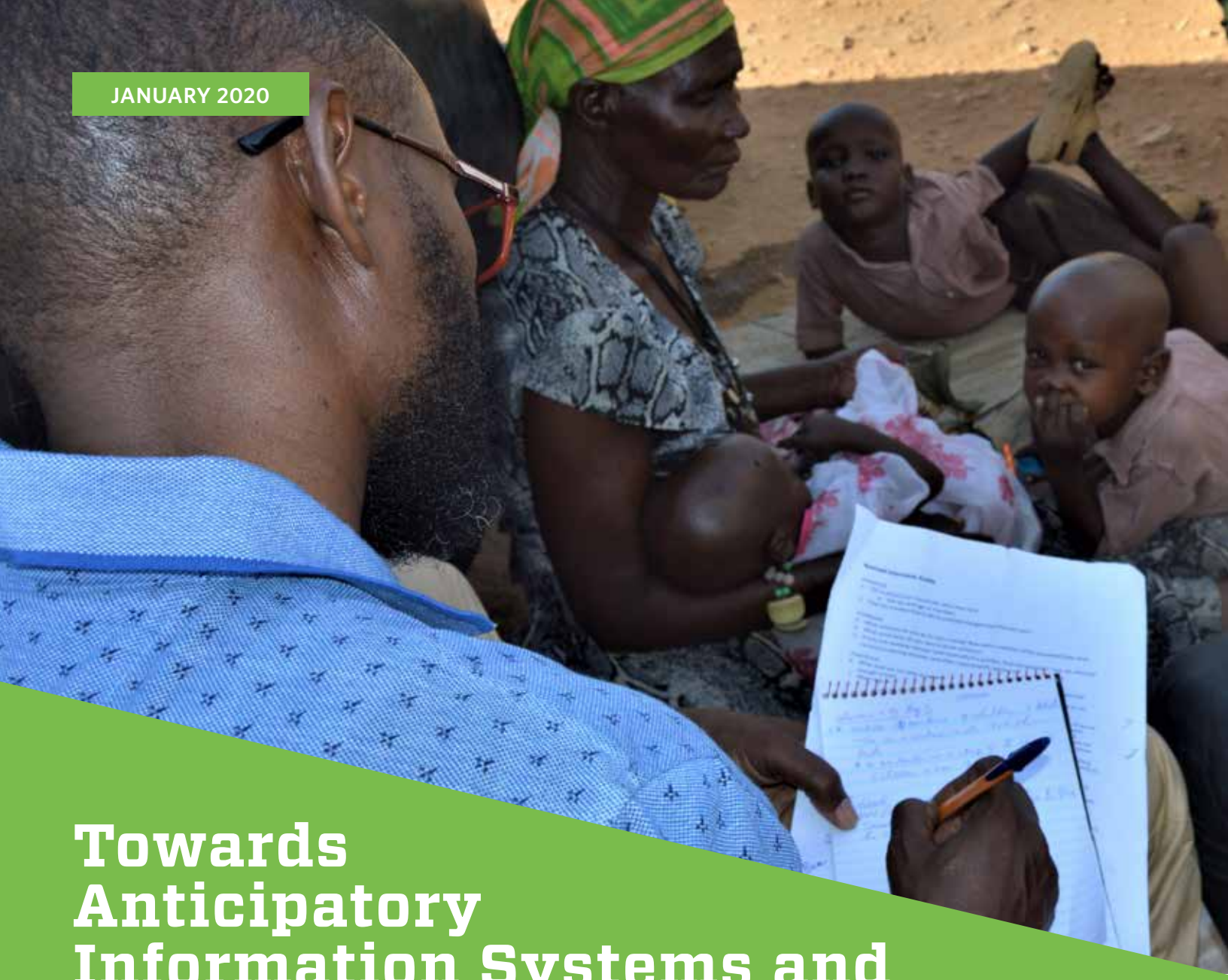


JANUARY 2020

Towards Anticipatory Information Systems and Action: Notes on Early Warning and Early Action in East Africa

A FEINSTEIN INTERNATIONAL CENTER AND CENTRE FOR HUMANITARIAN CHANGE PUBLICATION

Daniel Maxwell (Tufts University) and
Peter Hailey (Centre for Humanitarian Change)

Tufts University
Friedman School of Nutrition Science and Policy
Feinstein International Center
Centre for Humanitarian Change
Food and Agriculture Organization of the United Nations
IPC Integrated Food Security Phase Classification
Action Against Hunger
REACH Informing more effective humanitarian action

Cover photo: Joyce Maxwell

Citation: Daniel Maxwell and Peter Hailey. "Towards Anticipatory Information Systems and Action: Notes on Early Warning and Early Action in East Africa." Boston: Feinstein International Center, Tufts University; Nairobi: Centre for Humanitarian Change. 2020.

Corresponding author: Daniel Maxwell

Corresponding author email: daniel.maxwell@tufts.edu

Photo credits: Joyce Maxwell

Feinstein International Center, Friedman School of Nutrition Science and Policy

Tufts University
75 Kneeland Street
8th Floor
Boston, MA 02111

Tel: +1 617.627.3423

Twitter: @FeinsteinIntCen

fic.tufts.edu

Acknowledgements

We are grateful for feedback on an initial draft of these notes from (in alphabetical order by surname) Jannie Armstrong, Zacharey Carmichael, Sophie Chotard, Matthew Day, Yvonne Forsen, Barbara Frattaruolo, Laura Glaser, Gregory Gottlieb, Arif Hussain, Douglas Jayasekaran, Abdullahi Khalif, Kaija Korpi, Brenda Lazarus, Erin Lentz, Jose Lopez, Leila Oliveira, Daniel Pfister, Chris Porter, Anne Radday, Segio Regi, Katie Rickard, Vanessa Roy, Peter Thomas, Kamau Wanjohi, and Ellyn Yakowenko. We are grateful to Joyce Maxwell for copy editing.

Disclaimer

This work was undertaken while one author (Maxwell) was on sabbatical leave from Tufts University, affiliated with the Centre for Humanitarian Change (CHC) in Nairobi (Hailey). This work was part of a larger study supported by the UN Food and Agriculture Organization (FAO), REACH, and Action Against Hunger. The views expressed here are those of the authors and do not necessarily reflect the views of any of the supporting agencies, Tufts University, or the Centre for Humanitarian Change.

Contents

Acronyms

ACAPS Assessment Capacity Program
ACF Action Contre la Faim (Action Against Hunger in the US)
ACTED Agency for Technical Cooperation and Development (French)
AFIA Successor to SIFSIA
AI artificial intelligence
ALNAP Action Learning Network for Accountability and Performance
AoK area of knowledge
ASAL arid and semi-arid lands
BRCiS Building Resilient Communities in Somalia
BRE Building Resilience in Ethiopia
BVPA baseline vulnerability and poverty assessments
CERF Central Emergency Response Fund
CEWARN Regional Conflict Early Warning Project
CHC Centre for Humanitarian Change
CLiMIS Climate Information System
DEWS Drought Early Warning System
DFID Department for International Development (UK)
DISK Data and Information Subcommittee of the KFSSG
DPRMC National Disaster Risk Management Commission
EA early action
ECHO Office of European Civil Protection and Humanitarian Aid Operations
ENA emergency needs assessment
EPRDF Ethiopian People's Revolutionary Democratic Front
EW early warning
EW/EA early warning/early action
EWS early warning system
FAM World Bank Famine Early Action Mechanism
FAO UN Food and Agriculture Organization
FEWS NET Famine Early Warning System Network
FSL Food security and livelihoods
FSNAU Food Security and Nutrition Analysis Unit (Somalia)
FSNMS Food Security and Nutrition Monitoring System
FSNWG Regional Food Security and Nutrition Working Group
FSOM WFP Food Security Outcome Monitoring system
GAM Global acute malnutrition
GHACOF Greater Horn of Africa Climate Outlook Forum
HEA Household Economy Analysis
HRP Humanitarian Response Plan
IBLI Index-based livestock insurance
ICHA International Centre for Humanitarian Affairs
ICPAC IGAD Climate Prediction and Applications Centre
ICWG Inter-Cluster Working Group
IGAD Inter Governmental Authority on Development
INT Integrated Needs Tracking system
IPC Integrated Phase Classifications

JEOP Joint Emergency Operation (Ethiopia)
KFFSG Kenya Food Security Steering Group
KFSM Kenya Food Security Meeting
LEAP Livelihoods, Early Assessment and Protection
LIAS Livelihood Impact Analysis Sheet
LRA long rains assessment
M&E Monitoring and evaluation
MAAIF Ministry of Agriculture, Animal Industries, and Fisheries
MERIAM Monitoring Early Risks Indicators to Anticipate Malnutrition
MUAC Mid-upper arm circumference
NAWG Needs Assessment Working Group
NDMA National Drought Management Authority (Kenya)
NDRMC National Disaster Risk Management Commission (Ethiopia)
NDVI normalized difference vegetation index
NGO Non-Governmental Organization
NITWG Nutrition Information Technical Working Group
OCHA Office for the Coordination of Humanitarian Affairs
OTP Out-patient therapeutic feeding programs
PHEM Public Health Emergency Management
PRIME project operating in Somali Region
PSNP Ethiopia Productive Safety Net Program
SIFSIA Sudan Institutional Capacity: Food Security information and Analysis
SMART standardized methods for assessment of relief and transition
SRA short rains assessment
SSNPR Southern Nationalities and Nations People's Region (Ethiopia)
SSRRC South Sudan Relief and Rehabilitation Commission
UN United Nations
UNHCT UN Humanitarian Country Team
UNICEF UN Children's Fund
UNMISS UN Mission to South Sudan
USAID US Agency for International Development
WASH water, sanitation and hygiene
WFP World Food Programme
WFP VAM WFP Vulnerability Analysis and Mapping Unit
WRSI Water Requirement Satisfaction Index

Executive Summary

Despite early warning and humanitarian diagnostics information being more available than ever in history, confusion persists as to what it means and what to do with it. This review of early warning highlights several contemporary issues with humanitarian information and early warning (EW) systems. Cases are drawn from the East Africa region, but they have broader implications as well.

A number of points of confusion stand out in this review. These include the key question of how to clearly differentiate current status, projections of numbers in need, and early warning of threats along with the ability to rapidly identify deteriorating situations. A second point of confusion persists about the difference between "hard numbers" (which inevitably imply something that has already happened) and probabilistic estimates (about things that are likely to happen, but haven't happened yet). A third point of confusion regards linkages between information systems and action in terms of both policy and programs (this includes the much discussed lack of early warning/early action linkages but equally applies to longer-term actions and other parts of the program cycle as well). A fourth point is that conflict analysis is the weakest part of early warning, despite the fact that conflict is the common factor driving extreme humanitarian crises. Finally, the domain of early warning and humanitarian information systems is perceived to belong to data collection and analysis agencies as well as governments, donors, and humanitarian response agencies. There is limited recognition of the imperative of engaging with (or providing early warning information to) the communities that are at risk of shocks or resulting humanitarian crises.

The study highlights several key findings. First, the link of early warning to early action is not as effective as it could be. One key reason is a lack of clarity over what is a "projection," a "signal," and a "scenario." A "projection" is an estimate of the number of people in need of a particular kind of response (typically, but not necessarily, food assistance) at some point in the near-term future. A "signal" is an automatic trigger for some kind of rapid action. A trigger can be a single indicator, or a combination of factors that lead to a certain outcome. "Scenarios" are a more fleshed out analysis of what is likely to happen and inevitably involve turning lots of complex information into probabilistic descriptions of outcomes and priorities for response. These distinctions matter because different approaches to early warning shape different early action responses. The link of "scenarios" to early action include programmatic responses such as "crisis modifiers," "no regrets" programming, and surge approaches that build on already existing capacities—or in some cases, risk reduction and mitigation efforts. "Projections" may, at first glance, appear to simply be forecasting a needs assessment figure for early planning purposes, but projections are typically (if not always obviously) based on some kind of scenario analysis and may suggest different courses of action in addition to predicting the number of people in need in the near term future. A "signal" typically triggers a specific financial response, such as an insurance pay-out or a disaster bond—although these financial resources can also support no-regrets or surge programming. But early action can also consist of risk reduction and mitigation efforts. Each of these require different kinds of "early warning" information. In general, currently no single approach predominates—and, technically speaking, complementary approaches should be able to play out in concert. While most parties have preference for one or the other, ensuring that they work side by side would improve the overall humanitarian information system.

Second, conflict is a common driver of humanitarian crisis, but conflict early warning is weak, and discussion of conflict is often limited to being mentioned as a "contributing factor" and sometimes is missing from analysis altogether. The more specific humanitarian concern is not so much to predict conflict itself, as it is to systematically consider and incorporate the consequences of conflict into early warning for specific humanitarian outcomes. But scenario analysis would be significantly improved by better anticipation of conflict itself.

Third, beyond the analysis of conflict, political interests play a role in influencing the outcomes of humanitarian analysis—both of current status and early warning. Some of this relates to the role of governments in leading or managing information systems, but agency politics influence the analysis as well. Learning to better manage these political influences is a key challenge to humanitarian information systems almost without exception.

Fourth, new technologies involving remote sensing, satellite imagery, computational modeling, and artificial intelligence are all competing to improve early warning and humanitarian information systems. But it is not always clear whether these new technologies are being developed to address specific short-comings of existing systems or simply because technology developers are in search of applications and new markets (or some of both). New technologies can certainly help address some of the issues highlighted in this report but would bring with them some new issues that would require resolving.

Finally, the role and use of qualitative data, in early warning and information systems is unclear. Humanitarian information systems are heavily dominated by quantitative data and analysis systems. Yet, qualitative data is an important complement to quantitative efforts, both to aid in triangulation of findings but also because in highly dynamic and insecure situations, qualitative data may better capture the nature of crisis compared with quantitative data—and collecting quantitative data on a scale sufficient to be statistically representative may not be possible. As new quantitative approaches emerge, some major concerns have completely fallen through the cracks, including how and where any of these initiatives (traditional EW or computational modeling and artificial intelligence or AI) intersect with local realities and inform community action to prepare for or protect against shocks and hazards. A related question is about the role of human judgement in systems that purport to be "data driven" and analyzed by algorithms. Several recommendations grow out of these observations. Some of these are recommendations about *what* needs to be done; others are about *how* to do things differently.

1. **Focus on key issues, not institutions**. It makes little sense to scrap the systems we now have to start over from scratch. Given the wide range of actors in this arena (national governments, UN agencies, and international and national NGOs), concerns can be addressed within existing institutions or approaches.
2. **Think strategically about components of a "system."** Good early warning needs a variety of kinds of information. Estimates of current and future numbers of people in need are among these. So is the monitoring of risk and predictions about hazards. These must come together into an analysis of what the future may look like and the means to respond to human need in an anticipatory way.
3. **Build better linkages**. Within information systems greater integration both horizontally (between different systems) and vertically (across levels and time frames) is needed. Beyond information systems, much stronger mechanisms have to be built within decision-making and resource-allocation systems.
4. **Take a broader view of crisis and risk**. Current analytical approaches focus heavily on the severity of crisis and risk—dimensions of magnitude, duration, and spatial distribution are equally important.
5. **Build better mechanisms for "system accountability."** Accountability should focus on the accuracy of forecasts (were forecasts correct?), early action (did the forecast trigger action?), impact (did the action protect affected communities?), and learning.
6. **Broaden the scope of information**. To provide a more holistic understanding, a wider range of measures needs to be incorporated into existing information systems. These include coping and social connectedness, along with better information on WASH and health outcomes and a much stronger focus on causal factors. Better guidance is urgently needed for how to utilize qualitative information.
7. **Treat humanitarian information as a public good**. Humanitarian information is often not available for users and analysts to see. Donors can make this a requirement.
8. **Develop better methods to deal with politics**. This includes the politics of numbers of "people

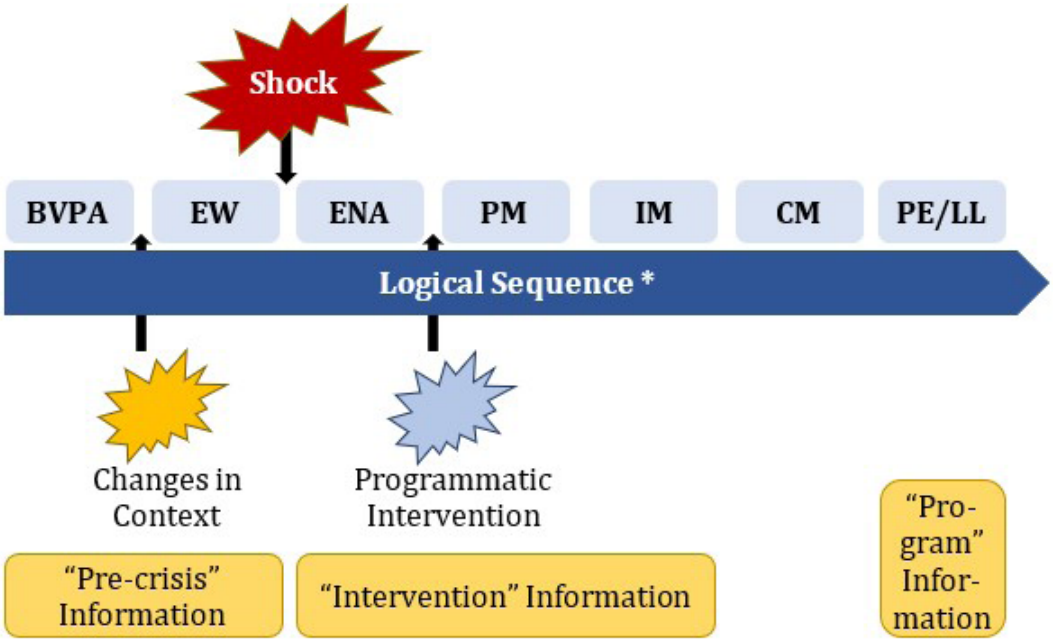
in need," the politics of famine, and accusations about undermining national sovereignty.

9. **Improve conflict information and conflict early warning in humanitarian information systems.** Conflict is a very common driver of crisis across the region. Conflict early warning is a field in its own right—it should be more systematically incorporated into humanitarian early warning, and its information should be fed into humanitarian scenarios and contingency plans.
10. **Clarify the role of government leadership**. Nearly all parties agree that government should lead on information systems, but this is problematic when non-state actors control much of the affected territory, or when government is one party in a conflict that is driving the humanitarian emergency. This brings up the inevitable question of sovereignty and the humanitarian imperative.
11. **Agencies engaged in information and EW initiatives have to work together.** Organizations implementing early warning initiatives urgently need to talk to each other. Many are attempting to address the same objectives or achieve the same outcomes but are unaware of the work that others are doing. Key priorities for dialogue would include (1) common problem identification, (2) identification of ways predictive modeling or AI analytics improve the quality of early warning information, (3) how will these very different approaches work together, and (4) whether predictive modeling can or cannot address political concerns, as well as (5) the accountability and transparency issues highlighted above.

1. Introduction

East Africa continues to be one of the most at-risk regions of the globe in terms of food insecurity, malnutrition, and poor health outcomes. Much of the region suffers from chronic poverty. Since at least the 1970s, some form of famine early warning has existed in the region, and this has become increasingly sophisticated—now relying on on-the-ground information collection systems combined with remote sensing, satellite imagery, complex modelling, and, increasingly, artificial intelligence. Yet significant challenges remain. This brief review notes some of these challenges, attempts to identify key questions, reviews existing systems and some of the constraints they face, and offers a modest analysis of the state of early warning/early action (EW/EA) in East Africa, with some reflections on systemic improvements.

Figure 1. The role of early warning in a humanitarian information system



* The logical sequence is depicted as linear, but in fact it rarely is.

BVPA	Baseline Vulnerability/Poverty Assessment (periodic)
EW	Early Warning (continuous)
ENA	Emergency Needs Assessment (as needed—or periodic for planning purposes)
PM	Program Monitoring (tied to programs)
IM	Impact Monitoring (tied to programs)
CM	Context Monitoring (continuous)
PE/LL	Program Evaluation/Lessons Learning (periodic)

Source: Maxwell and Watkins (2003)

2. Problem Statement

Practical Action (n.d.) defines early warning in relation to early action and defines early warning (EW) as "the provision of information on an emerging dangerous hazard that enables advance action to reduce the associated risks. Early warning systems exist for natural geophysical and biological hazards, complex socio-political emergencies, industrial hazards and personal health risks, among many others. . . . Early action can often prevent a hazard turning into a human disaster by preventing loss of life and reducing the economic and material impacts. To be effective and sustainable they must actively involve the communities at risk. . . . *The significance of an effective early warning system lies in the recognition of its benefits by local people*."[1]

To be effective, EW must take information about current events, trends, and signals (observable, empirical information), analyze that information to turn it into forecasts or scenario analyses (unobservable, probabilistic information), and link it directly to a decision-making mechanism that is accountable to act on the forecast or likely scenario. In other words, information—even if highly accurate in its forecast—is relatively useless if it is not acted upon, so it is critical to have an EW system that is tied, directly or indirectly, into a decision-making and action body. And as the Practical Action definition emphasizes, the basis on which to judge an EW system is not just its accuracy but its results—including to affected communities. This underlines the need for general information and targeted action and a strong link between the two.

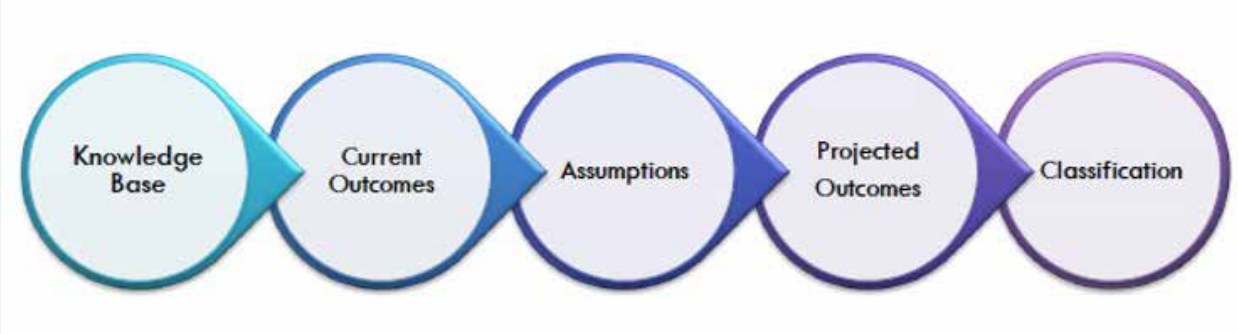
Sixteen years ago, Maxwell and Watkins (2003) tried to demonstrate the difference between early warning—as an information collection and analysis activity—from other activities in a humanitarian information system or framework. The assumption behind their framework was that the entire system (not just the early warning component) was linked to a decision-making and action body (Figure 1).

Although sometimes the entire left hand side of the diagram in Figure 1 may be referred to as an "early warning system," the point is that there is a difference between information about existing risks and hazards (baseline information), information about early warning (how those risks are changing and how they may affect human populations), and information about needs (current status). Both baseline vulnerability and poverty assessments (BVPA) and emergency needs assessments (ENAs) are *assessment* activities—they give a comprehensive, empirical picture about either baseline or current status. Early warning (EW), in this strict sense, is different—it is about prediction, about detecting adverse events as or before they occur and how they are likely to impact vulnerable populations. EW is intended to be indicative (not comprehensive) and probabilistic (not "hard" data).

Note that these different kinds of information have different purposes; all are necessary to enable early action. Baseline vulnerability/poverty assessments should enable preparedness planning and early action to manage risk and hazards; early warning should enable early action to mitigate the impact of a shock or stressor and also deployment of assessment resources in a rapid and well-targeted manner. Emergency (current status) needs assessments should enable rapid response—and provide information about how many people are affected, how badly, and for how long. (Note that the rest of Figure 1 is about monitoring interventions, so is tied to the monitoring and evaluation side of humanitarian information systems. This report is concerned with the "diagnostics" side of Figure 1). Diagnostics are based on indicators of food security, nutrition, health, and other humanitarian outcomes, as well as the drivers or factors that induce change in these outcomes. Indicators are frequently imperfect measures of those outcomes, and the way in which imperfect measures are interpreted is key to the analytical judgments that result. Thus at root, despite advances in data collection techniques, measurement, remote sensing, and modeling algorithms, the capacity of analysts to make sense of all this is still key.

[1] Practical Action, Policy and Practice: Early Warning Systems Project. (n.d.). Emphasis added. https://policy.practicalaction.org/projects/ews

Figure 2. "Approach to Early Warning Analysis and Outlook"



Source: FEWS NET

Nevertheless, early action rarely fails purely because of poor or insufficient early warning information. There have been numerous failures of early action over the years—few were directly caused by lack of information. This has become such a common problem that it even has a name—the early warning/early action "gap" or failure (Buchanan-Smith and Davies 1995, Hillbruner and Moloney 2012, Bailey 2013, Maxwell and Majid 2016). Recent initiatives have attempted to address this, but with limited success.

But the understanding of the nature of the output of early warning has subtly challenged. Some decision makers prefer the "hard numbers" of assessments rather than the probabilistic information of early warning. But that, by definition, means that action is delayed long past the point that mitigation or prevention actions are possible. Recognition of this has resulted in the attempt to merge hard numbers with forecasts, as manifested most clearly by Integrated Phase Classifications (IPC) projections. They provide one kind of useful information—expected numbers of people in need. But projections are not the only kind of early warning information. Early warning is a continuous activity, meant to give predictive information about changes in the situation, the impact of a shock, or the development of "hotspots"—areas of rapidly deteriorating humanitarian conditions—and the likely consequences thereof. Projections take the information available at a given point in time and project it into the future based on assumptions about seasonality and other contributing factors and how they influence each other. But projections remain fixed on the numbers of people that are now expected to need assistance in the future based on a range of assumptions. Figure 2 outlines the process. The third step in this process shown in Figure 2, "assumptions," is where early warning information becomes an input to the analysis. The output is the now-familiar IPC map—but rather than showing current status, the map depicts expected status three to six months in the future.[2]

Projections are one kind of early warning information. Numbers of expected people in need are important to donors and planners. But this process works well if the "assumptions" are updated very frequently. The other kinds of early warning information needed were highlighted in several recent examples in East Africa in 2019 (which was admittedly a difficult year to analyze and predict—but of course that is when good early warning is needed the most). At one point, for instance, the projections were all about the impact of the poor rainfall in the first months of the long rains (and rightfully so, because crop performance and livestock browse were strongly affected). At the same time, early warning information (for example, the Greater Horn of Africa Climate Outlook Forum) was noting a "positive" Indian Ocean dipole and was rightfully concerned about heavy rainfall and potential flooding. Both were "correct" in a sense, but it was difficult to get an overall sense of

[2] In some cases, attempts are being made to project future needs even farther into the future.

which mitigative action or response was appropriate to invest in.[3]

The crucial difference is the extent to which predictions of future need are tied to existing conditions and the extent to which they are influenced by predicted hazards that have yet to translate into a "shock" to human populations. Both are important, but one has tended to dominate. Several key issues in East Africa about relying primarily on projections need to be understood. The first is that while projections offer numbers about the future, they are rarely checked against future actual outcomes.[4] This kind of checking of predictions against actual outcomes is necessary for improving the system. In some cases, projections for the same populations at the same (projected) time period from related sources were quite different—meaning that the accuracy of a future check would depend on which projection was chosen.[5]

The second is that two different kinds of information are involved: outcomes and causal factors or drivers (sometimes called "contributing factors"). Causal factors by definition lead to changing circumstances; outcomes describe those circumstances. As depicted in Figure 2, assumptions (about the drivers or contributing factors) are the key link, and the frequency with which those assumptions are updated is key to the veracity of the early warning: in a dynamic, changing environment, information needs to be updated fairly constantly as risks change and hazards develop. Making a projection today that covers the next six months is different from monitoring and predicting constantly over the next six months (as demonstrated by the rapidly changing situation in the Horn of Africa in 2019). Current status is more or less the "short-term baseline" for any forecast about current status in the future. But whereas current status is based on "hard numbers," projections (and early warning information of any type) are inevitably based on probabilities, but these probabilities are often not explicit. Projections, including of highly specific numbers of people in different need classification categories (such as IPC phases) provide an illusion of certainty. The clear lack of certainty under the circumstances that prevailed in 2019 in East Africa led once again to confusion about how to respond.

A third issue, directly linked to the first two, is that while early warning *information* is routinely available—in fact may be *more* available now than at any time in history—it is often up to the individual consumer of that information to make sense of it (that is, to come up with a comprehensive analysis of it) or do anything about it (that is, to act on it). The major national sources of EW information in the region (the Somalia FSNAU dashboard, the Kenya NDMA early warning bulletins, etc.) provide a lot of information. The question is how that information is translated into an analysis—a forecast or a prediction—and then into action. Lots of information can be available about the rainfall, production estimates, livestock condition, prices, etc., without any particular good analysis of *what is actually likely to happen*. This requires synthesis of all that information—both process indicators and actual human outcomes—and the building of scenarios. FEWS NET uses scenario building as the means of conducting early warning analysis and routinely presents the "most likely" scenario in its analysis. With the latest version (V.3) of the IPC Technical Manual (IPC Partners 2019), IPC now has similar scenario building guidance.[6] The FSNAU dashboard identifies "alarms"—deviations of more than a set threshold compared to long-term means—and then counts up the number of alarms per district and maps these by increasingly deep colors of red depending on the number of "alarms." This gives a lot of information about the situation, but doesn't

[3] For instance, see "Early Warning-Early Action Dashboard Time Series Maps, January 2015-September 2019" (FSNAU, 2019) and "The Greater Horn of Africa Climate Outlook Forum" (The New Humanitarian, October 22, 2019).

[4] One notable exception to this is Chourlaton and Krishnamurthy (2019) on FEWS NETs projections.

[5] The most recent confusing example was the very different projections regarding large areas of southern and eastern Ethiopia following an IPC analysis there in mid-2019. Those differences were explained in terms of the timing of different products from the analysis, but for the consumer of information, they were confusing—and future checks of accuracy would depend on which version of the projection was chosen for comparison to observed outcomes.

[6] FEWS NET analysis is "IPC compatible"—meaning that FEWS NET staff follow IPC guidance, but conduct their analysis independently.

necessarily provide any early warning scenario or analysis.[7] Other national sources of early warning information raise similar issues. But at a minimum, this should suggest where current status assessment resources should be focused.

A fourth major concern is the one already mentioned—whether or not early warning information/analysis is ever systematically linked with early action (EA). This has been written about so much in recent years that there should be nothing else to say about it—but somehow the idea persists that the information or data "speaks for itself" and there shouldn't be any further need for linking information to action (i.e., decision makers are so hungry for information that they will automatically take it up and act on it). There are several problems with this: First, it obviously is not true—or this wouldn't be an issue. But second, and more importantly, EW information on its own is often confusing or even contradictory—without a nuanced analysis of what it means, it can be very difficult to act on. Probabilistic information about the future is just that—*the "most likely scenario" may in the end not turn out to be what actually transpires*. And of course, there is no escaping the politics of decision-making about the response.

This has given rise to the whole notion of "no regrets" programming or scalable safety nets—that intervention should proceed on the basis of the best information and analysis available and be based on actions that will mitigate negative humanitarian outcomes, but which should also be beneficial or developmental even if the situation does not deteriorate to the extent predicted. But the early warning/early action linkages are still tenuous, and this has helped create the demand for "triggers" or automated signals for pre-arranged responses or mitigative actions. On the other hand, other responses, including preventive or mitigative actions that might be taken under the banner of "crisis modifier" or livelihoods protection programming, do not require estimates of future populations in need—they should be triggered by drivers, not by outcomes.

And finally, this whole early warning information side of the early warning/early action question is often viewed as the preserve of data specialists and information analysts. Early action is left to program decision makers and donors. This emphasizes the need for stronger mechanisms to prompt early action. Questions that should arise here include whether or not the analysts actually provided the program staff with the information they needed, when they needed it, in a form accessible to them. Was that information understandable and usable? Was it timely? It is also useful to keep in mind the Practical Action definition about how one should judge the effectiveness of the entire system: does it protect the lives and livelihoods of at-risk communities (and does it do so in their perception, or just in the perceptions of data specialists, analysts, decision makers, and donors)?

None of this should be news to anyone! But somehow, the fact that this should all be common knowledge hasn't prevented the humanitarian community from confusing different types of information and analysis, and has not enabled better linking of information and analysis in fragile or at-risk environments to better preventive, mitigative, and resilience-building activities.

In a blistering critique written about the Sahel over a decade ago, Kent Glenzer noted that early warning/early action systems to detect and prevent famine are, at best, an institutionalized form of "partial success": some lives are saved and some livelihoods protected, but the whole system only kicks into gear when *some* lives have been lost and some livelihoods destroyed (Glenzer 2009, p. 224). Though writing about the 2005–06 crisis in Niger, his critique still stands today in terms of the engagement with—and accountability to—affected communities. What Glenzer generously called a "partial success" ten years ago, critics like Simon Levine more recently called a "system failure," or the inability to prevent substantial humanitarian loss (lives, livelihoods, dignity) through appropriate information and action (Levine et al. 2012). The demand for anticipatory humanitarian analysis and action has never been higher.

[7] This is changing. A task force is met regularly during 2019 to try to build on the initial FSNAU dashboard model to include more analysis and a stronger link to early action (EA). On the other hand, as noted above, the outcome of analysis can still be confusing. And it is not clear if communities in Somalia are being forewarned and helped to prepare.

So this brief review of EW/EA systems in the Greater Horn of Africa region[8] will proceed along the lines of several questions:

- What EW/EA systems are in place and how are they working?
- What kinds of information are being collected and utilized (with a focus on the missing role or lack of clarity around qualitative information)?
- What is the relationship between baseline, current status, projections of numbers in need, and early warning information?
- What is the link between early warning information, analysis, and early action?
- How are EW and EA accountable to, or engaged with, vulnerable communities they are meant to protect?

[8] Note that a parallel study, conducted by the Feinstein International Center and the Centre for Humanitarian Change, considered the political influences on information and analysis, including but not limited to early warning. Given that parallel study, these notes do not specifically address the politicization of humanitarian information, but most of the problems highlighted by that study remain unaddressed. See the Tufts University webpage on the Constraints and Complexities of Information Analysis research: https://fic.tufts.edu/research-item/the-constraints-and-complexities-of-information-and-analysis/.

3. Note on Methods

This report is based on a review of the literature on early warning, focused on East Africa but including analysis from other sources, and on key informant interviews with individuals working in early warning or food security information systems in East Africa or the users of that information. Thirty-nine interviews were conducted in four countries with some 60 individual key informants. From January to July, numerous processes of humanitarian analysis were also observed in several East African countries. In addition, in another part of the study, some 300 extremely hunger-affected households in three East African countries were interviewed. The purpose of those interviews was not directly related to the EW component of the study; however, the interviews provided an interesting backdrop to the question of the accountability of EW systems not only to donors and government decision makers, but also to affected communities. This study was approved for ethical clearance by the Institutional Review Board of Tufts University.

The report proceeds as follows. The next section (4) discusses existing EW/EA systems in East Africa, noting those at the regional as well as national level. Where appropriate it also notes more localized systems. Section 5 is a brief analysis of issues arising from the interviews, the mapping in Section 4. This constitutes the main section of the report. Section 6 is conclusions and recommendations for change. There is no stand-alone literature review here—appropriate literature is reviewed in the above sections. Annex 1 is devoted to a specific sub-question: the inclusion and analysis of qualitative information in EW systems, or humanitarian information systems more generally. Readers already broadly familiar with EW/EA systems in East Africa might prefer to skip the details of the mapping of regional and national systems, and focus only on the analysis and the conclusions.

4. Review of Current Systems

To understand some of the current issues, this section briefly reviews current systems, both at the regional level and at the level of individual countries for which information could be obtained. This is not intended to be a comprehensive review but rather a thumb-nail sketch to identify and highlight important topics for the analysis.[9]

Regional

Regionally, there are a number of actors. Nominally, the **Regional Food Security and Nutrition Working Group (FSNWG)** is the lead organization. Led by the IGAD Climate Prediction and Applications Centre (ICPAC) and FAO, its role and influence wax and wane, depending on how bad the year is. This year, 2019, has been a bad one, so people are attending meetings and paying close attention. In good rainfall years, it attracts less attention. It is very influential with regional donors and regional agency offices. But it does not have a working website and only communicates via email, so how far beyond Nairobi its reach extends is hard to say. Its eventual aim is to be a "one-stop shop" for information and early warning in the region.

ICPAC is the regional body charged with climate prediction and seasonal early warning. Nearly everyone interviewed noted that climate prediction has been increasingly difficult. The medium-term forecast—the "Greater Horn of Africa Climate Outlook Forum" (GHACOF)—is a probabilistic forecast of the likelihood of rainfall anomalies or failure. But many users tend to use it as if it were an iron-clad prediction—even though it provides percentage estimates of the likelihood of above-average, average, and below-average rainfall. Thus, users expect a good year if the likelihood of above-average rainfall is (even a little bit) higher than the likelihood of average or below-average rainfall. Read this way, the GHACOF has been "wrong" three out of the past four seasons, according to several observers. Indeed, the one component of early warning that used to be viewed as reliable (climate forecasting) is now increasingly doubted by many information users. But whether this is a problem of increased variability in seasonal rainfall outcomes (or climate change, as some observers would note) or is simply a result of information users failing to understand how to interpret probabilistic forecasts (or both) is a matter of debate. The GHACOF has recently changed its methodology in an attempt to improve forecasting.

FAO has an initiative to roll out the "dashboard" approach—first piloted in Somalia by FSNAU—across the region. However, this is still at the drawing-board stage. There are enough issues with the existing dashboard that it is not clear the "approach" is ready for a prime-time, region-wide "rollout." FAO was set to conduct a review of EW systems in the region, but the task got delayed by an increasingly bad season in the first half of 2019, requiring human resources to be deployed to response and advocacy tasks.

The **Integrated Food Security Phase Classification** (IPC) system is a consortium of fifteen agencies, hosted globally by FAO and usually funded as an FAO project in-country. It is now used—at least nominally—by all countries in the region except Eritrea. It was recently introduced in Ethiopia (2019) and is well established in many of the countries in the region. IPC was initially invented as a data amalgamation tool, relying on multiple sources of information but a standardized method for analyzing—and especially classifying—food security status

[9] The FAO regional office has plans to conduct a far more in-depth review and mapping of existing EW systems in the region. This was scheduled for the first half of 2019, but had to be delayed as the main rainy season turned out to be substantially worse than predicted, and efforts focused on ensuring that the EW message got out about the less-than-optimal season. This brief note is not intended to supplant that effort.

by geographic area (usually either livelihood zones or administrative units) based on a variety of sources of information. Increasingly, it has come to rely on—indeed even require—large-scale (frequently, but not always, nation-wide) household surveys of food security and nutrition status, supplemented by SMART surveys for a more in-depth portrait of malnutrition (and in some cases mortality) on a more limited scale. But a degree of confusion exists around IPC's role or products in terms of early warning that boils down to the issue already highlighted above about the difference between current status assessment, projections, and early warning of the threat or impact of new/changing hazards.[10] Note that IPC's primary function is to classify current-status conditions according to severity. The projection function of IPC was initially a secondary product of the analysis, although in recent years the projections have become at least as important an outcome as the current status classifications—in part because by the time the data are collected and analyzed, the results are already out of date. These notes don't address the current-status classification function of IPC, which is a tool for declarations and impartial allocation of resources, not early warning per se.[11]

The **World Bank** has several on-going EW initiatives, some at the national level. But the World Bank FAM initiative (Famine Early Action Mechanism) is focused on at least two countries in the region, and could have implications for more. FAM is partly focused on improved prediction and linking to early action through artificial intelligence-assisted early warning and partly focused on improved contingency planning and financing mechanisms—all as part of a single package, with World Bank leadership in collaboration with national governments. Increasingly, this initiative is also focused on resilience and monitoring the transition from IPC Phase 2 to Phase 3 (as the crisis intervention point). Initially, FAM was focused on generating a clear data signal for famine prediction that could be linked to insurance or disaster-bond type financing instruments, but it is now taking a somewhat broader approach and working with existing systems throughout the region. Somalia is the first of the "first mover" countries to take on the FAM initiative. South Sudan is also a "first mover" country, but the current context in South Sudan may not be as conducive to FAM's approach, which requires at least some degree of government leadership.

OCHA. Following a major speech on anticipatory humanitarian action by the Emergency Relief Coordinator in 2018 (Lowcock 2018), OCHA has invested heavily in improving predictive analytics through a newly formed Center for Humanitarian Data in The Hague and building links to existing financing mechanisms such as the Central Emergency Response Fund (CERF). Recently OCHA and the World Bank have begun collaborating in work on Somalia.

Household Economy Analysis (HEA) is still in use by some agencies and some countries (notably Ethiopia). While formally included in IPC protocols, it is not used very much in contemporary IPC analysis in the region—and it remains to be seen how it will continue to be incorporated in systems functioning in Ethiopia. HEA "outcome analysis" could be helpful for early warning if it were more broadly available. Respondents also note however that HEA outcome analysis is very vulnerable to manipulation by policy makers seeking to alter the numbers.

FEWS NET operates in all of the countries in the region except Eritrea, having staff and an office at the national level in some countries (Kenya, Somalia, South Sudan, Uganda, Ethiopia) and monitoring other countries remotely from a regional office (Rwanda, Burundi, Djibouti, Tanzania). FEWS NET attempts to collaborate with national partners, including governments and IPC teams, while maintaining the independence of its analysis. FEWS NET has long used a "most-likely" scenario approach to its forecasting. It uses baselines against which to measure variations in its short-term predictions, and it uses IPC-compatible classification for its mapping of both current status and predicted outcomes. It also has the longest-range forecasts.

Nutrition. All of the above are either food security classification and prediction or climate prediction mechanisms. Current status assessment for nutrition (SMART surveys)[12] has a much better established

[10] Attempting to sort out this confusion was part of the motivation for this study.

[11] For an in depth analysis of the current classification function, see Maxwell et al., 2019, "Determining Famine: Multi-Dimensional Analysis for the Twenty-First Century" (under submission to *Food Policy*).

[12] SMART stands for "standardized methods for as-

set of norms and practices, but early warning systems for malnutrition, or predicting the prevalence of malnutrition, is a significant deficit. Many systems rely on the rate of new admissions to nutrition programs as the "early warning" indicator, but senior nutritionists in the region point out that many things can lead to a rise in admissions. Even if *all of* these are controlled for, the number of admissions only works if existing programs have very good coverage. So some alternative mechanisms are being explored to predict prevalence of malnutrition—including the use of frequent mass screenings (which has its own problems). Several new initiatives are intended to be able to forecast wasting prevalence based on sophisticated modeling, including the MERIAM program (led by ACF) and a similar initiative led by the London School of Hygiene and Tropical Medicine. Both could substantially improve nutritional early warning, but neither have been rolled out yet. The team at LSHTM is also working on approaches to forecasting mortality based on similar modelling approaches.

A very different approach to forecasting food security outcomes, based on similar predictive modelling principles, was recently piloted by Lentz et al. (2019). Based on publicly available and relatively inexpensive information (prices, rainfall, and population demographics), they demonstrate a vastly improved means of identifying the most badly affected population clusters in Malawi in 2010–11, when compared to the existing EW/EA system in use at the time. Their approach has not yet been incorporated into any EW/EA system but has the same promise for food security outcomes that the MERIAM or LSHTM approach has for predicting the prevalence of global acute malnutrition or mortality. In many ways, predictive modeling seems to be the approach most favored to address some of the shortcomings of current EW approaches, but much of this work is still in its infancy.

There are numerous initiatives at the country level—described briefly below.

sessment of relief and transition." This is the current gold standard for nutrition assessment, but also often includes mortality, food security, health, and other indicators. SMART is strictly a current-status assessment instrument.

Kenya

Kenya has a long-established system for analyzing food security status and determining necessary actions. The Kenya Food Security Meeting (KFSM) consists of high-level actors (donors, government) who take final decisions on actions. The core of the system is the Kenya Food Security Steering Group (KFFSG), which has effectively taken the role that a Food Security and Livelihoods cluster would fill in other countries. It is led by the National Drought Management Authority (NDMA) and includes all relevant government line ministries and departments (agriculture, livestock, health, nutrition, water, etc.) as well as the main UN agencies (FAO, WFP, UNICEF) and FEWS NET. The core of the analytical capacity in KFFSG is DISK (Data and Information Subcommittee of the KFSSG)—which is just NDMA, the big three UN agencies, and FEWS NET.

The early warning system is operated by NDMA in conjunction with county governments—which have been significantly strengthened since devolution in 2013. There are 154 sentinel sites in 23 arid and semi-arid lands (ASAL) counties. Each site tracks 30 households per month, as well as markets and a handful of (3–5) key informants for specialized information. Rainfall, temperature, and normalized difference vegetation index (NDVI) data are also tracked. The Nutrition Information Technical Working Group (NITWG) oversees SMART surveys that feed into seasonal assessments (and nominally into IPC analysis), but these mostly operate independently. The NDMA collects MUAC data in its sentinel sites, and counties and NGOs often conduct mass screening exercises—particularly when they believe the situation might be getting worse. The MUAC findings often don't agree with the results of SMART surveys, but budgets don't allow for greater coverage with SMART surveys.

IPC is used in Kenya, but the established systems largely run on their own criteria and systems, into which IPC is only partially integrated. This may be changing after some recent efforts, but up until 2019, the main sources of information have been the seasonal (long- and short-rains) assessments, the

NDMA sentinel site surveillance information, and the WFP Food Security Outcome Monitoring system (FSOM). FSOM is being discontinued, and efforts are being made to incorporate other data sources and analysis into an IPC-compliant process. The most recent short rains assessment (SRA)—conducted in early 2019 after the short rains of October/November 2018—was deemed to not be IPC compliant because some procedures didn't comply with IPC requirements. This led to something of a crisis between the partners in DISK and the IPC team in Kenya. Accounts vary depending on who one speaks to, but it appears that several issues were raised. The process was not deemed consensus-driven; questions arose about the reliability of the data; data from the NDMA surveillance system didn't meet all of IPC's requirements for reliability; some members of the analysis team were not trained in IPC methodology; and finally, the means of coming up with numbers of people in need and the mapping of IPC outcomes didn't always seem to match. Members of KFSSG/DISK on the other hand noted that the IPC approach amounted to an analytical reversal—in effect, with classification preceding analysis. This situation was addressed by the heads of both FEWS NET and the IPC Global Support Unit. The outcome of these interventions was an IPC analysis in July 2019, but the incident also highlighted the differing views on the role of contextual knowledge and qualitative information in systems that are designed to be run not only on quantitative survey data—which is presumed to be globally comparable.

However, the bottom line is that most of the information generated by the various members of KFSSG is mostly about current status. The actual early warning information is generated by the NDMA's sentinel sites. The information is made available in EW bulletins from NDMA, which while reasonably complete, are based mostly on current information and require some interpretation for actual early warning purposes. However, NDMA does have a coding system that translates into general early warning classification: from "normal" to "alert" (meaning environmental factors like rainfall and water availability are low) to "alarm" (meaning production factors like crops and livestock are not doing well, or market prices are high) to "emergency" (meaning that humanitarian outcomes are bad) and finally to "recovery" (meaning that all factors are subsiding after a bad period).

Up to the July analysis, the IPC classifications for Kenya did not have population tables by phase classification for geographic areas (either livelihood zone or county). The phase classification in the seasonal assessment reports had only a single table showing populations in Phase 3 and above by county. As noted, this changed in 2019. Thus, the combination of information from the short- and long-rain assessment reports, the IPC projections, information from NDMA early warning bulletins, and other sources of information like FEWS NET reports and SMART survey results means that adequate EW information can certainly be found in Kenya. The evaluation of the ECHO response to the Horn of Africa drought of 2016–17 notes that the impetus to early action was not sufficiently speedy, but the lack of EW information was not the cause (Grunewald et al. 2019). Decentralization and devolution have increased responsiveness at local levels in many cases, and several national mechanisms built up in the aftermath of the 2011 drought emergency (the National Drought Contingency Fund and the Hunger Safety Net Programme) built better response capacity. There is enough of a dialogue and an awareness of the general situation in Kenya that key decision makers have a sense of what is happening, but there are still occasional oversights of developing hotspots.

One such incident occurred in March 2019 during a time of increased worry about a nation-wide drought when national newspapers began reporting "hunger deaths" in Turkana county. The government (both national and county) were caught off guard by the reports; both responded by ramping up repair of water infrastructure, water-tankering where necessary, and distributing some food. But the national government insisted all along that the reports were overblown and that all the indicators were within the "normal" ranges. National leaders (including the deputy president), insisted that if any deaths had occurred, they resulted from poverty, not from the drought. Some humanitarian organizations were chastised for saying anything about "hunger deaths."

Subsequent independent research (Centre for Humanitarian Change 2019) in late May indicated that problem was serious, however, and the food relief distributed by the county, while late, did indeed help to reduce a serious food security crisis among the poorest households in the county. No doubt chronic

poverty played a role in the deaths reported in the press, but so too did the deteriorating food security situation due to drought insofar as it weakened already highly stressed social support mechanisms. SMART surveys in Turkana in June and July confirmed levels of global acute malnutrition in excess of 30 percent (IPC Kenya 2019), but it is difficult to say for certain whether there was a failure of early warning in March, or the situation had deteriorated dramatically by late May/early June. However, the newspaper reports certainly triggered some action on the part of the authorities (underlining the role of a free press in a country like Kenya)—actions that CHC research confirmed was very helpful in dealing with hunger at the time.

Several issues are highlighted by the review of Kenya's EW system. The first is data sharing (or the lack of it). Actors have different views on data sharing. Some suggest it is not a problem; others suggest, like other countries, that whoever controls the data controls the narrative on decision-making (and therefore has the strongest influence over resource allocation). SMART survey data are available on request; but this is not always the case for food security information. The second issue, already highlighted, is the extent to which IPC is incorporated into the analysis in Kenya—and to what extent it should be. There is not a strong link between IPC projections and NDMA's EW system. And some observers believe the EW information is not as forward-looking as it could be. While KFSSG and KFSM serve as the link to early action, the communications are often fairly generic.

Third, some approaches to early action avoid EW altogether and simply rely on a "trigger" to activate a response. Several insurance-based approaches have been piloted in Kenya, which essentially tie a single indicator to a response—effectively replacing traditional early warning with index-based triggers. Index-based livestock insurance (IBLI) is one example that insures livestock losses against drought at the level of individual herders. Similar initiatives have been used for crop insurance at the farm level. These are hazard-specific initiatives (i.e., triggered by drought, but not other hazards such as livestock disease or fall army worm, to take the two most obvious hazards). Another approach (at least for agriculture) is area-yield-based micro insurance, which pays out on the basis of reduced yield, regardless of the cause of the reduction. These operate at the micro level. More macro approaches include initiatives like the Africa Risk Capacity initiative, which Kenya had bought into for several years, but which has been discontinued. Some of this work is being drawn together in the form of a National Disaster Risk Financing mechanism led by the World Bank. The Hunger Safety Net Program is intended to deal with the chronic cases that can't be insured by private sector mechanisms but has had mixed success with regard to targeting the chronically vulnerable (Fitzgibbon 2014, Kidd et al. 2017).

Another initiative is being convened informally by the International Centre for Humanitarian Affairs (ICHA)—a research center affiliated with Kenya Red Cross Society. In collaboration with government bodies, it is using a disaster risk-reduction framework to model risk at a local level and amalgamate data to track hazards and outcomes. While a new initiative, it promises to bridge some of the short-term/long-term gaps that have bedeviled other approaches. On the EA side of the equation a recent study found that while preparedness in Kenya is generally high, the ability of mitigation and response programs to adapt to rapidly changing conditions is still limited, and need to be more outcome-focused—whereas they are still more focused on inputs and activities (Obrecht 2019)

These efforts have been initiated to attempt to incorporate the risks of certain hazards, particularly drought, into a "regular" business model and not treat drought as a humanitarian crisis. This has been formal policy since the "Ending Drought Emergencies" initiative was announced in 2012. While promising, these initiatives were not sufficient to prevent the recurrence of a humanitarian emergency due to severe drought in 2017. So there is clearly a lot of activity and innovation in Kenya in the EW/EA space, some need to consolidate the learning and the gains made, and much to build on in a government-led system. Within nutrition and the health system the "surge" approach is now scaling up to cover all the ASAL counties. The approach aims to allow the government health and nutrition system to scale up its service delivery in response to increased demand caused by shocks such as drought impact on food, nutrition, health, and water security.

Ethiopia

Ethiopia has a long-existing national EW system that was linked first to an annual humanitarian response, but since 2006 has been linked first and foremost to the Productive Safety Net Program (PSNP). Ethiopia actually has numerous early warning systems, but the national system, run first by the Relief and Rehabilitation Commission under the Derg, then by the Disaster Preparedness and Prevention Commission under the early EPRDF government, and now the National Disaster Risk Management Commission (NDRMC), has overall responsibility for information and action.

Several early warning tools and systems exist for food security. These include the Livelihoods, Early Assessment and Protection (LEAP) tool and the Livelihood Impact Analysis Sheet (LIAS), which have been developed in Ethiopia (Dreschler and Soer 2016). LEAP is based on drought indicators (including planting date, rainfall, and the Water Requirement Satisfaction Index or WRSI) and their impact on crop production. It can be used to calculate yield reduction in the event of drought (which is the dominant—but by no means only—threat to food insecurity in Ethiopia). Combined with market and price information, the LEAP data is used to calculate beneficiary numbers for both PSNP, and ad hoc humanitarian programs under the national Humanitarian Response Plan (HRP). However, a major limitation of LEAP is "in the use of subjective information in the calculation of beneficiary numbers" (Dreschler and Soer 2016, p. 12). The health sector in Ethiopia has the Public Health Emergency Management (PHEM) system. It is mostly an epidemic information and response system but elements are connected to regular monitoring of non-epidemic morbidities and malnutrition.

The LIAS was developed as an input to HEA outcome analysis and is widely in Ethiopia in the calculation of beneficiary requirements and numbers. Along with the major seasonal assessments, these tools are the major cornerstones of what has come to be accepted at early warning in Ethiopia, and they feed into the Productive Safety Net Programme, which has been documented as an effective and more efficient response to both chronic and transitory food insecurity in Ethiopia (IFRC n.d.) During the major drought crisis in 2011 that affected Ethiopia and was one of the causes of the famine in Somalia, the PSNP was able to scale up to meet the needs of three million additional recipients and avoided the fate of people across the border in Somalia (World Bank 2019).

However, the primary function of providing the requirements and number of projected beneficiaries has confused the role of assessments and early warning. One respondent in Ethiopia noted that "as it now stands, 'early warning' is just the numbers from the seasonal assessments." The combination of the subjective calculation of numbers and the role of the political influences in determining such numbers results in substantial pressure to reconfigure EW in Ethiopia. The national system has been weakened by the retirement of a number of experienced leaders. The NDRMC itself—once reporting directly to the Office of the Prime Minister—now finds itself as a part of the Ministry of Peace, rather more distant from the center of decision-making in the government of Ethiopia.

The ECHO evaluation of response to the Horn of Africa drought in 2016–17 was especially critical of both the slowness of the response in Ethiopia and the extent to which the information system was politicized (Grunewald et al. 2019). Information was available, but often controlled in terms of what could be released and when. The report notes that the system in Ethiopia—even as recently at 2016–17—ran too much on trailing indicators (malnutrition or harvest data) rather than forecasts or the onset of rains; the system for processing data is too slow; and the process is too political—with different actors at different levels having competing interests to either downplay or inflate the figures (Grunewald et al. 2019). The report also notes that an "unofficial" early warning system exists that keeps independent records, passes information by word of mouth, and keeps key actors (especially international donors) informed—a finding that corroborates key informant information from interviews undertaken here.

A major concern is that Ethiopia's entire system is predicated on the assumption that drought is the major driver of food insecurity, and food insecurity the major driver of malnutrition. But in recent years localized conflict has driven substantial levels of internal displacement, which has become a lead-

ing cause of food insecurity alongside drought. But existing EW systems are ill-equipped to analyze conflict—and the government-led system is less able to address it. Among other factors, this has led to a number of NGO-led local EW systems that operate ostensibly alongside the national system and feed into it. Much of the residual capacity for HEA is with Save the Children—although much of it was built up by USAID projects in the 2000s. A consortium of NGOs known as the Joint Emergency Operation (JEOP) also has its own information system, as does Oxfam and a number of other NGOs.

FEWS NET uses an IPC-compatible process to classify current and predicted status in Ethiopia, but until 2019, an IPC Technical Working Group has not existed in Ethiopia. A survey was conducted in Ethiopia in 2019 for the first nation-wide IPC analysis (IPC Ethiopia 2019). Exactly how this will fit into an increasingly complicated EW or humanitarian information system in Ethiopia remains to be seen. In the meantime, the World Bank has been calling for a major assessment and perhaps reconfiguration of EW/EA systems in Ethiopia, building on LEAP and LIAS but recognizing some of the shortcomings of the current system. While identifying mostly technical constraints, the World Bank also notes the political influences within the existing system (World Bank 2019). DFID also has a related program, called Building Resilience in Ethiopia (BRE).

The "subjective" nature of calculating beneficiary numbers reflects a widespread problem with EW systems not only in Ethiopia but more generally in East Africa, and that is the lack of documented and standardized practices for incorporating qualitative information into systems that tend to be dominated by quantitative methods. The role of analytical judgment by human analysts, rather than analysis by machine algorithm, is labeled "subjective" in part because of this lack. Given the heavy dependence on extrapolation, human judgement, and consensus building to come up with needs and numbers, the process is subject to considerable political influence. Several key informants noted this issue, and it is implied in some of the World Bank documentation.

But this conflates two issues—political influences on the one hand and the use of qualitative information and human analytical judgement on the other.[13] The two overlap in this case, but should be separated: politics certainly influences quantitative processes too (Maxwell et al. 2018, Hailey et al. 2018), and irrespective of political influences, the use of qualitative information and human analytical judgment require better guidance (see Annex 1). Differentiating political influence from the role of human analytical judgment is critical in these systems: the former is damaging; the latter is not only valuable, it is absolutely necessary.

Ironically, the response is usually to ramp up (expensive!) quantitative data collection—and indeed this is how some observers interpreted the introduction of a large-scale household survey, conducted between the two major seasonal assessments, built to satisfy IPC quantitative data requirements. SMART surveys, run by the Emergency Nutrition Coordination Unit add quantitative nutrition information but only for a limited number of *woredas* (districts)—often the same ones year after year. To date little effort has been made to systematize the process of human analytical judgment—or the use of qualitative data or its incorporation into quantitative analysis-led processes. But with over 800 *woredas* nation-wide, and with the number of tools (HEA, LEAP, LIAS, IPC, and the seasonal assessments) and actors (NDRMC, WFP, FAO, UNICEF, the cluster system, FEWS NET, Save the Children and a number of other NGOs), the complexity of the information needs, and the "system" (perhaps "eco-system") itself are, in the words of one respondent, "overwhelming." One donor counted at least 20 major actors in the EW/EA arena, with "many stakeholders starting their own system since about 2012." And at this point, the number of actors and processes is still increasing—not consolidating.

For all that, EW/EA systems in Ethiopia have been functioning well enough to activate life-saving responses. Choularton and Krishnamurthy (2019) reviewed the accuracy of FEWS NET forecasts in Ethiopia between 2011 and 2017 in terms of food security outcomes by IPC classification. They found that predictions matched subsequent assessment of food

[13] Note that "subjective data" (e.g., perceptions or preferences, which get used quantitatively all the time) is different from "subjective analysis" (e.g., humans figuring out how to weigh complicated bits of data that can't be fed into an algorithm-driven model)—the latter is referred to here as "analytical judgement."

security outcomes 78 percent of the time, with about half of the errors in prediction being "false optimism" (actual outcomes being worse than predicted) and about half being "false pessimism" (actual outcomes being better than predicted). More importantly, they looked as transitions from IPC Phase 2 to Phase 3 (as indicative of the onset of a crisis) and found that for some areas of the country—notably the pastoral lowlands—the prediction of these transitions was highly accurate. But in the more densely populated and highly vulnerable SSNPR Region, only about half were predicted accurately—implying significant geographic variability in accuracy. FEWS NET is not immediately linked into any early action mechanism, although it informs USAID's response through JEOP and the PSNP. Choularton and Krishnamurthy did not judge which is preferable to avoid—false optimism or false pessimism—although there are clear trade-offs to be made in trying to prevent one or the other.

The NDRMC-led EW system—albeit primarily focused more on current status analysis—was good enough to activate interventions that avoided famine in both 2011 and 2016–17 (although broad areas were badly affected, particularly in 2017). But note that this was not early action to protect livelihoods, particularly in the pastoral lowlands. Some specific projects, for example the PRIME project operating in Somali Region, were able to introduce livelihoods-protecting interventions in 2015–16 (Smith et al. 2018). With shifting demands on the system—and new drivers of food insecurity in Ethiopia, particularly localized or even inter-regional political conflict—it remains to be seen how long this will continue to be the case. Small wonder that there is pressure for consolidation and stronger leadership by government—along with capacity building to restore some of what has been lost over the years. But key informants suggest that some actors have not yet stated a clear commitment to go along with such calls.

Somalia

Somalia has long had the premier capacity for food security and nutrition analysis of any country in the region in the form of the Somalia Food Security and Nutrition Analysis Unit (FSNAU). Operated by FAO-Somalia now, it started out in the 1990s as a joint WFP/FAO operation, but became an independent analysis unit, supervised by FAO later on. In recent years, it has moved back to being more completely within the FAO-Somalia portfolio. It is not an early warning unit per se but rather covers all elements of food security and nutrition analysis, with its major outputs being the semi-annual assessments and SMART surveys. FSNAU was also the birthplace of IPC analysis in the mid-2000s and, overall, has the best historical data set of any country in the world that has faced decades of conflict and crisis.

In the run-up to the famine in 2011—indeed from mid-2010 onwards—FSNAU and FEWS NET, with which FSNAU collaborates closely, put out a series of extraordinary early warning bulletins (meaning beyond their regular reporting) but at the time, no early action mechanism existed that was fully prepared to take on board the increasing seriousness of the warnings (Hillbruner and Moloney 2012). UNICEF and ICRC had access to affected areas and were able to scale up the nutrition response, but WFP had not been operational in South Central Somalia for over a year at that point. And donors had numerous restrictions on aid to Somalia due to the threat of severe legal and reputational consequences for aid that went astray and ended up in the hands of Al-Shabaab (Maxwell and Majid 2016). This combination of both causal factors and extraordinary constraints to the response meant that early warning was mostly not acted on—response didn't really scale up until after the declaration of famine. The impact on South Central Somalia was devastating, but it was also a searing experience for the humanitarian community in Somalia, given the death toll that resulted. Even though information about the impending crisis was not the constraint to early response, several efforts to improve EW and to link it more specifically to early action have since been made.

Perhaps the most significant of the developments in EW has been the FSNAU "dashboard." The dashboard is a different kind of data amalgamation platform. Unlike IPC, which focused heavily on outcome indicators, the intent with the dashboard is to focus on predictive indicators. These fall into fairly standard, recognizable categories: rainfall and vegetative cover (NDVI), market prices (including the price of water—a critical indicator in Somalia) and terms of trade, health indicators, and new admissions to OTPs

as the early warning indicator for nutrition (but see above for caveats about this). Given FSNAU's long-term availability of data in almost all these areas, its staff was able to calculate long-term means and deviations from those means in times of crisis. Thus they flag an indicator as "green" or normal if within usual bounds, yellow or "alert" if somewhat outside those bounds, and red or "alarm" if significantly outside those bounds.[14] These alarms or alerts can be traced by geographic unit of analysis (district) or by individual indicator country wide. The amalgamation function simply counts up the number of alarms and maps them by district, with greater numbers of alarm being depicted by deeper shades of red.

The dashboard was up and running in time to help predict the crisis of 2016–17, and indeed some donors attributed their earlier action in 2016–17 (compared to 2011) largely to the existence of the dashboard. That is not to imply that FSNAU would have somehow "missed" the onset of a crisis in the absence of the dashboard, but the "hard numbers" from the dashboard afforded donor governments greater confidence in the predictions (Dubois et al. 2018). The number of alarms went from a "baseline" (i.e., pre- and post-crisis "normal" levels)[15] number of around 100 to over 400 at the height of the crisis. Even just visually—looking at the map—the color gets distinctly darker in October of 2016 as the *deyr* rains failed and the extent of the crisis worsened significantly.

Still despite this relative success story, a number of issues were raised after the crisis about the dashboard approach. These were summarized by Oxfam (2017):

- **Lack of clarity about objective.** Was it to intended to offer a forecast or prediction—an analysis on which early action could be built—or simply to be a trigger for EA? This is summarized in the "signal" versus "scenario" discussion below, but this is a perfect example of the question.
- **Ensuring timely information.** The dashboard, like any data amalgamation platform, relies on data being submitted in a timely way—not all the data in the dashboard is generated by FSNAU itself. Late submission of data hampered the usefulness of the dashboard, even though it got the prediction mostly right
- **Ensuring data is accessible and understood.** It is one thing to count up alarms, but it is another to understand what they actually mean. (The dashboard simply presents a total number of alerts and maps them by district—it does not offer any narrative scenario. Scenario analysis may be found in other reports or bulletins from FSNAU or FEWS NET, who cooperate closely in Somalia.)
- **Building buy-in.** FSNAU led the process, but other actors had varying levels of engagement with both data inputs and analysis of the outputs. Critical to sort out here was the role of government.
- **Ensuring early action.** FSNAU is simply an information and analysis unit—it does not operate any programs or interventions. So the link to EA was still tenuous and, some believed, vague.
- And finally, the platform was **gender blind**—by not including any gender-disaggregated data or, for that matter, data disaggregated by any other social category. (The latter, however is a criticism that could be leveled against any of the systems reviewed, not just Somalia's.)

The Oxfam report stimulated significant discussion within the Somalia humanitarian community, and eventually a stakeholder consultation was held in mid-2018 (FSNAU 2018). The main upshot was to improve the analysis afforded by the dashboard, clarify the notion of triggers, and build an accountability framework for early action and an enforcement mechanism through the Inter Cluster Coordination Group in Somalia (FSNAU, Oxfam, and BRCiS 2018). By early 2019 and up to the time of writing, a task force has been examining these issues. The role of government was not clarified—indeed the debate over how to engage with, or even embed within, government was a discussion that went on through-

[14] Initially, these were percentage deviations from the long-term mean. Recent discussions to update the dashboard have focused on standard deviations or z-scores, but a final decision on this is still pending as of this writing.

[15] Observations like this raise the inevitable question of whether such usual levels of alert in Somalia perhaps mean that the thresholds for alerts are set too low, or that they simply reiterate the "normalization of crisis" in Somalia. Unfortunately, this research did not shed any light on that question.

out 2018-19 and involved not just the dashboard but indeed all operations of FSNAU. Part of the question was not whether to build greater government engagement, but with which unit of government (federal government or member state—and within either of those levels, which of potentially five or six different ministries).

A question not raised in the Oxfam report or the workshop was about intended users. The dashboard gives a "big picture" snapshot of what is happening in Somalia, but at the more granular level, it can be difficult to interpret. In early 2019, for example, only a handful of districts in the far northwest were in IPC Phase 4, and their overall status—as depicted by a dark red color on the dashboard map—matched their bright red color on the IPC map. But while the number of indicators that added up to that shade of red was relatively constant (in the range of four to six per month), the individual indicators varied from month to month, and it was difficult to detect any real trend in the indicators or and suggestion of what the actual problem was. If five alarms was the trigger, clearly action was called for—but it was not really clear what the problem was or what the response should be. So the question of analysis of the indicators needed to be addressed—back to the issue about offering some kind of "likely scenario" based on the amalgamated indicators. This also raised the question about who the dashboard was for. Initially it was intended for high-level response (UN humanitarian country team and donors); it was not intended for a local-level humanitarian programmer—whether or not it serves the population of affected districts is not clear.

A second issue not raised was the question of which of the array of indicators were genuinely forward looking and which simply depicted the real-time situation but were not necessarily predictive? A third issue is that, unlike IPC, the dashboard does not take into account existing humanitarian programming and the mitigating impact that this may have on the situation. And finally there was the issue that all the "alarms" are currently weighted equally; but, for example, should the price of maize be weighted equally with the price of water? Does a deviation in rainfall during the dry season "count" as much as a deviation during the (expected) rainy season? These questions were being addressed by the task force at the time of writing. The dashboard continues to amalgamate information even as these discussions go on, and its advocates continue to insist—correctly—that despite some of its shortcomings the dashboard does give good information at relatively low cost. Addressing the issues raised above will improve on that record.

Given the in-depth analysis of the IPC process in Somalia just conducted in a separate report (Hailey et al. 2018) this report will not reiterate those findings. However, two points are worth noting: First, the question of government engagement is similar. And second, some of the lessons learned about participation and inclusion from the Hailey et al. study have been taken on board by FSNAU—the process around revising and strengthening the dashboard has been much more broad-based and participatory than was found related to IPC analysis in 2018.

As noted in various country summaries, in many countries NGOs or other actors operate more localized or sub-national EW systems. A particularly notable one in Somalia is operated by the BRCiS consortium (Building Resilient Communities in Somalia). Although the actual information categories may be similar, these systems (using the BRCiS model as an example, albeit perhaps relatively more sophisticated than most) are different in a couple of ways. First, the granularity is much greater than national systems can afford—giving a much clearer picture of local dynamics. Second, the link to programming is much clearer since the analysts and the program staff are often in the same office—or at least the same organization. Third, and perhaps most critically, there is much greater participation from the at-risk community—in terms of both deciding what information to include and informing the community what the risk of a shock is so that they can prepare as well (not just the humanitarian agencies). These systems almost always have a community preparedness and contingency plan alongside the EW component and the early action planning of the agency (BRCiS 2019)

However, the sustainability of these smaller, usually NGO-led, systems can sometimes be a constraint (not all donors are as generous in this regard as BRCiS' donor is). And frequently questions arise about how information from localized information systems can be fed into national systems—and indeed frequent discussions occur about how the BRCiS system can interface with the dashboard.

South Sudan

For most of its existence, South Sudan has not had an actual early warning system. HEA was utilized heavily by WFP and Operation Lifeline Sudan in humanitarian operations during the civil war—particularly from the mid-1990s onwards, but was largely discontinued after the Comprehensive Peace Agreement in 2005. The IPC approach was adopted early on, managed by the then South Sudan Center for Census, Statistics, and Evaluation (now the National Bureau of Statistics), and this continues today. FAO has had two food security information system projects—formerly SIFSIA and currently AFIA (or CLiMIS as the website is titled)—that track prices, weather, livestock, and food security indicators. But neither of these is actually an early warning system that tracks the development of crises or provides probabilistic information about likely scenarios. As with other countries in the region, both WFP VAM and FEWS NET operate in South Sudan and do provide some information, but for the most part, donors and agencies rely on IPC projections for information about future trends.

IPC processes in South Sudan have been the source of some concern for several years. The analyses have been regularized and are now conducted twice yearly. The government leads on IPC judgement coordination, but this has politicized the discussion of food security in general and nowhere has this played out more powerfully than in the case of IPC projections—or in other words, in what passes for early warning in South Sudan. This issue has been reviewed elsewhere in detail (Maxwell et al. 2018) and will not be rehashed here. But suffice it to note that several of the issues highlighted by the South Sudan case are by no means unique to South Sudan. These issues have been noted and attempts are being made to address them.

The first concern is that projections have sometimes missed major deteriorations, at least in part because of the shortcomings in conflict analysis and the politicization of the process. Analysis in South Sudan continues to operate on the assumption that seasonality and climate are the major drivers of food insecurity, but nearly all the major outbreaks of crisis since 2014 have been driven by conflict, and there are strong indications that at least some of these crises were deliberately caused as a strategy of war (Anei et al. 2019). Aside from mentioning conflict as a "contributing factor," conflict analysis has often been only minimally incorporated formally into IPC classification in South Sudan. Attempts to increase its incorporation have been limited by government leadership (needless to say, many of accusations of Anei et al. about the use of food or starvation as a weapon are directed at this same government as well as some opposition groups). This is not to say that there are no conflict analysis or conflict EW mechanisms in South Sudan—it is simply to say that, beyond assumptions about seasonal patterns of different forms of violence, they cannot be systematically included in public IPC analysis.

Second, much of the information needed has often been missing or out of date. Mortality data is a huge concern in this regard (Maxwell et al. 2018). Several alternative methods are being experimented with, including rapid mortality surveys and the "capture-recapture" methods relying on key informant interviews. Nutrition information is provided both by SMART surveys and more recently by the large-scale Food Security and Nutrition Monitoring System (FSNMS) surveys, but early warning for acute malnutrition also relies on OTP admissions rates—with all the problems outlined above.

All of this has led to the development of alternative approaches—including but by no means limited to food security and nutrition. The most ambitious of these is the Integrated Needs Tracking (INT) system, funded by DFID—and thus with a strong link to the FSNAU dashboard initiative, but unlike the FSNAU dashboard, implemented by REACH (an NGO), not FAO. In terms of data amalgamation, the INT looks a lot like the FSNAU dashboard, albeit with a somewhat more complicated theory behind it and a greater number of indicators tracked—including food access, livelihoods, market prices, agriculture, health, WASH, nutrition, and mortality. It then classifies counties by the number of indicators that have been "triggered" and a five-level classification system depending on the severity. These are then mapped by trigger, with darker colors reflecting the severity. Unlike the dashboard, the INT does not claim to be an early warning mechanism.

The innovative part of the INT initiative is not so much the data amalgamation platform itself but rather its link to analysis and action. A Needs Assessment Working Group (NAWG) meets twice a month and maintains a "hotspot" list—locations (at the county level or perhaps more locally) of places where the situation is either deteriorating or very serious. This list is reviewed for new information every two weeks with an update issued—and perhaps removed from the list if the situation has improved. At the same time, based on the INT information, other areas may be added. The point is that an updated, consensus analysis is produced every two weeks. This updated analysis is then shared with the Inter-Cluster Working Group (ICWG), which is the decision-making body that decides on the appropriate action to take. Under current circumstance, and given the severity of the situation, much of this is response, not necessarily mitigation, but the possibility for both is built into the system. So INT is the data amalgamation platform, NAWG is the analysis and hotspot identification mechanism, and ICWG is the early action mechanism.

So in theory, even though this system doesn't claim to be an early warning system, it has the components needed to link analysis to rapid action in a complex emergency. Two concerns, however, arise. The first is that, unlike IPC, this is UN-led and completely run within the cluster system, and to date, government has not been a participant. On the one hand, this has made for a much more in-depth discussion of the drivers of the humanitarian crisis—particularly the conflict drivers. But the price paid for this more open discussion is that the key actor (government) is not present. It remains to be seen if the government will even let the system keep running. Individual agencies operating on the basis of the information and analysis generated by INT and NAWG reports separately to the SSRRC—the government body responsible for coordinating humanitarian response—so in theory government is still informed about actions. But the tenuous relationship with government is a concern.

The second concern with the INT system is simply the institutional complexity. A number of different agencies help to feed data into the INT (and, as in other countries, problems with data sharing have arisen); the INT/ NAWG/ ICWG nexus has so far run more on the strength of personal relationships and commitments than on institutionalized linkages; and some of the outward linkages go well beyond humanitarian mitigation and response. How sustainable such tenuous, but nevertheless critical, linkages are in the medium to long term remains to be seen. As of mid-2019 they were working, but again, mostly on the initiative of committed individuals, not necessarily institutions.

Uganda (Karamoja)

Uganda does not have a national EW/EA system. The focus of the systems that have existed in the past has mostly been Karamoja, in the northeast of the country. And there, despite intra-communal conflict and, at times, some conflict with the state, most of the EW systems have focused on drought, not conflict. The DEWS (Drought Early Warning System) was initially led by ACTED, but has been run by the Ministry of Agriculture, Animal Industries, and Fisheries (MAAIF) but includes participation by NGOs and UN agencies (IFRC n.d.). An unusual feature of the original DEWS was that it broadcast its messages on radio in local languages—something that most EW systems do not do. DEWS has now been closed down, although several options for renewed EW analysis are on the table.

FEWS NET operates in Uganda as in many other countries in the region, and CEWARN (the regional conflict early warning mechanism) still nominally operates in the "Karamoja cluster." Uganda Red Cross operates an EW/EA system independently in the Karamoja and Teso regions of Uganda.

A proposal to revamp early warning for Karamoja is under consideration. Led by WFP, the new system would focus on food security and nutrition, and set up triggers for early action; seek to explore the limitations of EW in informing nutrition programming; and present a finalized proposal to government and donors by mid- to late 2019 (WFP Uganda 2019).

5. Thematic Analysis of Information/Action Systems in East Africa

The following constitutes some analysis of the current status of early warning/early action systems in East Africa and how these need to evolve in the intermediate future. This is not intended to be an exhaustive review but rather a reflection on issues arising from reviewing the literature and especially from interviewing key informants. There is widespread agreement that something of an "early warning moment" is occurring now in East Africa. The general sense—across at least five countries—is that although progress has been made in some areas, systems aren't functioning as well as they could be, are not as well coordinated as they need to be, may be displaced by more automated systems, etc. For these reasons, the current moment also offers the potential to rethink systems, try to improve them, try to build in better accountability for linking to early action, and achieve better results. Some of this is already happening, as noted above. The recent crisis in 2016–17 was a bit of a wake-up call after five years of relatively favorable weather patterns—and thus perhaps a bit of complacency, at least in countries where drought is the primary hazard. While early action kicked into gear earlier in the 2016–17 crisis than it had in the 2010–11 famine, it was not early enough—and the lingering sense was, at least in Somalia, that even the earlier action that did ramp up was more the result of the efforts of specific individuals than of systemic improvements.

Nationally led systems have been under some stress in Kenya and Ethiopia. Kenya's system experienced some turmoil early in 2019—but it appears that this has led to a more nimble system. In Somalia, work is under way to make the FSNAU dashboard more nuanced and analytical, more participatory, and linked in more accountably with early action. As noted, South Sudan never really had an EW system, so INT/NAWG/ICWG is a major innovation though its sustainability remains a concern. So—there is a sense of existing systems struggling a bit and new activities coming along. But there is also a sense that some of this is going on in siloes—not necessarily linking up or progressing towards an agreed-upon set of objectives at the national level—and certainly not at the regional level.

Regionally, the sense is that early warning is becoming more difficult. The Regional Conflict Early Warning Project (CEWARN) exists on paper but is not very active in actual early warning terms (its most recent reports date from the 2010–2012 era), and existing systems note that they are struggling even with climatic early warning. To say the least, this is problematic in a region where conflict and drought are the major drivers of humanitarian crises. Recent research makes a strong case for the imperative of both resilience programming and early action—at least in drought emergencies—and that significant improvement in the EW/EA side of that equation is the first step to be taken (Cabot-Venton et al. 2012). Significant resilience gains have undeniably been made—at least in Kenya and in parts of Somalia and Ethiopia—but how much of this is because of better information and how much is just because of investment in much bigger budgets is not clear.

With regard to early warning, some agreement (but by no means consensus) seems to be converging around the following thumbnail sketch: early warning is dominated by food security analysis (OCHA is now trying to diversify beyond food security). Food security analysis has been mostly dominated by IPC. IPC analysis, despite some of its intent or potential capabilities, is still most applicable to current status assessment. IPC includes projections but, as stressed above, projections are only one kind of early warning. Many observers believe the projections are

the weakest link in the chain of IPC analysis—too focused on the numbers rather than the drivers and not sufficiently anticipatory in terms of early action. The new IPC Technical Manual (V.3) includes scenario planning for early warning, but emphasis is still on quantitative survey data regarding outcomes with a relatively high bar for the reliability of *current status* outcome data required for projections. None of this is meant as an attack on the IPC—it is simply an acknowledgement that the real strength of the IPC is current status analysis, and the projected numbers are one—but only one—early warning tool. And as noted above, to be more genuinely useful, projections need to be updated more frequently than they currently are.

This brief sketch is at least partly where the issue around qualitative information arises. Much of the information on "contributing factors" varies from absolute hearsay to well-triangulated assessment information, but is in the form of narrative notes, not numbers. Information represented by these two extremes—and everything in between—is classified as qualitative information (or sometimes "informal information"). Developing and implementing much stronger guidance for qualitative information and analysis is important across the boards (Annex 1).

These are not the only issues. Additional concerns are that, for example, some loss of capacity has been noted in national systems—partly because of retirement or "brain drain." In many countries, there are multiple players who are at best are coordinated only to a limited extent—everybody seems to want to control the early warning "space," but there is often less appetite for collaboration and no sign of different parties converging around a single system. IPC often presents itself as the system around which others should converge, but while IPC is definitely one of the strong actors, there is no consensus about early warning converging around IPC. In the meantime, numerous initiatives are ongoing, but mostly not resulting in a joined up analysis.

However, at least partly as a result of these trends, efforts are being made in every country visited to update and improve: In Somalia, the effort to improve the "dashboard" and its links to early action was slow to get going but gained momentum in 2019. Strengthening links to both scenario analysis and action will be important. In South Sudan, the INT/NAWG/ICWG initiative shows real potential (if also some worries about political and institutional sustainability). In Kenya and Ethiopia, several initiatives are underway to improve existing systems. And in Uganda, an effort is being made to set up a coordinated EW/EA system for Karamoja and Teso Regions. Greater nimbleness in early action and response is critical across the boards but beyond the ability of information systems alone to improve.

Some of the big themes emerging (in no particular order) are enumerated below:

1. In general, awareness of the early warning/early action gap is widespread, and everyone is talking about how to link data and analysis to judgment and action. Everyone wants to build in the link, but how to do this is neither clear nor streamlined:
    - Some want to build in "semi-automated" triggers from early warning data that will be linked directly to specific funding mechanisms and actions (the World Bank FAM initiative, or the OCHA link to the CERF for example).
    - Some want to leave EW out altogether and tie one single indicator (such as rainfall or crop yield) to preset insurance payouts (the index-based livestock insurance initiative or crop insurance).
    - Some are trying to build in relatively complex human judgment processes (for example the South Sudan INT linking to the Needs Assessment Working group linking to the Intercluster Working Group).
2. These different ideas can be summarized by the "signal" (a reliable trigger that sets in motion financing and intervention plans) versus "scenario" (an analysis of the problem and likely consequences) debate. "Signal" and "scenario" each have a role because any early warning/early action system has to accomplish several things:
    - Be light and flexible—adapt to unanticipated situations. This is the problem with "automatic" triggers—they may work for known and predictable hazards with known and predictable responses. They might not work very well for the "unknown unknowns" (like

unforeseen military offensives, militarized "cattle raiding" that displaces human populations, or extreme limits on humanitarian access). There could also be factors such as the ability of the affected to cope with the scale of deterioration in a crisis, or even the counter-terrorism restrictions that precluded early action in response to the Somalia famine in 2019, despite the fact that good early warning information was available. The point is that some things may not be predictable but can be monitored.

- Include "levels" or "triggers" but these have to be linked first to specific analysis and judgement (perhaps more in-depth analysis very rapidly) and link directly to human judgment and decision-making.
- Have a built-in accountability system. This means first and foremost accountability to affected communities (i.e., not just expert review but community consultation in judging the results of EW/EA), and accountability for learning from mistakes.
- Be linked with "no regrets" programming and "crisis modifiers"—as well as other options such as scalable safety nets and other forms of forecast-based early action. No single one of these has come to dominate, and there is no reason why all can't be used. The surge approach perhaps requires more institutionalized capacity in existing health care systems. Some donors are still cautious about "no regrets" programming or "crisis modifiers" that leave the decision-making at the local level because needs may be greater outside of the field of analysis of local-level decision makers. But at the same time, there is strong evidence that centralized decision-making systems are slower to respond (Grunewald et al. 2019).

3. Conflict early warning (and the ability to analyze conflict) remains a weak link in most early warning and information systems in the region. Recent efforts have been made to improve the incorporation of conflict into the analysis but, for the most part, conflict is left as a "contributing factor" rather than as a robust EW variable. This is clearly a major shortcoming in a region as wracked by violence as East Africa. But several caveats on this:
   - Predicting conflict may appear to be more difficult than predicting food security, but this may also simply reflect professional biases. Certainly groups are working on this with whom greater collaboration could be sought.
   - Increasingly good public datasets are available if people use them. Part of the issue is predicting the effects of conflict—not just predicting the onset of conflict itself. But the effects of conflict are not always predictable. One paper from Somalia directly ties higher levels of violence to higher prevalence of child malnutrition (Kinyoki et al. 2017) but another analysis of food insecurity finds the opposite relationship—conflict is correlated with lower levels of food insecurity (R. Choularton, personal communication). This underlines the need for more in-depth, qualitative analysis.
   - Conflict analysis of course is highly political—systems have to find a way to deal with the politics of incorporating conflict analysis into early warning if they are going to work. This has to be addressed, but is not easy in government-led systems where that same government is party to the conflict being analyzed (the case in several countries in the region). This problem requires leadership from the highest levels in the humanitarian community: Humanitarian country teams and global agency leaders—and may require independent analysis.
4. The lack of data sharing is a major constraint to good analysis everywhere! Humanitarian data should be treated as a public good and must be made available. At present, much of the data on which humanitarian analysis depends is kept private by the agency collecting it and only made public after its effective shelf-life for current analysis and relevant early warning has expired. Maintaining control over the data to some degree enables maintaining control over the "narrative" of early warning and crisis. Data shar-

ing, joint cleaning and analysis, and transparent procedures (instrument vetting, data collection, data entry) have to be standardized across all sectors of humanitarian assessment, analysis, M&E, and learning. The Nutrition Cluster in almost all countries has a means of ensuring data vetting and sharing. There is no reason why the rest of the humanitarian community cannot learn from nutrition. Donors have a strong role to play here, and strong incentives may be required.

5. The strong emphasis on systematization and institutionalization puts the cart ahead of the horse. While there are many sources of information at the moment, different systems or approaches are competing to show which one works best—but securing the "mantle" of institutionalization is taken as the way of "proving" performance. Greater cooperation among competing systems is a huge need—not competing between or simply ignoring other systems as characterizes some of the current situation. Much of this need for greater cooperation is horizontal—currently different systems are trying to do the same or similar things while ignoring all the others trying to do those things. But greater vertical integration systems operating at different levels, or different foci, are also needed. (For example, BRCiS focuses very much at the community or "area" level but has difficulty in linking up to FSNAU's systems.) And finally, of course, the greatest need for collaboration is on early action.

6. An early warning/early action system has several different and important needs, and these differences should be understood: the difference between "outcomes" (already-existing status or "hard data"—as per a seasonal assessment or an IPC current status analysis) and "forecasts" (probabilistic information about the future—which is what early warning is about). The projected number of people in need, as judged by an IPC analysis or a seasonal assessment is one form of early warning because it gives donors an estimate of needs in the coming season. But this is very different from monitoring to detect impending shocks or detecting developing "hotspots." Both are needed. Likewise, the difference between "triggering" events and "causes" (or "drivers") needs to be recognized (and of course—the term "trigger" is used in two rather different ways in this discussion as well: triggering a crisis as well as triggering a response). A baseline analysis of risk and potential hazards, linked to the means of managing risks and preventing shocks, is one component of a good early warning/early action system (perhaps the ICHA/government of Kenya initiative is the best example of this). A flexible and nimble early warning component, tied to targeted rapid appraisal and contingency planning that mitigates the impact of a shock and allocates resources accurately is a second component (perhaps the INT system is a good example of a foundation from which to build). And a needs assessment function that analyzes the impact of a shock, the people impacted, how badly and for how long is a third component (IPC continues to fulfill this role, as do seasonal assessments in Kenya and Ethiopia). Note that in each case, the information component is tied to a particular—but distinct—kind of early action. The amount of data collected over yearly assessments means that many systems now have a very accurate sense of the range of variation in indicators—indeed FSNAU's dashboard is built on the observation of these ranges.

7. A separate but similar issue is the way in which humanitarian action (response or programs) affect both current status and forecasts. IPC addresses this in terms of estimating the impact of humanitarian food assistance on currently observed and projected phase classifications, though these have proven to be fraught with difficulty (the projection maps with and without humanitarian food assistance, for example). But for the most part, other systems reviewed here have not taken on board the issue of the impact of humanitarian assistance on projected or forecasted outcomes, and even IPC is heavily focused primarily on one type of humanitarian assistance.

8. The role of government depends on the country. Governments have multiple roles in these processes, including data gathering, convening or overseeing analysis, and in some cases taking the lead on early action. How much each aspect is influenced by political considerations

varies substantially from country to country. On the one hand, most stakeholders agree with the standard view that early warning and humanitarian information systems should be run by governments, but quietly this notion is resisted in a number of contexts (some for good reason). In South Sudan, government plays a strong role in IPC but less so in other initiatives. In Somalia, so many branches of government exist—federal, state, etc.—that the role of "government" is atomized and difficult to characterize. In Kenya, there is a history of strong government leadership of the information and decision-making system, but some observers have suggested that other agencies have come to dominate the system. In Ethiopia, there is also a long history of a government-led information system—which may have lost some capacity and seems to have less and less control over a sprawling array of different systems. Ideally, government would be in the lead for all EW systems in the region, but that is clearly not always the case and not necessarily desirable in some cases. While governments should—as a rule—lead, a more deliberate system is needed to hold governments to account for their use of information (and on how to manage humanitarian information systems in countries with significant internal conflict).

9. Political influences need to be better managed. Throughout the region, political influences shape humanitarian information systems. This may be governmental politics, but agency and donor politics can play a role in this as well. Numbers may be inflated or deflated depending on the politics of the moment. Causes or contributing factors can be downplayed or highlighted because of political considerations. Agencies and individuals can be threatened or attacked for being too forthcoming with information that does not put governments or donors in a good light. All of this makes for a challenging environment for systems that would simply generate evidence for decision-making.[16] While isolating humanitarian information systems from politics entirely is an unrealistic goal, the political influences need to be better managed. Differentiating political influence from human analytical judgement is critical—politics can influence algorithmic processes just as much as it can human judgment processes. Both need to be isolated from political influence to the extent possible.

10. Sophisticated computational and econometric modeling and artificial intelligence are definitely in the near-term future for some components of early warning, but they are not here yet. Increasingly sophisticated means of modelling—and therefore predicting—food insecurity, malnutrition, and mortality are being developed. These will have major data requirements, so on-the-ground information systems are still needed (and will always be needed). Several important developments include the following:

    - Three-D photography is a promising technology for use in anthropometric measurement (rather than weighing and measuring the length of children). This technology has great promise but is also problematic from the perspective of protection of privacy.
    - Capacity to bring heretofore incompatible data sets together into an integrated analysis through improved algorithms is increasingly deployed.
    - On-line coaching involving computer prompts to a human analyst to not forget steps in an analysis process (such as codifying guidance for how to conduct IPC analysis) is a good next step.
    - Computational modelling is being conducted by at least half a dozen teams or institutions working on sophisticated modeling—some on food security, some related to malnutrition and mortality, some specifically on the prediction of famine or epidemics. But much of this is very siloed with little cross-fertilization between these teams.

[16] These have all been documented elsewhere. See the Tufts University webpage on the Constraints and Complexities of Information Analysis research (https://fic.tufts.edu/research-item/the-constraints-and-complexities-of-information-and-analysis/) for a series of individual country case studies on this. A synthesis of this study is due in early 2020.

- Artificial intelligence is being employed by the World Bank FAM initiative working on AI-assisted analytical processes and prediction. The FAM initiative has also couched the need to complement human decisions (not replace them) in its system as well.

11. Whether these new technologies are being developed to address a specific problem or just because they can be is not always clear. The role of both traditional early warning and human judgment and/or contextual knowledge in combination with these sophisticated modeling approaches is currently hotly debated. All of this increasingly recognizes the need for ethical guidelines on the use of AI and modeling. One observer described AI as the "wild, wild west right now." Putting some of these tools in the hands of governments engaged in civil wars with parts of their own populations raises serious ethical questions that, so far, few advocates have addressed.
12. At the same time, big gaps remain regarding the question of human judgement and especially with qualitative data (though as just noted, even the most sophisticated of these systems has recently recognized the need to incorporate human judgment into highly automated systems). One of the gaps is the lack of guidance. Another is the lack of means to adequately and publicly explain the analytical judgments that inevitably go into an analysis (and especially a forecast or projection).
13. As noted at the very beginning of this review, some major concerns have completely fallen through the cracks.
    - The most obvious—back to the Practical Action definition of early warning—is about exactly how and where any of these initiatives (whether traditional EW or computational modeling and AI) intersect with local realities and inform community action to prepare for or protect against shocks and hazards. These all need to be emphasized or, in some cases, brought back.
    - A second is that a previously used feature has since mostly fallen out of these systems: the whole notion of coping, and especially social connectedness. Relying on existing social networks is almost universally the first, and sometimes the only, kind of response that vulnerable households receive.
    - A third is that gender differences and other differences such as age or disability are rarely incorporated into forecasting or analyzing shocks.
    - A fourth, growing out of the above, is that EW/EA should move beyond food security and nutrition and broaden the range of outcomes to include, particularly, WASH and health outcomes but perhaps others as well.

6. Conclusions and Recommendations

Several recommendations grow out of these observations. Some have already been made but will be repeated here. Some of these recommendations are about *what* needs to be done; others are about ***how*** to do things differently. These are in no particular order and the ordering does not imply any prioritization.

1. **Focus on key issues, not institutions**. It makes little sense to scrap the systems we now have to start over from scratch. The key question is how to clearly differentiate between current status, projected numbers, and the ability to rapidly identify deteriorating situations. Different kinds of information and different kinds of instruments may have different links to early action. Mechanisms exist for all of these, but they are often confused by users.
2. **Think strategically about components of a "system."** Projections provide one kind of information for early warning. Others are the monitoring of risk and hazards and the incorporation of a rapidly changing landscape into an analysis of what the future may look like. Some components of a system may be in the form of triggers, other components might consist of a fleshed out scenario. Where does this come together at a system level? The quest for rigor may drive analysis back towards existing conditions, but consideration of rapidly changing causal factors may be more important for anticipating future problems. Early warning will always be probabilistic.
3. **Build better linkages.** Greater integration is needed, both horizontally (between systems) and vertically (across levels or time frames). Better linkages need to be built between information systems, resource allocation for decision-making, and program planning.
    - One system or approach might focus on the longer-term identification of risk and link that to risk reduction or risk management activities. Another system or approach might focus more in immediate outcomes and link them to crisis modifiers, no regrets programming, etc. that don't necessary attempt to reduce risk, but rather to mitigate a hazard. These kinds of analytical foci can and should be brought closer together in the same analytical system—and linked to the same response framework.
    - There are different levels (or perhaps different scales) of geographic/population granularity, including the global, regional, and national levels; the program level; and the community level.
    - Finally there are multiple ways to get to early action:

        semi-automated or index based systems

        parametric "dashboards"

        traditional EW indicators

        qualitative information

        human judgment

        current-status surveys

    All of these need to be integrated—or at least efforts need to be made to break down existing "siloes" of information and analysis.
4. **Take a broader view of crisis and risk**. All systems are now heavily slanted towards the analysis of the severity of current status and the risk of future severity. Much more consideration needs to be given to the magnitude of the problem, as well as temporal and spatial dimensions of the problem (Maxwell et al. forthcoming 2020).
5. **Build better mechanisms for accountability.** Accountability for the use of evidence *and* for

early action need to be built into systems. So rather than just checking for the accuracy of the forecasts (were they correct?), check for the information and the action (what was put in motion?), its impact (did the action protect affected communities), and the learning that resulted.

6. **Broaden the scope of information**. To provide a more holistic understanding, a wider range of measures needs to be incorporated into existing information systems. These include coping and social connectedness, along with better information on WASH and health outcomes and a much stronger focus on causal factors. Better guidance is urgently needed for how to utilize qualitative information.

7. **Treat humanitarian information as a public good.** Humanitarian information is often not available for users and analysts to see while it still has early warning and programmatic value. Changing this is imperative, but would require considerations of data privacy, who has first right of analysis, etc., but all these could be worked out. Establishing real-time access to humanitarian information will require different approaches in different places. In some places, governments are the ones keeping a tight lid on information; in other places it is the agencies that collect (and "safeguard") the information. Donors can easily make this a requirement.

8. **Develop better methods for dealing with politics**.[17] As noted, these can include government politics, but often also can include donor politics, humanitarian agency politics, and rivalries between information systems. Big concerns include the following:

    - The politics of numbers. The number of "people in need" is a huge issue in some countries, with numbers being inflated or deflated depending on politics.
    - The politics of famine. The use of the "f-word" is contentious nearly everywhere.
    - Accusations about undermining national sovereignty. If agencies declare emergencies (let alone famines!) before governments do, accusations of violating sovereignty arise; if agencies feel that governments are slow to declare emergencies, accusations about the humanitarian imperative and the lack of early action arise.
    - The politics of analyzing conflict. This becomes sensitive particularly if there is any hint of atrocities having been committed.

9. **Incorporate conflict analysis more systematically into humanitarian information systems**. Conflict is a very common driver of crisis across the region. Conflict early warning is a field in its own right—it should be more systematically incorporated into humanitarian early warning, and its information fed into humanitarian scenarios and contingency plans.

10. **Clarify the role of government leadership**. Nearly all parties believe that governments should lead on information systems, but this is problematic when non-state actors control much of the affected population, or when government is one party to a conflict that is driving the humanitarian emergency. A comprehensive review of the assumption that governments should always lead is needed. This brings up the inevitable question of sovereignty and includes waiting for governments to declare an emergency before acting (i.e., sovereignty versus the humanitarian imperative—see Grunewald et al. 2019). Principles guiding the relationship between the humanitarian community and government have to include cases that work well (such as independent government bodies, like NDMA, that are government-run but relatively shielded from political influences) and cases that don't work so well (such as government control over information systems in contexts where the government is party to the conflict that is driving food insecurity).

11. **Improve collaboration between multiple information and EW initiatives**. With so many initiatives going on currently, an urgent need is to get these different initiatives talking to each other. Many are attempting to address the same

[17] Again, this is dealt with in a separate study. See the Tufts University webpage on the Constraints and Complexities of Information Analysis research: https://fic.tufts.edu/research-item/the-constraints-and-complexities-of-information-and-analysis/.

objectives or achieve the same outcomes but are unaware of the work that others are doing. Some are working on "traditional" early warning or trying to improve the quality of IPC projections. Some are working with newly available remote sensing data or predictive modeling or artificial intelligence. There is little doubt that, in the medium term, all of these approaches are going to have to work together to improve overall outcomes. The sooner the dialogue across and between these different initiatives can be facilitated, the better. A couple of key questions should guide the dialogue:

- What is not working well now? Based on these notes, several suggestions could be made, at least for the East Africa region, but (1) the question is broader than East Africa and (2) there would be multiple perspectives on this question.
    - **Accuracy of forecasts.** How appropriate are the assumptions used for forecasts and transparency about the assumptions?
    - **Links to EA.** How can information and analysis link to anticipatory action and improve decision-maker responsiveness?
    - Systematic incorporation of violent conflict and better prediction of the impacts of conflict.
    - Politicization of the information and analysis.
    - Decision-makers' trust in the findings.
    - Prediction of "one-off" complicating factors.
    - Systematic links to—and inputs/analysis from—at-risk communities.
- How can predictive modeling or AI analytics improve the quality of early warning? Are there specific "gaps" that can be filled? Does the use of AI and predictive analytics constitute a wholly different approach? Can these different approaches be integrated and better linked to decision makers and at-risk communities?
    - How will these very different approaches (traditional EW and predictive modeling) work together? Can they work together to triangulate findings and forecasts?
    - Will predictive modeling take some of the politics out of early warning and information, or do the number simply provide a "veneer" of objectivity while obscuring how models are calibrated, how choices are made about inclusion or exclusion of data, etc.? (Or will it make the politics even more heated—since much of the predictive modeling is going on totally outside the affected context, sometimes with little or no engagement from national stakeholders?)
    - How can stronger links be built in to early action in crisis, as well as to preventive measure and risk reduction?

And there are a series of more technical questions within the predictive modeling community as well (Erin Lentz, personal communication).

References

Anei, Tong Deng, Alex de Waal, and Bridget Conley. 2019. "Accountability for Starvation Crimes: South Sudan." Policy brief. Somerville MA: World Peace Foundation.

Bailey, Rob. 2012. "Famine Early Warning and Early Action: The Cost of Delay." London: Chatham House.

BRCiS. 2019. "BRCiS Early Warning Early Action Strategic Plan and Roadmap." Nairobi: BRCiS.

Buchanan-Smith, Margie, and Susanna Davies. 1995. *Famine Early Warning and Response: The Missing Link.* London: Intermediate Technology Publications.

Cabot-Venton, Courtenay, Catherine Fitzgibbon, Tenna Shitarek, Lorraine Coulter, and Olivia Dooley. 2012. "The Economics of Early Response and Disaster Resilience: Lessons from Kenya and Ethiopia." Report for the Department for International Development (DFID). London: DFID.

Centre for Humanitarian Change. 2019. "Improving Food Security Analysis. Turkana Assessment: Differentiating IPC Phase. Nairobi: Centre for Humanitarian Change.

Choularton, Richard, and Krishna Krishnamurthy. 2019. "How Accurate Is Food Security Early Warning? Evaluation of FEWS NET Accuracy in Ethiopia." *Food Security* 11: 333–44.

Drechsler, Mareile, and Wolter Soer. 2016. "Early Warning, Early Action: The Use of Predictive Tools in Drought Response through Ethiopia's Productive Safety Net Programme." Markets Global Practice Group. Washington: World Bank.

Dubois, Marc, Paul Harvey, and Glyn Taylor. 2018. "Rapid Real-Time Review: DFID Somalia Drought Response." London: Humanitarian Outcomes.

Fitzgibbon, Catherine. 2014. "HSNP Phase II registration and Targeting Lessons Learned and Recommendations." Final Report for the Department for International Development (DFID). London: DFID.

Food Security Information Network. 2015. "Qualitative Data and Subjective Indicators for Resilience Measurement." Resilience Measurement Technical Working Group. Technical Series No. 4. Rome: Food Security Information Network.

FSNAU, Oxfam, and BRCiS. 2018. "Report of the Stakeholder Consultation Workshop on Linking Early Warning to Early Action in Somalia." Nairobi: FSNAU.

Glenzer, Kent. 2009. "We Aren't the World: The Institutional Production of Partial Success." In Xavier Crombé and Jean-Hervé Jézéque, eds., Niger 2005: Une Catastrophe Si Naturelle. Paris: MSF.

Grunewald, Francois, Valerie Leon, and Simon Levine. 2019. "Comprehensive Review of the 2016–17 ECHO Horn of Africa Drought Response." London: Groupe URD/ODI.

Hailey, Peter, Jeeyon Janet Kim, Erin McCloskey, Maria Wrabel, and Daniel Maxwell. 2018. "The Constraints and Complexities of Information and Analysis in Humanitarian Emergencies. Evidence from Somalia." Medford MA: Feinstein International Center and Nairobi: Centre for Humanitarian Change.

Hillbruner, Chris, and Grainne Moloney. 2012. "When Early Warning Is Not Enough—Lessons Learned from the 2011 Somalia Famine." *Global Food Security* 1 (1): 20–28.

IFRC. n.d. "Early Warning, Early Action: Mechanisms for Rapid Decision Making." Geneva: IFRC, Save the Children, Oxfam, FAO and WFP.

IPC Ethiopia. 2019. "Ethiopia: About Eight Million People Severely Food Insecure Due to Erratic Rains, Conflict and High Food Prices." July 2019-June 2020. Rome: IPC Info.

IPC Kenya. 2019. "Kenya: Food Insecurity Persists across ASAL Counties Exacerbated by Poor Seasonal Performance." July-October 2019. Rome/Nairobi: IPC Info and NDMA.

Kidd, Stephen, Bjorn Gelders, and Diloá Bailey-Athias. 2017. "An Assessment of the Effectiveness of the Proxy Means Test Poverty Targeting Mechanism." ESS Working Paper No. 56. Geneva: ILO and Development Pathways.

Kinyoki, Damaris, Grainne Moloney, Olalekan Uthman, Ngianga-Bakwin Kandala, Elijah Odundo, Abdisalan Noor, and James Berkley. 2017. "Conflict in Somalia: Impact on Child Undernutrition?" *BMJ Global Health* 2017 (2): 1–11.

Lentz, Erin, Hope Michelson, Katherine Baylis, and Yang Zhou. 2019. "A Data-Driven Approach Improves Food Insecurity Crisis Prediction." *World Development* 122: 399–409.

Levine, Simon, Alexandra Crosskey, and Mohammed Abdinoor. 2012. "System Failure? Revisiting the Problems of Timely Response to Crises in the Horn of Africa." HPN Network Paper #71. London: Humanitarian Practice Network.

Lowcock, Mark. 2018. "Casement Lecture: Towards a Better System for Humanitarian Financing." Lecture delivered at Iveagh House, Budlin, 23 March. New York: OCHA.

Maxwell, Daniel, Peter Hailey, Abdullahi Khalif, and Francesco Checchi. 2020 (forthcoming). "Determining Famine: Multi-Dimensional Analysis for the 21st Century." *Food Policy* (accepted for publication).

Maxwell, Daniel, Peter Hailey, Jeeyon Janet Kim, Erin McCloskey, and Maria Wrabel. 2018. "The Constraints and Complexities of Information and Analysis in Humanitarian Emergencies. Evidence from South Sudan." Medford MA: Feinstein International Center and Nairobi: Centre for Humanitarian Change.

Maxwell, Daniel, and Nisar Majid. 2016. *Famine in Somalia: Competing Imperatives, Collective Failures. 2011–2012.* London and New York: Hurst Publishers/Oxford University Press.

Maxwell, Daniel, and Ben Watkins. 2003. "Humanitarian Information Systems and Emergencies in the Greater Horn of Africa: Logical Components and Logical Linkages." *Disasters* 27 (1): 72–90.

Nissen, Lars-Peter. 2016. "Gandalfs and Geeks: Strengthening the Accountability of Humanitarian Decision-Making." In, *On the Road to Istanbul: How Can the World Humanitarian Summit Make Humanitarian Response More Effective?* Chapter 3. Geneva: CHS Alliance. Humanitarian Accountability Report. London: CHS Alliance.

Obrecht, Alice. 2019. "Adapting According to Plan: Early Action and Adaptive Drought Response in Kenya." London: ALNAP.

Oxfam. 2017. "From Early Warning to Early Action in Somalia. What Can We Learn to Support Early Action to Mitigate Humanitarian Crises?" Nairobi: Oxfam.

Practical Action. n.d. "Early Warning Systems." Policy and Practice. Warwickshire UK: Practical Action. https://policy.practicalaction.org/projects/ews.

Smith, Lisa, Tim Frankenberger, and Suzanne Nelson. 2018. "Pastoralist Areas Resilience Improvement and Market Expansion (PRIME) Project Impact Evaluation: Report of Recurrent Monitoring Survey 2 (2015/16)." Addis Ababa: Feed the Future/TANGO.

World Bank. 2019. "Technical Assistance to Strengthen the Early Warning System for Food Insecurity in Ethiopia and Its Links to a Scalable Rural Productive Safety Net Programme (RPSNP)." Addis Ababa: World Bank.

World Food Programme. 2019. "Strengthening Early Warning System in Karamoja for Nutrition Forecasting, Planning and Programming Support." Kampala: WFP.

Annex 1. Incorporation and Analysis of Qualitative Information

A recurrent theme throughout this review of early warning, early action, and humanitarian information systems more generally is the use (or exclusion) of qualitative data and qualitative analysis (and related questions) in these systems—whether the specific kind of information or analysis is more closely related to EW or to assessments of one kind or another. This theme arises across various systems and countries and perhaps warrants a separate discussion. But it is not all just about "qualitative" information—it is about a series of related but different issues. These are discussed below in no particular order.

The Problem

In an earlier time, the analytical problem—and indeed the analytical approach—was to make sense of disparate and incomplete strands of information to try to understand the bigger picture, come to conclusions that were consistent with the information available, and act on that information. That is still the task, but under pressure from donors and decisions makers demanding a verifiable evidence base for their policies and actions, systems have turned increasingly to quantitative survey data. This is because it can be argued that it is statistically representative and therefore can be extrapolated; it is defensible as rigorous if certain rules about sampling and analysis are followed (and there is general agreement on what those rules are); and it represents "hard numbers"—which has become synonymous with "hard facts." All of this is fine and indeed necessary, but the problem of disparate sources of information remains. In the particular endeavor of humanitarian information systems attempts to follow the rules of quantitative analysis frequently meet obstacles: blocked access or limited time on the ground make representative sampling of a whole population impossible; choices have to be made about different data that paint very different pictures; and indeed "objective" indicators may actually work very differently in different contexts, so sometimes "hard facts" need to be contextually interpreted.

The first problem encountered when incorporating qualitative information into the analytical approach is distinguishing valid qualitative information from hearsay (or distinguishing "hearsay" from "anecdotal evidence"). It is not un usual for narrative information from an in-depth field assessment of a rapidly worsening situaion to be dismissed from an analysis because it could not be boiled down to "hard numbers." But the same analysis might even accept as evidence a comment such as, "I have a colleague who drove through that district last week and he didn't see any hunger there" Both the narrative information and the comment are "qualitative" information in the sense that neither can be objectively verified numerically. But they are not equivalent pieces of information. The former is qualitative evidence; the latter is hearsay. The problem is that many current systems do not have the tools to distinguish between them—and indeed whether either is admitted into the analysis is more frequently a question of political power than of the information's reliability and validity. So "hearsay" is one of the problems (so too is the politicization of information systems that can admit hearsay but dismiss well-documented qualitative evidence).

The second issue is distinguishing qualitative information from subjective information. They are not the same: qualitative information exists in narrative rather than numeric form; subjective information is about people's perceptions, memories, and aspirations. Lots of subjective information is numeric, and well-documented qualitative information can be just

as "objective" as statistics. But the widely held belief in humanitarian information systems is that "numeric = objective" and "narrative = subjective."[18]

Another issue is that of contextual knowledge or "contextualization." In an insightful analogy written in the run-up to the World Humanitarian Summit is 2016, Lars-Peter Nissen, the director of ACAPS, wrote an article about "geeks and Gandalfs" that perfectly captured this problem. A "Gandalf" is someone with long experience and deep contextual knowledge—who tends to make decisions on "gut responses" based on that experience and knowledge. A "geek" is a dedicated number cruncher who thinks that "the data speak for themselves" and doesn't make decisions at all—just goes on what the numbers say. The problem is that, any analytical (or decision-making) process inevitably depends on contextual and experiential knowledge *as well as* on the results of quantitative analysis. The problem with relying solely on number crunching is obvious—questions might be asked incorrectly, assumptions about the meaning of the results may be incorrect, numeric analysis may show association but can be completely wrong about causality, etc. The problem with deep contextual knowledge is that contexts can change—in fact, by very definition, *are* changing dramatically (for the worse!) in situations where early warning or humanitarian assessment is needed. The point is that both are important—and both can be wrong. Nissen's point was that any humanitarian analysis has to "make explicit the evidence base, assumptions and options considered in coming to any given decision" (Nissen 2016, p. 31). The problem is that we don't have good guidance about how to balance contextual knowledge and statistical data—particularly when they suggest very different interpretations of a given empirical context. Purely data-driven conclusions may not take important interpretations into consideration. Contextual expert knowledge may trap the analyst into thinking s/he knows something for certain but cannot see it changing subtly over time (a kind of analytical "boiling frog" syndrome).

The related—and deeper—issue is the question of human analytical judgment incorporating disparate kinds of information. But this incorporation happens *all* the time. Almost never does a machine algorithm combine all the sources of information available and spit out not only an analysis but also a plan of action (that may be the holy grail of artificial intelligence, but has certainly not been achieved yet in the humanitarian world). By definition, human judgment is relied on in both early warning and assessment processes—both in terms of the analysis and the follow up decisions taken about early action (or *any* action). Yet, unlike strict statistical analysis, we don't have clear rules or guidance on how this should be done. Note that some guidance notes are out there that can help: IPC analysis has "evidence templates" for guiding and documenting the way in which pieces of information are combined; FEWS NET has a carefully worked-out process for developing scenarios, combining pieces of information, etc. All of these rely on human analytical judgment. And yet, analyzing "contributing" factors in an IPC analysis or developing scenarios of future status can both be processes fraught with methodological difficulties: What are the criteria for including or excluding evidence (the "hearsay" problem)? How should different factors be "weighted" in the final analysis? How can one balance statistically representative information with purposively collected information? How can one deal with contradictory outcomes, even if the data on which those are based is validated? In other words: what do analysts do when the data *do not* "speak for themselves"?

Significant progress has been made in this area, and some has been documented. However, much remains to be done. The rest of this annex is devoted to helping to define what has to be done and to documenting good practice.

Terminology

Part of the problem is a lack of agreement on terminology. Below is a partial list:

- **What is qualitative information?** Qualitative data is information in the form of narrative or words rather than in the form of numbers. Sometimes words can be converted into numbers, and of course frequently words are used to describe numeric information. But fundamentally quali-

[18] As noted above, "subjective data" (preferences, opinions, etc.) should not be conflated with human judgment in analysis.

tative information cannot easily be reduced to numeric representation. See Table A1.

- **What is subjective information?** Subjective information is about perceptions, opinions, preferences, and feelings—that are real to the provider of the information but may not be exactly equivalent between different respondents. Note that subjective information can be qualitative or quantitative (experiential quantitative indicators sometimes rely on subjective responses from respondents). See Table A1.
- **What is anecdotal information?** Anecdotes are stories—individual stories can sometime be quite revelatory of a broader phenomenon or can mislead if generalized from a very atypical case. Hence this kind of information is often dismissed as "anecdotal," particularly by parties who disagree with the obvious implications of such a story. But all information is to some degree based on an "anecdote." The result of a single qualitative household interview is an "anecdote"—a story about that household. So too are the results of a single household survey questionnaire—that too is an "anecdote" about that household.
- **What is hearsay?** "Hearsay" is information attributed to other people and which therefore cannot be substantiated (or refuted). It is *not* "data" and frequently is not only a distraction but a form of subterfuge.

The main question is about how all data is collected, judged, validated, used, tossed out, or ignored. But the problem is that the guidelines for how one judges quantitative information are clearer. The criteria by which to judge the reliability and objectivity of qualitative information is less strong—and sometimes non-existent.

The humanitarian system also lacks guidance for how to utilize qualitative information. This can be developed—in fact must be developed. This includes documenting existing good practice and reviewing the literature on qualitative methods. But it also includes countering political pressures that find it handy to throw out some qualitative information as "unreliable" while admitting convenient hearsay.

Good Practice Methods for Incorporating and Validating Qualitative Information

While there are perhaps fewer guidelines for collecting and analyzing qualitative information, there are some general guidelines:

1. **Diversification of data sources.** The best protection against a single story or anecdote is to deliberately diversify sources of qualitative information. This links to both triangulation and purposive sampling.
2. **Triangulation** is a process of verifying information from multiple and disparate source. It requires the interviewer or researcher to be skeptical of the results of a single interview or focus group, to seek counter factual or even contradictory information, and to seek to validate (or refute) tentative findings by repeating the same

Table A1. Differentiating Subjective and Qualitative Information

Data Type	Empirical Focus	
	Subjective	Objective
Numeric (Quantitative)	Examples: survey data on perceptions, preferences, self-assessment	Examples: survey data on events, behaviors, and material conditions
Textual (Qualitative)	Examples: interpretation and affective states, meaning/reason of preference or perception	Examples: political allegiances, social relations, decision-making, institutional forms

Source: FSIN 2015

questions or interviews with multiple sources. If information is broadly corroborated (triangulated) by different respondents it can generally be trusted as valid. If information is disputed, the responsibility of the interviewer or researcher is to understand why—which may require further interviewing or research. *Disputed* information doesn't necessarily mean *wrong* information—but it does require an explanation, which often has to do with the perceptions of the respondent (note that *perceptions* gets into the issue of subjective information, which as noted above is not necessarily the same as qualitative information). Of course triangulation can be used to cross-reference qualitative and quantitative information—in fact this is one of the most powerful uses of triangulation.

3. **Clarity about logic.** A study trying to validate a hypothesis uses *deductive* logic; a more open-ended question oriented at learning, but not necessarily validating a hypothesis, may use more *inductive* logic. Neither is "better" than the other—but identifying which is driving a particular question or information gathering/analysis exercise is important.
4. **Documentation** of the process of qualitative data collection, processing, and analysis is critical to the validation of qualitative information. This includes documenting:
    - sources of information (who, where, when, why);
    - processes (selecting the kinds of information gathering tools, selecting respondents, recording information, coding and analyzing information);
    - logic of the process (of determining key questions, of purposive sampling, of interpreting interviews, etc.); and
    - conclusions and the links between data, analysis, findings, and conclusions.
5. **Peer review.** Peer review processes can be used for qualitative information just as they can be for quantitative information—in fact may be more important for qualitative. But just as it takes a statistician or econometrician (i.e., someone with deep knowledge of the methods) to peer review a quantitative study, it requires someone with deep knowledge of qualitative methods to peer review a qualitative study.
6. **Key informant interviews** are conducted to get information about a specific topic from a person with deep knowledge of that specific topic. It can be a household interview (in which case the respondent of choice is the person in that household who knows the most about the topic of interest) or someone with lots of knowledge about a particular topic or geographic area. For example, local leaders or health workers have in-depth information about their community and can identify key informants to interview, etc. Key informants are therefore not selected randomly.
7. **Group interviews** may be used to get general information about a topic, a location, a trend, etc. Generally they are used when this kind of information is sought and where a single key informant is not likely to have this information.
8. **Focus group discussions** are similar in that they involve multiple people, but the difference is that focus groups are deliberately constructed to understand a given topic from multiple perspectives. So the objective with selection of informants for a focus group discussion is that they represent the maximum diversity of perspectives on a given topic about which they all share knowledge.
9. **Purposive sampling** (selecting a sample to include all perspectives on a phenomenon rather than on the basis of each individual in a population having the same statistical chance of being selected). Note that purposive sampling can be used with quantitative methods—for example the REACH Area of Knowledge (AoK) methodology relies on purposively identifying people who have recently been in an inaccessible area and interviews them using a quantitative questionnaire, resulting in quantitative information—but not information that is statistically representative of an entire population. FSNAU does the same thing with key informants in inaccessible areas.
10. **Saturation.** There is no way to statistically calculate the sample size for qualitative interviewing. Usually, when utilized with purposive sampling to ensure that questioning covers all possible

perspectives, interviewing stops when "saturation" is reached—meaning that an additional interview brings no new information.

11. **Interview "guides"** enable a researcher to follow a story line, including important categories of questions to be asked. These differ from "questionnaires," which require that every question be asked exactly the same way to multiple respondents. Note that key informant interviews can be about specific incidents, trends, or experience or about general, contextual knowledge. Triangulating the results of these interviews is critical for generalizable information. Qualitative information can be gotten from very formally structured interviews and questions, but frequently are semi-structured to enable an interview to follow a "story line" rather than presuming *a priori* that the interviewer knows what the "story" is.
12. Use of **qualitative analysis software** (NVivo, Dedoose, etc.). Analyzing qualitative information is done on the basis of coding similar or like pieces of information and analyzing those codes for emergent patterns of differences around the same or similar theme across a wide range of interviews or key informants. One can use either a deductive (hypothesis-testing) or inductive (hypothesis-generating) approach to both coding and analysis.
13. **Means of using qualitative and quantitative methods iteratively.** Perhaps the most powerful use of triangulated methods is to deliberately rely on both qualitative and quantitative approaches to information collection and analysis. This can be done in at least two separate ways:
    - Sequentially/iteratively: informing quantitative instrument development, ground truthing or cross checking often-speculative conclusions or "puzzles" from quantitative analysis, or understanding the "story behind the statistics"
    - In parallel: addressing the same set of questions simultaneously using different methods
14. **Other key factors** include the following:
    - Qualitative information is different from purposively sampled quantitative data. Many key informants in this study used the terms interchangeably—but they are not the same thing.
    - Qualitative data and perception-based subjective quantitative data are not the same. Many key informants equate qualitative information with subjective information, but they are different. See above: subjective information (information about individual perceptions, aspirations, experiences, or self-assessment) can be quantitative; qualitative information can be objectively verified if done properly (see above about triangulation and documentation).
    - Key informant interviews, group interviews, and focus group discussions have different purposes (and the kinds of information that can be gotten from each differ).
15. **Key questions** include the following:
    - What are the constraints to better use of qualitative data?
    - How does one "weight" very different kinds of information—especially when they suggest different conclusions? For example, where the qualitative data that can't be collected quantitatively suggest large-scale distress migration, but quantitative indicators don't depict the stress.
    - What are the best ways to get information from inaccessible areas?
        - rapid assessment methods (i.e., in IPC guidelines)?
        - through key informant interviews and even by phone-facilitated focus group discussions (i.e., as is done in FSNAU practice)?
    - What constrains better documentation of processes of human analytical judgment?
    - What are the best ways to go about what needs to be done?
        - Capacity building?
        - Developing guidance notes?
    - How can qualitative analysis findings be communicated to broader constituencies?

The Feinstein International Center is a research and teaching center based at the Friedman School of Nutrition Science and Policy at Tufts University. Our mission is to promote the use of evidence and learning in operational and policy responses to protect and strengthen the lives, livelihoods, and dignity of people affected by or at risk of humanitarian crises.

Twitter: @FeinsteinIntCen

fic.tufts.edu