

Systematic reviews: A brief historical overview

Quan Nha HONG^{a,b}, Pierre PLUYE^{a,b}

^aDepartment of Family Medicine, McGill University, 5858 Côte-des-Neiges, Suite 300, Montréal, QC, H3S 1Z1, Canada

^bMethod Development platform, Quebec SPOR SUPPORT Unit, 5858 Côte-des-Neiges, Suite 300, Montréal, QC, H3S 1Z1, Canada

Corresponding Author: Quan Nha Hong, Department of Family Medicine, McGill University, 5858 Côte-des-Neiges, Suite 300, Montréal, QC, Canada, H3S 1Z1; Tel: 1-514-398-8483, Fax: 1-514-398-4202. Email address: quan.nha.hong@mail.mcgill.ca

Abstract

Literature reviews, and more particularly systematic reviews are increasingly being produced and published. The past 40 years have been marked by considerable development of methodologies and methods in literature reviews. This paper aims to provide a brief historical overview of systematic reviews which will help to have a better understanding of what, why, when, and how they were developed. The paper is structured in three parts. The first part will provide a definition of systematic reviews and their main characteristics. The second part will present three main periods of the evolution of systematic reviews: foundation period (1970-1989), institutionalization period (1990-2000), and diversification period (2001-). These periods can be distinguished by the users of scientific evidence, the methodological influence, and the technological development. The last part will summarize the main elements in the history of the systematic reviews and present some direction for future research.

Keywords: Systematic review, literature review, history

1. Introduction

Literature reviews, and more particularly systematic reviews, are increasingly being produced and published. They can help keep up-to-date with scientific information since they consolidate information from several existing publications on a given topic (Baker, 2016; Cronin, Ryan, & Coughlan, 2008; Mulrow, 1994). Literature reviews represent a large part of the current published scientific literature and are the most cited type of paper (Patsopoulos, Analatos, & Ioannidis, 2005). Traditionally, literature reviews were aimed at summarizing prior research and establishing a state of knowledge on a given topic and used to justify and inform future research. However, with the increasing number of studies and the use of scientific information in decision-making, scientific methods have been developed to bring out new evidence from the synthesis of primary studies. The past 40 years have been marked by considerable development of synthesis methodologies and methods, and a proliferation of review types. Among these types, one that will be the focus of this paper is systematic review. This paper is the first of a three-part series on the topic of systematic reviews for information professionals.

Information professionals are more and more solicited to participate in the production of systematic reviews (Brettelle, 2009; Gore & Jones, 2015). A well-known role of information professionals in systematic reviews is expert searchers because of their skills in finding and managing information (Brettelle, 2009). However, they can play other roles and be involved in several other steps of systematic reviews. In a recent scoping review, Spencer and Eldredge (2018) identified 18 core roles of information professionals in conducting systematic reviews that range from planning to formulation of questions and evaluation. They can also be involved in developing tools to produce systematic reviews and teaching others on how to perform systematic reviews (Spencer & Eldredge, 2018).

Although systematic reviews have often been mainly associated with the field of health science, they have been applied in several other fields, from astronomy to zoology as claimed in Petticrew (2001). For example, a growing number of systematic reviews can be found in library and information studies (Xu, Kang, & Song, 2015) and information systems (Paré, Trudel, Jaana, & Kitsiou, 2015). Also, the first books on reviews were published by researchers in the field of education, social sciences, and political science (Cooper, 1984; Glass, McGaw, & Smith, 1981; Light & Pillemer, 1984; Ward & Reed, 1983).

Because information professionals are more and more involved in systematic reviews,

they need to have a better understanding of what systematic reviews are. This paper will provide a brief historical overview of systematic reviews, which will help to better understand what, why, when, and how they were developed. It is structured in three parts. The first part will provide a definition of systematic reviews and their main characteristics. The second part will present three main periods of the evolution of systematic reviews. Finally, the last part will summarize the main elements in the history of systematic reviews and present some directions for future research.

2. Definition

Literature review consists of synthesizing, summarizing, combining, analyzing, commenting, and criticizing studies on a given topic. It is considered to be secondary research, in which the unit of analysis is primary studies (Hong, Pluye, Bujold, & Wassef, 2017). There exists a variety of types of literature reviews. Grant and Booth (2009) suggest a typology of 14 different types that differ based on how the studies are identified, evaluated, synthesized and analyzed. The best known and used, especially by graduate students to justify the relevance of their project, is the so-called "traditional" literature review that presents a summary of knowledge on a given topic. Traditional literature reviews usually address broad review questions with no clear method for the selection, appraisal, and synthesis of studies (Pope, Mays, & Popay, 2007). This type of review was criticized for being subjective, scientifically unsound, and inefficient to extract information (Light & Pillemer, 1984). Traditional literature review is subject to several biases that can affect the results produced. Table 1 presents some biases that can be found when performing a literature review (Booth, Papaioannou, & Sutton, 2012; Higgins & Green, 2008; Light & Pillemer, 1984; Martin, Renaud, & Dagenais, 2013; Whiting et al., 2016).

Table 1. Main Biases in Literature Reviews

Bias	Definition	Strategies to minimize this bias
Identification	Not all studies on the subject of interest are found (not exhaustive search).	<ul style="list-style-type: none"> • Use several sources of information and bibliographic databases. • Involve a specialized librarian to develop a comprehensive search strategy.
Selection	Arbitrary selection of studies (e.g., a reviewer voluntary include or exclude studies to support position).	<ul style="list-style-type: none"> • Define clear selection criteria. • Involve a least two reviewers in the selection.
Reporting	The publication of research	<ul style="list-style-type: none"> • Use different sources of information.

	findings is influenced by the nature and direction of results (e.g., studies with positive results are more likely to be published).	<ul style="list-style-type: none"> • Include a variety of literature including the grey literature (e.g., theses and dissertations, reports, conference abstracts). • Contact authors of studies to obtain missing information.
Interpretation	The appraisal and synthesis of studies are influenced by a reviewer's subjectivity (e.g., preconceived ideas).	<ul style="list-style-type: none"> • Use critical appraisal tools. • Involve a least two reviewers in the synthesis and interpretation of findings.

To limit biases, a systematic approach to literature review was developed and advocated. Systematic review is defined as the “the application of strategies that limit bias in the assembly, critical appraisal, and synthesis of all relevant studies on a specific topic” (Chalmers, Hedges, & Cooper, 2002, p. 17). It follows an explicit, transparent, and reproducible process with the following characteristics (Pluye, Hong, Bush, & Vedel, 2016):

1. Specific/focused review question(s);
2. Pre-established precise eligibility criteria;
3. Exhaustive literature search using several sources of information;
4. Comprehensive and detailed search strategy designed with specialized librarians;
5. Reliable or dependable (performed by at least two researchers) selection of relevant studies, data extraction, and critical appraisal; and
6. Rigorous synthesis using specific qualitative, quantitative or mixed methods.

Nowadays, systematic reviews are considered the gold standard in literature reviews. They have raised the quality standards in the field of literature review and contributed to the emergence of a new discipline. They are one of the preferred methods for collecting scientific evidence to support the development of recommendations on best practices such as clinical practice guidelines (National Health and Medical Research Council, 1998), health technology assessments (Busse et al., 2002), and research synopses (Grad et al., 2008; Pluye et al., 2012). Numerous potential impacts of systematic reviews have been described in health research such as to (a) save lives by distinguishing what works and what is useless and even harmful, (b) save resources by avoiding unnecessary or unproven interventions as well as unnecessary duplication of studies, and (c) improve practice, clinical quality, and policies by providing up-to-date evidence on specific topics for decision-making (Bunn et al., 2015; Freeman, Williams, &

Dellavalle, 2006; Moynihan, 2004).

3. Historical overview

To better understand the origin and evolution of systematic reviews, a historical overview of the literature was performed. Systematic reviews have a long history. Back in the 18th century, Dr. James Lind emphasized the importance of producing a full and impartial critical view of the existing literature on the treatment of scurvy (Lind, 1757). Chalmers et al. (2002) traced some reviews published in the early 20th century conducted in several fields such as medicine, agriculture, physics, education, and social sciences. Although several papers advocating or using systematic methods to review the literature can be found in the first half of the 20th century, it is mainly in the 1970s that the science of research synthesis began to gain prominence from the need to apply explicit, transparent, and rigorous methods to enhance the validity of reviews (Chalmers et al., 2002). From this point on, it is possible to identify three major periods in the evolution of systematic reviews: (a) foundation (1970 - 1989), (b) institutionalization (1990 - 2000), and (c) diversification (2001 -). Figure 1 highlights some salient features of each period that are related to the users of scientific evidence, methodological influences, and technological development. The following sections will further describe each period.

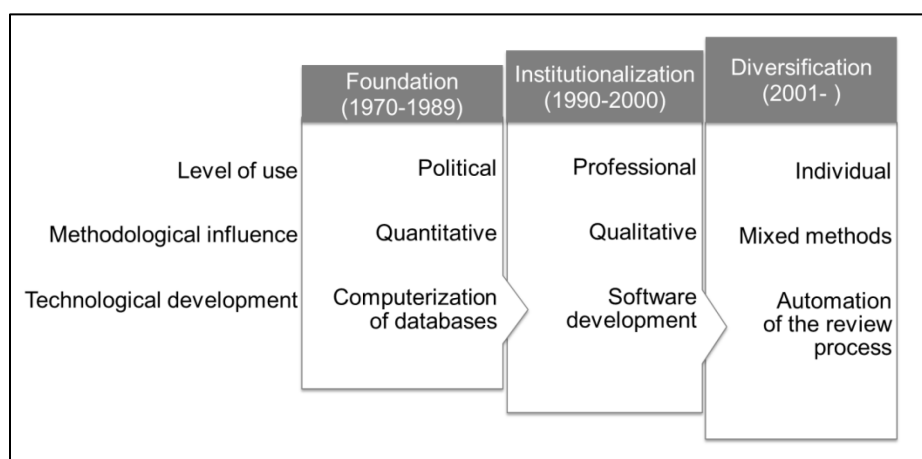


Figure 1. Main Periods in the History of Systematic Reviews

3.1 Foundation Period (1970 - 1989)

A growing interest in the use of reviews for policy development can be seen in the late 1960s, especially in educational policy (Light & Smith, 1971). This interest has had a major

impact on the methodological development of systematic reviews. In research, a literature review is an essential step for gathering information about what has been done on a subject and what results have been achieved. The results of a review can lead to identify areas of disagreement within the current state of knowledge, which researchers can use to justify the relevance of a project. The identification of contradictory studies on a subject is not necessarily perceived as problematic by researchers. However, when applied in the political field, inconsistencies between studies can be considered as a source of confusion and a hindrance to the development of clear policies (Light & Smith, 1971). To cope with this difficulty, there was a need to systematize the review process to limit the arbitrary selection of studies and increase rigor in the analysis of studies.

The growing interest in supporting the development of policies in healthcare based on scientific evidence had resulted, among other things, in the development of a multidisciplinary field called health technology assessment (HTA). In particular, HTA conducts systematic reviews to evaluate various dimensions of a technology (e.g., safety, effectiveness, cost-effectiveness, ethical implications) and to make policy recommendations about its use (Busse et al., 2002). Health technology is a very broad concept and can include drugs, devices, medical, and surgical procedures as well as the organizational and support systems within which health care is delivered (Busse et al., 2002). One of the first HTA agencies was established in 1972 in the United States (US Congressional Office of Technology Assessment) whose mandate was to advise the government on the use and application of health technologies (Banta, 2003).

Another milestone in the history of this foundation period is the computerization of bibliographic indexes that occurred primarily in the 1960s and 1970s. For example, one of the most popular bibliographic databases in medicine, the Index Medicus, was developed in 1879. Its computerized version was introduced in 1964 (MEDLARS), an online version was developed in 1971 (MEDLINE), and a free online version was made available in 1996 (PubMed) (Office of Health Technology Assessment, 1982). A similar development can be observed for other bibliographic databases in sciences such as the Sciences Citation Index inaugurated in 1963 (Office of Health Technology Assessment, 1982). This computerized access to bibliographic catalogues greatly simplified the task of searching for relevant studies and facilitated the production of systematic reviews.

At the beginning of this period, Light and Smith (1971) reported four main approaches for

combining studies: (a) listing all the factors that demonstrated an effect on the dependent variable in at least one study; (b) choosing one or two studies that stand out from the set of studies; (c) computing overall averages of results obtained in selected studies, and (d) taking a vote (vote-counting procedure), which consists in categorizing the studies in three groups (results that are significantly positive, significantly negative or no significant difference), counting the number of studies in each category and identifying the “winner” (i.e., the category with the largest number of studies). Some of these approaches were considered irrelevant for drawing comprehensive and representative conclusions from the studies reviewed. These approaches have paved the way for the development of aggregative synthesis approaches that consist in combining, statistically or not, the results of studies (Gough, Thomas, & Oliver, 2012).

During this period, there was a growth in the importance and number of randomized controlled trials (RCTs). In the 1960s and early 1970s, Dr. Donald Campbell (in social sciences) and Dr. Archie Cochrane (in medicine) advocated the use of RCTs as a standard for evaluating the effectiveness of an intervention; they considered RCTs as the most reliable source of scientific evidence (Campbell & Stanley, 1963; Cochrane, 1972). To synthesize RCTs, a statistical synthesis method to aggregate effect sizes of studies was developed. This method was named ‘meta-analysis’ for the first time in 1976 (Glass, 1976). Another synthesis methods developed during this period is quantitative case survey, which consists in analyzing the content of case studies by using closed-ended questions (Yin & Heald, 1975). This aggregative synthesis method was developed in political science in which case study is a commonly used methodology.

In summary, this foundational period was influenced by the need for reliable and reproducible reviews for policy development, the expansion of RCTs, and the computerization of bibliographic databases. It is also characterized by the development of aggregative synthesis methods to combine results of studies.

3.2 Institutionalization Period (1990 - 2000)

The institutionalization period is marked by the development of the evidence movement, which was introduced in the early 1990s and advocates the “conscientious, explicit and judicious use of the current best evidence in making decisions about the care of individual patients” (Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996, p. 71). Initially developed in the field of medicine (evidence-based medicine), this movement has spread to other areas than health such as

management (Lomas, Culyer, McCutcheon, McAuley, & Law, 2005a), librarianship (Eldredge, 2000), and education (Davies, 1999). In this evidence movement, the findings from systematic reviews are not only useful at a political level, but also at a professional level that uses evidence for decision-making by practitioners.

During this period, several organizations dedicated to the production of systematic reviews were established to promote evidence-based practice. One of the most important organizations developed in response to Dr. Archie Cochrane's call for systematic reviews of RCTs is the Cochrane Collaboration, which was founded in 1993 (Chalmers, 1993). This non-profit, non-governmental organization works with volunteer experts around the world to conduct systematic reviews of health interventions. Currently, more than 37,000 contributors from 130 countries are involved in the Cochrane Collaboration, making it the world's largest producer of systematic reviews in health care (The Cochrane Collaboration, 2017). A similar organization, the Campbell Collaboration, was inaugurated in 2000 in honour of Dr. Donald Campbell. This organization focuses on the effect of intervention in the areas of education, crime and justice, international development, and disability and social welfare (Schuerman et al., 2002).

Part of this organizational development, several guidelines on systematic reviews and tools to facilitate their production were created. For example, several critical appraisal tools have been developed to help reviewers judge the quality of RCTs (Moher et al., 1995). Also, levels of evidence (also named hierarchy of evidence) has been developed to help assess the scientific evidence (Evans, 2003). In these levels, systematic reviews with meta-analysis are assigned the highest level of evidence. Moreover, computer software has been created to facilitate the production of reviews (e.g., software for managing bibliographic reference, for coding studies, and for performing meta-analysis). These developments have contributed to the standardization of the systematic review procedure and to the growth in the number of reviews published. For instance, fewer than 10 systematic reviews could be annually found in the 1980s in the PubMed database, while their number has steadily increased since the 1990s, reaching more than 5,000 in 2000 and more than 28,000 in 2014 (Ioannidis, 2016). Also, Ioannidis (2016) compared the annual rate of systematic reviews published from 1991 to 2014. He found that the annual publication rate of systematic reviews increased by 2,728% in PubMed compared to only 153% of increase for all PubMed indexed items.

At the methodological level, an important methodological development was the use of

qualitative studies in reviews. At the end of the 1980s, researchers in social sciences were interested in including qualitative studies in reviews (Noblit & Hare, 1988). This led to the development of interpretive synthesis methods, which aim at generating new ways of understanding a phenomenon (Gough et al., 2012). Several interpretive (also called configurative) synthesis approaches were developed such as meta-ethnography (Noblit & Hare, 1988), meta-synthesis (Jensen & Allen, 1996; Sandelowski, Docherty, & Emden, 1997), and meta-study (Zhao, 1991).

Further methodological development of aggregative approaches, in particular of meta-analysis methods were also observed during this period (Sutton, Jones, Abrams, Sheldon, & Song, 1999). For example, one development consisted of testing Bayesian meta-analysis for incorporating qualitative and quantitative evidence in the context of the uptake of childhood immunization (Jones, Dixon-Woods, Abrams, & Fitzpatrick, 1999). Also, because RCTs are not always available and feasible in some fields, other aggregative synthesis methods have been developed such as cross-design synthesis for combining the results of studies with complementary designs (Droitcour, Silberman, & Chelimsky, 1993). Aggregative approach was also applied to the synthesis of qualitative studies. For example, techniques were developed to transform qualitative results into quantitative data such as quantitative content analysis that consists in assigning codes to qualitative results from a pre-established coding grid and performing statistical analysis on coded data (Neuendorf, 2002).

In summary, the period of institutionalization is marked by the establishment of various organizations promoting evidence-based practice, and developing standards and tools for producing systematic reviews. Besides from RCTs, other types of studies are considered in reviews, especially in social sciences, such as qualitative studies. Both aggregative and interpretive synthesis methods were developed.

3.3 Diversification Period (2001 -)

This period can be characterized by the diversification of synthesis methods and types of reviews as well as diversification of users. In addition to political and professional levels, the use of systematic reviews can also be seen at an individual level. For instance, consumers (e.g., patients) play a more active role in decision-making as part of person-centred care (Olsson, Jakobsson Ung, Swedberg, & Ekman, 2013). Also, during this period, more emphasis was put on

knowledge translation to close the gap between practice and research (Sudsawad, 2007). This trend influenced how systematic reviews are produced and disseminated. For example, studies have been conducted on the effectiveness of dissemination strategies to facilitate the uptake of systematic reviews (Tricco et al., 2015). Also, increasing involvement of stakeholders (e.g., consumers, clinicians, and policy-makers) in the production of systematic reviews has been advocated (Cottrell et al., 2014; Morley, Norman, Golder, & Griffith, 2016).

During this period, the number of scientific documents available has considerably increased. Systematic reviews are used to keep up-to-date and cope with the huge volume of scientific publications (Bastian, Glasziou, & Chalmers, 2010). For example, 10 million documents were indexed in PubMed in 2006 while this number exceeded 20 million in 2014. This means that it took less than 10 years (2007-2014) to reach the same number of documents (10 million) that took more than 100 years to index (1865 - 2006) (Bastian et al., 2010; U.S. National Library of Medicine, 2013). The number of scientific publications worldwide is estimated at 50 million and about 2.5 million articles are published every year in peer-reviewed journals (Bjork, Roos, & Lauri, 2009; Jinha, 2010). There are more than 33,000 peer-reviewed academic journals and this number continues to grow exponentially (Gu & Blackmore, 2016). For example, from 1986 to 2013, it was found that the growth rate of academic journals was 4.7% on average (Gu & Blackmore, 2016). This growth creates new challenges for the production of systematic reviews (e.g., increased time and resources for the selection of studies). With more literature to cover and more complex issues to consider, it is also required to produce and disseminate systematic reviews faster to inform policy-makers. New types of reviews have been developed to reduce their production time such as rapid review (i.e., a review produced in a shorter timeframe by simplifying or omitting some steps of a systematic review) (Tricco, Langlois, & Straus, 2017) and review of systematic reviews (i.e., a review compiling evidence from existing systematic reviews) (Robinson et al., 2016). Moreover, to reduce the burden of identifying and categorizing studies, researchers have been interested in studying the use of text mining tools, which involve machine learning to allow automation of specific systematic review tasks (Thomas, McNaught, & Ananiadou, 2011).

In the previous periods, systematic reviews mainly focused on RCTs. However, these reviews are problematic in areas where research is dominated by non-trial quantitative evidence such as public health, rehabilitation, and primary care (Goldsmith, Bankhead, & Austoker, 2007).

Also, these reviews only address effectiveness questions such as ‘Does it work?’ and ‘What works for whom?’. Besides from the effectiveness of an intervention, it was advocated to consider other issues such as its acceptability, applicability, feasibility, and transferability in different contexts (Petticrew et al., 2013; Shaw, Larkin, & Flowers, 2014). Other review questions can be asked such as ‘Why does it work?’, ‘How does it work?’, ‘In what context and when?’, and ‘What are the factors that promote or hinder the implementation of an intervention?’. These questions can be addressed by reviewing other types of studies such as qualitative and mixed methods studies. A new type of reviews was developed to provide a more complete picture of the evidence and to address different questions: mixed studies reviews, i.e., reviews of qualitative, quantitative, and mixed methods studies (Pluye & Hong, 2014). Mixed studies reviews combine the strengths of both qualitative and quantitative evidence, and provide in-depth answers to review questions involving complex phenomena (Petticrew et al., 2013; Pluye, Gagnon, Griffiths, & Johnson-Lafleur, 2009; Whitemore, Chao, Jang, Minges, & Park, 2014).

In the previous periods, the synthesis methods were mainly designed for one type of data (quantitative or qualitative). In this period, new interpretive methods for synthesizing qualitative and quantitative evidence were developed. The most common synthesis methods of qualitative and quantitative evidence are as follows:

- Meta-narrative synthesis is aimed at clarifying a construct by highlighting commonalities and differences among the multiple research traditions in which it has been studied (Greenhalgh et al., 2005).
- Critical interpretive synthesis is aimed at generating theory from a large body of evidence using meta-ethnographic methods such as reciprocal translational analysis (translate concepts across included studies), refutations within and across included studies, and lines of argument (building a general interpretation for representing all studies by constant comparisons between individual studies) (Dixon-Woods et al., 2006).
- Realist synthesis is middle range theory-driven; a middle range theory potentially explains all similar programs whereas a macro-theory explains something bigger than programs and the micro-logical model explains one individual program in one setting. It is aimed at testing a candidate theory or candidates, and producing a revised/improved middle range theory explaining why a program works. The latter is based on demi-regularities (patterns) identified

across ‘context-mechanism-outcome’ configurations observed within included studies (Pawson, Greenhalgh, Harvey, & Walshe, 2005).

Also, several papers on the conceptualization of the synthesis in mixed studies reviews were written (Frantzen & Fetters, 2015; Heyvaert, Maes, & Onghena, 2013; Hong et al., 2017; Sandelowski, Voils, Leeman, & Crandell, 2012). The conceptualization of mixed studies reviews was highly influenced by the development of mixed methods research, which combines methods of collecting and analyzing qualitative and quantitative primary data using experimentation, observation, and simulation (Abbott, 1998; Creswell & Plano Clark, 2018).

In summary, the conceptual and methodological development of systematic reviews during this diversification period was influenced by the importance of knowledge translation to all users (at political, professional, and individual levels), the explosion of the number of available documents, and the development of mixed methods research. New tools have been explored to automate the review process. Also, new types of reviews and synthesis methods have been developed to address different review questions, to accelerate the review process, and to deal with different types of data.

4. Conclusion

The purpose of this paper was to trace important milestones in the history of systematic reviews in order to better understand what, how, when, and why they were developed throughout the past decades. From this historical overview, it is possible to identify some salient features in each period that can be distinguished by the users of scientific evidence, methodological influence, and technological development (Figure 1). It should be noted that the periods overlap with each other, creating a cumulative effect. To conclude, we will summarize these historical milestones and present some research avenues that are currently pursued.

Over the past years, systematic reviews have been used for providing scientific knowledge in decision-making. This has been observed especially for policy-making (political level) in the foundation period and for decision-making of practitioners (professional level) in the institutionalization period with the development of the evidence movement. In the current period of diversification, reviews are read by anyone involved in the decision-making process including politicians, practitioners, and consumers. More and more, it is advocated to close the gap between researchers and stakeholders in order to increase the uptake of research findings and improve

their impact. However, there are tensions between research and practice. For example, the scientific values (e.g., rigor, precision, comprehensiveness) differ from the users' values (e.g., timescale, fitness for purpose, value for money) (Greenhalgh et al., 2005). To foster interaction between the worlds of research and practice, several strategies have been developed such as involving users in the different stages of the review process, from the conception to implementation, in order to guide the choice of review questions and interpretation of results emerging from the reviews. To this end, some have proposed the use of a deliberative process aimed at reaching consensus on recommendations from stakeholder consultation (practitioners and researchers) (Lomas, Cuyler, McCutchenon, McAuley, & Law, 2005b). This cooperation and co-production of knowledge makes it possible to ensure greater relevance of reviews and to promote their use.

At the methodological level, the evolution of research approaches has greatly guided the development of synthesis methods. Initially, the research methods were mostly quantitative in nature (in particular, RCT). Over time, qualitative research and mixed methods have developed to address different and complementary research questions. The development of the synthesis methods applied in systematic reviews also followed this evolution, starting with aggregative synthesis of quantitative studies, then interpretative synthesis of qualitative studies, and more recently, mixing different synthesis methods and types of studies. The next years are likely to be rich in conceptual and methodological development in systematic reviews, especially on how to combine various types of research as seen in mixed studies reviews. Also, the development of hybrid methods are needed for combining scientific and colloquial evidence (i.e., evidence issued from the expertise, views and realities of stakeholders) (Lomas et al., 2005a).

Finally, technology has played an important role in the history of systematic reviews. The web allows access to a mass of information in bibliographic databases as well as personal and institutional sites. This has helped to increase the volume of scientific information and has contributed to the need to synthesize it. The tremendous growth in the number of published documents is likely to continue. For example, it was estimated that at least 114 million English-language scholarly documents are available on the web (Khabisa & Giles, 2014). Aside from the increasing number of researchers and scientific journals (Ware & Mabe, 2015), this can also be explained by the pressure put on researchers and graduate students to publish ("Publish or Perish" (Rawat & Meena, 2014)), the complexity of the research questions studied, and the

multiplication of publications on a same study. Also, access to scientific knowledge has been facilitated not only for researchers, but also for the non-scientific community. In addition, computers and software make it possible to compile a large amount of data in a short period of time, which has facilitated the production of reviews and the analysis of complex and multiple data. The most recent technological contributions concern the automation of review production. Researchers are interested in speeding up the production of reviews using automation tools such as text mining and machine learning. Studies on these tools are being conducted to test their performance in executing several tasks in systematic reviews such as for the selection of relevant studies, the appraisal of the quality of the studies, and the extraction of data (Beller et al., 2018).

This first paper of this three-part series on systematic reviews for information professionals provided a historical overview of systematic reviews. The next two papers will focus on one type of review for complex issues: mixed studies reviews. More specifically, the second paper will present a toolkit for conducting and reporting mixed studies reviews (Pluye, Hong, Granikov, & Vedel, 2018) and this third paper will focus on a critical appraisal tool that was developed for use in this type of review (Hong et al., 2018).

5. Acknowledgement

Quan Nha Hong was a PhD candidate at the time of writing this manuscript and held a Doctoral Fellowship Award from the Canadian Institutes of Health Research (CIHR). Pierre Pluye holds a Senior Investigator Award from the Quebec Health Research Funds (FRQS) and is the Director for Method Development at the Quebec SPOR-SUPPORT Unit, which is funded by the CIHR, the FRQS, and the Quebec Ministry of Health.

Several parts of this paper are derived from one of the authors' dissertation: Hong, Q. N. (2018). *Revision of the Mixed Methods Appraisal Tool (MMAT): A Mixed Methods Study* (Doctoral dissertation). McGill University, Montréal, Canada.

6. References

- Abbott, A. (1998). The causal devolution. *Sociological Methods & Research*, 27(2), 148-181.
- Baker, J. D. (2016). The purpose, process, and methods of writing a literature review. *AORN Journal*, 103(3), 265-269.
- Banta, D. (2003). The development of health technology assessment. *Health Policy*, 63(2), 121-132.
- Bastian, H., Glasziou, P., & Chalmers, I. (2010). Seventy-five trials and eleven systematic reviews a day: How will we ever keep up? *PLoS Medicine*, 7(9), e1000326.
- Beller, E., Clark, J., Tsafnat, G., Adams, C., Diehl, H., Lund, H., et al. (2018). Making progress with the automation of systematic reviews: Principles of the International Collaboration for the Automation of Systematic Reviews (ICASR). *Systematic Reviews*, 7(1), 77.
- Bjork, B.-C., Roos, A., & Lauri, M. (2009). Scientific journal publishing: Yearly volume and open access availability. *Information Research: An International Electronic Journal*, 14(1), 1-14.
- Booth, A., Papaioannou, D., & Sutton, A. (2012). *Systematic approaches to a successful literature review*. London, UK: SAGE Publications.
- Brettell, A. (2009). Systematic reviews and evidence based library and information practice. *Evidence Based Library and Information Practice*, 4(1), 43-50.
- Bunn, F., Trivedi, D., Alderson, P., Hamilton, L., Martin, A., Pinkney, E., et al. (2015). The impact of Cochrane Reviews: A mixed-methods evaluation of outputs from Cochrane Review Groups supported by the National Institute for Health Research. *Health Technology Assessment*, 19(28), 1-100.
- Busse, R., Orvain, J., Velasco, M., Perleth, M., Drummond, M., Jørgensen, T., et al. (2002). Best practice in undertaking and reporting health technology assessments. *International Journal of Technology Assessment in Health Care*, 18(02), 361-422.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin Company.
- Chalmers, I. (1993). The Cochrane Collaboration: Preparing, maintaining, and disseminating systematic reviews of the effects of health care. *Annals of the New York Academy of Sciences*, 703(1), 156-165.
- Chalmers, I., Hedges, L. V., & Cooper, H. (2002). A brief history of research synthesis. *Evaluation and the Health Professions*, 25(1), 12-37.
- Cochrane, A. L. (1972). *Effectiveness and efficiency: Random reflections on health services*. London: Nuffield Provincial Hospitals Trust.
- Cooper, H. M. (1984). *The integrative research review: A systematic approach*: SAGE Publications.
- Cottrell, E., Whitlock, E., Kato, E., Uhl, S., Belinson, S., Chang, C., et al. (2014). *Defining the benefits of stakeholder engagement in systematic reviews*. Research white paper. Rockville, MD: Agency for Healthcare Research and Quality (AHRQ).
- Creswell, J. W., & Plano Clark, V. (2018). *Designing and conducting mixed methods research* (3rd ed.). Thousand Oaks, CA: SAGE Publications.
- Cronin, P., Ryan, F., & Coughlan, M. (2008). Undertaking a literature review: A step-by-step approach. *British journal of nursing*, 17(1), 38-43.
- Davies, P. (1999). What is evidence-based education? *British Journal of Educational Studies*, 47(2), 108-121.

- Dixon-Woods, M., Cavers, D., Agarwal, S., Annandale, E., Arthur, A., Harvey, J., et al. (2006). Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. *BMC Medical Research Methodology*, 6(35), 1-13.
- Droitcour, J., Silberman, G., & Chelimsky, E. (1993). Cross-design synthesis: A new form of meta-analysis for combining results from randomized clinical trials and medical-practice databases. *International Journal of Technology Assessment in Health Care*, 9(03), 440-449.
- Eldredge, J. D. (2000). Evidence-based librarianship: An overview. *Bulletin of the Medical Library Association*, 88(4), 289.
- Evans, D. (2003). Hierarchy of evidence: A framework for ranking evidence evaluating healthcare interventions. *Journal of Clinical Nursing*, 12(1), 77-84.
- Frantzen, K. K., & Feters, M. D. (2015). Meta-integration for synthesizing data in a systematic mixed studies review: Insights from research on autism spectrum disorder. *Quality & Quantity*, 50(5), 2251-2277.
- Freeman, S. R., Williams, H. C., & Dellavalle, R. P. (2006). The increasing importance of systematic reviews in clinical dermatology research and publication. *Journal of Investigative Dermatology*, 126(11), 2357-2360.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5(10), 3-8.
- Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research* (Vol. 56): Beverly Hills, CA: Sage Publications.
- Goldsmith, M. R., Bankhead, C. R., & Austoker, J. (2007). Synthesising quantitative and qualitative research in evidence-based patient information. *Journal of Epidemiology & Community Health*, 61(3), 262-270.
- Gore, G. C., & Jones, J. (2015). Systematic reviews and librarians: A primer for managers. *Partnership: The Canadian journal of library and information practice and research*, 10(1), 1-16.
- Gough, D., Thomas, J., & Oliver, S. (2012). Clarifying differences between review designs and methods. *Systematic Reviews*, 1(28), 1-9.
- Grad, R. M., Pluye, P., Mercer, J., Marlow, B., Beauchamp, M.-E., Shulha, M., et al. (2008). Impact of research-based synopses delivered as daily e-mail: A prospective observational study. *Journal of the American Medical Informatics Association*, 15(2), 240-245.
- Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information and Librarians Journal*, 26(2), 91-108.
- Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., Kyriakidou, O., & Peacock, R. (2005). Storylines of research in diffusion of innovation: A meta-narrative approach to systematic review. *Social Science & Medicine*, 61(2), 417-430.
- Gu, X., & Blackmore, K. L. (2016). Recent trends in academic journal growth. *Scientometrics*, 108(2), 693-716.
- Heyvaert, M., Maes, B., & Onghena, P. (2013). Mixed methods research synthesis: Definition, framework, and potential. *Quality & Quantity*, 47(2), 659-676.
- Higgins, J. P., & Green, S. (2008). *Cochrane handbook for systematic reviews of interventions*. Chichester, UK: Wiley Online Library.
- Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., et al. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information* (Special Issue).

- Hong, Q. N., Pluye, P., Bujold, M., & Wassef, M. (2017). Convergent and sequential synthesis designs: Implications for conducting and reporting systematic reviews of qualitative and quantitative evidence. *Systematic Reviews*, 6(61), 1-14.
- Ioannidis, J. (2016). The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Quarterly*, 94(3), 485-514.
- Jensen, L. A., & Allen, M. N. (1996). Meta-synthesis of qualitative findings. *Qualitative Health Research*, 6(4), 553-560.
- Jinha, A. E. (2010). Article 50 million: An estimate of the number of scholarly articles in existence. *Learned Publishing*, 23(3), 258-263.
- Jones, D., Dixon-Woods, M., Abrams, K., & Fitzpatrick, R. (1999). *Meta-analysis of qualitative and quantitative evidence*. Leicester, UK: University of Leicester.
- Khabisa, M., & Giles, C. L. (2014). The Number of Scholarly Documents on the Public Web. *PLOS ONE*, 9(5), e93949.
- Light, R. J., & Pillemer, D. B. (1984). *Summing up: The science of reviewing research*: Harvard University Press.
- Light, R. J., & Smith, P. V. (1971). Accumulating evidence: Procedures for resolving contradictions among different research studies. *Harvard Educational Review*, 41(4), 429-471.
- Lind, J. (1757). *A treatise on the scurvy. In three parts. Containing an inquiry into the nature, causes and cure of that disease. Together with a critical and chronological view of what has been published on the subject*. London: A. Millar in the Strand.
- Lomas, J., Culyer, T., McCutcheon, C., McAuley, L., & Law, S. (2005a). *Conceptualizing and combining evidence for health system guidance* Ottawa, ON: Canadian Health Services Research Foundation.
- Lomas, J., Culyer, T., McCutcheon, C., McAuley, L., & Law, S. (2005b). *Conceptualiser et regrouper les données probantes pour guider le système de santé*. Ottawa, ON: Fondation canadienne de la recherche sur les services de santé (FCRSS).
- Martin, V., Renaud, J., & Dagenais, P. (2013). *Les normes de production des revues systématiques: Guide méthodologique*. Montréal, QC: Institut national d'excellence en santé et en services sociaux (INESSS).
- Moher, D., Jadad, A. R., Nichol, G., Penman, M., Tugwell, P., & Walsh, S. (1995). Assessing the quality of randomized controlled trials: An annotated bibliography of scales and checklists. *Controlled Clinical Trials*, 16(1), 62-73.
- Morley, R. F., Norman, G., Golder, S., & Griffith, P. (2016). A systematic scoping review of the evidence for consumer involvement in organisations undertaking systematic reviews: Focus on Cochrane. *Research Involvement and Engagement*, 2(1), 36.
- Moynihán, R. (2004). *Evaluating health services: A reporter covers the science of research synthesis*. New York: Milbank Memorial Fund.
- Mulrow, C. D. (1994). Systematic reviews: rationale for systematic reviews. *British Medical Journal*, 309(6954), 597-599.
- National Health and Medical Research Council. (1998). *A guide to the development, implementation and evaluation of clinical practice guidelines*. Canberra, AU: National Health and Medical Research Council (NHMRC).
- Neuendorf, K. A. (2002). *The content analysis guidebook*. Thousand Oaks, CA: SAGE Publications.
- Noblit, G. W., & Hare, R. D. (1988). *Meta-ethnography: Synthesizing qualitative studies* (Vol. 11). Thousand Oaks, CA: SAGE Publications.

- Office of Health Technology Assessment. (1982). *MEDLARS and health information policy: A technical memorandum*. Washington, DC: US Government Printing Office.
- Olsson, L. E., Jakobsson Ung, E., Swedberg, K., & Ekman, I. (2013). Efficacy of person-centred care as an intervention in controlled trials – A systematic review. *Journal of Clinical Nursing*, 22(3-4), 456-465.
- Paré, G., Trudel, M.-C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, 52(2), 183-199.
- Patsopoulos, N. A., Analatos, A. A., & Ioannidis, J. P. (2005). Relative citation impact of various study designs in the health sciences. *Journal of the American Medical Association*, 293(19), 2362-2366.
- Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2005). Realist review - A new method of systematic review designed for complex policy interventions. *Journal of Health Services Research & Policy*, 10 (Suppl 1), 21-34.
- Petticrew, M. (2001). Systematic reviews from astronomy to zoology: Myths and misconceptions. *British Medical Journal*, 322(7278), 98.
- Petticrew, M., Rehfuess, E., Noyes, J., Higgins, J. P., Mayhew, A., Pantoja, T., et al. (2013). Synthesizing evidence on complex interventions: How meta-analytical, qualitative, and mixed-method approaches can contribute. *Journal of Clinical Epidemiology*, 66(11), 1230-1243.
- Pluye, P., Gagnon, M. P., Griffiths, F., & Johnson-Lafleur, J. (2009). A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in mixed studies reviews. *International Journal of Nursing Studies*, 46(4), 529-546.
- Pluye, P., Grad, R., Granikov, V., Theriault, G., Fremont, P., Burnand, B., et al. (2012). Feasibility of a knowledge translation CME program: Courriels Cochrane. *Journal of Continuing Education in the Health Professions*, 32(2), 134-141.
- Pluye, P., & Hong, Q. N. (2014). Combining the power of stories and the power of numbers: Mixed methods research and mixed studies reviews. *Annual Review of Public Health*, 35, 29-45.
- Pluye, P., Hong, Q. N., Bush, P. L., & Vedel, I. (2016). Opening-up the definition of systematic literature review: The plurality of worldviews, methodologies and methods for reviews and syntheses. *Journal of Clinical Epidemiology*, 73(5), 2-5.
- Pluye, P., Hong, Q. N., Granikov, V., & Vedel, I. (2018). The wiki toolkit for planning, conducting and reporting mixed studies reviews. *Education for Information* (Special Issue).
- Pope, C., Mays, N., & Popay, J. (2007). *Synthesizing qualitative and quantitative health research: A guide to methods*. Maidenhead, UK: Open University Press.
- Rawat, S., & Meena, S. (2014). Publish or perish: Where are we heading? *Journal of Research in Medical Sciences*, 19(2), 87.
- Robinson, K. A., Chou, R., Berkman, N. D., Newberry, S. J., Fu, R., Hartling, L., et al. (2016). Twelve recommendations for integrating existing systematic reviews into new reviews: EPC guidance. *Journal of Clinical Epidemiology*, 70(2), 38-44.
- Sackett, D. L., Rosenberg, W. M., Gray, J., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. *British Medical Journal*, 312(7023), 71-72.
- Sandelowski, M., Docherty, S., & Emden, C. (1997). Focus on qualitative methods - Qualitative metasynthesis: Issues and techniques. *Research in Nursing and Health*, 20, 365-372.

- Sandelowski, M., Voils, C. I., Leeman, J., & Crandell, J. L. (2012). Mapping the mixed methods–mixed research synthesis terrain. *Journal of Mixed Methods Research*, 6(4), 317-331.
- Schuerman, J., Soydan, H., Macdonald, G., Forslund, M., de Moya, D., & Boruch, R. (2002). The Campbell collaboration. *Research on Social Work Practice*, 12(2), 309-317.
- Shaw, R. L., Larkin, M., & Flowers, P. (2014). Expanding the evidence within evidence-based healthcare: Thinking about the context, acceptability and feasibility of interventions. *Evidence-Based Medicine*, 19(6), 201-203.
- Spencer, A. J., & Eldredge, J. D. (2018). Roles for librarians in systematic reviews: A scoping review. *Journal of the Medical Library Association*, 106(1), 46.
- Sudsawad, P. (2007). *Knowledge translation: Introduction to models, strategies, and measures*. Austin, TX: Southwest Educational Development Laboratory, National Center for the Dissemination of Disability Research.
- Sutton, A. J., Jones, D. R., Abrams, K. R., Sheldon, T. A., & Song, F. (1999). Systematic reviews and meta-analysis: A structured review of the methodological literature. *Journal of Health Services Research & Policy*, 4(1), 49-55.
- The Cochrane Collaboration. (2017). *About us*. Retrieved June 18, 2017, from <http://www.cochrane.org/ca/about-us>
- Thomas, J., McNaught, J., & Ananiadou, S. (2011). Applications of text mining within systematic reviews. *Research Synthesis Methods*, 2(1), 1-14.
- Tricco, A. C., Cardoso, R., Thomas, S. M., Motiwala, S., Sullivan, S., Kealey, M. R., et al. (2015). Barriers and facilitators to uptake of systematic reviews by policy makers and health care managers: A scoping review. *Implementation Science*, 11(4), 1-20.
- Tricco, A. C., Langlois, E. V., & Straus, S. E. (2017). *Rapid reviews to strengthen health policy and systems: A practical guide*. Geneva: World Health Organization.
- U.S. National Library of Medicine. (2013, February 20). *Fact sheet - MEDLINE®*. Retrieved April 5, 2014, from <http://www.nlm.nih.gov/pubs/factsheets/medline.html>
- Ward, S. A., & Reed, L. J. (1983). *Knowledge structure and use: Implications for synthesis and interpretation*. Philadelphia, PA: Temple University Press.
- Ware, M., & Mabe, M. (2015). *The STM report: An overview of scientific and scholarly journal publishing*. The Netherlands: STM: International Association of Scientific, Technical and Medical Publishers.
- Whiting, P., Savovic, J., Higgins, J. P., Caldwell, D. M., Reeves, B. C., Shea, B., et al. (2016). ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *Journal of Clinical Epidemiology*, 69(1), 225-234.
- Whittemore, R., Chao, A., Jang, M., Minges, K. E., & Park, C. (2014). Methods for knowledge synthesis: An overview. *Heart & Lung*, 43(5), 453-461.
- Xu, J., Kang, Q., & Song, Z. (2015). The current state of systematic reviews in library and information studies. *Library & Information Science Research*, 37(4), 296-310.
- Yin, R. K., & Heald, K. A. (1975). Using the case survey method to analyze policy studies. *Administrative Science Quarterly*, 20(3), 371-381.
- Zhao, S. (1991). Metatheory, metamethod, meta-data-analysis: What, why, and how? *Sociological Perspectives*, 34(3), 377-390.