

WHO VERBAL AUTOPSY REFERENCE GROUP VIRTUAL
WORKSHOP REPORT

RESPONSE PATTERN ANALYSIS FOR
THE REVISION OF THE 2016 WHO
VERBAL AUTOPSY INSTRUMENT

8,9 AND 15,16 JULY 2020



**World Health
Organization**

ACKNOWLEDGEMENTS

The realization of the virtual workshop would not have been possible without the collaboration and support of numerous organizations, institutions and individuals. The organizations contributing to this work include World Health Organization (WHO), Bloomberg Philanthropies Data for Health (D4H), US Centres for Disease Control and Prevention (CDC), London School of Hygiene & Tropical Medicine (LSHTM), Swiss Tropical and Public Health Institute, Ghana Health Service (GHS), Ohio State University, South African Medical Research Council, Morocco Ministry of Health, Ifakara Health Institute, the Aga Khan University, Senegal Ministry of Health and Social Action, University of Zambia, University of Zambia, the Kenya Medical Research Institution, Thailand Ministry of Public Health and University of the Basque Country.

Acknowledgment is given to the WHO Verbal Autopsy Reference Group. In particular, we would like to also acknowledge the comprehensive inputs of Edward Fottrell, University College London; Arantza Casillas, University of the Basque Country; Hermon Gebrehiwet, CDC; Jordana Leitão, WHO consultant; Richard Li, OpenVA; Erin Nichols, CDC; Kristen Pettrone, CDC; Clarissa Surek-Clark, The Ohio State University; Yue Chu, OpenVA; Brent Vickers, CDC; Samuel Clark, The Ohio State University; Alicia Perez, University of the Basque Country; Jason Thomas, OpenVA; Daniel Cobos, Swiss Tropical and Public Health Institute; Owen Trigueros, University of the Basque Country; Peter Choi, OpenVA and Tyler McCormick, University of Washington.

EXECUTIVE SUMMARY

The WHO Verbal Autopsy Reference Group (VARG), with support from the U.S. Centers for Disease Control and Prevention (CDC), the CDC Foundation, and the Bloomberg Philanthropies Data for Health Initiative, organized a technical virtual workshop on “Response pattern analysis with 2016 WHO verbal autopsy data”. The workshop was structured in two parts, with the first taking place July 8-9, and the second on July 15-16.

The virtual workshop brought together 78 participants, including the members of the VARG, officials of several LMIC that are using VA, experts in the field and key partners.

The current 2016 version of the WHO VA Tool was developed to be fully compatible with publicly available automated algorithms (Tariff 2.0 – Smart VA, InterVA, and InSilicoVA) and has been subjected to training, testing and extensive field use. Breaking down complex questions for clarity has resulted in a considerable number of questions, and users have requested a shorter questionnaire. However, that approach clarified the meaning of questions and enables the assessment of what detail can truly be reported in VA interviews.

A considerable amount of sites have conducted VA with the 2016 WHO VA questionnaire, allowing the assessment of the feasibility of questions.

The methodology and results of the analyses with collected 2016 WHO VA data were presented through quantitative and qualitative approaches for the item reduction of the WHO VA instrument. The agreed process involves the data-driven review of the feasibility of questions and expert-led assessment about the actions to be taken. It was agreed that additional data will be sought for inclusion for above analyses, either to be analysed centrally or in the relevant country, and the resulting update to the WHO VA instrument would be targeted to be finalized by December 2020.

For the assessment of the importance of particular items for the different target causes, it was highlighted that a very large reference death dataset with wide geographic, epidemiological and historical representation would need to be compiled. The preliminary results also highlighted the importance of triangulating different methods to enable robust assessments over whether a VA item should be kept, removed or changed to address any clarity or redundancy issue(s).

The workshop also informed about the questions added for COVID-19. Countries that had integrated VA as part of their routine ascertainment of causes of death had started to formulate additional questions by themselves, or were asking WHO VARG for advice. It was decided to add specific questions in order to standardize the approach and align with rapid mortality surveillance.

TABLE OF CONTENTS

Acknowledgements	2
Executive Summary	3
Table of Contents	4
Background	5
Objectives and Expected Outcomes.....	6
Proceedings.....	6
1. Quantitative item response pattern results over full dataset	7
Discussion points and outcomes	7
2. Quantitative and qualitative (mixed-methods) results for known VA issues	9
Quantitative analysis methods.....	9
Qualitative analysis – Cognitive interviewing methodology.....	10
An Inter-cultural approach to VA implementation.....	11
Discussion points and outcomes	11
3. Targeted maximum likelihood estimation (TMLE)	13
Discussion points and outcomes	14
4. Entropy coefficient and VA narrative contributions.....	15
Discussion points and outcomes	16
5. Triangulation of response pattern and item significance analysis – openVA team..	18
Discussion points and outcomes	19
6. Implementation feedback	20
7. Data sharing	21
8. VA in the context of COVID-19	22
Summary, recommendations and next steps.....	24
Appendix A – List of participants.....	26
Appendix B – Agenda.....	33
Appendix C – WHO VARG maintenance process of the WHO VA instrument.....	35
Appendix D – List of proposed questions for the identification of possible COVID-19 deaths.....	36

BACKGROUND

Reliable and comparable data on the levels and causes of mortality are cornerstones for building a solid evidence base for health policy, planning, monitoring and evaluation. Verbal Autopsy (VA) is an important strategy for addressing the gaps in population-level data on cause-specific mortality, in the absence of medical certification of the causes of death. With VA, a structured interview is conducted with those best informed on the circumstances preceding death. The outcomes of such an interview are then analysed in a standard fashion to establish the cause of death.

Since the 1970s, the WHO has been developing continuously improved VA instruments. Research has seen special adaptations and new developments over time. A systematic analysis of causes identifiable with VA, and of questions relevant to identify these causes led to the development of the 2012 WHO VA instrument. The goal was also to have a VA instrument that was feasible in routine environments, parsimonious, and where the answers can be analysed by automated analytical software for assigning causes of death. The current 2016 version was developed to be fully compatible with publicly available automated algorithms (Tariff 2.0 – Smart VA, InterVA, and InSilicoVA) and has been subjected to training, testing and extensive field use.

Since the release of the 2016 version, issues reported with its use have been compiled (e.g. skip patterns, unreliable questions) and a major revision of the instrument planned for 2020 based on users' feedback and evidence from the field. The following encapsulate the rationale for the revision of the WHO VA instrument:

- Shorter and more practical instrument to further facilitate routine applications
- Improve collection of adequate data
- Increase acceptability of VA by respondents and communities
- Enhance the value and specificity of individual questionnaire indicators
- Improve the validity and utility of the process

The WHO VA Reference Group (VARG) developed a protocol to revise the WHO VA instrument and generate a questionnaire that is as short, concise and efficient as possible, and that works well in the field with currently available algorithms and physician-certified VA (PCVA). Overall, the proposed procedure first requires VA interview data representative of community deaths in LMICS to identify what questions in the instrument are getting reliable responses and to resolve known issues with the 2016 WHO VA instrument. At a subsequent phase, with a high-quality repository containing VA survey data with independently assigned causes of death, questions considered irrelevant for cause assignment are identified based on a ranking of symptoms by importance.

The WHO VARG, with support from the U.S. Centers for Disease Control and Prevention (CDC) and CDC Foundation under the Bloomberg Data for Health Initiative, organized a technical virtual workshop on “Response pattern analysis with 2016 WHO verbal autopsy data”. The workshop was structured in two parts, with the first taking place July 8-9, and the second on July 15-16.

The virtual workshop brought together an average of 78 participants across its four day duration, including members of the VARG, country officials, experts in the field and representatives of key partners. A detailed list of participants can be found in Appendix A.

OBJECTIVES AND EXPECTED OUTCOMES

The workshop’s key objectives were to:

1. Run item by item frequency distribution to identify unusual response patterns (i.e. lack of variability in responses);
2. Investigate response patterns for enumerated problematic items and issues with the 2016 WHO VA instrument; and
3. Determine and consolidate remaining steps to complete revision of 2016 WHO VA instrument.

Expected outcomes from the workshop, included:

- Item by item frequency distribution tables from datasets with 2016 WHO VA instrument;
- Enumeration and categorization of issues with the 2016 WHO VA instrument;
- Draft recommendations for the resolution of identified issues with the WHO VA instrument; and
- Proposed protocol for item reduction with demonstration reference dataset to evaluate significance of VA items.

PROCEEDINGS

The workshop was chaired by Daniel Chandramohan and the agenda (Appendix B) covered 4 days (8,9 and 14,15 July) with a continuous working program starting from 12:00 pm until 4:00 pm (UTC). The following report sections are structured around the key sessions of the workshop.

1. QUANTITATIVE ITEM RESPONSE PATTERN RESULTS OVER FULL DATASET

The main objective of the item response pattern analysis was to summarize the “missing information” across collected VA data with the 2016 WHO VA instrument, in the shape of:

- Item non-response (no value included in the data)
- “Don’t know” responses
- “Refused to answer” responses

The analysis was carried out on 5 datasets, totalling 20,276 VAs from:

- CHAMPS (N = 1,313)
- COMSA (N = 6,600)
- Kenya (N = 4,230)
- South Africa (N = 5,387)
- Zambia (N = 2,746)

The process of merging the different datasets involved the following steps:

- Data cleaning
- Harmonizing the column names
- Harmonizing the values (e.g. yes -> y, no -> n)
- Finding the right age columns
- Checking for missing columns/values
- Fitting into WHO VA instrument format

A key challenge for the analysis was having to account for the various skip patterns embedded within the WHO VA instrument (ODK XLSForm), as items are arranged in hierarchical groups (groups of questions nested within groups) - which varies between versions and adaptations by users. By default, the openVA assumed that all datasets used v1.5.1 of the 2016 WHO VA instrument.

DISCUSSION POINTS AND OUTCOMES

The response pattern analysis showed that there are very few refusals to VA questions in the data; that 90% of questions have fewer than 13% of “don’t know” responses (max is 62%); and several Yes/No questions that almost always take a single value.

Other key findings highlighted and discussed included:

- There could be deviations in the forms used by teams from the standard v1.5.1 which could be leading to some problems with the current response distribution – useful to give closer look at some items.
- The openVA team attempted to follow skip patterns, but some are very complex and could explain unusual missing information.
- The data preparation stage may influence % missing (frequency of missing values) and % asked (the relative number of times questions are asked)
- Even with a dataset of 20,276 deaths some questions were asked too infrequently; many missing (e.g. maternal section).

The slide with the key results¹ from the response pattern analysis includes:

- List of VA items with the 10 largest values for % refused to answer
- List of VA items with the 10 largest values for % of don't know
- List of VA items with the 10 largest values for % of missing response
- List of VA items with the 11-20 largest values for % of missing responses
- List of VA items with the 21-30 largest values for % of missing responses

¹ Available at: <https://www.dropbox.com/sh/c6lo76d6g0ggjnr/AADU3IfHRuVT5SDKCkiOVe26a?dl=0>

2. QUANTITATIVE AND QUALITATIVE (MIXED-METHODS) RESULTS FOR KNOWN VA ISSUES

Since the publication of the 2016 WHO VA instrument, reported issues and feedback by users have been compiled and managed by the VARG through the Public GitHub platform: https://github.com/SwissTPH/WHO_VA_2016

From the VARG maintenance process of the WHO VA instrument (see Appendix C), 14 issues were identified and assigned for mixed-methods review:

1. Tobacco use
2. Swallowing
3. Sores and ulcers
4. Swelling, lump, ulcers, pits in the breast
5. Other female health related questions
6. Diagnosis by a health professional vs symptom report
7. Vaccinations
8. Injury questions
9. Urine
10. Abdominal problem
11. Lumps
12. Vomiting
13. Violence
14. Baby size

The kinds of problems these 14 issues represent can be categorized into the following:

- Item sequence
- Redundancy/indicator overlap
 - Indicators eliciting similar information
 - Identify and clarify key constructs of interest and utilize minimal set of questions to elicit this information
 - Achieved with shortened sequence
- Frame of reference
- Clarity
 - Confusion between constructs
 - Respondent not able to accurately differentiate between the constructs
 - Constructs not clearly phrased or understood
 - Lack of understanding of the conditions or medical terminology

QUANTITATIVE ANALYSIS METHODS

For the quantitative side of the analysis, frequency tables, cross tabulations and prevalence ratios were calculated over two sources of datasets with the 2016 WHO VA instrument. Key characteristics of the datasets used are listed below.

1. **WHO aggregated dataset (N= 19,150 VAs)** – 5 VA teams with data-sharing agreement for confidentiality, limited use

- Child Health Mortality Prevention Surveillance (CHAMPS)
- COMSA
- Kenya
- South Africa
- Zambia

Age distribution:

- 13,736 adults
- 2,916 children
- 2,498 neonates

Gender distribution:

- 10,280 males
- 8,861 females

2. **South Africa Cause of Death Validation Study utilizing PCVA for cause of death assignment (N=5,387 VAs)**

Age distribution:

- 102 neonates
- 187 children
- 5100 adults

Gender distribution:

- 2,579 females
- 2,808 males

QUALITATIVE ANALYSIS – COGNITIVE INTERVIEWING METHODOLOGY

For the qualitative analysis, cognitive interviewing, used normally for question evaluation research, was used to provide context and to assess the validity of constructs captured (vs as intended). Through semi-structure interviews, the method identifies: constructs captured by question; specific phenomena that account for respondents' answers (Yes/No); and patterns across groups. Table 1 shows the three difference sources of data used for the qualitative analysis.

Table 1 – Data sources for cognitive testing of 2016 WHO VA instrument

	Zambia	Morocco
Locations	Lusaka	Rabat
Time period	2019	2019
Sites	City hospital (morgue)	Ministry of the Interior's Vital Registration Office
Interviewers	8	12

Interview period	4 months	12 months
Language	Nyanja	French and Arabic
Total interviews	65	84

The approach used in the different sites can be synthesized into the following:

- Cognitive testing conducted immediately following VA interview at same location;
- All “Proxy respondents” from one of three age categories (neonate, child, adult)
- Testing was performing for a list of specific questions from VA instrument (i.e. not all questions in the WHO VA instrument tested)
- Retrospective probing (e.g. “When you were asked this question, what were you thinking about?”)
- Open/unstructured probing

AN INTER-CULTURAL APPROACH TO VA IMPLEMENTATION

As VA interviews are rarely conducted in English, another component of the revision process of the 2016 WHO VA instrument will involve looking at translations and trying to understand how any differences on intended constructs and other related issues might be impacting performance.

DISCUSSION POINTS AND OUTCOMES

During the workshop’s course, participants were divided into 6 breakout groups that carefully reviewed and discussed the mixed methods results ² - and drafted recommendations on how best the issues could be resolved. The outcomes were also presented by each group and discussed in plenary.

For most of the issues, groups considered that the question or question series was lacking clarity. In terms of completeness of data provided for assessment of the issue and recommendation of solution, 36.4% considered that the following missing aspects were key for decision-making on the issue: sex of the respondent; how algorithms are using the question; impact of removing/retaining on COD assignment (incl. for PCVA); sub-question response patterns and some cognitive testing aspects.

Some of the changes recommended by the breakout groups to resolve the issues, included: to remove part of question series; to combine or to drop questions; to provide more clarification (e.g. hints, definition, QbyQ); interviewer training; to add a screening question.

² <https://drive.google.com/file/d/1etatSCcinoUA7qFJbdnjlJv8KxUZaoQ/view>

Other comments made by the groups included the influence of respondent selection on the quality of the interview and the type of information accessible; considerations for how algorithms would accommodate changes; and how more time to review carefully the issues and to deliberate on answers would have been beneficial.

Besides the specific recommendations for the 14 VA issues, the following key conclusions were drawn:

- Mixed-methods analysis with quantitative and qualitative results is available and can help understand why questions are underperforming.
- Analysis was completed for the 14 previously-enumerated issues; and important feedback has been compiled from the participants that requires further and additional consideration before reaching a final decision on the outcomes of the issues.
- Triangulation with mixed-methods results should be done with items of concern that are flagged from the quantitative analyses (item response patterns and item significance).

3. TARGETED MAXIMUM LIKELIHOOD ESTIMATION (TMLE)

In the context of item reduction, the OpenVA team developed the Targeted Maximum Likelihood Estimation (TMLE) method, in order to:

- Quantify the importance of each VA symptom for identifying each cause of death;
- Conduct this analysis in a general, algorithm-agnostic way so that the results do not depend on the assumptions, symptom-cause information, or logic of a specific algorithm; and
- Rank the symptoms by importance.

The method requires reference deaths that have both VA symptoms and a cause assigned through a trusted method and not assigned using a VA cause-coding algorithm. Additionally, it is essential that: **(a)** there are enough reference deaths to have observations in all cells defined by combinations of specific causes and symptoms; and **(b)** the reference deaths come from a wide enough variety of populations that represent all the epidemiological conditions and historical periods that are of interest to VA users. For detailed description of the analytical method to quantify item importance, see the resources available online at:

<https://www.dropbox.com/sh/c6lo76d6g0ggjnr/AADU3IfHRuVT5SDKckjOVe26a?dl=0>.

To the workshop's date, the VARG did not have access to sufficient reference deaths for the adequate and accurate application of the TMLE, hence the results of the analysis presented are only from two datasets (CHAMPS and South Africa validation study – Table 2) for demonstration purposes only (i.e. being of incomplete and of limited generalizability).

Table 2 – Available reference deaths for TMLE demonstration. Datasets: CHAMPS – child deaths; South Africa VA validation study by the South African Medical Research Council (MRC) – mostly adult deaths.

WHO cause categories , excluding neonatal and external causes - 19 causes (with at least 59 observations)	3,859 observations
Aggregate WHO cause categories , excluding neonatal and external causes - 8 causes	4,001 observations
Neonatal causes – 6 causes	1,034 observations

DISCUSSION POINTS AND OUTCOMES

Key outputs from the TMLE analysis³:

- Cause distributions using WHO all cause categories (except neonatal and external causes), aggregated cause categories and neonatal causes.
- TMLE heat maps for all maps and causes - three lists for all causes, aggregate causes and neonatal causes.

As the TMLE is an empirical, data-driven approach and is therefore entirely dependent on the reference death dataset – it requires enough reference deaths to have sufficient observations for all the questions for all causes in the 2016 WHO VA instrument; and as much variability as possible in its epidemiological and historical representation. The limitations of the currently available “small” reference death dataset for demonstration purposes, was visible in the grey boxes/zones of the heat maps presented.

Having more reference deaths in number and variety would likely change and improve clarity over the results– however, the take away message from the demonstrated analysis - that many symptoms are not that important for cause assignment is still valid. The initial analysis indicated potentially around 25% of the questions as candidates for removal. The TMLE will be one of the components/processes involved in the revision process – the revision process of the WHO VA instrument will centre on triangulation of different methods, including structured discussions with physicians.

Other key important aspects to be considered for subsequent TMLE analysis include:

- Most causes rely on key combinations of symptoms and a sole data driven approach is not feasible for the identification of interdependency between symptoms. For this purpose, physicians will need to identify important combinations of symptoms to then apply with the TMLE. HIV was proposed as case study to test the effect on importance with symptoms’ combinations.
- Important to crosscheck with the distribution of causes on the reference death dataset to be able to pinpoint cases where a low importance score can be due to insufficient cases in the reference dataset (e.g. association between unable to open mouth, that scored poorly, and tetanus).
- Future analysis should highlight/distinguish the root from the follow up questions.

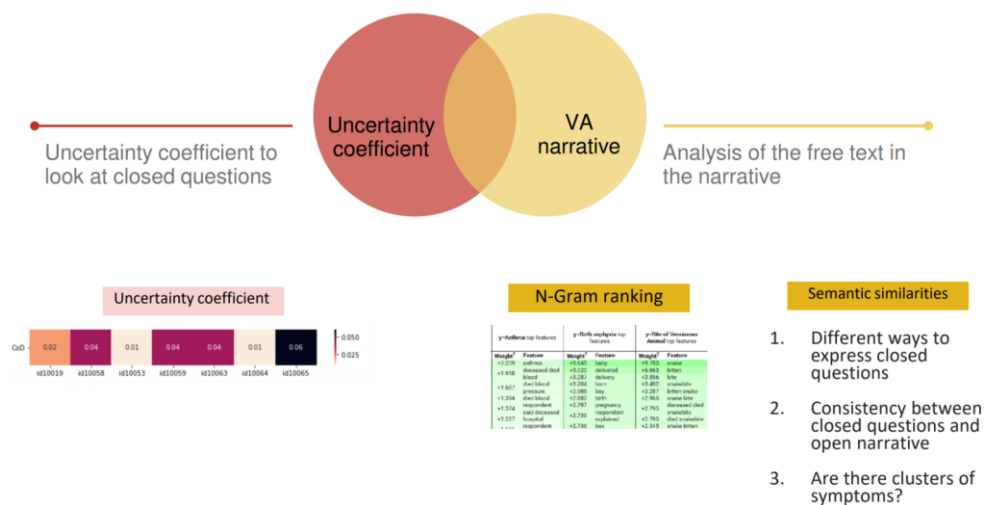
³ Results available at:

<https://www.dropbox.com/sh/c6lo76d6g0ggjnr/AADU3IfHRuVT5SDKCkjOVe26a?dl=0>

4. ENTROPY COEFFICIENT AND VA NARRATIVE CONTRIBUTIONS

On the basis that most information in the VA narrative is largely missed by algorithms, the Entropy coefficient and Natural language processing (NLP) are proposed to identify variables in the 2016 WHO VA questionnaire that add little value to the ascertainment of the cause of death using VA narratives. The approach focuses on the most important variables instead of the less important; and complements and triangulates the results of the other analyses.

Figure 1 – Uncertainty coefficient and VA narrative contributions (N-Gram ranking and Semantic similarities).



Detailed descriptions of the analytical methods can be found in X. The approach can be succinctly delineated as the following:

Find terms and questions that are associated with a cause of death via:

1. Correlation between cause and closed questions
 - How informative is a closed question with respect to a given cause?
 - Approach: measure the correlation between closed questions and cause of death by means of uncertainty coefficient
2. Open response n-gram ranking wrt the cause
 - Which are the most relevant terms for each cause of death?
 - Linear classifiers assign a weight to every feature
 - We explore the usage of n-grams in the narrative as features in order to: predict the cause of death; and rank the weights that the model gives to them
3. Identification of relevant words wrt the cause
 - Explainable Artificial Intelligence (XAI)
 - Usage of attention mechanisms to mark the text fragments that motivated the prediction of a cause of death

- Gain in model interpretability

The models require reference deaths (incl. narratives and closed questions), and the process with VA teams essentially comprises the following steps:

1. Get the necessary agreements and approvals
2. Review data management plan and create a platform to transfer the data
3. Run the analysis
4. Debrief with country team and review the results
5. Share the results once agreed

For demonstration purposes, results of the models with the PHMRC corpus adult data (a) experimental results for logistic regression and BiGru models; (b) uncertainty coefficient in CHAMPS and South Africa – are available at:

<https://www.dropbox.com/sh/c6lo76d6g0ggjnr/AADU3IfHRuVT5SDKCkjOVe26a?dl=0>

In the presented analysis, the analytical models were trained in three ways: (1) open responses only; (2) closed questions only; and (3) open responses + closed questions.

DISCUSSION POINTS AND OUTCOMES

Methods that automatically analyse the VA narrative (e.g. machine learning and natural language processing) can be used to: (a) improve the ascertainment of the cause of death; and (b) to identify relevant questions in the WHO VA questionnaire – helping to inform decisions on item removal, retention, adaptation or addition.

The demonstration with the PHMRC dataset showcased that even with narratives of limited quality, containing only few sentences in most cases – there was improvement in the performance compared to closed questions in isolation.

In the South African MRC VA validation study, in the VA interviews the narrative is asked first, before the closed questions and the approach was reported as facilitating rapport with respondents. The team identified in the narrative information not contained in the WHO VA instrument's closed questions – especially relevant for HIV and TB cases (e.g. respondents facilitated information that deceased had discontinued treatment).

Narratives depend on the quality of the interviewers which is chiefly influenced by the interviewers and the quality of training. Important to account for the inherent variability in the process with interviewers, that is also brought in from the objectives of implementation that affect training (e.g. all-cause mortality vs focus on maternal deaths). Additionally, VA interviews and the narratives are in most cases not conducted in English but are reported in English. For it to be of significant value for algorithms – need verbatim level of interpretation of narratives. Current technology allows such

applications and the presented analytical models are not dependent on the language of the narrative – narratives in any language can be used in the analysis.

The importance of the narrative was not questioned, and the presented analytical models were found useful to contribute to the revision of the WHO VA instrument. It will be particularly interesting to compare the measures of importance for cause assignment coming from the two different sources of information within VA interviews – the closed questions vs the narrative.

Additionally, as a future research interest the potential combined use of natural processing and machine learning and use of the narrative at the start of VA interviews were discussed as potential ways to shorten the interview. Conceptually, from the account in the narrative an embed algorithm/model could select related closed-questions to cross-check and complement information already captured in the narrative.

5. TRIANGULATION OF RESPONSE PATTERN AND ITEM SIGNIFICANCE ANALYSIS – OPENVA TEAM

This session focused on a potential strategy for combining the use of the quantitative methods proposed by the OpenVA team for item selection. The first phase in the process is preparation of the reference death dataset. The following gives an overview of the steps involved in data preparation:

1. Separate VA questionnaire items into groups
 - 1st level and root questions – questions not dependent on any other question
 - First child questions – direct descendants of root questions
 - Age groups (including all) and maternal questions
2. Calculate by each group:
 - **Response rate** of each question: proportion of response don't know, refused to answer, and missing responses
 - **Measure of variability** using entropy value⁴: a larger value means more information/variability/diversity in the distribution of responses.
 - **Importance measure**: as the largest effect size of the associated InterVA indicator in the TMLE analysis across all causes
 - Relevance of an item to differentiating each of the causes compared to all of the others. Get a score for each cause and look at the maximum score across all causes.

Combining the three metrics

On the following phase, the numerical scores for each metric are turned into percentiles ranging from 0 to 100 - smaller values associated with larger response/variability/importance. Next, percentiles are turned into tertiles and each VA item is labelled on each dimension with either 0 (first tertile – least informative), 1 (second tertile), or 2 (third tertile – most informative). Lastly, the 3 tertile scores are combined to create the final scores:

- **'Candidates for dropping'**: At least two 0's or (0, 1, NA).
- **'Could be useful but require justification'**: At least two 1's.
- **'Probably want to keep'**: At least two 2's.
- **'Hard to judge'**: The rest of the combinations - score values that include all of 0, 1, and 2

⁴ For details into how to calculate the entropy value see [x](#).

Draft results

As previously noted, the insufficient number of reference deaths restricts the analysis to demonstration purposes and demands caution in the interpretation of results. Due to small sample sizes, the importance scores could not be calculated for 175 VA items (i.e. excluded from the analysis).

Table 3- Summary of draft results from triangulation of response pattern and item significance

	Category	Frequency
	Candidates for dropping	132
	Could be useful but require justification	103
	Probably want to keep	98
	Hard to judge	122

Symptom rankings were shown for: all symptoms; 1st level questions; root questions; adult questions; neonatal questions; and maternal questions. Whereas, for each VA item, a summary was given of all the quantitative scores and what group the question is finally categorized into. These preliminary results are available at:

<https://www.dropbox.com/sh/c6lo76d6g0gqjnr/AADU3IfHRuVT5SDKCkjOVe26a?dl=0>

The same six breakout groups and plenary discussion were used to:

1. Review quantitative analysis – item response pattern and item significance;
2. Review mixed methods findings – identify common issues; and
3. Incorporate group feedback.

DISCUSSION POINTS AND OUTCOMES

The process and outcomes of the discussions led by the breakout groups can be found at:

<https://www.dropbox.com/sh/c6lo76d6g0gqjnr/AADU3IfHRuVT5SDKCkjOVe26a?dl=0>

The session provided a first step into understanding what the method is capable of providing, and again brought focus to the very large amounts of reference data required for it to be truly informative for effective item reduction. Exercise also highlighted the importance of using different methods and types of analysis when trying to shed light into complex VA issues, associations between symptoms and causes and symptom interdependency.

6. IMPLEMENTATION FEEDBACK

For the revision of the instrument, a listing of teams that were understood to have at least 1000s VAs carried out with the 2016 WHO VA instrument was compiled. These teams shared their present experience working with the instrument and any expectations or concerns regarding the revision process of the WHO VA instrument.

Countries that were available to provide feedback on their experiences with the 2016 WHO VA instrument included: Mozambique; Kenya - KEMRI; Zambia and Morocco.

Mozambique - COMSA

- Started with VA - once satisfied with implementation, social autopsy was added, integrated with COMSA the questionnaire - VASA social autopsy integrated interview.
- Government wants to maintain the social autopsy and VA.
- Concern about the length of interview. But don't know the source: VA or social autopsy or both.

Kenya - KEMRI

- Raised issues with the VA instrument that were fixed with the updates to the 2016 WHO VA instrument.
- Interviewers had to be retrained to better record the narrative.
- Challenges answering immunization questions.
- The question on medical diagnosis of dengue- not common in Kenya's context.
- Still to see if and how the issues in the maternal section have been addressed with the updated version of the 2016 WHO VA instrument (v1.5.2).

Zambia

- Translation into 3 languages – potential for more.

Ghana:

- Long story of research-based VA.
- With support from D4H initiated community-based VA.
Community health workers trained as VA interviewers in 2017.

7. DATA SHARING

While the analysis presented during the virtual works is informative, the results will be more informative if the dataset is more representative; countries were therefore invited to contribute available data to the exercise. For data sharing, data sets need to be deidentified and anonymised – data should not contain any personal information. Specific sets of meta data for analysis purposes are required to be able to compare between data sets and identify specific issues or problems within the data sets. For data sharing, the first step would be to send an inquiry to the verbal autopsy WHO email address (verbalautopsy@who.int). The VARG will then reach back to provide access to a secure portal on the WHO Sharepoint for the datasets. The relevant links will be made available.

8. VA IN THE CONTEXT OF COVID-19

Countries that had integrated VA as part of their routine ascertainment of causes of death had started to formulate additional questions by themselves, or were asking WHO VARG for advice. It was decided to add specific questions in order to standardize the approach and align with rapid mortality surveillance.

Based on core symptom patterns emerging from exiting evidence (that continues to involve), the VARG is adding a few questions (n=7; 6 root questions and 1 follow up question) as an addendum to the current 2016 v.1.52. Of the most common symptoms of COVID-19 infection, the 2016 v1.5.2 questionnaire captures information on the following: fever, cough, shortness of breath, and headache. There are no questions that address loss or change of smell/taste, myalgia (or muscle aching) or fatigue. The set of questions added to the 2016 WHO VA instrument v1.5.3 (soon to be published on the WHO VA standards webpage) to enable the detection of probable COVID-19 deaths is shown in Appendix D. Besides the release of an updated version of the electronic VA instrument, the publication will also include updates to the COD list, included in the WHO VA standards manual, and to the QbyQ (i.e. guidance for the newly introduced questions).

Considerations for use of COVID-19 questions in VA:

- Research, field experience and testing required before adding questions to standard VA questionnaire
- Potential differences in symptom experience across disease severity and population factors

Until the automated algorithms are updated to include COVID-19, PCVA is the recommended method of assigning the cause of death for probable COVID-19 deaths from VA data. Updating the algorithms, requires evidence and validation to define the symptom-cause relationship (work underway to collect cases for InterVA and InSilicoVA updates).

The following key points were raised regarding the use of VA for COVID-19 epidemic awareness:

- Consider information goals and resources available to inform appropriate selection of questionnaire and cause of death interpretation methods
- Rapidly detect deaths possibly associated with COVID-19?
- Screening deaths for further diagnostic testing or clinical review?
- Substitute standard processes where capacity is strained?
- Maintaining routine, standard VA practices for statistical purposes?
- Measurement of excess mortality (all-cause and cause-specific)
- Screen for post-mortem testing
- Circumstances Of Mortality CATegories (COMCATs from InterVA-5)
- Further guidance will be compiled as experience expands

Implementation considerations

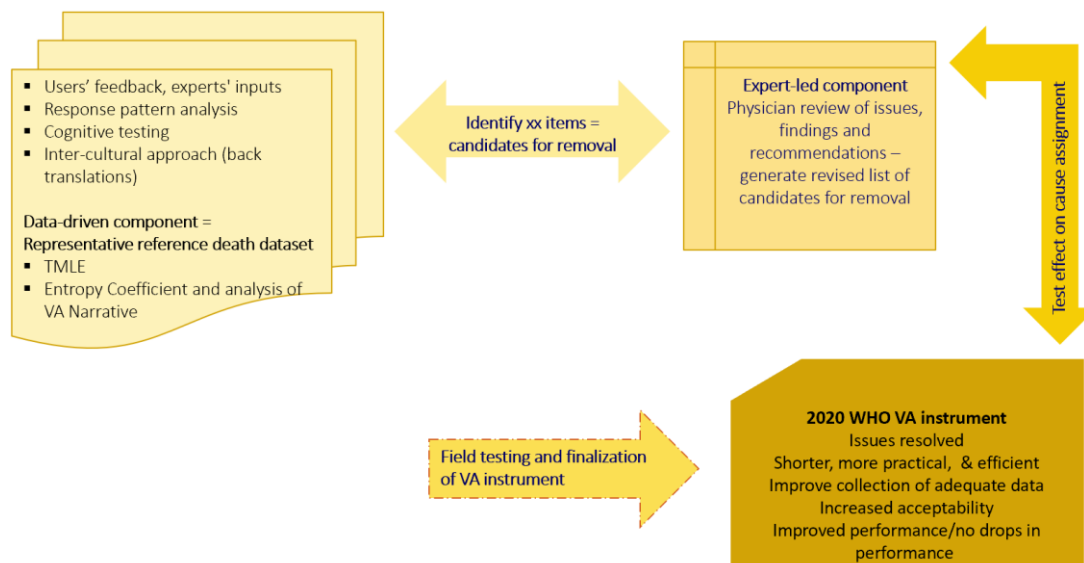
The COVID-19 pandemic demands specific implementation considerations. In terms of safety – its essential to protect field workers, trainees with appropriate infection prevention and control measures (see RMS guide). The impact on field activities and routine death management systems could be reflected in the availability and accessibility of respondents, and disruptions to normal community worker activities. Possible solutions include:

- Continue tracking deaths, but following up later for VA;
- Use of mobile communication; and
- New death verification processes as opportunity to collect cause of death information.

SUMMARY, RECOMMENDATIONS AND NEXT STEPS

Through the course of the workshop, the methodologies and respective full or preliminary results were presented for all the quantitative and qualitative approaches planned for the revision of the 2016 WHO VA instrument. The final outcomes of the workshop were an agreed process (Figure 2) involving data-driven and expert-led components and a timeline (Figure 3) for the revision of the instrument.

Figure 2 – Consensus protocol for the revision of the 2016 WHO VA instrument



The presentation of the quantitative analytical methods proposed to measure VA item importance, using two different sources within VA data (closed questions vs. narratives) – highlighted their dependency on a very large reference death dataset with wide geographic, epidemiological and historical representation. Looking at the preliminary results also highlighted the importance of triangulating different methods to enable robust assessments over whether a VA item should be kept, removed or changed to address any clarity or redundancy issue(s). Furthermore as depicted in Figure 2, VA items identified as candidates for removal will not be effectively removed from the WHO VA instrument before: (a) their cause of death assignment importance is cross-checked by physicians with VA experience; (b) the effect of removal is tested on algorithms; and (c) the revised simplified and improved instrument is field tested to ensure there has been no drops in the performance of the instrument.

Figure 3 – Timeline for the revision of the 2016 WHO VA instrument



Per the agreed timeline, the data sharing agreement with collaborators should be completed and the quantitative analysis of individually anonymized data sets available at a central repository by September 2020. Alternatively, the data-driven analysis is performed in-country with support from the OpenVA team and results shared with the WHO VARG. The option of conducting the analysis in-country, incurs longer times in the revision process and the non-trivial analysis is likely to pose other challenges. In case insufficient reference deaths are not obtained by September 2020, the revision process will be carried on as specified in the timeline - relying on the triangulation of the other methods to resolve the reported issues with the 2016 WHO VA instrument and produce a revised version of the instrument by December 2020. Subsequently, when a sufficient number of the reference deaths becomes available, the data-driven stage of the revision process can resume to better infer on the importance of VA items and conclude the item reduction process.

To address the issues of priority with the WHO VA instrument and of feasibility over the large reference dataset - a multi-stage and stepwise approach will allow the revision of the instrument by the end of 2020. The timeline represents a tight schedule but also represents the momentum on cooperation and the expressed needs for an improved and simplified VA instrument.

APPENDIX A – LIST OF PARTICIPANTS

Names of those who attended are listed in bold font.

VARG members

Alan Lopez

Director, Global Burden of Disease Group
Director, Bloomberg Initiative for Civil
Registration and Vital Statistics
Melbourne School of Population and
Global Health
The University of Melbourne, Level 5, 207
Bouverie St, Carlton 3053 VIC, Australia
alan.lopez@unimelb.edu.au

Arvind Pandey

Ex-Director, National Institute of Medical
Statistics
ICMR, Ansari Nagar, New Delhi 110 029
INDIA
arvindp.nims@gmail.com

Aurelio Di Pasquale

Software Project Coordinator
Health Systems Research and Dynamical
Modelling Unit
Dept. Epidemiology & Public Health
Swiss Tropical and Public Health Institute
P.O. Box
4002 Basel, Switzerland
aurelio.dipasquale@swisstph.ch

Azza Badr

Civil Registration and Vital Statistics
WHO HQ
20 Av. Appia
1211 Geneva 27, Switzerland
badra@who.int

Carla AbouZahr

Data for Health Initiative
Independent Consultant
Geneva, Switzerland
abouzahr.carla@gmail.com

Carine Alsokhn

Technical Officer
Classifications and Terminologies Unit
Division of Data, Analytics and Delivery for
Impact
WHO HQ
20 Av. Appia
1211 Geneva 27, Switzerland
alsokhnc@who.int

Chalapati Rao

Associate Professor
Department of Global Health, Research
School of Population Health
Australian National University
Canberra, Australia
chalapati.rao@anu.edu.au

Chomba Mwango

D4H
cmwango@gmail.com

Daniel Chandramohan

Professor of Public Health
Department of Infectious and Tropical
Diseases
London School of Hygiene & Tropical
Medicine
Keppel Street
London WC1E 7HT, UK
Daniel.Chandramohan@lshtm.ac.uk

Daniel Cobos

Senior Scientific Collaborator
Swiss Tropical and Public Health Institute
P.O. Box
4002 Basel, Switzerland
daniel.cobos@swisstph.ch

Don de Savigny

Professor
Health Systems and Policy
Swiss Tropical and Public Health Institute
P.O. Box
4002 Basel, Switzerland
D.deSavigny@swisstph.ch

Doris MaFat

Statistician
Health Statistics and Evidence,
Measurement and Health Information
Systems
WHO
20 Av. Appia, 1211 Geneva 27, Switzerland
mafad@who.int

Edward Fottrell

Senior Lecturer
Institute for Global Health
University College London
30 Guilford Street
London WC1N 1EH, United Kingdom
e.fottrell@ucl.ac.uk

Erin Nichols

International Statistics Program, NCHS
Centers for Disease Control and
Prevention
3311 Toledo Rd., MS P08
Hyattsville, MD 20782, USA
igd1@cdc.gov

Henry Kalter

Associate Professor
Johns Hopkins Bloomberg School of Public
Health
hkalter@jhsph.edu

Jordana Leitão

WHO Consultant
Angola/Portugal
leitaojordana@gmail.com

Khem R.Sharma

Assistant Professor,
School of Public Health & Community
Medicine,
BP Koirala Institute of Health Sciences,
Dharan
Nepal
khemraja@gmail.com

Lalit Dandona

Distinguished Research Professor
Director, Wellcome Trust Capacity Building
Programme
Public Health Foundation of India
Plot 47, Sector 44, Gurgaon – 122002,
National Capital Region, India
lalit.dandona@phfi.org

Luciana Tannos

Montpellier WHO CC
luciana.tanno@gmail.com

Patricia Soliz

WDC
PAHO
solizpat@paho.org

Peter Byass

Professor of Global Health
Director Umea Centre for Global Health
Umea, Sweden
peter.byass@epiph.umu.se

Philip Setel

Vital Strategies
Civil Registration and Vital Statistics
Improvement Program
Seattle, WA, USA
PSetel@vitalstrategies.org

Martin Bratschi

Vital Strategies
Civil Registration and Vital Statistics
Improvement Program
146 Robinson Road, #06-01, Singapore
068909, Singapore
mbratschi@vitalstrategies.org

Rakesh Rastogi

Technical Officer
Health System Development Department
South-East Asia Regional Office - SEARO
World Health Organization - WHO
rastogir@who.int

Riley Hazard

Melbourne School of Population and
Global Health
The University of Melbourne, Level 5, 207
Bouverie St, Carlton 3053 VIC, Australia
riley.hazard@unimelb.edu.au

Robert Jakob

Team Leader Classifications (ICD, ICF, ICHI,
VA)
Information, Evidence and Research
WHO HQ
20 Av. Appia
1211 Geneva 27, Switzerland
jakobr@who.int

Robert Mswia

Senior Technical Advisor, Verbal autopsy
Vital Strategies
New York, NY, USA
rmswia@vitalstrategies.org

Samuel Cheburet

Senior Health Information Manager
Ministry of Health
Afya House, Cathedral Road,
P.O. Box:30016–00100, Nairobi, Kenya
samuelcheburet@gmail.com

Samuel Clark

Professor
Department of Sociology
The Ohio State University
Columbus, OH 43210, USA
work@samclark.net

Shams El Arifeen

Senior Director, Maternal and Child Health
International Centre for Diarrhoeal Disease
Research (ICDDR)
Dhaka, Bangladesh
shams.el.arifeen@gmail.com

Soewarta Kosen

Country Coordinator, Civil Registration and
Vital Statistics Component
Data for Health
Indonesia
soewarta.kosen7@gmail.com

Tita Rosita

Center for Health Management &
Humanities
Indonesia Agency for Health research &
Development
Jalan Percetakan Negara 29
Jakarta 10560, INDONESIA
tita_wiguno@yahoo.com

Vishnu Vardhan Rao

Director, National Institute of Medical
Statistics
ICMR, Ansari Nagar, New Delhi 110 029
INDIA
dr_vishnurao@yahoo.com

Country and field teams

Anupap Somboonsavatdee

Assistant Professor of Statistics
Director
Center of Statistical Consulting and
Research
Chulalongkorn Business School
Chulalongkorn University,
Bangkok, Thailand
anupap@cbs.chula.ac.th

Beth Barr

Director
Western Kenya
Nairobi, Kenya
btippettbarr@cdc.gov

Chabila C. Mapoma

D4H consultant
University of Zambia/D4H-Zambia
Zambia
chabilamapoma@gmail.com

Debbie Bradshaw

Director
Burden of Disease Research Unit
South African Medical Research Council
Cape Town, South Africa
debbie.Bradshaw@mrc.ac.za

Dianna Blau

Managing Director for Science
CHAMPS
Centers for Disease Control and
Prevention
3311 Toledo Rd., MS P08
Hyattsville, MD 20782, USA
JKOPLAN@emory.edu

Fidelia Dake

Ghana Country Coordinator
Vital Strategies
fdake@dataforhealth.org

Hicham Darfouf

Ministry of Health
Morocco
darfouf-hicham@hotmail.fr

Isaac Lyatuu

Ifakara Health Institute
Tanzania
ilyatuu@gmail.com

Jamila El Mendili

Division of planification and studies
Planification and financial resources
department
Ministry of Health
Morocco
Elmendilij@yahoo.fr

Kanitta Bundhamcharoen

Researcher, Lead investigator
Burden of Disease Program
International Health Policy Program
Ministry of Public Health
Nonthaburi, 11000, Thailand
kanitta@ihpp.thaigov.net

Monique Maqungo

South African Medical Research Council
Cape Town, South Africa
Monique.Maqungo@mrc.ac.za

Nuttapat Makka

Research assistant
Burden of Disease Program
International Health Policy Program
Ministry of Public Health
Nonthaburi, 11000, Thailand
nuttapat@ihpp.thaigov.net

Odhiambo Collins Ochieng

Programmer

KEMRI VASK

Kenya

Coodhiambo@kemricdc.org

gollinsochieng@gmail.com

Pamela Groenewald

Burden of Disease Research Unit

South African Medical Research Council

Cape Town, South Africa

Pamela.Groenewald@mrc.ac.za

Patrick Larbi-Debrah

Ghana Health Service

Accra, Ghana

Sajid Soofi

Associate Professor

Department of Paediatrics and Child
Health

The Aga Khan University

Karachi, Pakistan

sajid.soofi@aku.edu

Steve B. O. Odhiambo

Data Analyst

HDSS KEMRI

Kenya

stevephilomena@gmail.com

SBOdhiambo@kemricdc.org

Youssoupha Ndiaye

Director for Planning, Research and
Statistics

Ministry of Health and Social Action

Dakar, Senegal

youndiaye2000@gmail.com

Zulfiqar Bhutta

Director

Institute for Global Health and
Development

Aga Khan University

Karachi, Pakistan

zulfiqar.bhutta@aku.edu

Special contributors**Alicia Perez**

University of the Basque Country

alicia.perez@ehu.eus

Arantza Casillas

University of the Basque Country

arantza.casillas@ehu.eus

Carmen Diva Saldiva de André

University of São Paulo, Brazil

carmensaldiva@gmail.com

Clarissa Clark

The Ohio State University

Columbus, OH 43210, USA

clarissa@clarissasurekclark.name

Diane Morof

CDC South Africa office

igd6@cdc.gov

Fátima Marinho

University of São Paulo, Brazil

mfmsouza@gmail.com

Hermon Gebrehiwet

CDC/NCHS

hgebrehiwet@fasttrackresearch.com

Jason Thomas

openVA

jarathomas@gmail.com

Jonathan Brent Vickers

CDC/NCHS

nnx9@cdc.gov

Kristen Pettrone

CDC/NCHS

pho3@cdc.gov

Lucia Barosso

University of São Paulo, Brazil

lpbarroso@gmail.com

Mohamed Jelle

University College London

30 Guilford Street

London WC1N 1EH, United Kingdom

m.adan@ucl.ac.uk

Owen Trigueros

University of the Basque Country

otrigueros001@ehu.eus

Paulo Afonso de André

University of São Paulo, Brazil

paulo.afonso@fm.usp.br

Peter Choi

openVA

peter1125@gmail.com

Richard Li

openVA

lizehang@gmail.com

Susan Gacheri

Public health consultant

Kenya

smgacherism@gmail.com

Tyler H. McCormick

University of Washington

tylermc@uw.edu

Yue Chu

OpenVA

chu.282@osu.edu

WHO

Benson Droti

WHO – AFRO Office
drotib@who.int

David Novillo Ortiz

WHO – EURO Office
dnovillo@who.int

Eman Aly

WHO – EMRO Office
alye@who.int

Henry Doctor

WHO – EMRO Office
doctorh@who.int

Hillary Kipruto

WHO – AFRO Office
kiprutohi@who.int

Jun Gao

WHO – WPRO Office
gaoj@who.int

Karapet Davtyan

WHO – EURO Office
kdavtyan@who.int

Vilma Gawryszewski

WHO – PAHO Office

APPENDIX B – AGENDA

DAY 1: ITEM RESPONSE PATTERNS

8th July (start at 8:00 am EST/12:00 UTC/14:00 Geneva)

Presentation and discussion of the quantitative and qualitative findings of response patterns with the 2016 WHO VA instrument

45 mins	Introduction Kick-off Introductions Logistics	<i>Michelle Panneton</i> <i>Daniel Chandramohan</i>
1 hr	Quantitative item response pattern results over full dataset Intro/approach Results Questions/discussion	<i>Jason Thomas</i> <i>Samuel Clark</i>
15 mins	Break	
2 hrs	Quantitative and qualitative (mixed-methods) results for known VA issues Intro/approach Example results Questions/discussion Breakout groups to discuss specific issues	<i>Erin Nichols</i> <i>Kristen Pettrone</i> <i>Brent Vickers</i>

DAY 2: ITEM RESPONSE IMPORTANCE

9th July (start at 8:00 am EST/12:00 UTC/14:00 Geneva)

Presentation of the set of item importance assessment methods to be applied on VA questions

1 hr	Targeted Maximum Likelihood Estimation (TMLE) Intro/approach Results Questions/discussion	<i>Jason Thomas</i> <i>Samuel Clark</i>
1 hr	Entropy coefficient and VA narrative contributions Intro/approach Results Questions/discussion	<i>Daniel Cobos</i> <i>Owen Trigueros</i> <i>Alicia Perez</i> <i>Arantza Casillas</i>
15 mins	Break	
2 hrs	Discussion, conclusions and next steps	<i>Daniel Chandramohan</i>

DAY 3: TRIANGULATION OF FINDINGS

15th July (start at 8:00 am EST/12:00 UTC/14:00 Geneva)

Combination of approaches introduced in the first part of the virtual workshop and cross-verification of findings for item reduction

10 mins	Introduction Kick-off Introductions Logistics	<i>Michelle Panneton Daniel Chandramohan</i>
80 mins	Mixed-Methods Analysis Follow-up An intercultural approach to VA implementation Breakout group recap (5 mins/group) Triangulation Discussion	<i>Clarissa Surek-Clark Group notetakers Erin Nichols</i>
30 mins	Triangulation of Response Pattern and Item Significance Analysis	<i>Sam Clark</i>
10 mins	Break	
105 mins	Triangulation of Response Pattern and Item Significance Analysis (cont'd) Breakout groups Breakout group recap (5 mins/group) Discussion	<i>Group notetakers</i>

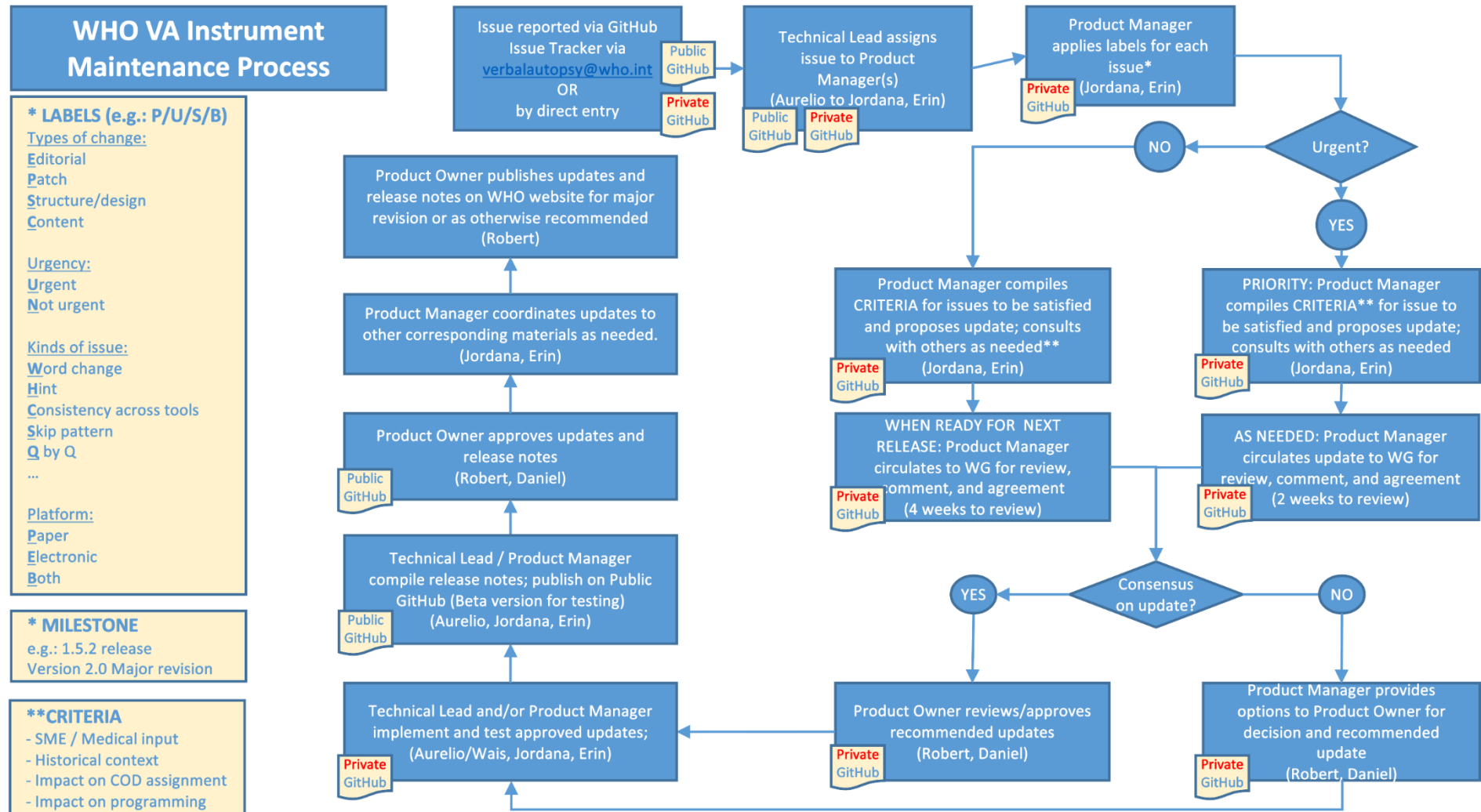
DAY 4: COVID-19, RECOMMENDATIONS AND NEXT STEPS

16th July (start at 8:00 am EST/12:00 UTC/14:00 Geneva)

Discussion and synthesis of findings and next steps

10 mins	Introduction Kick-off Logistics	<i>Michelle Panneton Daniel Chandramohan</i>
35 mins	VA in the Context of COVID-19	<i>Erin Nichols</i>
60 mins	Summary Recommendations and Next Steps Wrap up	<i>Daniel Chandramohan Daniel Chandramohan Robert Jakob</i>

APPENDIX C – WHO VARG MAINTENANCE PROCESS OF THE WHO VA INSTRUMENT



APPENDIX D – LIST OF PROPOSED QUESTIONS FOR THE IDENTIFICATION OF POSSIBLE COVID-19 DEATHS

1. (Id10482) Was there any diagnosis by a health professional of COVID-19?

- Yes
- No
- Don't know
- Refused to answer

2. (Id10483) Did s(h)e have a recent test by a health professional for COVID-19?

- Yes
- No
- Don't know
- Refused to answer

2.1. (Id10484) What was the result? *(Hint for interviewer: Prompt for the result of the most recent test in case the deceased had more than 1 test performed)*

- Positive
- Negative
- Unclear
- Don't know
- Refused to answer

3. (Id10485) Did s(h)e suffer from extreme fatigue? *(Hint to interviewer: Probe whether the deceased felt so tired that (s)he found it hard to get out the bed and do the routine things like taking a shower or changing clothes)*

- Yes
- No
- Don't know
- Refused to answer

Restriction: Only for adults.

4. (Id10486) Did (s)he experience a new loss, change or decreased sense of smell or taste?

- Yes
- No
- Don't know
- Refused to answer

Restriction: Only for adults

5. (Id10487) In the two weeks before death, did (s)he live with, visit, or care for someone who had any COVID-19 symptoms or a positive COVID-19 test? *(Hint to interviewer: COVID-19 symptoms include fever, difficulty breathing, cough, extreme fatigue, and changes in sense of smell or taste. In case of neonates or young children, please omit "care for".)*

- Yes
- No
- Don't know
- Refused to answer

6. (Id10488) In the two weeks before death, did (s)he travel to an area where COVID-19 is known to be present? *(Hint to interviewer: Based on self-report of the respondent. If there is doubt, note the location in the narrative and check with the respective supervisor.)* Yes

- No
- Don't know
- Refused to answer