

An Overview of the Adolescent Data Hub (ADH)

The Adolescent Data Hub (ADH) is a catalog of open-access datasets on adolescents and young people from low- and middle- income countries (LMICs). It contains links to data sources, from which users can download or request to download data.

How Does it Stand out in the Adolescent Health and Development Space?

The Population Council's GIRL Center created the ADH in 2018 with the goal of improving availability and accuracy of gender disaggregated data. The first and largest of its kind to catalog adolescent-focused data in LMICs, the tool aims to facilitate data sharing, connect researchers to existing data, and provide insight into geographic and topic data availability. Additionally, the ADH can serve as a tool for donors and implementers who want to integrate data into their practice and decision-making.

Why Open Data?

Open data catalogs facilitate secondary data analysis by reducing the time and effort researchers must put into identifying open access data on topics of interest. This may be particularly helpful to early-career researchers who may not yet have the institutional funding or networks to gain access to databases or collect their own data. Additionally, open data catalogs enable further analysis of existing data, encourages research collaborations, promotes research transparency, and ensures resources are additive rather than duplicative.

How Does the ADH Foster Transparency and Collaboration in Research?

Ultimately, the ADH serves to further a transparent and collaborative research environment. Other contributions include:

- Data sources span disciplines and development sectors, inspiring cross-sectoral research and drawing connections across domains
- Offers robust data that also identifies experimental or intervention-based studies to stimulate consideration for what works—and what does not work or what has not yet been tested
- Offers longitudinal evidence, which is more conducive to investigating causality and change in adolescents' experiences over time
- Enables researchers to compare patterns, trends and prevalence across diverse contexts while using standardized instruments and processes

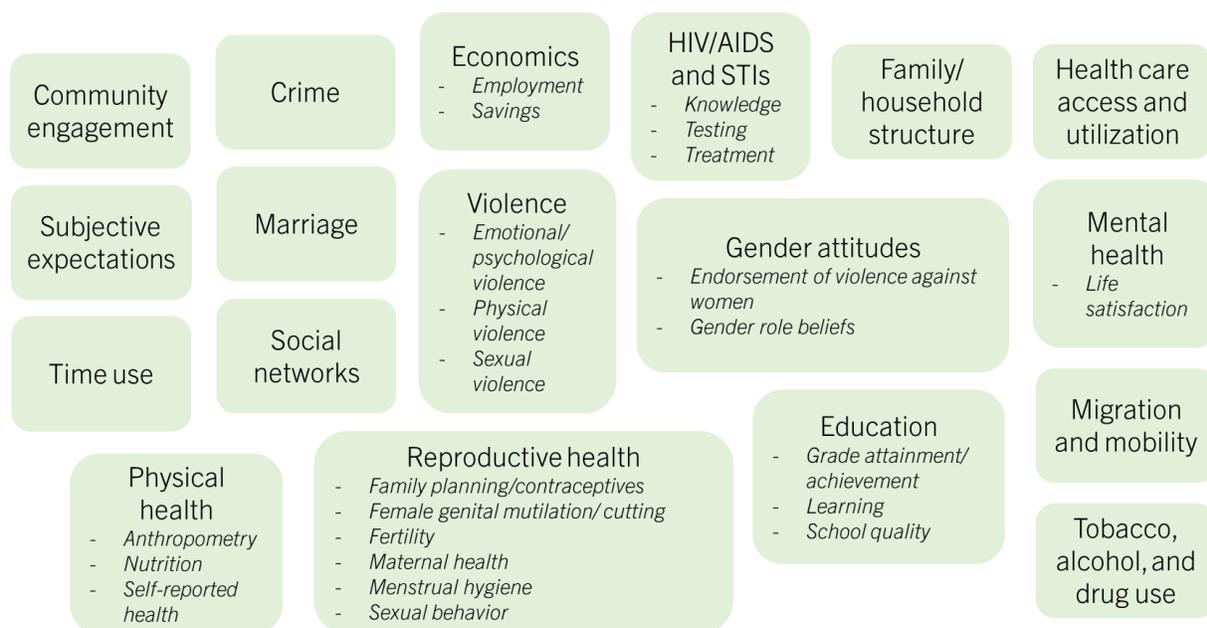
What Data Sources are Included in the ADH?

Sources for data in the ADH were included based on four criteria:

1. the data is self-reported from females and/or males between ages 10 and 24 years;
2. one or more rounds of the data was collected in the year 2000 or later;
3. the data was collected in one or more LMICs;
4. the data is already available or can be requested for download and used by the public.

We identified open-access data sources through three search methods: 1) a scan of existing, open-access databases using keywords related to the inclusion criteria; 2) a broad internet search using keywords related to the inclusion criteria; 3) or direct and indirect outreach to researchers studying adolescents in LMICs. As of January 2022, there are links to 763 data sources on the ADH including single and multi-country cross-sectional and longitudinal studies.

We first formed thematic topics based on assessments of subsets of open-access data sources and their questionnaires, as well as overarching priorities emerging from adolescent-focused research. Sub-topics of specific themes were also drawn from the questionnaires. We identified a total of 18 topics and 23 sub-topics (Figure 1).



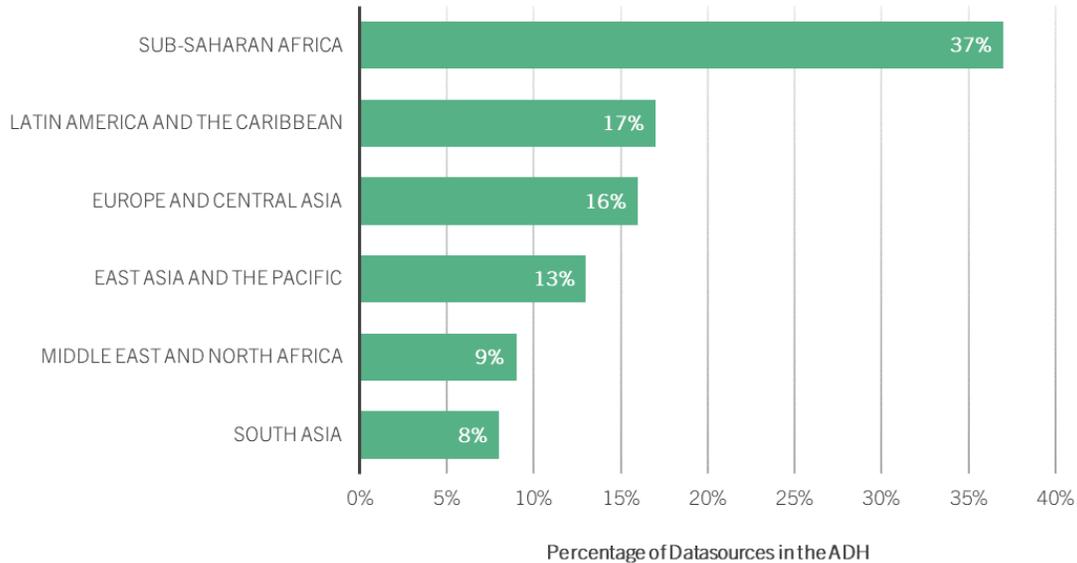
How Can the Included Survey Series and Study Types Add Nuance to Research?

Standardized questionnaires, survey instruments and implementation guidelines make it possible for researchers to compare findings within and/or across countries and subpopulations. About 92% of the data in the ADH are from survey series including the Multiple Indicator Cluster Survey (MICS), Demographic and Health Surveys (DHS), and Violence Against Children Surveys (VACS). Such survey series facilitate international subnational, and sometimes allow for subnational comparison. About 80% of studies are nationally representative; we note that subnational data offers greater nuance, such as identifying the differences among urban and rural youth.

About 10% of data sources are longitudinal studies, which are valuable to researchers seeking to better understand causative factors or the timing of life transitions during adolescence. About 3% are experimental studies, which are valuable in understanding interventions implemented in a context.

What Are the Gaps in Data Availability?

ADH data sources are categorized by study design, type of data, region, country, age, and sex of respondents and key thematic areas.



Though countries from all regions are represented, country-level representation varies widely, leaving gaps in nation-level coverage. For example, within Sub-Saharan Africa's regions (n = 282), Kenya (n = 21), Ghana (n = 17), and South Africa (n = 16) are strongly represented, while Angola, Central African Republic, and Eritrea each have only two sources that meet the inclusion criteria. Countries such as the Syrian Arab Republic (n = 3) and South Sudan (n = 2) similarly had far fewer studies.

The distribution of longitudinal and experimental studies is varied by region. For example, there are no experimental studies in Latin America and the Caribbean or Europe and Central Asia, and only three longitudinal surveys in Latin America and the Caribbean and one in Europe and Central Asia that met our inclusion criteria.



Finally, there are gaps in adolescent age coverage. Specifically, 39% of data are on young adolescents (aged 10-14), 99% on older adolescents (aged 15-19), and 62% on young adults (aged 20-24). While most datasets have overlapping coverage of adolescent age groups, 99% cover some subset of the older adolescent population (ages 15-19), comparatively fewer datasets, 62%, contained information on young adults (ages 20-24), and only 39% included younger adolescents

(ages 10-14). Here we can observe an abundance of data on older adolescents with least coverage of very young adolescents.

Within the themes of education, HIV/AIDS and STIs, and reproductive health, we can also observe overrepresentation and underrepresentation of certain data subtopics. In education, almost all data sources have data on school enrollment and grade attainment; however, only about 25% of sources have data on learning outcomes, collected through survey instruments measuring literacy and/or numeracy. About 40% of data on HIV/AIDS and STIs is related to respondents' knowledge of transmission, but less than 10% of sources share information about respondents' received treatment.

Learn more about the Adolescent Data Hub from the GIRL Center:

- The Adolescent Data Hub: The Largest Catalog of Open Access Data on Adolescents Living in Low- and Middle-Income Countries ([link](#))
- Adolescent Data Hub User Tutorial ([link](#))

Brief prepared by Yeeva Cheng, Cara Kraus-Perrotta