A step by step tutorial for putting together a minimalistic evidence map and interpret the data

Arindam Basu[1]

[1]University of Canterbury

September 12, 2020

What is an evidence map and how does it differ from an systematic review?

Evidence on the effectiveness or prevalence of a health condition or an environmental exposure and health outcome can be presented in many different ways. In those situations where primary studies are available, it is possible to conduct evidence synthesis using the primary studies based on systematic reviews and meta-analyses. In systematic reviews, such syntheses follow the steps of framing an answerable question using the population under consideration, the specific intervention or the exposure, the comparison groups, and the specific outcomes. Beyond identifying the studies and a narrative summary of the key findings, no specific summary statistic of the findings from the primary studies is presented (Khan et al., 2003). In a meta-analysis, on the other hand, in addition to narrative summaries, findings from the primary studies are combined to produce summary estimates; besides, in a meta analysis a formal statistical examination of the biases involved in selecting only those studies that were large in size and had positive outcomes is also conducted – this is referred to as "publication bias". Further, a meta analysis provides an opportunity to examine how the findings would differ or examination of the relationships between different entities in the analyses, referred to as sensitivity analyses or meta-regression (Deeks et al., 2019). In meta-analysis and systematic review, the analyst aims to collect and process the primary sources and then critically appraise the body of the literature to identify biases in the conduct of the study or whether a cause and effect association can be justified based on the nature of the studies, and also attempts to conduct a narrative or a statistical summary of the key findings.

Sometimes this is not necessarily the aim of the research, in particular when information about an exposure and an outcome is emerging on a particular topic. In those circumstances, the strategy is to identify what literature are available out there without necessarily conducting either a full critical appraisal of the body of the literature, or statistically summarising the findings. The goal in these situations is to identify the body of evidence that has accrued and the populations studied, the interventions or the exposures assessed, the comparison groups studied, and the key outcomes. In addition, analysts may also be interested to study the data sources, and other details that are not necessarily the main aims of a formal systematic review. Here, rather the aim is to "map" the "landscape" of the research. This is where an "evidence map" or "evidence gap map" is useful, as these information are useful for the analyst or the researcher to identify "what is out there" and what gaps in the available evidence exists.

Hence an evidence map is about laying the "landscape" of what evidence exists about a research question and they provide a visual overview of existing evidence. A World Bank document on evidence gap maps states that the idea is to conduct a visual summary of the evidence. In the context of environmental health, defining evidence maps, the journal "Environmental Evidence" states,

> Systematic Maps are overviews of the distribution and abundance of evidence in relation to multifaceted elements of a broad question of policy or management relevance. The process and

rigour of the mapping exercise is the same as for systematic review except that no evidence synthesis is attempted to seek an answer to the question.

For example, suppose you are interested to find out what recent evidence exists in terms of heart disease risks of exposure to ambient air pollution in the studies published in the past one year, but you do not want to critically appraise each article or formally synthesise the literature, you can construct an evidence map about how many studies on heart health effects of air pollution were reported in the scholarly literature databases since first January of 2019 (and up to a certain date of your choice, say 31st December 2019), in order to understand the landscape of the current research and then on the basis of your findings, you could construct a visual overview of these topics. These might include for example, how many studies were published on the relationship between ambient air quality and exposure to particulate pollution and people suffering from heart attacks (acute myocardial infarction), how many studies investigated the association between long term exposure to air pollution and ischaemic heart disease. You would then construct the maps of these various considerations and would be able to obtain a snapshot of the research and policy landscape. What you would not do would be to summarise the results of these studies. That would be the objective of a scoping review rather than an evidence map (Munn et al., 2018). So in summary:

- An evidence map is a "map" like structure
- It uses some principles of systematic review
- It provides a list of articles and resources
- It provides a visual summary in the form of tables and graphs as to what evidence are available on a particular topic
- There is no data analysis or narrative summary of the findings involved unlike that of a scoping review
- There may be some quality appraisal, but it is not the main purpose of an evidence map unlike that of systematic review or meta-analysis.

In this tutorial, I'd like to present the steps of a rapid evidence map construction. For full details and more detailed work, please review the list of citations and references I have included at the end of this document. Here, we will aim to construct a rapid and "miniaturised" evidence map on an environmental topic. Using the skills of searching literature and some data analyses such as using a spreadsheet or an open source statistical tool such as R or Python, you can construct a visual or knowledge map or evidence map.

Steps of constructing an evidence map

Step 1: Frame a PECO or PICO (population, exposure or intervention, comparison group, outcomes) formatted question

This is the first step. Develop a question based on the four criteria:

P: represents population or people.

Who are the human beings or which population will benefit from this question? For example, if you are interested in learning about what evidence exists that exposure to air pollution is associated with increased heart disease risks among elderly, then the population should reflect elderly (those above 65 years old). You can further narrow down the population to mean certain ethnicity (Whites or Maori in New Zealand), or only men or only women, or restrict the age band to above 65 years and below 85 years to exclude the old olds, or those who live in cities and towns and not in villages and so on.

E or I: If you are studying the effects of interventions, then the preferred entity would be interventions, otherwise E.

Suppose you want to study the effectiveness of improved stoves on respiratory illnesses as improved stoves are known to reduce indoor air pollution (Smith et al., 2007). Besides, you know that indoor air pollution is a risk factor for heart disease (Fatmi & Coggon, 2016). In that situation, you will use an "intervention".

Alternatively, you can also use E or exposure, if all you are interested in studying what are the heart health risks of those who are exposed to air pollution.

Comparator or comparison group or C:

The comparison group will be in **comparison with the exposure.** So, if you are studying exposure to air pollution and heart disease among elderly people, then your comparison group will be elderly people who are either NOT exposed to air pollution at all, or if that is not possible to lesser extent of air pollution (this is perhaps more realistic). If you are studying effectiveness of improved stoves as intervention and reduced heart disease risk in the elderly, then improved stoves is the intervention whose effect you want to study. Your comparison group will be those who do not use improved stoves in the household.

O is for outcomes

After you have defined P, E or I, and C, define the outcomes that you want to study. For example, in our case we want to study what evidence exists in the literature since 2019 about the excess heart disease risk among people who are exposed to air pollution? Outcome for us can be defined in terms of heart disease risk or death from heart disease. This might include acute myocardial infarction (also abbreviated as AMI, or known as heart attack or stroke), or hospital admission with AMI. We have mentioned this as "heart disease". However, the more precise you can be in defining the outcome, the more focused will be your review.

Once you have settled on the P, E or I, C, and O, it is time for you to put them together in the form of a table and a question. The question is referred to as PICO (or PECO) question. You accompany this question by developing a table (Table 1). To give you an example, we have put together a question and a table for our evidence map of all studies on air pollution and heart disease published since 2019.

Here is the PICO formatted question

> "*Compared with those with lower levels of exposure to air pollutants, those who are exposed to higher levels of air pollution, what is the risk of heart disease among older adults?* "

Here is the accompanying table:

Person	Elderly people above 65 years, both sex, all ethnicity
Exposure	High index of air pollution
Comparator	Lower index of air pollution
Outcome	Heart disease any type (acute myocardial infarction, chronic heart disease, others)

Table 1: PICO criteria

After you set up this table, set up some more criteria that will help you to narrow down your search and retrieval of literature that you will use for your evidence map. For example, in our study we have decided that we will include only the following:

- All studied published since 2019
- Every study must only be published in English language or an English translation must be available
- All studies must be peer reviewed and standard publications (this means we will exclude letters to the editor, editorials, popular non-peer reviewed magazine articles, books, book chapter, monographs that are not primary studies, etc).
- All studies must be freely available in full text format

You can create further restrictions or you can relax restrictions as the case may be. The less restrictions you will have , the more wide you can cast your "net" for fishing articles, and the more number of articles and primary sources you will use.

Step 2: Search literature

After you complete the first step of framing a pico formatted question, you will need to formulate search terms. The search terms can be framed in plain English, and you will also need to access the Pubmed MeSH vocabulary (link: https://www.ncbi.nlm.nih.gov/mesh) to identify what might be correct term to search. The search terms can be anywhere in the following areas of a paper or a monograph or a research report:

- Terms can be found in the title, abstract, or text body
- Terms can be combined using a logic referred to as Boolean logic (if you do not know what is meant by Boolean logic, make sure to read this article: https://www.lotame.com/what-is-boolean-logic/)
- You will need to use synonyms of the terms. For example, instead of heart disease, you can use "acute myocardial infarction", or "AMI", or cardiac illness and other terms. It is a good idea to spend time to identify the various synonyms that you are likely to come across.

After you set up the terms, think of which databases you will search. If you are working with environmental health related data sets, I recommend you search at least:

- Google Scholar:
- Pubmed/Medline
- University Library (University of Canterbury Library databases for example)
- Microsoft Academic
- Agricola
- Web of Science
- Other databases that are relevant to your discipline
- Campbell Collaboration database
- Cochrane Collaboration database

For example, for our study, we used the following search terms and we only searched Google Scholar. But in your case, you may want to search more than one database and want to include more search terms in various combinations. Note that the boolean operators "AND", "NOT", and "OR" when they are used, are used in capital letters or upper case letters.

These are our search terms

"air pollution" OR "criteria pollutants" OR "PM*"

AND "heart disease" OR "acute myocardial infarction" OR chronic heart disease"

Step 3: Read the abstract and select studies for reviewusing the ROSES form

Once you have obtained a set of titles and abstracts based on your literature database search, read the titles and abstracts of every study one by one. When you do so, read them with your inclusion and exclusion list in hand. Exclude the studies one by one if they do not meet the criteria of focusing on the topic of your interest, or if they have wrong population, of if they have wrong intervention or exposure, and so on. Make a note of how many studies you originally selected and how many studies you ended up in your data set that you will analyse.

When you have identified the studies, download or obtain their full texts and read the full texts of the papers. Then, read their reference lists and identify any potential study that will be useful. Obtain the title and abstract of that study and check if that paper can be included. If so, then obtain the full text and read. This is an iterative process, and as you do so, keep a tally of how many studies you started with, how many studies were rejected based on the various grounds of rejection, and how many you finally ended up with. For environmental health, use the ROSES (Reporting Standard for Systematic Evidence Synthesis) Protocol (URL: https://www.roses-reporting.com/systematic-review-protocols) . Download the spreadsheet from the above site and fill in the ROSES checklist and the flowchart maker.

In our study, we did not exactly follow this sequence as we identified 10 studies published since 2019 and we decided to use all of them. We also did not read the reference lists of all the studies, and hence we did not exactly follow the process we outlined above.

Step 4: Create a data extraction form

Once you have settled on the exact number of studies, create a data extraction form. You can use a spreadsheet for this purpose, typically something like an Excel spreadsheet or a Google Sheets. The spreadsheet or the sheets app will enable you to create forms that you can distribute to research students or others such as you colleagues to enter data there. Or you can directly use a spreadsheet to enter data.

What elements will go into the form is subjective but below we recommend a minimum set of variables that must go into the form.

For this project we have the following items to be extracted from each study:

- **study_id**: based on the study outcome, last name of the first author, and the year
- **author**: name the first or corresponding author
- **outcome**: specify the outcome studied
- **population**: specify what population was studied
- **exposure**: specify and describe the particular exposure studied
- **comparison** group: what was compared with respect to the exposure (the comparator must be described with respect to the exposure variable)
- **year** of the study: specify which year the study was completed
- **finding**: summary finding (one word)
- **study design**: what epidemiological study design was used to study the association between exposure to air pollution and heart disease

After this step, we obtained the following spreadsheet

study_-id	au-thor	outcome	popula-tion	exposure	com-pari-son	year	finding	design	eth-nic-ity
warb_-hrv_-2019	war-bur-ton	heart rate variability	Unde-fined, those with CAD	particu-late air pollution	vari-ous de-grees	2019	heart rate vari-ability	SR	mis-cel-la-neous
wang_-los_-2019	wang	length of stay	those with ihd	criteria pollutants	high ver-sus low	2020	length of stay	poisson regression	china
kon-duraka_-adm_-2019	kon-duracka	admissions	those with MI	criteria pollutants	high ver-sus low	2019	admis-sions	time series regression with poisson model	poland
bai_-ami_-2019	bai	ami	general members	criteria pollutants	high vs low	2019	ami in-cidence	cox ph	canada
bassig_-ihd_-2020	bassig	ihd	women	indoor air pollution from coal	smoke-less coal	2019	ihd in-cidence	cox ph	china
ljung-man_-ihd_-2019	ljung-man	ihd	general members	PM10, PM2.5, BC	high vs low	2019	ihd risk	cox ph	swe-den
zheng_-ami_-2020	zheng	ami admission	general members	criteria pollutants	high vs low	2020	ami admis-sion	poisson regression	china
liu_-ashd_-2019	liu	atheroscle-rotic heart disease deaths	those with ashd	criteria pollutants	high vs low	2017	deaths due to ashd	GAM time series	china
ho_-ami_-2019	ho	ami incidence	general members	criteria pollutants	high vs low	2015	ami in-cidence	case crossover study with poisson regression	sin-ga-pore

Table 2: Evidence map: master data

Step 5: Analyse the spreadsheet to identify various summary patterns

Steps 1 through 4 will provide you with a comprehensive, possibly large spreadsheet. You are now prepared to look in details into the nature and characteristics of the studies that you have included in your evidence map. In this step, you will summarise the characteristics of the papers you obtain.

You can do this using the spreadsheet or you can export the spreadsheet into a comma separated value file (CSV file), then import the spreadsheet into a statistical data analysis package and you will create tables, and charts to examine different aspect of the evidence that you have gathered. Here is a brief but incomplete list of the tables and graphs you will generate:

- A basic tabulation of the important variables (see Figure 1)

- Cross tabulation of important variables (Figure 2 and Figure 3)
- Different types of plots (Figure 4) of variables: this could be in the form of histograms, barplots, and plots of categorical variables.

design	n
\<chr\>	\<int\>
case crossover study with poisson regression	1
cox ph	3
GAM time series	1
poisson regression	2
SR	1
time series regression with poisson model	1

Figure 1: Distribution of the studies by types of data analysis

A tibble: 6 × 7

population	canada	china	miscellaneous	poland	singapore	sweden
\<chr\>	\<dbl\>	\<dbl\>	\<dbl\>	\<dbl\>	\<dbl\>	\<dbl\>
general members	1	1	0	0	1	1
those with ashd	0	1	0	0	0	0
those with ihd	0	1	0	0	0	0
those with MI	0	0	0	1	0	0
Undefined, those with CAD	0	0	1	0	0	0
women	0	1	0	0	0	0

Figure 2: Cross-tabulation of population by ethnicity of the studies conducted

A tibble: 8 × 7

finding	canada	china	miscellaneous	poland	singapore	sweden
<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
admissions	0	0	0	1	0	0
ami admission	0	1	0	0	0	0
ami incidence	1	0	0	0	1	0
deaths due to ashd	0	1	0	0	0	0
heart rate variability	0	0	1	0	0	0
ihd incidence	0	1	0	0	0	0
ihd risk	0	0	0	0	0	1
length of stay	0	1	0	0	0	0

Figure 3: Cross-tabulation of findings with ethnicity

Some simple two way plots are also possible (in this case, we have sparse data to work with but in most cases there are plenty of data where such plotting is possible):
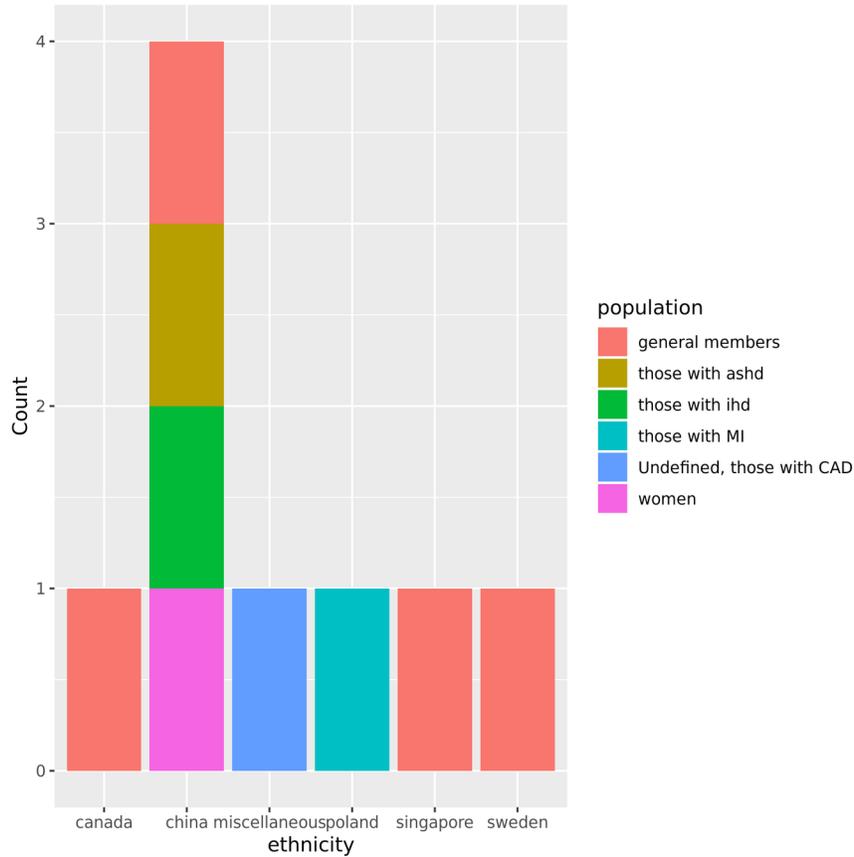
Figure 4: Plot of population with ethnicity

Structure of writing up the evidence map

Once you have generated the plots and tables, and charts, it is time to write the results. In the results section, write the distribution of the different types of studies that you have seen, and the results. Then identify and comment on the pattern of the evidence that you get to see. Finally, comment on the limitations and advantages of your evidence map.

For example, in our study, we see that most studies used Cox Proportional Hazards modelling to study the incidence of different types of heart disease in response to exposure to air pollution. Cox Proportional Hazards is a modelling strategy where "time to event" or hazard of events are analysed. These are useful analytical strategy for prospective and retrospective cohort studies and where events (such as hospital admissions or heart disease incidences) are reported (Thomas & Reyes, 2014). We also see majority of the studies came out of Chinese population (Figure 4) . A shortcoming of this rapid evidence map is that it is not comprehensive enough to provide a detailed break up of the different kinds of studies on air pollution and heart disease published throughout the world as it relied on only one source and was too restrictive. So while it may be useful for a short demonstration, it is neither comprehensive nor be useful to provide a detailed picture of the nature of evidence on the association between air quality and heart disease. A detailed ROSES map was also not provided.

9

Further reading

So, this was a short introduction to how to conduct an evidence map. We recommend you learn more about these in the following pages:

Systematic map at Environmental Evidence Site: https://environmentalevidencejournal. biomedcentral.com/submission-guidelines/preparing-your-manuscript/systematic-map provides detailed instructions on how to create evidence maps for publications

James and Haddaway (2016) provide a detailed methodology of writing a systematic map (James et al., 2016)

GEMS or the Global Evidence Mapping Initiative provides more detailed examples (Bragge et al., 2011)

List of papers used for our study:

Bai, Li, Saeha Shin, Richard T. Burnett, Jeffrey C. Kwong, Perry Hystad, Aaron van Donkelaar, Mark S. Goldberg, Eric Lavigne, Ray Copes, and Randall V. Martin. "Exposure to Ambient Air Pollution and the Incidence of Congestive Heart Failure and Acute Myocardial Infarction: A Population-Based Study of 5.1 Million Canadian Adults Living in Ontario." *Environment International* 132 (2019): 105004.

Bassig, Bryan A., H. Dean Hosgood, Xiao-Ou Shu, Roel Vermeulen, Bingshu E. Chen, Hormuzd A. Katki, Wei Jie Seow, Wei Hu, Lützen Portengen, and Bu-Tian Ji. "Ischaemic Heart Disease and Stroke Mortality by Specific Coal Type among Non-Smoking Women with Substantial Indoor Air Pollution Exposure in China." *International Journal of Epidemiology* 49, no. 1 (2020): 56–68.

Cohen, Gali, David M. Steinberg, Lital Keinan-Boker, Yuval, Ilan Levy, Shimon Chen, Rakefet Shafran-Nathan, Noam Levin, Tal Shimony, and Guy Witberg. "Preexisting Coronary Heart Disease and Susceptibility to Long-Term Effects of Traffic-Related Air Pollution: A Matched Cohort Analysis." *European Journal of Preventive Cardiology*, 2020, 2047487320921987.

Ho, Andrew Fu Wah, Huili Zheng, Arul Earnest, Kang Hao Cheong, Pin Pin Pek, Jeon Young Seok, Nan Liu, Yu Heng Kwan, Jack Wei Chieh Tan, and Ting Hway Wong. "Time-Stratified Case Crossover Study of the Association of Outdoor Ambient Air Pollution With the Risk of Acute Myocardial Infarction in the Context of Seasonal Exposure to the Southeast Asian Haze Problem." *Journal of the American Heart Association* 8, no. 6 (2019): e011272.

Konduracka, Ewa, \Lukasz Niewiara, Bartosz Guzik, Maksymilian Kotynia, Piotr Szolc, Grzegorz Gajos, Jadwiga Nessler, Piotr Podolec, and Krzysztof Żmudka. "Effect of Short-Term Fluctuations in Outdoor Air Pollution on the Number of Hospital Admissions Due to Acute Myocardial Infarction among Inhabitants of Krakow, Poland." *Polish Archives of Internal Medicine* 129, no. 2 (2019): 88–96.

Liu, Guangcong, Baijun Sun, Lianzheng Yu, Jianping Chen, Bing Han, Bo Liu, and Jie Chen. "Short-Term Exposure to Ambient Air Pollution and Daily Atherosclerotic Heart Disease Mortality in a Cool Climate." *Environmental Science and Pollution Research* 26, no. 23 (2019): 23603–23614.

Ljungman, Petter LS, Niklas Andersson, Leo Stockfelt, Eva M. Andersson, Johan Nilsson Sommar, Kristina Eneroth, Lars Gidhagen, Christer Johansson, Anton Lager, and Karin Leander. "Long-Term Exposure to Particulate Air Pollution, Black Carbon, and Their Source Components in Relation to Ischemic Heart Disease and Stroke." *Environmental Health Perspectives* 127, no. 10 (2019): 107012.

Lunny, Carole, Sue E. Brennan, Steve McDonald, and Joanne E. McKenzie. "Toward a Comprehensive Evidence Map of Overview of Systematic Review Methods: Paper 1—Purpose, Eligibility, Search and Data Extraction." *Systematic Reviews* 6 (November 21, 2017). https://doi.org/10.1186/s13643-017-0617-1.

Wang, Xuyan, Yong Yu, Chuanhua Yu, Fang Shi, and Yunquan Zhang. "Associations between Acute Exposure to Ambient Air Pollution and Length of Stay for Inpatients with Ischemic Heart Disease: A Multi-City Analysis in Central China." *Environmental Science and Pollution Research*, 2020, 1–12.

Warburton, Darren ER, Shannon SD Bredin, Erin M. Shellington, Christie Cole, Amanda de Faye, Jennifer Harris, David D. Kim, and Alan Abelsohn. "A Systematic Review of the Short-Term Health Effects of Air Pollution in Persons Living with Coronary Heart Disease." *Journal of Clinical Medicine* 8, no. 2 (2019): 274.

Zheng, Murui, Yanhong Zhang, Wenru Feng, Yuliang Chen, Lin Huan, Shu Ye, Jiagang Wu, Jun Huang, Yifu Liao, and Pi Guo. "Short-Term Exposure to Ambient Air Pollution and Acute Myocardial Infarction Attack Risk." *Journal of Public Health* 28, no. 4 (2020): 367–374.

References

Five steps to conducting a systematic review. (2003). *Journal of the Royal Society of Medicine*, *96*(3), 118–121.

Analysing data and undertaking meta-analyses. (2019). *Cochrane Handbook for Systematic Reviews of Interventions*, 241–284.

Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. (2018). *BMC Medical Research Methodology*, *18*(1), 143.

Monitoring and evaluation of improved biomass cookstove programs for indoor air quality and stove performance: conclusions from the Household Energy and Health Project. (2007). *Energy for Sustainable Development*, *11*(2), 5–18.

Coronary heart disease and household air pollution from use of solid fuel: a systematic review.. (2016). *Br Med Bull*, *118*, 91–109.

Tutorial: Survival Estimation for Cox Regression Models with Time-Varying Coefficients UsingSASandR. (2014). *Journal of Statistical Software*, *61*(Code Snippet 1). https://doi.org/10.18637/jss.v061.c01

A methodology for systematic mapping in environmental sciences. (2016). *Environmental Evidence*, *5*(1). https://doi.org/10.1186/s13750-016-0059-6

The Global Evidence Mapping Initiative: scoping research in broad topic areas.. (2011). *BMC Med Res Methodol*, *11*, 92.