviews* (Cochrane Review), viewed 10 June 2007, (online Wiley InterScience/ Cochrane Database of Systematic Reviews).


