Rüegg et al.
CABI One Health (2022)
https://doi.org/10.1079/cabionehealth.2022.0007

CABI One Health

ORIGINAL RESEARCH

OPEN ACCESS

Guidance for evaluating integrated surveillance of antimicrobial use and resistance

Simon R. Rüegg[1], Nicolas Antoine-Moussiaux[2], Cécile Aenishaenslin[3], Lis Alban[4,5], Marion Bordier[6,7], Houda Bennani[8], Birgit Schauer[9], Jean-Christophe Arnold[8], Isobel Gabain[8], Carola Sauter-Louis[10], John Berezowski[11], Flavie Goutard[7], Barbara Häsler[8] and on behalf of the CoEvalAMR consortium[†]

Abstract

Antimicrobial resistance (AMR) resulting from antimicrobial use (AMU) is an emerging threat to global health. One of the key elements for a better understanding and management of AMU and AMR is to develop effective and efficient integrated surveillance systems that consider the complex epidemiology of these issues and the impacts of resistance on humans, animals and the environment. Consequently, for this project, an international consortium of experts from multiple fields called CoEvalAMR was formed with the objectives to study user needs, characterise and compare existing tools for the evaluation of integrated AMU and AMR surveillance, apply them to case studies, and elaborate guidance on the purpose-fit selection and the use of the tools. For the comparison of evaluation tools, questions were extracted from existing tools and attributed to themes, to assess the user needs, interviews were conducted with national key stakeholders, and we applied a series of different evaluation tools to understand and document their strengths and weaknesses. The guidance was refined iteratively. From 12 evaluation tools, 1117 questions/indicators were extracted and attributed to seven emerging themes. Twenty-three experts were interviewed, who suggested to increase the ease-of-use, grant open access, provide web-based interfaces and allow results to be automatically generated. Respondents also wished for tools providing the flexibility to conduct a rapid review, or an in-depth analysis of the surveillance system, depending on the evaluation objectives. The case studies emphasised that proper evaluations require adequate resources, typically requiring the involvement of several assessors and/or stakeholders, and can take weeks or months to complete. The resulting web-based guidance comprises six main sections: 1. Introduction to surveillance evaluation, 2. Evaluation of surveillance for AMU and AMR, 3. Evaluation tools, 4. Support for selecting an evaluation tool, 5. Case studies and 6. Directory of existing tools. The audience for the guidance is personnel working in public, private, and non-governmental organisations, from public health, animal health, plant health and environmental health, at local, national and international levels. We conclude that the field is challenged by opposing user needs for reduction and simplicity versus system approaches allowing the synthesis of that knowledge to sufficiently reflect the complexity of AMU and AMR ecology for real-world decisions. The CoEvalAMR web platform allows a better understanding of the different evaluation tools and assists users in the selection of an approach that corresponds to their evaluation needs. The CoEvalAMR consortium continues to address remaining gaps and consolidate evaluation tools and approaches in the future.

One Health Impact Statement

Antimicrobial resistance (AMR) resulting from antimicrobial use is a threat to human and animal health and may have unknown impacts on the environment. To understand how governance and human behaviour relate to AMR, it is important to implement integrated surveillance across the human, animal and environmental sectors. In this work, we describe the development of guidance on how to evaluate such surveillance, what the available tools cover and what gaps remain.

Correspondence: [1]Section of Epidemiology, Vetsuisse Faculty, University of Zürich, Winterthurerstrasse 270, 8057 Zürich, Switzerland. [2]Fundamental and Applied Research for Animals and Health (FARAH), University of Liège, Liege, Belgium. [3]Centre de recherche en santé publique de l'Université de Montréal et du CIUSSS du Centre-Sud-de-l'Île-de-Montréal, Montreal, QC, Canada. [4]Department of Veterinary and Animal Sciences, University of Copenhagen, Frederiksberg, Denmark. [5]Department for Food Safety, Veterinary Issues and Risk Analysis, Danish Agriculture & Food Council, Copenhagen, Denmark. [6]ASTRE, Univ Montpellier, CIRAD, INRAE, Montpellier, France. [7]CIRAD, UMR ASTRE, Montpellier, France. [8]Veterinary Epidemiology Economics and Public Health Group, Department of Pathobiology and Population Sciences, Royal Veterinary College, Hatfield, United Kingdom. [9]Institute for Community Medicine, University Medicine Greifswald, Germany. [10]Institute of Epidemiology, Friedrich-Loeffler-Institut, Federal Research Institute for Animal Health, Südufer 10, D-17493 Greifswald-Insel Riems, Germany. [11]Scotland's Rural College, Inverness, United Kingdom. [†]CoEvalAMR consortium: The authors plus Jessica Boname, Ilias Chantziaras, Alexis Delabouglise, Marie Hallin, Madelaine Norström, Gerdien van Schaik, Katharina Stärk, Marianne Sandberg, Victor del Rio Vilas, André Ravel, Lian Thomas, and Olivier Vandenberg

Corresponding author: Simon R. Rüegg srueegg@vetclinics.uzh.ch

Received: 13 June 2022 | Accepted: 20 September 2022 | Published: 4 November 2022

© The Author(s) 2022. Open Access. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data

The integration of surveillance is essential to address the interrelated pathways across the system. Consequently, the work with the reporting interdisciplinary consortium revealed that the broad diversity of users and their needs imply that a sustained exchange is needed to address the evolving challenge. This will be important to develop meaningful measures of impact and to avoid fixing one sector's problems at the expense of another.

Keywords: antimicrobial resistance, antimicrobial use, integrated surveillance, evaluation, decision support

Introduction

Antimicrobial resistance (AMR) constitutes a serious risk to health worldwide with major economic impacts. Pathogens carrying resistance genes are present in humans, animals and the environment, and resistance genes may spread between them (González-Zorn and Escudero, 2012). Antimicrobial use (AMU), through medical, veterinary, and agronomic use and disposal, may create selective pressures, which drives pathogenic populations to evolve and disseminate AMR genes (Hedman et al., 2020; Sharma et al., 2018). Resistance genes also spread between bacterial species, meaning that genes may be transferred from commensal to pathogenic flora (González-Zorn and Escudero, 2012).

One of the key elements for a better understanding and management of AMU and AMR is to develop effective and efficient integrated surveillance systems that consider the complex epidemiology of these issues and the impacts of resistance on humans, animals and the environment. Such integrated surveillance strategies are expected to identify emerging resistance (Avery et al., 2014; Donado-Godoy et al., 2015), fuel our understanding of AMR ecology (Banerji et al., 2019), allow the estimation of the AMR burden (Schnall et al., 2019) and inform strategies for AMR reduction.

International organisations have called for collaboration across public health, veterinary, agri-food and environmental sectors and the adoption of a multi-sectoral One Health (OH) approach, to provide a more complete picture of AMU and AMR. The World Health Assembly adopted a Global Action Plan (GAP) for AMR (WHO, 2015a), urging member states to develop National Action Plans (NAPs), which explicitly includes the implementation and evaluation of systematic, integrated monitoring and surveillance of AMU and AMR (WHO, 2015b). Even if nearly half of the WHO member states have currently developed a NAP, implementing an integrated OH surveillance system for AMR and AMU remains a challenge for many countries, in particular in low- and middle-income countries (WHO, 2022). It has been highlighted in the One Health Joint Plan of Action of the Quadripartite Alliance consisting of the World Health Organisation (WHO), the World Organisation for Animal Health (WOAH, formerly OIE), the UN Food and Agricultural Organisation (FAO) and the UN Environmental Programme (UNEP) that there is a need for capacity building to carry out integrated surveillance (FAO et al., 2022).

Several integrated surveillance strategies exist globally, but their effectiveness and economic efficiency remain to be evaluated (Queenan et al., 2016; Rabinowitz et al., 2013). The need for better evidence on the value of integrated OH surveillance has been underlined by experts and decision-makers from different fields (Aenishaenslin et al., 2019; Bhatia, 2019). Besides, monitoring and evaluation are requirements of the NAPs to assess the progress made in achieving the intended goal(s), objectives and targets (FAO et al., 2016). Without regular and systematic evaluation in line with the policy cycle, there is a risk that resources will not be used efficiently and that opportunities to create better information for the targeted and effective management of AMU and AMR will be missed (Ogyu et al., 2020).

The evaluation of integrated surveillance for AMU and AMR is challenging given its complex and multifaceted nature. OH integration applies to many activities of AMU and AMR surveillance from data collection to analyses, interpretation and dissemination, and different potential surveillance outcomes and impacts can be observed over long time periods (Aenishaenslin et al., 2021). The capacity for knowledge integration or 'OH-ness' is determined by six dimensions, namely (1) systems thinking, (2) holistic planning, (3) transdisciplinary working, supported by an enabling environment to allow for (4) sharing, (5) learning, endorsed through and (6) a systemic organisation (Rüegg et al., 2018). Several of these dimensions underpin effective collaboration, cooperation and coordination, and capacity building across sectors and disciplines as postulated by the One Health High-Level Expert Panel (Adisasmito et al., 2022). The quality and appropriateness of collaboration can be measured using a recently developed tool, called ECoSur (Bordier et al., 2019; Peyre et al., 2022, chap. 9), based on the semi-quantitative evaluation of attributes related to the organisation, functioning and functionalities of the collaboration.

Despite the implicit assumption that more integration in surveillance generates better outcomes, this may not always be the case—while increased integration may come at a greater cost. Improved outcomes may take multiple forms and be conceptualised at multiple levels ranging from close-to-real-time data to long-term impacts. Examples are up-to-date information on AMU and AMR in people, animals, and ecosystems; immediate outcomes such as enhanced understanding of the AMR epidemiology or early warning signals for emerging resistance; intermediate outcomes such as policies or behaviour changes leading to lower AMU and ultimate outcomes such as healthier people and animals (Aenishaenslin et al., 2021; Aenishaenslin et al., 2019). The evaluation of these interrelated outcomes requires complex, long-term study designs and data collection plans that would ideally form part of the NAPs from the outset.

The evaluation of integrated surveillance for AMU and AMR draws from the fields of evaluating animal and public health surveillance and the evaluation of OH and knowledge integration in the broadest sense. The evaluation of animal health surveillance has a history of progressive expansion to consider more surveillance aspects and to harmonise and bring together data (Calba et al., 2015; Peyre et al., 2019). In contrast, the challenge of integrating data, information and knowledge between disciplines and stakeholders occurs in various contexts, and the evaluation of these processes has only recently received more attention. Examples can be found in interdisciplinary research (e.g. GCRF interdisciplinary research hubs[1]), teamwork (e.g. Nancarrow et al., 2013), sustainable development (Lélé and Norgaard, 2005), education (Linn, 2005) or OH (Bordier et al., 2021; Hitziger et al., 2019; Rüegg et al., 2018). Bringing together these concepts with regard to integrated AMU and/or AMR surveillance, researchers have been looking at different perspectives of the evaluation challenge, such as the Theory of Change (Aenishaenslin et al., 2021), business cases (Queenan et al., 2016), definition of integration levels (Aenishaenslin et al., 2019), institutional collaboration, governance structures (Bordier et al., 2019), progress indicators for the NAP (Kakkar et al., 2017), the FAO-ATLASS tool (FAO, 2020) or AMU metrics as drivers of behaviour (Ibrahim and Polk, 2014), among others. While these different approaches generate important feedback for evolving AMU and AMR surveillance, their disconnected nature can result in disjointed recommendations. This can make it difficult for policy,

[1] https://www.ukri.org/research/global-challenges-research-fund/interdisciplinary-research-hubs-to-address-intractable-challenges-faced-by-developing-countries/

Downloaded from https://cabidigitallibrary.org by 128.106.115.228, on 11/08/22.
Subject to the CABI Digital Library Terms & Conditions, available at https://cabidigitallibrary.org/terms-and-conditions

research and implementation groups to select relevant approaches to fulfil their evaluation goals. Given the diversity of strengths among the different evaluation tools, a holistic and comprehensive appreciation appears beneficial. Also, the sheer quantity of possible information associated with animal health surveillance has been found to be overwhelming, and therefore, European stakeholders have advocated for a centralised platform with curated advice and exchange (Häsler et al., 2019).

Consequently, for this project, an international consortium of AMU and AMR surveillance evaluation experts from multiple fields was formed. It was named 'Convergence in evaluation frameworks for integrated surveillance of AMU and AMR' (CoEvalAMR). Its objectives were to study user needs, characterise and compare existing tools for the evaluation of integrated AMU and AMR surveillance, apply them to case studies and elaborate guidance on the fit-for-purpose selection and the use of the tools. In this article, we describe the implementation of the project and the resulting online selection tool with illustrating case studies.

Materials and methods

The development of the guidance started with a workshop with the CoEvalAMR network, during which, participants discussed current evaluation frameworks and tools, approaches, and methods, identified synergies and gaps, as well as critical elements that should support the selection of the appropriate approach to evaluation. The workshop was followed by online consultations with CoEvalAMR members, a survey of user needs, case studies and an external review process. Finally, the structure, content and mode of dissemination of the guidance were proposed.

COMPARISON OF EVALUATION TOOLS

A preliminary list of evaluation tools was developed by the Viet Nam One Health Partnership for Zoonoses (OHP Office, 2014) through a literature review complemented with an expert opinion elicitation. This list was presented to the CoEvalAMR consortium for discussion and further completion with tools that may have been missed, particularly with regard to AMU and AMR monitoring, integrative or OH initiatives. To warrant comprehensive coverage, the consortium consisted of people with expertise in surveillance evaluation and/or AMR/AMU and/or system thinking and working in various professional fields, that is government, academia, research institutes, international organisations and professional organisations. The final list contained 18 evaluation tools.

The relevant documents to implement the tools were either downloaded from publicly accessible sources or provided through personal contacts in the network. For comparison of the tools, the evaluation questions or indicators were extracted and included in a table. The questions were classified in a process based on grounded theory, a systematic methodology frequently used in qualitative research. In essence, the method posits that as researchers review the data, ideas or concepts become apparent to them. These ideas/concepts are said to 'emerge' from the data. In the present process, questions were interpreted by three interacting researchers (SR, NAM and JB). Resting on the internal logic of the questions and the explanatory text, distinctive themes were built and ascribed to the questions. In an iterative process, the themes were built by reading the assigned questions and determining one or several descriptors representing the focus of interest within the realm of surveillance attributes and performances. Subsequent questions could be assigned to existing themes, and if needed, further descriptors were added to the theme definitions. If none of the preceding themes proved appropriate to express the focus of the question, a new theme was coined with relevant descriptors. Each question was subject to direct debates in order to ascribe or adapt themes and definitions. Progressively themes were adapted and complemented to incorporate the attributes and performances effectively covered by the tools until saturation, that is stabilisation of the number of themes, their headings and definitions, was reached. For comparison of the tools, the proportion of questions attributed to each theme was computed. As an alternative, a combination of text mining tools and hierarchical cluster analysis was applied to search for a meaningful semantic structure in the database (see supplementary data).

SURVEY OF USER NEEDS

To assess the user needs for anticipated guidance, CoEvalAMR members conducted interviews. From each country, one member conducted interviews with the national key AMR/AMU stakeholders. The central points of interest were how evaluations of AMU and AMR surveillance were conducted, which tools were used and which needs evaluators had identified. The interviews were focused on surveillance activities for AMU and/or AMR in any human, non-human animal or environment. At the start of the interview, the interviewees were informed about the study objectives, and how the information was going to be used and asked for consent. The ensuing structure of the conversation was provided with a questionnaire that included the following sections: (1) demographic and professional information of the interviewee, (2) description of the surveillance programme or component under scrutiny (AMU/AMR, populations covered, integration points in the system), (3) decision elements that led to the specific surveillance design and evaluation, (4) responsibilities, processes, tools and outcomes of completed evaluations, (5) personal experiences with the applied tools such as user-friendliness, complexity, efficiency or knowledge of other existing tools and (6) factors promoting or preventing future evaluations such as drivers, barriers and potential-added value. The questions were provided online *via* Google forms and as a pdf document (supplementary materials).

The interviews were conducted with managers or coordinators of private and public AMU and AMR surveillance systems, or any other person including internal or external evaluators, who had high-level knowledge of the surveillance programme, individual components or its evaluation by means of evaluation tools. If an evaluation had been conducted, but the interviewee was not aware of the evaluation methodology and outcome, then it was recommended to seek an interview with the evaluator. Because conditions and infrastructure varied in different countries (e.g. some countries had established strong public-private partnerships, and in others, surveillance was state driven), it was left to the interviewers to invite the appropriate experts. The goal for each country was to interview at least four respondents covering the relevant institutions from the animal and human health sector.

The survey questions were shared with interviewees before the interview on request. The interviews were conducted in person or on the phone to facilitate the understanding of the questions and thinking through the more complex aspects of integration. Where more convenient, the national language was used, but the responses were entered online in English by the interviewers.

For analysis, data were downloaded from Google forms and subjected to summary statistics using Microsoft Excel. Where respondents were asked to rank a range of criteria (e.g. on a scale of 1 to 5, and 'no rank', rank 1 being most important and 5 the least), each factor was subsequently given an *importance measure* according to the following formula:

$$importance\,measure = \frac{number\ of\ times\ the\ factor\ is\ ranked^2}{(mean\ rank\ when\ cited)} \quad (1)$$

The number of times the factor was ranked was set to the power of 2 to assign greater importance to the factors that were cited more often, independent of their mean rank (Häsler et al., 2014). For answers to open questions, an interpretative summary was provided using themes and topics.

Downloaded from https://cabidigitallibrary.org by 128.106.115.228, on 11/08/22.
Subject to the CABI Digital Library Terms & Conditions, available at https://cabidigitallibrary.org/terms-and-conditions

Table 1. Ten functional aspects used to report user experiences with evaluation tools.

Functional aspect	Scoring instructions (1–4)
User-friendliness	1: not satisfactory, 4: completely satisfactory
Meets evaluation needs/requirements	1: not satisfactory, 4: completely satisfactory
Efficiency	1: not satisfactory, 4: completely satisfactory
Overall appearance	1: not satisfactory, 4: completely satisfactory
Generation of actionable evaluation outputs	1: not satisfactory, 4: completely satisfactory
Allows evaluation of One Health-aspects	1: not satisfactory, 4: completely satisfactory
Workability in terms of required data	1: very complex, 4: simple
Workability in terms of required people to include	1: many, 4: few
Workability in terms of analysis to be done	1: difficult, 4: simple
Time taken for application of tool	1: > 2 months,, 2: 1–2 months, 3: 1 week – 1 month, 4: < 1 week

CASE STUDIES

The guidance should address the concrete needs of evaluators and use a language easy to understand by users working with competent authorities. Hence, members of CoEvalAMR set out to gain first-hand experience with a series of different evaluation tools to understand and document their strengths and weaknesses. For reference, a standardised scoring matrix was established to evaluate the aspects presented in Table 1. The aspects were developed in a process based on grounded theory, coining new aspects in discussions as the case studies seemed to require them. Each aspect was rated between 1 = not satisfactory and 4 = completely satisfactory. This was complemented with a strengths-weaknesses-opportunities-threats (SWOT)-like approach, considering the following four questions: (S) elements that I liked, or that the evaluation tool covered well; (W) elements that I struggled with; (O) (beneficial) elements people should be aware of when using this tool; and (T) elements that this tool covers insufficiently. Finally, the evaluators described, how well the tools covered the seven themes established according to the process described in the previous paragraph, and additionally, the aspect of AMU/AMR governance. The methods including two case studies were described in detail by Nielsen et al. (2020)) and Sandberg et al. (2021).

ONLINE GUIDANCE AND SELECTION TOOL

The initial version of the guidance content was shared as a document to all CoEvalAMR members by email with the invitation to provide unstructured feedback. This feedback was used to refine the structure and content, and to implement a preliminary version of the guidance in the form of a web-based platform. The platform guides the user through a sequence of steps. After an introduction to the evaluation of surveillance and the specific challenges of evaluating surveillance of AMU and AMR, the available evaluation tools are introduced. Subsequently, the users reach a page in which they can indicate the importance given to the seven themes established during the comparison of the evaluation tools by attributing a proportion of 100% to each. To support the selection of the best fit-for-purpose evaluation tool(s), the page computes the relative suitability score of a specific tool by summing the products of the proportion of questions in the tool for each theme with the corresponding importance given by the user. Thus, themes with a high proportion of questions and high attributed importance by the user receive the highest scores. A tool obtains a high relative suitability score if the distribution of questions in the themes corresponds to the weighting of the user. The formula is as follows:

$$\textit{Rel. suitability score} = \sum_{n=1}^{7} (\textit{Prop. of question in theme } n \textit{ X user importance given to theme } n) \quad (2)$$

This is followed by the presentation of the eight case studies in which different evaluation tools were applied. These should further clarify which tool is best suited for the user's needs. The guidance concludes with a directory of all tools and corresponding references.

The web platform was again submitted to all CoEvalAMR members for comments, and additionally to two external reviewers with expertise in AMU and AMR surveillance, who were invited to provide structured feedback on the following items:

- Is the content of the different sections clearly presented and useful?
- Do the short descriptions of each evaluation tool represent well the tools that you know? Do you see any mistakes or misinterpretations of existing tools?
- Is the decision-support tool easy to understand and useful (including the description of each theme)?
- Is the online guidance meeting its objectives (listed above)? What is missing? What would require more details or other steps?
- Is the online guidance responding to a need, according to your experience? What is missing? What would require more details or other steps?

All reviews were collected and synthesised, and used to implement the final version.

Results

COMPARISON OF EVALUATION TOOLS

During the initial workshop, 18 tools were considered potentially relevant for the purpose of this project (Table 2). After closer analysis, six tools were not included: The Health Security Financing Assessment Tool and the 'One Health Tool' were excluded, because they did not include questions relevant to AMU and AMR. The OH Systems Mapping and Analysis Resources Toolkit was excluded because it proved to be better suited towards facilitation, rather than evaluation. Several other tools were derived from one another and thus covered in essence the same concepts. To avoid duplications, SURVTOOLS (derived from SERVAL), SET (the origin of ATLASS) and the WHO-OIE Handbook for the assessment of capacities at the human-animal interface (included in the IHR Framework) were not further considered. Because the Joint External Evaluation (JEE) framework is very comprehensive, only the AMR section within the 'prevent' chapter and the full 'detect' chapter were included in the analysis. Thus, for the comparison of tools, 1117 questions or indicators were extracted from 12 evaluation tools. The manual co-construction of themes

Downloaded from https://cabidigitallibrary.org by 128.106.115.228, on 11/08/22.
Subject to the CABI Digital Library Terms & Conditions, available at https://cabidigitallibrary.org/terms-and-conditions

Table 2. Alphabetical list of identified tools which were relevant for the evaluation of the surveillance of AMU and AMR. Some tools were not considered in the present study for the reasons indicated in the footnotes.

Acronym	Name of the tool (authors)	References
ATLASS	Assessment Tool for Laboratory and AMR Surveillance Systems (FAO)	http://www.fao.org/antimicrobial-resistance/resources/tools/atlass/en/
EcoSur	Evaluation of collaboration for OH surveillance tool (CIRAD)	https://www.frontiersin.org/articles/10.3389/fvets.2019.00109/full
HACHAI[a]	WHO-OIE Handbook for the assessment of capacities at the human-animal interface (WHO and OIE)	https://www.who.int/ihr/publications/handbook_OMS_OIE_HD.pdf
HSFAT[b]	Health Security Financing Assessment Tool (World Bank Group)	https://www.ghsagenda.org/docs/default-source/default-document-library/archive---vietnam-zdap-files/day-1/s3-3---wbg_zdap-conference-in-vietnam--hsfat_final.pdf
ISSE	Evaluation framework for Integrated Surveillance Systems for AMR (Aenishaenslin *et al.*)	https://www.frontiersin.org/articles/10.3389/fvets.2021.611931/full
IHR	International Health Regulations core capacity monitoring framework (WHO)	https://www.who.int/ihr/publications/WHO_HSE_GCR_2015.8/en/
JEE	Joint External Evaluation tool (2nd edition, WHO)	https://extranet.who.int/sph/joint-external-evaluation-tool-2nd-edition
NEOH	OH-ness tool of the Network for Evaluation of OH evaluation framework (multiple mainly research institutions)	https://www.wageningenacademic.com/doi/book/10.3920/978-90-8686-875-9 and http://neoh.onehealthglobal.net/
OASIS	Outil d'Analyse des Systèmes de Surveillance (CIRAD)	https://www.jstor.org/stable/41262696?seq=1"\l"page_scan_tab_contents and https://www.plateforme-esa.fr/sites/default/files/images/documents/oasis/rapport_oasis_maj2013.pdf
OH-APP	OH Assessment for Planning and Performance (USAID)	https://www.onehealthapp.org/about
OH-SMART[c]	OH Systems Mapping and Analysis Resources Toolkit (University of Minnesota and USDA)	https://vetmed.umn.edu/centers-programs/global-one-health-initiative/one-health-systems-mapping-and-analysis-resource-toolkit
OHT[d]	OH Tool (WHO)	https://www.who.int/choice/onehealthtool/en/
PMP-AMR	Progressive Management Pathway for AMR (FAO)	http://www.fao.org/europe/events/detail-events/en/c/1197882/
PVS	Tool for the Evaluation of the Performance of Veterinary Services (OIE)	http://www.oie.int/fileadmin/Home/eng/Support_to_OIE_Members/pdf/A_PVS_Tool_Final_Edition_2013.pdf
SERVAL	Surveillance Evaluation Framework (RVC, AHVLA – now APHA, SAC)	https://www.rvc.ac.uk/Media/Default/VEEPH/Documents/SERVAL.pdf
SET[e]	Surveillance Evaluation Tool (FAO)	http://www.fao.org/AG/AGAInfo/home/en/news_archive/2017_FAO_Surveillance_evaluation_tool-SET.html
SURV-TOOLS[f]	Surveillance Tools—previously called RISKSUR (multiple, mainly research institutions)	https://survtools.org/user/login
SurF	Surveillance Evaluation Framework (New Zealand Ministry for Primary Industries)	https://www.mpi.govt.nz/dmsdocument/18091/send

[a]Excluded, because considered through the IHR framework.
[b]Excluded, because it does not cover AMU or AMR.
[c]Excluded, because designed for facilitation, not suitable for evaluation.
[d]Excluded, because it is a health financing tool without relation to one health.
[e]Excluded, because derived from ATLASS.
[f]Excluded, because derived from SERVAL.

resulted in a set of seven main concepts that are relevant when choosing an evaluation framework (Table 3). The proportion of affiliated questions to each theme was computed and is presented in Table 4. The text mining and hierarchical clustering approach failed to provide a meaningful structure to the questions (the reasons are explained in the supplementary material).

SURVEY OF USER NEEDS

Twenty-three experts were interviewed for the project. The key findings from the survey are summarised in the following text, and the full results are available in the supplementary material.

Seventeen of 23 respondents (74%) had conducted some kind of surveillance evaluation. Of the 17 evaluations, the most common primary evaluation objective was functionality and OH-ness/ the strength of OH, with 4/17 (24%) of respondents each citing these objectives. The secondary evaluation objectives mostly mentioned were functionality with 6/17 (35%) respondents and performance with 3/17 (18%) of respondents. Multiple domains and species were covered, including livestock, food chain, humans, companion animals, equids, camelids/deer, aquaculture, bees and wildlife. For the 6/23 respondents who did not conduct an evaluation, the most frequent reason was 'I/we never considered an evaluation', followed by an evaluation not being considered necessary.

Downloaded from https://cabidigitallibrary.org by 128.106.115.228, on 11/08/22.
Subject to the CABI Digital Library Terms & Conditions, available at https://cabidigitallibrary.org/terms-and-conditions

Table 3. Seven themes emerged from 1117 questions in an iterative process based on grounded theory. The questions originated from 12 evaluation frameworks considered in the CoEvalAMR analysis.

Theme	Description	Example question
Technical operations of surveillance	Questions on technical features of surveillance operations (surveillance design, laboratory capacities, management of specimens, tests applied, data management and analysis, etc.), their quality management (SOP, traceability, etc.), and the assessment of their performance (sensitivity, specificity etc.	How representative of the target population is the surveillance system? (Extracted from Serval) The sensitivity of the case or threat definition. (Extracted from OASIS)
Resources	Questions quantitatively addressing human, physical and financial resources. Questions on the training level of human resources are also considered in this category.	Are resources for rapid response during public health emergencies of national or international concern accessible? (Extracted from IHR) Adequacy of the central level's material and financial resources (Extracted from OASIS)
Output and use of information	Questions on surveillance outputs that are provided to inform public and private stakeholders, their use to inform decision-making and the benefits from this use (expected, perceived or measured).	Consider how the benefits are distributed among stakeholders, including producers, consumers, the livestock industry or society. (Extracted from SurF) How do OH outputs (OH team, information, and network) impact on decision-making? (Extracted from ISSE)
Integration	Questions considering three levels of integration: integration of data systems (within organisations and at national, regional or international level, data systems interoperation, and adherence to international testing and data standards) integration between sectors and disciplines (knowledge integration, shared decision-making and planning, and formulation of common goals) integration in the national and international context motivating the need for surveillance (link to decision-making, shared decision-making and planning between countries).	How is the interaction between people organised to foster collaboration across the initiative? (Extracted from NEOH) Are there official agreements with labs outside of the country for specialised testing, not available in-country? (Extracted from JEE)
Collaboration	Questions on the framework of collaboration (organisation of roles and responsibilities) and the object of collaboration (exchange of data, information and knowledge, sharing of capacities). This category also covers questions about the inclusive participation of stakeholders (e.g. considering gender).	The formalisation of roles and responsibilities of surveillance actors involved in collaborative modalities. (Extracted from ECoSur) Does the multi-sectoral coordination mechanism have a current One Health Strategy developed in a participatory manner with its stakeholders? (Extracted from OH-APP)
Progress and adaptivity	Questions on any structural elements allowing for the surveillance system to adapt and evolve. This may include tools, plans and agreements to evolve (e.g. continuous learning programs, external evaluation, etc.), as well as the features of management and governance allowing for regular evaluation and adaptation of operations (e.g. frequency of meeting, regularity of progress reports, etc.).	How flexible is the project design and timeline to respond to internal or external changes at long-term? (Extracted from NEOH) Implementation of supervision by the intermediary level. (Extracted from OASIS)
Surveillance items specific to AMU/ AMR	Questions that are specifically addressing AMU (recording and management) or AMR (occurrence, prevention or response).	Which structure is responsible for AMR data collection, analysis and interpretation? (Extracted from ATLASS) Are data available on the magnitude and trends of antimicrobial resistance? (Extracted from IHR)

Nine respondents used existing evaluation tools. Table 5 shows the usage and familiarity with existing tools and the reason for selecting a tool (where applied). A total of 4/9 (44%) respondents additionally used other tools not listed, including two who developed their own internal evaluation tool, one who used an EU Commission Survey and one who did not further specify the tool. One respondent applied three different evaluation tools and three respondents applied two different tools, whilst the five remaining respondents used one tool. Of the eight participants who utilised their own evaluation approach, two stated that the evaluation was mainly qualitative in design whilst another used a risk-based approach. Another explained that they developed their own evaluation because none of the pre-existing tools met their evaluation needs, and one performed an evaluation as part of EU audits. Regarding the generation of evaluation information, 4/8 (50%) respondents said that the evaluation generated the expected information, 3/8 (38%) did to some extent and 1/8 (13%) did not. Regarding the use of the evaluation results, 3/17 (18%) respondents said that recommendations were fully addressed, and 7/17 (41%) said that they were partially addressed; for the others, results were still being addressed and were not communicated, or the person did not remember.

Respondents reported that some existing tools were cumbersome and time consuming to operate, questions were poorly phrased, and the process is difficult to follow. It was suggested to improve the ease-of-use, grant open access, provide web-based interfaces and allow results to be automatically generated. Furthermore, the respondents wished for tools providing the flexibility to conduct a rapid review, or an

Downloaded from https://cabidigitallibrary.org by 128.106.115.228, on 11/08/22.
Subject to the CABI Digital Library Terms & Conditions, available at https://cabidigitallibrary.org/terms-and-conditions

Table 4. Relative importance (proportion of affiliated questions) of each theme for each tool assessed.

	Framework											
Theme	**IHR (%)**	**ISSE (%)**	**JEE (%)**	**EcoSur (%)**	**NEOH (%)**	**OASIS (%)**	**PMP-AMR (%)**	**PVS (%)**	**SurF (%)**	**OH-APP (%)**	**SERVAL (%)**	**ATLASS (%)**
Technical operations of surveillance	28	6	28	13	6	50	10	17	49	3	46	28
Resources	2	3	11	12	5	12	2	10	7	3	2	3
Output and use of information	10	30	4	8	3	8	13	10	13	0	9	6
Integration	12	30	19	19	36	6	10	7	13	51	13	7
Collaboration	12	9	14	40	30	1	1	7	8	34	4	8
Progress and adaptivity	14	3	6	8	20	23	15	4	9	6	9	4
Surveillance items specific to AMU/ AMR	1	18	16	0	0	0	48	26	0	0	0	45
Not attributed*	21	0	1	0	0	0	2	19	1	3	17	0
Total	100	100	100	100	100	100	100	100	100	100	100	100

*Questions that were not attributed to any of the seven themes.

Table 5. Usage and familiarity of surveillance evaluation tools by survey participants who conducted evaluations using existing tools (n = 9).

Evaluation Tool	**Applied**	**Familiar, but have not applied**	**Heard of, but not familiar**	**Never heard of/unsure**	**If applied, reason for selecting the tool**
ATLASS	2	2	3	2	- Involved in development of tool - Used by other countries or programme managers, Recommended by relevant institution
EcoSur	1	1	3	4	- Involved in development of tool
HACHAI	0	3	4	2	
HSFAT	0	1	2	6	
ISSE	0	2	2	5	
IHR	0	4	2	3	
JEE	1	3	1	4	- Used by other countries or programme managers
NEOH	0	0	2	7	
OASIS	2	1	1	5	- Recommended by colleagues
OH-APP	1	1	2	5	- Participant was part of a stakeholder workshop organised by P&R project (USAID) which is the developer of the tool
OH-SMART	0	2	2	5	
OHT	0	1	3	5	
PMP-AMR	0	1	4	4	
PVS	1	4	1	3	- Involved in development of tool - Used by other countries or programme managers
SERVAL	0	0	4	5	
SET	1	1	2	5	- Recommended by relevant institution
SURVTOOLS	1	2	3	3	- Recommended by colleagues
SurF	0	0	0	9	
Other Tools	4	-	-	5*	

*For 'other tools', where participants answered 'Never heard of/unsure' it meant that they did not use other tools in their evaluation

Downloaded from https://cabidigitallibrary.org by 128.106.115.228, on 11/08/22.
Subject to the CABI Digital Library Terms & Conditions, available at https://cabidigitallibrary.org/terms-and-conditions

in-depth analysis of the surveillance system, depending on evaluation objectives. It was suggested that time and resources may be saved by offering a 'modular' framework, where the user could select the modules of interest. To warrant confidence in research and for informing decisions, survey participants advised that the evaluation output should provide evidence of the quality of data integration across sectors.

The most important factor influencing whether an/another evaluation would be conducted within the next five years was the requirement by a national governing body. Further reasons for their rank are given in Table 6. Of the four respondents who provided a rank in the 'other' category, responses included having adapted tools to evaluate integrated surveillance systems for AMR, the necessity to assess the new system, based on an in-house or inter-governmental decision or a national AMR strategy group, whilst one respondent did not specify the other criterion.

In terms of the added value generated through an integrated surveillance evaluation, respondents stated that it was instrumental to raise awareness about the importance of the OH approach for AMU and AMR and identifying gaps in knowledge that may hinder fulfilment of the full potential of integrated surveillance systems. The gaps identified by the respondents are presented in Table 7.

The analysis of the data revealed that integration was evaluated primarily at the level of data collection, interpretation, and reporting, while less than 50% of evaluations identified harmonisation across sectors, sharing of data and the consideration of multiple perspectives. Most of the evaluation tools we found were not known to the majority of the survey participants. Evaluations tended to be geared towards the internal validity of surveillance and its processes, and rather low importance was given to the impact of the surveillance, or the links between knowledge integration and performance or impacts. The exposure to other frameworks seemed to have triggered some reflections about the narrow scope of custom-made tools in relation to the broader challenges of integrated AMU/AMR surveillance.

CASE STUDIES

In eight countries, case studies were conducted. The choice of surveillance programme and tool to evaluate were based on the case group's individual needs and interests. In total, six evaluation tools (ISSE, EcoSur, ATLASS, PMP-AMR, NEOH and SURVTOOLS) were assessed after being applied to the national AMU/AMR surveillance systems. Detailed reports along the outlined standardised evaluation procedure can be found online (https://guidance.fp7-risksur.eu/welcome/case-studies/). For an overview of the case studies undertaken and the users' answers to a SWOT-like analysis, please see Table 8. Here, we summarise the most prominent findings. The evaluators liked the scoring systems producing semi-quantitative results provided by the PMP-AMR, ATLASS, EcoSur and NEOH. In

Table 6. Factors influencing whether an evaluation would be conducted within the next five years, with up to five factors being ranked from greatest importance measure* (rank 1) to smallest importance measure (rank 5) (n = 23).

Factor	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	No Rank	Count of ranks given	Mean rank	Importance measure
Requirement by national government body	12	3	0	1	0	7	16	1.4	186.2
Allocation of human resources for internal evaluation	2	5	5	4	2	5	18	2.9	110.0
Generation of evaluation capacity	3	6	2	2	1	9	14	2.4	80.7
Increase in evaluation awareness	4	3	3	2	3	8	15	2.8	80.4
Allocation of financial resources for internal evaluation	3	3	3	2	4	8	15	3.1	73.4
Requirement by international government body	4	2	2	1	1	13	10	2.3	43.5
Allocation of financial resources for external evaluation	2	1	3	3	3	11	12	3.3	43.2
Training on evaluation	1	1	3	3	4	11	12	3.7	39.3
Requirement by funder	4	1	0	0	2	16	7	2.3	21.4
Change in institutional culture	1	1	3	1	2	15	8	3.3	19.7
Other	2	0	0	0	2	19	4	3.0	5.3
Requirement by private company	1	1	0	0	1	20	3	2.7	3.4
Requirement by private standard setting body	0	0	0	1	1	21	2	4.5	0.9

*Refer to Formula 1 on how to compute the importance measure.

Table 7. Gaps for integrated surveillance of AMU and AMR elicited from 23 key stakeholders and decision-makers.

Gap	Potential/ reason/ benefit
Standardisation of evaluations across sectors	Better inform policy-makers and enable prioritisation of the most effective mitigation strategies
Guidance to aid in choosing an appropriate evaluation tool for integrated AMU and AMR surveillance, depending on the evaluation goals	Time gain
Simple, clear and detailed training for the evaluation process	More consistent application between evaluators, thus minimising personal bias in evaluation results
Gaps remain for all aspects of AMU and AMR, from the measurement of impact to integration across sectors	

Downloaded from https://cabidigitallibrary.org by 128.106.115.228, on 11/08/22.
Subject to the CABI Digital Library Terms & Conditions, available at https://cabidigitallibrary.org/terms-and-conditions

contrast, ISSE and SURVTOOLS were found to be a step away in terms of usefulness, as they resulted in a plan for how to conduct evaluation(s). The evaluators commented that ISSE, EcoSur, NEOH and SURVTOOLS allowed for in-depth analysis and therefore required more complex data, information and specific training of evaluator(s). Furthermore, NEOH and ISSE were perceived by the users as the best tools for evaluation of broader OH aspects, and EcoSur and ISSE as best for evaluation of the quality of collaboration. Users praised the ease of use of the PMP-AMR tool and found that ATLASS was particularly well-designed for risk managers. ATLASS was the only tool focusing specifically on laboratory activities. Only ISSE included the production of a direct measure for the 'integration' and 'impact on decision-making'.

The experiences emphasised that proper evaluations require adequate resources. The typical evaluation would require the involvement of several assessors and/or stakeholders and take weeks or months to complete. As the tools differ in focus, it is important to consider carefully the choice of tool(s) before embarking on the evaluation, so that it fits with the objective of the evaluation as well as the resources available.

ONLINE GUIDANCE

Twenty experts participated in the initial CoEvalAMR workshop, eight consortium partners provided feedback on the proposed structure and content of the guidance, and seven provided structured feedback on the preliminary version of the guidance web platform, in addition to the two external reviewers. The final web platform (https://guidance.fp7-risksur.eu/) comprises six main sections: (1) Introduction to surveillance evaluation, (2) Evaluation of surveillance for AMU and AMR, (3) Evaluation tools, (4) Support for selecting an evaluation tool, (5) Case studies and (6) Directory of tools. The audience of the guidance is personnel working in public, private and non-governmental organisations, from public health, animal health, plant health and environmental health, at local, national and international levels.

The first section briefly describes the needs and rationales for surveillance evaluation in general. The second section presents the goals of AMU and AMR surveillance evaluation and explains the specificity of evaluating integrated surveillance systems. This section describes a Theory of Change for integrated surveillance systems of AMU and AMR and five evaluation levels based on the ISSE framework, adapted from Aenishaenslin et al. (2021). The evaluation tools included in the comparison analysis are mapped according to their suitable evaluation levels: NEOH and EcoSur for OH integration; ATLASS, IHR, JEE, OASIS, Surf and SERVAL for the production of information and expertise; ATLASS, PVS and PMP-AMR for knowledge generation; and SERVAL and SurF for decision-making, and SERVAL and SurF. The third section provides a rapid interactive overview of all tools with a brief description, a reference

Table 8. Results of SWOT-like analysis of six different evaluation tools used in eight country*-based case studies.

Framework	ATLASS	ECoSur	ISSEP	NEOH	PMP-AMR	SURVTOOLS
Applied in*	DK	VN, DK	CA, UK	DK, BE, IT, NO, NL	BE, DK, IT, NO	DK, NL
Like	Automated analyses Progress monitoring Easy to communicate results	Comprehensive evaluation of collaboration Participatory evaluation Provision of a clear guidance	Provision of a conceptual model for integrated surveillance of AMU and AMR surveillance	Allows for comprehensive and multi-facet-ted OH assessment Evaluation of implementation quality	Easy to use Progress monitoring Participatory evaluation Evaluation of the implemen-tation levels	Objectivity Comprehensive framework for different evaluation aspects
Difficulty	Why need for such detailed data?	Evaluation of collaboration only	No provision of guidance to collect and analyse of data	Cumbersome to use	Subjectivity when scoring Crude scoring method	Requirement of training for conducting evaluation Time consuming for evaluation of complex aspects
Be aware of	Not possible to measure minor progress of epidemiological performance	Characterisation and evaluation of integration regarding collaborative objectives and context	Necessary to combine with other tools depending on the evaluation question	Requirement of training for application Resource-demanding	Complexity in terms of people to include Self-assess-ment tool Results are not comparable across countries	Provision of an evaluation plan only, not specific to AMU and AMRc
Not covering	Environment and plant sector specifically	Surveillance performance	Guidance for conducting evaluation	Progress monitoring Surveillance performance	One Health assessment Distinction between ongoing and incomplete activities Evaluation of quality of activities	Laboratory aspects One Health assessment

*BE: Belgium, CA: Canada, DK: Denmark, IT: Italy, NL: the Netherlands, NO: Norway, UK: United Kingdom, and VN: Vietnam.

Adapted after Sandberg *et al.* (Sandberg et al., 2021).

Downloaded from https://cabidigitallibrary.org by 128.106.115.228, on 11/08/22.
Subject to the CABI Digital Library Terms & Conditions, available at https://cabidigitallibrary.org/terms-and-conditions

link, the evaluation framework from which it derives and a link to case studies in which it was applied. In the fourth section, the users indicate their priorities for the seven themes (Table 3) and receive a bar chart with the relative suitability score for the 12 frameworks (Fig. 1). To support the user in prioritising each theme, examples of evaluation questions are provided. The fifth section presents the case studies to elaborate on the experience of users with selected evaluation tools. The last section provides a detailed description of each tool, including its purpose, scope, evaluation process, output and references.

Discussion

SURVEILLANCE EVALUATION

The complex, multifaceted ecology of AMR and the multiple intertwined connections to AMU in humans, animals and agriculture represent a typical wicked problem (Wilson et al., 2020). For its appreciation, it would be important to gain a better understanding of the occurrence of AMU and AMR in different sectors at national and global scales. However, only six of the 17 evaluations mentioned by survey respondents actually covered the use of as well as the occurrence of resistance to antimicrobials. For each tool used, respondents made recommendations for improvements resulting in a diverse set of statements (Table 7). This reflected the diverse evaluation needs and user expectations depending on their context and emphasised the need for guidance expressed by many survey participants. More specifically, knowledge of the most effective and efficient OH surveillance strategies for AMU and AMR is lacking (Schnall et al., 2019), making the development of benchmarks and best practices challenging. Secondly, clear guidance on how to evaluate the impact and socio-economic value of integrated surveillance systems is still scarcely provided through evaluation tools. This was in contrast to the view of stakeholders and decision-makers considering this aspect to be of critical

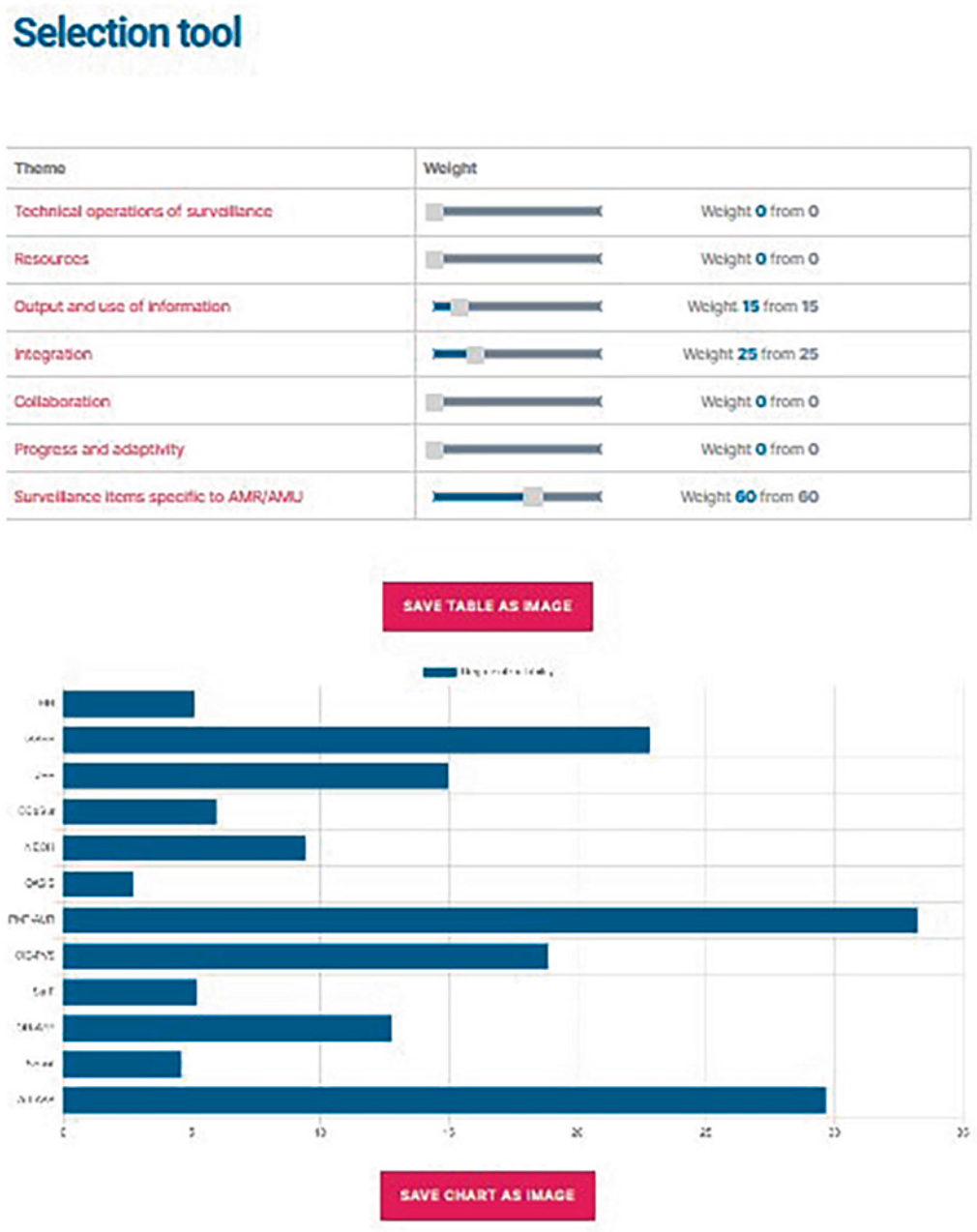

Fig. 1. Screen for user preference input (top) and relative suitability output (bottom) of the online decision-support tool (Section 4) created to guide the selection of evaluation frameworks for integrated AMU and AMR surveillance. The user prioritises the seven themes by distributing a total of 100% over all themes. The page then returns the relative suitability according to formula 1 for each evaluation tool. https://guidance.fp7-risksur.eu/welcome/decision-support/.

Downloaded from https://cabidigitallibrary.org by 128.106.115.228, on 11/08/22.
Subject to the CABI Digital Library Terms & Conditions, available at https://cabidigitallibrary.org/terms-and-conditions

importance. Thirdly, the stakeholders wished for a way to assess the quality of data integration across sectors.

One explanation for these gaps may be the dominance of veterinary professionals and the partial absence of social scientists in this field of research (Robinson et al., 2016). Nevertheless, in our selection, we found several promising evaluation tools (EcoSur, NEOH, ISSE), which require further case studies to assess their strengths, weaknesses and usefulness for surveillance stakeholders in different contexts. In this line of thought, further tools beyond the realm of AMU/AMR and surveillance may be interesting and worth bringing to the attention of the community.

THE THEORY OF CHANGE OF AMR SURVEILLANCE

Throughout the project, the realisation emerged that a generic theory of change (TOC) for integrated AMR surveillance as proposed by Aenishaenslin et al. (2019) plays a pivotal role in facilitating a common understanding of the multiple expected outcomes from implementing an OH approach in AMU and AMR surveillance systems. While looking for source attribution in a conventional, linear approach is inappropriate for a problem such as AMR, and this TOC allowed a more holistic view of elements involved in the socio-ecological system of AMU and AMR. The need for a central sense-making TOC was further supported by the observation that the automated text mining and hierarchical clustering did not produce intelligible categories. The co-construction of themes based on the question texts appeared to reveal the implicit TOCs (Table 3), which contained descriptions in line with the TOC of Aenishaenslin et al. (2019). However, with this procedure, the governance of AMU/AMR did not emerge as a specific theme due to the framing of the question: because the intention was to organise questions in relation to the evaluation of AMR surveillance, not the governance of AMU, several questions were classified in the themes 'collaboration' or 'progress and adaptivity' that would typically concern the governance of AMU. Governance of surveillance *per se* did not emerge as a theme because the three researchers interpreted governance elements (e.g. number of meetings, definition of roles and responsibilities) as means to an end, and thus assigning them to the theme including this end, such as 'collaboration'. In the case studies, different evaluation tools held varying strengths and weaknesses, but no single tool covered all the aspects of AMU/AMR.

The most pragmatic approach for better understanding, to bring the available tools into relation and to provide some oversight appeared to map them to the TOC. Embedded in the online guidance, this insight can support technical staff, evaluators and advisors working in AMU and AMR surveillance in selecting the themes relevant for their evaluations. While a single tool may not be perfect for a specific evaluation, the guidance will reveal the best fit for purpose and complementary tools, saving the evaluator precious time testing and examining the variety of available tools. By providing a comprehensive frame of reference, this may also contribute to the development of improved evaluation protocols, approaches and metrics.

EVALUATION TOOL SELECTION

Survey participants highlighted that there was a lack of resources and/or assistance for end users in choosing an evaluation tool for integrated AMU and AMR surveillance. Participants wished to see an easy-to-use and tailorable tool to fit all evaluations. Unfortunately, no tool can fulfil this request, because of the complexity of the AMR ecology and consequently of the diverse and dynamic aspects of surveillance evaluation needs. As is typical for wicked problems, the system in which they occur evolves rapidly, goals change and people engage (Hester and Adams, 2017, pp. viii–x; Wilson et al., 2020), and thus surveillance requirements frequently change, as do the functioning and the goal(s) of evaluation(s): Evaluation questions are shaped by their context and a protocol is influenced by the evaluators and their experience (Williams, 2019, p. i), surveillance systems may improve through their own evaluation, and human judgement is applied at every step to capture such information. Consequently, evaluation tools will always need to offer the flexibility of a contextual adaptation and cover a range of different evaluation needs. Admittedly, there is a risk of information overload when too much flexibility and too many tools and options are available (Eppler, 2003; Häsler et al., 2019). Therefore, this online guidance was built from the user's perspective, eliciting evaluation needs first and then attempting to narrow the field of options to a manageable task. The veterinary background of the three researchers developing the themes may have caused some bias and a potential next version of the tool should consider broadening the expertise when adapting the tool to the needs and expectations of the largest possible audience. Despite having CoEvalAMR members from low- and middle-income countries (LMICs) a majority of the information was collected in high-income countries due to a lack of capacities in LMICs and the funding structure. Thus, further needs may arise when such evaluations are applied more broadly. Also, further case studies from LMICs would be welcome. The case studies provide a prior user experience with a specific evaluation tool and surveillance system. Undecided evaluators may combine the selection tool with these experiences, to enhance their ability to choose a tool suiting their ambitions and objectives. Ultimately, the challenge for the evaluators remains to consider carefully where their surveillance programme for AMU/AMR stands and what it should achieve considering inputs, activities, outputs and outcomes. This is embedded in the broader challenges stated in a survey of AMR stakeholders, namely the inadequacy of one-size-fits-all policies, evaluation of the impact of policy, a policy that is reactive rather than proactive, a plurality of stakeholders, unclear ownership of the problem, differing risk perceptions, inadequate resources to address the problem and research not being translated into action, among others (Sargeant et al., 2007). It also reflects the conundrum with which the evaluator is left alone, that a holistic (One Health) evaluation should identify the role of surveillance in its context, in the broader goal of salutogenesis for people, non-human animals and ecosystems (Berezowski et al., 2019; Lindström and Eriksson, 2006). Our survey revealed that citizens were rarely considered as stakeholders and evaluation was perceived rather as a prescribed control mechanism than a tool to facilitate adaptive governance of AMU and AMR. Our project further highlights that before the launch of an evaluation, the assessors need to be selected and trained carefully, as they have a significant impact on the scope and the outcomes.

WEB PLATFORM EVOLUTION

We did not intend the guidance as a comprehensive one-way prescription. Rather it should provide a platform open to the AMU and AMR surveillance and evaluation community, for further development, learning and exchange of ideas. Like the wicked problem of AMU/AMR itself, the development and use of this guidance is a cyclic process. The guidance website remains curated and provides a form for users to share their experiences, enabling a flow of information back and forth. In our survey, participants stated that they did not use certain tools in previous evaluations, because they were unaware of their existence. This resonates with the findings of other researchers that in the context of AMU/AMR knowledge was poorly communicated, research was not being translated into action, and platforms for knowledge exchange are needed (Häsler et al., 2019; Sargeant et al., 2007). The web platform responds to these calls and facilitates knowledge dissemination effectively and in a complexity-aware format. In the mid-term, we expect that the application of the tool at national and international levels will facilitate a collection of case studies providing evidence for the suitability of evaluation tools in respect to different aspects of integrated AMU and AMR surveillance.

Downloaded from https://cabidigitallibrary.org by 128.106.115.228, on 11/08/22.
Subject to the CABI Digital Library Terms & Conditions, available at https://cabidigitallibrary.org/terms-and-conditions

Consequently, gaps in concepts, methods and data should emerge and offer opportunities to increase the effectiveness and efficiency of current practices in AMU and AMR surveillance. In the long term, we hope that this will lead to healthier human and non-human animal populations through effective governance of antimicrobials.

Conclusion

This project has contributed an important point of reference in the complex and critical space of evaluation of integrated surveillance systems for AMU and AMR. The tools differ, for example regarding objectives, and no single tool is comprehensive, and some critical gaps remain. The field is challenged by opposing needs for reduction and simplicity to generate scientific knowledge, and for the synthesis of that knowledge to sufficiently reflect the complexity of AMU and AMR ecology for real-world decisions. The CoEvalAMR web platform allows a better understanding of the different evaluation tools and assists users in the selection of an approach that corresponds to their evaluation needs. Its usefulness in providing guidance in a growing diversity of tools and raising awareness for needs in the future will be determined by its uptake in the community. We hope for broad engagement on the website to advance our knowledge and direct integrated AMU/AMR surveillance towards the most adequate tools. The CoEvalAMR consortium continues to address remaining gaps in integrated surveillance evaluation and to consolidate knowledge about evaluation tools and approaches.

CONFLICT OF INTEREST

LA works for an organisation which gives advice to farmers and meat-producing companies. The other authors have no conflicts of interest to declare.

ETHICS STATEMENT

For the survey of user needs, ethical approval was sought from and granted by the Social Sciences Research Ethical Review Board (SSRERB) of the Royal Veterinary College, London, United Kingdom (Approval number URN SR2019-0291).

AUTHOR CONTRIBUTIONS

The project was conceived and managed by CA and BH. The workshops were hosted and coordinated by CA, BH, FG and MB. Working packages were implemented by CA, BH, SR, NAM, JB, HB, LA, BS, CSL, MB and JCA. The manuscript was drafted and revised by SR and IG with active support through complementation and refinement from all co-authors.

FUNDING STATEMENT

The project was funded by the EU Joint Programming Initiative on Antimicrobial Resistance (JPIAMR; https://www.jpiamr.eu/).

References

Adisasmito, W.B., Almuhairi, S., Behravesh, C.B., Bilivogui, P., Bukachi, S.A., et al. (2022) One health: a new definition for a sustainable and healthy future. *PLoS Pathogens*, 18, e1010537. DOI: 10.1371/journal.ppat.1010537

Aenishaenslin, C., Häsler, B., Ravel, A., Parmley, E.J., Mediouni, S., et al. (2021) Evaluating the integration of one health in surveillance systems for antimicrobial use and resistance: a conceptual framework. *Frontiers in Veterinary Sciences*, 8. DOI: 10.3389/fvets.2021.611931

Aenishaenslin, C., Häsler, B., Ravel, A., Parmley, J., Stärk, K., et al. (2019) Evidence needed for antimicrobial resistance surveillance systems. *Bulletin of the World Health Organization*, 97, 283–289. DOI: 10.2471/BLT.18.218917

Avery, B., Parmley, E., Reid-Smith, R., Daignault, D., Finley, R., et al. (2014) Canadian integrated program for antimicrobial resistance surveillance: retail food highlights, 2003–2012. *Canada Communicable Disease Report*, 40, 29–35. DOI: 10.14745/ccdr.v40is2a05

Banerji, A., Jahne, M., Herrmann, M., Brinkman, N. and Keely, S. (2019) Bringing community ecology to bear on the issue of antimicrobial resistance. *Frontiers in Microbiology*, 10, 15. DOI: 10.3389/fmicb.2019.02626

Berezowski, J., Rüegg, S.R. and Faverjon, C. (2019) Complex system approaches for animal health surveillance. *Frontiers in Veternary Sciences* 6, 1–11. DOI: 10.3389/fvets.2019.00153

Bhatia, R. (2019) Implementation framework for One Health approach. *The Indian Journal of Medical Research*, 149, 329. DOI: 10.4103/ijmr.IJMR_1517_18

Bordier, M., Delavenne, C., Nguyen, D.T.T., Goutard, F.L. and Hendrikx, P. (2019) One health surveillance: a matrix to evaluate multisectoral collaboration. *Frontiers in Veternary Sciences*, 6, 1–12. DOI: 10.3389/fvets.2019.00109

Bordier, M., Goutard, F.L., Antoine-Moussiaux, N., Pham-Duc, P., Lailler, R., et al. (2021) Engaging stakeholders in the design of one health surveillance systems: a participatory approach. *Frontiers in Veternary Sciences*, 8. DOI: 10.3389/fvets.2021.646458

Calba, C., Goutard, F.L., Hoinville, L., Hendrikx, P., Lindberg, A., et al. (2015) Surveillance systems evaluation: a systematic review of the existing approaches. *BMC Public Health*, 15, 448. DOI: 10.1186/s12889-015-1791-5

Donado-Godoy, P., Castellanos, R., León, M., Arevalo, A., Clavijo, V., et al. (2015) The establishment of the colombian integrated program for antimicrobial resistance surveillance (COIPARS): a pilot project on poultry farms, slaughterhouses and retail market. *Zoonoses and Public Health*, 62, 58–69. DOI: 10.1111/zph.12192

Eppler, M.J. (2003) A framework for information overload research in organizations insights from organization science. *Accounting, Marketing, MIS, and Related Disciplines*. 42 pp.

FAO (2020) FAO Assessment Tool for Laboratories and AMR Surveillance Systems (FAO-ATLASS). Available at: https://www.fao.org/antimicrobial-resistance/resources/tools/fao-atlass/en/ (accessed 15 March 2022).

FAO, OIE, WHO (2016) *Antimicrobial Resistance: A manual for developing National Action Plans, Version 1*. World Health Organization (WHO), Food and Agriculture Organization of the United Nations (FAO) and World Organisation for Animal Health (OIE), Geneva.

FAO, UNEP, WHO, and WOAH (2022) One Health Joint Plan of Action (2022–2026). Working together for the health of humans, animals, plants and the environment. Food and Agriculture Organization of the United Nations (FAO), Rome. DOI: 10.4060/cc2289en.

González-Zorn, B. and Escudero, J.A. (2012) Ecology of antimicrobial resistance: humans, animals, food and environment. *International Microbiology*, 15, 101–109. DOI: 10.2436/20.1501.01.163

Häsler, B., Taylor, N., Hardstaff, J., Bisdorff, B., Vergne, T., et al. (2014) Mapping of surveillance systems, animal populations, trade flows, critical infrastructure and decision-making processes in seven European countries. This is a video recording with ppt at: https://www.fp7-risksur.eu/progress/surveillance-symposium-2015

Hedman, H.D., Vasco, K.A., and Zhang, L. (2020) A review of antimicrobial resistance in poultry farming within low-resource settings. *Animals*, 10(8), 1264.

Hester, P.T. and Adams, K.M. (2017) *Systemic Decision Making. Topics in Safety, Risk, Reliability and Quality*, 2nd edn. Springer International Publishing, Cham. DOI: 10.1007/978-3-319-54672-8

Hitziger, M., Aragrande, M., Berezowski, J.A., Canali, M., Del Rio Vilas, V., et al. (2019) EVOLvINC: EValuating knOwLedge INtegration Capacity in multistakeholder governance. *Ecology and Society*, 24, art36. DOI: 10.5751/ES-10935-240236

Ibrahim, O.M. and Polk, R.E. (2014) Antimicrobial use metrics and benchmarking to improve stewardship outcomes: Methodology, opportunities, and challenges. *Infectious Disease Clinics of North America*, 28, 195–214. DOI: 10.1016/j.idc.2014.01.006

Kakkar, M., Sharma, A., and Vong, S. (2017) Developing a situation analysis tool to assess containment of antimicrobial resistance in South East Asia. *BMJ* 358, j3760. DOI: 10.1136/bmj.j3760

Downloaded from https://cabidigitallibrary.org by 128.106.115.228, on 11/08/22.
Subject to the CABI Digital Library Terms & Conditions, available at https://cabidigitallibrary.org/terms-and-conditions

Lélé, S. and Norgaard, R.B. (2005) Practicing Interdisciplinarity. *Bioscience*, 55, 967. DOI: 10.1641/0006-3568(2005)055[0967:PI]2.0.CO;2

Lindström, B. and Eriksson, M. (2006) Contextualizing salutogenesis and Antonovsky in public health development. *Health Promotion International*, 21, 238–244. DOI: 10.1093/heapro/dal016

Linn, M.C. (2005) The knowledge integration perspective on learning and instruction. In: *The Cambridge Handbook of the Learning Sciences*. Cambridge University Press, Cambridge, pp. 243–264. DOI: 10.1017/CBO9780511816833.016

Nancarrow, S., Booth, A., Ariss, S., Smith, T., Enderby, P., et al. (2013) Ten principles of good interdisciplinary team work. *Human Resources for Health*, 11, 19. DOI: 10.1186/1478-4491-11-19

Nielsen, L.R., Alban, L., Ellis-Iversen, J., Mintiens, K., Sandberg, M. (2020) Evaluating integrated surveillance of antimicrobial resistance: experiences from use of three evaluation tools. *Clinical Microbiology and Infection*, 26, 1606–1611. DOI: 10.1016/j.cmi.2020.03.015

Ogyu, A., Chan, O., Littmann, J., Pang, H.H., Lining, X., et al. (2020) National action to combat AMR: a One-Health approach to assess policy priorities in action plans. *BMJ Global Health*, 5, e002427. DOI: 10.1136/bmjgh-2020-002427

OHP Office (2014) The Viet Nam One Health Partnership for Zoonoses. Available at: https://onehealth.org.vn/

Peyre, M., Hoinville, L., Njoroge, J., Cameron, A., Traon, D., et al. (2019) The RISKSUR EVA tool (Survtool): a tool for the integrated evaluation of animal health surveillance systems. *Preventive Veterinary Medicine*, 173, 104777. DOI: 10.1016/j.prevetmed.2019.104777

Peyre, M., Roger, F., Goutard, F.L., Salman, M.D., Häsler, B., et al. (2022) *Principles for Evaluation of One Health Surveillance: The EVA Book*. Springer International Publishing, Cham. DOI: 10.1007/978-3-030-82727-4

Queenan, K., Häsler, B. and Rushton, J. (2016) A One Health approach to antimicrobial resistance surveillance: is there a business case for it? *International Journal of Antimicrobial Agents*, 48, 422–427. DOI: 10.1016/j.ijantimicag.2016.06.014

Rabinowitz, P.M., Kock, R., Kachani, M., Kunkel, R., Thomas, J., et al. (2013) Toward proof of concept of a One Health approach to disease prediction and control. *Emerging Infectious Diseases*, 19, e130265. DOI: 10.3201/eid1912.130265

Robinson, T.P., Bu, D.P., Carrique-Mas, J., Fèvre, E.M., Gilbert, M., et al. (2016) Antibiotic resistance is the quintessential One Health issue. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 110, 377–380.

Rüegg, S.R., Rosenbaum Nielsen, L., Buttigieg, S.C., Santa, M., Aragrande, M., et al. (2018) A systems approach to evaluate One Health initiatives. *Frontiers in Veternary Sciences*, 5, 1–18. DOI: 10.3389/fvets.2018.00023

Sandberg, M., Hesp, A., Aenishaenslin, C., Bordier, M., Bennani, H., et al. (2021) Assessment of evaluation tools for integrated surveillance of antimicrobial use and resistance based on selected case studies. *Frontiers in Veternary Sciences*, 8. DOI: 10.3389/fvets.2021.620998

Sargeant, J.M., Ramsingh, B., Wilkins, A., Travis, R.G., Gavrus, D., et al. (2007) Constraints to microbial food safety policy: Opinions from stakeholder groups along the farm to fork continuum. *Zoonoses and Public Health*, 54, 177–184. DOI: 10.1111/j.1863-2378.2007.01042.x

Schnall, J., Rajkhowa, A., Ikuta, K., Rao, P., and Moore, C.E. (2019) Surveillance and monitoring of antimicrobial resistance: limitations and lessons from the GRAM project. *BMC Medicine*, 17, 176. DOI: 10.1186/s12916-019-1412-8

Sharma, C., Rokana, N., Chandra, M., Singh, B.P., Gulhane, R.D., et al. (2018) Antimicrobial resistance: its surveillance, impact, and alternative management strategies in dairy animals. *Frontiers in Veterinary Sciences*, 4, 237.

WHO (2022) *Library of National Action Plans on Antimicrobial Resistance*. Available at: https://www.who.int/teams/surveillance-prevention-control-AMR/national-action-plan-monitoring-evaluation/library-of-national-action-plans (accessed 6 July 2022). World Health Organization (WHO), Geneva.

WHO (2015a) *68th World Health Assembly—Resolutions and Decisions*. World Health Organization (WHO), Geneva.

WHO (2015b) *Global Action Plan on Antimicrobial Resistance*. WHO Regional Office for South-East Asia, Geneva.

Williams, B. (2019) *Systemic Evaluation Design: A Workbook*, 2nd edn. Bob Williams, Wellington, New Zealand.

Wilson, J.B., Salman, M., Janzen, E., Sparagano, O., Speer, N., et al. (2020) Community network integration: an approach to alignment of One Health partners for solutions to 'Wicked' problems of antimicrobial resistance. *Preventive Veterinary Medicine*, 175, 104870. DOI: 10.1016/j.prevetmed.2019.104870

Downloaded from https://cabidigitallibrary.org by 128.106.115.228, on 11/08/22.
Subject to the CABI Digital Library Terms & Conditions, available at https://cabidigitallibrary.org/terms-and-conditions