

Exploring all that is Grey in the Health Sciences: What is Grey Literature and How to use it for Comprehensive Knowledge Synthesis

Vaska M¹, Chowdhury MZI^{2,3}, Naidu J⁴, Baig K², Turin TC^{2,3,5}

¹Knowledge Resource Service, Tom Baker Cancer Centre, Alberta Health Services, Calgary, Alberta, Canada. ²Department of Family Medicine, University of Calgary, Calgary, Alberta, Canada. ³Department of Community Health Sciences, University of Calgary, Calgary, Alberta, Canada. ⁴Alberta Health Services, Calgary, Alberta, Canada. ⁵Department of Epidemiology & Research, National Heart Foundation Hospital & Research Institute, Dhaka, Bangladesh.

Abstract

Grey literature is defined as “a field in library and information science that deals with the production, distribution, and access to multiple document types produced on all levels of government, academics, business, and organization in electronic and print formats not controlled by commercial publishing, i.e. where publishing is not the primary activity of the producing body” [GreyNet International, 2016]. This paper will provide an overview of grey literature, along with definitions, examples of sources, and suggestions on how to search for this material. In addition, it will describe how best to document grey literature search strategies, as well as evaluate and critically appraise sources of information located in the grey literature realm. Finally, it will outline why the inclusion of grey literature in systematic literature reviews is important in order to reduce publication bias, provide a balanced picture of the available evidence, and increase the comprehensiveness of the review.

Conflict of Interest: None

Keywords: Grey literature, systematic review, environmental scan, comprehensive search (JNHFB 2019; 8 : 14-19)

Background

What is Grey Literature?

On March 29, 2016, GreyNet, the international grey literature community, announced the launch of a new portal dedicated to unveiling “good practices and resources in grey literature”¹. Further, the 18th international conference on grey literature, held at the New York Academy of Medicine in November 2016, focused on a central theme of leveraging the diversity of literature for innovation and change². The concept of grey literature is a body of information that has surfaced in recent years due to the explosion in literature dissemination as a result of the current digitalization process worldwide. It is now seen as supplementary to literature found in academic databases.

While several synonyms have been used in place of ‘grey’ to describe this literature, including non-conventional, informal, informally published, fugitive, and invisible, ‘grey’ is undoubtedly the moniker by which this concept is most widely known. As demonstrated in Table 1, several definitions of grey literature have been put forth over the

years, with a 2010 conference in Prague discussing the incorporation of new elements and additions³. Nevertheless, a definition developed more than a decade earlier during the third international grey literature conference in 1997 is perhaps the most widely accepted: “a field in library and information science that deals with the production, distribution, and access to multiple document types produced on all levels of government, academics, business, and organization in electronic and print formats not controlled by commercial publishing, i.e. where publishing is not the primary activity of the producing body”⁴.

In academia, it is essential that a researcher remains objective in the pursuit of information. Researchers must therefore include grey literature along with published material when conducting literature searches, particularly with regard to systematic reviews, in order to limit bias³. The following paper will discuss examples of grey literature documents, as well as provide tips and suggestions for how to search for this material. In addition, it will describe how best to document grey literature search strategies, and evaluate and critically appraise sources of information located in the grey literature realm.

Corresponding Author

Tanvir C. Turin MBBS MS PhD,
Department of Family Medicine, University of Calgary
Room G012F, Health Sciences Center,
3330 Hospital Drive Northwest, Calgary, Alberta T2N 4N1, Canada.

Table 1: Definitions of Grey Literature

Source	Definition
Cochrane Handbook^a	“Literature that is not formally published in sources such as books or journal articles”
Wikipedia^b	“Materials and research produced by organizations, outside of commercial or academic publishing and distribution channels”
GreyNet International^c	A field in library and information science that deals with the production, distribution, and access to multiple document types produced on all levels of government, academics, business, and organization in electronic and print formats not controlled by commercial publishing
University of Toronto Libraries^d	Any literature that has not been published through traditional means. It is often excluded from large databases and other mainstream sources
Dalhousie University Libraries^e	Literature produced by any organization whose central purpose is not publishing. These organizations include governments, businesses, notfor- profits, health organizations and associations, and more.

^aCochrane Handbook: https://handbook-5-1.cochrane.org/chapter_6/6_2_1_8_grey_literature_databases.htm;

^bWikipedia: https://en.wikipedia.org/wiki/Grey_literature

^cGreyNet International: <http://www.greynet.org/>

^dUniversity of Toronto Libraries: <https://guides.library.utoronto.ca/c.php?g=250454&p=1670693>

^eDalhousie University Libraries: <http://dal.ca.libguides.com/systematicreviews/greylit>

Technique

Typology of Grey Literature

Approximately 150 types of grey literature have been identified in a comprehensive list on GreyNet, and this is growing at an exponential rate as more potential sources of grey are discovered⁵. In addition to traditional publications such as theses and dissertations, conference papers and proceedings, newsletters, and government reports, various forms of informational communication, such as telephone calls, meeting minutes, e-mails, and interviews, are forms of grey literature that have come to be seen as valid sources of information in recent years³. While much of this information is increasingly available online, electronic access is not a factor in determining whether or not a resource is deemed grey. In fact, pamphlets, posters, tickets, and timetables are all considered grey literature, despite existing in both print and electronic formats³. Referred to as ephemera, this material “carries a verbal message and is produced by printing or illustrative processes, but not in the standard book, periodical, or pamphlet format.”⁶ Social media information deserves mention as a potential source of grey literature. For instance, while Twitter has been scrutinized with regard to information trustworthiness, it played a key role in the 2009 H1N1 outbreak, particularly during the early months when healthcare officials were scrambling to organize a mass vaccination program. In fact, a number of tweets were actually quite informative, accessible to the public long

before a summary of daily events was published in a local newspaper³. Table 2 and Figure 1 show the various sources and examples of grey literature discussed above, as well as others.

Table 2: Sources of Grey Literature

Type of source	Example(s)
Technical documentation	Manual
Thesis	Bachelor's Thesis, Dissertation, Habilitation Thesis, Master's Thesis
Website	Company website; retail website; health website
Article	Essay, Interview, Preprint, Review
Proceedings	Conference Proceedings
Conference materials	Call for papers, Conference paper, Conference program, Poster
Tertiary documents	Bibliography, Codebook, Directory, Glossary, Index, Timeline, Catalogue
Corporate literature	Product Catalogue, Minutes
Journal	House Journal, Bulletin, Newsletter
Course material	Course Text, Curriculum, Exam Topics, Syllabus
Image material	Map
Informative material	Announcement, Brochure, Chronicle, Leaflet, Memorandum
Monograph	Annual, Anthology, Guidebook, Handbook
Normative document	Government Document, Green Paper, Legal Document, Standard
Report	Annual Report, Press Release
Research plan	Analysis, Dataset, Proposal, Protocol, Questionnaire, Study
Software	Productivity software; tax software

NB. Adapted from: GreyNet Document Types in Grey Literature <http://www.greynet.org/greysourceindex/documenttypes.html>

A fundamental feature of grey literature is its almost instantaneous, rapid publication; information is posted as soon as it is written³. Some may argue that this also serves as one of its biggest disadvantages. Lack of bibliographic control, non-professional format, and search interfaces that are often no more than a single text box in an obscure location on a page, are among the many challenges of searching grey literature effectively. This reality would coincide well with Auger’s view of the minimalist approach to indexing, providing “few details, no abstracts, and no indexes.”⁷

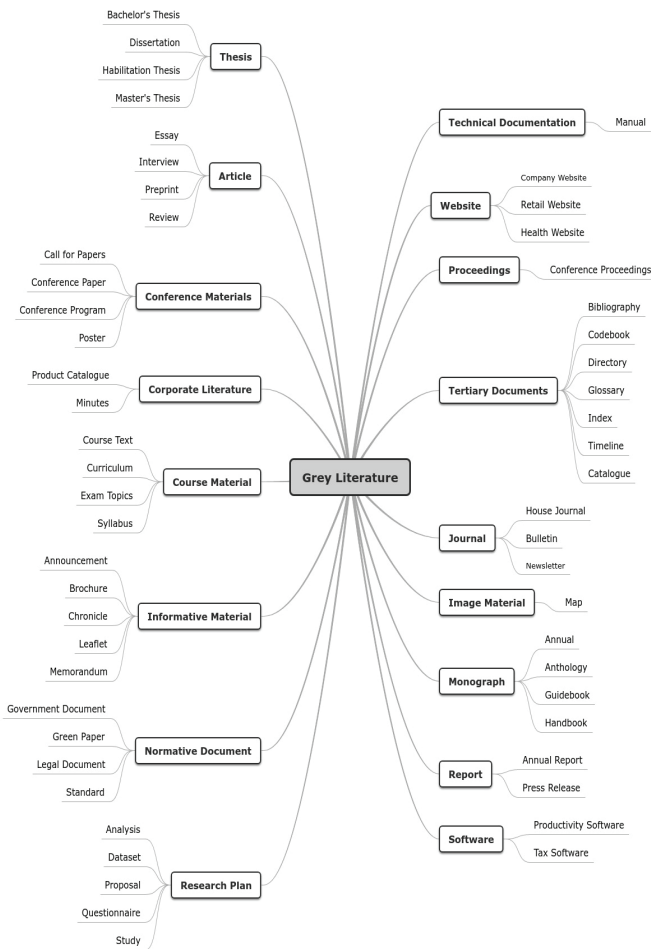


Figure 1: Grey Literature typology based on sources

Documenting Grey Literature Search Strategies

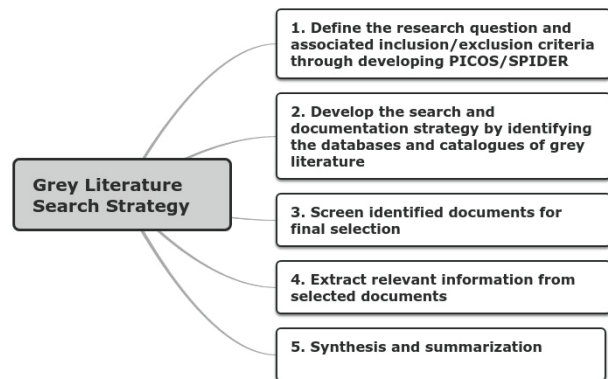
How then, does one go about searching the realm of grey?

Searching for grey literature can be an iterative process. As a hard and fast rule, keyword searching dominates any search strategy when seeking out the grey. As many websites are powered by the Google search engine, common search strategy tips, such as truncation/wildcard, phrase searching in quotes, and the Boolean operators AND as well as OR can be universally applied. The difference lies in how search strings are structured. While an academic database uses a consistent algorithm to retrieve articles matched by the inputted keywords, grey literature search strategies often need to be tailored to each website, “accounting for the characteristics of the search interfaces and the areas of focus of the host organization.”⁸ In 2011, the Canadian Institute for Health Information (CIHI) released a literature search methodology paper, describing how best to search the wealth of rather unwieldy papers. The authors demonstrated a clear distinction between total results indicated for a search string compared to the actual number of relevant results retrieved and cited.

Further, they recommended that given the relevancy algorithm of Google searches, “only the first 100 hits [are] retained for each search [string]” in addition to any

topic-specific resources that may be introduced, due to the breadth and volume of material stored in Google⁹. The methodology behind searching the grey literature is similar to the Logic Model approach undertaken when conducting a search in databases and can be broken down into five key steps (see Figure 2 for a visual representation). The first step, fundamental to any research pursuit, pertains to formulating an answerable research question, using a number of available frameworks (PICO, SPIDER, etc.) to help guide this process¹⁰. Once the research question has been determined, the search strategy is developed (Step 2), where appropriate resources and search terms are established. (11) Step 3 involves screening search yields to identify documents for final selection, while step 4 consists of critically appraising each selected document in order to extract information relevant to the research question. Finally, step 5 involves synthesizing and summarizing the critically appraised documents in step 4 into a final product.

Figure 2: Grey Literature Search Strategy Logic Model



While the ability to generate complex search strings is rather severely limited when searching for grey literature, even a single search box on a website can be used effectively if one follows a recommended guide of search strategy tips. With its Grey Matters checklist, the Canadian Agency for Drugs and Technologies in Health (CADTH) has streamlined the search parameter conundrum by including scope notes below each resource mentioned, indicating evidence-based search strategy tips compiled by the CADTH team to help retrieve the most relevant results for every resource consulted. For instance, one can only filter results by year or category when perusing the Alberta College of Family Physicians website, whereas search results retrieved from CADTH’s own website can be filtered down to product line, result type, publication date, and several additional options. (12) CADTH states that the purpose for providing searching guidance for each resource is two-fold, namely to ensure that grey literature search standards are parallel to that of academic search standards, and also that thorough documentation of the grey literature search process increases “transparency and the potential for reproducibility.”¹² Table 3 presents an outline of the differences between grey literature and academic literature.

Table 3: Difference between academic and grey literature

	Grey Literature	Academic Literature
Speed of publication	Instant, few to no steps	Slow, many steps
Available in / through	There are no dedicated databases. Captured using search engines, surfing through various websites	Dedicated databases are available for academic journals, academic databases (MEDLINE, PubMed, EMBASE)
Accessibility	Free access	Expensive subscriptions (some are freely available)
Controlled by	Various organizations, very decentralized	Publishing companies, research inventory or cataloguing entities, more centralized
Searchability	Difficult to navigate, little to no standardization, lacks advanced search features	Easier to navigate, more standardized and advanced search features
Quality	Wide variety due to little to no review	Generally, consistently good due to peer-review process
Change over time	Content in grey literature resources might frequently change.	Largely, content of the resources does not change over time
Reproducibility of search	There are problems with reproducibility, especially if the search documentation is not done with greater detail	Following a systematic and comprehensive approach leads to good reproducibility
Cost	Cheap to free	Expensive

Evaluating & Critically Appraising Grey Literature

With so many unconventional document types comprising the grey literature realm, recognizing a trustworthy source from an imposter site containing suspicious or unverifiable information can be a daunting task. While suggestions on how to judge the reputability of information found on the Internet has been around almost since the invention of the Web itself, there has, until recently, been no official, universal standard to judge the credibility of a grey literature resource.

This all changed in 2010 with Tyndall’s invention of the AACODS checklist, shown below¹³. Consisting of six broad-level categories (authority, accuracy, coverage, objectivity, date, and significance), the checklist encourages the reader to carefully scrutinize the reputability, validity, currency, and credibility of each grey literature document s/he comes across before rendering a final decision as to whether or not it is worthy of inclusion. While Tyndall admits that the significance of a grey literature resource is, in essence, a value judgment made by the researcher, the purpose of the AACODS checklist is to uphold the argument that “grey (unpublished) studies and RCTs should be appraised using the same tools as their (published) counterparts.”¹³.

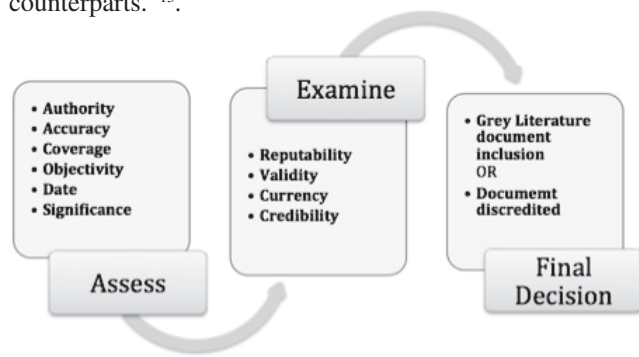


Figure 3: Checklist of Appraising Grey Literature (graphical adaptation of Tyndall’s AACODS checklist available in table format https://dspace.flinders.edu.au/xmlui/bitstream/handle/2328/3326/AACODS_Checklist.pdf)


GreyHub.org and OAISter: Supporting the Open Access Movement and Grey Literature

In order to provide a synopsis of the various grey literature resources that are available to locate that elusive non-commercially published document, it may be best to represent the search journey as a case-based scenario, as the following example illustrates, where a search for electronic health/medical records in a clinical setting was conducted. As a cardiovascular surgeon, you and your team have been tasked with determining the benefit and impact of installing portable defibrillators in sports arenas. You would like to locate a few policy papers on how communities are approaching this issue.


Undoubtedly, despite the profusion of academic databases that exist today, Google continues to be the first choice to quench the desire to quickly find information on a topic. However, a Google search will often lead to a dead-end or information disconnect, such as retrieving the citation for an article, but not the full-text due to licensing restrictions, publication embargos, etc. OpenGrey, a repository system for information on grey literature in Europe, was launched in 2011 with a desire to open up boundaries and keep grey literature open, free, and accessible to all. This notion strongly coincides with the Open Access Movement, providing “unrestricted access and unrestricted re-use”¹⁴. to a diverse range of material. One shade of grey that has gained momentum as a legitimate grey literature source is the open access federated search engine or institutional repository. Two of these information warehouses, namely, GreyHub.org and OAISter, act as gatekeepers for a wealth of repository material, such as theses, manuscripts, conference proceedings, and government documents.

Launched in 2018 as a discovery service for Grey Literature, GreyHub is a “catalog of thousands of curated grey literature repositories.”¹⁵. The result of a joint effort between Deep Web Technologies (DWT), GreyNet, and the Grey Literature community, GreyHub searches through reports, government documents, theses, data, and other types of grey literature¹⁵. While the search yields of queries performed in GreyHub are large, results are clustered according to Topic, Author, Publication, Document Type, Source, and Document Format. As an example, the search string, defibrillator* AND ("sports hall" OR "sports arena" OR "community arena" OR arena OR stadium) yields 578 total results, which can be clustered by topic to focus on study (36 hits), research (25 hits), cardiac arrest (20 hits), care (12 hits), sudden cardiac death (11 hits), and type of sport, i.e. football (6 hits). A sample record is included below (Box 1):

Box 1: A sample record derived from GreyHub search

1. [Football stadium defibrillators for cardiac arrest](#) 

★★★★★



WorldWideScience.org 

2009-01-01 Federal Science Library - Canada
 ISSN: 0195-668X Volume: 30 Issue: 7
 Categories: **Defibrillators** - supply & distribution, Genetic Testing, Soccer, Humans, Heart Arrest - genetics, Heart Arrest - therapy ... CardioPulse European Heart Journal (2009) 30, 743749
 doi:10.1093/eurheartj/ehp083 Football stadium debrillators for cardiac arrest A player collapsing...

The Open Archives Initiative (OAI)ster database states that it contains “millions of digital resources from thousands of contributors”.(16) It is a union catalog developed by the Online Computer Library Centre (OCLC), that in fact, contains more than 30 million records. Utilizing the same search strategy as per GreyHub, 9 unique records are retrieved, including a consensus document on cardiovascular safety at sports arenas from Oxford University (Box 2).

Box 2: A sample record derived from OAIster search

Select All Clear All Save to: [New List] Save Sort by: [Library and Relevance] Save Search

1.  [Consensus document regarding cardiovascular safety at sports arenas: position stand from the European Association of Cardiovascular Prevention and Rehabilitation \(EACPR\), section of Sports Cardiology.](#)
 by Borjesson, Mats; Serratos, Luis; Carre, Francois; Corrado, Domenico; Drezner, Jonathan; Dugmore, Dorian L; Heidbuchel, Hein H; Mellwig, Klaus-Peter; Panhuyzen-Goedkoop, Nicole M; Papadakis, Michael; Rasmussen, Hanne; Sharma, Sanjay; Solberg, Erik E; Van Buuren, Frank; Pelliccia, Antonio
 Downloadable archival material
 Publication: European Heart Journal
 Database: OAIster
 Libraries that own this item: WorldCat Libraries

Discussion

Importance of Grey literature search for comprehensiveness of information gathering

There are several benefits of including grey literature in systematic reviews, all of which contribute to the comprehensiveness of the review. The first is that it reduces publication bias, which refers to the “propensity for only studies reporting positive findings to be published, and may skew the results of the meta-analysis and systematic review.”¹⁷ Gray literature sources are more likely to include neutral or negative results, which may provide the reviewer with a more accurate understanding of the evidence at hand⁸. The inclusion of grey literature also allows reviewers to draw a balanced picture of available evidence on a given topic. For instance, it is noted that published trials in health sciences report greater treatment effects than those found in grey literature. Removing grey literature from some meta-analyses results in larger estimates of treatment effects, less precise effect-size estimates, and more significant results. Paez (2017) notes that as a result, “excluding unpublished studies may compromise the validity and reliability of meta-analyses and the specificity of systematic reviews”¹⁷.

Grey Literature Searching vs. Internet Scan or Online Web Searching

With increasing digitization of archived material,

the Internet is indeed the go-to-place to locate grey literature resources. Nevertheless, there is an important distinction between a dedicated grey literature search and an online/web search. Grey Literature searching often involves perusing an organization’s website (whether via a keyword search or browsing) to locate material pertinent to a topic. These websites are dedicated to retrieve as much content as possible that is produced directly by that organization. The searches are thus more structured and yield higher quality, relevant documents. In contrast, an online/web search involves “using search engines and websites which are not dedicated web-based literature resources.” (18) These searches yield results that are more general and overarching in nature, as crawling takes place across a vast number of resources to retrieve potential websites worthy of consideration. The so-called ‘black and white’ literature that we get when searching most online databases during our day-to-day routine is managed by commercial entities (e.g. Yahoo, Google, Bing, etc.). They have their commercial / advertisement interest intrinsic in their search engines which introduces bias on derived search results. On the contrary, grey literature is produced by entities whose main task is not publishing or managing search engines. So, the question now is, if grey literature is not produced by commercial publishing entities then who is producing Grey literature? A number of organizations or entities can be producers of Grey literature such as all levels of governmental departments / sections, think tanks, non governmental organization or non-profits, as well as scholarly societies and associations. Table 4 shows the features of a grey literature search and an Internet scan / online search in a comparative manner.

Table 4: Features of grey literature search and Internet scan / online search

	Grey Literature search	Internet scan / Online search
Availability through	Search engines (e.g. Google), various websites (e.g. GreyHub)	Search engines (e.g. Google, Yahoo, Bing)
Accessibility	Free access	Free access
Hosted by	Various organizations, very decentralized	Anybody can publish any information without any scrutiny
Searchability	Difficult to navigate, little to no standardization	No rules/regulations associated with information hosting
Credibility of host	Varies widely	No control whatsoever on credibility of host
Reproducibility of search	There are problems with reproducibility, especially if the search documentation is not done with greater detail	The results of web searches will have very low reproducibility because web content and search algorithms often change. Nonetheless, it is still very important to report the detailed search process in order to ensure the methods used are transparent.
Reliability of information	Wide variety due to little to no review process in place	No control whatsoever on quality of information hosted
Cost for accessing	Cheap to free	Free

To illustrate the distinction between a grey literature and online/web search, let’s refer back to our case-study of the cardiovascular surgeon determining the feasibility of porta-

ble defibrillators in sports arenas. A targeted grey literature search on this topic may involve consulting the Fédération Internationale de Football Association (FIFA), which produced a paper detailing discussions among FIFA's 208 member associations on best practices of ensuring availability of defibrillators in all major football (soccer) events. One of the fundamental traits of grey literature is ensuring that bias should be accounted for in documents that are produced; the paper authored by FIFA certainly accounts for this distinction, as it is a blend of both factual and expert opinion. On the other hand, interest in more quantitative factual information, such as how many defibrillators have been installed in sporting arenas in Canada, would lead to information about a program funded by the Public Health Agency of Canada whereby 3,234 defibrillators were installed in sporting arenas and recreation centres across the country in 2016. An online/web search is thus more exploratory than a grey literature search; consulting the Public Health Agency of Canada may not have been the obvious first choice resource, however perusing web sites via the aid of a search engine leads one to find the answer.

Conclusion

This paper has provided a mere snippet of the world of grey; raising awareness of this information medium is a never-ending task, one that is far from complete. Nevertheless, grey literature is increasingly being seen as a required supplement to academic databases and is often mentioned in search methodologies. In an interview conducted at the turn of the millennium with Dr. Dominic Farace, founder of GreyNet, the vision of grey literature presented then still holds true today: "just as in the last decade of the twentieth century came to dominate the information supply side, so well before the end of the first decade of this twentyfirst century society will also come to dominate the information demand side."¹⁹ Anyone who has followed the grey literature movement over the past 16 years knows that these words certainly ring true.

References

1. GreyGuide: Repository and guide to good practices and resources in grey literature. [Internet]. 2016 [cited 2019 Mar 25] Available from: <http://greyguide.isti.cnr.it/>
2. TextRelease. GL18: Eighteenth international conference on grey literature [Internet]. 2015 [cited 2019 Mar 25]. Available from: <http://www.textrelease.com/home.html>
3. Vaska M. Introducing grey literature. In: Pejsova P, editor. Grey Literature Repositories [Internet]. Zlin, Czech Republic: VeRBuM; 2010 [cited 2019 Mar 25]. p. 11-9. Available from: http://repozitar.techlib.cz/record/285/files/idr-285_2.pdf
4. GreyNet International [Internet]. 2016 [cited 2019 Mar 25]. Available from: <http://www.greynet.org/home.html>.
5. GreyNet International. Document types in grey literature [Internet]. 2016 [cited 2019 Mar 25]. Available from: <http://www.greynet.org/greysourceindex/documenttypes.html>.
6. Auger CP, Auger CP. Use of reports literature. Information sources in grey literature. 3rd ed. London; New Providence, NJ: Bowker-Saur; 1994.
7. Auger CP, Auger CP. Use of reports literature. Information sources in Grey literature. 2nd ed. London; New York: Bowker-Saur; 1989.
8. Bellefontaine SP, Lee CM. Between black and white: examining grey literature in metaanalyses of psychological research. *J Child Fam Stud*. 2014;23:1378-88.
9. Canadian Institute for Health Information (CIHI). Urban physical environments and health inequalities: Literature search methodology paper [Internet] 2011 [cited 2019 Mar 25]. Available from: https://www.cihi.ca/sites/default/files/cphi_upe_lit-search_method_en_0.pdf
10. Ahmed S, Vaska M, Turin TC. Conducting a literature review in health research: basics of the approach, typology and methodology. *JNHFB*. 2016;5(2):44-51
11. Ahmed S, Vaska M, Turin TC. Comprehensive systematic search process of health literature: hunting pearls out of the sea. *JNHFB*. 2016;5(1):12-16
12. Canadian Agency for Drugs and Technologies in Health (CADTH). Grey Matters: a practical tool for searching health-related grey literature [Internet]. 2015 [cited 2019 Mar 25]. Available from: <https://www.cadth.ca/resources/finding-evidence/grey-matters>.
13. Tyndall J. The AACODS checklist: designed to enable evaluation and critical appraisal of grey literature [Internet]. 2010 [cited 2019 Mar 25]. Available from: https://dspace.flinders.edu.au/jspui/bitstream/2328/3326/4/AACODS_Checklist.pdf.
14. Shockey N. International Open Access week [Internet]. 2019 [cited 2019 Jun 6]. Available from: <http://www.openaccessweek.org/>
15. Deep Web Technologies. GreyHub.org [Internet]. 2018 [cited 2019 Mar 25]. Available from http://greyguide.isti.cnr.it/wp-content/uploads/2018/12/GL20_ProgramBook.pdf.
16. The OAISter database [Internet]. 2016 [cited 2019 Mar 25]. Available from: <http://www.oclc.org/oaister.en.html>.
17. Paez, A. Gray literature: An important resource in systematic reviews. *J Evid Based Med*. 2017 Aug; 10(3):233-240.
18. Bethel A, Briscoe S, Rogers M. Searching the grey literature [Internet]. n.d. [cited 2019 May 31]. Available from: <https://medicine.exeter.ac.uk/media/universityofexeter/medicalschoo/research/pentag/docu19>. Gelfand J. Interview with Dominic Farace, founder of GreyNet. *The International Journal on Grey Literature*. 2000;1(2):73-6.