Organisation for Economic Co-operation and Development

COM/GOV/PGC/SBO/RD(2022)1

Unclassified **English - Or. English**

4 October 2022

PUBLIC GOVERNANCE DIRECTORATE
PUBLIC GOVERNANCE COMMITTEE COMMITTEE OF SENIOR BUDGET OFFICIALS

Policy evaluation in times of crisis: key issues and the way forward

OECD Conference Centre, Paris
10-11 October

This paper is shared as background to Session 6 on 'Conducting evaluations during crises' of the 4th Annual meeting of the Policy Evaluation Expert Group. It provides an overview of the challenges policy evaluators faced in the context of COVID-19, an overview of OECD governments' evaluative practices during the period, and sets out lessons for the field in this regard.
Comments from delegates are welcome until 28 October.

Claire.Salama@oecd.org
Silvia.Picalarga@oecd.org

JT03503943

Introduction

1. The COVID-19 pandemic has demanded urgent decision-making in the face of substantial uncertainties. In this fast-paced context, governments' demand for evidence increased substantially given the complexity of the situation they were facing (a so-called 'wicked challenge') and the many policies required to tackle it. But while the speed, scope and depth of innovative government responses to the pandemic made evaluation more important than ever to ensure accountability and generate policy learning, the pandemic also strained government resources. As a result, policy evaluation became both technically challenging to conduct and less of a priority for crisis-managing authorities. In this context, governments had to innovate in order to continue to evaluate their policy responses to the pandemic.

2. This paper provides an overview of the challenges policy evaluators faced in the context of COVID-19, both due to pandemic-specific hurdles and resource constraints within governments. Then, the paper provides an overview of OECD governments evaluation practices during COVID-19, with a specific emphasis on the actors, the aims and the methods involved. Finally, the third and final section sets out lessons for future policy evaluations in light of advances made during the period, both for evaluating crisis responses, as well as for the evaluation field in general.

Policy evaluation in times of crisis: why does it matter and what challenges?

3. OECD governments have implemented policy measures unprecedented in scope and depth to contain and mitigate the effects of COVID-19. While healthcare solutions took centre stage, every area of human and economic activity was affected. In many countries, emergency measures have been accompanied by a restriction of civic freedoms and rights (OECD, 2021[1]). For instance, at the start of the pandemic, 20 out of 38 OECD countries declared a state of emergency to give the executive special powers to prevent the spread of COVID-19 and to mitigate its effects on society (OECD, 2021[2]). While lockdowns, quarantines, contact-tracing and facemask mandates, were rooted in scientific advice, evidence suggests many governments also operated with lower standards of consultation, transparency, oversight, or control during COVID-19.

4. Yet, trust and transparency are crucial for citizens to understand and comply with extraordinary measures in extraordinary times. They are also key to government's capacity to absorb and bounce back from shocks (OECD, 2021[2]). In a context where 'states of emergency' curtailed the scope or time for democratic deliberation, resulting in what some authors refer to as the loss of "input legitimacy", policy evaluations can offer "results-based legitimacy" (Bekker, Ivankovic and Biermann, 2020[4]). Policy evaluations thus not only provide evidence on the results of the policies adopted during the pandemic but also serve as critical instruments to increase **accountability and transparency**.

5. *Ex post* policy evaluation, in particular, is also an essential instrument to generate **policy learning** and to draw lessons from the many policies that were adopted in this uncertain context. Discerning which solutions delivered the desired results, but also discovering their unanticipated consequences (for example the indirect impact of mobility restrictions on cancer early detection rates) is only possible with adequate evaluation. Policy evaluation can save governments from making the same mistakes during future crises.

6. During the COVID-19 crisis, however, governments encountered important challenges in conducting policy evaluations. First, **resources and time** during the crises **were scarce**. As governments were overwhelmed by the amount of challenges facing them, they were also required to access and analyse vast quantities of complex data and information in order to inform decision-making and prioritise action (OECD, 2021[2]). Even prior to the crisis, many OECD countries felt that they had insufficient analytical skills to conduct policy evaluations (OECD, 2020[7]). As the existing analytical skills were mobilised to analyse monitoring data on the epidemiological situation or collect scientific evidence on what

public health measures to adopt, governments did not have many resources at their disposal to conduct evaluations.

7. Another related challenge is that of time. Policy evaluations require significant time. Depending on the design, the resources and the needs, evaluations generally can take from several months to several years (Gawaya, Terrill and Williams, 2022[8]). In an emergency such as COVID-19, governments needed to adopt immediate policy answers to mitigate the major, and rapidly changing, impacts from the pandemic. As a response, evidence indicates that governments have innovated and altered their processes to deliver COVID-19 responses quicker, including by lowering standards of consultation, transparency, oversight or *ex ante* evaluation.

8. Second, government evaluators may have faced **technical challenges** relating to collecting data for evaluations. During COVID-19, collecting evaluation specific data (see Box 1 to understand the differences between different data sources) may have been affected by restrictions imposed by COVID-19, such as travel restrictions or limitations to group gatherings. In other words, lockdown measures would have impacted access to evaluation specific qualitative data. Moreover, survey data may also have been deprioritised in a crisis context where resources were mobilised for the immediate response.

Box 1. Sources of data for evaluations

Conducting quality evaluation requires quality data, which may come from various sources:

- **Survey data**: commonly used in research, it corresponds to census data or more generally to information on a given population collected through national or international surveys.
- **Administrative data**: this data is generally collected through administrative systems managed by government departments or ministries, and usually concerns whole sets of individuals, communities and businesses that are concerned by a particular policy. For instance, it includes housing data, tax records and data from public administrations.
- **Big data**: mainly drawn from a variety of sources such as citizen inputs and the private sector, big data is most often digital and continuously generated. It has the advantage of coming in greater volume and variety.
- **Evaluation data**: this data is collected for the purpose of the evaluation. It can take the form of qualitative questionnaires, on-site observations, focus groups, or experimental data. See further down for a description of impact evaluation methods to collect and analyse data.

Combining different data sources also has the potential to unlock relevant insights for policy evaluation.

Source: (OECD, 2021[6]) (OECD, 2020[4])

9. Even when survey and evaluation specific data were available, existing administrative data may have been difficult to use for evaluation purposes. Countries have collected a significant amount of new data during COVID-19, often at a high frequency. Yet, the data that they collected was not always useful for robust evaluations, and specifically for impact evaluations. For instance, data collected on contact cases may suffer from a number of biases due to different testing behaviours across time and places, which may skew results. Additionally, whilst OECD countries did develop data collecting standards to ensure a certain level of uniformity across institutions, the way that they are applied is always up to human interpretation. This is the case for example when it comes to classifying a death as COVID-19 related in a patient suffering from pre-existing conditions (Stoto et al., 2022[7]). Performing evaluations ignoring these potential biases can generate unreliable results. For this reason, poor quality of data might have slowed down some evaluation attempts.

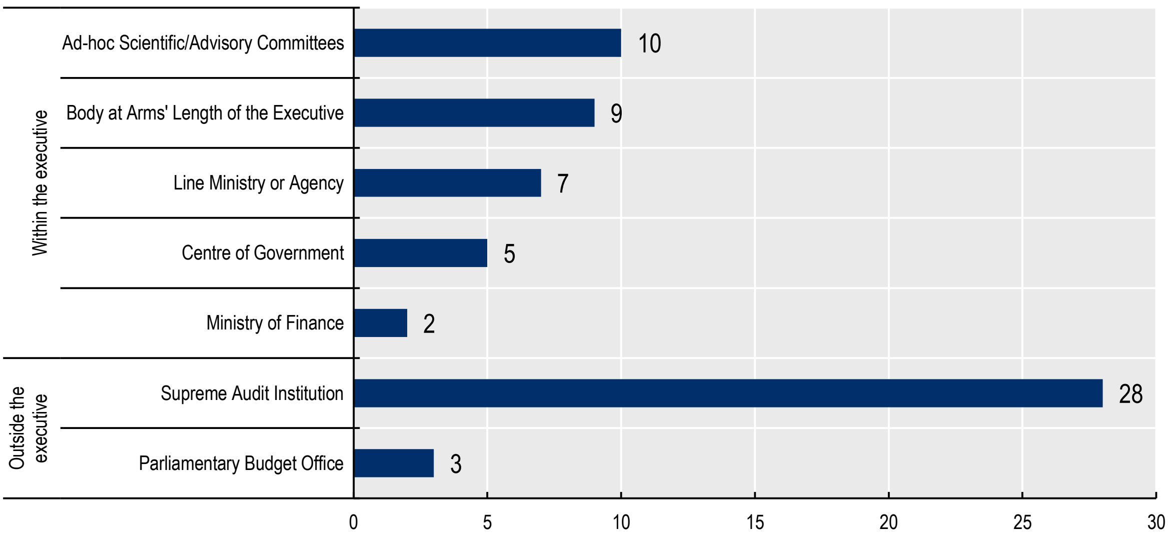
How did governments evaluate during Covid-19?

10. Despite the important challenges discussed above, OECD governments pursued efforts to evaluate their COVID-19 responses. Based on evidence from 18 OECD countries,[1] this section offers an overview on the actors involved in conducting policy evaluations during the pandemic, and the methods used to this end, in terms of timing (*ex ante*, *ex durante* or *ex post*) and criteria. The last part of the section discusses the innovative approaches that countries have used to overcome some of the challenges presented above.

The main actors to conduct evaluations of COVID-19 responses sat outside, or at arm's length, of the executive

11. The need for credible and trusted policy advice, as well as the limited resources available to governments, may explain the important role played by autonomous and independent evaluators in government. Indeed, actors at arm's length of the executive have taken on an increased evaluative role since the beginning of the COVID-19 crisis, compared to traditional actors within the executive (OECD, 2020[7]). Only three countries in the sample (Colombia, Switzerland, United States, see Figure 1) have conducted evaluations within their Centre of Government. Similarly, data collected by the OECD indicates that only one country conducted evaluations in its Ministry of Finance (Italy).

Figure 1. Number of evaluations conducted related to the COVID-19 pandemic conducted per type of institution



Note: n= 67 evaluations amongst 18 OECD countries.
Source: OECD (2022), "First lessons from government evaluations of COVID-19 responses: A synthesis", https://read.oecd-ilibrary.org/view/?ref=1125_1125436-7j5hea8nk4&title=First-lessons-from-government-evaluations-of-COVID-19-responses.

12. Conversely, actors at arm's length of the executive, whether they be permanent institutions – such as autonomous agencies – or *ad hoc* institutions, have been quite active in conducting evaluations of governments' COVID-19 responses. Autonomous agencies and other permanent **institutions at arm's length of the executive** conducted evaluations (OECD, 2020[7]) of their government's COVID-19 response in 6 countries (9 evaluations). Three countries also set up *ad hoc* institutions, specifically to provide robust and credible advice to the government during the crisis. This is the case for instance of the "Norway

Towards 2025" Commission, set up by the Ministry of Finance and the Office of Government of Norway to assess the efficiency and effectiveness of economic countermeasures to the pandemic and their impact on the Norwegian economy. Together, these two types of actors at arm's length of the executive are responsible for the majority of evaluations that were conducted by institutions of the executive.

13. Institutions outside of the executive have also played an important role in evaluating COVID-19 responses and specifically **Supreme Audit Institutions** (hereafter, SAIs) as SAIs from 8 out of 18 countries conducted evaluations (see Box 2 for an example). This share is important and may show an increased interest from SAIs to conduct evaluations. Data from an INTOSAI survey shows that in normal times, only one third of responding SAIs perform evaluations or audits of an evaluative nature (INTOSAI, 2010).

Box 2. The United States Government Accountability Office (GAO)'s evaluations of COVID-19 responses

The US Government Accountability Office (GAO) has played a very important role in the assessment of the Federal Government's response to COVID-19. This role was formally attributed by the Coronavirus Aid, Relief, and Economic Security Act of 2020, which required the GAO to issue regular reports on the impact of COVID-19. Since then, the Office has issued bi-monthly reports addressing recommendations to the Government. As of March 2022, 279 recommendations were issued and 61 of them were fully implemented. These recommendations included improvements in a variety of areas such as better reporting COVID-19 nursing home vaccination data and targeting vaccine outreach to veterans (GAO, 2022[8]).

Source: (GAO, 2022[9]) (GAO, 2022[8])

14. In Denmark, it is the Parliament that set up a specific expert group to evaluate the COVID-19 response (see Box 3).

Box 3. The Evaluation expert group for evaluating COVID-19 in Denmark

In July 2020, the **Danish Parliament** tasked an independent expert group with evaluating the first four months (January to April 2020) of the government's policy response to the pandemic. The expert group was headed by Professor Jørgen Grønnegård Christensen, of the political science department at Aarhus University. The remainder of the group included five university professors, with expertise in policy evaluation either, as economists, political scientists, or jurists. It also included the chief physician of Aarhus University. The evaluation group reported to Parliament.

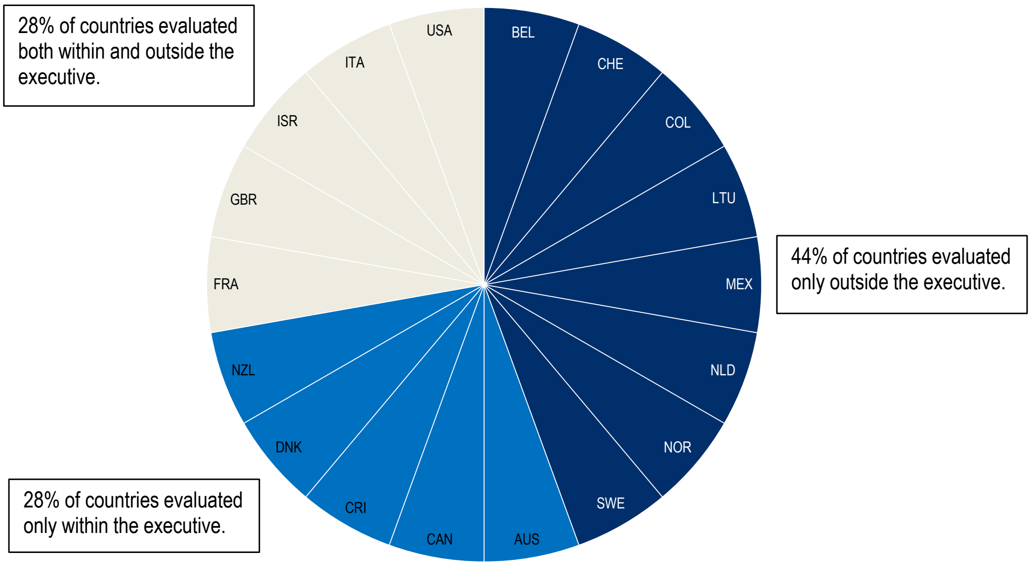
Source: (Folketinget, 2021[10]) (Jørgen Grønnegård Christensen et al., 2021[11])

15. This atypical distribution of roles can, in part, be explained by the fact that providing independent and credible evaluations is particularly important in times of crisis when trust in decision-making becomes an important issue. Indeed, as mentioned in *The Changing Face of Strategic Crisis* Management, "*when confronted with unprecedented emergency, [governments need to be] be able to quickly identify and mobilise the most relevant and trustworthy evidence to help make sense of the crisis* (OECD, 2015[12])." In that sense, distance to decision-making and policy implementation may increase the credibility and legitimacy of certain evaluative actors such as autonomous agencies. Another possible explanation of this distribution of roles is the important time and resource constraints faced by Centre of Government and

Ministries of Finance in this period. Centres of government have played an important role in tackling the crisis caused by the coronavirus (COVID-19) pandemic (OECD, 2020[19]) – especially in the immediate response and crisis management phases (OECD, 2020[20]). As a result, they may not have had the time or resources to dedicate to conducting *ex ante* assessments or collecting data for evaluations. Similarly, Ministries of Finance were quickly and massively mobilised to design response measures aimed at supporting businesses and households (Working Party of Senior Budget Officials, 2020[21]), and evaluations may not have been a priority. On the other hand, SAIs and autonomous agencies in the executive were not mobilised for crisis response and management. Simply put, these actors had more resources to invest in evaluation.

16. Less than one-third of surveyed countries report having conducted evaluations of their COVID-19 responses *within* and *outside* the executive (see Figure 2). An exception in this regard is the United Kingdom, where the Department of Health and Social Care published an "analysis of the health, economic and social effects of COVID-19 and the approach to tiering" (HM Government, 2020[25]) and the National Audit Office also conducted a report on the "implementation of employment support schemes in response to the COVID-19 pandemic" (National Audit Office, 2020[26]). Similarly, in the United States, the Office of Management and Budget published two quarterly reports on the economic impact of COVID-19 response funds and the Government Accountability Office published several evaluations that covered a variety of policy issues, such as health response measures (e.g. vaccines) or crisis communication.

Figure 2. Type of institutions having conducted evaluation(s) of COVID-19 responses



Note: n= 67 evaluations amongst 18 OECD countries. 28% of countries only conducted evaluations outside the executive (in light blue). 44% of countries only conducted evaluations within the executive (in dark blue). 28% of countries (in grey) conducted evaluations both within and outside the executive.
Source: OECD (2022), "First lessons from government evaluations of COVID-19 responses: A synthesis", https://read.oecd-ilibrary.org/view/?ref=1125_1125436-7j5hea8nk4&title=First-lessons-from-government-evaluations-of-COVID-19-responses.

17. This distribution of roles may suggest that independent and autonomous bodies could take on an increased role in the future (OECD, 2017[23]). Institutions that provide scientific policy advice to

governments, including in the form of evaluations, have an important role to play in all phases of the crisis response cycle: preparedness, crisis management response and recovery (OECD, 2018[24]). Their advice and evaluations can be particularly useful when a crisis first occurs, which is when sense-making is the most needed. Furthermore, many OECD countries increasingly rely on structural and functional independence, such as that which autonomous agencies can provide, in order to safeguard the quality and credibility of evaluations (OECD, 2020[7]).

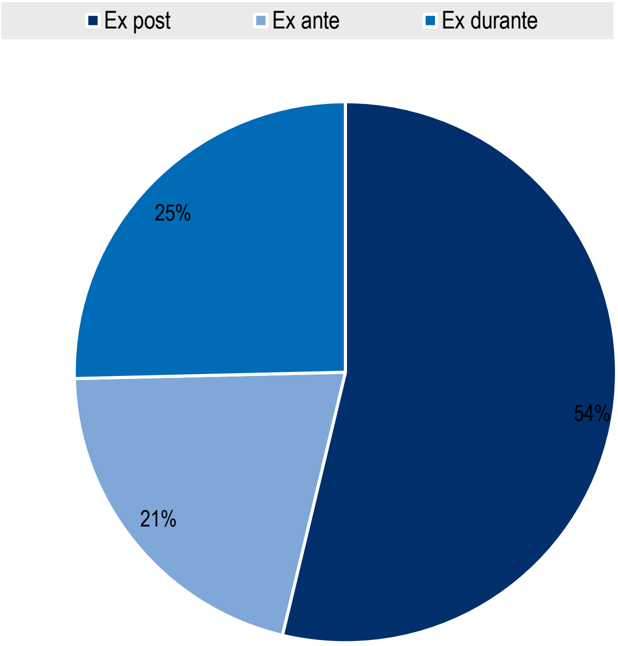
18. Yet, relying on these actors for evidence is not without its challenges. First, in so far as these different types of institutions tend to produce evaluations that respond to different policy aims (i.e. accountability, learning, transparency), and that there are complex trade-offs between evaluations' independence and their impact, conducting evaluations both within the executive and outside the executive is necessary (OECD, 2020[7]). Second, in countries where autonomous bodies or commissions were created specifically for the purpose of the crisis, a key challenge is to ensure trust in these bodies. Trust is often built over time, and it is important for these autonomous bodies to demonstrate that they are governed in a way that preserves their autonomy (OECD, 2022[20]).

Governments conducted fewer ex ante evaluations than ex post

19. The type of evaluations that were conducted by respondent institutions are consistent with the important time and resource constraints faced by governments at the onset of the crisis. Indeed, only one fifth of evaluations in the sample are *ex ante*. This is likely because governments did not have the time to conduct such assessments in the acute phase of the crisis (Figure 3.), especially because *ex ante* evaluations are typically performed by the executive and their resources were focused on the COVID-19 response.

20. Still, an important effort was made by many countries to conduct evaluations. The share of *ex durante* evaluations received by the OECD (25% of the 67 evaluations), reflects the concerted efforts by countries to evaluate the effectiveness of their COVID-19 responses in real time, as they were implementing them. Durante evaluations also allow governments to make adjustments of the policy while it is still in place, to better target the population or to fix any implementation issues. For example, the National Audit Office in the UK has conducted one of its *ex post* audits on the employment support scheme at a relatively early stage, recognising that the full impact of the scheme was not fully clear at the time of the review (UK National Audit Office, 2020[16]). Nevertheless, the majority of evaluations in the sample were conducted *ex post*, as can be seen in Figure 3.

Figure 3. Timing of evaluations



Note: n=67 evaluations amongst 18 OECD countries
Source: OECD (2022), "First lessons from government evaluations of COVID-19 responses: A synthesis", https://read.oecd-ilibrary.org/view/?ref=1125_1125436-7j5hea8nk4&title=First-lessons-from-government-evaluations-of-COVID-19-responses.

Evaluations focus on relevance, efficiency and effectiveness mostly, while issues of coherence and sustainability are under explored

21. Evaluation criteria are "standards used in evaluation as the basis for evaluative judgement." (OECD, 2021[2]) Each criteria, provides a framework to assess a policy according by answering a specific question (see Box 4).

Box 4. Policy evaluation criteria

- **Relevance** — to what extent do the (original) objectives (still) correspond to needs and issues?
- **Effectiveness** — to what extent did a policy/public intervention generate observed effects and changes? To what extent do the observed effects correspond to the objectives?
- **Efficiency** — were the costs involved justified, given the changes and effects achieved?
- **Coherence** —to what extent does the intervention fit?
- **Sustainability** — does the policy/public intervention present net benefits at the long term?
- **Impact** — what are the effects produced by an intervention (i.e. positive or negative, primary and secondary long-term effects produced, directly or indirectly, intended or unintended)?

Note: These criteria were defined by the DAC Network on Development Evaluation. They provide a normative framework to determine the merit or worth of an intervention (policy, strategy, programme, project or activity) and serve as the basis upon which evaluative judgements are made.

Source: (OECD, 2020[4]) (OECD-DAC, 2002[22])

22. Evaluations received by the OECD have mostly assessed some aspects of policies' effectiveness, but also their impact and efficiency. However, there is less evidence about the effectiveness and impact of response and recovery measures, as countries are only now starting to look at the medium and long-term effects of these policies (for an exception, see Box 5). Similarly, half of the evaluations that look at impact evaluations are actually *ex ante* assessment that look at expected or intended impacts.

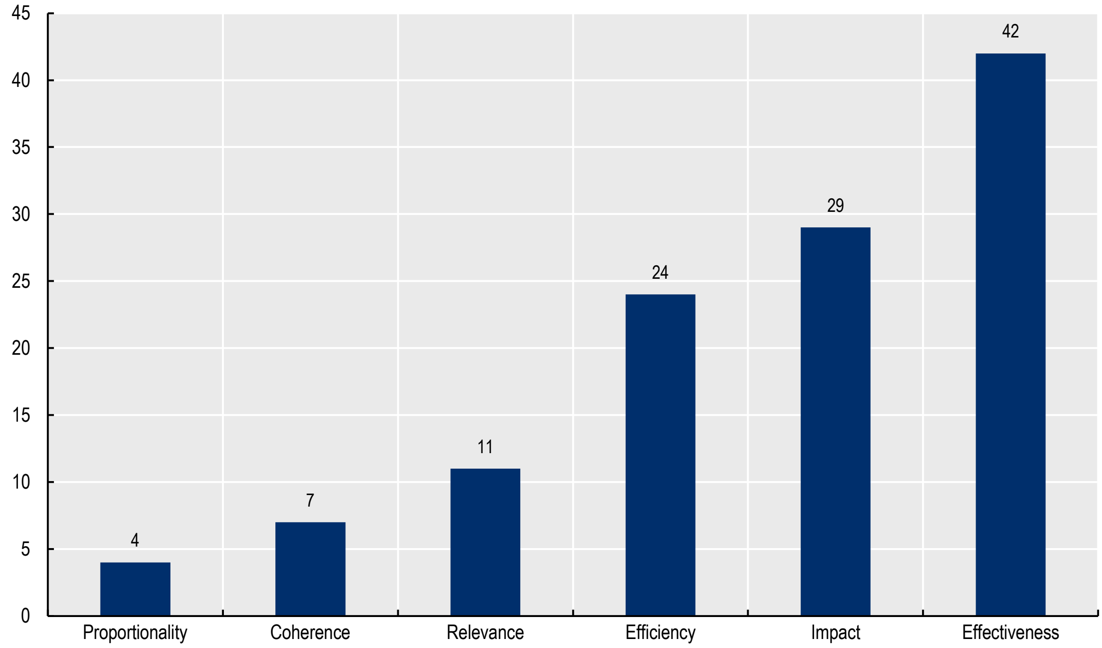
Box 5. *Ex post* evaluation of Luxembourg's COVID-19 responses

The government of Luxembourg asked the OECD to conduct an evaluation of its COVID-19 response. The *ex post* evaluation:

- **Analyses all 3 stages of the response cycle**: 1) Crisis preparation prior to the pandemic, 2) Crisis management during the initial phase of the pandemic and 3) Response and recovery policies related to public health, education, economic and fiscal measures, and social and employment policies.
- **Reviews 5 criteria**: Relevance, Coherence, Efficiency, Effectiveness, and Impact.
- **Relies on mixed-method**: the evaluation relies on qualitative data from interviews with stakeholders, and from results from OECD evaluations of other countries, as well as quantitative data from surveys administered to central government authorities and territorial actors (the 102 cities, 171 educational establishments and 4 hospitals), and from administrative micro data.

Sources: OECD, 2022, Evaluation of Luxembourg's Covid-19 Policy Response

Figure 4. Main criteria evaluated



Note: N= 67 evaluations among 18 OECD countries. One evaluation may look at different criterions.
Source: OECD (2022), "First lessons from government evaluations of COVID-19 responses: A synthesis", https://read.oecd-ilibrary.org/view/?ref=1125_1125436-7j5hea8nk4&title=First-lessons-from-government-evaluations-of-COVID-19-responses.

23. Proportionality (4 evaluations) and coherence (7) are still largely under explored by government evaluations, whilst their usefulness for policy debate may be heightened in times when resources are scarce and cross-government coordination is crucial (See Box 6). This is particularly true also in a context where normal deliberative democratic processes, as well as control and transparency mechanisms were eased to allow governments to make decisions quicker. In particular, the notion of proportionality is crucial in the debate regarding restrictions on individual freedoms such as lockdowns.

Box 6. Assessing proportionality: UK's equalities duty impact assessment of Coronavirus Act

To ensure that the **Coronavirus Act 2020** was allowing Ministries to fulfil their non-discriminatory requirements as expressed by the **Equality Act 2010**, the **Department of Health & Social Care** performed an **impact assessment**. The objective of the impact assessment was to assess whether the Coronavirus Act 2020 would have affected people in protected categories in different manners and whether this was justifiable in the context of the pandemic. The impact assessment verified also if the provisions could constitute indirect discriminations. The overall assessment stated that the provisions could, in some cases, generate more significant impacts on specific protected groups. However, in all cases these were considered proportionate to the legitimate aim of protecting the general public. Moreover, the provisions were recognised to have a positive impact on protected characteristics as age or disability which were otherwise more exposed to negative impacts of Coronavirus.

Source: (Department of Health & Social Cae, 2020[23])

Governments also adapted their evaluation methods

24. In the face of the crises, countries also adapted their evaluation plans, increased their use of evidence synthesis methods and relied on innovative evaluation or data collection methods to produce the evidence needed in these circumstances. Firstly, institutions have adapted their evaluation plans to provide timely evidence to decision-makers. For instance, the evaluation department at Global Affairs Canada had to review its evaluation schedule to provide tailored evidence to inform decision-makers (Global Affairs Canada, 2020[27]).

25. Similarly, in order to balance timing and impact considerations, several countries relied on "rapid evaluations" or decided to rely on existing evidence and evaluations of past crises instead of commissioning a new one. Rapid evaluation methods are, in this context, an interesting alternative to classic evaluations (See Box 7). They can provide meaningful insights to policymakers in a timeframe more in line with the fast pace of (political) decision-making, while the explicit involvement of key, local stakeholders allows for more tailored policy recommendations.

Box 7. Understanding Rapid Evaluation Methods

Rapid Evaluation Methods (REMs) is an expression used to refer to several types of evaluations designed to quickly and systematically provide evidence when time or resources are limited. Many terms are used to describe these approaches, including real time evaluations, rapid feedback evaluation, rapid evaluation methods, rapid-cycle evaluation and rapid appraisal. These methods share the following characteristics:

- Short time frames: the evaluations are conducted under 6 months.

- Mixed-methods design: REMs combine sources such as surveys, semi-structured interviews, and quantitative data.
- Participatory and team based: REMs involve local stakeholders, from institutions to end users to enable faster and targeted data collection. Members of the evaluation team are involved collectively at each stage of the policy assessment.
- Iterative data and analysis: data are analysed while they are collected, and initial findings are used to guide decisions on further data collection.

There are a number of reasons that governments may choose to use rapid methods as opposed to more standard evaluation approaches including where:

- A short-term or immediate outcome is expected, for example in a public health crisis such as a pandemic where improvement in infection control is expected to happen within a short timeframe;
- Near-term decision-making is needed, for example where there is an impending decision about whether to maintain, expand or terminate a pilot program; or
- There are limited resources and/or available for the evaluation.

Source: (McNall and Foster-Fishman, 2007[25]); (Williams, E., 2021[26])

26. One example of OECD country that used rapid evaluation methods during COVID-19 is Australia. The state of Victoria's Centre for Evaluation and Research Evidence conducted Rapid Evaluations to evaluate innovative public health policy solutions deployed during COVID-19 (See Box 8).

Box 8. Rapid Evaluation Methods in the state of Victoria, Australia

The Centre for Evaluation and Research Evidence (CERE) in the State of Victoria's Health Department relied on Rapid Evaluation Methods (REM) to evaluate 15 policy solutions delivered during COVID-19. The solutions ranged from telehealth delivery in pregnancy and afterbirth, to a family violence helpline web-chat function.

- The 15 policy innovations were evaluated in 6–8-weeks spread over 4 tranches.
- Beyond surveys and group discussions, tightly scoped 15–30-minute interviews with 15 lead providers were carried out face to face, via telephone or via conference with the providers.
- CERE prepared standardised templates for the evaluation plan, data matrixes, guides for the interview and focus group discussions, and final reporting templates to streamline evaluation.
- At least two fulltime-experienced evaluators were involved the evaluation of each policy solution, as well as a part-time evaluator in most cases.
- The evaluations helped identify unanticipated costs of the rapid adoption of telehealth in some areas.

Source: (Gawaya, Terrill and Williams, 2022[27])

27. Despite the interesting benefits arising from such methods, REMs do carry the risk of sacrificing scientific rigour for speed. This is why thorough preparation is of critical importance. Unlike traditional evaluations, survey respondents in REM are sometimes selected by evaluators for their diversity and proximity to the policy solution (Rubin and Rubin, 2005[28]). Concerns of internal validity are then at odds with those of researcher bias. Countries and evaluators should therefore invest in producing more thorough

evaluations during normal times to revisit or supplement evidence that produced rapidly for the purpose of crisis management.

28. Instead of producing new evaluations, even if rapid ones, some countries have relied on evidence synthesis methods. Evidence synthesis methods are a type of research method that allows researchers to bring together all relevant information on a research question (Gough, Thomas and Oliver, 2012[29]). These methods rely on aggregating learnings from already existing studies. While they do not provide new evidenced on current policies, they can be useful to inform policymakers and practitioners on the potential impacts of policies based on past experience or that of other countries. Colombia's Department of National Planning conducted an evidence gap map in order to first identify where evidence was needed on the effects of COVID-19, with special attention to economic and social issues (such as employment, social and human capital, mental health, gender issues, etc.) to better prioritise its evaluative efforts. It also used these methods to collect existing evidence and recommendations from other countries and past crises on these issues (See Box 9). Evidence gap maps can be very useful instruments as they have a double function of providing information on what we know and providing information on what we don't know yet.

Box 9. Use of evidence gap maps to inform COVID-19 in Colombia

In seeking to promote evidence-based decision-making in the context of COVID-19, the Directorate of Monitoring and Evaluation of Public Policy of the Department for National Planning has reviewed 372 documents and studies, selected from a base of more than 1000 documents from different parts of the world. These studies analysed how COVID-19 affected different policy domains and which were the action taken in areas such as Food Safety, Educational Assistance/Permanence, Social Protection, Health Care, Gender Issues, Migration, Mortality, Social and Human Capital, Mental Health, Economic Reactivation, Employment, Liquidity, Entrepreneurship, Softening of Consumption, Institutional Modernization and Digital Transformation.

The map divides the interventions in short, medium and long term and also by different typology of study and the robustness of the evidence (systematic reviews, theoretical studies, impact assessments).

The map was not only used as a useful exercise to identify areas for research investments but the Directorate produced a list of Recommendations to inform policy-makers on the potential impact of several policies (Colombian government, Department for National Planning, 2021[30]).

Source: (Colombian government, Department for National Planning, 2021[30]) (Colombian government, Department for National Planning, 2021[31])

29. Countries have also been able to rely on different sources of data for evaluations, as gathering data during periods of crisis can be challenging. Countries may have used high-frequency data (e.g. New Zealand, Australia) on policy and programme performance during the crisis (Working Party of Senior Budget Officials, 2020[21]) (See Box 10). Others may have capitalised on existing sources of data that they normally do not tap into, including geospatial, financial or social media data (IEG, 2020[28]). This use of big data for evaluation remains relatively rare in governments. Although the potential for the evaluative field is immense, this could also pose ethical issues related to the use of private personal data by governments.

Box 10. New Data sources for policy evaluation during COVID-19

With constraints on traditional evaluation data sources, governments resorted to alternative data:

- Australia's National Reserve Bank used high-frequency data from digital payments and reservations to examine the impact of COVID-19 mobility restrictions like lockdowns on goods and services consumption. Among other things, the study revealed the lockdown induced a sharp drop in gambling losses (Bishop, Boulter and Rosewall, 2022[30]).
- The Italian Ministry of Innovation turned to big data analytics to analyse the compliance with lockdown restrictions (Mehta and Shukla, 2021[31]).
- In Italy, Spain, Greece, and Portugal, telecoms operator Vodafone's analytics platform, drawing on big data and AI techniques helped track population flow to understand how policy decisions were impacting the spread of the virus (Mehta and Shukla, 2021[31]).

Source: (Bishop, Boulter and Rosewall, 2022[33]) (Caselli, Fracasso and Scicchitano, 2022[35]) (Mehta and Shukla, 2021[34]) (Gould, Joshi and Tang, 2020[36])

The way forward: evidence-based policy making in time of crisis

30. The COVID-19 era has seen profound changes in how governments made decisions and were accountable to the public. It is therefore not surprising that this period also saw important changes in the practice of policy evaluation. Even if more time is needed to understand the extent to which these changes will be lasting, some interesting lessons can be learnt for future crises and for the way that governments conduct evaluations overall.

Policy evaluations methods must be proportionate to use

31. In crises contexts, governments can consider the use of alternative evaluation methods, such as Rapid Evaluations, or evidence syntheses, to produce the information they need with limited resources and time. These methods provide governments with precious evidence that they can use real time to improve their decision-making on the crisis response. Nevertheless, while rapid evaluations and evidence syntheses represent interesting practices in times of crises, they cannot replace traditional evaluations. Indeed, rapid evaluations can offer a snapshot of a specific context in a specific period, but are unable to explain more long-term complex dynamics (Gawaya, Terrill and Williams, 2022[8]). Similarly, evidence synthesis can provide interesting insights but might not capture some context specific elements, which would require further analysis. In other words, governments face a trade-off between speed (and cost), and scientific rigor of results. As governments look to normal times, they will need to invest further in conducting more in depth *ex post* evaluations of the policies that were adopted during the crisis.

Investing in data is essential

32. Investing in more integrated, real time and accessible data systems is essential to be able to evaluate public policies. This remained a problem in several OECD countries as there were issues of harmonisation of data. Indeed, even though during COVID-19 the amount of published data was by far

greater than in any previous pandemic or disease outbreak, the volume, variety, and velocity of this data output typically did not allow for the quality checks required for rigorous research (Stoto et al., 2022[10]). As discussed above there were also significant challenges with harmonisation of data collection practices. Investing in integrated data collection systems, will not only allow governments to make better sense of upcoming risks, but also evaluate their crises responses. In this regard, the United Kingdom's Situation Center, created in 2021 to allow the government to monitor risks and model their effects, is an interesting practice for OECD governments.

Trust in evaluators is built over time

33. As discussed in this paper, most OECD countries relied on institutions outside of the executive to perform policy evaluations. This, of course, is reasonable as many civil servants were involved in the design and elaboration of policies in the first place. However, many governments also relied on *ad hoc* autonomous bodies, such scientific committees, to conduct evaluations. Yet, trust in evaluators is built over time, and proper governance of these bodies is key to ensuring their autonomy. The COVID-19 period therefore reinforced the importance of a rich evaluation ecosystem composed of different actors, with different roles (learning, accountability, planning, etc.).

Skills for using and reviewing evidence are more needed than ever

34. Evidence production skyrocketed during COVID-19. Already in August 2020, a study counted 23,634 unique documents published in major research databases on COVID-19 (Teixeira da Silva, Tsigaris and Erfanmanesh, 2020[33]). Similarly, data was produced at unprecedented rates during the crises. This represents to some extent the importance of this event in all fields of research, but it can also pose a challenge for decision-makers who have to deal with an enormous amount of information. Being equipped with the right skills to understand and use evidence has thus become essential. This is even more important as not all evidence is equal, and decision-makers not only need to understand evaluations but also to know what information they can trust. In this context, investing in skills for use of evidence will not only ensure that decisions are rooted in the best available scientific evidence, but will also contribute to clarifying the role of science in decision-making by re-empowering decision-makers in the face of scientific advice.

References

Bekker, M., D. Ivankovic and O. Biermann (2020), "Early lessons from COVID-19 response and shifts in authority: public trust, policy legitimacy and political inclusion", *European Journal of Public Health*, Vol. 30/5, pp. 854-855, https://doi.org/10.1093/eurpub/ckaa181. [3]

Belgian Health Care Knowledge Centre (2020), "Assessing the management of hospital surge capacity in the first wave of the COVID-19 pandemic in Belgium". [47]

Bishop, R., J. Boulter and T. Rosewall (2022), "Tracking Consumption during the COVID-19 Pandemic" March, https://www.rba.gov.au/publications/bulletin/2022/mar/tracking-consumption-during-the-covid-19-pandemic.html (accessed on 23 September 2022). [33]

Canadian Office of the Auditor General (2021), "Pandemic Preparedness, Surveillance, and Border Control Measures". [49]

Caselli, M., A. Fracasso and S. Scicchitano (2022), “From the lockdown to the new normal: individual mobility and local labor market characteristics following the COVID-19 pandemic in Italy”, *Journal of Population Economics*, Vol. 35/4, pp. 1517-1550, https://doi.org/10.1007/s00148-022-00891-4. [35]

Colombian government, Department for National Planning (2021), *Evidence Maps*, https://sinergiapp.dnp.gov.co/#MapasEvidencia (accessed on 27 September 2022). [30]

Colombian government, Department for National Planning (2021), “Evidencia para política pública recomendaciones del mapa de brechas de evidencia sobre intervenciones para la crisis del covid-19”, https://sinergiapp.dnp.gov.co/#MapasEvidencia (accessed on 27 September 2022). [31]

Cour des comptes (2021), “L’hébergement et le logement des personnes sans domicile pendant la crise sanitaire du printemps 2020”, http://www.ccomptes.fr (accessed on 27 September 2022). [42]

Department of Health & Social Cae (2020), *Coronavirus Act 2020: the public sector equalities duty impact assessment*, https://www.gov.uk/government/publications/coronavirus-act-2020-equality-impact-assessment/coronavirus-act-2020-the-public-sector-equalities-duty-impact-assessment#introduction (accessed on 27 September 2022). [23]

Di Girolamo, N. and R. Meursinge Reynders (2020), “Characteristics of scientific articles on COVID-19 published during the initial 3 months of the pandemic”, *Scientometrics*, Vol. 125/1, pp. 795-812, https://doi.org/10.1007/s11192-020-03632-0. [43]

Emery Tom, Pollock Gary and Scherpenzeel Annette (2019), “Securing High-quality Data on Populations”*, Policy Brief No. 23*, Max Planck Society for the Advancement of Science on behalf of the collaborative network “Population Europe”. [52]

Folketinget (2021), *Report on the authorities’ handling of the first phase of the COVID-19 pandemic*, https://www.ft.dk/da/aktuelt/nyheder/2021/01/udredning-om-covid_19 (accessed on 28 September 2022). [10]

GAO (2022), *Coronavirus Oversight*, https://www.gao.gov/coronavirus (accessed on 28 September 2022). [9]

GAO (2022), *COVID-19: Current and Future Federal Preparedness Requires Fixes to Improve Health Data and Address Improper Payments*, https://www.gao.gov/products/gao-22-105397 (accessed on 28 September 2022). [8]

Gawaya, M., D. Terrill and E. Williams (2022), “Using rapid evaluation methods to assess service delivery changes: Lessons learned for evaluation practice during the COVID-19 pandemic”, *Evaluation Journal of Australasia*, Vol. 22/1, pp. 30-48, https://doi.org/10.1177/1035719x211057630. [5]

Gawaya, M., D. Terrill and E. Williams (2022), “Using rapid evaluation methods to assess service delivery changes: Lessons learned for evaluation practice during the COVID-19 pandemic”, *Evaluation Journal of Australasia*, Vol. 22/1, pp. 30-48, https://doi.org/10.1177/1035719X211057630/FORMAT/EPUB. [27]

Global Affairs Canada (2020), *Five-Year Departmental Evaluation Plan 2020-21 to 2024-25*, https://www.international.gc.ca/gac-amc/publications/evaluation/2020/five-year-plan-quinquennal.aspx?lang=eng. [24]

Gough, D., J. Thomas and S. Oliver (2012), "Clarifying differences between review designs and methods", *Systematic Reviews*, Vol. 1/1, https://doi.org/10.1186/2046-4053-1-28. [29]

Gould, M., I. Joshi and M. Tang (2020), *The power of data in a pandemic*, https://healthtech.blog.gov.uk/2020/03/28/the-power-of-data-in-a-pandemic/ (accessed on 23 September 2022). [36]

Haber, N. et al. (2021), "Impact Evaluation of Coronavirus Disease 2019 Policy: A Guide to Common Design Issues", *American Journal of Epidemiology*, Vol. 190/11, pp. 2474-2486, https://doi.org/10.1093/aje/kwab185. [44]

HM Government (2020), *Analysis of the health, economic and social effects of COVID-19 and the approach to tiering*, https://www.gov.uk/government/publications/the-health-economic-and-social-effects-of-covid-19-and-the-tiered-approach. [16]

IEG (2020), *Adapting evaluation design in times of COVID-19 (coronavirus): four questions to guide decisions*, https://ieg.worldbankgroup.org/blog/adapting-evaluation-designs-times-covid-19-coronavirus-four-questions-guide-decisions. [32]

Independent Evaluation Group (IEG) (2020), *Adapting evaluation designs in times of COVID-19 (coronavirus): four questions to guide decisions | Independent Evaluation Group*, https://ieg.worldbankgroup.org/blog/adapting-evaluation-designs-times-covid-19-coronavirus-four-questions-guide-decisions (accessed on 28 September 2022). [41]

Jørgen Grønnegård Christensen, P. et al. (2021), "Håndteringen af covid-19 i foråret 2020". [11]

McNall, M. and P. Foster-Fishman (2007), "Methods of Rapid Evaluation, Assessment, and Appraisal", *American Journal of Evaluation*, Vol. 28/2, pp. 151-168, https://doi.org/10.1177/1098214007300895. [25]

Mehta, N. and S. Shukla (2021), "Pandemic Analytics: How Countries are Leveraging Big Data Analytics and Artificial Intelligence to Fight COVID-19?", *SN Computer Science*, Vol. 3/1, https://doi.org/10.1007/s42979-021-00923-y. [34]

National Audit Office (2020), *Implementing employment support schemes in response to the COVID-19 pandemic*, https://doi.org/9781786043399. [17]

OECD (ed.) (2022), *Enhancing public trust in COVID-19 vaccination: The role of governments*, https://read.oecd-ilibrary.org/view/?ref=1094_1094290-a0n03doefx&title=Enhancing-public-trust-in-COVID-19-vaccination-The-role-of-governments (accessed on 23 September 2022). [55]

OECD (2022), "First lessons from government evaluations of COVID-19 responses: A synthesis", https://read.oecd-ilibrary.org/view/?ref=1125_1125436-7j5hea8nk4&title=First-lessons-from-government-evaluations-of-COVID-19-responses (accessed on 22 September 2022). [50]

OECD (2022), *Scientific advice in crises: Lessons learned from COVID-19*. [20]

OECD (2021), *Better criteria for better evaluation*. [2]

OECD (2021), *Government at a Glance 2021*, OECD Publishing, Paris, https://doi.org/10.1787/1c258f55-en. [1]

OECD (2021), *Mobilising Evidence at the Centre of Government in Lithuania: Strengthening Decision Making and Policy Evaluation for Long-term Development*, OECD Public Governance Reviews, OECD Publishing, Paris, https://doi.org/10.1787/323e3500-en. [6]

OECD (2021), *Survey on Building a Resilient Response within Centres of Government*. [45]

OECD (2020), *Building a coherent response for a sustainable post-COVID-19 recovery*, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0381. [14]

OECD (2020), *Building resilience to the Covid-19 pandemic: the role of centres of government*. [13]

OECD (2020), *Improving Governance with Policy Evaluation: Lessons From Country Experiences*, OECD Public Governance Reviews, OECD Publishing, Paris, https://doi.org/10.1787/89b1577d-en. [4]

OECD (2020), “The Covid-19 crisis: A catalyst for government transformation?”*, OECD Policy Responses to Coronavirus (COVID-19)*, OECD Publishing, Paris, https://doi.org/10.1787/1d0c0788-en. [38]

OECD (2018), *Scientific Advice During Crises: Facilitating Transnational Co-operation and Exchange of Information*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264304413-en. [19]

OECD (2017), *Policy Advisory Systems: Supporting Good Governance and Sound Public Decision Making*, OECD Public Governance Reviews, OECD Publishing, Paris, https://doi.org/10.1787/9789264283664-en. [18]

OECD (2015), *The Changing Face of Strategic Crisis Management*, OECD Reviews of Risk Management Policies, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264249127-en. [12]

OECD-DAC (2002), *Glossary of Key Terms in Evaluation and Results Based Management*. [22]

Results for all (2017), “100+ Government Mechanisms to Advance the Use of Data and Evidence in Policymaking: A Landscape Review”. [40]

Rubin, H. and I. Rubin (2005), *Qualitative Interviewing (2nd ed.): The Art of Hearing Data*, SAGE Publications, Inc., 2455 Teller Road, Thousand Oaks California 91320 United States , https://doi.org/10.4135/9781452226651. [28]

Secretary General (2021), *Agricultural trade & policy responses during the first wave of the COVID-19 pandemic in 2020*, FAO, https://doi.org/10.4060/cb4553en. [53]

Srivastava, A. (2022), “Challenges for evaluation practices and innovative approaches: Lessons during COVID-19 pandemic”, *Evaluation and Program Planning*, Vol. 92, p. 102095, https://doi.org/10.1016/j.evalprogplan.2022.102095. [39]

Stoto, M. et al. (2022), “COVID-19 data are messy: analytic methods for rigorous impact analyses with imperfect data”, *Globalization and Health*, Vol. 18/1, p. 2, https://doi.org/10.1186/s12992-021-00795-0. [7]

Swiss Federal Chancellery (2020), “Chancellerie fédérale ChF Rapport concernant l’évaluation de la gestion de crise pendant la pandémie de COVID-19”. [46]

Teixeira da Silva, J., P. Tsigaris and M. Erfanmanesh (2020), "Publishing volumes in major databases related to Covid-19", *Scientometrics*, Vol. 126/1, pp. 831-842, https://doi.org/10.1007/s11192-020-03675-3. [37]

UK National Audit Office (2020), *Implementing employment support schemes in response to the COVID-19 pandemic*, https://www.nao.org.uk/report/implementing-employment-support-schemes-in-response-to-the-covid-19-pandemic/. [21]

US Government Accountability Office (2021), "COVID-19: Critical Vaccine Distribution, Supply Chain, Program Integrity, and Other Challenges Require Focused Federal Attention". [48]

Vázquez, E. et al. (2021), "Personal Freedom and Social Responsibility in Slowing the Spread of COVID-19: A Rapid Qualitative Study", *Health Education & Behavior*, Vol. 49/1, pp. 26-34, https://doi.org/10.1177/10901981211033241. [54]

Williams, E. (2021), *Rapid evaluation*. [26]

Working Party of Senior Budget Officials (2020), *The use of performance information during the COVID-19 crisis*. [15]

World Bank (2021), *Lessons from the Review of Health and Social Innovations in the Coronavirus (COVID-19)*. [51]

35.

[1] 67 evaluations, covering the period from March 2020 to June 2021.