Rapid, responsive and relevant? A systematic review of rapid evaluations in healthcare

Cecilia Vindrola-Padros¹, Eugenia Brage² and Ginger A. Johnson³

1. Department of Applied Health Research, UCL, London, UK
2. Centro de Estudos da Metrópole, Faculdade de Filosofía, Letras e Ciencias Humanas, Universidade de Sao Paulo, Brazil
3. Health Section, UNICEF, Islamabad, Pakistan

ABSTRACT (150 words)

Changing healthcare climates mean evaluators need to provide findings within shorter timeframes, but challenges remain in the creation of rapid research designs capable of delivering quality data to inform decision-making processes. We conducted a review of articles to grapple with these challenges and explore the ways in which rapid evaluations have been used in healthcare. We found different labels being used to define rapid evaluations and identified a trend in the design of evaluations, where researchers are moving away from short studies, to longer evaluations with multiple feedback loops or cyclical stages. Evaluators are using strategies to speed-up evaluations: conducting data collection and analysis in parallel, eliminating the use of transcripts, and utilizing larger research teams to share the workload. Questions persist in relation to the suitability of rapid evaluation designs, the trustworthiness of the data and the degree to which evaluation findings are used to make changes in practice.

Key words: rapid evaluation methods, rapid feedback evaluations, rapid cycle evaluations, data validity, healthcare
INTRODUCTION
Rapid evaluations are becoming more frequent in healthcare contexts. Changing healthcare climates mean evaluators need to be responsive to changing priorities and deliver evaluation findings within shorter timeframes (McNall and Foster-Fishman 2007). Timeliness has become a feature that might determine if and how findings can be used to inform decision-making processes (McNall et al. 2004). However, the delivery of findings at a time when they can be actionable remains a challenge and research and evaluations continue to lag behind the needs of evidence-based decision-making (Riley et al. 2013; Glasgow et al. 2014). Riley and colleagues (2013) argued that in order for research and evaluations to have an impact on healthcare organization and delivery, they would need to align to the 3 Rs: rapid, responsive and relevant. Alignment with the 3 Rs requires the creation of ‘rapid-learning research systems’, which bring together researchers, funders, practitioners and community partners to ask relevant questions and use efficient and innovative research designs (Riley et al. 2013). The challenge that remains is the creation of research and evaluation designs capable of delivering findings to these systems when they can inform decision-making processes.

Researchers have been experimenting with different types of research designs to make evaluations more efficient and to organize regular feedback loops so findings can be shared at key points in time. Rapid assessment procedures (RAP), rapid appraisals, rapid ethnographic assessments (REA) and rapid ethnographies were developed as research approaches (Beebe 2014; Johnson and Vindrola-Padros 2017; Vindrola-Padros and Vindrola-Padros 2018), but rapid evaluation designs were also created through approaches such as rapid evaluation methods (REM), real-time evaluations (RTE), rapid feedback evaluations (RFE) and rapid cycle evaluations (RCE). Over 10 years ago, McNall and Foster-Fishman (2007) reviewed the landscape of rapid evaluation and appraisal methods (REAM), documenting the diversity of rapid approaches and highlighting the challenges they shared. They identified an intrinsic tension between speed and trustworthiness and argued that rapid approaches would need to address issues of validity and data quality to gain greater popularity in the evaluation landscape (McNall and Foster-Fishman 2007).

Despite evident advances in the field of rapid evaluations, challenges remain in the way we define, design and implement rapid evaluations. For example, there is variability in the ways in which we define rapid timeframes as well as what we mean by evaluation (Nunns 2009; McNall et al. 2004). Short timeframes are often associated with evaluations that might appear to be rushed, less rigorous and lacking engagement with theory (McNall and Foster-Fishman 2007). These assumptions can influence how evaluation findings are viewed and, ultimately, used in practice (McNall and Foster-Fishman 2007). Evident trade-offs are present in relation to the breadth and depth of data and, some services, interventions or contexts might be more amenable to rapid evaluations, but this might not be true for others (Nunns 2009). Finally, gaps remain in our understanding of the value and use of rapid evaluations in comparison to longer-term evaluations.

The purpose of this review was to grapple with these challenges and explore the ways in which rapid evaluations have been used in healthcare. We considered common rapid evaluation designs, issues in implementation and strategies used for the sharing of findings and offer recommendations for the design of evaluations over short periods of time. We identified gaps in knowledge and future areas of exploration.

METHODS
Design
This is a systematic review of the literature based on peer-reviewed academic articles. The Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement was used to guide the reporting of the methods and findings (Moher et al. 2009). The review protocol was registered with PROSPERO (Ref: CRD42017078530).

Research questions
The review sought to answer the following questions:

1. When are rapid evaluations used in healthcare contexts?
2. How are rapid evaluations defined?
3. What are the evaluation designs (including study timeframes)?
4. How are findings disseminated and used in practice?
5. What are the benefits, limitations and challenges of carrying out rapid evaluations?

Search strategy
We used the PICOS (Population, Intervention, Comparison, Outcomes, Setting) framework (Gough et al. 2012) to develop the search strategy. We conducted a review of published literature using multiple databases: MEDLINE, CINAHL Plus, Web of Science, and Proquest Central. The searches were conducted in July 2018 and updated in June 2019. Results were combined into RefWorks and duplicates were removed. The reference lists of included articles were screened to identify additional relevant publications.

Selection
Two of the authors (blinded) screened the articles in three phases (title, abstract, and full text) based on the following inclusion criteria: 1) article was published in a peer-reviewed journal, 2) study focused on rapid evaluations, and 3) study presented findings from empirical studies or models for rapid evaluations based on case studies. Since we were interested in the use of the rapid evaluation ‘label’, we included articles that self-identified as rapid evaluations. We did not include articles on rapid assessment as we followed the lineage of rapid evaluation and appraisal methods (REAM) proposed by McNall and Foster (2007), which groups rapid assessments in a separate category to rapid evaluations due to their roots in anthropology and links with rapid ethnographic assessments. We did, however, include articles that reported findings from short evaluations as well as those that used iterative or cyclical approaches to rapid evaluation such as rapid feedback evaluation and rapid cycle evaluation (i.e. multiple short periods of data collection, analysis and feedback over time). Disagreements were discussed until consensus was reached. We did not apply any restrictions in terms of language or date of publication.

Data extraction and management
The included articles were analyzed using a data extraction form developed in REDCap (Research Electronic Data Capture). The categories used in the data extraction form are listed in Appendix 1. The form was developed after the initial screening of full-text articles. It was then piloted independently by two of the authors (blinded) using a random sample of five articles. Disagreements were discussed until consensus was reached. The form was changed based on the findings from the pilot.

Data synthesis
Data were exported from REDCap and the main article characteristics were synthesized. The REDCap report presented quantitative summaries of some of the entries in our data extraction form (for details see Appendix 1). The information entered in the free text boxes of the data extraction form was exported from REDCap and analyzed using framework analysis (Gale et al.)
The themes were based on our research questions, but we were also sensitive to topics emerging from the data.

Quality assessment
We used the Mixed Methods Appraisal Tool (MMAT) to assess the quality of the articles (Pluye et al. 2012; Pluye et al. 2014) as the review included studies with quantitative, qualitative and mixed-methods research designs. Two of the authors rated these articles independently. In cases of disagreement, the raters discussed their responses until consensus was reached. Inter-rater reliability was calculated using the kappa statistic (Landis and Koch 1977).

RESULTS
Identification of articles
The initial search yielded 2667 published articles (Figure 1). These were screened based on title and type of article, resulting in 249 articles. These articles were further screened on the basis of their abstracts, which left 40 articles for full-text review. Full-text of these articles led to 11 articles that met the inclusion criteria. One additional article was identified by reviewing the bibliography, ultimately leading to 12 articles included in the review.

We excluded articles that used rapid feedback methods for clinical testing or described rapid evaluations carried out in non-healthcare sectors. We also excluded study protocols and systematic reviews.

INSERT FIGURE 1

The 12 studies were subjected to a quality appraisal process undertaken by two of the authors using the MMAT tool (see Table 1 for study-specific appraisal results). Inter-rater agreement was 90% with a Cohen’s Kappa indicating substantial agreement (k=0.80). Overall, most studies covered three out of four criteria. Common limitations found in the articles was lack of reporting some aspects of the research design (i.e. sample composition and size) and lack of reflexivity in qualitative evaluations.

Characteristics of included studies
The characteristics of the 12 articles included in the review are presented in Table 1. The articles were published between 1993 and 2019, but half of the articles (n=6) were published before 2010. A significant amount of the studies took place in the USA (n=6), and one each in the UK, Bangladesh, India and Brazil. One study was based on a comparative study drawing from evaluations in Botswana, Madagascar, Papua New Guinea, Uganda, and Zambia and another compared services in Malawi, Uganda and Kenya.

INSERT TABLE 1

Definitions of rapid evaluations
The articles used different labels to describe their rapid evaluation designs. Six articles identified their studies as using rapid evaluation methods (REM), three used rapid feedback evaluation (RFE), and three used rapid cycle evaluation (RCE). The definitions of each of these terms are presented in Table 2. REMs were the oldest approach to rapid evaluations, followed by RFEs and RCEs. There was an overlap in definitions between RFEs and RCEs, but studies using RCEs tended to adapt the concept of rapid cycles to common iterative processes used in quality improvement (i.e. Plan-Do-Study-Act cycles).

Table 2. Definitions of rapid evaluation labels
<table>
<thead>
<tr>
<th>Articles reviewed</th>
<th>Rapid evaluation label</th>
<th>Definitions/ Key features</th>
<th>Cited literature (to support definition)</th>
</tr>
</thead>
</table>
| Anker et al. 1993; Chowdhury et al. 2004; Aspray et al. 2006; Felisberto et al. 2008; Grant et al. 2011; Munday et al. 2018 | Rapid evaluation methods (REM) | • Set of observation and survey-based diagnostic activities, which provide a basis for identifying operational problems and taking action  
• Active involvement and participation of all stakeholders | WHO 2002; Pearson 1989 |
| Bjorson-Benson, et al. 1993; McNall et al. 2004; Zakocs et al. 2015               | Rapid feedback evaluation (RFE) | • Data are continually collected, analyzed and used to inform action within a short time period  
• Aims at providing program managers with focused, timely evaluation conclusions | McNall et al. 2004; Hargreaves 2014; Sonnichsen 2000; Wholey 1983 |
| Schneeweiss et al. 2015; Keith et al. 2017; Skillman et al. 2019                 | Rapid cycle evaluation (RCE) | • Provides timely feedback to funding organizations and program staff and care providers  
• Offers support for continuous quality improvement and allows observations of changes over time | McNall and Foster-Fishman 2007; Shrank 2013 |

**Reasons for using rapid evaluations**

All reviewed articles included a brief description of the reason why the authors decided to use a rapid evaluation approach. The most frequent reason was the need for the quick turnaround of findings to inform decision-making, programs or service delivery (Schneeweiss et al. 2015; Keith et al. 2017). Related to this reason, some authors argued that rapid evaluations allowed for mid-course program corrections (Skillman et al. 2019; Anker et al. 1993; Bjornson-Benson et al. 1993). Anker et al. (1993) argued that rapid evaluations tended to have a flexible study design, allowing researchers to adapt to changing circumstances. Zakocs et al. (2015) indicated that rapid evaluations, particularly RFEs, facilitated communication with stakeholders, increasing buy-in into the project and creating a culture of learning.

**Research designs**

One of the aims of this review was to explore the ways in which rapid evaluations were carried out in practice. In terms of duration, the evaluations varied considerably, from 6 days to 3 years.
(with 20 feedback cycles). Seven of the evaluations had a mixed methods design (although one article only reported findings from the qualitative strand), three were qualitative, and two were quantitative. Four evaluations presented detailed information on the design of formative evaluations with multiple feedback loops or cycles (Zakocs et al. 2015; McNall et al. 2004; Schneeweiss et al. 2015; Skillman et al. 2019). We have synthesized the main steps used in each of these four articles (see Table 3).

**Table 3.** Examples of the steps involved in rapid feedback or rapid cycle evaluations

<table>
<thead>
<tr>
<th>Rapid Feedback Evaluation (RFE)</th>
<th>Rapid Cycle Evaluation (RCE)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Zakocs et al. (2015)</strong></td>
<td><strong>McNall et al. (2004)</strong></td>
</tr>
<tr>
<td>1. Clarify intent: Purpose, questions, study protocol</td>
<td>1. Collect existing data on program performance</td>
</tr>
<tr>
<td>2. Collect “good enough” data: Collect and analyze data quickly</td>
<td>2. Collect new data on program performance</td>
</tr>
<tr>
<td>4. Engage in reflective debrief: Discuss findings with project team</td>
<td>4. Share findings/recommendations with project team</td>
</tr>
<tr>
<td>5. Decide if more information is needed, take action or take no action</td>
<td>5. Develop and analyze alternative designs for full-scale evaluation</td>
</tr>
<tr>
<td>Repeat feedback loops (steps 2-5)</td>
<td>6. Assist in developing policy and management decisions</td>
</tr>
</tbody>
</table>

The RCE model proposed by Schneeweiss et al. (2015) began once findings were shared with decision-makers and considered the steps involved in using the findings to inform decisions. The other three models involved earlier stages of evaluation and followed similar steps: 1) preliminary work to identify the aims, existing knowledge and design, 2) an initial data collection stage, 3) an analysis and reflection stage, 4) sharing of emerging findings with key stakeholders, 5) decisions about how to move forward (e.g. informing changes in the intervention/program being evaluated or the data collected for the evaluation), and 6) cycle or feedback loop beginning again with another round of data collection.

**Data collection and analysis**
Qualitative evaluations tended to combine interviews, focus groups and observations, while quantitative evaluations and those with mixed-methods designs relied on the analysis of routinely collected data (with the exception of one article, which used questionnaires to collect quantitative data, Munday et al. 2018). Sample sizes varied and, in the case of five evaluations,
these were not reported (Anker et al. 1993; McNall et al. 2004; Felisberto et al. 2008; Schneeweiss et al. 2015; Zakocs et al. 2015).

**Strategies for speeding up evaluations**

Some of the articles included in the review reflected on the strategies used to carry out evaluations over a short period of time. These strategies could be grouped in the following categories: 1) strategies to reduce evaluation duration; 2) strategies to increase engagement, and 3) strategies for quality control. The strategies used to speed up the evaluation entailed carrying out data collection and analysis in parallel and, in the case of qualitative evaluations, engaging in multiple stages of coding as data were being collected. The synthesis of data in manageable formats (e.g. tables) was also mentioned as a strategy to make sense of the findings quickly and share them with stakeholders. Skillman et al. (2019) also decided to eliminate the transcription of interviews and replace these with detailed note-taking to reduce time and burden on the researchers. Large research teams were also used to cover more ground in shorter amounts of time (Anker et al. 2013; Chowdury et al. 2004; Grant et al. 2011; Keith et al. 2017; Skillman et al. 2019).

Anker et al. (1993) described the benefits of establishing a ‘core team’ of stakeholders during early stages of the evaluation to guarantee engagement and the use of findings. This group of stakeholders took responsibility over the evaluation and participated in decisions related to design, implementation, timeline and logistics. They also facilitated access to research sites and participants to prevent delays in the implementation of the evaluation (Anker et al. 1993).

As mentioned before, one of the key areas of concern in rapid evaluations is the validity of the findings. Some of the evaluations included a series of quality control measures to ensure the data reflected the realities being studied and to maintain consistency across teams of researchers. Anker and colleagues (1993) did this by piloting data collection instruments in the field and cross-checking the data collected by multiple researchers as data collection was ongoing. Other evaluations relied on the cross-checking of coding during the analysis phase, the development of a codebook and the training of coders (Keith et al. 2017; Skillman et al. 2019).

**Dissemination**

One of the areas we were interested in exploring in the review were common formats used to share the findings and descriptions of how the findings were used in practice. Unfortunately, not all of the evaluations included information on the dissemination of findings. The few evaluations that did share dissemination strategies most often described sharing data in the form of a written report with a selected group of stakeholders (Chowdhury and Moni 2004; Aspray et al. 2006; Skillman et al. 2019). Aspray et al. (2006) provided a detailed description of the report they generated at the end of their evaluation (20-page report shared with stakeholders). Bjornson-Benson et al. (1993), drawing from a rapid feedback design, developed short weekly reports to inform rapid decisions affecting the program under evaluation, while Anker et al. (1993) developed tables with critical indicators within seven to ten days of the evaluation start date and fuller reports after several weeks. Zakocs et al. (2015) shared information in the form of a brief seven-page memo, which was distributed two weeks after data collection had ended and included information on implementation processes and recommended changes.

**Challenges and limitations in the implementation of rapid evaluations**

In this review, we were also interested in identifying any factors the authors highlighted as barriers or challenges in the implementation of the rapid evaluations. One of the main challenges was making sure the quality of the data was adequate enough to inform decision-
making. Zakocs et al. (2015) argued that their team had to make difficult decisions on the best action to take based on the available data. The authors frequently came across the question “how do you decide when there is enough evidence to make a change?” (Zakocs et al. 2015: 478). They included a stage in their rapid feedback cycle titled collecting “good enough” data which alluded to the need to consider the extent to which the team could make decisions based on data available at the time (even if the data quality or depth could eventually be better) (Zakocs et al. 2015).

Balancing project resources, mainly in the sense of covering researcher time, was also mentioned in two of the articles (Skillman et al. 2019; Zakocs et al. 2015). According to Skillman et al. (2019) successful rapid evaluations required a relatively stable team of researchers. This stability helped maintain consistency in data collection and analysis, and particularly in the case of evaluations with qualitative research designs, it meant new researchers did not have to secure access to sites and build relationships with research participants mid-way through the evaluation.

As mentioned above, some of the articles presented strategies they used to speed up data collection and analysis. In the case of one qualitative evaluation, Skillman and colleagues (2019) relied on a combination of inductive and deductive approaches for data analysis. The authors created an analytical framework based on findings from the literature as well as their research questions and applied it to their data. Even though they were still open to topics emerging from the data, they identified not being able to use a purely inductive approach as a limitation of their evaluation (Skillman et al. 2019).

DISCUSSION
In this review, we aimed to explore the ways in which rapid evaluations have been used in healthcare, identify gaps in knowledge and highlight future areas of exploration. We were interested in identifying the extent to which rapid evaluations respond to the agenda set out in the healthcare research landscape, where research that can be rapid, responsive and relevant – the 3 Rs – has been identified as being better suited to inform changes in policy and practice (Riley et al. 2013).

We found that while the majority of studies took place in the US, a range of other countries around the world – many of which are using a multi-sited comparative approach – are also exploring the functionality of rapid evaluations. The most common reason for using rapid evaluations was the need to inform decision-making in relation to a service, program or intervention, but authors also indicated that these types of designs granted additional flexibility to the evaluators (if changes needed to be made in the design mid-way through the study) and facilitated communication and engagement with stakeholders (particularly designs with feedback loops). We found that three main labels are currently being used to define rapid evaluations: rapid evaluation methods, rapid feedback evaluations and rapid cycle evaluations. We identified a trend in the design of rapid evaluations, where evaluators are moving away from studies that are short, to longer studies with multiple short stages with feedback loops or cycles. This change in design leads to evaluation approaches that are more centered on stakeholder engagement and the continuous dissemination of findings.

While the articles we reviewed varied in their research design – qualitative, quantitative or both (i.e. mixed-methods) – qualitative evaluations tended to utilize primary data such as interview, focus group and observations while quantitative and mixed-method evaluations concentrated upon analyzing routinely collected data. Strategies to reduce data analysis timeframes included
conducting data collection and analysis in parallel, eliminating the production of transcripts in favor of detailed note-taking, and utilizing larger research teams to share the workload.

Unfortunately, not all of the articles reviewed shared details of their data dissemination strategy. Those that discussed how their findings were disseminated mentioned developing a written report to be shared with select stakeholders. These reports were often described as ‘short’ reports or ‘memos’ ranging from 7-20 pages in length with the inclusion of summary tables to aid stakeholders in better understanding study findings and their associated recommendations.

Going back to the 3 Rs proposed by Riley and colleagues (2013), we can see that the field of evaluations in healthcare is responding to the need for rapid findings in the following ways:

1. **Rapid**: Study timeframes are either short or have built-in feedback loops/cycles for the continuous sharing of findings.
2. **Responsive**: Rapid evaluation designs were identified as flexible designs that could be adapted to changes in the healthcare climate or the needs of stakeholders. Research teams tended to maintain close relationships with the evaluation users and other relevant stakeholders to keep abreast of these changes.
3. **Relevant**: Several of the evaluations included in the review involved the participation of a ‘core’ group of stakeholders who were involved in different stages of the evaluation. In some cases, these stakeholders advised on the evaluation design and implementation, while, in other instances, findings were shared with this group on a continuous basis (through feedback loops or cycles). This allowed teams to make sure the aims of the evaluation responded to the needs of stakeholders and future users of the findings.

These findings seem promising, but additional work needs to be carried out to strengthen the development of rapid evaluations. We found gaps in the reporting of information, particularly in relation to sample size. Our quality assessment also indicated that evaluations using qualitative designs rarely engaged in a process of reflection of the role of the evaluator and how their presence might influence the collection of data. We also need more information on how dissemination is built into evaluation designs (i.e. how feedback loops are negotiated with stakeholders), the formats that are effective for the sharing of findings and, ultimately, the impact of sharing findings rapidly on decision-making processes.

We also need to carry out additional work to critically examine the design and implementation of rapid evaluations and address concerns over data quality and the validity of the findings. Questions such as, “Are longer evaluations always better?” need to be fully explored and could be answered with comparative analyses of rapid and long-term evaluation designs. For example, recent work in qualitative research has compared rapid and more ‘traditional’ longer term approaches to qualitative data analysis, finding that both approaches can lead to similar, valid, results (Gale et al. 2019; Taylor et al. 2018).

The review is limited in a number of ways. The literature search for academic articles was carried out in July 2018 and updated in June 2019, but any articles published after this date were not included. Furthermore, although we employed multiple broad search terms, it is possible that we missed articles that did not use these terms. Some articles might have used other terms to describe rapid evaluations that we are not aware of. We did not include grey literature, thus potentially excluding an important number of rapid evaluations that have not been published in academic journals. For instance, we are aware that the literature on real-time evaluations (RTEs) has mainly been published in the form of reports (Sandison 2003). The tool
we used to assess the quality of the studies, the MMAT, also has limitations and these have been discussed elsewhere (Crowe and Sheppard 2011; O’Cathain et al. 2008; O’Cathain 2010).

**Conclusions**

Our review pointed to a wide range of approaches currently being used to design rapid evaluations within the healthcare landscape. Evaluators grappling with the need to deliver findings at a time when they can be used to inform decision-making are turning to rapid evaluation methods, rapid feedback and rapid cycle evaluations to make sure their evaluations are rapid, responsive and relevant. Nevertheless, questions still remain in relation to the suitability of rapid evaluation designs, the trustworthiness of the data and the degree to which evaluation findings are used to make changes in practice. Future studies could compare different rapid evaluation designs and explore the impact of rapid evaluations on the use of findings to inform decision-making.
References


