Adolescent Topics

Adolescents and HIV: Definitions and Disaggregation

FULL REPORT

Author: Mannion Daniels
Contents

Executive summary ........................................................................................................................................3
Defining adolescence ................................................................................................................................3
Challenges of disaggregating adolescent data.......................................................................................3
Recommendations .......................................................................................................................................4

Acronyms .........................................................................................................................................................6

1. Introduction .............................................................................................................................................7
   1.1. The issue of disaggregation ..............................................................................................................7
   1.2. Purpose of this discussion paper ......................................................................................................8
   1.3. Scope ..............................................................................................................................................9
   1.4. Methods ..........................................................................................................................................9

2. Current definitions ................................................................................................................................. 11
   2.1. Overlapping definitions .................................................................................................................. 11
   2.2. The 10-19 age range ...................................................................................................................... 13
   2.3. Definitions beyond age .................................................................................................................. 15
   2.4. A useful compromise for data ........................................................................................................ 16
   SUMMARY OF KEY POINTS COVERED IN SECTION 2..................................................................................... 21

3. Current indicators, data sources and disaggregation ............................................................................... 22
   3.1. How indicators are used ................................................................................................................ 22
   3.2. Common disaggregations ............................................................................................................... 25
   SUMMARY OF KEY POINTS COVERED IN SECTION 3..................................................................................... 29

4. Challenges and Solutions ........................................................................................................................ 30
   4.1. Standard grouping strategies ......................................................................................................... 30
   4.2. Disaggregation factors ................................................................................................................... 32
   4.3. Exclusion of 10-14 year olds ........................................................................................................... 34
   4.4. Sampling........................................................................................................................................ 36
   4.5. Confounding.................................................................................................................................. 37
   4.6. Modelled estimates ....................................................................................................................... 38
   4.7. Routine service data/HMIS ............................................................................................................ 39
   4.8. Data accessibility ........................................................................................................................... 40
   4.9. Need for nuanced understanding ................................................................................................... 42
   4.10. Muddling two distinct groups ..................................................................................................... 42
   SUMMARY OF KEY POINTS COVERED IN SECTION 4..................................................................................... 43

5. Conclusions ............................................................................................................................................ 45
   5.1. Defining adolescence for data ........................................................................................................ 45
   5.2. Disaggregating adolescent data - recommendations ................................................................. 45
Executive summary

There is currently a lack of clarity and consistency in how adolescents are defined and how HIV data about adolescents is disaggregated. This report aims to summarise the considerations for how adolescents should be defined and disaggregated when it comes to quantitative data including surveys, surveillance, routine data and models. It also outlines the challenges to disaggregating adolescent HIV data, and the corresponding solutions available. Disaggregation refers to both analysis within smaller age ranges and by other factors such as sex and school attendance.

Clear definitions and disaggregated analysis of quantitative data are useful for designing and implementing impactful programs and policies. Disaggregated data has the potential to help stakeholders:

- identify the sub-groups in most need or at most risk
- identify the sub-groups that are currently under-served
- determine which geographical locations to focus resources on
- understand how well a program is performing
- estimate impact of a program or policy

Defining adolescence

Challenge – broad and overlapping age ranges

The definition of adolescence commonly given for use with quantitative data is 10-19 years old. The more commonly used age range when presenting data about adolescents or young people is 15-24 years old. These two age ranges have the advantages of being well known and of conforming to the common data analysis approach of grouping people into multiples of 5-years of age. Their shortcomings include:

- the characteristics and needs of 10 year olds are very different to those of 19 year olds. Since so much change occurs so rapidly during adolescence, it is questionable whether such wide age ranges are the most useful way to look at data for programming and policy
- Adolescence is more than an age range – it is a collection of life transitions that include changes in biology and in social roles. These transition points occur at different ages for different people and in different contexts.

Recommendations

Other definitions of adolescence that go beyond crude age ranges and attempt to capture the experience of adolescence in terms of life transitions exist, but may not be practical for quantitative data.

A common and useful compromise is to use more refined age ranges which do not overlap with each other and which separate adolescence into three meaningful stages – early adolescence, middle adolescence and young adulthood. There are various options for the age ranges that should correspond to these stages. One option that is consistent with common data analysis practices is to use 5-year age bands such that early adolescents are 10-14 year olds, middle adolescents are 15-19 year olds and young adults are 20-24 year olds. Where possible, further subdivision of these bands has been shown to yield useful results.

Challenges of disaggregating adolescent data

Adolescent HIV data is currently disaggregated to varying degrees depending on the data source. Each data source presents different challenges to achieving more disaggregated data. The following is a summary of some of the main challenges covered in the report:
Survey & surveillance sample sizes

National surveys generally provide the most scope for disaggregated data. HIV-related indicators are analysed by various disaggregation factors, usually including age, education level, wealth and urban/rural residence. Surveillance data should also allow for some disaggregation of adolescent data, but such analysis is not yet consistently done. Adolescents however are often a subset of a broader dataset which includes adults. It is not always possible to sub-divide this subset of a sample and retain enough respondents for disaggregated analysis.

Defaulting to 5 and 10-year age bands

It is common to separate adults into 5 or 10-year age bands. Since change occurs rapidly during adolescence, and at specific life transition points, these broad ranges hide a lot of important information.

Using disaggregation factors appropriate to adults

The factors by which data should be disaggregated also differ for adolescents. For example, it is particularly useful to disaggregate adolescent data by school attendance and by household structure where possible.

Routine service data lost in aggregation

Routine service data is limited by the need to keep data collection simple and practical within the context of service delivery. However, age and sex disaggregated data should be possible since this information is often in the patient records. The challenge is that most management information systems require that patient-level data is aggregated on paper before being sent up the reporting chain. It is rare to find service data specific to adolescents, let alone data disaggregated by sex or into smaller age groups.

Models limited by their underlying data and assumptions

Models provide us with estimates of many indicators that direct data collection currently cannot, including numbers of new infections and deaths due to AIDS. Models however are limited by the data they are based on. If this data cannot be disaggregated, models that use the data generally cannot accurately provide accurate disaggregated estimates.

The methods used by the models also can limit their ability to provide disaggregated data. For example, if a model calculates estimates within 5-year age bands, its ability to produce estimates for other age ranges may be undermined.

The exclusion of 10-14 year olds

10-14 year olds are under-represented in much of the data. This is thought to be partly due to the ethical requirements of collecting data from children of this age. Including younger adolescents can also require the data collection team to use specific interview methods and additional training is often required, which is not always feasible.

Recommendations

Adolescent-specific data collection

Ideally, it would be possible to establish adolescent-specific data sources using adolescent-appropriate data collection methods, samples and analysis, including the use of qualitative and longitudinal study designs where needed, to address the issues outlined above. Adolescent-specific data collection projects in various forms do exist, and can be learned from and replicated.
Make existing data go further

The scarcity of resources available often makes adolescent-specific data collection unrealistic. Fortunately, much can be done with existing data sources. Re-analysis of existing data can be performed using analysis plans based on an understanding of the needs of those that work on adolescent HIV. This would include disaggregation by factors most relevant to adolescents such as school attendance and analysis by year of age rather than only using 5-year bands. Data repositories that make this data more widely available can facilitate this further analysis. This kind of adolescent-specific re-analysis has proven to be powerful in the past.

Methods also exist to obtain individual patient level routine service data even where patient level databases are not yet possible.

Advances in current practices

To increase representation of 10-14 year olds in data, it is useful to support survey teams and ethics committees to ensure that younger adolescents are not systematically excluded from data collection.

Technological advancements, including the establishment of patient level databases for routine service data, can facilitate analysis of adolescent specific and disaggregated service data.
### Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIDS</td>
<td>Acquired Immune Deficiency Syndrome</td>
</tr>
<tr>
<td>ART</td>
<td>Anti-retroviral Therapy</td>
</tr>
<tr>
<td>DHS</td>
<td>Demographic and Health Survey</td>
</tr>
<tr>
<td>DFID</td>
<td>Department for International Development</td>
</tr>
<tr>
<td>DREAMS</td>
<td>Determined, Resilient, Empowered, AIDS-free, Mentored and Safe Women</td>
</tr>
<tr>
<td>EHPSA</td>
<td>Evidence for HIV Prevention in Southern Africa</td>
</tr>
<tr>
<td>ESA</td>
<td>East and Southern Africa</td>
</tr>
<tr>
<td>HAVEG</td>
<td>HIV AIDS Vaccines Ethics Group</td>
</tr>
<tr>
<td>HMIS</td>
<td>Health Management Information System</td>
</tr>
<tr>
<td>HIV</td>
<td>Human Immunodeficiency Virus</td>
</tr>
<tr>
<td>MIS</td>
<td>Malaria Indicator Survey</td>
</tr>
<tr>
<td>PEPFAR</td>
<td>United States President’s Emergency Plan for AIDS Relief</td>
</tr>
<tr>
<td>PHIA</td>
<td>Population-based HIV Impact Assessment</td>
</tr>
<tr>
<td>SDGs</td>
<td>Sustainable Development Goals</td>
</tr>
<tr>
<td>STI</td>
<td>Sexually Transmitted Infection</td>
</tr>
<tr>
<td>THMIS</td>
<td>Tanzania HIV and Malaria Indicator Survey</td>
</tr>
<tr>
<td>TRREE</td>
<td>Training and Resources in Research Ethics Evaluation</td>
</tr>
<tr>
<td>UNAIDS</td>
<td>Joint United Nations Programme on HIV/AIDS</td>
</tr>
<tr>
<td>UNICEF</td>
<td>United Nations Children’s Fund</td>
</tr>
<tr>
<td>UNESCO</td>
<td>United Nations Educational, Scientific and Cultural Organisation</td>
</tr>
<tr>
<td>UNFPA</td>
<td>United Nations Population Fund</td>
</tr>
<tr>
<td>VMMC</td>
<td>Voluntary Male Medical Circumcision</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
</tbody>
</table>
1. Introduction

Recent years have seen a global recognition of the importance of tackling HIV among adolescents\(^1\). Adolescence is a period of transition from childhood dependence to adult independence. It is a time of rapid change in terms of biology, psychology and social roles. Adolescents are often in a particularly vulnerable position in society due to this transition. During this period, many adolescents begin to experiment with their sexuality and other behaviours including drug use\(^2\). They may also experience various forms of exploitation, or suffer from inadequate access to information and services due to their social status, economic circumstances or social stigmas.

HIV prevalence data indicate that adolescents are a group left behind, particularly in sub-Saharan Africa, by HIV prevention programming. An estimated 670,000 15-24 year olds were newly infected with HIV in 2015, including 250,000 15-19 year olds. More of half of the world’s newly infected 15-19 year olds were in Eastern and Southern Africa\(^3\), AIDS-related deaths among adolescents grew from 71,000 in 2005 to 110,000 in 2012 globally despite a decline among other age groups in the same period\(^4\).

However, adolescence is not only a period of vulnerability and risk, and it is more than a transition into adulthood. It is an important life stage in itself. The changes that can result in vulnerability also make it a time of discovery and opportunity. It represents a particularly important opportunity for public health interventions, as a formative period when health trajectories can be shaped towards positive health outcomes\(^5\).

1.1. The issue of disaggregation

It is known that the broad trends in HIV among adolescents mask variations and sub-groups amongst adolescents at particular risk or with specific needs. For example, adolescent girls in Eastern and Southern Africa are at a higher risk than boys: 59% of 10-19 year olds with HIV were female in 2012. In South Africa, more than 80% of new adolescent HIV infections are among girls. Other sub-groups among adolescents have differing vulnerability, such as the relatively poor, those in rural areas versus urban areas, those in- or out-of-school, or married and unmarried adolescents\(^6\).

Our ability to disaggregate data on adolescent HIV, both by age and by other factors, is limited and so prevents a better understanding of sub-groups among adolescents in ESA. We also lack clear information on the proportions of adolescents with HIV that acquired the virus perinatally or during infancy via breastfeeding and those that acquired it during adolescence via unprotected sex or non-sterile needles\(^7\). This hampers our understanding of who is at particular risk, specific needs of sub-groups of adolescents and what groups, contexts or behaviours might most beneficially be targets for programmatic interventions.

---


\(^3\) https://data.unicef.org/topic/hivaids/adolescents-young-people/#, 2016


\(^7\) Ibid, S144-53
In recognition of this, global HIV prevention policies from UNAIDS, WHO and other agencies have urged stakeholders to improve data systems to allow for more disaggregation of the adolescent population:

*It has become clear that data on the epidemic and service delivery must be disaggregated by age and sex in order to better define the impact of HIV on adolescents and other age groups, and to better inform policy and advocacy at the national level. Due to the rapid physical, psychological, emotional and cognitive changes of adolescence, it is essential that data be stratified in ways that facilitate the design and implementation of interventions and services that are appropriate for adolescents aged 10–14 and 15–19 years in particular.*

There is currently a lack of clarity and consistency around how surveillance systems, surveys and monitoring and evaluation efforts should be modified to cater to this underserved group. A further complicating factor is that the group referred to as ‘adolescent’, a term used throughout the HIV prevention field, is defined in different ways by different stakeholders. Furthermore, people going through adolescence are also often lumped in with other broad categories that may not reflect their needs, such as ‘children’, ‘adults’ or ‘youth’, depending on the focus of the organisation.

1.2. Purpose of this discussion paper

In order to synthesise the evidence generated by the large body of research taking place in the East and Southern Africa Region, the Evidence for HIV Prevention in Southern Africa (EHPSA) Programme contracted MannionDaniels Limited to produce two discussion papers aimed at contributing to, and broadening, the regional debate on issues affecting the design of research, policy and programming for HIV prevention for adolescents. This paper tackles the issue of disaggregation of data on adolescents.

**The problem:** There is currently a lack of clarity and consistency around how adolescence should be defined, and limited disaggregation of adolescents in quantitative data

**The key questions:**

- What are the issues in defining adolescence?
- What are the challenges in disaggregating adolescent data such that programmers and policy-makers can better identify and serve key groups?

**The first section** of this paper summarises the definitions of adolescence currently in use and how they relate to quantitative data. This section will offer some ways forward on a common approach to defining adolescence in Eastern and Southern Africa (ESA).

**The second section** of this paper summarises the main HIV indicators and their data sources and what they can tell us about adolescents in ESA. This includes the disaggregation of the broad adolescent group that is currently available.

**The third section** outlines the challenges that prevent more detailed breakdowns of adolescent HIV data from being used, and the solutions that can help us overcome those challenges.

The objectives of each section are to cover the following:

**Current definitions of adolescence for quantitative data:**

- the various ways that adolescents and young people are categorised by different agencies and programs
- the pros and cons of the ways adolescents are currently defined
- a useful approach to defining adolescents for quantitative data

---


9 [www.manniondaniels.com](http://www.manniondaniels.com)
Current data:

- The uses of different types of data for understanding the adolescent HIV epidemic and service provision
- The ways that data are collected and processed and their implications for adolescent disaggregation, including:
  - Surveys & surveillance
  - SPECTRUM and other models
  - Research studies
  - Routine service data
- The disaggregation that current data provides us – both in terms of age ranges and sub-groups within adolescents, such as being married or unmarried

Disaggregation challenges and solutions:

- The challenges and limitations that limit the disaggregation of adolescent data presented by:
  - each data source (surveys, surveillance, routine service data, models & research)
  - accessibility and use of the data
  - the habits and traditions of analysts
- The solutions being employed to overcome each of these challenges

1.3. Scope

There is a large body of literature on measuring HIV amongst adolescents in Eastern and Southern Africa. This discussion paper therefore seeks to provide a palatable summary of the issues around defining adolescence and disaggregating adolescent data drawing on the literature, an analysis of statistical models being used to inform decision makers and the perspectives of experts working across Eastern and Southern Africa.

This paper covers disaggregation of quantitative indicators about people as opposed to supply-side indicators such as health facility indicators or other types of data.

1.4. Methods

This discussion paper is based on a triangulation of information from three sources: a review of published literature, discussions with experts on the topics covered and case studies focused on three selected countries - Malawi, South Africa and Tanzania.

Consultations with experts

The experts were consulted both via telephone and through in-person meetings. The authors visited South Africa and Malawi to allow for in-person meetings in those countries. Experts were identified via the networks of the authors of this paper and through recommendations of the experts themselves. This has been included as a supporting annex.

Prior to the discussions, a list of topics to be covered was prepared. The topic guide is included as a supporting appendix. While these lists of topics were used to ensure all questions were covered, the discussions were kept organic and free-flowing, to allow for emerging themes to be identified.

Literature review

The full list of documents reviewed are listed in the bibliography. The types of documents reviewed were as follows:
• Global and regional literature
• UNAIDS, WHO, PEPFAR and GFATM M&E toolkits and surveillance guidelines
• Major agency and national government HIV progress reports
• Synthesis of different data sources to produce disaggregated data analysis on adolescents for HIV
• HIV modelling documentation (e.g. Spectrum methodology)
• National AIDS Progress Reports from countries in the region
• Journal articles on adolescence and adolescent HIV specific to the sub-topics discussed in this paper (for example, articles on the age of onset of puberty)

From the three focal countries (Malawi, South Africa and Tanzania):

• Routine data reports from national Health Management Information Systems
• Documentation on electronic HMIS systems such as DHIS
• National survey and surveillance reports
• Research on adolescence in ESA recommended or referred to by experts
• Youth policies from governments and donors

Country case studies

Three countries were selected for more in-depth investigation – Tanzania, Malawi and South Africa.

The selection of countries was based on their diverse HIV prevalence levels and trends and their ability to provide a range of insights into the topics. The characteristics used in the selection process are detailed in the supporting appendix. Key factors about each of the selected countries that influenced country selection are summarised below:

• Malawi has been able to integrate patient-level databases to its health management information system. It has a recent Demographic and Health Survey that included HIV testing\(^{10}\) as well as recent key population surveillance on MSM and sex workers\(^{11}\) among other groups
• Tanzania has multiple national surveys with HIV testing and other relevant data on adolescents\(^{12}\), which were synthesised into an adolescent specific analysis widely used in the country for program planning\(^{13}\)
• South Africa has a large body of research into HIV as well as a series of in-depth national surveys which differ to the more common Demographic and Health Surveys and AIDS Indicator Surveys\(^{14}\). Probably the most extensive in the region. These include HIV data on under 15s, and laboratory test-based incidence estimates. South Africa’s ANC surveillance is also well documented\(^{15}\)

Limitations

This paper focuses on definitions of adolescence and the challenges of disaggregating quantitative data. It does not attempt to provide a comprehensive summary of which adolescents in the region should be targeted

---


with programming, or provide a nuanced understanding of the adolescent experience and its associated vulnerabilities.

The experts we consulted for this paper were identified via our own networks, from recommendations of other experts or based on their involvement in relevant projects or literature they had published. The opinions of the experts we spoke to may not be representative of all those working in this field.

Similarly, the breadth of issues covered in this paper, combined with the huge amount of documentation and literature available on each of these subjects covered in this paper, means that not all literature and research relating to these topics could be reviewed.

2. Current definitions

2.1. Overlapping definitions

Adolescents are defined as 10-19 year olds by the WHO and UNICEF. This definition is shared by many governments in the region and has been adopted by various NGOs and research institutions.

UNAIDS often refers to ‘adolescents and young people’ defined as 15-24 year olds, and this is the age range used for the UNGASS indicator relating to young people. Data relating to HIV commonly reflect the 15-24 year old population. Despite AIDS Indicator Surveys being described as surveys of ‘adolescents and adults’, their samples do not include 10-14 year olds. Various UN agencies and USAID are focused on ‘youth’ or ‘adolescents and youth’, including for HIV programming, which they define as 15-24 year olds or ‘young people’ defined as 10-24 year olds.

Most ESA countries have set the age of majority at 18, which means that under 18s are legally classified as children. UNICEF similarly categorises children this way. In routine data, service indicators are often divided between 0-14 year olds and 15-year olds and up, effectively including 15-year olds as adults.

10-24 year olds are therefore included in a variety of groupings, but these groupings may be based more on the needs of agencies and data analysts than a clear rationale based on the actual needs and risks faced by different sub-groups within the age range.

Depending on the agency or governmental body, a 15-year-old in Eastern and Southern Africa will be categorised as a child, an adult, an adolescent, a young person and a youth. Figure 1 below summarises the definitions covering the major age ranges.

---

Being categorised into various broad groups has a number of implications. Adolescents are likely to be the target of a range of services, policies and programmes designed to include anyone from infants to the elderly.

“Too old for paediatrics and often deemed too young for adult health services, many adolescents fall through the cracks at a time they most need our attention.” - Anthony Lake

Many progressive service providers are modifying standard practices in recognition of this. In Malawi, for example, the Lighthouse Trust and Dignitas among others have designed teen clubs specific to the needs of adolescents and separate to paediatric and adult services. This has successfully improved adherence to HIV treatment among adolescents. 

---


Categorisation should add value, not confusion

Overlapping definitions are necessary. Younger adolescents are also children, and should benefit from the same protections. Older adolescents are also adults and must bear adult responsibilities. However, since adolescents are already subsumed into these other broad categories, it is important that the definitions of adolescence that we use to distinguish them from other children and adults do represent their specific needs. Where possible, categorisations layered on top of ‘children’ and ‘adults’ should avoid being overly broad, and should not overlap with each other. The issue is not that the 10-19 and 15-24 age ranges overlap with childhood and adulthood, but that they themselves are broad and overlap with each other.

Categories that span the child-adult cut-off point of 18 are also problematic. 18-19 year olds are adults that are categorised as adolescents and grouped in with children, while 15-17 year olds are children that are categorised as youth and grouped in with adults.

2.2. The 10-19 age range

The standard definition of adolescents as 10-19 year olds used by WHO and UN agencies is the period of life that encompasses the majority of the main changes and transition points that encompass the adolescent experience. It is useful because it has been widely taken up by other organisations and projects as their official definition of adolescence. Consistent terminology facilitates communication. This definition also conforms to the norm among data analysts of using 5-year bands to group people by age. A standard age range allows for comparability in data, policies and program designs.

However, as with any strict definition of a diverse group, it also has a number of important limitations in addition to the issue of overlapping with other age categories described above.

**Limitation 1:** the wide age range groups together people at very different life stages

Due to the rapid biological, social and psychological changes that occur from the ages of 10-19, the attributes and needs of 10 year olds are dramatically different to those of 19 year olds. The factors that will determine adolescent policy and programmes – knowledge, influenced by cognitive ability, risk behaviours influenced by social roles and biology, and even the appropriate approach to service provision - cannot be known for the 10-19-year-old age group as a whole. The prevalence of HIV among early adolescents, middle adolescents and young adults is also very different.

*Figure 2 - HIV Prevalence by Sex and Age, South Africa 2012* ^25

---


**Limitation 2:** the 10-19 definition is rarely used in presentation of data

While this definition has been adopted by a variety of international agencies and national governments, it is sporadically used in practice – largely negating the benefit of having a standard definition.

HIV-related data for 10-19 year olds as a group were very rarely found among the online data repositories, national HIV survey reports, surveillance reports both for ANC and key populations, country MIS databases, M&E guidelines, models, national AIDS reports and NGO program strategies purporting to be about adolescents that were reviewed for this report.

Age groups commonly reported on in the literature reviewed were 0-14 year olds, 15-19 year olds, 15-24 year olds, 15-49 year olds, 15 and over, or 18 and over. Early adolescence – defined in terms of age ranges similar to 10-14 year olds, are generally grouped in with smaller children or are neglected entirely (see section 4.3 for more on this). This may be one reason that data on 10-19 year olds are rarely presented.

**Limitation 3:** the start and end points of adolescence are not only determined by age

Adolescence is generally considered to start with the biological changes of puberty. Puberty is a complex interconnected set of changes within the body, which tend to occur at different points between the ages of 10 and 15-years old.

The onset of puberty varies between individuals and contexts. In South Africa for example, the median age of menarche among black urban girls was found to have decreased from 14.9 in 1956 to 12.4 years old in 1990. In Northern Malawi, the median age of menarche was found to be 15. Thus the age at which the biological processes of adolescence begin is not only a function of age, but of the conditions in which the person is growing up.

Changing social roles, which depend on the context in which the adolescent lives, are as important to understanding adolescence and its associated vulnerability as biological changes. Adolescence is thought to end with physical maturation combined with the transition to adult social roles. Thus while finding a specific year of age that corresponds to the start of adolescence is difficult, determining an end-point is even more of a challenge.

The transitions that are thought to mark the end of adolescence vary between cultures. Even if we could agree adolescence should be marked by some combination of transition points such as physical maturation, employment, marriage, or childrearing, these also vary by context, and to make matters more complex, they appear to be moving targets.

During our consultations with experts, we found mixed views on whether the variation by context undermines the use of any standard age range to define adolescence. The box below provides an illustration of the fact that while the age of transition points can vary by context, it is not always clear whether this is sufficient to...
undermine the use of standard age ranges. To properly determine whether a standard age range is so out of tune with the actual processes and transition points of adolescence would require a comprehensive analysis of the variation in the ages at which individuals within each society experience the physical, psychological and social transition points that define the experience of adolescence. The findings from a range of societies would then need to be compared to determine whether a regional standard is achievable.

Transition points in context – examples from Tanzania, Malawi and Kenya

The median ages for sexual debut, marriage and age at first birth are given for three countries in Table 1 by urban/rural residence. The ages illustrate that a number of changes occur in adolescent’s lives within the 15-19-year-old age range.

While there is some variation between contexts in the median ages given in the table, it is debatable whether it is sufficient to undermine the use of standard age ranges. To properly answer this question requires a comprehensive analysis of the available data including consideration of the distribution around the median.

Table 1 - Median ages at life transitions as reported by 25-49 year olds

<table>
<thead>
<tr>
<th></th>
<th>Kenya</th>
<th>Tanzania</th>
<th>Malawi</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Urban</td>
<td>Rural</td>
<td>Urban</td>
</tr>
<tr>
<td>Age at sexual debut</td>
<td>18.8</td>
<td>17.3</td>
<td>17.8</td>
</tr>
<tr>
<td>Age at first marriage</td>
<td>21.5</td>
<td>19.5</td>
<td>19.8</td>
</tr>
<tr>
<td>Age at first birth</td>
<td>21.6</td>
<td>19.4</td>
<td>19.9</td>
</tr>
</tbody>
</table>

Sources: Malawi DHS Data 2010, Kenya DHS Data 2014, Tanzania DHS Data 2010

2.3. Definitions beyond age

The use of the 10-19-year-old definition of adolescence is not the result of a lack of understanding of adolescence. WHO, UNICEF, USAID and other agencies that use and promulgate age range definitions of adolescence and youth also recognise their shortcomings. They supplement these definitions with nuanced ways of understanding adolescence, taking into account the varying biological changes, the social transitions and the cultural factors that define people’s actual experience of adolescence.

Standard age ranges continue to be used because the practical benefits for the purposes of data analysis and presentation, and for clarity in communications and consistent terminology, outweigh the limitations. It is worth considering whether an alternative approach could be developed. DFID’s youth policy provides an example of a progressive approach that goes beyond age in their definition of youth:

“DFID takes a ‘lifecycle approach’ to youth, defining it as ‘the period of time during which a young person goes through a formative transition into adulthood’... Taking a lifecycle approach is valuable because it allows us to go beyond defining young people by age to identifying what it means to become an adult by other defining

---

moments - such as getting a job, marrying or becoming a parent. It also enables us to take into consideration cultural and country contexts, and to examine the broader social, economic, political, physical, emotional and cultural changes that the transition to adulthood involves.\textsuperscript{36}

Using such an approach to defining adolescence resolves some of the limitations of standard age ranges. This approach adjusts for the variation of the timing of adolescent transition points by context. It also ensures that people are grouped together by meaningful characteristics.

To be usable as a definition of adolescence for use in quantitative data, a clear set of criteria would be needed to categorise people. The starting point of female adolescence, for example, could conceivably be defined by menarche rather than a specific age. However, given the complex interplay of factors that demarcate adolescence from adulthood, it is unclear whether it would be possible to create a strict definition of the endpoint of adolescence. Furthermore, there would be the risk that comparable data could not be presented on adolescent HIV from different contexts using this context-specific and nuanced definition.

A life cycle approach is useful for developing a qualitative understanding of the true nature of adolescence. It can also help us to interpret quantitative data by understanding how factors such as education and marriage fit into adolescence as a complex period of transition. However, it cannot replace age range definitions used to categorise people as adolescents for quantitative data.

2.4. A useful compromise for data

The limitations of both the age-based and life stage-based definitions of adolescence illustrate that there is a tension between the complex and context-specific nature of adolescence and the need for a simple and standard definition for the purposes of quantitative data.

It is useful to recognise that adolescence is not something that can easily be categorised with a strict set of criteria. Adolescence is about age, biology, sexual activity, marital status, school, relationships and more. To understand the multiple facets of adolescence, we will need multiple variables, of which age is just one. This is why disaggregation of adolescent data, explored later in this paper, is particularly important.

There is no current solution which will both fully capture adolescence and be usable for categorising people in quantitative data. Instead, we can aim to address the other limitations of the 10-19 definition described above – that it is too broad, that it overlaps with other age range categories and that in practice it is rarely used to present data.

A common approach to achieving more nuance while retaining the practical benefits of standard age ranges is to focus on three stages – early adolescence, middle adolescence and young adulthood – and to use these stages in place of the broad and overlapping groups ‘adolescence’ and ‘youth’. Some of the key characteristics of these stages from the experts we talked to for this paper\textsuperscript{37} and from literature on adolescents are summarised below. These stages vary by context and by individual, so generalisations such as those below will not apply in all cases.

**Early adolescence** – Members of this group are entering puberty, and experience the associated rapid changes in biology and social roles and expectations. Most in this group have a developing interest in sex but not reached sexual debut, with some important exceptions. Relationships with peers become increasingly important\textsuperscript{38}. Programs working with uninfected early adolescents aim to positively influence their trajectories.
and prevent future risky behaviour. This group is more likely to be in school. Many of those living with HIV in this group are thought to have been infected during infancy\footnote{39}.  

**Middle adolescence** – The physical changes are less dramatic for this group, though changing social relationships and roles become increasingly important\footnote{40}. Members of this group are starting to have sex. This causes incidence to increase, though it tends to remain relatively low. Programs aim to reduce the risks associated with vulnerability and risky behaviour. This group is divided into those still in school and those out of school, which can determine both their vulnerability and how programs can reach them.

**Young adulthood** (sometimes referred to as late adolescence) – This group is largely sexually active and out of school. They are becoming more emotionally stable and achieving increased independence. They are more likely to be married, have children, and work. These factors affect their vulnerability and risk of infection. In most contexts, HIV incidence among this group is higher than for early and middle adolescents.

---

**Dividing by stage of adolescence in practice - DREAMS**

The Dreams program in Tanzania has been designed with the use of three age ranges – 10-14 year olds, 15-19 year olds and 20-24 year olds. This breakdown has allowed for strategies that reflect the needs of these distinct groups.

For 10-14 year olds, the program reaches its target population via schools, with the aim of influencing their future ‘trajectory’. It is thought that this is a time of life where the knowledge, attitudes and practices that will determine risk of HIV transmission during middle adolescence can be pre-emptively improved.

The strategy for 15-19 year olds and 20-24 year olds meanwhile are focused on mitigating current risk. These adolescents may not be in school, which can adversely affect their risk of HIV transmission. These adolescents also may be engaged in risky behaviours such as sex work. Thus the way that the program seeks to reach those at risk in these age groups is very different to that for 10-14 year olds.

---


\footnote{40} Ibid.
There is no standard set of age ranges corresponding to these three stages. Different organisations and experts use different age ranges. The table below present three age range options with a rationale for each. The first is the most common.

**Table 2 - options for adolescent sub-division of age ranges**

<table>
<thead>
<tr>
<th>Early adolescent</th>
<th>Middle adolescent</th>
<th>Young adulthood</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Option 1 – standard 5-year age bands</td>
<td>10-14</td>
<td>15-19</td>
<td>20-24</td>
</tr>
<tr>
<td>Option 2 – life transition points</td>
<td>12 - 14 Starting age for puberty</td>
<td>15-17 Age of increasing sexual debut, lower school attendance</td>
<td>18-24 Legal age of majority, increasing marriage and independence</td>
</tr>
</tbody>
</table>

---

\(^{41}\) Ibid.


### Option 3 - Country specific transition points

<table>
<thead>
<tr>
<th>Early adolescent</th>
<th>Middle adolescent</th>
<th>Young adulthood</th>
<th><strong>Rationale</strong></th>
</tr>
</thead>
</table>
| Context and sex specific start of puberty | Context and sex specific age with increased school dropouts | Context and sex specific | Adolescence, and the changing risks that we are interested in, depend on age in conjunction with context and sex.  
Age ranges that reflect transition points specific to each context and sex are one way to achieve ranges that reflect people's actual life stages and transitions, and therefore may be more useful in some cases that a regional standard.  
The huge downside to this is that data between countries would be less comparable, and terminology would mean different things in different places, creating confusion. |

---


Figure 3 reflects option 2. Compare this graphic to Figure 1 to see how the use of mutually exclusive age ranges that match the child-adult cut-off could reduce the risk of people being lost in definition overlap.

**Figure 3 - ages and proposed definitions**

<table>
<thead>
<tr>
<th></th>
<th>Child</th>
<th>Early adolescent</th>
<th>Middle adolescent</th>
<th>Young adult</th>
<th>Adult</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
SUMMARY OF KEY POINTS COVERED IN SECTION 2

- There is considerable overlap between the definitions of a child, an adolescent, youth/young person and an adult.
- The two principal definitions employed by stakeholders, the 10-19 and 15-24 age ranges, are broad and overlap with each other, further complicating analysis and comparability.
- The most widely used definition, the 10-19 age range, has a several limitations to note:
  - Limitation 1: the wide age range groups together people at very different life stages
  - Limitation 2: the 10-19 definition is rarely used in presentation of data
  - Limitation 3: the start and end points of adolescence are not only determined by age but influenced by qualitative contextual factors which make quantitative analysis tricky.
- A life cycle approach is useful for developing a qualitative understanding of adolescence, however it cannot replace age range definitions used to categorise people as adolescents for purposes of quantitative data analysis.
- Further sub-dividing the 10-19 age bracket into early adolescence, middle adolescence and young adulthood presents an important solution to the above. However different approaches to this solution have resulted in slightly different age groupings being employed by stakeholders.
3. Current indicators, data sources and disaggregation

In this section, we will look at indicators and data sources, and what disaggregated information they can provide on adolescents. The first part provides an overview of how the approach to disaggregation relates to how that data will be used. The second part presents an overview of data sources and their strengths and weaknesses relating to adolescent disaggregation. The third part presents the common ways in which adolescent data are disaggregated.

3.1. How indicators are used

Indicators tell us about the direction and scale of the epidemic. The way that an indicator should be disaggregated depends on how the data is to be used.

Below we have outlined some of the considerations of how different uses of indicators affect how they should be disaggregated. Useful information on indicators can be found in the Indicator Registry.

Prevalence and incidence data

Disaggregated prevalence data help us understand who has a current need for treatment or which groups may drive further infection.

Prevalence data is less useful for understanding current risk of infection among specific age groups or sub groups, as other factors such as past exposure and survival also affect prevalence. This is particularly true for older groups, but can also affect prevalence among adolescents. For example, high prevalence among adolescents may be the result of people infected at infancy aging into adolescence rather than due to infection during adolescence. Age-disaggregated incidence and new infection numbers are therefore important to understanding the current risk of infection faced by different age groups.

Multiple measurements of prevalence over time can be used to estimate incidence.

Prevalence versus incidence, an example from South Africa

Table 3 provides an example from South Africa, where prevalence and incidence were estimated in 2012. While prevalence is much lower among 15-24 year olds compared to the whole 15-49-year-old age group, incidence was similar indicating a similar level of risk of current infection. The number of new infections show that 15-24 year olds represent about a third of new infections among 15-49 year olds.

---

Table 3 - prevalence and incidence for 15-24 year olds and 15-49 year olds from the South African National HIV Prevalence, Incidence and Behaviour Survey 2012

<table>
<thead>
<tr>
<th>Age Group</th>
<th>HIV Prevalence</th>
<th>HIV Incidence</th>
<th>Number of New Infections</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-24</td>
<td>7.1%</td>
<td>1.49%</td>
<td>139,000</td>
</tr>
<tr>
<td>15-49</td>
<td>18.8%</td>
<td>1.72%</td>
<td>396,000</td>
</tr>
</tbody>
</table>

Causal inferences from disaggregated prevalence data are also challenging, as the point of infection is unknown. For example, if adolescents out of school have higher prevalence than those in school, it may not be clear whether this is because school is actually protective, or because HIV status is linked with factors that affect both HIV risk and school attendance simultaneously. Prevalence data disaggregated by school attendance may be useful however for understanding whether or not going to schools is a useful strategy for reaching HIV positive adolescents, or whether risk at the time of transition from school needs specific attention in programming or further research.

Data on knowledge and behaviours

Data on knowledge and behaviours serve as both indicators and disaggregation factors. Disaggregated data about knowledge can tell us who is in need of further education and who has been left behind in education efforts.

Disaggregated data on behaviours such as condom use and multiple sexual partners inform us of who is likely to be at heightened risk of infection. It also provides a better understanding of the dynamics of the epidemic by giving some indication of how HIV is transmitted between individuals or groups. As well as self-reported behaviours, risk behaviour can also be understood in some cases using proxy factors such as STIs or pregnancies. Disaggregating HIV prevalence data by knowledge or behaviour can provide some indication of whether knowledge and specific behaviours affect the risk of infection, though such analyses can be affected by confounding, as explored in this report’s challenge box on the subject.

Service data

Routine data on service provision is used to inform management decisions at all levels, ranging from procurement to performance monitoring. Routine data are also important for estimating the coverage of services, factors affecting uptake of preventive services such as MMC, and can also indicate the extent to which people are starting and adhering to treatment or PreP. Routine service data is also important for modelling the epidemic with tools such as SPECTRUM, because treatment affects the dynamics of the epidemic.

Since this data has so many uses at all levels of health systems, more detailed disaggregation would be powerful. Understanding how specific groups are receiving services would allow the health system to better serve those groups. However, there are important challenges to achieving this such that currently most systems cannot provide figures specific to adolescents, let alone disaggregated data about them. These challenges are explored further in section 4.7.

Data sources

Each data source can provide information for different indicators and allows for disaggregation and grouping in different ways. Table 4 provides an outline of the types of indicators each data source commonly provides, how the data is disaggregated and some of the main limitations of each data source in terms of disaggregation.

Common indicators are the indicators that all or most reports from these data sources present. Common disaggregation factors are the factors by which these indicators are disaggregated in all or most of the reports. Actual lists of disaggregation factors from recent reports from ESA are presented in the following section. Many of the limitations listed here are explored in detail in the challenges section.

**Table 4 - Data source summary**

<table>
<thead>
<tr>
<th>Data source</th>
<th>Common indicators</th>
<th>Factors by which adolescents are commonly disaggregated</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>National surveys</td>
<td>HIV prevalence, Knowledge, Behaviours</td>
<td>Age groups, sex, Education, occupation, urban/rural residence, region, sexual behaviour, wealth, marital status, sometimes by sex, for 15-24 year olds.</td>
<td>Resource intensive, cannot be done often Sample sizes need to be large enough for sub-group analysis</td>
</tr>
<tr>
<td>ANC Surveillance</td>
<td>HIV prevalence and trends</td>
<td>5-year age groups, pregnant adolescents only Region by age group.</td>
<td>Includes only pregnant women receiving ANC Samples, where applicable, need to be large enough for sub-group analysis</td>
</tr>
<tr>
<td>Key population surveillance</td>
<td>HIV prevalence, knowledge and behaviours for key populations</td>
<td>5-year age groups</td>
<td>Often irregular Limited geographic coverage Sensitivities around collecting data from under 18 year olds Inconsistent approaches to analysis and reporting</td>
</tr>
<tr>
<td>Routine data on service provision</td>
<td>ART uptake and adherence HIV testing VMMC</td>
<td>Child and adult</td>
<td>Data collection must fit into a busy clinic workflow Multiple uses of the data, such as stock management, mean only essential data should be collected Reliable cohort data often not available for ART Specific data on adolescents often nor reported routinely Denominator population for HIV testing and HAV often unclear</td>
</tr>
<tr>
<td>Studies and research projects</td>
<td>Any</td>
<td>Specific to research requirements</td>
<td>In general, studies are not standardised, making consistency difficult to achieve</td>
</tr>
<tr>
<td>Spectrum estimates</td>
<td>New infections, Deaths due to AIDS Need for treatment</td>
<td>15-24 year olds 5-year age bands possible but not presented as standard</td>
<td>Based on data using 5-year age bands A range of input data, assumptions and standards are used which are difficult to determine for sub-groups Partially based on assumptions, each with varying implications for accuracy</td>
</tr>
</tbody>
</table>
3.2. Common disaggregations

The figures below provide examples of the disaggregation of adolescent or youth data presented in national HIV progress reports, survey reports, surveillance reports and routine data reports from Malawi, Tanzania and South Africa. National survey reports tend to present the most disaggregation, and have been more standardised across countries than other data source reports. Key population surveillance presents inconsistent disaggregation and often excludes under 18s. Routine service data is the least disaggregated – generally data specific to adolescents is not provided.

National surveys present standard and detailed disaggregation tables for 15-24 year olds, partly because of the standardised nature of DHS and AIS which represent most of the national surveys in the region. The factors by which the data is disaggregated are similar to those used for adults, except that narrower age bands are also presented.

**Table 5 - HIV prevalence among young people by background characteristics**

<table>
<thead>
<tr>
<th>Background characteristic</th>
<th>Women</th>
<th></th>
<th>Men</th>
<th></th>
<th>Total</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percentage HIV positive</td>
<td>Number</td>
<td>Percentage HIV positive</td>
<td>Number</td>
<td>Percentage HIV positive</td>
<td>Number</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15-19</td>
<td>4.2</td>
<td>1,545</td>
<td>1.3</td>
<td>1,703</td>
<td>2.7</td>
<td>3,248</td>
</tr>
<tr>
<td>15-17</td>
<td>3.4</td>
<td>1,036</td>
<td>2.0</td>
<td>1,102</td>
<td>2.7</td>
<td>2,138</td>
</tr>
<tr>
<td>18-19</td>
<td>5.7</td>
<td>510</td>
<td>0.1</td>
<td>601</td>
<td>2.7</td>
<td>1,110</td>
</tr>
<tr>
<td>20-24</td>
<td>6.4</td>
<td>1,401</td>
<td>2.8</td>
<td>1,176</td>
<td>4.7</td>
<td>2,577</td>
</tr>
<tr>
<td>20-22</td>
<td>5.6</td>
<td>824</td>
<td>1.8</td>
<td>752</td>
<td>3.8</td>
<td>1,576</td>
</tr>
<tr>
<td>23-24</td>
<td>7.5</td>
<td>577</td>
<td>4.6</td>
<td>424</td>
<td>6.2</td>
<td>1,001</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never married</td>
<td>3.8</td>
<td>1,312</td>
<td>1.5</td>
<td>2,361</td>
<td>2.3</td>
<td>3,673</td>
</tr>
<tr>
<td>Ever had sex</td>
<td>7.6</td>
<td>400</td>
<td>1.4</td>
<td>1,445</td>
<td>2.8</td>
<td>1,846</td>
</tr>
<tr>
<td>Married/living together</td>
<td>2.1</td>
<td>912</td>
<td>1.6</td>
<td>916</td>
<td>1.9</td>
<td>1,827</td>
</tr>
<tr>
<td>Divorced/separated/widowed</td>
<td>5.5</td>
<td>1,441</td>
<td>3.5</td>
<td>480</td>
<td>5.0</td>
<td>1,921</td>
</tr>
<tr>
<td>Currently pregnant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregnant</td>
<td>3.8</td>
<td>300</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>Not pregnant or not sure</td>
<td>5.4</td>
<td>2,645</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>Residence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>11.2</td>
<td>589</td>
<td>2.9</td>
<td>666</td>
<td>6.8</td>
<td>1,255</td>
</tr>
<tr>
<td>Rural</td>
<td>3.7</td>
<td>2,357</td>
<td>1.6</td>
<td>2,213</td>
<td>2.7</td>
<td>4,570</td>
</tr>
<tr>
<td>Region</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northern</td>
<td>2.9</td>
<td>368</td>
<td>1.1</td>
<td>313</td>
<td>2.1</td>
<td>680</td>
</tr>
<tr>
<td>Central</td>
<td>3.5</td>
<td>1,245</td>
<td>1.7</td>
<td>1,278</td>
<td>2.6</td>
<td>2,523</td>
</tr>
<tr>
<td>Southern</td>
<td>7.5</td>
<td>1,333</td>
<td>2.4</td>
<td>1,289</td>
<td>5.0</td>
<td>2,622</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No education</td>
<td>9.1</td>
<td>185</td>
<td>1.2</td>
<td>70</td>
<td>6.9</td>
<td>255</td>
</tr>
<tr>
<td>Primary</td>
<td>3.9</td>
<td>1,994</td>
<td>2.0</td>
<td>1,904</td>
<td>3.0</td>
<td>3,898</td>
</tr>
<tr>
<td>Secondary</td>
<td>7.8</td>
<td>719</td>
<td>1.7</td>
<td>847</td>
<td>4.5</td>
<td>1,566</td>
</tr>
<tr>
<td>More than secondary</td>
<td>(5.6)</td>
<td>48</td>
<td>(2.4)</td>
<td>58</td>
<td>3.8</td>
<td>106</td>
</tr>
<tr>
<td>Wealth quintile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lowest</td>
<td>3.2</td>
<td>486</td>
<td>1.3</td>
<td>428</td>
<td>2.3</td>
<td>914</td>
</tr>
<tr>
<td>Second</td>
<td>3.5</td>
<td>575</td>
<td>1.0</td>
<td>528</td>
<td>2.3</td>
<td>1,103</td>
</tr>
<tr>
<td>Middle</td>
<td>4.0</td>
<td>601</td>
<td>3.2</td>
<td>519</td>
<td>3.6</td>
<td>1,120</td>
</tr>
<tr>
<td>Fourth</td>
<td>6.2</td>
<td>535</td>
<td>1.3</td>
<td>570</td>
<td>3.7</td>
<td>1,105</td>
</tr>
<tr>
<td>Highest</td>
<td>8.1</td>
<td>749</td>
<td>2.5</td>
<td>835</td>
<td>5.1</td>
<td>1,583</td>
</tr>
<tr>
<td>Total</td>
<td>5.2</td>
<td>2,946</td>
<td>1.9</td>
<td>2,879</td>
<td>3.6</td>
<td>5,825</td>
</tr>
</tbody>
</table>

Note: Figures in parentheses are based on 25-49 unweighted cases.
na = Not applicable

---

Table 6 - Tanzania’s THMIS 2011-12 report – HIV prevalence among 15-24 year olds disaggregated by sexual behaviour

<table>
<thead>
<tr>
<th>Sexual behaviour characteristic</th>
<th>Women</th>
<th></th>
<th>Men</th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percentage HIV positive</td>
<td>Number</td>
<td>Percentage HIV positive</td>
<td>Number</td>
<td>Percentage HIV positive</td>
</tr>
<tr>
<td>Multiple sexual partners and partner concurrency in past 12 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>4.2</td>
<td>265</td>
<td>12</td>
<td>277</td>
<td>4.0</td>
</tr>
<tr>
<td>1</td>
<td>4.7</td>
<td>148</td>
<td>2.0</td>
<td>468</td>
<td>2.7</td>
</tr>
<tr>
<td>2+</td>
<td>1.2</td>
<td>1,423</td>
<td>2.4</td>
<td>1,423</td>
<td>2.4</td>
</tr>
<tr>
<td>Had concurrent partners 1</td>
<td>4.1</td>
<td>123</td>
<td>2.0</td>
<td>252</td>
<td>0.8</td>
</tr>
<tr>
<td>None of the partners were concurrent</td>
<td>4.7</td>
<td>148</td>
<td>2.0</td>
<td>468</td>
<td>2.7</td>
</tr>
<tr>
<td>Condom use at last sexual intercourse in past 12 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Used condom</td>
<td>4.4</td>
<td>576</td>
<td>0.7</td>
<td>780</td>
<td>2.3</td>
</tr>
<tr>
<td>Did not use condom</td>
<td>2.9</td>
<td>1,731</td>
<td>2.2</td>
<td>915</td>
<td>2.6</td>
</tr>
<tr>
<td>No sexual intercourse in past 12 months</td>
<td>4.2</td>
<td>267</td>
<td>0.7</td>
<td>289</td>
<td>2.4</td>
</tr>
<tr>
<td>Total 15-24</td>
<td>3.4</td>
<td>2,574</td>
<td>1.4</td>
<td>1,986</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Note: An asterisk indicates that a figure is based on fewer than 25 unweighted cases and has been suppressed. Total includes 1 woman and 83 men for whom information on multiple sexual partners and partner concurrency in the past 12 months is missing, and 2 men for whom information on condom use at last sexual intercourse in past 12 months is missing. A respondent is considered to have had concurrent partners if he or she had overlapping sexual partnerships with two or more people during the 12 months before the survey. (Respondents with concurrent partners include polygynous men who had overlapping sexual partnerships with two or more wives).

The 2012 National Antenatal Sentinel HIV & Herpes Simplex Type-2 Prevalence Survey in South Africa presents a variety of useful disaggregations, including by age group. Age group prevalence is presented for each region as in the dot plots below.

Figure 4 - Comparison of age specific HIV prevalence profiles between provinces

ANC surveillance does not follow a standard template, so while South Africa’s ANC presents the detailed disaggregation below, reports from other countries will present results in a different way and may disaggregate by different factors.

While the ANC surveillance report presents useful disaggregation, disaggregated results are not presented for adolescents or young people specifically. As with most survey and surveillance exercises, adolescents and young people are part of a broader data collection exercise and age is one of many disaggregation factors. It is challenging to present more than one layer of disaggregation, for example splitting by marital status within age groups, since there are a large number of combinations that can result in complexity.

Table 7 - Association between the demographic and background characteristics and HIV outcome status of survey participants, 2010 to 2012

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Population group</td>
<td>African</td>
<td>28 533</td>
<td>32.5</td>
<td>30 053</td>
<td>31.4</td>
<td>30 345</td>
<td>31.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asian</td>
<td>144</td>
<td>7.1</td>
<td>148</td>
<td>8.8</td>
<td>186</td>
<td>4.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Coloured</td>
<td>2 930</td>
<td>7.0</td>
<td>3 047</td>
<td>7.8</td>
<td>3 069</td>
<td>7.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>White</td>
<td>166</td>
<td>3.0</td>
<td>179</td>
<td>1.1</td>
<td>185</td>
<td>2.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level of Education</td>
<td>None</td>
<td>562</td>
<td>33.5</td>
<td>1 093</td>
<td>32.8</td>
<td>539</td>
<td>33.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Primary</td>
<td>3 786</td>
<td>33.3</td>
<td>3 667</td>
<td>33.7</td>
<td>3 717</td>
<td>33.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Secondary</td>
<td>24 627</td>
<td>30.2</td>
<td>25 846</td>
<td>28.8</td>
<td>25 329</td>
<td>29.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tertiary</td>
<td>2 576</td>
<td>22.7</td>
<td>2 757</td>
<td>22.1</td>
<td>3 108</td>
<td>22.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marital Status</td>
<td>Single</td>
<td>24 802</td>
<td>31.2</td>
<td>26 435</td>
<td>30.1</td>
<td>25 877</td>
<td>30.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Married</td>
<td>6 317</td>
<td>25.1</td>
<td>6 736</td>
<td>24.6</td>
<td>6 687</td>
<td>24.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Widowed</td>
<td>72</td>
<td>41.7</td>
<td>76</td>
<td>50.0</td>
<td>88</td>
<td>51.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Divorced</td>
<td>105</td>
<td>39.0</td>
<td>102</td>
<td>19.6</td>
<td>86</td>
<td>22.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parity</td>
<td>None</td>
<td>12 805</td>
<td>20.7</td>
<td>13 338</td>
<td>19.0</td>
<td>13 839</td>
<td>18.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of live born children</td>
<td>1</td>
<td>10 531</td>
<td>35.9</td>
<td>10 977</td>
<td>34.9</td>
<td>10 972</td>
<td>34.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5 324</td>
<td>37.7</td>
<td>5 543</td>
<td>38.3</td>
<td>5 821</td>
<td>38.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2 146</td>
<td>34.9</td>
<td>2 236</td>
<td>35.3</td>
<td>2 237</td>
<td>38.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>763</td>
<td>33.4</td>
<td>816</td>
<td>32.2</td>
<td>852</td>
<td>33.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>301</td>
<td>26.9</td>
<td>318</td>
<td>28.6</td>
<td>294</td>
<td>29.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>110</td>
<td>28.2</td>
<td>125</td>
<td>22.4</td>
<td>98</td>
<td>27.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>More than 6</td>
<td>71</td>
<td>21.4</td>
<td>55</td>
<td>20.0</td>
<td>63</td>
<td>31.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age of Partner</td>
<td>&lt;15</td>
<td>2</td>
<td>0.0</td>
<td>3</td>
<td>0.0</td>
<td>2</td>
<td>0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>15-19</td>
<td>1559</td>
<td>8.3</td>
<td>1 692</td>
<td>7.7</td>
<td>1 895</td>
<td>6.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20-24</td>
<td>6642</td>
<td>17.9</td>
<td>7 054</td>
<td>15.9</td>
<td>7 160</td>
<td>16.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>25-29</td>
<td>8510</td>
<td>28.9</td>
<td>9 174</td>
<td>28.4</td>
<td>8 929</td>
<td>27.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30-34</td>
<td>6516</td>
<td>37.3</td>
<td>6 941</td>
<td>35.7</td>
<td>7 159</td>
<td>36.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>35-39</td>
<td>4287</td>
<td>41.1</td>
<td>4 621</td>
<td>40.4</td>
<td>4 756</td>
<td>41.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>40-44</td>
<td>2101</td>
<td>39.4</td>
<td>2 256</td>
<td>39.2</td>
<td>2 584</td>
<td>37.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;45</td>
<td>1303</td>
<td>35.9</td>
<td>1 491</td>
<td>36.0</td>
<td>1 456</td>
<td>38.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aware of their HIV status</td>
<td>Yes</td>
<td>20 943</td>
<td>31.5</td>
<td>24 429</td>
<td>29.9</td>
<td>24 800</td>
<td>30.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>10 173</td>
<td>26.9</td>
<td>8 852</td>
<td>26.4</td>
<td>7 982</td>
<td>25.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

54 Ibid
Key population surveillance reports also do not follow a standard template and present varying disaggregation tables. Data for under 18 year olds, or at least under 15-year olds, is often excluded, which undermines the potential for presenting any disaggregated data on adolescents.

Table 8 - Socio-Demographic Characteristics

<table>
<thead>
<tr>
<th>No.</th>
<th>Indicator</th>
<th>Female Sex workers</th>
<th>Clients of Sex workers</th>
<th>Truck Drivers</th>
<th>Female Estate workers</th>
<th>Male Estate workers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Age Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 – 19</td>
<td></td>
<td>13.6</td>
<td>10.5</td>
<td>5.1</td>
<td>5.1</td>
<td>0.0</td>
</tr>
<tr>
<td>20 – 24</td>
<td></td>
<td>39.5</td>
<td>40.8</td>
<td>27.2</td>
<td>4.4</td>
<td>4.8</td>
</tr>
<tr>
<td>25 – 29</td>
<td></td>
<td>28.3</td>
<td>23.8</td>
<td>27.5</td>
<td>16.8</td>
<td>22.6</td>
</tr>
<tr>
<td>30 – 34</td>
<td></td>
<td>12.7</td>
<td>12.6</td>
<td>19.6</td>
<td>22.8</td>
<td>26.2</td>
</tr>
<tr>
<td>35+</td>
<td></td>
<td>5.9</td>
<td>12.2</td>
<td>20.6</td>
<td>55.6</td>
<td>46.4</td>
</tr>
<tr>
<td>Highest Education Level Attained</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None/Primary</td>
<td></td>
<td>70.4</td>
<td>73.1</td>
<td>41.8</td>
<td>29.2</td>
<td>33.5</td>
</tr>
<tr>
<td>Secondary +</td>
<td></td>
<td>29.6</td>
<td>26.9</td>
<td>58.2</td>
<td>70.8</td>
<td>66.5</td>
</tr>
<tr>
<td>Marital Status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td></td>
<td>4.7</td>
<td>7.8</td>
<td>44.4</td>
<td>81.4</td>
<td>83.9</td>
</tr>
<tr>
<td>Single</td>
<td></td>
<td>28.8</td>
<td>14.6</td>
<td>40.1</td>
<td>13.3</td>
<td>15.3</td>
</tr>
<tr>
<td>Divorced/Widowed/Separated</td>
<td></td>
<td>66.5</td>
<td>77.6</td>
<td>15.5</td>
<td>5.3</td>
<td>0.8</td>
</tr>
<tr>
<td>Religion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Christian</td>
<td></td>
<td>87.2</td>
<td>86.7</td>
<td>85.2</td>
<td>86.7</td>
<td>83.1</td>
</tr>
<tr>
<td>Muslim</td>
<td></td>
<td>11.2</td>
<td>13.3</td>
<td>10.9</td>
<td>11.5</td>
<td>16.9</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td>1.6</td>
<td>3.9</td>
<td>0.8</td>
<td>1.8</td>
<td>0.5</td>
</tr>
<tr>
<td>Alcohol Intake</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heavy</td>
<td></td>
<td>28.3</td>
<td>40.5</td>
<td>42.7</td>
<td>17.0</td>
<td>15.3</td>
</tr>
<tr>
<td>Light</td>
<td></td>
<td>40.7</td>
<td>41.8</td>
<td>40.7</td>
<td>24.1</td>
<td>27.0</td>
</tr>
<tr>
<td>Never</td>
<td></td>
<td>31.0</td>
<td>17.7</td>
<td>16.6</td>
<td>58.9</td>
<td>57.7</td>
</tr>
<tr>
<td>N</td>
<td></td>
<td>686</td>
<td>294</td>
<td>567</td>
<td>435</td>
<td>248</td>
</tr>
</tbody>
</table>

In ESA most routine service data does not present numbers for adolescents as a group. They are either merged in with children or with adults. For example, in Malawi’s quarterly routine service number reports, ART figures are given for children 2-14 years old and adults 15-years old and above, but can be split by period and geographic area. This is common for routine service data. Malawi’s routine service reports do however include HCT figures for 15-24 year olds specifically.

Table 9 - Snapshot of a Malawi quarterly report on ART provision

<table>
<thead>
<tr>
<th>District</th>
<th>Site</th>
<th>Facility type</th>
<th>Quarter</th>
<th>Adults 15+ yrs</th>
<th>Children 2-14 yrs</th>
<th>Children below 24 mths</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chitipa</td>
<td>District</td>
<td>Public</td>
<td>2015 Q4</td>
<td>105</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Chitipa</td>
<td>District</td>
<td>Public</td>
<td>2015 Q4</td>
<td>3719</td>
<td>263</td>
<td>66</td>
</tr>
<tr>
<td>Chitipa</td>
<td>Mission</td>
<td>Public</td>
<td>2015 Q4</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Chitipa</td>
<td>Mission</td>
<td>Public</td>
<td>2015 Q4</td>
<td>317</td>
<td>26</td>
<td>6</td>
</tr>
<tr>
<td>Chitipa</td>
<td>Health</td>
<td>Public</td>
<td>2015 Q4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

SUMMARY OF KEY POINTS COVERED IN SECTION 3

- The way that an indicator should be disaggregated depends on how the data is to be used.
- Disaggregated prevalence data help us understand who has a current need for treatment or which groups may drive further infection.
- Age-disaggregated incidence and new infection numbers are important to understanding the current risk of infection faced by different age groups.
- Data on knowledge and behaviours help target prevention efforts although disaggregating HIV prevalence data by knowledge or behaviour can lead to confounding.
- Further disaggregation of routine data on service provision would provide an important opportunity for increased understanding issues around coverage and barriers to access. This should be explored further to overcome the specific challenges detailed in section 4.8 of this report.
- Different data sources present varying levels of disaggregation:
  - National survey reports tend to present the most disaggregation, and have been more standardised across countries than other data source reports.
  - Key population surveillance presents inconsistent disaggregation and often excludes under 18s.
  - Routine service data is the least disaggregated – generally data specific to adolescents is not provided.

---

Table 10 - Snapshot of a Malawi quarterly report on HCT provision

<table>
<thead>
<tr>
<th>District</th>
<th>Site</th>
<th>Facility type</th>
<th>Quarter</th>
<th>Children below 12 mths (Age group A)</th>
<th>Children 12 mths - 14 yrs (Age group B)</th>
<th>Young adults 15-24 years (Age group C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chitipa</td>
<td>District</td>
<td>Public</td>
<td>2015 Q4</td>
<td>1</td>
<td>130</td>
<td>565</td>
</tr>
<tr>
<td>Chitipa</td>
<td>District</td>
<td>Public</td>
<td>2015 Q4</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Chitipa</td>
<td>District</td>
<td>Public</td>
<td>2015 Q4</td>
<td>0</td>
<td>41</td>
<td>164</td>
</tr>
<tr>
<td>Chitipa</td>
<td>District</td>
<td>Public</td>
<td>2015 Q4</td>
<td>0</td>
<td>110</td>
<td>622</td>
</tr>
<tr>
<td>Chitipa</td>
<td>District</td>
<td>Public</td>
<td>2015 Q4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Chitipa</td>
<td>District</td>
<td>Public</td>
<td>2015 Q4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Chitipa</td>
<td>District</td>
<td>Public</td>
<td>2015 Q4</td>
<td>17</td>
<td>108</td>
<td>439</td>
</tr>
</tbody>
</table>

---

57 Ibid
4. Challenges and Solutions

This section outlines a series of challenges that undermine the use of disaggregated data on adolescents. The underlying causes of the challenges and possible solutions are also given.

4.1. Standard grouping strategies

**Challenge:** The standard approaches to grouping people in analysis may not be appropriate to adolescents. It is common to provide disaggregation of results by demographic factors such as age, sex, marital status, race, education and occupation. These are still relevant factors for grouping adolescents, but there are differences in how we should disaggregate by these factors for adolescents. The most important example is age groups. Data is commonly presented for 15-24 year olds or 15-19 and 20-24 year olds. Compared to adults, adolescents undergo considerable social, developmental and psychological changes that are lost or masked when data is grouped into 5-year age bands.

**Underlying causes:** (see sections 2.1 and 2.2)

**Solutions:** One solution is to first analyse data by year of age, rather than automatically creating multi-year bands. The analysis by year of age can be presented to allow users of the data to see any specific points in people’s lives where risks appear to change. The analysis can in some cases also be used to create more relevant age groupings, and comparisons with behaviours, knowledge or other characteristics specific to the population in question, if that information has been collected and sample sizes are large enough.

Age grouping is not the only case where standard grouping practice can cause losses in information.

**What can be lost by jumping to standard age groupings**

In the Kenya’s AIDS Indicator Survey (AIS) report, HIV prevalence for 15-24 year old females was 3% and for males was 1.1%. Disaggregating further, we see that prevalence for 15-19 year old females is 1.1% and for males is 0.9%, while for 20-24 year olds, female prevalence was 4.6% compared to 1.3% for males. This shows us that the gap between the sexes is more pronounced among 20-24 year olds.

Even this further disaggregation masks some important information. Kenya’s AIS report also presents prevalence by year of age. In this graph we can see that female prevalence dramatically increases at the age of 23.

*Figure 5 - HIV prevalence among women and men aged 15-24 years, KAIS 2013*

This may also affect factors such as type of residence, being sexually active, having multiple partners among others. For example, it is common to put geographical areas into urban and rural categories. In reality, there is
a spectrum of types of neighbourhoods rather than a simple dichotomy as illustrated in the box titled ‘What can be hidden in an urban/rural categorisation?’

What can be hidden in an urban/rural categorisation?

In South Africa, grouping areas all settlements into simple ‘urban’ or ‘rural’ categories means that we lose important distinctions between different kinds of urban and rural communities. Recognising this, the South African National Prevalence, Incidence and Behaviour Survey 2012 separates both the urban and rural categories into formal and informal. In the table below, we can see that incidence is the ‘urban formal’ category is much lower than the ‘urban informal’ category.

*Table 11 - HIV incidence (%) and number of new infections by race and locality type among respondents aged two years and older, South Africa 2012*

<table>
<thead>
<tr>
<th>Variables</th>
<th>HIV incidence % (95% CI)</th>
<th>Estimated number of new infections per year (95% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Locality type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban formal</td>
<td>1.06 (0.84–1.28)</td>
<td>227 000 (180 000–274 000)</td>
</tr>
<tr>
<td>Urban informal</td>
<td>2.46 (1.98–2.94)</td>
<td>80 000 (64 000–96 000)</td>
</tr>
<tr>
<td>Rural informal</td>
<td>0.87 (0.69–1.05)</td>
<td>143 000 (113 000–173 000)</td>
</tr>
<tr>
<td>Rural formal</td>
<td>0.84 (0.65–1.03)</td>
<td>19 000 (15 000–23 000)</td>
</tr>
<tr>
<td>Total</td>
<td>1.07 (0.87–1.27)</td>
<td>469 000 (381 000–557 000)</td>
</tr>
</tbody>
</table>

*Table 12 - Percentage having multiple sexual partners in the past 12 months by age and sex among respondents aged 15-years and older, South Africa 2012*

<table>
<thead>
<tr>
<th>Variable</th>
<th>n</th>
<th>%</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Locality type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban formal</td>
<td>9 397</td>
<td>12.9</td>
<td>11.3–14.7</td>
</tr>
<tr>
<td>Urban informal</td>
<td>1 752</td>
<td>12.8</td>
<td>10.4–15.6</td>
</tr>
<tr>
<td>Rural informal</td>
<td>2 943</td>
<td>13.3</td>
<td>11.5–15.2</td>
</tr>
<tr>
<td>Rural formal</td>
<td>1 666</td>
<td>6.0</td>
<td>4.5–8.0</td>
</tr>
</tbody>
</table>
4.2. Disaggregation factors

**Challenge:** Important factors by which we should disaggregate adolescents are not commonly presented.

While the standard demographic factors by which data are commonly disaggregated for adults are also relevant to adolescents, there are other factors that are of particular importance. The majority of the program implementers, researchers and policy-makers we spoke to for this paper mentioned *school attendance* and *household structure* as important factors by which to disaggregate adolescent-related data.\(^{58}\)

### The importance of disaggregating by school attendance

There are two important reasons that we should understand HIV among adolescents in terms of their school attendance. The first is that school is thought to be a protective factor that can reduce their risk of infection [NCBI - A conceptual framework for early adolescence: a platform for research]. As noted above, the disaggregation of cross-sectional data is not well suited to finding such causal links. Longitudinal data would be more powerful for this purpose.

The second is that school is a useful channel by which programs can reach adolescents. HIV-related data on school attendance therefore help program designers understand whether they can reach their target audience via schools.

*Table 13 - Condom use at last sex by current school attendance among female 15-17 year olds that had had sex in the last 12 months*

<table>
<thead>
<tr>
<th>Condom use at last sex</th>
<th>Tanzania (N=664)</th>
<th>Malawi (N=770)</th>
</tr>
</thead>
<tbody>
<tr>
<td>In school</td>
<td>55.1%</td>
<td>42.7%</td>
</tr>
<tr>
<td>Out of school</td>
<td>29.6%</td>
<td>11.6%</td>
</tr>
</tbody>
</table>

---

\(^{58}\) Tanzania THMIS 2011 dataset and Malawi DHS 2010 dataset, available at dhsprogram.com
The importance of disaggregating by relationship to household head

Adolescents live in households with a variety of structures and heads of household. In many cases, a household may be headed by someone other than their parents, including siblings and grandparents. The relationship of the adolescent to the household head were thought by program implementers we spoke with to influence how vulnerable an adolescent is and therefore their risk of infection.

The table below provides an example of how condom use can differ according to the head of the household. As one would expect, condom use was lower among those who were married to the household head. An unexpected finding from this example analysis was that condom use was lower among those in households headed by their parents than those headed by other relatives such as siblings and grandparents, indicating that further investigation of the relationship between household structure and condom use may be useful.

Table 14 - Condom use at last sex by relationship to household head among female 15-19 year olds

<table>
<thead>
<tr>
<th>Condom use at last sex</th>
<th>Tanzania (N=1638)</th>
<th>Malawi (N=1798)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wife</td>
<td>7.9%</td>
<td>8.5%</td>
</tr>
<tr>
<td>Daughter or daughter-in-law of HH head</td>
<td>40.9%</td>
<td>23.8%</td>
</tr>
<tr>
<td>Other relationship to household head (sibling, grandchild, other relative)</td>
<td>47.8%</td>
<td>33%</td>
</tr>
</tbody>
</table>

Source: Tanzania THMIS 2011

Underlying causes: Data on adolescents often come from a source where they are a sub-group of a larger population, such as adults. As such, their data is analysed using the same standards as those used for the rest of the dataset, which includes disaggregation by standard factors such as occupation, urban/rural residence and educational attainment rather than factors that are relevant to only adolescents.

Solutions: The first step of a solution is to ensure that the required data, such as school attendance and household structure, are collected where this would be possible and useful – especially in national surveys, key population surveillance and relevant research studies. The second step is to make disaggregation of adolescent data by these factors standard by including them in report templates and M&E guidelines.

The benefits of making raw data publically available

It is useful to make raw data publically available, so that analysts interested in adolescents can perform adolescent-appropriate analyses on the data in addition to the main report. This would include the use of disaggregation factors such as school attendance, analysis by year of age and other solutions mentioned in this paper.

DHS and AIS datasets are made available to the public, which has allowed for extensive additional analysis for a variety of purposes. One study in 2011 found that more than 1000 studies had been published involving an analysis of DHS data. Over 1000 requests for raw data received by The DHS Program from 2008 to 2013 were for research into adolescents or youth. (Sources59)

59 Tanzania AIS 2011 dataset and Malawi DHS 2010 dataset, available at dhsprogram.com
Other data repositories have also increased access to raw data, such as APHRC’s Microdata Portal and the Life Histories, Health and HIV/AIDS Data Laboratory.

4.3. Exclusion of 10-14 year olds

Challenge: 10-14 year olds are often excluded from surveys, surveillance, models and research.

National surveys such as DHS and AIS make a huge contribution to our broader understanding of HIV epidemics. It has been standard practice to include those 15 and above in these surveys. Under 15s are not well represented in ANC surveillance as they are less likely to become pregnant. Spectrum also presents data on youth defined as 15-24 year olds, with 10-14 year olds merged into a broader child category. Most key population surveillance does not collect data for under 15s, or in some cases all under 18 year olds.

Since most 10-14 year olds have not reached sexual debut in most contexts, it may be considered justifiable that data on them are not collected. However, there are a number of reasons that data on 10-14 year olds are important:

1. Many adolescents with HIV, including 10-14 year olds, were infected at infancy. This group has specific needs, an issue which is explored further later in this report. While it is possible to estimate the size of this group using models based on perinatal transmission data and survival estimates, actual measurement is likely to be more accurate and will allow for more detailed analysis including disaggregation
2. Approximately 10%, 15% and 14% of early adolescents in Tanzania, Kenya and Malawi respectively have reached sexual debut. A systematic review found consistent associations between early sexual debut and HIV. This increased risk was hypothesised to be due to a longer period exposed to risk, physiological immaturity or genital trauma, a higher likelihood of future risky behaviour and of having an age-disparate and HIV-infected partner
3. Early adolescence is an important opportunity to influence future trajectories. It is important that programs working with this group are armed with knowledge and behaviour data about them

Underlying cause: Three possible causes were given by experts for the lack of data on 10-14 year olds:

- Demographic tradition – DHS, which is also the basis of the AIS, began as a data collection exercise for demographics, for which reproductive age is relevant. Reproductive age for women is traditionally taken to be 15-49 years old

---

62 It is worth noting that 15-year olds can also be excluded from studies in some contexts, on the basis that the age of consent is 16.
• Political and cultural acceptability – collecting data with implications for sexual activity is particularly sensitive when the subject is children
• Ethics – there is a perception that including younger adolescents in data collection will result in a much more complicated ethical approval process, which acts as a disincentive to researchers

A paper from the DHS Program however notes some other important factors that make it difficult to include 10-14 year olds in their surveys. These issues may also affect other data sources. The first is that including 10-14 year olds requires an adjustment to the way that data are collected, thereby complicating the already challenging data collection effort. Questions suitable for adults are not always suitable to early adolescents, who may find questions on topics such as sexual activity difficult to discuss. The paper speculates that having to interview children this way may also contribute to interviewer fatigue.

Secondly, including 10-14 year olds would increase the required sample sizes by around 25% (or reduce the number of respondents that are 15 and above, thereby reducing the precision of estimates for older people). Samples are already large, which leads to increased costs and timeframes and more complex training, logistics and data quality requirements.

The paper also describes the risk that including 10-14 year olds would increase call-back rates (where interviewers must revisit the household because the child is at school during their first visit), thereby further increasing the costs and timeframes of the survey.

**Solutions:** A number of strategies are helping ensure that data on 10-14 year olds are collected. Examples include the Population-Based HIV Impact Assessments (PHIA), large-scale national surveys currently being implemented in Malawi, Zimbabwe, Zambia, Uganda and Swaziland that will include data on under 15-year olds as part of data collection encompassing the whole household, and The Global Early Adolescent Study whose focus is 10-14 year olds.

The DHS Program paper on this topic outlines how retrospective data from older adolescents about their early adolescence may be more accurate than data collected via interviews of early adolescents themselves, since they are more comfortable talking about sensitive topics.

The ethical barrier to data collection with 10-14 year olds is multifaceted. Ethical considerations are particularly important when data is to be collected from children, as they may not be able to fully understand the project and their role within it, and they may be especially vulnerable to any negative consequences from participation.

The ethical review process does not render data collection with adolescents impossible in most cases. However, it does mean that data collection must be designed such that adolescent participants are protected and that the benefits are deemed to outweigh any downsides for the subjects. This can affect timeframes, as data collection team training, consent protocols and other factors must meet ethical standards. It can also result in longer ethical review processes, which may be more stringent when child subjects are involved.

Ethics committees are increasingly taking more balanced approaches to research with adolescents, based on better understanding of their risks and needs, and the need to ensure that they are not excluded from research. The HIV AIDS Vaccines Ethics Group (HAVEG) in South Africa has helped facilitate such progress by producing memos that can be attached to research proposals that inform both the researchers and ethics committees.

---


70 Ibid


74 Ibid
committees about the nuances of research with adolescents – an area that not all ethics committees will have expertise in.

The other aspect to this issue is about researchers’ perceptions of the ethical review process. In some cases researchers may avoid the inclusion of adolescents due to the perception that ethical requirements will add to the costs and burden of data collection, ethical review timeframes and other complications, which may not always be as restrictive as they perceive. Guidance from organisations such as HAVEG and initiatives such as TRREE (Training and Resources in Research Ethics Evaluation), an online ethics training resource, help to address these perceptions by providing insight into the actual requirements and steps involved in ethically collecting data from children.

The legal and ethical considerations vary by context. The guidance and information produced by an initiative such as HAVEG may not be appropriate to guiding ethics processes outside of the context in which they work. Audits of the local ethical and legal requirements for research with children and adolescents are a useful way to ensuring that all stakeholders are fully informed.

4.4. Sampling

**Challenge:** Some of the desired disaggregation would require enormous sample sizes.

Sample sizes are one of the determinants of how precise the results of a survey or surveillance will be. In some cases, the sample sizes used are not large enough for the disaggregation that would be most useful to program designers and policy makers. Analyses specific to a sub-national geographic area are particularly problematic.

**Underlying causes:** While other key populations may have data collection activities targeted specifically on them, meaning the sample only needs to represent them, the national surveys and surveillance used to gather data on adolescents need to represent the general population. This means that adolescents are only a subset of the larger sample.

Adolescents are also often underrepresented in samples. HIV tests for adolescents may require the consent of guardians who are not present at the time of data collection, or who do not wish to give consent. Members of this age group are often mobile and not living in the type of established household that is usually included in survey samples. This serves to reduce the sample size available for adolescents.

A third issue is that the multifaceted nature of adolescence means we need multi-layered disaggregation, where we want to disaggregate by more than one factor. Since the dynamics of adolescent HIV tend to be very different for males and females, we need to divide our adolescent sample into two. Adolescence is a time of rapid change, so we need smaller age sub-groups, so we now require four or more divisions. If we are interested in other factors such as school attendance, we will need eight or more divisions. At a national level, samples with this many divisions may still be large enough to be informative, but tend to become too small if we are interested in specific geographic areas. Even if we only need to split the data between five regions, if we are interested in school attendance by sex and stage of adolescence, we now have divided our adolescent sample into 40 sub-divisions.

**Solutions:** One solution to this issue is much larger sample sizes for national surveys and surveillance. The new PHIAs are gathering data with enormous samples to allow for analysis specific to different geographic areas and sub-groups. However, this is a very expensive solution.

---

78 http://phia.icap.columbia.edu/
A second possibility is to oversample adolescents specifically. This would involve including a disproportionately high number of adolescents in the sample to ensure that we can analyse them as a separate group, without inflating the number of other respondents in the sample. This is more economical than larger overall samples. This solution may not be practical for household surveys where samples are based on interviews with all eligible members of selected households. Adding respondents from a particular age group would require visits to additional households, including data collection about those households, at which point the advantages of oversampling one specific group are negated.

How disaggregation can make even large sample sizes inadequate

The table below illustrates how a large survey or surveillance sample can allow for age and sex disaggregated analysis and possibly an additional disaggregation factor such as school attendance, but become inadequate when disaggregating by multiple factors, particularly when we are interested in analysis by region, using the dataset from Tanzania’s 2011-12 THMIS.

Table 15 - Tanzania THMIS Regional Data

<table>
<thead>
<tr>
<th></th>
<th>Sample – 20,811</th>
<th>Successfully completed HIV test – 18,342</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age group</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15-19 year olds</td>
<td>4,323</td>
<td>18-24 year olds – 3,302</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1,991</td>
<td>Female – 2,332</td>
</tr>
<tr>
<td>Female</td>
<td>1,378</td>
<td>Male – 1,924</td>
</tr>
<tr>
<td><strong>School attendance</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In school – 1,132</td>
<td>Out of school – 883</td>
<td>In school – 1,130</td>
</tr>
<tr>
<td>In school – 269</td>
<td>Out of school – 1,062</td>
<td>In school – 185</td>
</tr>
<tr>
<td><strong>Number per region</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 to 62</td>
<td>12 to 66</td>
<td>11 to 61</td>
</tr>
<tr>
<td>16 to 84</td>
<td>11 to 76</td>
<td>3 to 22</td>
</tr>
<tr>
<td>1 to 20</td>
<td>31 to 100</td>
<td></td>
</tr>
</tbody>
</table>

Source: Tanzania THMIS 2011-12 dataset, available at dhsprogram.com

A third solution is to conduct data collection exercises specific to adolescents, rather than attempting to capture data on them as part of a wider national survey effort. Adolescent specific data collection allows for samples to be designed with the specific requirements of adolescents (for example, ensuring those in-school and out-of-school are both represented). It also allows for solutions to some of the other challenges presented in this report, such as the need for adolescent appropriate data collection methods. The Global Early Adolescent study79 is an example of such an initiative. However, separate data collection for adolescents can be an expensive solution.

The final option is to recognise that large-scale national surveys may not be suitable to achieving anything more than a big picture snapshot of adolescent HIV, for which the existing samples are sufficient. This idea is discussed further in the challenge 'The Need for Nuanced Data'.

4.5.Confounding

Challenge: Causal relationships cannot always be inferred from simple disaggregation.

Many of the program implementers we talked to wished to determine the factors that resulted in adolescent vulnerability through disaggregation. For example, they wanted to see whether being relatively poor resulted in higher vulnerability and therefore higher HIV prevalence. The simple cross-tabulations usually presented do not easily allow for such inferences, since confounding is not adjusted for. Basic crosstabulation of wealth and HIV prevalence indicated that the wealthiest were at a higher risk of HIV than poorer groups. However, when adjusted for urban/rural residence, this relationship disappeared. This implies that living in urban areas is the real risk factor, not being wealthier.

Underlying causes: This issue comes back to the variety of ways in which data is used. There is nothing incorrect about showing that those in higher quintiles have higher HIV prevalence. For example, it may be useful to know that to reach people living with HIV with certain messages, it is important to use communication channels that are relevant to both the wealthier and poorer parts of the population. However, there is a problem if this data is used to determine the causes of higher HIV prevalence. Data should be presented in ways that suit the specific question that is being asked.

Solutions: There many various ways to adjust for confounding, including multivariate analysis techniques. Such techniques are widely used and documented. However, the difficulty is that the way that sample sizes are determined, data is analysed and presented should be based on how the data is to be used, but HIV data from surveys and surveillance is used in a variety of ways. Unlike research studies focused on a specific question, they are general purpose data. To present analyses that would cover all the potential uses of data would make reports unwieldy. There are two strategies that can be employed to overcome this:

1) Ensure that reporting templates and guidelines present results that are appropriate to the priority needs of the data’s audience. This requires consultation with users of data to determine the most important ways that the data is used.
2) Provide access to raw data. This allows analysts to produce results appropriate to their needs. Despite each DHS having a long standard report, DHS datasets have been reanalysed thousands of times for varying purposes because the raw data is publically available. In some cases, other data sources such as surveillance datasets or datasets from research studies have also been made available in data repositories. Expanding such efforts would ensure that data can be analysed and presented to meet the myriad needs of programmers and policy-makers.

4.6. Modelled estimates

Challenge: The disaggregation presented by models such as Spectrum is limited.

Spectrum is a commonly used model in the world of HIV. It has a variety of useful functions and is easy to use. It is used to estimate the numbers of new infections occurring in a population and the numbers of deaths due to AIDS among other indicators. As we have seen, both incidence and deaths due to AIDS are particularly relevant to understanding the risks faced by adolescents. Spectrum’s widespread adoption has allowed for global, regional and national estimates on HIV epidemics to be made consistently. Spectrum as standard includes results for 15-24 year olds as a separate group. However, more detailed breakdowns of estimates into smaller age groups or by other factors are more difficult to obtain.

Underlying causes:

As with most models, Spectrum must employ assumptions to make its calculations possible. Some of the data and assumptions that Spectrum uses limits its ability to present disaggregated data. While Spectrum is focused on here, these issues apply to many other modelling exercises.

Spectrum must use the available survey, surveillance and routine service data to create estimates that this source data cannot provide. It produces figures for years where direct measurements were not taken, and produces indicators that also are difficult to measure routinely, such as new infections and deaths due to AIDS. To fill in the gaps between actual measurements via surveillance and other sources and the indicators it needs to produce, Spectrum must make assumptions. One of the most important is that results from research can be generalised to the population in question. For example, to estimate the number of perinatally HIV-infected adolescents, Spectrum uses information from research into survival of perinatally infected adolescents and assumes that the survival is similar for the population in question.

Spectrum estimates the incidence of HIV by age by applying a distribution of incidences across 5-year age groups to estimates of overall incidence. These distributions are usually based on estimates from cohort studies in the ALPHA network, which are assumed to apply to all generalised epidemics. In other words, the same pattern of infection by age is assumed to occur in different contexts. Since the distributions are based on 5-year age bands, any specific years of age where incidence spikes are lost.

Spectrum also estimates sex-specific incidence by applying a standard ratio of male to female incidence. This is assumed to be the same for all age groups, so differences in the sexes by age are not reflected in the estimates.  

The results that Spectrum produces can be evaluated against actual measurements to determine whether the assumptions made were reasonable and the figures can be adjusted when necessary.

Another issue with disaggregating modelled estimates is that they are only as good as the data on which they are based. Since models often combine estimates from different sources, errors in those sources are compounded. Surveillance and surveys are affected by sampling error, which is increased when looking at subgroups within the population as the sample size effectively becomes smaller. Furthermore, we have seen that routine service data specific to adolescents is usually not available. The limitations of the underlying data for disaggregation combined can heavily undermine the potential for modelled estimates to be disaggregated.

**Solutions:**

Spectrum is regularly improved upon and updated. The assumptions and source data that models can use are also progressing and becoming more detailed and accurate. For example, the PHIA surveys will use large samples that include adolescents, providing us with a powerful data source with which to calibrate modelled estimates.

In the region, SACEMA is pioneering various new methods for incidence estimation. A modelling approach has been developed that does not use 5-year age bands for estimating the age distribution of incidence. This approach uses data on the level of individual people rather than aggregated data, allowing for more accurate estimates particularly when the aim is disaggregated information.

Another innovation is the use of laboratory tests that determine whether an HIV positive person was recently infected. These methods are being developed into new approaches to estimating incidence. In terms of disaggregation, this could be extremely useful both for direct indications of risk and incidence, and for modelling trends in adolescent risk. Incidence information can be gained from surveys, which as we have seen are well suited to disaggregating data.

**4.7 Routine service data/HMIS**

**Challenge:** Most HMIS do not produce data specific to adolescents.

Most HMIS in ESA provide data on testing and treatment services for children and adults (usually children are categorised as those aged 0-14 years old, and adults are those 15-year olds and above). Deviations from this are rare but do exist – for example Malawi’s quarterly reports include HTC figures for 15-24 year olds specifically. In general, there are no figures specific to adolescents, let alone data split by sex or other factors.

**Underlying causes:** Data on age, sex and in some cases other factors are collected at the facility level for each patient. This means that the data required for age and sex disaggregation at least exists. The problem is that this data generally does not go beyond the facility.

---

82 PMC, Updates to the Spectrum model to estimate key HIV indicators for adults and children, available at [http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4247263/](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4247263/), 2014


The facility will send data up the reporting chain to the local government using standard reporting templates. At this stage in the reporting chain, the age and sex information is lost. More complex reporting templates may allow for more disaggregated data to be reported up. However, the reporting system needs to be kept simple and practical, as it needs to fit into the day to day activities of facility staff. The priority activity for these staff members is to deliver services rather than fill in time consuming report templates.

**Solutions:** Complexity in an MIS can undermine data quality and completeness. Complicating existing paper reporting systems with additional fields specific to adolescents is not ideal. There are three other solutions being employed that can provide adolescent-specific data.

**Revisit original records**

One option is to conduct a data extraction exercise where a sample of facilities are visited and patient records are reviewed to create a dataset with age and sex variables. This dataset can then be analysed with any combination of age- and sex- disaggregated data required, as well as disaggregation by geographical area. Such exercises are being undertaken in both Malawi and Tanzania as a collaboration between UNICEF and the national governments.

**Sentinel sites**

While national management information systems must be kept simple and practical, and ideally would work in the least resourced facilities in the country, it is possible to select a sample of sites for more intensive data collection. These sites are called sentinel sites. Additional resources can be put into technology, training and other resources to make the collection of additional data easier, and provide access to un-aggregated raw data.

**Patient level databases**

The third solution is to use computer-based management information systems in facilities such that patient-level databases are created. Such systems do not require facility staff to spend time aggregating figures to send up the reporting chain. Individual level data can be accessed to produce reports. If these databases include a demographic record for each individual patient with their sex and age, it should be possible to produce service figures disaggregated by age and sex, and by geographic area since the clinic’s location will be known. Other disaggregation factors can also be used if they exist in the databases.

An important example of such a system being successfully implemented in the region can be found in Malawi. The Baobab Health Trust, a locally run organisation specialising in management information systems, in partnership with the government have created electronic systems that currently captures data on a large proportion of the HIV services in the country.

Keys to Malawi’s success were given as development of local capacity and focusing the design of the system on improved service provision for patients. It is important to keep in mind that adding complexity to management information systems can undermine their success, to a greater extent than other data sources like surveys and research studies. Adding fields required for better disaggregation of data, for example, can be costly. Since data on age, sex and geographic area are usually included in patient level databases for various purposes, disaggregation by these factors should be possible. Other electronic patient information management systems have been developed in ESA, creating potential for this approach in other countries too.

### 4.8. Data accessibility

**Challenge:** *Useful data is not seen by those that need it.*

---

85 [http://baobabhealth.org/](http://baobabhealth.org/)
The data on adolescents that is easily accessed and that has been widely disseminated is only the tip of the iceberg of the data that has been collected. The disaggregation of adolescent or youth related data in national progress reports is inconsistent as can be seen in appendix 1. Data hubs such as UNAIDS’ excellent AIDSinfo\textsuperscript{86} now include some results specific to 10-19 year olds split by sex, such as number of new infections. Many knowledge and behaviour indicators are presented for 15-24 year olds. However, more detailed breakdowns are unavailable.

Reports from national surveys and surveillance do include chapters or sections dedicated to adolescents or youth with varying degrees of detail. It is very positive that a chapter is dedicated to 15-24 year olds in the standard DHS and AIS reports, or that age group data is presented in South Africa’s ANC survey and surveillance reports. However, these are often a relatively small part of longer reports, and did not appear to be well known from discussions with program implementers. We have seen that adolescents are a group with a unique set of needs. Policy-makers and programmers would benefit from access to adolescent-specific information without having to go through a variety of long adult-focused documents to find it.

**Underlying causes:** Adolescent data are often analysed and presented as a subset of a broader adult population. Reporting the findings of such general purpose data sources requires that a balance be struck between meeting the needs of their broad audiences and not overloading the reports and hubs with details on the many sub-groups included in the datasets. Unless more adolescent-specific analysis is specifically prioritised, time and resources available to do them may also be limited by the need to cover a wide range of ages and issues in survey reports.

**Solutions:** Various efforts are being undertaken to rectify this. Global and regional reports specific to adolescents, such as UNICEF’s Synthesis report of the Rapid Assessment of Adolescent and HIV Programme Context in Five Countries, provide an accessible source of detailed information\textsuperscript{87} specific to adolescents. Idele et al’s publication on adolescents provides an excellent summary of the current data on adolescents globally. However, while excellent as global or regional overviews, due to their broad geographical scope, the information contained in these high-level reports may not always be specific enough to the contexts that program implementers work in.

An excellent example of a solution to this is presented in the box below. This re-analysis of existing data in Tanzania demonstrates what is possible when an adolescent-specific analysis and dissemination plan are applied to existing data sources. There are other ways of providing access to data beyond reports. One example is StatCompiler\textsuperscript{88}, which allows the public to produce a large variety of tables and graphs based on DHS data, depending on their requirements. Such tools can be enhanced to cater to the requirements of adolescent-focused users\textsuperscript{89}.

---

\textsuperscript{86} UNAIDS, AIDSinfo, available at: \url{http://aidsinfo.unaids.org/}


\textsuperscript{88} \url{http://www.statcompiler.com/}

4.9. Need for nuanced understanding

**Challenge:** *The standard large-scale data sources used for understanding HIV epidemics may not be what is needed to further our current understanding HIV among adolescents.*

A view expressed by some of the experts we consulted was that there is a need to develop a deeper understanding of the dynamics between the multiple facets of adolescence and HIV, but that we often hope to achieve this with data sources that are not well suited to this purpose – bigger and better surveys, more sophisticated management information systems and surveillance mechanisms, all with better disaggregation.

**Underlying causes:** We rely on indicators to understand and evaluate HIV and its programming at a high level. However, indicators only indicate— they tell us about scale and general directions. Many may tell us only limited amounts about important outliers and particular sub-groups because they generally report "the average" for any population or group. Indicators do not replace in-depth investigations of issues.

Most surveys and surveillance provide detailed data about one moment in time. Snapshots of indicators, even with the largest samples, useful disaggregation and best data collection methods, can only provide limited insights into the complex dynamics of adolescent vulnerability. Routine service data includes basic facts about a person's characteristics and service uptake. Neither is well suited to providing insight into the interaction of biological, psychological and social transitions that can result in adolescent vulnerability.

**Solutions:** The broader takeaway point is to manage our expectations of what each data source can provide us. Many alternative approaches such as qualitative methods and longitudinal studies are used successfully to answer these questions.

4.10. Muddling two distinct groups

---

Challenge: Routine data and survey data cannot usually distinguish between adolescents infected during infancy and during adolescence.

Whether an adolescent was infected during infancy or adolescence has different implications for risk, treatment adherence and prevention efforts. Distinguishing between them in survey, surveillance and routine service data would therefore be useful. However, global reporting currently does not distinguish between the groups.

A modelling exercise in 2009 estimated that 36% of HIV infected infants are 'slow progressors', surviving to a median age of 16. The number of deaths among these slow progressors in South Africa if untreated was estimated to increase to 23,000 per year by the year in 2030.

Underlying causes: There are no definitive characteristics or tests that would help us determine when an adolescent with HIV was infected unless they have a history of being diagnosed and treated previously. There is evidence that those infected at infancy that have aged into adolescence will suffer some physiological disadvantages. However, these may not be clear enough to allow service providers and data collectors to ascertain when an adolescent was infected. Thus standard routine data systems, cross-sectional studies, surveys and standard surveillance are not well suited to tracking perinatally HIV-infected adolescents.

Solutions: Currently the main method available to us to estimate the number of perinatally HIV infected adolescents by modelling the survival of those infected during infancy using a tool like Spectrum. There are also longitudinal cohort studies that have furthered our understanding of this group. However, in many countries there is still considerable uncertainty about estimates of the number of HIV infected adolescents. Calls have been made for surveillance systems designed to systematically track perinatally HIV infected adolescents.

SUMMARY OF KEY POINTS COVERED IN SECTION 4

- Standard age grouping strategies mask important variations in the adolescent age groups. This can be overcome by first analysing data by year of age to enable a more nuanced analysis of risk.
- In addition to the standard demographic factors by which data are commonly disaggregated, school attendance and household structure are important additional disaggregations for adolescents.
- Strategies to increase the focus on the 10-14 age group include: adolescent-specific data collection as well as data collection encompassing the whole household; retrospective interviewing of older adolescents; attaching memos to research proposals for ethical approval to explain the nuances of adolescent data collection; and providing additional guidance for researchers.
- Issues of small sample sizes can be overcome through several means: increasing overall sample sizes (although this may not be cost effective); over-sampling adolescents (although this may not always be feasible); adolescent specific data collection; and simply recognising that large-scale national surveys only provide a snapshot view.
- To address the complexities of confounding stakeholders should produce audience-specific reporting templates and guidelines and/or provide access to raw data.
- Several innovations are being employed to overcome the limited disaggregation under SPECTRUM such as: improving source data and new methods for incidence estimation, including the use of laboratory tests that determine whether an HIV positive person was recently infected.

92 Ibid.
The lack of disaggregation in HMIS data is being overcome through: Revisiting original records for a sample of facilities; selecting a sample of sites for more intensive data collection; and creating patient-level databases from which specific disaggregations can easily be extracted for analysis.

Data accessibility barriers can be overcome by adolescent-specific publications; re-analysis of existing data; and tools to make data accessible beyond reports.

There is a need to manage expectations of what each data source can provide us. Alternative, more qualitative approaches may be better suited to telling us what we want to know.

Distinguishing between perinatally HIV infected adolescents and those infected during adolescence is hard and although tools like Spectrum or longitudinal cohort studies help, stakeholders are calling for surveillance systems designed to systematically track perinatally HIV infected adolescents.
5. Conclusions

The response to the challenges of defining and disaggregating adolescence will be largely driven by the needs of advocates, planners, programmers and service providers. Adaptations will have to balance what is desirable with what is feasible given the practical challenges raised by the diversity of data collection methodologies, increased workloads, increased costs and lower cost effectiveness when we seek to refine information further. Thus developing agreement on what questions are priorities to answer and which produce “nice to have” information will be a key first step to deciding what disaggregation of routine or other data and analyses should be undertaken.

5.1. Defining adolescence for data

Overlapping and broad additional categories layered on top of the child and adult categories make the specific needs of adolescents less visible and less clearly defined. This makes it more difficult to target and design interventions to use scarce resources to maximum effect.

For most data collection and reporting purposes, rather than using the ‘adolescent’ and ‘youth’ categories, we can standardise and more widely use three stages of adolescence – early adolescence, middle adolescence and young adulthood. These stages better group together people with similar needs and experiences, and are mutually exclusive.

The standard age ranges corresponding to these stages can either adhere to the current standard 5-year age bands (10-14, 15-19, 20-24) or we can attempt to base them on years of age that correspond to generalisable life transition points.

It is debatable whether there is too much variation between individuals and contexts to allow us to identify years of age that would correspond to generalisable life transition points – i.e. a year of age at which people start puberty, leave school, reach sexual debut, get married and have children. This is a question that may warrant further investigation.

If it is not possible to determine useful and generalisable life transition points, it would probably be most practical to stay with 5-year age bands for most purposes, as they already so commonly used. However, to address certain issues, data and analyses of risk and needs in relation to transition stage rather than age, or in addition to it, may provide useful insights to programmers and service providers. Where policy and planning questions require, and samples can be large enough, it may also be useful to collect data and produce analyses for individual years or groupings of years. Strategic choices will also have to be made about where to change routine data collection systems, and where to use non-routine studies or sentinel sites to explore more nuanced disaggregation and inform decisions about whether it is worthwhile to change routine systems.

5.2. Disaggregating adolescent data - recommendations

The calls for more disaggregated data on adolescents can be partially met by pushing existing data further:

- Data on adolescence should be analysed based on the needs of the programmers and policy-makers that work with adolescents, rather than using the standard approaches applied to adult analysis
- Where possible, there should be consistency in how adolescent data is disaggregated, analysed and presented
- Analysis of adolescent data should be disseminated to those that work with adolescents
- National survey data should be disaggregated by key adolescent-related variables such as school attendance and household structure
- Presenting important figures by year of age can identify changes within age groups that can be masked when defaulting to 5-year age bands
- Where we are interested in the relationships between dependent and independent variables, adjusting for confounding should be done wherever possible.
Survey planners and funders should consider strategic provision of extra resources to increase sample sizes, numbers of variables and analyses to resolve persisting uncertainties about trends and relationships between key variables such as schooling, pregnancy and HIV risk.

Ensure that datasets are made available for further analyses to explore more issues and trends than might have been possible in initial studies.

Many of these improvements have however been achieved in the past with efforts that could be replicated:

- Consult the programmers and policy-makers through well designed processes to identify analyses that would most benefit their work
- Re-analyse or revisit existing data sources to create adolescent-specific publications using adolescent-appropriate disaggregation factors and methods
- Re-analysis of survey, surveillance and research data occurs when the raw data is accessible – something that data repositories are seeking to achieve
- Disaggregation of routine service data is made possible through the use of data extraction exercises, sentinel sites or patient level databases
- Disseminate the adolescent-specific information to the programmers and policy-makers that need it

Further improvements could be gained with additional data collection exercises, though these are likely to require more resources.

- Support survey teams and ethics committees to ensure particularly that 10-14 year olds and 15-16 years old are represented, when appropriate, in datasets
- Conduct adolescent-specific data collection exercises where adolescent-appropriate data collection methods, samples and analysis can be used, including the use of qualitative and longitudinal study designs where needed
This document has been prepared as part of the Evidence for HIV Prevention in Southern Africa (EHPSA) programme which is supported by UK aid from the Department for International Development (DFID) and Sweden, through the Swedish International Development Agency (Sida) - mandated to represent the Norwegian Agency for Development Cooperation (NORAD).

The content and opinions as expressed within this document are those of the authors and do not necessarily reflect the opinion of UK aid, DFID, Sida, NORAD or that of the programme managers, Mott MacDonald.