Weight of Evidence:
Participatory Methods and Bayesian Updating to
Contextualize Evidence Synthesis in Stakeholders’ Knowledge

Anna Dion1, Alessandro Carini-Gutierrez1, Vania Jimenez1, Amal Ben Ameur2, Emilie Robert3,4, Lawrence Joseph1, and Neil Andersson1,5

Abstract
Mixed methods research is well-suited to grapple with questions of what counts as valid knowledge across different contexts and perspectives. This article introduces Weight of Evidence as a transformative procedure for stakeholders to interpret, expand on and prioritize evidence from evidence syntheses, with a focus on engaging populations historically excluded from planning and decision making. This article presents the procedure’s five steps using pilot data on perinatal care of immigrant women in Canada, engaging family physicians and birth companions. Fuzzy cognitive mapping offers an accessible and systematic way to generate priors to update published literature with stakeholder priorities. Weight of Evidence is a transparent procedure to broaden what counts as expertise, contributing to a more comprehensive, context-specific, and actionable understanding.

Keywords
knowledge synthesis, mixed methods research, Bayesian updating, fuzzy cognitive mapping, realism

Background
Recent advances in mixed methods reviews demonstrate the value of combining qualitative and quantitative findings, often derived from differing perspectives and epistemologies, in evidence

1Department of Family Medicine, McGill University, Montréal, Québec, Canada
2CARE Canada, Ottawa, Ontario, Canada
3SHERPA University Institute, Montreal, Quebec, Canada
4ICARES Montreal, Quebec, Canada
5Autonomous University of Guerrero, Acapulco, Mexico

Corresponding Author:
Anna Dion, Department of Family Medicine, McGill University, 5858 chemin de la Cotes-des-Neiges, Montréal, Québec H3S 1Z1, Canada.
Email: anna.dion@mail.mcgill.ca
syntheses (Batalden & Davidoff, 2007; Greenhalgh et al., 2016; Harris et al., 2016; Pluye & Hong, 2014). Knowledge synthesis approaches such as critical interpretive synthesis, realist reviews, and narrative reviews, offer rich interpretations sometimes across different paradigms (Dixon-Woods et al., 2006; Pawson et al., 2006; Popay et al., 2006; Sandelowski, 1991). This can be complex when translating and synthesizing evidence from differing perspectives, requiring a diversity of concepts, theories and methods that maybe at odds with one another (Di Ruggiero, 2018; Mertens & Hesse-Biber, 2013).

Wrestling with questions of what counts as valid knowledge requires examining evidence in context and engaging multiple perspectives on complex social problems, to which mixed methods research is uniquely suited (Di Ruggiero, 2018; Mertens & Hesse-Biber, 2013). There are several tools and approaches to assess quality and trustworthiness of evidence during evidence synthesis. These approaches are often driven by adherence to methodological steps to assure the quality and rigour of evidence with little questioning of who decides what is good evidence for a specific context and how.

This article introduces Weight of Evidence as a transformative procedure to contextualize evidence in the understanding of relevant stakeholders. Weight of Evidence presents a transparent and systematic procedure for stakeholders to interpret, expand on and prioritize evidence from synthesis, with a particular focus on engaging those historically excluded from planning and decision-making. Weight of Evidence can inform local service improvements, program design or evaluation as well as to refine syntheses efforts. This procedure may be interest to mixed methods researchers as it draws on and operationalizes Greene and Hall’s dialectic stance, where differences in understanding are generative, contributing to causal understanding and ultimately lead to more informed and consultative decision making (Befani & Stedman-Bryce, 2017; Greene & Hall, 2010).

This article begins with a description of Weight of Evidence’s philosophical orientations and methodological approaches. Each of the procedure’s five steps is outlined together with a brief description of how it was applied through a pilot study with family physicians and birth companions around unmet perinatal care needs of recent immigrant women in Canada. Finally, the contributions to mixed methods research accompany considerations for researchers interested in applying this procedure in their own work.

Method

Philosophical Pillars of Weight of Evidence

Three complementary approaches orient Weight of Evidence. The first is rooted in transformative participatory research that enacts the principle that people have a right to be involved in decisions that shape their lives (Borda, 1996; Mertens & Hesse-Biber, 2013; L. T. Smith, 2012). Participation makes research processes and interventions more relevant to local needs and priorities, and therefore, more effective (George et al., 2017). Commitment to participatory research stems from the belief that people make better decisions when they benefit from both scientific and more informal forms of knowledge. This includes contributions from evidence transferred through theoretical or statistical inferences, often developed through empirical studies or syntheses. It also includes context-specific understanding, meaning knowledge claims based on local settings, experience, and tacit understanding of practice and organizational “know how” (Mertens, 2003; Oliver et al., 2018). Weight of evidence incorporates stakeholder perspectives on even footing with evidence synthesized from the literature.
A second orientation of the Weight of Evidence is a critical realist philosophy, which recognizes there is a real world with which we interact, though may never truly know, and that one’s social position and context affect how they understand and navigate through the world (Bhaskar, 2008; Maxwell & Mittapalli, 2010). Weight of Evidence is guided by critical realist reasoning in leveraging qualitative, quantitative, and stakeholder-derived understanding to develop causal explanations (Maxwell & Mittapalli, 2010).

A third orientation is standpoint theory, recognizing that all knowledge is socially situated and that one’s social position and context affect how they understand and navigate through the world. Standpoint theory suggests that disadvantaged groups have critical perspectives to offer in understanding the status quo as they navigate more mainstream systems from a position of marginalization (Collins, 1986; Harding, 2003). This is consistent with arguments long-made by leading feminist, Indigenous, disability-rights and working-class academics, activists, and communities, as well as within mixed methods research (Collins, 1986; Harding, 2003; Lavell-Harvard & Anderson, 2014; Maxwell & Mittapalli, 2010; Tuck, 2008). While Weight of Evidence seeks to engage all relevant stakeholders around a particular issue, it emphasizes meaningful engagement of groups historically excluded from contextualization and decision-making opportunities.

Methodological Pillars of Weight of Evidence

Weight of Evidence is guided by three complementary methodologies. Bayesian statistics provide a formal procedure to learn from data (or knowledge) outside of conventional epidemiological models and incorporate this data together with established models (Gelman et al., 2013; Joseph, 2000a; Sprenger & Hartman, 2019). Statistically incorporating knowledge external to an empirical study is called Bayesian updating (Goldstein, 2006). Several studies have used Bayesian updating to combine qualitative and quantitative findings, either by quantifying qualitative data, coding the presence and absence of themes in both the qualitative and quantitative literature, or by drawing on qualitative data to create prior distributions (Crandell et al., 2012; Roberts et al., 2002; Voils et al., 2009). Disciplines outside of health science have also drawn on Bayesian statistics to contextualize published literature in end-user perspectives (Badampudi et al., 2019; Badampudi & Wohlin, 2016). Weight of Evidence applies Bayesian statistics to juxtapose and to combine stakeholder knowledge with empirical evidence reflected in a synthesis.

A second methodological pillar is boundary critique, a systems thinking concept that argues broadening system boundaries can help identify underlying values and offer insights about how a particular system works (Flaspohler et al., 2003; Foster-Fishman et al., 2007; Midgley, 2000). Examining multiple perspectives on complex issues can generate understanding through the comparison and sometimes combining of different ways of knowing and experiencing the same phenomenon (Greene & Hall, 2010; Ulrich, 1998). Weight of Evidence operationalizes boundary critique to broaden what counts as relevant expertise.

A third methodological pillar of Weight of Evidence is graph theory, building on a long history of using diagrams to represent how things relate to one another, spatially or conceptually (Andersson & Silver, 2019). In Weight of Evidence, fuzzy cognitive maps serve as a critical translation tool to make both epidemiological data and knowledge from stakeholders available to decision makers (Kosko, 1988). Through data transformation, combining fuzzy cognitive maps brings different knowledge types in conversation with one another (Fetters et al., 2013; Kosko, 1988; U. Ózesmi & Ózesmi, 2004).

An overview of the Weight of Evidence procedure is provided in Table 1.
Preparatory Step: Identify Focus and Synthesize Published Literature

General description: Weight of Evidence begins once a systematic synthesis of published evidence is complete. While Weight of Evidence can be used with any type of question, it may be most useful for questions where considerable contextual understanding might influence intervention success and decision making, such as complex interventions (Booth et al., 2019).

In practice: We searched for primary studies describing perinatal (during pregnancy to 12 months postpartum) health outcomes and experiences of immigrant women within 5 years of...
their arrival to Canada published in French or English after the year 2000 in Medline, CINAHL, and Web of Science. Our search identified 91 publications. The lead author (AD) assessed all abstracts to determine eligibility and appraised eligible articles using the Joanna Briggs Institute (JBI) checklists for case control or cohort designs for quantitative studies and Interpretative & Critical Research (JBI QARI) for qualitative studies (Higginbottom et al., 2014; JBI, 2017a, 2017b, 2017c). Our protocol specified that studies would be excluded only if they had fatal flaws and scored below 60% on any scale; however, no studies met this criteria. We extracted data from 39 relevant articles (24 quantitative, 15 qualitative) using inductive thematic synthesis for qualitative findings and descriptive statistics for quantitative findings (Braun & Clarke, 2006; Pluye & Hong, 2014; Popay et al., 2006).

Among eligible articles, eight publications (five quantitative, three qualitative) focused on unmet postpartum care needs among recent immigrant women. This was selected as the outcome of interest for this pilot study because it is a clinically relevant outcome, addresses a core interest of our research group (the intersection of marginalization and perinatal health) and is well-described in the literature (Gagnon, Carnevale, et al., 2013; Gagnon, Dougherty, et al., 2013; Gagnon et al., 2007; Gagnon et al., 2010; Gagnon et al., 2012; Higginbottom, 2013; Higginbottom et al., 2014; Merry et al., 2011; Munro et al., 2014). Recent immigrant women to Canada are at increased risk of postpartum health concerns, including postpartum depression, abnormal blood pressure, maternal pain, and lack of access to contraception compared with their Canadian-born counterparts (odds ratio [OR] = 1.69, 95% confidence interval [CI: 1.46, 1.96]; Gagnon, Dougherty, et al., 2013). They are more than twice as likely to have these concerns unaddressed at 1 week (OR = 2.24, 95% CI [1.73, 2.9]) and 4 months postpartum (OR = 2.36, 95% CI [1.75, 3.17]) compared with Canadian-born women (Gagnon, Dougherty, et al., 2013).

**Step 1: Represent Evidence as Fuzzy Cognitive Map**

*General description:* Findings from the evidence synthesis are translated into a fuzzy cognitive map (Giles et al., 2008). Cognitive maps are made up of concepts or nodes (factors affecting the issue) and links that describe the relationships between factors that can be weighted by relative importance (Giles et al., 2007; U. Özesmi & Özesmi, 2004). To facilitate comparison, all effect estimates are converted to a shared format (e.g., odds ratio, relative risk; Bornstein & Hedges, 2019; Giles et al., 2008). Qualitative themes are included as “unattached” nodes when the included studies suggest a that a theme or concept contributes to the outcome of interest.

**Step 1 in Practice: Summarize Evidence as Fuzzy Cognitive Map**

All effect estimates were converted to an odds ratio (OR). When other statistics were presented, they were converted to the standardized mean difference (d), and then to an OR (Bornstein & Hedges, 2019). If multiple effect estimates described the same relationship, estimates were pooled using a Bayesian hierarchical random effects model with noninformative priors to account for within- and between-study sources of variation (Joseph, 2000a). Figure 1 shows the fuzzy cognitive map from the evidence synthesis of factors contributing to unmet postpartum care needs among recent immigrant women in Canada.

**Step 2: Stakeholders Generate Cognitive Maps**

*General description:* This step begins by identifying stakeholders—people (or groups of people) that have an interest or stake in the selected outcome. Who to include as stakeholders is a
question of relevance and expertise balanced with access, resource and equity considerations (Midgley, 2000; Wallerstein & Duran, 2006). This step builds on existing protocols to generate fuzzy cognitive maps (described in more detail below) but advances them in several ways (Andersson & Silver, 2019; U. Özesmi & Özesmi, 2004). Giles et al. (2008) had previously used fuzzy cognitive maps to represent published evidence. This approach is adapted in Weight of Evidence to make both quantitative and qualitative evidence accessible to stakeholders while not requiring extensive training in evidence synthesis methods. Stakeholders are invited to represent their own knowledge and understanding on that same issue in a systematic way (Giles et al., 2008). Data integration occurs as stakeholders adapt the map of published literature, adding or removing factors and/or the relationships between them.

Interviews start by asking participants to identify what they consider to be important factors contributing to the outcome of interest. They generate ideas independently and may be prompted to identify relevant social and structural influences based on factors identified in the literature. Participants record each of their ideas on small, laminated magnets and are introduced to the literature-derived cognitive map (created in the previous step) on a magnetic white board.

Stakeholders are invited to adapt the map, incorporating their own ideas, removing factors they consider irrelevant and labelling more magnets if necessary (Andersson & Silver, 2019; U. Özesmi & Özesmi, 2004). After grouping similar or synergistic factors, participants assign a weight and direction of effect (+ve or −ve, from 1 to 5) to each relationship in their map. Detailed notes or a recording capture the discussion during map construction and a photo is taken of each cognitive map to facilitate analysis.

Final stakeholder maps represent quantitative and qualitative data from the literature review, together with stakeholder-identified factors, with relationships between factors weighted for their importance in relation to the outcome. As maps are created by different stakeholders and
stakeholder groups, they can be combined together and/or compared with one another as well as with the fuzzy cognitive map of evidence from the literature (Kosko, 1986, 1988).

**Step 2 in practice:** We selected stakeholders for our pilot study based on expertise and interest in contributing to our pilot study. We interviewed three family physicians specialized in perinatal care of immigrant and refugee women with between 5 years to over 20 years, as well as two birth companions, with 5- and 8-years experience in perinatal support to refugee and recent immigrant women. Birth companions provide physical, emotional and informational support, and help families navigate health and social services related to their pregnancy and early postpartum period (Mahoney & Mitchell, 2016). The focus of this work was to demonstrate the Weight of Evidence procedure rather than to contribute directly to service improvements. Given there was no budget or organizational commitment to action the recommendations emerging from this study, our research group decided not to directly involve recent immigrant women. While not representing the perspectives of recent immigrant women, the inclusion of family physicians and birth companions provided a helpful example of how including different perspectives about a complex issue can provide valuable insight into understanding a problem and how these understandings together generate a more comprehensive understanding of the problem at hand.

The lead author carried out all semi-structured mapping interviews with birth companions and was joined by ACG (a family medicine resident) for the interviews with family physicians. Each mapping interview took approximately 2 hours. Fuzzy cognitive maps generated by family physicians and birth companions (Figure 2A and 2B, respectively) identified many factors not in the literature. Both stakeholder groups described the experience of discrimination and a patient’s feeling of not having a voice as important contributors to unmet postpartum care needs. Both groups also identified clinical factors, such as differences in the perceived value of referrals, socioeconomic determinants, including poverty and social isolation, and structural or organizational factors, linking high provider workloads and fragmentation between health and social services.

**Step 3: Contextualize Evidence-Base in Stakeholder Perspectives**

**General description:** This step operationalizes what Greene and Hall refer to as a “dialectic stance” by engaging knowledge generated under different paradigms or mental models in respectful dialogue. In line with Greene and Hall, this step seeks to develop a more comprehensive understanding of a phenomena by comparing and combining knowledges, drawing on differences as generative, while documenting each perspective in a way that is transparent and ultimately traceable as to how each contributes to a more comprehensive understanding (Greene & Hall, 2010).

This dialectic stance is operationalized by first comparing and then combining different types of knowledge about a shared phenomenon. Fuzzy cognitive maps are converted to adjacency matrices to facilitate analysis, where factors are listed by row and column and the corresponding assigned weight at their intersection. To compare knowledge types, weights in both literature-based and stakeholder maps are normalized to the same scale of 0 (of having no influence) to 1 (of having the strongest possible association; Andersson et al., 2017). This creates a shared measure to represent the weight (or strength) of factors identified from the literature and stakeholder perspectives. While this puts weights on an equivalent relative scale, this does not change what those values represent, in that the effect estimates from quantitative literature remain measures of likelihood based on observed patterns of outcomes, and stakeholder weights represent a measure of perceived importance of what contributes to an outcome. Both represent measures of
importance: one measured by patterns of observable outcomes, and the other through experiential knowledge by those with intimate understanding of the issue.

Normalizing maps to a probability-based scale (e.g., from 0 to 1) allows for the application of analytical tools that draw on probability science. In graph theory, reachability refers to the ability to get from one node to another within a graph. In directed graphs, this is calculated using transitive closure (Niesink et al., 2013). This algorithm accounts for all connections between pairs of concepts, as well as all implied connections between two concepts as a result of their being part of the same indirect pathway (Giles et al., 2008; Morzaria & Šajna, 2016). Transitive closure accounts for how some factors might have small individual influence but may contribute as part of a sequence of events that have an important overall influence (Andersson et al., 2017; Giles et al., 2008; Niesink et al., 2013). After applying a transitive closure algorithm to all maps, comparing factors, weights, and walks between knowledge sets through a pattern matching table (Table 2) is a simple yet insightful way to identify areas of agreement and differences across knowledge sources.

Once all the maps are on the same scale and relationships accounted for, maps can be put into conversation to formally update one knowledge source with another using Bayesian statistics (Goldstein, 2006). This use of weights from fuzzy cognitive maps as Bayesian priors to

---

Figure 2A. Fuzzy cognitive map of created by family physicians describing factors contributing to unmet postpartum care needs among recent immigrant women in Canada.

Note. Dashed lines indicate a negative relationship; dash-dot-dashed lines indicate an indirect relationship.
update published literature with stakeholder perspectives is an innovation unique to Weight of Evidence.

A Bayesian model begins with a likelihood function over a set of parameters, a conventional measure of plausibility assigned to each parameter. In Weight of Evidence, these are the individual effect estimates and their measures of uncertainty (e.g., CI) identified from the literature. In conventional Bayesian analysis, expert opinion, or other sources of data (e.g., observational studies) contribute to estimating a measure of certainty for each parameter. In Weight of Evidence, these are the stakeholder-assigned weights from the mapping process described in Step 2. Stakeholder-assigned weights for each factor combine to create a central measure and a distribution, representing the variability in stakeholder weights for that factor. These are represented as a normal distribution as it has an easily interpretable measure of central tendency and uncertainty (or variance). This forms the prior distribution, which when multiplied by the likelihood function, updates the parameters identified from the literature (Gelman et al., 2013). The strength of Bayesian analysis lies in its ability to learn from the data in question by combining it with other forms of relevant knowledge, while explicitly accounting for the uncertainty in all types of knowledge (Kruschke, 2015). The resulting posterior distribution represents updating on a conceptual, rather than probabilistic basis and contributes to generating explanatory accounts in the next step.

**Step 3 in practice:** Odds ratios from the literature-derived map (created in Step 1) were normalized using the formula \(1 - \frac{2}{(OR + 1)}\); Andersson et al., 2017). Stakeholder generated maps (created in Step 2) were transformed by dividing the stakeholder assigned weights by the maximum weight possible (in our case 5). A transitive closure algorithm (ProbTC) was applied to each individual stakeholder map and the separate literature-based map. The resulting weight of any identified walk was calculated as the product of the weights of the component arcs or links (Morzaria & Šajna, 2016; Niesink et al., 2013).
Figure 3 shows the walks identified through transitive closure across each of the three knowledge types. Figure 3B shows the relationship between being an immigrant and low social support, living in poverty and having a Caesarean section identified by family physicians. Birth companions highlighted care responsibilities for other family and/or community members as a central element shaping unmet postpartum needs, particularly in the context of a high-needs infant, having clinical signs of depression and not seeking out early prenatal care (Figure 3 C.i). Birth companions identified additional challenges around provider workloads, access to culturally safe care and having access to accessible and appropriate information (Figure 3 C.ii). While some factors were also mentioned in the qualitative literature, they were often done so without the explicit connection and weighting with respect to other factors influencing unmet care needs among recent immigrant women.

Weights averaged across each stakeholder group generate an average stakeholder-specific map, one representing clinicians and another representing perspectives of birth companions. Concepts and weights were compared across knowledge sources using a pattern-matching table (see Table 2). We calculated the level of disagreement between each of the different knowledge groups (literature, family physicians, and birth companions) by calculating an average degree of disagreement between two knowledge sources, according to:

| Concept                                      | Literature | Family physicians | Birth companions |
|----------------------------------------------|------------|-------------------|------------------|
| Similarities                                 |            |                   |                  |
| Being an immigrant                           | 0.23       | 0.79              | 0.99             |
| Poor relationship with provider              | 0.53       | 0.9               |                  |
| Having a caesarean section                   | 0.17       | 0.46              | 0.8              |
| Provider workload                            | 0.27       | 0.62              |                  |
| Lack of respectful care                      | 0.33       | 0.5               |                  |
| Perceived value of care                      | 0.4        | 0.5               |                  |
| Poverty                                      | 0.4        | 0.47              | 0.5              |
| Low social support                           | 0.47       | 0.48              |                  |
| Patient has no voice                         | 0.27       | 0.32              |                  |
| Perceived discrimination                     | 0.6        | 0.22              |                  |
| Fragmentation between health and social services | 0.47   | 0.5               |                  |
| Less than high school                         | 0.18       | 0.27              |                  |
| Differences                                  |            |                   |                  |
| Lack of multidisciplinary teams               | 0.73       |                   |                  |
| Communication misunderstandings              |            |                   | 0.6              |
| Family responsibilities                       |            |                   | 0.4              |
| History of trauma                            |            |                   | 0.4              |
| Experience of delivery                        |            |                   | 0.4              |
| Risk for depression                           |            |                   | 0.32             |
| Not knowing who to trust                      |            |                   | 0.32             |
| Lack of access to mental health services      |            |                   | 0.26             |
| Degree of consensus between family physicians and birth companions | 0.31 | |
| Degree of consensus between family physicians and the literature | 0.37 | |
| Degree of consensus between birth companions and the literature | | 0.45 |

Figure 3 shows the walks identified through transitive closure across each of the three knowledge types. Figure 3B shows the relationship between being an immigrant and low social support, living in poverty and having a Caesarean section identified by family physicians. Birth companions highlighted care responsibilities for other family and/or community members as a central element shaping unmet postpartum needs, particularly in the context of a high-needs infant, having clinical signs of depression and not seeking out early prenatal care (Figure 3 C.i). Birth companions identified additional challenges around provider workloads, access to culturally safe care and having access to accessible and appropriate information (Figure 3 C.ii). While some factors were also mentioned in the qualitative literature, they were often done so without the explicit connection and weighting with respect to other factors influencing unmet care needs among recent immigrant women.

Weights averaged across each stakeholder group generate an average stakeholder-specific map, one representing clinicians and another representing perspectives of birth companions. Concepts and weights were compared across knowledge sources using a pattern-matching table (see Table 2). We calculated the level of disagreement between each of the different knowledge groups (literature, family physicians, and birth companions) by calculating an average degree of disagreement between two knowledge sources, according to:
Figure 3 shows the walks identified through transitive closure across each of the three knowledge types. Figure 3B shows the relationship between being an immigrant and low social support, living in poverty and having a Caesarean section identified by family physicians. Birth companions highlighted care responsibilities for other family and/or community members as a central element shaping unmet postpartum needs, particularly in the context of a high-needs infant, having clinical signs of depression and not seeking out early prenatal care (Figure 3 C.i). Birth companions identified additional challenges around provider workloads, access to culturally safe care and having access to accessible and appropriate information (Figure 3 C.ii).

While some factors were also mentioned in the qualitative literature, they were often done so without the explicit connection and weighting with respect to other factors influencing unmet care needs among recent immigrant women.

Weights averaged across each stakeholder group generate an average stakeholder-specific map, one representing clinicians and another representing perspectives of birth companions. Concepts and weights were compared across knowledge sources using a pattern-matching table (see Table 2). We calculated the level of disagreement between each of the different knowledge groups (literature, family physicians, and birth companions) by calculating an average degree of disagreement between two knowledge sources, according to:

$$\bar{d} = \frac{\sum |d|}{N},$$  \hspace{1cm} (Eq1)

where a higher value for $\bar{d}$ represents greater differences between knowledge groups (Andersson & Silver, 2019). Consensus was greater between family physicians and birth companions ($\bar{d} = 0.31$) and lower between published literature and the birth companions ($\bar{d} = 0.45$).
Published literature was updated with prior weighting densities from family physician perspectives, and separately from birth companion perspectives for each nonzero cell in the adjacency matrices (Bernardo & Smith, 2000; Joseph, 2000b). To simulate a full-scale implementation of this procedure, this analysis was also conducted with duplicated stakeholder maps to simulate a more appropriate sample size across each stakeholder group. A custom computer program using open access software was developed to carry out three approaches to Bayesian updating (see online Supplementary Additional File 1 & 2). R statistical software was used to plot the weighting distributions for selected relationships identified in the data, from stakeholder perspectives and from data updated by each stakeholder perspective. All maps were created using the open-access software yEd (www.yworks.com/yed).

Each combining procedure generated new weights and therefore a new map architecture. Each updated weight was represented by a central value (used to generate the updated maps in Figure 4) and a normal density around that central value (used to specify the density for each of the factors in Tables 3 & 4). The variance around the central weight decreased around factors where there was agreement between published literature and stakeholder perspectives and broadened (or increased) where there were areas of disagreement. In other words, agreement led to increased confidence and lower variability, and vice versa. Figure 4 shows the literature map updated (using duplicated data; Equation 4 in Additional File 1) by family physicians (Figure 4a) and by birth companions (Figure 4b).
Published literature was updated with prior weighting densities from family physician perspectives, and separately from birth companion perspectives for each nonzero cell in the adjacency matrices (Bernardo & Smith, 2000; Joseph, 2000b). To simulate a full-scale implementation of this procedure, this analysis was also conducted with duplicated stakeholder maps to simulate a more appropriate sample size across each stakeholder group. A custom computer program using open access software was developed to carry out three approaches to Bayesian updating (see online Supplementary Additional File 1 & 2). R statistical software was used to plot the weighting distributions for selected relationships identified in the data, from stakeholder perspectives and from data updated by each stakeholder perspective. All maps were created using the open-access software yEd (www.yworks.com/yed).

Each combining procedure generated new weights and therefore a new map architecture. Each updated weight was represented by a central value (used to generate the updated maps in Figure 4) and a normal density around that central value (used to specify the density for each of the factors in Tables 3 & 4). The variance around the central weight decreased around factors where there was agreement between published literature and stakeholder perspectives and broadened (or increased) where there were areas of disagreement. In other words, agreement led to increased confidence and lower variability, and vice versa. Figure 4 shows the literature map updated (using duplicated data; Equation 4 in Additional File 1) by family physicians (Figure 4a) and by birth companions (Figure 4b).

Figure 4a. Fuzzy cognitive maps of (A) the literature updated by the fuzzy cognitive maps created by family physicians.

Note. Dashed lines indicate a negative relationship; dash-dot-dashed lines indicate an indirect relationship identified by transitive closure.

Figure 4b. Fuzzy cognitive maps of (B) the literature updated by the fuzzy cognitive maps created by birth companions.

Note. Dashed lines indicate a negative relationship; dash-dot-dashed lines indicate an indirect relationship identified by transitive closure.

Table 3 provides an example of factor-level analysis across knowledge sources, comparing weights assigned to having a Caesarean section among recent immigrant women according to the literature, family physicians and birth companions. The literature-based weight was drawn from one study with an OR 1.42 (95% CI [1.03, 1.96]), subsequently normalized to a weight of 0.17 with a density of 0.01 to 0.32 (Gagnon et al., 2007). Having a Caesarean section was weighted by family physicians (M = 0.46, $\sigma^2 = 0.1$) and by birth companions (M = 0.8, $\sigma^2 = 0.06$). Updating using Bayesian analysis (Equation 4, see online Supplementary Additional File 1) resulted in an updated point estimate and credibility interval, accounting for variability in both the literature-based estimate and across stakeholder groups, shown in the figure embedded in Table 3. A comparison of updated results between original data and duplicated data is provided in online Additional File 2.

Table 4 provides an example of when stakeholders include factors not reported in the quantitative literature. Both family physicians and birth companions identified discrimination as an important contributor to unmet postpartum care needs among recent immigrant women. Discrimination was identified in qualitative, but not quantitative, literature (Gagnon, Carnevale, et al., 2013; Merry et al., 2011; O’Mahony & Donnelly, 2010). Table 4 compares stakeholder weights and distributions, offering some insight into how each stakeholder group interpreted the importance of discrimination on unmet postpartum care needs among recent immigrant women.

Step 4: Describe Explanatory Processes Leading to Priority Outcomes

General description: This step draws on published literature, stakeholder explanations, and the analyses carried out in the previous step to develop candidate explanatory accounts. Relationships described in stakeholder cognitive maps show how stakeholders make sense of
Table 3. Original Weights and Updated Values Form the Literature, Family Physicians and Birth Companions for the Influence of Having a Cesarean Section on Unmet Postpartum Care Needs Among Recent Immigrant Women.

| Source of estimate or weight | Odds ratio or weight | Lower 95% CI (or CrI) | Upper 95% CI (or CrI) |
|-----------------------------|---------------------|------------------------|------------------------|
| Literature                  | 1.42                | 1.03                   | 1.96                   |
| Normalized value            | 0.17                | 0.01                   | 0.32                   |
| **Updated by family physician knowledge** |                     |                        |                        |
| Average weight assigned by Family Physicians: 0.46 |                   |                        |                        |
| Bayesian                    | 0.39                | 0.17                   | 0.61                   |

**Updated by birth companion knowledge**

| Average weight assigned by Birth Companions: 0.8 |                   |                        |                        |
| Bayesian updating | 0.65                | 0.41                   | 0.89                   |

i) Values from the literature (in black), weighted by family physician (in blue), and when updated (in red).

ii) Values from the literature (in black), weighted by birth companion (in green), and when updated (in red).
their experience in the context of evidence from the literature (Mingers, 2005; U. Özesmi & Özesmi, 2004) Weighting by stakeholders helps prioritize factors or processes that contribute to an outcome, whether from published evidence or those added to the maps by stakeholders. If stakeholders identify novel factors contributing to the outcome, a brief literature search is done to identify relevant findings, which are then included in the evidence base to develop explanatory accounts.

Consistent with participatory research, explanatory accounts and/or consolidated themes can be shared back with stakeholders to be adjusted to align with their own understandings. This strengthens the trustworthiness of the explanatory accounts and can be done with multiple different stakeholder groups. This is especially important when working with marginalized communities, where theories and explanations generated outside the community may reinforce erroneous stereotypes (Tuck, 2008).

**Step 4 in practice:** Following the process outlined by Pearson and colleagues, the lead author generated explanatory accounts from published literature and stakeholder accounts to describe how specific factors and relationships between them may contribute to unmet postpartum care needs among recent immigrant women (Pearson, 2015). Explanatory accounts were informed by identified literature, the relationships identified across different stakeholder groups (for practical reasons, all relationships weighted 0.2 or above in any map were included), direct or indirect relationships between two or more factors (identified as walks by transitive closure), as well as differences in weights between knowledge sources about how specific factors contribute to unmet postpartum care needs. These accounts were also informed by notes from mapping interviews (Step 2) where stakeholders verbally rationalized their selection and weighting of relationships while making the maps (Pawson, 2008; Pearson, 2015). Ninety explanatory accounts were consolidated down to 20, guided by questions to specify the conceptual clarity of each explanatory account (Pearson, 2015). Accounts were grouped by common mechanism to explain factors and relationships that contribute to the outcome, and then grouped across four related themes as shown in the online Additional File 4. Original, consolidated explanatory

---

**Table 4.** Original Weights and Updated Values From the Literature, Family Physicians and Birth Companions for the Influence of Perceived Discrimination on Unmet Postpartum Care Needs Among Recent Immigrant Women.

| Source of estimate or weight | Odds ratio or weight | Lower 95% CI (or CrI) | Upper 95% CI (or CrI) |
|-----------------------------|---------------------|-----------------------|-----------------------|
| Literature                  | Not measured        |                       |                       |
| Normalized value            | 0                   |                       |                       |
| **Updated by family physician knowledge** |                     |                       |                       |
| Average weight assigned by Family Physicians: 0.6 |                     |                       |                       |
| Bayesian updating            | 0.49                | 0.2                   | 0.79                  |
| **Updated by birth companion knowledge** |                     |                       |                       |
| Average weight assigned by Birth Companions: 0.22 |                     |                       |                       |
| Bayesian updating            | 0.18                | 0.0                   | 0.37                  |

Weight assigned by family physician (in blue) and birth companions (in green).
accounts and overarching themes were reviewed by a senior author with experience in realist analysis (ER). Consolidated themes and explanatory accounts are shown in Table 5 and the online Additional Files 3 and 4. Each explanatory account is attributed to a particular stakeholder group, the literature (through reference to the specific study), or a combination of these. A narrative of findings is included in the online Additional File 5.

**Step 5: Identify Local Interventions and Implications for Evidence Syntheses**

This step asks stakeholders to make recommendations to address the outcome based on the understanding facilitated by Weight of Evidence. Given the breadth of factors included in the maps, it can be helpful to draw on key outputs from Weight of Evidence, such as maps generated by stakeholders and the evidence updated by stakeholder perspectives (from Step 2 and Step 3, respectively) and the consolidated explanatory account from Step 4. Each stakeholder group can recommend strategies that they can carry out themselves, as well as those that would be most effective in addressing the issue in both the short and long term (Abelson et al., 2003). While this step was not fully implemented as part of this pilot of Weight of Evidence, other applications of this procedure have contributed to evidence-based stakeholder-led recommendations for perinatal care for adolescents (Dion & Andersson, in press).

**Discussion**

Weight of Evidence facilitates the inclusion of complex and often informal stakeholder knowledge alongside evidence from qualitative and quantitative literature in a transparent and reproducible way. Thus, informing a richer understanding of the issue at hand and giving users an opportunity to reconceptualize a problem and its potential solutions. Weight of Evidence provides an accessible way to represent published literature to a diversity of stakeholders and draws directly on stakeholder knowledge and causal understanding to determine how to prioritize and combine different types of evidence.

Our research group developed Weight of Evidence as a rigorous procedure to incorporate stakeholders’ perspective in the design and evaluation of health services, with a focus on incorporating perspectives from marginalized populations, often underrepresented in health services and biomedical literature (Serrant-Green, 2011). The systematic and analytical approach to integrating different understandings around a single issue are what make Weight of Evidence particularly relevant for evidence-informed decision-making, whether at national or regional levels, for policy development or institutional networks.

Recognizing that someone with different lived experience may understand and solve a problem differently can be useful information. This is not to suggest that one type of knowledge has a hierarchy over others (Harding, 2003; Oliver et al., 2018). Consistent with the aims of mixed methods research, Weight of Evidence offers a way to bring these different forms of knowledge together, in support of a more complete understanding of an issue (Creswell & Plano Clark, 2011; Mertens, 2003). Few knowledge synthesis approaches preserve divergent or conflicting perspectives, often homogenizing across studies thus losing the richness within difference. By holding space for multiple ways of understanding the same issue, Weight of Evidence creates opportunities to unpack differences, identifying how and when these differences arise and with what consequences (Keller, 1992; Maxwell, 2012).

In this demonstration case, stakeholder-identified factors were more diverse and more actionable than factors identified in the literature. They focused less on individual “risk factors” of recent immigrant women (e.g., high school education or specific health behaviours) and more on the contextual supports and systems around them. Stakeholder-identified factors also
Table 5. Consolidated Explanatory Accounts Describing Factors Contributing to Unmet Postpartum Care Needs Among Recent Immigrant Omen in Canada (MD = Physician, BC = Birth Companion).

| Processes contributing to outcomes | Source explanatory accounts | Primary evidence |
|-----------------------------------|----------------------------|-----------------|
| **Theme: Economic and social exclusion** | | |
| Poverty’s influence on health navigation and access | 1, 2, 3, 4, 50, 53, 64, 69 | (Gagnon, Carnevale, et al., 2013; Munro et al., 2014); Stakeholders (MD) |
| Social and physical isolation, lack of social support | 5, 6, 22, 23, 24, 37, 47, 51, 62, 70, 80, 81, 82 | (Gagnon et al., 2010; O’Mahony & Donnelly, 2010); Stakeholders (MD&BC) |
| **Theme: Building shared understanding** | | |
| Social understanding of illness | 7, 11, 17, 18, 20, 21, 44, 45, 67, 68, 72, 84, 88 | (Gagnon, Carnevale, et al., 2013; Higginbottom, 2013); Stakeholders (MD&BC) |
| Culture as an asset for health | 43, 57, 89 | (Gagnon, Carnevale, et al., 2013; Higginbottom et al., 2014); Stakeholders (BC) |
| **Theme: Respectful communication** | | |
| Commitment to culturally safe care | 9, 10, 26, 38, 39, 48, 65, 71, 78, 83, 85 | (Gagnon, Carnevale, et al., 2013; Gagnon et al., 2010); Stakeholders (MD&BC) |
| Commitment to trauma-informed care | 27, 28, 29, 42, 49, 55, 56, 87 | (O’Mahony & Donnelly, 2010); Stakeholder (MD&BC) |
| **Theme: Supporting self-determination** | | |
| Access to information | 19, 25, 32, 36, 52, 58, 63, 79 | (Gagnon, Carnevale, et al., 2013; Gagnon et al., 2010; O’Mahony & Donnelly, 2010); Stakeholders (MD&BC) |
| Building trust | 8, 33, 86, 90 | (Gagnon, Carnevale, et al., 2013; Gagnon et al., 2007; O’Mahony & Donnelly, 2010); Stakeholders (BC) |
| **Health services organization** | | |
| Fragmentation | 40, 41, 54, 61, 62, 73 | Stakeholders (MD) |
| Centralized decision-making | 59, 74, 76 | Stakeholders (MD) |
| Individualization of care responsibilities | 34, 46, 60, 75, 12, 13, 66 | (Gagnon, Carnevale, et al., 2013; Gagnon et al., 2010; O’Mahony & Donnelly, 2010); Stakeholders (MD) |
highlighted larger structural and organizational issues, such as resource allocation, workforce planning, and social isolation, in relation to unmet postpartum care needs, demonstrating how the inclusion of relevant perspectives in problem definition can broaden the menu of possible interventions to address unmet postpartum care needs.

Contextualizing evidence is often framed as part of knowledge translation, contributing to evidence-based practice and decision making, whereby once primary research is synthesized, it is assessed and interpreted for local relevance. Effective translation requires that evidence be understood and applied in relation to local context and practical wisdom, incorporating insights from clinical experience and patient experience to identify and define meaningful interventions (MacDermid & Graham, 2009). Several authors have highlighted the challenges of contextualization and stakeholder involvement in evidence synthesis more broadly, including poor operationalization and a lack of understanding of how stakeholder input influences evidence interpretation and translation (Haddaway et al., 2017; Pollock et al., 2018). The majority of stakeholder engagement in evidence syntheses has focused on those who have some understanding of research evidence and already have a seat at decision-making tables. Two recent systematic reviews of stakeholder and community engagement highlight that stakeholder involvement in evidence synthesis has focused primarily on the engagement of health professionals, academics, and decision makers with only 30% of studies including patients and communities as stakeholders (George et al., 2017; Pollock et al., 2018).

Weight of Evidence introduces a systematic method to assess the legitimacy of reviews by those most affected by the issue at hand without expecting stakeholders to become highly versed in methods of evidence synthesis (Langer et al., 2017). This has important implications for accessibility and shapes who has the opportunity to participate in contextualization processes. By maintaining a high-level of transparency and clarity in how and where stakeholder perspectives prioritize and adapt available evidence base, Weight of Evidence presents a reliable and verifiable procedure to contextualize evidence syntheses (Haddaway et al., 2017). Making explicit when and how we value different forms of knowledge opens up analysis and decision-making processes to greater scrutiny and discussion, facilitating a more collaborative conceptualization of priority issues (Befani & Stedman-Bryce, 2017).

This application of Weight of Evidence also highlights potential future applications to refine and legitimate systematic reviews based on the contextualizing of evidence by those most directly affected by the issue. Identifying factors not included in the literature as well as adapting or introducing relationships between factors not previously accounted for, Weight of Evidence offers a procedure to identify what evidence matters, to whom and with what consequences. These are all critical questions that shape how evidence syntheses contribute to the translation of evidence into practice and decision making (Barkham & Mellor-Clark, 2003; Booth et al., 2019; MacDermid & Graham, 2009). While our pilot application of Weight of evidence did not include stakeholders in setting the question for the evidence synthesis, subsequent application of this method included an evidence-based participatory approach to the initial problem definition (Dion et al., 2021).

Contributions to mixed methods research: Weight of Evidence presents a concrete example of how advanced statistical tools applied within a transformative participatory research perspective can contribute to innovations in mixed methods research (Andersson, 2018). Weight of Evidence offers a way to move beyond a monolithic view of evidence and expertise and advance evidence syntheses to incorporate contextual understanding of which evidence is most relevant for a particular setting (Mertens, 2003). This may be particularly relevant when the evidence base is sparse or contradictory, making it difficult to make conclusion or recommendations (Badampudi et al., 2019). Although poor and socially marginalized groups are often very clear about how marginalization affects their health, the issue is often poorly reflected in
the available evidence bases, making Weight of Evidence particularly suited to contextualizing evidence with underserved populations (Serrant-Green, 2011).

This article demonstrates that people’s understanding of root causes and priority factors can be described in detailed and systematic ways, facilitating a transparent and rigorous combining with evidence. Combining cognitive maps across different perspectives operationalizes the dialectic stance described by Greene and Hall (2010), providing insights into how perspectives complement and differ from one another. This can act as an important lever to identify potentially shared priorities as well as areas of disagreement or difference.

Fuzzy cognitive maps provide an accessible way to represent different forms of knowledge to be understood and adapted across perspectives and paradigms (Kosko, 1988; Popay & Williams, 1996). Making epidemiological data, or knowledge from other stakeholders, accessible through fuzzy cognitive mapping invites stakeholders to engage with the full scope of evidence often available to other decision makers. Maps demonstrate how different knowledge is taken into account in the identification of priorities in consideration of, and not in isolation, of all available evidence. Practically, fuzzy cognitive maps provide an architecture for data integration during the mapping interview (Step 2) as stakeholders indicate how and with what weight qualitative data should be integrated with quantitative evidence. This process also catalyzes new data from stakeholders that is incorporated into the cognitive maps. This expands on exiting protocols and applications of fuzzy cognitive maps by transforming qualitative data from themes (either identified from qualitative literature or identified by participants) into a relational structure together with quantitative data, led by participant expertise rather than researcher expertise (Fetters et al., 2013; Giles et al., 2007; U. Özesmi & Özesmi, 2004). As conceptual rather than probabilistic models, cognitive maps then reflect interdependence between factors that can serve as a decision aid in complex settings like clinical care, where artificially isolating associations within a de facto network can diminish the contextual understanding and relevance of decisions (Napier et al., 2014).

Weights generated through fuzzy cognitive mapping invite stakeholders to analyze problems in context, while generating transparent and meaningful measures of influence for each factor that can be easily analyzed by stakeholder group, or “in conversation” with other types of knowledge. This expands conceptual boundaries by incorporating issues formerly rejected or seen as outside of the system of influence (Midgley, 2000; van Bertalanffy, 1968).

Examining weights within and across stakeholder groups offers an accessible and systematic tool to address the long-standing concern in Bayesian statistics of generating meaningful and representative priors. Graphic representations and pattern matching tables encourage a meta-level comparison between knowledge sources. A factor-level analysis is possible by comparing the distribution of weights for specific relationships, visually representing a density of credible weights assigned by different stakeholder groups instead of plotting a single average value (Kruschke, 2015).

Weight of Evidence integrates stakeholder-identified, qualitative and quantitative data through fuzzy cognitive mapping and Bayesian updating. This integration is then leveraged in the generation of explanatory accounts (Step 4) as qualitative, quantitative, and stakeholder-generated data contribute to identify candidate causal processes (Pawson, 2008; Pearson, 2015). Stakeholder involvement led to the identification of factors beyond those identified by the original literature review, prompting a broader examination of literature to confirm, expand and refine our explanatory accounts.

**Lessons and limitations:** Duplicating our data skewed results toward stakeholder values and artificially reduced variance. Yet, comparing results between the original and duplicated data provided several insights. The first is that it matters how we present stakeholder perspectives. In this demonstration case, individual stakeholders each created their own map. In other
applications of this procedure within our research group, a group of stakeholders collectively created one map (Sarmiento et al., 2020). These differences have important implications for the development of the maps and ensuring that the final stakeholder generated maps are representative of differences in power and lived experience across and within stakeholder groups (S. A. Gray et al., 2013). It also has implications for how updating occurs, as each additional stakeholder map contributes data to the updating procedure.

The second insight is that the densities of stakeholder weights are helpful representations of the extent of agreement both within and between stakeholder groups. Differences in opinion within stakeholder groups, as evidenced by broad or multimodal densities, suggest the need for larger sample sizes, greater differentiation between stakeholder groups, or as a jumping-off point for further exploration of observed differences.

Various forms of bias may also influence how weights generated through the fuzzy cognitive mapping process represent the beliefs of stakeholders. We drew on the recommendations in the literature to develop robust priors including carefully structured elicitation procedures, drawing on multiple experts to generate an aggregated prior estimate and the opportunity for stakeholders to provide feedback and corrections to their maps (Burgman et al., 2011). Several forms of uncertainty remain inherent to the process of generating stakeholder weights, including linguistic and epistemic uncertainty. Adaptations to the elicitation process can address these concerns, however, we did not have the opportunity to conduct sensitivity analysis around adaptations.

That this demonstration case focused on methods development rather than implementing solutions introduced some limitations to this study and has several implications. Since systematic review methods are well-established and therefore not part of the methodological innovation of this procedure, the lead author independently conducted the mixed methods scoping review. To use the procedure for real-time service improvement, a conventional review following established protocols would be more appropriate to provide comprehensive and robust evidence in the literature-based map.

A second important limitation of this study is that it did not involve a sufficient number or diversity of stakeholders to draw definitive conclusion about the implications of this work for community perinatal care. For researchers planning to apply Weight of Evidence in their work, it is recommended they follow existing sample size recommendations for fuzzy cognitive map (U. Özesmi & Özesmi, 2004).

**Conclusions**

Weight of Evidence combines context-specific stakeholder knowledge with quantitative and qualitative data from published studies, reconciling several perspectives and translating these into statistically interpretable and actionable results. It advances the concept and practice of stakeholder engagement and sets out a procedure to contextualize evidence synthesis toward more comprehensive and relevant findings. This article describes how the application of participatory methods and advanced statistics contributes to a more comprehensive, context-specific understanding of an issue while also presenting novel approaches to data transformation and integration.

Weight of Evidence was developed to strengthen the voice of marginalized groups to inform interventions in health services. It also has broader applicability in supporting the contextualization of evidence syntheses and introducing more participatory and rigorous approaches to the design of quality improvement and evaluation initiatives. Weight of Evidence offers a transparent procedure to unpack differences, to identify how and when these differences arise and with what consequences, contributing to more comprehensive, relevant, and effective interventions and recommendations.
Acknowledgments
This work was carried out on the traditional territories of the Kanien'kehá:ka and Algonquin Nations. We are grateful to the physicians and birth companions that contributed to this work. We are grateful for the contributions of Ivan Sarmiento for his contributions to advancing the use of Weight of Evidence. We also thank Patrick Belisle for his contribution to the development of our custom software.

Acknowledgement and Role of Funding Sources
AD holds a Pierre Elliott Trudeau Foundation Doctoral Scholarship; NA received a grant from the Québec Strategy for Patient Oriented Research (SPOR) Methodological Development Group; ER received a training grant from the Canadian Institutes of Health Research. Funding agencies had no role in the pilot study or the writing of the manuscript.

Authors Contributions
AD, ER, LJ and NA developed the method and conceived of the pilot study, AD drafted the article, provided substantive revisions; AD led and ABA, ACG and VJ contributed to the pilot study. All authors reviewed and approved final version of manuscript.

Availability of Data and Material
All data generated or analyzed during this study are included in this published article and its supplementary information files.

Ethics Committee Approval
The pilot work described in this study was carried out as a quality improvement initiative with contributing family physicians ACG and VJ and birth companion ABA included as authors. The full implementation of this method has been approved by the McGill Faculty of Medicine Institutional Review Board (A09-B51-17A).

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD
Anna Dion https://orcid.org/0000-0003-2122-222X

Supplemental Material
Supplemental material for this article is available online.

References
Abelson, J., Forest, P.-G., Eyles, J., Smith, P., Martin, E., & Gauvin, F.-P. (2003). Deliberations about deliberative methods: Issues in the design and evaluation of public participation processes. Social Science & Medicine, 57(2), 239-251. https://doi.org/10.1016/S0277-9536(02)00343-X
Andersson, N. (2018). Participatory research: A modernizing science for primary health care. *Journal of General and Family Medicine, 19*(5), 154-159. https://doi.org/10.1002/jgf2.187

Andersson, N., Beauchamp, M., Nava-Aguilera, E., Paredes-Solís, S., & Sajna, M. (2017). The women made it work: Fuzzy transitive closure of the results chain in a dengue prevention trial in Mexico. *BMC Public Health, 17*(1), Article 408. http://doi.org/10.1186/s12889-017-4301-0

Andersson, N., & Silver, H. (2019). Fuzzy cognitive mapping: An old tool with new uses in nursing research. *Journal of Advanced Nursing, 75*(12), 3823-3830. https://doi.org/10.1111/jan.14192

Badampudi, D., & Wohlin, C. (2016, August 31-September 2). Bayesian synthesis for knowledge translation in software engineering: Method and illustration [Conference]. 42th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), Limassol, Cyprus. https://doi.org/10.1109/SEAA.2016.45

Badampudi, D., Wohlin, C., & Gorschek, T. (2019). Contextualizing research evidence through knowledge translation in software engineering. In *Proceedings of the evaluation and assessment of software engineering* (pp. 306-311). Association for Computing Machinery. https://doi.org/10.1145/3319008.3319358

Barkham, M., & Mellor-Clark, J. (2003). Bridging evidence-based practice and practice-based evidence: Developing a rigorous and relevant knowledge for the psychological therapies. *Clinical Psychology & Psychotherapy, 10*(6), 319-327. https://doi.org/10.1002/cpp.379

Batalden, P. B., & Davidoff, F. (2007). What is “quality improvement” and how can it transform healthcare? *BMJ Quality & Safety, 16*(1), 2-3. https://doi.org/10.1136/qshc.2006.022046

Befani, B., & Stedman-Bryce, G. (2017). Process tracing and Bayesian updating for impact evaluation. *Evaluation, 23*(1), 42-60. https://doi.org/10.1177/1356389016654584

Bernardo, J., & Smith, A. (2000). *Bayesian theory*. Wiley. https://statisticalsupportandresearch.files.wordpress.com/2019/03/jos3ca9-m.-bernardo-adrian-f.-m.-smith-bayesian-theory-wiley-1994.pdf

Bhaskar, R. (2008). *A realist theory of science*. Routledge.

Booth, A., Moore, G., Fleming, K., Garside, R., Tuncalp, Ö., Rollins, N., & Noyes, J. (2019). Taking account of context in systematic reviews and guidelines considering a complexity perspective. *BMJ Global Health, 4*(1), 1-15. https://doi.org/10.1136/bmjgh-2018-000840

Borda, O. F. (1996). Power/knowledge and emancipation. *Systems Practice, 9*(2), 1-5. https://doi.org/10.1007/BF02172931

Bornstein, M., & Hedges, L. (2019). *Handbook of research synthesis and meta-analysis*. Sage.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. https://doi.org/10.1191/147808706qp06oa

Burgman, M., Carr, A., Godden, L., Gregory, R., McBride, M., Flander, L., & Maguire, L. (2011). Redefining expertise and improving ecological judgment. *Conservation Letters, 4*(2), 1-7. https://doi.org/10.1111/j.1755-263X.2011.00165.x

Collins, P. H. (1986). Learning from the outsider within: The sociological significance of black feminist thought. *Social Problems, 33*(6), S14-S32. https://doi.org/10.2307/800672

Crandell, J., Voils, C., & Sandelowski, M. (2012). Bayesian approaches to the synthesis of qualitative and quantitative research findings. In K. Hannes & C. Lockwood (Eds.), *Sylthesizing qualitative research choosing the right approach* (1st ed., chap. 7, pp. 137-159). https://doi.org/10.1002/9781119959847.ch7

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research*. Sage.

Di Ruggiero, E. (2018). The relaunch of IJPH’s knowledge synthesis, translation, and exchange section. *International Journal of Public Health, 63*(3), 309-310. https://doi.org/10.1007/s00038-018-1089-3

Dion, A., & Andersson, N. (under review). How adolescent mothers interpret and prioritize evidence about perinatal child protection involvement: Participatory contextualization of published evidence. *Child and Youth Studies Review.*

Dion, A., Klevor, A., Nakajima, A., & Andersson, N. (2021). Evidence-based priorities of under-served pregnant and parenting adolescents: Addressing inequities through a participatory approach to contextualizing evidence syntheses. *International Journal of Equity in Health, 20*, Article 118. https://doi.org/10.1186/s12939-021-01458-7
Dixon-Woods, M., Cavers, D., Agarwal, S., Annandale, E., Arthur, A., Harvey, J., Hsu, R., Katbamna, S., Olsen, R., Smith, L., Riley, R., & Sutton, A. J. (2006). Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. BMC Medical Research Methodology, 6(1), Article 35. https://doi.org/10.1186/1471-2288-6-35

Fetters, M., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods design-principles and practices. Health Services Research, 48(6, Pt. 2), 2134-2156. https://doi.org/10.1111/1475-6773.12117

Fetters, M., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods design-principles and practices. Health Services Research, 48(6, Pt. 2), 2134-2156. https://doi.org/10.1111/1475-6773.12117

Flaspohler, P., Wandersman, A., Keener, D., Maxwell, K. N., Ace, A., Andrews, A. B., & Holmes, B. (2003). Promoting program success and fulfilling accountability requirements in a statewide community-based initiative: Challenges, progress, and lessons learned. Journal of Prevention and Intervention in the Community, 6(2), 37-52. https://doi.org/10.1300/J005v02n02_04

Foster-Fishman, P. G., Nowell, B., & Yang, H. (2007). Putting the system back into systems change: A framework for understanding and changing organizational and community systems. American Journal of Community Psychology, 39(3-4), 197-215. https://doi.org/10.1007/s10464-007-9109-0

Gagnon, A. J., Carnevale, F., Mehta, P., Rousseau, H. L. N., & Stewart, D. E. (2013). Developing population interventions with migrant women for maternal-child health: A focused ethnography. BMC Public Health, 13(1), Article 471. https://doi.org/10.1186/1471-2458-13-471

Gagnon, A. J., Carnevale, F., Saucier, J.-F., Clausen, C., Jeannotte, J., & Oxman, J. (2010). Do referrals work? Responses of childbearing newcomers to referrals for care. Journal of Immigrant and Minority Health, 12(4), 559-568. https://doi.org/10.1007/s10903-009-9242-y

Gagnon, A. J., Dougherty, G., Platt, R., Wahoush, O., George, A., Stanger, E., Oxman-Martinez, J., Saucier, J.-F., Merry, L., & Stewart, D. E. (2007). Refugee and refugee-claimant women and infants post-birth. Canadian Journal of Public Health, 98(4), 287-291. https://doi.org/10.1016/j.bjph.2007.07.002

Gagnon, A. J., Dougherty, G., Wahoush, O., Saucier, J.-F., Dennis, C.-L., Stanger, E., Palmer, B., Merry, L., & Stewart, D. E. (2013). International migration to Canada: The post-birth health of mothers and infants by immigration class. Social Science & Medicine, 76(January), 197-207. https://doi.org/10.1016/j.socscimed.2012.11.001

Gagnon, A. J., Van Hulst, A., Merry, L., George, A., Saucier, J.-F., Stanger, E., Wahoush, O., & Stewart, D. E. (2012). Cesarean section rate differences by migration indicators. Archives of Gynecology and Obstetrics, 287(4), 633-639. https://doi.org/10.1007/s00404-012-2609-7

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., & Rubin, D. (2013). Bayesian data analysis (3rd ed.). Chapman & Hall/CRC. https://doi.org/10.1201/b16018

George, A., Mehra, V., Scott, K., & Srimam, V. (2017). Community participation in health systems research: A systematic review assessing the state of research, the nature of interventions involved and the features of engagement with communities. PLOS ONE, 10(10), e0141091. https://doi.org/10.1371/journal.pone.0141091

Giles, B. G., Findlay, C. S., Haas, G., LaFrance, B., Laughing, W., & Pemberton, S. (2007). Integrating conventional science and aboriginal perspectives on diabetes using fuzzy cognitive maps. Social Science & Medicine, 64(3), 562-576. https://doi.org/10.1016/j.socscimed.2006.09.007

Giles, B. G., Haas, H., Sajna, M., & Findlay, C. S. (2008). Exploring aboriginal views of health using fuzzy cognitive maps and transitive closure. Canadian Journal of Public Health, 99(5), 411-417. https://doi.org/10.1010/j.1542-4276.2008.00655.x

Goldstein, M. (2006). Subjective Bayesian analysis: Principles and practice. Bayesian Analysis, 1(3), 403-420. https://doi.org/10.1214/06-BA116

Gray, S. A., Zaunre, E., & Gray, S. R. J. (2013). Fuzzy cognitive maps as representations of mental models and group beliefs. In Fuzzy cognitive maps for applied sciences and engineering (Vol. 54, pp. 29-48). Springer. https://doi.org/10.1007/978-3-642-39739-4_2

Greene, J., & Hall, J. (2010). Dialectics and pragmatism: Being of consequence. In A. Tashakkori & C. Teddlie (Eds.), Handbook of mixed methods in social & behavioral research (2nd ed.). Sage. https://doi.org/10.4135/9781506335193.n5

Greenhalgh, T., Jackson, C., Shaw, S., & Janamian, T. (2016). Achieving research impact through co-creation in community-based health services: Literature review and case study. Milbank Quarterly, 94(2), 392-429. https://doi.org/10.1111/1468-0009.12197
Haddaway, N., Kohl, C., da Silva, N. R., Schiemann, J., Spök, A., Stewart, R., Sweet, J. B., & Wilhelm, R. (2017). A framework for stakeholder engagement during systematic reviews and maps in environmental management. *Environmental Evidence, 6*, Article 11. https://doi.org/10.1186/s13750-017-0089-8

Harding, S. (2003). A socially relevant philosophy of science? Resources from standpoint theory’s controversiality. *Hypatia, 19*(1), 25-47. https://doi.org/10.1111/j.1527-2001.2004.tb01267.x

Harris, J., Croot, L., Thompson, J., & Springett, J. (2016). How stakeholder participation can contribute to systematic reviews of complex interventions. *Journal of Epidemiology & Community Health, 70*(2), 207-214. https://doi.org/10.1136/jech-2015-205701

Higginbottom, G. M. A. (2013). “I have to do what I believe”: Sudanese women’s beliefs and resistance to hegemonic practices at home and during experiences of maternity care in Canada. *BMC Pregnancy and Childbirth, 13*(1), Article 51. https://doi.org/10.1186/1471-2393-13-51

Higginbottom, G. M. A., Hadziabdic, E., Yohani, S., & Paton, P. (2014). Immigrant women’s experience of maternity services in Canada: A meta-ethnography. *Midwifery, 30*(5), 544-559. https://doi.org/10.1016/j.midw.2013.06.004

Joanna Briggs Institute. (2017a). Critical appraisal checklist for cohort studies. Author. https://joannabriggs.org/sites/default/files/2019-05/JBI_Critical_Appraisal-Checklist_for_Cohort_Studies2017_0.pdf

Joanna Briggs Institute. (2017b). Critical appraisal checklist of qualitative research. Author. https://joannabriggs.org/sites/default/files/2019-05/JBI_Critical_Appraisal-Checklist_for_Qualitative_Research2017_0.pdf

Joanna Briggs Institute. (2017c). Critical appraisal for case control studies. Author. https://joannabriggs.org/sites/default/files/2019-05/JBI_Critical_Appraisal-Checklist_for_Case_Control_Studies2017_0.pdf

Joseph, L. (2000a). A brief introduction to meta-analysis. In L. Rosenberg, L. Joseph, & A. Barkun (Eds.), *Surgical arithmetic: Epidemiological, statistical and outcome-based approach to surgical practice* (pp. 143-158). CRC Press.

Joseph, L. (2000b). Introduction to biostatistics: Describing and drawing inferences from data. In L. Rosenberg, L. Joseph, & A. Barkun (Eds.), *Surgical arithmetic: Epidemiological, statistical and outcome-based approach to surgical practice* (1st ed., pp. 14-62). CRC Press.

Keller, E. F. (1992). Gender and science: An update. In *Secrets of life secrets of death: Essays on science and culture* (pp. 1-13). Routledge.

Kosko, B. (1986). Fuzzy cognitive maps. *International Journal of Man-Machine Studies, 24*(1), 65-75. https://doi.org/10.1016/S0020-7373(86)80040-2

Kosko, B. (1988). Hidden patterns in combined and adaptive knowledge. *International Journal of Approximate Reasoning, 2*(4), 377-393. https://doi.org/10.1016/0888-613X(88)90111-9

Kruschke, J. K. (2015). Bayes’ rule. In *Doing Bayesian data analysis* (2nd ed., chap. 5, pp. 99-120). Elsevier Inc. https://doi.org/10.1016/B978-0-12-405888-0.00005-2

Langer, L., Erasmus, Y., Tannous, N., & Stewart, R. (2017). How stakeholder engagement has led us to reconsider definitions of rigour in systematic reviews. *Environmental Evidence, 6*, Article 20. https://doi.org/10.1186/s13750-017-0098-7

Lavell-Harvard, D. M., & Anderson, K. (2014). *Mothers of the nations* (D. M. Lavell-Harvard & K. Anderson, Eds.). Demeter Press.

MacDermid, J., & Graham, I. (2009). Knowledge integration: Putting the “practice” in evidence-based practice. *Hand Clinics, 25*(1), 125-143. https://doi.org/10.1016/j.hcl.2008.10.003

Mahoney, M., & Mitchell, L. (2016). *The doula*. Feminist Press.

Maxwell, J. A. (2012). Diversity is real. In *A realist approach for qualitative research* (pp. 49-67). Sage.

Maxwell, J. A., & Mittapalli, K. (2010). Realism as a stance for mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioral research* (pp. 124-167). Sage. https://doi.org/10.4135/9781506335193.n6

Merry, L., Gagnon, A. J., Kalim, N., & Bouris, S. (2011). Refugee claimant women and barriers to health and social services post-birth. *Canadian Journal of Public Health, 102*(4), 286-290. https://doi.org/10.1007/BF03404050
Mertens, D. (2003). Mixed methods and the politics of human research: The transformative-emancipatory perspective. In A. Tashakkori & C. Teddlie (Eds.), Handbook of mixed methods in social & behavioral research (pp. 135-164). Sage.

Mertens, D., & Hesse-Biber, S. (2013). Mixed methods and credibility of evidence in evaluation. In D. Mertens & S. Hesse-Biber (Eds.), Mixed methods and credibility in evaluation new directions for evaluation (Vol. 138, pp. 5-13). https://doi.org/10.1002/ev.20053

Midgley, G. (2000). Boundary critique. In Systemic intervention: Philosophy, methodology, and practice (pp. 135-158). Kluwer. https://doi.org/10.1007/978-1-4615-4201-8_7

Mingers, J. (2005). A critique of statistical modelling in management science from a critical realist perspective: Its role within multimethodology. Journal of the Operational Research Society, 57(2), 202-219. https://doi.org/10.1057/palgrave.jors.2601980

Morzaria, J., & Šajna, M. (2016). User manual for ProbTC. Author.

Munro, K., Jarvis, C., Kong, L., DSouza, V., & Graves, L. (2014). Perspectives of family physicians on the care of uninsured pregnant women. Journal of Obstetrics and Gynecology of Canada, 35(7), 599-605. https://doi.org/10.1016/S1701-2163(15)30880-X

Napier, A. D., Ancarino, C., Butler, B., Calabrese, J., Chater, A., Chatterjee, H., Chatterjee, H., Guesnet, F., Horne, R., Jacyna, S., Jadhav, S., Macdonald, A., Neuendorf, U., Parkhurst, A., Reynolds, R., Scambler, G., Shamsadani, S., Smith, S. Z., Stougaard-Nielsen, J., & . . Woolf, K. (2014). The Lancet Commissions culture and health. Lancet, 384(9954), 1607-1639. https://doi.org/10.1016/S0140-6736(14)61603-2

Niesink, P., Poulin, K., & Šajna, M. (2013). Computing transitive closure of bipolar weighted digraphs. Discrete Applied Mathematics, 161(1-2), 217-243. https://doi.org/10.1016/j.dam.2012.06.013

O’Mahony, J., & Donnelly, T. (2010). Immigrant and refugee women’s post-partum help-seekings experiences and access to care: A review and analysis of the literature. Journal of Psychiatric and Mental Health Nursing, 17(10), 917-928. https://doi.org/10.1111/j.1365-2850.2010.01625.x

Oliver, S., Roche, C., Stewart, R., Bangpan, M., Dickson, K., Pells, K., Cartwright, N., Hargreaves, J., & Gough, D. (2018). Stakeholder engagement for development impact evaluation and evidence synthesis (Vol. 3). CEDIL.

Özesmi, U., & Özesmi, S. L. (2004). Ecological models based on people’s knowledge: A multi-step fuzzy cognitive mapping approach. Ecological Modelling, 176(1-2), 43-64. https://doi.org/10.1016/j.ecolmodel.2003.10.027

Pawson, R. (2008). Causality for beginners. National Centre for Research Methods Research Methods Festival 2008: Frameworks for Research and Research Designs, Southampton, United Kingdom. http://eprints.ncrm.ac.uk/245/

Pearson, M. (2015). Using realist review to inform intervention development: Methodological illustration and conceptual platform for collaborative care in offender mental health. Implementation Science, 10, Article 134. http://doi.org/10.1186/s13012-015-0321-2

Pluye, P., & Hong, Q. N. (2014). Combining the power of stories and the power of numbers: Mixed methods research and mixed methods reviews. Annual Review of Public Health, 35, 29-45. https://doi.org/10.1146/annurev-publhealth-032013-182440

Pollock, A., Campbell, P., Struthers, C., Synnot, A., Nunn, J., Hill, S., Goodare, H., Morris, J., Watts, C., & Morley, R. (2018). Stakeholder involvement in systematic reviews: A scoping study. Systematic Reviews, 7, Article 208. https://doi.org/10.1186/s13643-018-0852-0

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., & Duffy, S. (2006). Guidance on the conduct of narrative synthesis in systematic reviews: A product from the ESRC methods programme. https://www.lancaster.ac.uk/media/lancaster-university/content-assets/documents/fhm/dhr/chir/NSsynthesisguidanceVersion1-April2006.pdf

Popay, J., & Williams, G. (1996). Public health research and lay knowledge. Social Science & Medicine, 42(5), 759-768. https://doi.org/10.1016/0277-9536(95)00341-X

Roberts, K. A., Dixon-Woods, M., Fitzpatrick, R., Abrams, K. R., & Jones, D. R. (2002). Factors affecting uptake of childhood immunisation: A Bayesian synthesis of qualitative and quantitative evidence. Lancet, 360(9345), 1596-1599. https://doi.org/10.1016/S0140-6736(02)11560-1
Sandelowski, M. (1991). Telling stories: Narrative approaches in qualitative research. *Image: Journal of Nursing Scholarship, 23*(3), 161-166. https://doi.org/10.1111/j.1547-5069.1991.tb00662.x

Sarmiento, I., Paredes-Solis, S., Loutfi, D., Dion, A., Cockcroft, A., & Andersson, N. (2020). Fuzzy cognitive mapping and soft models of indigenous knowledge on maternal in Guerrero, Mexico. *BMC Medical Research Methodology, 20*, Article 125. https://doi.org/10.1186/s12874-020-00998-w

Serrant-Green, L. (2011). The sound of “silence”: A framework for researching sensitive issues or marginalised perspectives in health. *Journal of Research in Nursing, 16*(4), 347-360. https://doi.org/10.1177/1744987110387741

Smith, L. T. (2012). *Decolonizing methodologies*. Zed Books.

Sprenger, J., & Hartman, S. (2019). *Bayesian philosophy of science*. Oxford University Press. https://doi.org/10.1093/oso/9780199672110.001.0001

Tuck, E. (2008). Theorizing back. In J. Anyon (Ed.), *Theory and educational research toward critical social explanation* (pp. 111-133). Routledge.

Ulrich, W. (1998). Systems thinking, systems practice, and practical philosophy: A program of research. *Systems Practice, 1*(2), 137-163. https://doi.org/10.1007/BF01059855

van Bertalanffy, L. (1968). *General system theory*. George Braziller.

Voils, C., Hasselblad, V., Crandell, J., Chang, Y., Lee, E., & Sandelowski, M. (2009). A Bayesian method for the synthesis of evidence from qualitative and quantitative reports: The example of antiretroviral medication adherence. *Journal of Health Services Research & Policy, 14*(4), 226-233. https://doi.org/10.1258/jhsrp.2009.008186

Wallerstein, N. B., & Duran, B. (2006). Using community-based participatory research to address health disparities. *Health Promotion Practice, 7*(3), 312-323. https://doi.org/10.1177/1524839906289376