GUIDANCE TOOL

GETTING IT RIGHT FROM PLANNING TO REPORTING:

A GUIDANCE TOOL FOR WOMEN'S LAND RIGHTS DATA AND STATISTICS

MAY 2021

GLTN GLOBAL LAND TOOL NETWORK

UN HABITAT FOR A BETTER URBAN FUTURE

UN WOMEN

This work is co-published by the United Nations Entity for Gender Equality and the Empowerment of Women (UN Women), the United Nations Human Settlements Programme (UN-Habitat) and the Global Land Tool Network (GLTN).

UN Women's editorial style has been used in this joint publication.

Editor (UN Women): Andy Quan
Production Coordination (UN Women): Carla Kraft
Design (UN Women): Melanie Doherty Design

GETTING IT RIGHT FROM PLANNING TO REPORTING:

A GUIDANCE TOOL FOR WOMEN'S LAND RIGHTS DATA AND STATISTICS

ECONOMIC EMPOWERMENT SECTION
UN WOMEN
New York, May 2021

GLTN
GLOBAL LAND TOOL NETWORK

UN HABITAT
FOR A BETTER URBAN FUTURE

UN WOMEN

ACKNOWLEDGEMENTS

UN Women, UN-Habitat and the Global Land Tool Network (GLTN) gratefully acknowledge the lead author of this guidance tool, Gianluigi Nico, and appreciate and recognize all the peer-reviewers and contributors who made this publication possible. This includes colleagues in the UN Economic Commissions and National Statistical Offices, the Africa Land Policy Centre, Kadaster International, Landesa, International Land Coalition, the Food and Agriculture Organization of the United Nations and the World Bank, who provided their technical expertise during the development of the tool. Substantive inputs and technical support was gratefully received from colleagues in UN-Habitat and GLTN, including Oumar Sylla, Robert Lewis-Lettington, Dr. Robert Ndugwa, Clinton Omusula, and Eric Gachoka under the overall coordination of Everlyne Nairesiae. The guidance tool was produced under the supervision of Seemin Qayum, with coordination and substantive support from Venge Nyirongo, Jocelyn Chu and Carla Kraft from UN Women.

TABLE OF CONTENTS

ABBREVIATIONS

CAPI	Computer-assisted personal interviewing
CI	Cognitive interview
CSOs	Civil society organizations
CT	Cognitive testing
DHS	Demographic and Health Surveys
EA	Enumeration area
FAO	Food and Agriculture Organization of the United Nations
GIS	Geographic information system
GLTN	Global Land Tool Network
GLII	Global Land Indicators Initiative
IAEG-SDGs	Inter-agency and Expert Group on SDG Indicators
INGO	International non-governmental organization
LSMS	Living Standards Measurement Study
GSBPM	Generic Statistical Business Process Model
LR	Land rights
MAPI	Mobile-assisted personal interviewing
NGO	Non-governmental organization
NSO	National Statistical Office
PAPI	Paper-and-pencil interviewing
PSU	Primary sampling unit
SDGs	Sustainable Development Goals
SD	Standard deviation
SSU	Secondary sampling unit
UN-Habitat	United Nations Human Settlements Programme
UN Women	United Nations Entity for Gender Equality and the Empowerment of Women
VGGT	Voluntary Guidelines on the Responsible Governance of Tenure of Land, Fisheries and Forests
WB	The World Bank Group
WLR	Women's land rights

ABOUT THIS GUIDANCE TOOL

To ensure a better and more sustainable future for all, the 2030 Agenda for Sustainable Development ("the 2030 Agenda") has identified 17 Sustainable Development Goals (SDGs) to be achieved by 2030. SDGs range from poverty eradication, zero hunger, decent work and reduced inequalities to quality education, clean water and sanitation, and gender equality, only to name some of them.[1] To monitor progress made towards the SDGs, a total of 244 statistical indicators have been developed accordingly, including indicators to monitor women's land rights (WLRs). Women's full realization of their land rights (LRs) is not only a key factor in reducing gender disparities and improving women's living conditions—including the full realization of human rights—but also a key vehicle to eradicate extreme poverty and hunger and ensure a path to sustainable development.[2]

This guidance tool aims to explain the practical steps towards enhancing the quality of WLRs data and statistics for data producers, analysts and researchers. In doing so, it addresses critical gaps in the quality of the design, collection, analysis, management and dissemination of data and statistics on women's land rights. The tool recognizes the diverse relationships between women and land with respect to their land rights, which vary from country to country and even context to context, based on legal, policy and institutional regimes.

In particular, this tool provides guidance on survey design, the data collection and statistics development process, and analysis and reporting on WLRs data. It suggests good practices for enhancing the quality of data and statistics on women's land rights in the context of measuring, monitoring and reporting on Sustainable Development Goals indicators 1.4.2, 5.a.1 and 5.a.2.[3]

The data value chain process is structured around four main stages: a) planning and design; b) collection; c) analysis and d) reporting. Each stage is further separated into additional substeps, which all include checklists of quality measures.

Although data on women's land rights can be collected through a number of statistical instruments, including agricultural censuses, land registries and satellite imagery, this tool specifically provides guidance on survey data collection instruments, which represent the best and most cost-effective method to collect data on women's land rights since they allow gathering all the necessary information that are required to construct WLRs-related indicators. This is because administrative data on land rights are often not disaggregated by sex, nor do they contain an individual identification number (ID) that is associated with the sex of the individual to produce gender statistics. Therefore, they do not always support the identification of whom these rights are associated with. For instance, the use of cadastral maps would allow for an even more reliable measurement of parcels and plots of land, as compared to self-reported data from surveys, but they generally do not allow for a breakdown of these data by sex.

Who can use this guidance tool?

This tool is for use by data producers and data users alike. It is a useful tool for government data and statistics authorities, the private sector, civil society, research and grassroots organizations that generate and/or use WLRs data and statistics. The tool is developed by taking into account the data elements that are required to construct the SDG indicators[4] on women's land rights but has further use. It aims to complement and support data and statistics efforts of National Statistical Offices (NSOs) and national land institutions, human rights and research organizations in the monitoring of the three SDG indicators mentioned above through the responsible custodian agencies, and also supports the collection of data and monitoring of other non-SDG indicators on women's land rights. Building on the global methodologies for monitoring these SDG indicators, this tool profiles and strengthens responsiveness to WLRs issues in the generation of data and statistics to ensure quality and reliable data for evidence-based decisions

are achieved. Although this tool is aligned to the approved methodology for the relevant SDG indicators, its content also addresses key issues that affect the quality of WLRs data and statistics more broadly. For more guidance on the approved methodologies, including tools for data collection on SDG land indicators, visit indicator 5.a.2[5] and access the joint module for collecting data and reporting on SDG indicators 1.4.2 and 5.a.1.[6]

Why women's land rights data and statistics matter

In the context of women's land rights, the goal of collecting data is to best describe the status of women's tenure rights, with the ability to show any inequalities and trends faced by women in securing land rights, at one point in time and over time. More specifically, data will show whether women are granted equal rights to land as men, if they can exercise those rights, if these rights are enforceable, and if women risk losing them arbitrarily. Data will also give more nuances into the effects of a specific land tenure system, on how those rights are granted and secured and who are those individuals or groups that face more disadvantages in securing land rights are (e.g. indigenous versus non-indigenous people, young or old, men or women, urban or rural settlers, and their main activity for deriving their livelihoods).

Collecting, analysing and reporting on WLRs data and statistics are key to planning and policy decisions that impact women and girls. Data and statistics on women's land rights are important to:

1. Identify legal and institutional challenges faced by women in accessing, using and securing their land rights, as compared to men.
2. Provide objective and empirical foundations for informing decisions about national policies to be enforced, developed, reviewed or repealed.
3. Be used as a baseline for policymaking and to set targets.
4. Enable comparison of the situation of women's land rights in different contexts including customary, indigenous, urban, and informal settlements, among others.
5. Inform strategies for closing gaps and serve as a tool for advocacy and influence around women's land rights linked to global, regional and national development frameworks including the SDGs, the Voluntary Guidelines on the Responsible Governance of Tenure of Land, Fisheries and Forests (VGGTs), the New Urban Agenda and other state obligations to international conventions on human rights and gender equality.

THE SDGs AND WOMEN'S LAND RIGHTS

The SDGs have provided unprecedented opportunities to advance tenure security for women, local communities and vulnerable groups, regardless of tenure type. SDG targets 1.4 and 5.a put women's land rights on the global agenda for sustainable development, with three specific indicators for monitoring tenure security provided to measure progress on their implementation using comparable data and statistics.

While SDG indicators 1.4.2 and 5.a.1, which cover all tenure types, allow countries to track the position of women in absolute terms as well as in relation to men, SDG indicator 5.a.2 is a legal indicator that monitors whether a country-specific legal framework (including customary laws) guarantees women's equal rights to land ownership and/or control. In more detail:

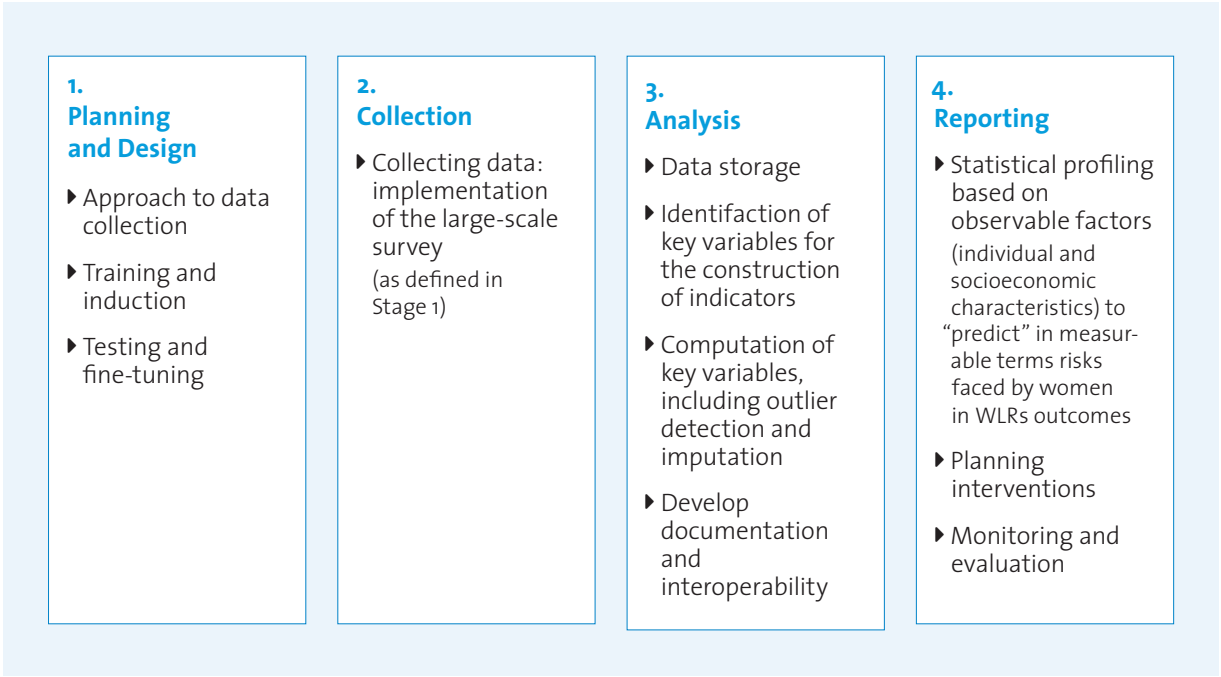
- Indicator 1.4.2 covers all adults and focuses on whether they have legally recognized documentation to prove their rights as well as on whether they perceive their rights as secure.
- Indicator 5.a.1 looks at similar issues but looks more closely at the population that depends on agricultural land for their livelihoods, including the ability to transfer their rights (by selling or bequeathing land). It aims to profile the proportion of women among the agricultural population who hold these rights. It is noteworthy that SDG 5.a.1 is intended to better capture issues around access, decision-making and control, which are essential from a gender perspective. This is because a woman may have documentation, i.e. land is under her name, but less ability to manage and make decisions regarding the land, i.e. to sell or bequeath it.
- Indicator 5.a.2 assesses the legal frameworks including customary law provisions to secure land tenure rights for women and girls.

These indicators enable countries to generate quantitative and qualitative data and statistics for measuring progress and informing policy decisions. It should be noted that although quantitative and qualitative data are strongly complementary, they are generated through different data and statistical processes.

FROM PLANNING AND COLLECTION TO ANALYSIS AND REPORTING ON WOMEN'S LAND RIGHTS

This guidance tool covers four stages intended to improve and ensure the quality of data and information on women's land rights, as represented in Figure 1 below.

FIGURE 1
The four stages of the guidance tool



Planning and design

The starting point of any data value chain process concerns the identification of the problem: what do we want to find out by collecting data? and how do we plan data collection activities? The answers to these questions inform the planning and design stage where the objective and purpose are established, consent is sought and data collection activities are planned, including survey design, training of enumerators and survey testing. A key requirement of this stage concerns the identification of why data are to be collected, particularly from governments and stakeholders which must support data collection to enable the creation of a strong evidence base that can be used for discussions and impact change at the policy level. This identification of the needs when deciding which data to collect is meant to be operationalized in the planning stage by accounting for existing information about potential barriers faced by women in securing land rights (e.g. from land registries). The engagement of experts, stakeholders and governments in this stage constitutes, therefore, a key issue to ensuring that the key inputs for the quality and reliability of the data being collected are considered in the design stage.

Collection

Consequently, the second stage of the data value chain process concerns data collection. Regarding the collection of quantitative and qualitative information on women's land rights, data collection activities are typically conducted through multi-topic household surveys, which represent the most prominent and authoritative data collection method, while the day-to-day administration of land (e.g. title creation and transactions) may provide an additional source of data to complement survey data and that can be used for statistical purposes.

Survey instruments represent the best and most cost-effective way to collect data elements that are used to compute SDG indicators on women's land rights, and a survey questionnaire administered to a representative sample of the population of interest is the best way to gather the complex mix of information that is required to construct SDG indicators 1.4.2 and 5.a.1. Survey instruments, if implemented on a continuous basis, facilitate the regular production of official statistics for better, evidence-based policies, thus supporting a continuous monitoring process. It is worth noting that a questionnaire module to collect the required data for SDG indicators 1.4.2 and 5.a.1 has already been developed by custodian agencies. Thus, countries that have planned to conduct data collection activities on women's land rights are strongly encouraged to administer this guidance tool alongside the respective modules.

In many countries, data collection through survey instruments is still conducted using the traditional paper-and-pencil interviewing (PAPI) method, while the use of computer-assisted personal interviewing (CAPI) or mobile-assisted personal interviewing (MAPI) is still limited.

Analysis

The third stage is data analysis. Before data analysis can be carried out, the first substage is data storage. This applies particularly to data collected through PAPI techniques, even though it must also be performed when data are collected through CAPI or MAPI. Once data collection activities in the field have been finalized and before data are analysed, data must be rigorously verified, organized, transformed and integrated; data must finally be extracted in an appropriate form to be properly packaged for subsequent data use and integrity. This includes validating, coding and analysing the data based on key variables and to make it possible to provide key analytical fields in responding to the overall objectives for which the data were collected.

Reporting

The fourth and final stage concerns reporting. This final stage provides relevant information needed for the design, monitoring and evaluation of polices, programmes and projects aimed at reducing gender gaps in land rights and tenure security. There should be a minimal time lag between the collection and reporting of data on women's land rights. This ensures that indicators on women's land rights report current, rather than historical information.

Moreover, ensuring data collection on a continuous basis is key to more frequent monitoring and reporting of potential progress made toward WLRs outcomes. The reporting stage should therefore also inform on whether or not changes in WLR outcomes are actually achieved and, in parallel, should allow for the evaluation of the quality, efficiency and effectiveness of the implemented interventions—policies and programmes—at various stages of implementation.

The tables below summarize quality issues to be considered during the design, collection, analysis and reporting of each of the SDG indicators. They also consider additional quality issues that go beyond SDG indicators, along the four stages of the data value chain. While the methodologies for SDG indicators 1.4.2, 5.a.1 and 5.a.2 have already been approved and the checklist of quality issues already consolidated in their methodologies, a number of additional indicators that are not part of the SDGs can be further computed.

As summarized below, data collection efforts should lead to the collection of details on women's land rights, such as: a) who owns the parcel of land, b) who makes the decisions about these rights, c) who has the right to sell and rent the land or use it as collateral, and d) the potential joint ownership/management of the land.

Yet, when data are meant to capture the reality about potential gender inequalities in the agricultural context, data collection efforts should be aimed at collecting information on a) who makes the decisions about what to plant, b) who controls the agricultural production, c) who decides whether to sell the final agricultural output, and d) who controls the agricultural income.[7]

Finally, in relation to the aim of collecting additional indicators on women's land rights, it is also worth adding questions on the value and area of each parcel of land that is owned, managed and/or controlled by a given individual. When the value and/or area of land that is owned by women, men or jointly is collected and translated into measurable indicators, it is possible to provide sound information on potential gender inequalities that goes beyond the SDG indicators.

In parallel to survey data, administrative sources represent a second data collection strategy, at least in those countries having a well-functioning land information system (cadastres and registries). However, a well-functioning land information system is a necessary but not sufficient condition to extract data on secure tenure rights. In order to construct meaningful and reliable indicators on women's land rights, the following data items should also be available from the country's land cadastres and registries: a) sex; b) tenure type (e.g. freehold, leasehold) and the corresponding legally recognized documentation; and c) land use type (e.g. residential, agricultural, pastoral). Data extracted from administrative sources have the advantage of facilitating the reporting phase of some SDG indicators on women's land rights, particularly because of their low costs and high frequency of data. Nonetheless, they are limited in scope: administrative data only cover ownership, not the other forms of tenure; or measure perception of tenure security. in this regard, only the first component of SDG 1.4.2, i.e. "legally recognized documentation", can be computed using these sources of data.[8]

TABLE 1

Key quality issues for SDG and non-SDG indicators along the data value chain – Planning and Design

Quality issues for Indicator 1.4.2	**1. Who to survey?** All of the adult population living in both rural and urban areas. **2. Characteristics of sampled individuals:** a) gender b) age c) indigenous people versus people who have settled, occupied or colonized the area/region d) relationship of sampled individuals. **3. Which women?** Women's multiple and intersecting identities, including legal wife, cohabiting partner, polygamous spouse, widow, divorced, single, daughter-in-law, rural/urban, agricultural producers, religious, ethnic background, etc. **4. Which land use?** All categories of land: residential, agricultural, pastoral, public, etc. **5. Which land rights?** The indicator purposely does not limit which rights. Its operationalization through the survey asks for use rights: "Any rights, but at least the right to use." **6. Which tenure rights?** Indicator 1.4.2 explicitly says that documents have to be legally recognized and tenure arrangements have to be those that the government plans to report on, therefore also tenure types the government recognizes. Rights backed by a customary structure alone would not be sufficient unless the government backs that structure. **7. How will secure tenure rights be proxied?** Through legally recognized documents and through people's perceptions of their own tenure security.
Going beyond the SDGs	Since it is not necessarily the case that the same person holds all of the rights[9], it is advisable to: **1. Include in the survey questions:** a) whether the parcel of land can be sold, rented or used as collateral b) who, within the household, has the above rights c) who makes the decisions d) who manages the parcel of land, e.g., who makes decisions about how to use it e) who controls the output, who decides whether or not to sell the produce, and who controls the income if it is sold f) reasons for fear of involuntarily losing their land tenure rights. **2. Include information on:** a) the area of the parcel of land, as measured in either standard units (acres, hectares, etc.) or non-standard units of measurement b) the value of the parcel of land: i.e. the monetary value of each parcel of land.

Quality issues for Indicator 5.a.1	**1. Who to survey?** All of the agricultural population living in both rural and urban areas. **2. Characteristics of sampled individuals**: All women as long as they are in a household that depends on agriculture for its livelihood. **3. Which women?** Women's multiple and intersecting identities, including legal wife, cohabiting partner, polygamous spouse, widow, divorced, single, daughter-in-law, rural/urban, agricultural producers, religious, ethnic background, etc. **4. What land?** Only agricultural (under temporary or permanent crops, meadows and pasture). **5. Which land rights?** The indicator purposely does not designate specific rights. Its operationalization through the survey asks for use rights: "Any rights, but at least the right to use." **6. Which tenure rights?** Indicator 5.a.1 explicitly says that documents have to be legally recognized and tenure arrangements have to be those that the government plans to report on, therefore also tenure types the government recognizes. Rights backed by a customary structure alone would not be sufficient unless the government backs that structure. **7. How will secure tenure rights be proxied?** Through legally recognized documents and through rights to bequeath or rent the land. **Important note:** This indicator does not account for perception of tenure security, nor does it leverage administrative data.
Going beyond the SDGs	Since it is not necessarily the case that the same person holds all of the rights,[10] it is advisable to: **1. Include in the survey questions on:** a) whether the parcel of agricultural land can be sold, rented or used as collateral b) who, within the household, has the above rights c) who makes the decisions about whether and what to plant d) who manages the parcel of land, e.g. who makes decisions about how to use it e) who controls the final agricultural output, who decides whether or not to sell the agricultural production, and who controls the agricultural income if it is sold. **2. Include information on:** a) the area of the parcel of agricultural land as measured in either standard units (acres, hectares, etc.) or non-standard units of measurement b) the value of the parcel of agricultural land: i.e. the monetary value of each parcel of agricultural land.
Quality issues for Indicator 5.a.2	**Proxy 1:** The joint registration of land is compulsory or encouraged through economic incentives. **Proxy 2:** The legal and policy framework requires spousal consent for land transactions. **Proxy 3:** The legal and policy framework supports women's and girls' equal inheritance rights. **Proxy 4:** The legal and policy framework provides for the allocation of financial resources to increase women's ownership and control over land. **Proxy 5:** If the legal systems recognize customary land tenure, does the legal and policy framework explicitly protect the land rights of women? **Proxy 6:** Does the legal and policy framework mandate women's participation in land management and administration institutions?

Going beyond the SDGs	**During the planning and design phase, it is important to consider the following:** • Legislation that is under discussion • Donor-funded programmes, project documents or memoranda of agreements because they don not fall within the legal and policy framework • Judicial decisions or advisory opinions of courts or official bodies to determine whether a proxy exists in the legal or policy framework • Customary and religious frameworks.

TABLE 2

Key quality issues for SDG and non-SDG indicators along the data value chain – Collection

Quality issues for Indicator 1.4.2	**1. Which instrument for data collection?** Survey questionnaire with a dedicated module on land. **2. Administrative sources** Cadastres and registries. **3 How to ask questions?** a) At land parcel b) Self-reported responses. **4. Who to ask questions** Either all adults in the household or a randomly selected adult in the household.
Going beyond the SDGs	Approach as per methodology for Indicator 1.4.2: Administrative data (cadastral data) for ownership of land. It is important that the land records contain an individual ID. Otherwise, it is not possible to aggregate to the person-level and data cannot show how women are doing. Please note that even having a "gender column" would not be enough because it will miss how many and what type of women are the ones who have the rights.
Quality issues for Indicator 5.a.1	Approach as per methodology for Indicator 1.4.2. **Important Note:** There needs to be caution on the use of administrative data, which might not identify agricultural land.
Going beyond the SDGs	Approach as per methodology for Indicator 1.4.2: Administrative data (cadastral data) for ownership of agricultural land.
Quality issues for Indicator 5.a.2	**How to collect?** Identify whether the country's policy and legal framework supports and implements gender-equitable land tenure arrangements.
Going beyond the SDGs	**How to collect?** Identification of a national legal expert to assess any potential progress made to close the gender gap which does not fall within the legal and policy framework.

TABLE 3

Key quality issues for SDG and non-SDG indicators along the data value chain – Analysis

Quality issues for Indicator 1.4.2	• Parcel-level data have to be aggregated to person-level information to determine which adults have secure tenure rights to land with (1) legally recognized documentation and (2) perception of secure rights to land. • Surveys will then have to be combined to cover as much of the country's population as possible. • Still, it is likely that many will not be covered, either because their households were not captured by the current portfolio of surveys gathering data on land, or because they were captured by a household survey that did not include questions on land rights in its coverage. • Indicators must be disaggregated by the type of tenure under which people have rights to land and by sex. • Special considerations should apply if the data come from household surveys which systematically interview the same type of adult , i.e. longitudinal studies. Unless they randomized who they interviewed (or they interviewed all adults), the data gathered on tenure security (and especially the data on perceptions of tenure security) will NOT be representative of all adults but rather representative of the type of adult included in the survey. For all other adults, there will be no information and therefore they cannot be considered tenure secure.
Going beyond the SDGs	**The unit of analysis:** a) land area b) value of the parcel of land c) individual person.
Quality issues for Indicator 5.a.1	Indicators can be constructed for various subgroups of people, including tenure system, acquisition type, primary land use and other land characteristics.
Going beyond the SDGs	**The unit of analysis:** a) agricultural land area b) value of the plot of the agricultural land c) individuals who engage in the agricultural sector.
Quality issues for Indicator 5.a.2	• Checklist of policy and legal instruments • List of policy and legal instruments for reporting the six proxies under Indicator 5.a.2.
Going beyond the SDGs	Checklist of whether legislation is under discussion and whether documents or memoranda of agreements do not fall within the legal and policy framework.

TABLE 4

Key quality issues for SDG and non-SDG indicators along the data value chain – Reporting

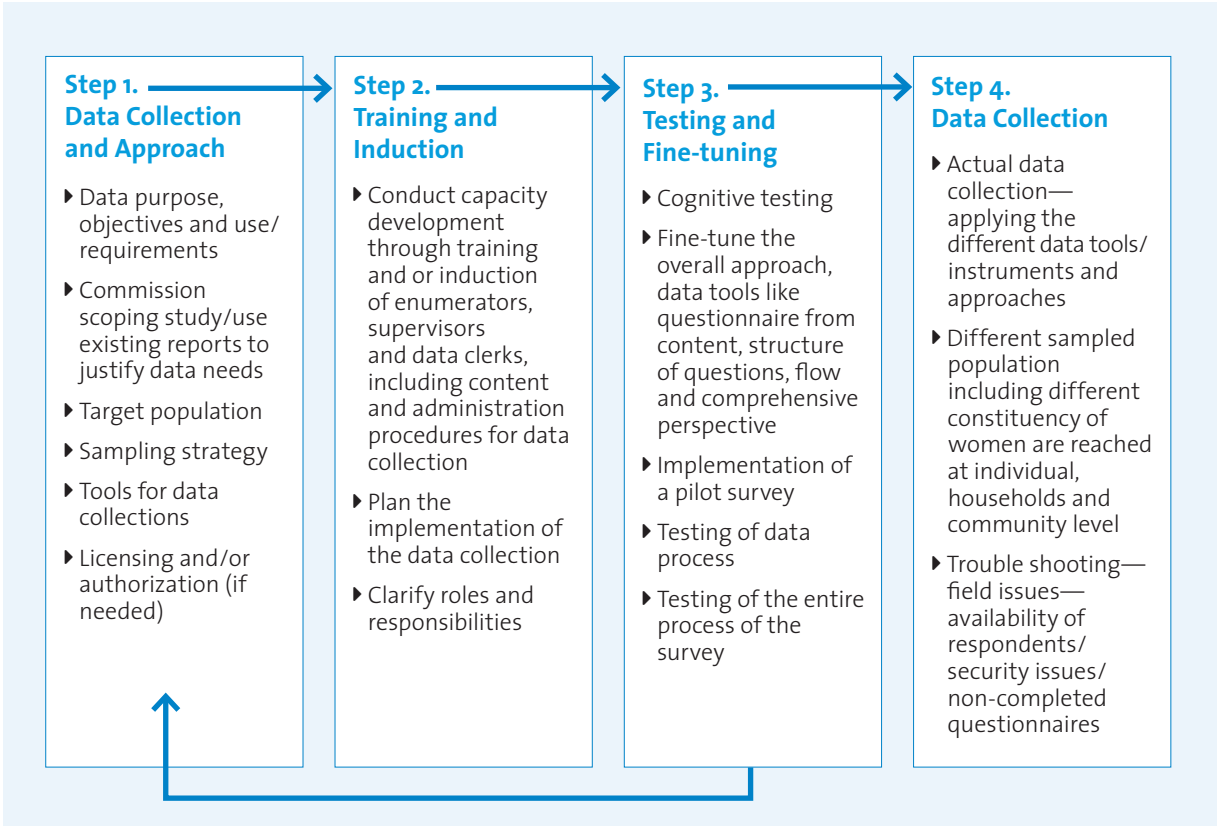
Quality issues for Indicator 1.4.2	**Proportion of the total adult population with secure tenure rights over land:** a) through legally recognized documentation, **by sex and type of tenure** b) who perceive their rights over land to be secure, **by sex and type of tenure.**
Going beyond the SDGs	**Indicators can be reported as:** • Distribution of land area owned/managed (as expressed in hectares of land, by men, women or joint ownership) • Distribution of land value owned/managed (monetary value), by sex, i.e. men, women and jointly • Disaggregation by sex, age, type of tenure and women living in different situations, e.g. conflict (internally displaced persons, returnees, refugees), informal, urban/rural/slums.
Quality issues for Indicator 5.a.1	**Report ownership and rights over lands in terms of:** a) Percentage of people with ownership or secure rights over agricultural land (out of total agricultural population), by sex b) Share of women among owners or rights-bearers of agricultural land, by type of tenure.
Going beyond the SDGs	**Indicators can be reported as:** • Distribution of agricultural land area owned/managed (as expressed in hectares of land, i.e. men, women and joint ownership) • Distribution of agricultural land value owned/managed (monetary value), by sex, i.e. men, women and jointly • Disaggregation by sex, age, type of tenure, agricultural women living in different situations, e.g. conflict (internally displaced persons, returnees, refugees), informal, urban/rural/slums.
Quality issues for Indicator 5.a.2	**What is the country's policy and legal context in terms of women's land rights?** Provide a detailed description of the country's legal and policy context, including reforms recently implemented to give women equal rights to economic resources, as well as access to ownership and control over land and other forms of property, financial services, inheritance and natural resources, in accordance with national laws.
Going beyond the SDGs	Provide a detailed description of whether legal and customary issues are *de facto* addressed but not yet encompassed in a formal political process which is expected to be translated in a plan of action, ultimately resulting in primary or secondary legislation.

STAGES 1 AND 2:
FROM PLANNING AND SURVEY DESIGN TO SURVEY IMPLEMENTATION

Before embarking on fieldwork operations, the data collection process is preceded by three interrelated steps: (a) planning, (b) survey design and (c) implementation of data collection.

The planning phase must involve key stakeholders in the country working on women's land rights. The planning stage has to secure needed permission from relevant authorities or offices and must ensure that enumerators are trained with relevant technical knowledge on the data collection process and the skills needed to manage the sensitivity issues of which they need to be aware. A testing phase to check that the entire process of the survey has properly been developed must precede the implementation

FIGURE 2
Planning, survey design and survey implementation



of the actual survey, i.e. from the cognitive testing to check the validity, reliability and comprehension of questions administered to respondents, to the implementation of a pilot survey (which only includes a small sample) and the testing of the processing of collected data. This is important to allow room for any necessary adjustments to the data collection exercise.

Figure 2 (previous page) shows the four key steps of Stage 1 for embarking on a successful data collection process.

Stage 1 – Step 1: Survey design for data collection

What is the purpose and objectives of collecting data?

Three elements are central to the design of the data collection approach on land rights issues, and specifically on women's land rights. The first comprises the definition of the objectives and purpose for which data are being collected. This informs the drafting of the key questions to ask. The second is the identification of who should ask the questions and the third element concerns how to ask questions (i.e. the unit of data collection). These elements ensure reliability of information and quality.

Data are often collected through various approaches and instruments. As highlighted above, data on women's land rights are often collected through surveys that use questionnaires administered directly to women as respondents or at the household level through the selection of the most knowledgeable household member or their proxy. A survey questionnaire consists of a series of questions for the purpose of gathering information from respondents.[11]

Administrative data on women's land rights are mainly collected through land registries and national cadastre systems at the country level, which are often not disaggregated by sex and frequently do not contain an individual ID. Without individual IDs, it is not possible to determine who are the people who have documented rights and who do not. For instance, cadastral maps allow for a more reliable measurement of parcels and plots of land, as compared to self-reported data from surveys, indicating the size of each parcel and plot. In some instances (and depending on the country land information system), cadastral maps may also contain further data elements that are important to inform land rights and women's land rights, thus providing evidence-based statistics for informing policymaking. However, they are limited in scope and unlikely to contain all those elements that are needed to compute SDG indicators on women's land rights. In this regard, in order to use data on land parcels from cadastres and registries for the computation of SDG indicators on women's land rights, it is necessary that the country develops a multipurpose cadastre to provide a full picture of the land information at the parcel level, including sex-disaggregated data on all potential land rights in each parcel of land.[12] There is also a variety of other methods that use mobile technologies to collect data on land rights and women's land rights, which include geographic information systems (GIS), according to which land areas can be measured and identified by a grid of cells.

Researchers, academics and evaluators, among others, conduct field and systematic reviews to mine data that support decision-making on women's land rights. Regardless of the data collection approach used, ensuring that key steps are followed in the planning and design stage is critical to enhancing the quality and reliability of the data produced. It is critical, however, to ensure that all data collection approaches support the fundamental principles of official statistics by ensuring data security and allowing disaggregation of data by sex, tenure and multiple and intersecting identities of women, e.g. rural, urban, indigenous and in conflict or post-conflict situations. This guarantees the quality of the data collected in adequately representing and profiling the issues that affect women's land rights.

Another critical element to ensure the quality of the collected data concerns pre-survey scoping research. Pre-survey scoping research is instrumental in ensuring that the data elements to be included during the phase of the survey design are aligned to the country context.

Recommendations for target population and sampling criteria

The precision and reliability of estimates is highly dependent on who and what is to be included in the sampling frame. A central element during the implementation of the sampling strategy is to ensure that:

a) Representation of the entire population is covered at the time of sampling, including marginalized and hidden groups (who to include).

b) All relevant tenure types and types of land use are included in the sample (what to include).

The implementation of a stratified two-stage sample strategy[13] is generally suitable to collect household survey data on land rights and women's land rights and ensure the statistical reliability of collected data. This strategy is based on a two-stage selection, which is: 1) to conduct a random selection of primary sampling units (PSUs), i.e. enumeration areas (EAs)[14] and 2) to conduct a random selection of secondary sampling units (SSUs), i.e. households.[15] It is important, however, to discuss some considerations on the stratification of both PSUs (what to include) and SSUs (who to include), to ensure that all key groups of the population are included in the final sample.

What to include?

- At the **first stage**, the smallest administrative geographical areas within a given stratum (typically administrative villages), i.e. enumeration areas, must be selected randomly, after stratification by a) ***type of residence (urban–rural)***, b) ***land use type***, and c) ***land tenure***, crossed by administrative districts.

Who to include?

At the **second stage**, SSUs, i.e. a fixed number of households, are selected from each of the selected enumeration areas, as previously stratified.

Gender differences in land rights for the entire population in a country are better assessed when key and relevant subpopulation groups are included in the sample of households. The inclusion of a large enough number of cases of relevant subpopulation groups will ensure that the extent of gender discrepancies in land rights are properly estimated. Therefore, it is strongly advisable that—at the second stage selection—key subpopulations that are more concerned with land rights are not missed out, particularly: those living in a) ***informal contexts***, b) ***conflict areas***, c) ***indigenous women*** and those living in d) ***communities***, among others. The number of cases for each key subpopulation group, as listed, should be large enough to ensure the precision and reliability of estimated statistics. A large enough number of cases for key subpopulation groups will reduce the margin of error of estimated indicators of women's land rights accordingly. **Annex I** contains a simplified scheme on key subgroups in each enumeration areas.

Concerning the selection of respondents within sampled households, two approaches are advised. The first and preferred approach is based on interviewing of household members (all age 18 years and over). This ensures that the information gathered through the survey will be self-reported rather than reported by proxy. The second approach consists in interviewing only one person in the sampled household, while the information for the other household members will be reported by the proxy respondent. In this case, it is recommended to select one person at random from all adult household members. This will ensure that women and men will have a non-zero probability to be interviewed.

Stage 1 – Step 2: Training enumerators, planning and deciding on roles and responsibilities

Training of enumerators and supervisors

One of the key requirements to ensure high-quality data is to ensure that those who collect the data are fully informed of the objective, content and structure of the data collection instruments. To do this, targeted training to supervisors, field enumerators (data collectors) and data clerks is critical to ensure quality control in the generation, coding and feeding of the data in the system before analysis. This training can be conducted by the data collection agency and/or outsourced to a service provider. A well-structured training manual should be prepared to this end.

Specific recommendations include:

- At least a **two-day training** is carried out before field operations take place to explain the key concepts and relate indicators and to make sure that enumerators are fully familiar with the survey questionnaire.
- Periodic or annual trainings should also be conducted in case new enumerators are selected or revision is made in the questionnaire design. In case the data collection process on WLRs statistics is expected to be implemented on a yearly basis, it is strongly advisable to conduct refresher training (or retraining) even if enumerators remain the same and there is no revision in the questionnaire. The retraining of enumerators is aimed at the recall and reinforcement of previously acquired knowledge and skills about data collection on women's land rights.
- Trainings are structured to allow for substantial time for discussion around the conceptualization of issues related to women's land rights, gender concepts and the translation of the concepts into measurable indicators, as well as potential tensions that may arise around the most challenging questions.
- Problems of communication during the phase of interview are addressed during the training: enumerators must approach respondents gently and speak the same language as the respondent.
- Enumerators work in small teams that are each guided by a well-trained supervisor.

During the training:

- Enumerators need to be extremely clear on the different tenure types and the many ways in which people may refer to each tenure arrangement. For example, in case of paper-based data collection, enumerators must be provided with an "enumerator manual", i.e. a document which details each of the questions contained in the survey questionnaire, which clarifies the meanings underlying each question and the way enumerators should record the corresponding information, including which questions to ask next, or which questions must be skipped. If data collection is implemented through the use of mobile devices, devices should include the function of dynamic help which allows enumerators to access help at any time and at any stage during the interview.
- Enumerators must also be provided with visual aids to identify country-specific types of legally recognized documents that are potentially held by household members that would prove whether the respondent is the owner or use right holder of the land (e.g. a photo of an actual document or the reproduction of a facsimile). These visual aids are meant to be shown to the respondent during the interview.

Stage 1 – Step 3: Survey testing and testing of data process

Once the training is completed, the field supervisors and enumerators must do a field test of the data collection instruments (e.g. questionnaires) as well as the entire survey process on a small sample in order to check the validity and reliability of the data generated and the correctness of questions asked and different survey stages. Two types of tests can be applied before the implementation of the actual survey, depending on the financial and operational capacity of the implementing agency:

- The **cognitive test** addresses problems related to the wording of survey questions and ensures a timely revision of questions contained in the designed survey. This ensures that aspects of comprehension of questions by the respondents, language ambiguity and translation issues are identified and properly addressed before the actual data collection starts. Even though the cognitive test is strongly recommended to ensure the correctness of questions asked and to assess respondents' general cognitive ability to answer questions, it can be costly. As such, this test should be implemented only if the implementing agency has the sufficient financial and operational capacity to implement the test.

BOX 1.

The importance of cognitive testing for probing perception of tenure security

Why cognitive testing?

Perception of tenure security: assessing the cognitive ability of respondents

The concept of perception of tenure security is well-defined within the metadata of SDG indicator 1.4.2. Perceived security of tenure is defined as an *"individual's perception of the likelihood of involuntary loss of land, such as disagreement of the ownership rights over land or ability to use it, regardless of the formal status and can be more optimistic or pessimistic. Although those without land rights' documentation may frequently be perceived to be under threat, and those with documentation perceived as protected, there may be situations where documented land rights alone are insufficient to guarantee tenure security."*[16]

It is important to emphasize that perceptions of tenure security need to be self-reported i.e. reported by those holding the right to a piece of land. They cannot be reported by someone else in the household. This issue is absolutely critical.

For the purpose of constructing sub-indicator 1.4.2(b), perceptions of tenure security are based on the fear of involuntary loss of the land within the next five years.

This question, designed for the land module developed for indicator 1.4.2 by custodian agencies, ensures that the survey captures perceptions of tenure security. The question asks sampled respondents, *"On a scale from 1 to 5, where 1 is not at all likely and 5 is extremely likely, how likely is [NAME of owner/useright holder] to involuntarily lose ownership or use rights to this [PARCEL] in the next 5 years?"*

Recognizing that perceptions of tenure security are inherently context-specific requires probing this question during cognitive interviews to explore the potential drivers of these perceptions and the way the term "perception" is understood by respondents. Cognitive testing of this question informs the redesign of the survey questionnaire by making the question on perception of tenure security less general and more country-specific to ensure that collected data are more aligned to the context where the survey is conducted.

A number of elements concerning the "perception of tenure security" are suggested to be tested during cognitive interviews, as reported below. The results from cognitive tests can ultimately be used by survey designers to reformulate the question.

- What does "perception" mean to you? Could you repeat this question in your own words? These questions help to probe whether the respondent understood the subjectivity element behind the survey question, regardless of the existence of any document.
- Please list all the threats that you perceive as influencing the involuntary loss of your land.
- How can you make sure that no threats will be experienced in the next five years?
- Please describe, using examples, each of the main threats that you may experience in the next five years.
- Please explain why you gave a score of "x". What has driven your decision for that score?

- **Pilot survey.** The pilot survey presents advantages to ensure the identification of data elements that might have been missed out during the design phase. The pilot survey can be administered to a purposive sample of a limited amount of households, in a few selected districts, using the sampling strategy that will be applied to the actual large-scale survey. It is important that the testing of the pilot survey is not limited to data collection tools, i.e. testing of the questionnaire, but it also extends to the testing of the **full survey** operations that are implemented to determine whether problems exist and need to be addressed prior to putting the final actual survey in the field. In case of data collection with laptop computers or handheld devices, the pilot survey should test relevant issues like data transmission, data storage and security and processing programmes. It is important that these tests are accurately and properly conducted before the implementation of the actual, final survey since, if not functioning, they can generate serious problems that will overall affect the reliability of the entire survey process.

It is advisable that survey design and data collection, including the cognitive and/or pilot tests and the actual large-scale data collection, embrace participatory approaches that involve and incorporate the voices of other stakeholders in key decision making. This includes coordination of the implementing institution—with agencies such as non-governmental organizations (NGO), universities, grassroot organizations—and the national statistics offices, with technical support from custodian agencies of SDG indicators, if needed.

BOX 2

The importance of conducting cognitive and pilot tests

Main differences between cognitive and pilot tests

To promote efficiency in conducting surveys, researchers generally perform a pilot survey, i.e. a small study conducted over a subsample of typically 300 to 400 households. While survey pilot testing (see Hassan, Schattner and Mazza 2006) focuses on issues related to the research protocols, appropriateness of data collection instruments (i.e. the questionnaire), sample strategies and operational and field-related matters, cognitive testing is mainly implemented to focus on problems related to the wording of survey questions (see ILO 2018).

The overarching goal of the cognitive testing is to check whether questions asked capture what they intend to measure. This especially applies to questionnaires aimed at collecting complex information, as in the case of questions asked to construct SDG and non-SDG indicators on land rights and women's land rights. This is because respondents may have different interpretations of the meanings of the questions or have memory and recall problems, and these issues are potential sources of errors in surveys and affect the overall reliability of the collected data.

The implementation of both cognitive tests and pilot surveys will therefore increase the reliability of data and indicators, and they represent key steps prior to the implementation of the final large-scale survey.

According to Tourangeau (1984), cognitive testing is aimed at assessing the correctness of the survey questions vis-à-vis the four elements that constitute a potential source of response error. These are:

1. **Comprehension:** to ascertain whether the questions asked are fully understood by the respondents. For example, problems of comprehension arise if respondents are not familiar with some of the key terms used in the survey questionnaire, or if questions asked are long and/or complex.
2. **Recall:** to ascertain the respondent's potential difficulty in recalling the requested information.

For example, respondents might not recall with precision the information requested, thus leading to recall errors.

3. **Judgemental:** to explore a potential alteration in the response due to perceived social expectation. The judgemental element of the cognitive test is of particular importance in the context of data relating to women's land rights, particularly in situations where social and cultural norms prescribe female submission to the men.
4. Whether the response options **fully fit with the respondent's desired answer**. For example, to check for incomplete response options, ambiguous options and mismatched options.

Concerning the implementation of the cognitive test:

- It is typically conducted by selecting key questions from the designed questionnaire, which are tested through in-depth, semi-structured interviews. The suggested interviewing technique is the "verbal probing" approach (Beatty et al. 2007), which is defined as paraphrasing, e.g., "could you repeat this question in your own words?" and a variety of probes such as, "what does 'x' mean to you?" or "you hesitated during answering, could you tell me why?"
- The cognitive test is administered to a very small subsample, typically 30 respondents, to enable assessment of how well questions work for different types of respondents. This implies that different profiles are required to test the soundness of the designed questionnaire. The respondent's characteristics should include, but are not limited to age, sex and level of education attained.
- As the cognitive test involves a small number of respondents, the geographical spread of the sample may be restricted to limited number of districts/regions (i.e. a maximum of 1 or 2 districts).
- Inputs gathered from the small subsample of respondents during the cognitive test will finally be used to fine-tune the questionnaire from a design, flow and comprehension perspective. This will ensure that questions asked during the implementation of the survey are sufficient and fully understood by all sampled respondents.

Key roles of a lead data agency

The lead data agency could be an organization or institution such as a national statistics office, a private sector organization, a civil society organization, an international non-governmental organization, a research entity, or a partnership between these bodies. In general, the **lead data collection agency is responsible for:**

- **Overall design of the data collection approach** and tools and securing authorization for the research.
- **Making sure that interview respondents are stratified** appropriately and a large enough case of key groups are included in the sample.
- **Ensuring translation** of the survey questionnaire into local languages (if needed).
- **Identifying and ensuring the recruitment of the field workforce**, i.e. interviewers, enumerators and supervisors.
- **Training of field personnel** including enumerators.
- **Coordinating test activities and survey implementation at the country level**, including coordinating logistics and assigning interviewers to households in each cluster or primary sampling unit, making travel arrangements for data collectors, informing local authorities about the survey, supervising the interview process and recording daily activities.
- **Overseeing the testing exercises** and ensuring their timely completion.
- **Ensuring the quality** of the fieldwork.

- **Consolidating the results** from the cognitive and pilot tests to inform actual data collection.
- **Evaluation and quality control:** the supervisors should check that data have been recorded properly by the enumerators in their team and ensure proper follow-up for unavailable selected participants and unfinished interviews.
- **Analysis.**
- **Reporting.**
- **Data sharing.**
- **Data archiving and security,** i.e. moving data to a separate storage device for long-term retention and ensuring the protection of data that have been archived in a database from destructive forces and from the use of unauthorized users who should not access confidential information about respondents.

The **implementing agency** should be responsible for:

- Coordinating the daily work, i.e. tracking the participants' recruitment and selection.
- Planning for interviews to be conducted in a neutral environment.
- Reporting on the daily updates of the fieldwork back to the focal point/national consultant.

In the context of the SDGs, **custodian agencies responsible for land indicators** have specified roles and responsibilities as defined by IAEG-SDGs.[17]

Stage 2 – Step 1: Data collection

Once the survey has been designed and survey testing and training activities conducted, the final large-scale survey has to be implemented. Before and during implementation of this survey, it is important that:

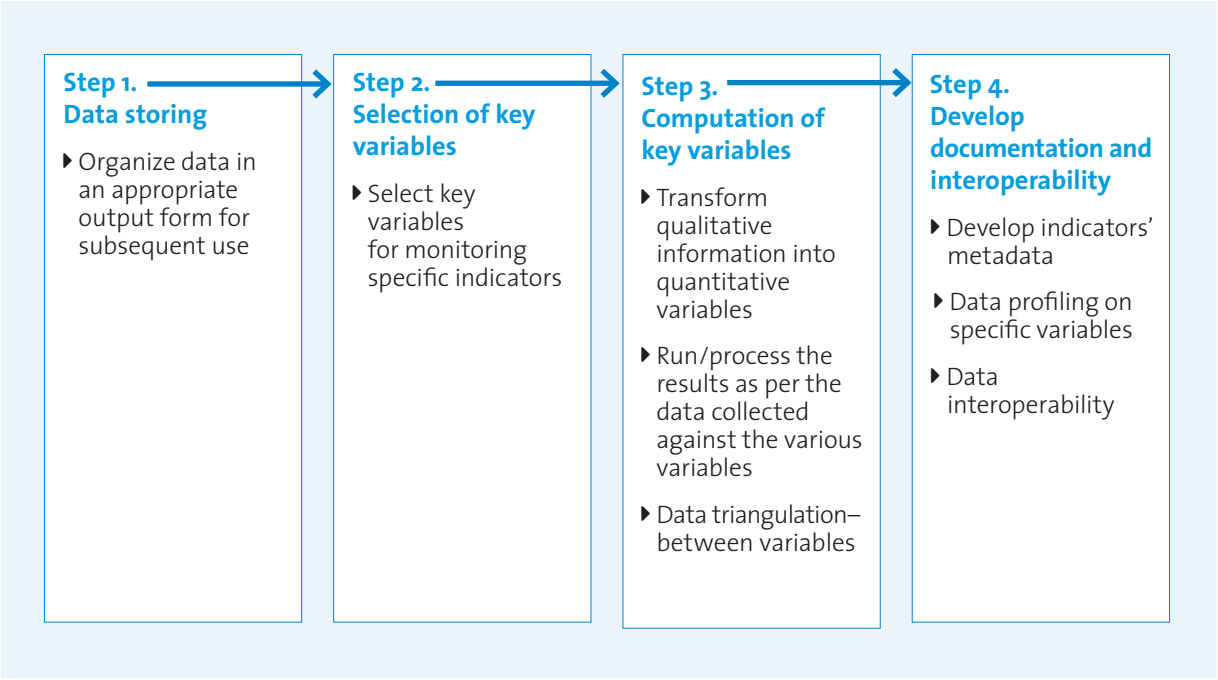
- Survey teams are well trained on the survey content (see Step 2).
- There is clear agreement on the time frame for completing the survey.
- Roles and responsibilities among team members are well defined (see Step 3).
- Supervisors of each survey team ask permission from the enumeration area authorities (e.g. the village chiefs) prior to implementing the survey. Supervisors should inform the village head, clarify that data collected will remain strictly confidential and explain the purpose of the survey, the way the survey will be conducted, the outcome of the survey and how results will be used.
- A list of replacement households is selected to avoid potential delays while conducting the survey.
- **Enumerators interview women separately from men** and in places where they are free to talk in order to avoid situations where women will risk reprisals from their husbands or relatives for having been interviewed, typically in situations where social and cultural norms prescribe male dominance.

The time frame for completing the data collection should be agreed in advance. It will largely depend on availability of human and financial resources and informed by the purpose of which data are being collected.

STAGE 3:
DATA PROCESSING, MANAGEMENT AND ANALYSIS

Once data have been collected from the field and validated, it is then coded and fed into a data system (e.g. SPSS, EPI Data, Microsoft Excel, Stata). These data processing systems are also used for analysing and presenting data in ways that can easily be understood. Often, data go through four steps after they have been collected from the field, as presented in Figure 3.

FIGURE 3
From data storing to data analysis



Firstly, once field operation activities have been completed, data from the field must be verified, organized, transformed, integrated and extracted in an appropriate output form for subsequent use. Secondly, key variables should be selected for monitoring specific indicators. Thirdly, primary variables to construct selected SDG and non-SDG indicators on land rights and women's land rights must be computed using raw data from the field, as organized in a specific output form (see Step 1 in Figure 2). Finally, appropriate metadata for the selected indicators must be developed to inform data users about the content of the indicators themselves.

Stage 3 – Step 1: Data storing

This step of data storing only applies to data collected through paper-and-pencil interviewing (PAPI).

Data collection conducted through computer-assisted personal interviewing (CAPI) and mobile-assisted personal interviewing (MAPI) methods is an interviewing technique in which the respondent and interviewer use an electronic device (a laptop or handheld device) to ask and answer the questions. Survey implementation through CAPI and MAPI interviewing has a number of advantages as compared to questionnaires on paper (PAPI). The most important one is that data collected through CAPI and MAPI are immediately available in electronic format, thus avoiding the overbounding operations of data entry. These interview methods can, however, be more expensive as they require the purchase of electronic devices (i.e. laptops for CAPI and handheld devices for MAPI) before embarking on data collection activities.

By contrast, data entry operations must be performed for implementing the PAPI method. Once data collection activities have been completed, questionnaires are returned to the NSO (or the implementing agency) so that data entry and data storing operations can start. Before starting the data entry operation and immediately upon receipt of the questionnaires from the field, the supervisor in charge of data collection operations must verify the contents of the shipment and questionnaires must be accurately stored. Upon arrival, there should be a designated area within the storage room, or a small room adjacent to the storage room, that can accommodate these questionnaires.

For each lot of surveys from a new enumeration area received, the enumerators' supervisor should conduct the following checks, in the order indicated:

- Ensure all filled questionnaires from a given enumeration area are sorted in ascending order by designated numbers.
- Ensure all geographic and administrative identification codes are correctly completed on the cover sheet of the questionnaire, for example, by ensuring that no household shares the exact combination of identifying codes with another household.
- Ensure the correct number of data collection instruments (questionnaires) allocated to specific enumeration areas are accounted for, including those that were partially completed, refused or otherwise unfinished.

Once the checks above have been performed, data must be entered into a computerized database or spreadsheet. This operation is performed only if data have been collected through face-to-face interview questions through PAPI.

Data entry operations must also be handled with care to ensure accurate information is filled in. This exercise is often done by data clerks who are trained in various computer packages for data management and processing. The double data entry technique is sometimes implemented. To ensure data entry is well managed, data agencies need to ensure that:

- The data entry forms to be used by enumerators for entering the collected data are well-developed in advance to ensure they are organized according to the flow of the questionnaire.
- Data entry forms should contain all variables that must be filled with data collected and accompanied by appropriate documentation, such as a codebook explaining:

1. *Variable:* i.e. the variable name as reported in the data entry spreadsheet
2. *Question number:* i.e. the question number as reported in the paper questionnaire
3. *Question:* i.e. the verbatim question as reported in the paper questionnaire
4. *Variable description:* a concise description of the variable
5. *Type of variable:* whether the variable must be filled in the data entry template as a string or as a numeric variable
6. *Numeric codes allowed:* each question asked in the questionnaire has a corresponding numeric code. For instance, female takes on value 2 while male takes on value 1. The numeric code is described for each variable and prevents data entry mistakes. It also informs data entry operators

whether (or not) missing values are allowed for the variable under investigation

7. *Labels:* associated with numeric codes, i.e. the value label that allows converting the numeric code associated with a given variable with the corresponding label.

Annex 2 contains an example of the Excel spreadsheet for data entry operations and the corresponding codebook.

Important note: For CAPI, MAPI or GIS data, computer systems must design and respond to data capture, management and processing; and must also be fully tested before deployment for actual data collection.

Stage 3 – Steps 2 and 3: From formulation to computation of variables for analysis

The qualitative information collected through surveys (properly organized after data collection operations) must then be transformed in appropriate quantitative primary variables, ultimately used to compute the final indicator(s) on land rights and women's land rights. While the methodology to measure SDG indicators 1.4.2 and 5.a.1 has already been developed and approved,[18] qualitative information collected through surveys has the potential to derive other variables (other than those used for constructing SDG indicators) which, in turn, can be used to compute other non-SDG indicators (see Tables 1 to 4).

Data analysis needs to reflect the original aims and objectives for which data and statistics are being collected. To help structure the analysis, the components of selected indicators, i.e. variables, must be identified *a priori as informed by the original objectives of the surveys,* to ensure that the qualitative information from the survey is properly transformed into quantitative variables. The computation of variables for constructing and measuring women's land rights indicators is better organized according to two steps, as follows:

1. In the first step, a set of scripts and procedures, typically done with statistical software such as Stata, are applied to the survey data for obtaining primary variables. Primary variables are vectors of n observations, with n capturing the total number of individuals in the sample for which the variable is constructed. The computation of primary variables is typically done by combining the qualitative information collected through questions asked in the survey. For instance, in the context of SDG 1.4.2 (sub-indicator b), the aim is to compute a primary variable that accounts for the total adult population who perceive their rights to land as secure. To this aim, the qualitative information collected from questions asked through the survey must be combined. Such a primary variable is computed by inferring information on a) fear of involuntary loss of the land within the next five years, and b) the landholder's right to bequeath the land, and it aims at capturing individuals who have a perception of tenure security on at least one plot. Figure 4 below shows an example of how to compute the primary variable on "perception of tenure security".

   The computation of primary variables must reflect the components associated with each indicator being monitored.

   It is important that a number of key variables—such as age, rural/urban residence, educational attainment, labour force participation, participation in the agricultural sector and marital status—that are relevant to profiling the most vulnerable and excluded in land rights—from a gender perspective—are also computed. The computation of the above-listed variables is a crucial step in data analysis: such variables allow producing WLRs indicators that can be further cross-tabulated in order to produce indicators that not only illustrate the main gaps in land rights between women and men, but also among various subgroups of women, especially those that are vulnerable or not visible in general statistics.

2. In the second step, WLRs indicators are computed by combing the primary variables. This is to say that a given WLRs indicator is generally comprised of more than one primary variable, whose combination allows deriving the final indicator.

FIGURE 4

From variable formulation to variable computation

Step 1.

- Extract qualitative information collected through survey questions

Question 1: Do you have the right to bequeath this [PARCEL], either alone or with someone else?

Question 2: On a scale from 1 to 5, where 1 is not at all likely and 5 is extremely likely, how likely is [NAME of owner/use right holder] to involuntarily lose ownership or use rights to this [PARCEL] in the next five years?

Step 2.

- Compute primary variable(s): «perception of tenure security»

Combine information from questions # 1 and 2 to compute primary variable on «perception of tenure security»

Step 3.

- Check the content of the computed primary variable.

For example «perception of tenure security equals 1» if:

1. s/he has the right to bequeth the parcel; and/or
2. s/he is not at all likely to involuntarily lose ownership or use rights to the parcel
3. otherwise 0

The final primary variable will look like this:

	Value code	Value label
Individual identifier (code)	Var Perception	Tenure Security
1	1	yes
2	1	yes
3	0	no
...	0	no
101	0	no
102	1	yes

How to clean up potential errors from collected data

Household survey data are prone to various data entry mistakes, which may result in a considerable amount of outliers and inconsistencies in the raw data. Outliers can be defined as extreme values that deviate from other observations on data; they may indicate a variability in a measurement, experimental errors or a novelty. Simply stated, an outlier is an observation within a given variable that diverges from an overall pattern on a sample, i.e. single data points that lay far from the rest of the distribution. Observations that are detected as outliers may seriously affect data reliability and, therefore, must be replaced with appropriate and more reliable values. A natural concern that arises in the analysis of data points that are outliers is the evaluation of the sensitivity of the variables with respect to a given threshold, above or below which these data points can be considered outliers. Since an unavoidable degree of arbitrariness is generally involved in this kind of analysis—what is the threshold that must be set in order for a data point to be an outlier?—the analysis of outliers is generally done according to the following procedure, which ensures a more scientific way to outlier detection and imputation. In particular, the detection of outliers only applies to SDG data that are not binary variables[19] and is implemented by benchmarking each data point against the median value: an outlier is defined as a value higher or lower

than 3 Standard Deviations (SD) from the median,[20] as specified below:

$Outlier_{i,v} < Median_{v} - 3 * SD_{v}$ [1]
$Outlier_{i,v} > Median_{v} + 3 * SD_{v}$ [2]

Where is the value of variable "v" for the i-th individual, is the standard deviation of the value of variable "v", Median is the value of variable "v" corresponding to the 50th percentile, and finally, 3 is a constant that multiplies the standard deviation. Values of variable "v" higher or lower than 3 Standard Deviations from the median are identified as "outliers". Data points identified as outliers are typically replaced with the median value of the variable of interest, even though the replacement procedure rests upon the decision of the data analyst.

Procedures for outlier detection and imputation should be implemented on all continuous variables, e.g. the value of land or the area of land at the lowest level of aggregation (i.e. after computation of primary variables).

Stage 3 – Step 4: Develop data documentation and interoperability

Once indicators on women's land rights have been computed, it is important that they are accompanied by metadata for interoperability for data users to understand a) the context of data collection, b) the data collection methodology, c) data validation and quality assurance, d) data manipulations through data analysis of raw data, and e) data confidentiality, access and use conditions. The development of metadata documentation is generally suitable for this aim. The metadata instrument should be used to provide detailed information about the computed indicators and describe the way indicators on land rights and women's land rights stored in a database were acquired and the method of compilation and processing, among other information.

The following elements are suggested to be associated while developing LRs and WLRs indicators' metadata:

Main data characteristics

1. *Data description*: What does the data set contain (i.e. the complete list of LRs and WLRs indicators that can be found)?
2. *Name of the indicator*
3. *Statistical concepts and definitions:* Concepts and definitions associated with each indicator and with each indicator's component. It is important to clearly explain the indicators' components that are country-specific.
4. *Data description:* Detailed information on what the indicator is meant to monitor.
5. *Time coverage:* The years covered by the indicator, e.g. from 2001 to 2018.
6. *Periodicity:* The recall period of the indicator, e.g. annual, monthly.
7. *Reference period:* Can be the calendar year, the last 12 months prior to the date of the interview, etc.
8. *Reference area:* The country(ies)/enumeration area(s) covered by the indicator.
9. *Unit of measure:* The household, the individual, the land, the value of land, etc.
10. *Unit of measure codes:* e.g. hectares, acres. The unit of measure is not applicable to households and individuals.

Frequency of disseminated data

1. *Frequency of dissemination:* When data are going to be disseminated, e.g. "every year" (if data collection is conducted continuously) or "not available" if frequency depends on the source of the survey in a given country.

Source of the indicator

1. *Type of source:* i.e. primary or secondary data.
2. *Specific source:* Administrative, census or survey data.
3. *Name of source:* e.g. NSOs, LSMSs/DHS, UN agencies, CSOs.

Method of computation

1. *Calculation of the final indicator:* e.g. the indicator is constructed as the weighted mean of total adults with secure tenure rights divided by the total adult population in the country.
2. *Data adjustment:* Description of any outlier detection and imputation procedure applied to the computation of final indicators. It is, however, important to clarify that data points can be imputed as long as data analysts are comfortable in assuming that the missing information is similar to the information in the data set. For example, women's tenure security need not mirror their husbands' tenure security. If and until we ask women, we will not know how much they differ.

STAGE 4:
REPORTING AND UPTAKE

Once data have been collected and indicators constructed, the final stage of the data value chain process involves sharing data and using them to track progress and inform planning and key decision-making processes.

Data can be presented in various formats that are easy to understand, including their integration in the country profiles or fact sheets on land rights and women's land rights. It is important that the country profile clearly sets out the indicators that are used to monitor women's land rights and that the selected indicators focus particularly on the continuum of land rights, rather than solely on land ownership. This will also ensure that land access, control and ownership rights are documented and monitored including women's land rights in the context of informal, cultural and other tenure regimes where ownership may not be applicable; and that no one is left behind.

It is advisable to start the country profile by means of information contained in SDG Indicator 5.a.2, which provides a basis for policy measures aimed at securing equal opportunities and access to rights and resources. The country profile is better developed through description of the country's legal and policy framework of reference, including customary law, which implies answering the following questions:

What are the main national policies and laws that reflect good practices in guaranteeing women's equal rights to land ownership and/or control in the country, if any?

Has any reform to give women equal rights to economic resources been recently implemented?

Has any reform that gives women equal rights in terms of access to, ownership and control over land and other forms of property, financial services, inheritance and natural resources, in accordance with national laws, been implemented in the country?

- What land tenure types do the policy and legal framework recognize (e.g. public, private, customary, communal, tribal)?
- Does the policy and legal framework recognize the equal rights of men and women?
- Is any legal mechanism in place that makes it illegal to discriminate based on gender in relation to land rights?
- Do secure tenure rights vary widely between different regions and areas (rural/urban) in the country/enumeration area?
- How are secure tenure rights to land correlated with sex, age, urban and rural, racial, or ethnic characteristics?
- What are the socioeconomic, legal, cultural and behavioural constraints limiting women's ability to claim their land rights?
- What are the socioeconomic characteristics of those women facing disadvantages in securing land rights?
- How are women's land rights linked to malnutrition or educational outcomes?
- Are certain women groups at a higher risk of facing disadvantages in securing land rights than others are? Who are they (e.g. indigenous/non-indigenous people)?
- What land do the poor own?
- Do legally recognized property rights over such land exist? Do women and men have the same legally recognized property rights? How secure is their access to, and tenure over, land?

The complex mix of information related to both the country's legal and policy framework and women's land rights outcomes—documented in the reporting stage—are expected to convey accurate information of "what works" and "for whom", e.g. performance monitoring and impact evaluation of the policies and programmes implemented. It is pointless, in fact, to have a sound profiling system able to "predict" risks faced by women in securing land rights with a certain accuracy and to then refer to ineffective policies and programmes.

How to present the country profile

The country profile can be presented in two different ways. The first way comprises presenting the incidence of access to land and secure tenure rights for men and women, by land tenure type and land use type. The second way involves presenting the characteristics of those men and women who can and cannot access land and have or do not have secure tenure rights over land. In order to provide a universal picture of the disadvantages faced by women in accessing land and securing land rights as compared to men, and monitoring the evolution over time, this will be useful as it allows targeting women more specifically. The second way, instead, is more concerned with the characteristics of those women who face disadvantages in securing land rights. As such, it allows targeting a specific subgroup of the female population. For example, it is useful to disaggregate WLRs statistics by income quintile, poverty status, women who do or do not engage in agriculture, rural/urban women, etc.

In general, it is advisable that both ways are included in the country profile in order to enable decision makers to be fully informed by data.

How to incentivize the use of data for decision-making

The final step of the data value chain process involves incentivizing the uptake of data by decision and policymakers. Incentives can be either direct or indirect. The direct way is based on the organization of workshops or seminars to raise awareness about land rights and women's land rights in the country or enumeration area under scrutiny. The indirect way comprises the publication of indicators and country profiles in a dedicated data repository or website.

CONCLUSION

This guidance tool has described the data value chain process for SDG and non-SDG indicators on land rights and women's land rights, in particular.

In particular, it explains the practical steps that should be implemented for the production of reliable data and indicators on land rights and women's land rights, ranging from planning and design of the survey to data collection, analysis and reporting.

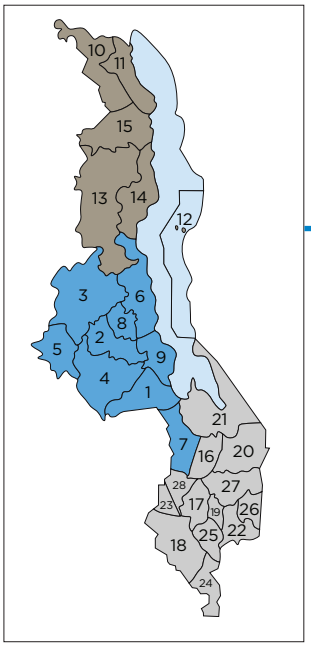
To understand women's land rights in a given country and/or context, key stakeholders should be involved during the planning phase, and the objectives, purpose and sampling strategy of the survey must be clearly defined and implemented during the design stage. Once the survey has been designed and activities planned, it is recommended that enumerators are properly trained, and roles and responsibilities fully agreed, before embarking on field operations for data collection. Ensuring high-quality data on women's land rights also requires the correctness of data processing, management and analysis, which must be done according to those basic principles, ranging from the validation and organization of collected data in appropriate output form, to the construction of the final WLRs indicators.

Quantitative and qualitative indicators constructed by using data collected from the field serve the purpose of reporting on the country's situation about women's land rights while considering recently implemented policies and reforms and their potential impact on the reduction of gender gaps in land rights and tenure security.

ANNEXES

ANNEX I
Sampling strategy

Total number of geographical districts (e.g. 32 districts in Malawi)



Stratify each district by rural and urban areas

Stratify each district by land tenure type and land use, crossed by rural/urban areas

First stage random selection:
Select a number of enumeration areas (EA) within each stratum (e.g. rural area of district #4 with agricultural land under customary land tenure)

Second stage random selection:
Select a fixed number of households from each EA and ensure key subgroups of households are included, especially:

- Households living in community
- Indigenous households
- Households living in informal context

ANNEX II

Example of spreadsheet for data entry operations and related code book

Spreadsheet for data operations

	A	B	C	D	E
1	Variable_name				
2	Household_id	Respondent_id	Parcel_identification_numb	Tenure_system	Parcel_use
3			1		
4			2		
5			3		
6			4		
7			5		

Code book

	C	D	E	F	G
1	**Question**	**Variable description**	**Type of variable**	**Numeric codes allowed**	**Labels associated with numeric codes**
2	Household unique identification number	Code associated with each sampled household	Numeric/general	From 1 to N	Not applicable
3	Respondent unique identification number	Code associated with each sampled individual within the household	Numeric/general	From 1 to N	Not applicable
4	Respondent unique identification number	Please tell me about each parcel for which you currently own or hold use rights for, either alone or with someone else. Please describe or give me the name of each parcel, starting with the parcel you reside on, if applicable.	Numeric/general	From 1 to N	Not applicable
5	Under which tenure system is this [PARCEL]?	Type of tenure associated with the parcel	Numeric/general	From 1 to 7	CUSTOMARY ..... 1 FREEHOLD ..... 2 LEASEHOLD ..... 3 STATE ..... 4 COMMUNITY/GROUP RIGHTS ..... 5 COOPERATIVES ..... 6 OTHER (SPECIFY) 7
6	What is the primary current use of this [PARCEL]?	Current use of the parcel	Numeric/general	From 1 to 6	RESIDENTIAL ..... 1 AGRICULTURAL ..... 2 PASTORAL ..... 3 FOREST ..... 4 BUSINESS/COMMERCIAL ..... 5 OTHER (SPECIFY) 6

ENDNOTES

1 See UN, undated, "Sustainable Development Goals". https://www.un.org/sustainabledevelopment/sustainable-development-goals/

2 UN Women 2013

3 The Custodian Agencies of the three SDG indicators discussed in this guidance tool—the United Nations Food and Agriculture Organization (FAO), the United Nations Human Settlements Programme (UN-Habitat) and the World Bank—have developed methodologies for data collection, which have been approved by the Inter-agency and Expert Group on SDG Indicators (IAEG-SDGs). These methodologies provide a comprehensive analysis of the critical issues, making this guidance tool a complementary one to ensure that key gaps are addressed in practice (links to joint module for 1.4.2, 5.a.1 and 5.a.2).

4 For more details, see https://sdgs.un.org/

5 Specific metadata for 5.a.2 and most updated is available via: https://unstats.un.org/sdgs/metadata/files/Metadata-05-0A-02.pdf

6 See for example Global Land Tool Network's "Measuring Individuals' Rights to Land; An Integrated Approach to Data Collection for SDG Indicators 1.4.2 and 5.a." available at: https://gltn.net/2019/08/27/measuring-individuals-rights-to-land/

7 UN 2019

8 EGM 2017

9 See for example, Slavchevska et al. 2017.

10 Slavchevska et al. 2017

11 National statistical organizations and other survey agencies including Demographic Health Surveys (DHS), Living Measurement Study Surveys (LSMS) and the Agricultural Integrated Survey (AGRIS) apply survey methods including computer-assisted personal interviewing (CAPI) to collect data on land rights.

12 Mondal et al. 2016

13 Two-stage stratified random selections are typically implemented in both LSMS and DHS. See Aliaga and Ren 2006.

14 Enumeration areas represent the smallest geographical areas within a stratified stratum.

15 The details behind the implementation of a two-stage sampling strategy can be found at: https://dhsprogram.com/pubs/pdf/WP30/WP30.pdf

16 World Bank and UN Habitat 2017.

17 For more details, visit: https://unstats.un.org/sdgs/indicators/indicators-list/

18 For more details, visit: http://www.fao.org/3/ca4885en/CA4885EN.pdf and https://unstats.un.org/sdgs/metadata/files/Metadata-01-04-02.pdf) and https://unstats.un.org/sdgs/metadata/files/Metadata-05-0a-01.pdf

19 It is important to note that outlier detection cannot be implemented when the scrutinized variable is a dummy one taking on value 1 or 0. It only applies to continuous variables.

20 In the literature, a number of outlier detection techniques exist that can be implemented. The one described in this guidance tool is the most common but can be substituted by many others. The choice of the outlier detection technique remains, therefore, up to the discretion of the data analyst.

REFERENCES

Aliaga, A. and R. Ren. 2006. Optimal Sample Sizes for Two-stage Cluster Sampling in Demographic and Health Surveys. DHS working papers 2006, No. 30.

Beatty, P. et al. 2007. "Research synthesis: The practice of cognitive interviewing". *Public Opinion Quarterly* 71(2): 287–311. doi:https://doi.org/10.1093/poq/nfm006

FAO, World Bank and UN-Habitat. 2019. Measuring Individuals' Rights to Land: An Integrated Approach to Data Collection for SDG Indicators 1.4.2 and 5.a.1. Washington, DC: World Bank.

Hassan, Z.A., P. Schattner and D. Mazza. 2006. "Doing A Pilot Study: Why Is It Essential?" *Malays Fam Physician* 1 (2-3): 70–73.

International Labour Organization. 2018. ILO LFS pilot studies – Cognitive interviewing tests: Methodology, process and outcomes. Geneva: ILO Department of Statistics.

Mondal S., D. Chakravarty, J. Bandyopadhyay and K. Maiti Kanti. 2016. GIS based Land Information System using Cadastral model: A case study of Tirat and Chalbalpur rural region of Raniganj in Barddhaman district. Switzerland: Springer International Publishing.

OHCHR and UN Women. 2013. R*ealizing Women's Rights to Land and Other Productive Resources*. New York and Geneva: UN Women.

Slavchevska, V., A.P. De la O Campos, C. Brunelli and C. Doss. 2017. Beyond Ownership: Tracking progress on women's land rights in Sub-Saharan Africa. Paper presented at the World Bank Land and Poverty Conference, Washington DC, March 20–24.

Tourangeau, R. 1984. "Cognitive science and survey methods: a cognitive perspective". In *Cognitive Aspects of Survey Methodology: Building a Bridge between Disciplines*. T. Jabine and J. Tanur (Eds). Washington DC: National Academy Press: 73–100.

United Nations. 2019. Guidelines for Producing Statistics on Asset Ownership from a Gender Perspective. Department of Economic and Social Affairs, Statistical Division. Studies in methods. Series F No. 119. New York: United Nations.

World Bank, UN Habitat, IPRA-CINDER, University Pompeu Fabra (UPF) and Global Land Indicators Initiative (GLII)/Global Land Tool Network (GLTN). 2017. Expert Group Meeting: Using Administrative Data to Monitor SDG land Indicator. Proceedings.

GLTN
GLOBAL LAND TOOL NETWORK

NOF, UN Complex, UN Avenue, Gigiri
P.O. 30030, Nairobi 00100, Kenya
Tel: +254 20 76 5199
Fax: +254 20 762 4256

www.gltn.net
www.twitter.com/GLTNnews
www.facebook.com/landgltn
https://www.youtube.com/LandGLTN
http://www.flickr.com/photos/landgltn/

UN HABITAT
FOR A BETTER URBAN FUTURE

NOF, UN Complex, UN Avenue, Gigiri
P.O. 30030, Nairobi 00100, Kenya
Tel: +254 20 76 23120
Fax: +254 20 762 4266
www.unhabitat.org
www.facebook.com/unhabitat
www.twitter.com/unhabitat
https://www.youtube.com/
unhabitatglobal
https://www.flickr.com/unhabitat/

UN WOMEN

220 East 42nd Street
New York, New York 10017, USA
Tel: 212-906-6400
Fax: 212-906-6705

www.unwomen.org
www.facebook.com/unwomen
www.twitter.com/un_women
www.youtube.com/unwomen
www.flickr.com/unwomen