2015 HUMANITARIAN NEEDS OVERVIEW

GUIDANCE
Humanitarian Needs Comparison Tool

August 2014

Table of Contents

Introduction	1
At a glance	2
How it works	3
Limitations & appropriate use	4
Process outline	4
Joint analysis	6
Model construction	7
Annex 1: Technical notes	12
Annex 2: Technical setup	13

Introduction

Generating structured information products that can facilitate joint inter-sectoral analysis of humanitarian needs is a major challenge in many humanitarian emergencies. At the same time, ensuring consensus on the key humanitarian needs of a crisis-affected population, and how these vary between sectors and geographic areas, can be invaluable in supporting both strategic and operational analysis of the response required. This becomes particularly important when considering setting priorities as part of strategic response planning, which the needs analysis must inform.

At any given moment, there is significant diversity in the type, quality, sources and formats of humanitarian data available in the field, especially during a rapidly-evolving emergency. There is no simple solution to the challenge of comparing large volumes of such information on humanitarian needs, but the adaptation and use of the tool documented in this guidance can help this process, aggregating and combining data into a format that assists comparison and discussion. If developed collaboratively and transparently with the advice from technical experts, the outputs from this tool can serve as a central frame of reference and promote a common view of basic humanitarian needs across time, locations and sectors. This then provides both the starting point and the framework for a joint approach to inter-cluster/sector analysis and validation.

The tool itself is merely a generic framework, built in Microsoft Excel, to simplify the rapid structured comparison and visualisation of data from a number of sources. It could potentially be applied to a number of different purposes. However, its usefulness and applicability depends heavily upon its appropriate use. This guidance covers one specific purpose of the tool: its use in informing the sectoral and inter-sectoral discussions taking place as part of the Humanitarian Needs Overview (HNO) process, in particular facilitating a comparison of needs across geographic areas. **Field practitioners intending to use this tool should be aware of the considerable potential for inappropriate use, and ensure that they are fully familiar with this guidance before proceeding[1].**

This documented approach is based on lessons learnt from countries which have previously developed and implemented similar versions of such a comparison tool, which has been used in natural disaster settings as well as protracted crises. In some countries the tool has been used for a different or additional purpose beyond the HNO process, including for strategic and/or operational planning at the inter- or intra-cluster level, and for continuous situation monitoring. This guidance does not cover such use cases. Also, this guidance recognises that there are different methodologies to inform needs analysis and subsequent response prioritization. This tool is one possible approach; country operations already using another methodology should evaluate the different approaches and employ the methodology that is most appropriate for their context and experience.

[1] Products such as maps, summary tables and charts generated from the tool can inform operational decisions but must be accompanied by specific metadata citing analytical reliability and quality of information. Additionally, extreme caution should be used when disseminating products publicly as they could be misinterpreted or used out of context.

AT A GLANCE: USING THE HUMANITARIAN NEEDS COMPARISON TOOL

1. **Purpose**. The tool provides a structured way to compare humanitarian needs, impacts and vulnerabilities across geographic areas and sectors, combining data on different aspects of the crisis and from different sources together in a consistent way. The resulting heat maps provide an easy-to-understand overview of how needs vary, and these can inform (not replace) the needs analysis component of the HNO and the prioritisation discussions of the SRP.

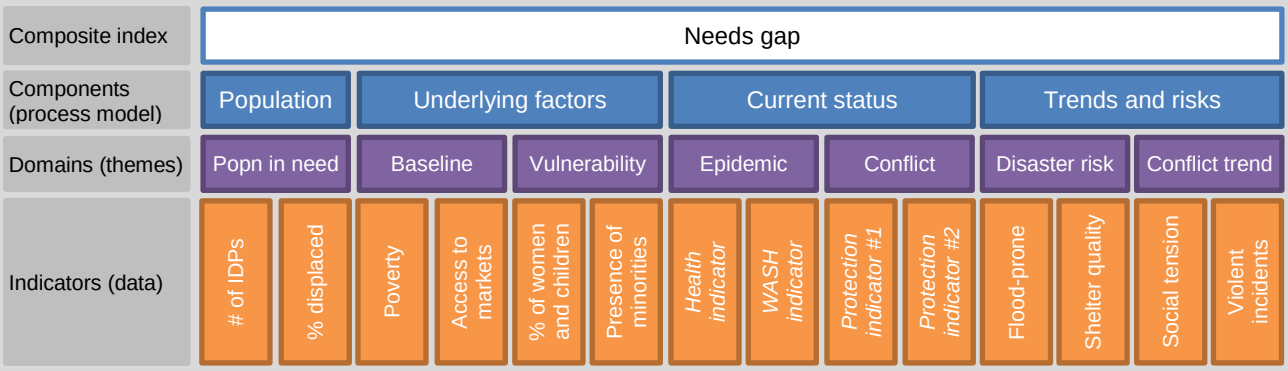
2. **How it works**. The HCT and ICCG define the construction of a 'composite index' which combines data from many different indicators or datasets, grouped in domains (themes), into a single score: a short-hand for the overall humanitarian needs. This score is then calculated for each geographical subdivision and plotted onto a 'heat map' - a visualisation of how needs differ across areas. Per-domain scores can also highlight differences between themes.

3. **Process**. This dovetails with the HNO process steps described in the HNO guidance.

1. Model structure and setup	2a. Indicator selection and definition	2b. Data consolidation and output preparation	3. Joint analysis
ICCG or smaller working group of experts agree on how the index will be structured (see example diagram below). This is adapted to each country context. • **Main components** (e.g. underlying factors, current status, trends and risks) • **Domains to include** (e.g. grouping by issues or themes instead of by cluster) • **Responsibilities** OCHA conducts technical setup of tool.	Domain experts evaluate available secondary data and identify indicators which best illustrate the underlying situation, taking into account • **Data availability** • **Reliability** Appropriate definitions, thresholds and weightings are defined for each indicator. OCHA and IMWG supports with advice on data availability and correct definition of thresholds	Data is obtained and populated as appropriate at the sub-national or national level, according to the defined indicators. Data is reviewed by domain experts. Where necessary, indicators and definitions are revised or replaced to accommodate field realities. OCHA consolidates data from all domains and produces heat map outputs	Heat maps are validated and then used by the HNO joint analysis team to inform their identification of key humanitarian issues. They are used only in conjunction with other evidence (such as in-depth needs assessments, views of affected population, expert knowledge.

4. **How to use**. Technical experts must be involved in setting up the model in accordance with this guidance to ensure its relevance and robustness (within the situational constraints). Outputs must be used carefully to minimise inappropriate use, with appropriate disclaimers and direct acknowledgement of their limitations. The impact of variances in data reliability and availability must not be considered at each stage. The tool serves as a starting point only for a joint analysis, and does not replace it. It is designed to be used as an integral part of a wider HNO and joint analysis process and should not be used as a standalone product.

5. **Support**. Contact OCHA (country office or Programme Support Branch in Geneva) for technical & setup support. Global clusters can advise on indicator selection & definition, and appropriate use.



Example of the composition of a needs index. For illustrative purposes only. Each country develops and adapts the model to their own context.

How it works

The comparison tool framework makes straightforward the creation and management of a composite measure[2], which combines information from a number of different user-selected indicators into one index. This index can then be visualised and its values compared to highlight trends and differences. This is a widely-used technique to permit multi-dimensional comparison, employed for example in the formulation of the Human Development Index (HDI) and its related indices.

The index (or indices) generated by the tool in this case refers to a combined 'score' based on a set of humanitarian needs indicators chosen carefully to illustrate the different dimensions and aspects of a particular humanitarian context, either within a particular theme (or 'domain') or across several. By calculating this score for each of a set of geographical subdivisions, a visual comparison of the situation is obtained which then informs further analysis and validation.

Key to the approach's strength is that it allows the 'distillation' of large volumes of needs assessment reports, expert judgement and other data sources to their key components. Providing that these components – and how they are combined – are wisely chosen through a collaborative approach based on an analysis of available data, this then allows the results to be compared across geographical areas and themes. In this way the tool's outputs can be viewed as a 'shorthand' common view of humanitarian needs – easily understood and an excellent place to begin the sectoral and inter-sectoral needs analysis that is required to support strategic decision making.

Key features

- The tool is provided without pre-defined indicators. Users are required to develop their own indicators in consultation with the stakeholders of the tool.
- Indicators are grouped together by 'domains' – typically a thematic grouping of humanitarian indicators. A domain may represent a humanitarian cluster or sector, or a cross-cutting theme such as displacement or access. The tool generates index scores for each domain, as well as a combined score across all domains.

Example

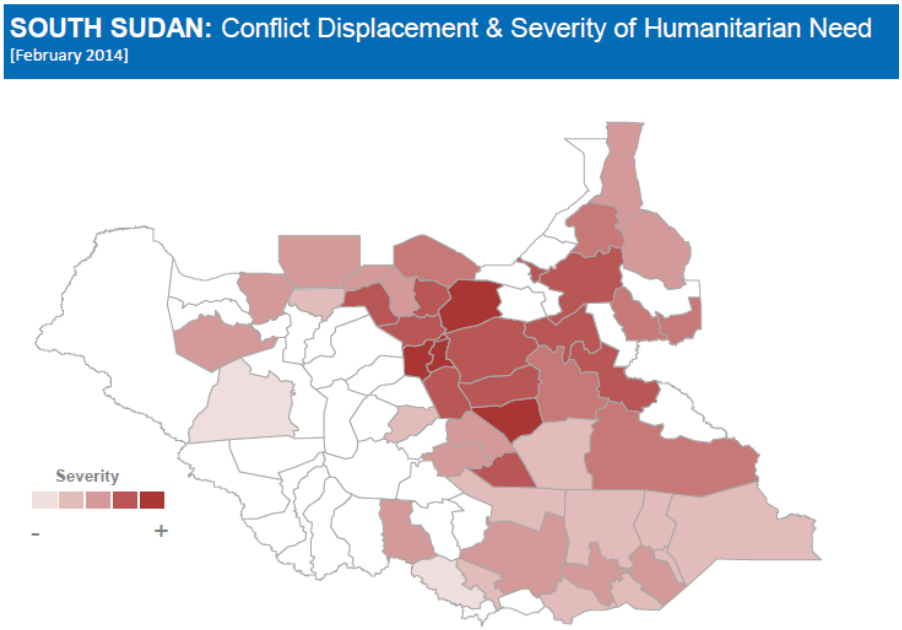


Figure 1 shows the standard output of this indicator-based comparison tool: a 'heat map'. In this case, individual indicators of humanitarian needs in South Sudanese counties with IDPs have been scored and combined to create an overall index – reflecting the severity of needs - and compared at the county level.

Figure 1: Comparison of severity of humanitarian needs, South Sudan 2014.

[2] A composite measure combines together several different indicators or measures into a single number. It provides an overall 'score' calculated using these different inputs, which thus allows a numerical comparison of more complex concepts than can be expressed in a single indicator.

Limitations and appropriate use

As explained above, the Excel tool itself merely provides the architecture to enable the comparison process. It is constructed generically and is flexible enough to allow it to be tailored to the technical capacity and information environment where it is being deployed. However, the usefulness of its outputs does depend on its correct use, which is underpinned by two components:

Managing inputs

To ensure that the tool's heat map outputs reflect the actual humanitarian situation as closely as possible, a careful and appropriate selection of its inputs – the domains and indicators – is vital. This initially requires an inter-cluster process to decide on the overall purpose of the tool (for example, measuring current needs, vulnerability, risks etc.)[3]. Once this is agreed, consensus is built around which domains to include by evaluating the particular humanitarian context (sudden-onset disaster, continuing conflict etc.), and highlighting key thematic issues.

Following this, the technical expertise of domain experts (within each cluster, for example) is used to decide on the indicators themselves. This includes a consideration of relevance, data availability and reliability for each potential indicator, a determination of the appropriate thresholds to represent different levels of need, and the relative weighting to apply to each chosen indicator.

Interpreting outputs

Appropriate use depends on a careful analysis and interpretation of the results, both at the domain and the overall level. **The outputs of the model serve as a useful starting point for strategic comparison and can be used to provide the basis for an overall, informed decision-making process where it is deemed to be useful. It should never replace a subsequent in-depth inter-sector analysis, but be used as an input to it.** If this process highlights the inappropriateness of the model or major limitations within the data which make the results less relevant, the model should be revised and improved, for example through the reconsideration of the choice of domains and indicators.

Because of the high level of variation in data quality and availability in humanitarian contexts, the indicators chosen to represent humanitarian needs may vary from highly-robust measurements following international standards (such as infant mortality or levels of severe acute malnutrition), to less-robust 'expert judgements' which serve as a proxy where measured data is scarce. These differences between 'data-rich' and 'data-poor' environments will affect the robustness and precision of the resulting index. Therefore, the analytical process particularly in the latter context should focus on identifying and explaining the larger trends, similarities and differences which emerge from the tool, and not place too great an importance on individual results at the detailed level.

The tool's outputs should always be interpreted bearing in mind the parameters with which it was set up. For example, if relative rankings rather than absolute measurements have been used[4], the 'least affected areas' on a map may still be severely affected. For this reason, it is essential that a concise overview of the relevant parameters and an indication of robustness always accompanies the outputs, to limit their out-of-context and potentially inappropriate use.

Process outline

The key steps required to implement the comparison tool in support of the HNO are as follows:

1. **Concept presentation**. Successful implementation of the tool at a country level requires full ownership of cluster/sector lead agencies and their members and the Humanitarian Country Team. As soon as possible after the launch of the HNO process (or at the opening workshop), an inter-sectoral meeting is held to introduce the comparison tool and discuss whether and how it will be used. This discussion should summarise the purpose and the limitations of the tool, as outlined in this guidance, and allow ample consultation on the underlying concepts and level of commitment required to carry out the exercise.

[3] The later section on 'process model (tool parameters)' provides greater detail on recommended approaches.
[4] The later section on 'selecting indicators' provides further information on relative vs absolute measurements

2. **Timeline**. The recommended timeframe for the exercise is about 3 weeks, but can also be shortened if required. A sequence of steps similar to that illustrated in the table below should be agreed upon.
3. **Tool parameters**. Decisions are made either through consensus at the inter-cluster meeting, or a smaller inter-cluster working group is formed to study and take decisions as appropriate to decide on the overall purpose, establish which domains to include and assign domain responsibilities. See 'Process model (tool parameters)' below.
4. **Setup**. OCHA conducts the technical setup of the tool, and shares it with clusters and technical experts.
5. **Indicator selection**[5]. Within each domain, technical experts meet to review secondary data (an essential part of the HNO process irrespective of the use of this tool), and decide on the best indicators which conform to the guidelines established in the tool parameters, balancing theoretical indicator relevance with actual availability of data, and seeking to reuse previously-used indicators where possible (to support monitoring and multi-year needs analysis). Referring to recommendations made at the global level (or consulting directly with global cluster experts), they then agree on categories, thresholds, weights and reliability scoring (the measurement model). The indicators are inserted directly into the tool itself. See 'Measurement model (indicator selection)' below.
6. **Data provision**. Data managers (often cluster IM focal points) within each domain populate the tool with available data. This can occur at the sub-national or national level, or both. However, if done by different groups at national and sub-national level it needs to be ensured that they are using the same methodology/criteria for populating the tool.
7. **Consolidation**. Each domain group returns the tool to OCHA, where it is then consolidated with the data from the other domains and used to publish both domain-level and inter-domain analyses (heat maps) as the principal outputs from the model. The results are disseminated to all stakeholders.
8. **Joint analysis**[6]. At an inter-cluster meeting to conduct joint analysis, the results from the tool (heat maps, for example) are presented, discussed and validated. The findings are then analysed and interpreted in conjunction with all other available data, including cluster-level secondary data reviews, situational analyses, crisis impact evaluations, in-depth assessments etc. The principal outputs of this meeting are sectoral and inter-sectoral analyses of needs, which may include consideration of absolute or relative severities across geographic regions, differentiation of key needs by areas and some form of ranking or prioritisation which can inform the strategic response planning process. See 'Joint analysis' below.
9. **Drafting**. The HNO report drafting process, including the sectoral chapters, may also be informed by the tool outputs.

	1	2	3
Presenting the concept of the comparison tool by inter-cluster coordination mechanism and the Humanitarian Country Team			
Adaptation of the comparison tool a. Inter-sectoral agreement on tool parameters, included domains b. Domain experts (often clusters) evaluate available secondary data and identify 1-3 relevant indicators for assessing needs c. OCHA formats the templates and advises on consistency between domains			
The comparison tool shared with responsible group of subject matter experts for each domain The subject matter experts with support of the domain (cluster) IM fill in the matrix and return the template to OCHA			
OCHA CO compile information from all domains and produce outputs			
Joint analysis with cluster representatives and technical experts where outputs are presented and discussed, and comparison and prioritisation decisions are made as appropriate; if possible representatives from the affected population should be invited to this analysis process			

[5] See 'Indicator selection' on p.6 for more detail.
[6] See 'Joint Analysis' section on p.11 for additional detail.

Joint analysis

Joint analysis is an integral part of the Humanitarian Needs Overview process, regardless of whether this tool or an alternative comparative approach has been used. Further guidance on the format and procedure for conducting this analysis can be found in the HNO guidance document. Main considerations specific to the use of the comparison tool during the joint analysis of the HNO process are as follows (see also 'interpreting outputs' above):

- Presentation of the results and inter-sectoral discussion of both the domain-specific and the overall outputs. Depending on the context, available data and applicability of the domain-level analysis to predominant humanitarian themes and/or clusters, it may be decided that one or some of the domain-specific heat maps (rather than the overall, inter-domain maps) are the most useful and relevant in illustrating needs. Maps can be used as appropriate. However, it should made clear that the heat maps should never be used as stand-alone product(s) and without an appropriate narrative, legend and disclaimer to highlight the limitations/methodology used in a transparent way
- Comparison of needs. Discussion should focus on highlighting and explaining the main differences, similarities and trends which the heat maps illustrate. Reasons for differences should be discussed and agreed, with the appropriate contextualisation and expert interpretation that can inform a more nuanced analysis. This may focus on a comparison of geographic areas as well as of different domains, or of underlying factors vs current impact vs risks[7]. This exercise may culminate in an agreed ranking of severity, and may already incorporate non-needs-based considerations such as access or response capacity.
- This should ultimately contribute to the identification of 'key humanitarian issues' – a central output of the HNO process which will inform the response analysis component of the strategic response plan.
- Prioritisation of response is an important component of the strategic planning process which normally follows an HNO, in order to ensure both that assistance is delivered as a priority to areas and/or domains where it is most needed, and that scarce funds are allocated appropriately. When used in conjunction with information on access and response capacities and within the overall framework of the agreed humanitarian issues and strategic objectives, the outputs from this tool can directly inform this prioritisation process.

[7] More information on these distinctions can be found in the 'Process model (tool parameters)' section

Model construction

The following section provides more technical detail on the construction of the process, data and measurement models. It should be used to guide the inter-cluster or technical working group which establishes the tool parameters, as well as the technical experts within each domain who select the indicators.

Even greater technical detail suitable for data managers and statisticians is provided in the Annex.

Key features

The following features of the tool should be kept in mind when in the planning phase:

The tool is provided with no pre-defined indicators. The tool is built with functioning placeholders for the indicators. Users are required to develop their own indicators in consultation with the stakeholders of the tool. As guidance, the IASC maintains a list of standardized humanitarian indicators[8] that can serve as a starting point, or existing humanitarian indicators currently being tracked in-country can be integrated with the tool.

Indicators are grouped by domains: the tool allows the user to group indicators by domain (the orange boxes in Figure 2). Typically each domain is a thematic grouping of humanitarian indicators. When the value for each indicator is entered, the tool combines their scores to generate the index (an overall score) for the domain. A domain may represent a humanitarian cluster or sector, or a cross-cutting theme. Domain scores can then be combined with other domains to give an overall index, or considered individually.

Indicators within each domain can have different relative importance. By adjusting the possible range of scores for individual indicators within a domain, the user can control the relative influence of each indicator on the domain's index (i.e. variable weighting of indicators within a domain). The indicators are combined additively for this purpose.

Domain indices are normalized to a percentage score before combination with other domains. Domain indices are always represented as a percentage of the maximum possible score for the domain. In this way no single domain has more influence in the overall index which combines multiple domains (i.e. equal weighting of each domain relative to other domains). The domains are combined multiplicatively to further ensure equal balance. If users wish to weight domains relative to each other the tool can be adjusted in individual circumstances to include this functionality.

'No data' and a score of 0 are handled differently- meaning that there is no bias towards information-rich locations over those that are missing data points[9]. Areas of missing information are automatically mapped and highlighted in order to see limitations in the model and inform assessment planning. Missing indicator scores does not reduce the score for that domain.

Process model (tool parameters)

Agreement must be reached on the tool parameters before indicators are selected. Whilst it may seem an obvious choice to define the model around the traditional humanitarian sectors or clusters and then proceed immediately to indicator selection, this approach bypasses the importance of the inter-sectoral themes and issues which define a humanitarian crisis. Key decisions should therefore first be made at an inter-cluster level – either through larger consensus or the proposal of a smaller working group – on the following:

A. What the tool will be used to measure. In simple terms, the process model defines what you want to measure, and the basic formula behind it. It explains the overall purpose of the tool. Each component of the process model will then link to one or several domains in the data and measurement models. The MIRA framework[10] provides a useful guide. Common alternatives are presented below.

1. A process model focusing on **current needs only** can nevertheless distinguish between the underlying factors driving the crisis (pre-existing conditions, baseline indicators and vulnerabilities), and the main crisis characteristics in terms of impact on status (health, protection etc.). It should also

[8] The humanitarian indicator registry, available at http://ir.humanitarianresponse.info
[9] Ensuring consistency between application of '0' and 'no data', requires appropriate definition of thresholds. See 'Measurement model (indicator selection) for more detail.
[10] The MIRA Framework is available at https://www.humanitarianresponse.info/programme-cycle/space/document/mira-framework

incorporate some measure of the size of affected population, key to assessing the relative importance of needs:

*Basic needs gap = Population * Underlying factors * Current status*

2. In **sudden-onset disasters**, this model should be expanded to include a measure of the exposure to disruption – hurricane wind-speeds, for example, or earthquake intensity.

*Sudden-onset needs gap = Population * Exposure to disruption * Underlying factors * Current status*

3. Particularly in situations of **continuing conflict** or for **preparedness** purposes, the model can also include trends and risks, which evaluate how humanitarian needs are likely to evolve over the subsequent time period[11].

*Continuing needs gap = Population * Underlying factors * Current status * Trends & risks*

4. For certain analyses, or as a next step, the model can be extended to incorporate measures of '**abilities to respond**' such as access or response capacities to inform strategic response planning.

*Response gap = Needs gap * Abilities to respond*

The choice of model may differ depending on context as well as data availability, though models 2 and 3 are recommended for sudden-onset and continuing conflict situations respectively. Model 1 may be most appropriate for data-scarce or relatively static situations, while model 4 should only be used as a subsequent step after the HNO, when considering response prioritisation as part of the SRP process.

B. Which domains will be included. Each component of the process model can be interpreted as a 'domain' (or several domains) within which indicators are used to measure the contribution of each domain to the overall score. Common inter-sectoral themes or issues may serve as better proxies of underlying needs. For example, if a major humanitarian issue is cholera, it may be worth defining this as one domain and populating it with non-sector-specific proxy indicators known to be correlated closely with cholera (such as flooding risk) – which may then inform strategic planning for both WASH and Health. Displacement (as measured by IDP numbers, for example) may be another common theme. The following table provides guidance on the appropriate selection of domains. Each component should be represent by as few domains as possible, to balance a clarifying reduction to key components with a consideration of the complexity of the humanitarian situation.

Process model component	Example domains
Population	Total population
	People affected or in need
	Population subgroups (women, IDPs, refugees, returnees etc.)
Underlying factors	Vulnerable groups
	Baseline indicators
Current status	Inter-sectoral theme or issue (conflict, epidemic, etc.)
	Sector (Health, Protection etc.)
Trends and risks	Disaster risk
	Conflict/security trends
Abilities to respond	Humanitarian access
	Response capacities

The decision on domains to include, particularly in the 'current status' component, may then mirror the structure and framework provided by the key humanitarian issues of the HNO and/or the strategic objectives of the SRP, and can serve to break down cluster silos.

[11] The InfoRM risk management framework is a good starting point for assessing risk. More information at http://inform.jrc.ec.europa.eu/

Data model

Before proceeding to select indicators for each domain, it is worth reviewing at the inter-cluster level the data model which will be used to assemble the domains and indicators together. This is illustrated in Figure 2.

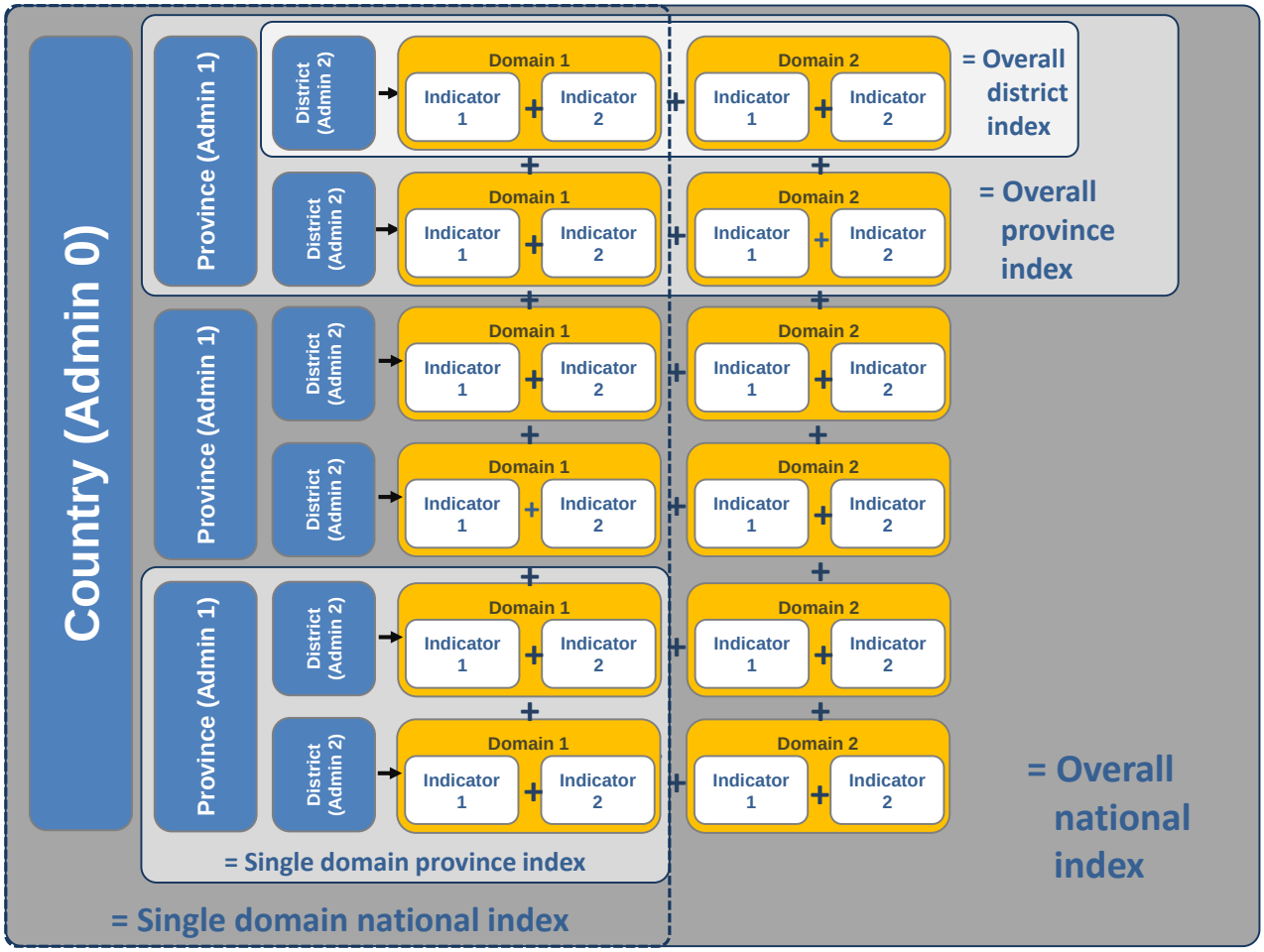


Figure 2: the basic data model of the comparison tool.

In this example, the tool is configured to receive data entry at the 2nd administrative level. The white boxes represent individual indicators or datasets, and the orange boxes represent a single domain (a set of indicators or datasets that together characterize a single sector or theme). Each indicator has a score, and the combination of the scores in a domain provide an index. The grey rectangles represent the levels at which indexes can be derived. In this example indexes can be calculated for a single domain (the orange boxes) or an overall index can be generated by amalgamating all the domains into a single index at all geographic levels from level 2 up to national (level 0).

The key decision on the data model to be made at an inter-cluster level is on **the targeted geographic resolution** – with the understanding that a general, large-scale resolution (at the Admin-1 level, for example) may be useful to do comparisons for strategic planning, but may require the averaging of data which is available at smaller-scale resolutions; whereas too small a scale may make it very difficult to find consistently available data and result in large pockets of 'no data', limiting the tool's general applicability. An appropriate balance must be found.

Measurement model (indicator selection)

The success of the comparison tool depends heavily on an appropriate selection of indicators in each domain (the measurement model). Before selecting indicators, instructions must be given by the inter-cluster group on how indicators must be chosen to ensure consistency of indicators across domains. In particular:

- Indicate whether indicators should measure magnitude (independent of population size) or intensity (i.e. adjusted for population size). This depends upon whether population has been included as a separate component in the process model, in which case **indicators of magnitude** should be selected. If intensity

indicators are used, then the impact of different population sizes will already be taken into account within the indicators and should not be included as a separate component in the process model. All indicators in all domains should be one or the other.

- Specify that indicators should be **absolute**, i.e. measure severity of the situation in absolute terms, rather than relative (where results are ranked). This is important because relative indicators can obscure the true severity of needs, understating them for low-ranked areas in a region of consistently high severity, or vice versa.

The inter-cluster group should also allocate specific responsibilities for each domain to a person or group who will select the domain's indicators as well as take the lead in providing the data.

Following these instructions, domain (subject matter) experts in cooperation with data/IM officers must carry out a review of secondary data and for each potential indicator or dataset evaluate the following[12]:

1. **Suitability as a domain-wide proxy**. An analysis of composite indices shows that incorporating a large number of separate indicators adds little accuracy for the extra effort involved. It is better to choose a small number of suitable indicators (one or two) which prior experience and expert consensus agree track closely the situation across the whole domain, and thus act as a suitable proxy. This also increases the likelihood of being able to find indicators with good availability of data (or expert knowledge) availability.
2. **Non-correlation**. Choosing two indicators which are known to correlate closely with each other effectively leads to redundancy, adding no value to the model and reducing its sensitivity. Identify independent indicators which are not known to be closely associated with any other indicator being included in the model (either in the same domain or others). If two domains are closely related, such as WASH and Health, this may require checking with colleagues in other domains to ensure one domain's indicators are reasonably independent of the other's.
3. **Data availability**. Indicators with a comprehensive recent dataset at the appropriate geographic level are preferred. If there is a lack of available data across all areas, the discussion needs to focus either on using a different indicator for which better data is available, or creating an "estimation indicator" based on expert opinion or qualitative evidence – with non-quantitatively-defined categories, or in the worst case a simple Yes/No. A guide to illustrate this process:
    - If the context is data-rich, clusters may choose to identify quantitative indicators (i.e.: Number of schools used to shelter IDPs)
    - If the context is data-challenging, clusters can use ranges (i.e.: Between 1-25% of schools are used to shelter IDPs=1; Between 26-50% of schools are used to shelter IDPs=2)
    - If the context is data-poor, clusters can use incrementing graduation (i.e.: No schools are used to shelter IDPs=0, Some schools are used to shelter IDPs=1, Most schools are used to shelter IDPs=2)

Once suitable indicators have been identified, thresholds need to be defined for each indicator as follows:

4. **Category thresholds**. Appropriate thresholds need to be selected to define the boundaries between different categories. These thresholds can be chosen according to static or global standards, or dependent upon the specific humanitarian context. Each threshold must be clearly defined and recorded, so that an accurate and consistent categorisation can be carried out by those responsible for data collection. Also, thresholds should be set such that a normal situation is represented by a 0 score, and such that all possible indicator values are represented by exactly and only one category (i.e. there is no possible value which would fall outside the defined thresholds, or within several). For example:

[12] A shorthand way to check the quality of indicators is given by the acronym 'TURC': [T]echnically sound (robust enough to be comparable across time, countries, and crises); [U]nderstandable (simple to grasp and immediately apparent how indicator changes with the underlying situation); [R]elevant (a good proxy of the underlying situation, i.e. closely correlated) and [C]ost-effective (can be collected easily and regularly)

- An Education indicator on school attendance may be grouped into categories according to local impact, e.g. 0-25% = no concern (0), 26-50% equals some concern (1), >50% equals significant concern (2).
- A Yes/No WASH indicator may follow Sphere minimum standards, scoring 1 if average water use per person per day is at least 15 litres, or 0 otherwise.
- A qualitative categorical indicator may follow a severity scale, where 0=no impact, 1=minor impact without threat to life, 2=significant suffering but not life-threatening, 3=life-threatening if no assistance provided, 4=deaths are already reported. Data is categorised according to expert consensus.

5. **Reliability thresholds**. Similar to the well-defined category thresholds, each indicator must have well-defined reliability criteria, again to ensure consistent application. This means that the decision to assign 'Low', 'Medium' or 'High' reliability to any particular measurement can be made according to these criteria, which should incorporate consideration of the data source (type of organisation), age (when it was collected, compared with the timeframes over which this indicator is likely to change), and collection methodology (if known). Especially for qualitative or expert-opinion based data, these criteria can include information on the level of corroboration between sources, for example "5 experts agree", or "two studies conducted within last year show similar findings." All data sources should be clearly referenced.

6. **No data thresholds**. In data poor environments, data availability for a particular indicator may vary across areas. In areas where the data is considerably poorer than others, a threshold should be set to determine at what level of reliability the available data is discarded and recorded as 'no data' instead. For example, one household survey in a geographic area may not be considered adequate and thus recorded as 'no data', whereas a minimum of three household surveys would be required for the data to be included. For areas of no data, there can also be a discussion on whether values for these areas can be estimated or extrapolated consistently, for example by substituting the mean or median value, or should simply be left as 'no data'.

A final step, once all indicators for a domain have been chosen and their thresholds defined, is to consider:

7. **Weighting**[13]. Determining the weight of each indicator within a domain requires careful consultation and consideration with concerned stakeholders, and should be based on expert consensus. Each indicator should be assigned a weight reflecting their importance in relation to the indicators within the same domain. For example, in a domain with three indicators it could be decided that one of the three is twice as important in reflecting the underlying needs as the other two; the other two are of equal importance. This would result in giving the more important indicator a weight of '2', with the others receiving a weight of '1' each.

All information on the chosen domains, indicators, thresholds and weights is reflected directly in the Excel tool, which is then populated by data managers responsible for each domain according to these criteria and definitions.

[13] Some clusters, such as WASH at the global level, are developing dedicated tools to assist this weighting process and help establish consensus.

Annex 1: Technical notes

Creation of composite index

- Recognising that different models for constructing the composite index could impact on the resulting index values and any comparison conclusions and decisions based on them, the mathematical model used in this comparison tool is based on lessons learned and best practice from similar indices, while remaining as straightforward as possible to a) facilitate comprehension for users without statistical expertise; b) be broadly applicable to the types of indicators likely to be chosen according to this guidance without requiring modification in each different circumstance; and c) to be easily calculable in Excel. It aims to achieve the right balance between statistical robustness and the realities and necessities of the field.
- The composite index is created through the definition of one or several indicators for each domain. Each indicator can be either boolean (Yes/No) or categorical (e.g. Low/Medium/High/ Severe), Categories need to be mutually exclusive, i.e. multiple selection of several categories is not allowed. Continuous indicators (e.g. in a continuous range from 0.0 to 5.9) may be supported in the future, but currently these need to be split into distinct categories.
- Each indicator score is then normalised to a range of 0 to 1(expressed as % of maximum score). The lower-bound score of 0 is equivalent to 'normal / no concern' and the upper-bound score of 1 is equivalent to 'largest impact / worst concern'. The definition of upper and lower bounds can be adjusted, so that they are based on (for example) observed maxima and minima from the current dataset; maxima/minima from previous datasets in the same context, global observed maxima/minima, or nationally or globally agreed thresholds (such as literacy rates or nutrition scores)
- Within each domain (which may, for example, represent a sector or a cross-cutting theme), the indicators can be weighted by importance, and are then combined additively (using a *weighted arithmetic mean*) to create a composite index in the range 0 to 1 (0-100% of maximum score, as above) for that domain. Arithmetic combination is used both to permit an easy-to-interpret weighting, and to ensure that a value of 0 in one indicator does not lead to an overall domain index of 0, i.e. ignoring the values of other indicators.

$$\text{Domain index } D_a = W_1 * i_1 + W_2 * i_2 + W_3 * i_3 + \ldots$$

- An overall index across all domains is then obtained by the multiplicative combination of the domain indices with equal weighting, i.e. the *geometric mean*. This method is used to ensure a more balanced consideration of each domain, such that a doubling of one domain index would be balanced by the halving of another. To ensure that domain indices with a zero value (indicating 'normal') do not lead to an overall index of zero, all zero values are offset by +0.05 before calculating the geometric mean.

$$\text{Overall index } O = (D_a * D_b * \ldots * D_n)^{1/n}$$

- The equal weighting of domains has been implemented deliberately to prevent potentially long and fruitless discussions which could limit the tool's ability to achieve consensus. If strategic decisions require a consideration of the relative weight / importance of each domain, this should occur during the joint analysis stage, using the domain-specific outputs of the comparison tool as a guide. If variable domain-weighting within the model itself is absolutely essential, it could be adjusted in individual circumstances.

Calculation of reliability score

- The reliability score provides a first approximation of the validity / applicability of the composite index. It is estimated by assigning a 'Low', 'Medium' or 'High' reliability value to each indicator measurement.
- These values are then interpreted as 0, 0.5 and 1 respectively, and combined arithmetically alongside the indicator data, using the same weights, to create a domain reliability score between 0 and 1, which is still visually displayed as 'L' (0-0.33), 'M' (0.34-0.66) and 'H' (0.67-1).
- As with the domain indices, the domain reliability scores are combined multiplicatively to generate an overall reliability score across domains. As above, the geometric mean is calculated using an offset of +0.05 to account for zero-value domain reliability scores.

Treatment of 'No data'

- Where no data is available for a particular indicator in a particular area, this is recorded as a separate value (such as 'ND' or '?'), rather than a '0'.
- When calculating the domain index, any indicators with 'no data' are ignored and the weightings adjusted. For example, with three indicators weighted as 0.2, 0.6 and 0.2 respectively, if the third indicator has 'no data' the other weights are adjusted to 0.2/0.8=0.25 and 0.6/0.8=0.75. This is to avoid the index being biased towards data-rich areas.
- Unlike the domain index, the domain reliability score does not ignore 'no data' indicators but assigns them a reliability value of 'Low' (0). This is to ensure that data-poor areas have lower reliability scores than data-rich areas.
- If all indicators within a domain are 'no data', the domain index itself becomes 'no data', and is excluded from the calculation of the overall index. The domain reliability score in such a case is treated as 0 and NOT excluded in the calculation of the overall reliability score, again to ensure adequate representation of data-richness.

Annex 2: Technical setup instructions

Detailed step-by-step guidance as well as hands-on support is available for the initial setup and configuration of the Excel tool (for example, creating polygons from shape files) as well as its subsequent use (configuring domains and indicators, populating with data and generating outputs).

IM focal points from OCHA or other agencies responsible for setting up the Excel tool should contact the OCHA Programme Support Branch (Kashif Rehman, rehman@un.org) for further information and technical support.