Achieving good health and well-being in Africa by 2030 using multi-state models, survival analysis, statistical methods for evidence-based medicine, diagnosis and determination of risk factors

Humphrey E. Misiri
BBSC, Box 2742, Blantyre, Malawi
Tel.: +265 994 194 469; E-mail: misiri.humphrey@gmail.com

Abstract. Seventeen Sustainable Development Goals (SGDs) were adopted by the World Health Organization (WHO) in 2015 for the 2030 Agenda for Sustainable Development. Sustainable Development Goal 3 (SDG3) is 'Better health and well-being by 2030'. According to WHO, good health in the context of SDG3 is assessed with respect to the level and distribution of individuals' and communities' healthy life, conditions that affect health and well-being and risk factors whose presence would affect health and well-being. The overall aim is that each SDG target is achieved by 2030. In 2018 the WHO used statistical methods to assess the state of health in Africa in the context of SDG3. Their analysis revealed successes and shortfalls towards attaining SDG3. Backed by public health and other activities, statistics play an important role in improving the health and well-being of Africa. This paper explains how statistics can be used to help African countries to attain SDG3, in its role in modeling event histories, diagnosis, evidence-based medicine, determination of risk factors of exposures of morbidity and mortality, determination of risk factors of morbidity and mortality, the computation of the level and distribution of vital events, measuring disease frequency and progress, quantification of life expectancy and monitoring and evaluation.

Keywords: Health, data, sustainable development, goal, well-being

1. Background

Sustainable Development aims to maximize the welfare of the current generation without jeopardizing the starting position for future generations or living at the expense of people in other countries [1]. In 2015, the United Nations adopted 17 Sustainable Development Goals (SDGs) to be achieved by member countries by 2030 [2]. This is the 2030 Agenda for Sustainable Development. Each Sustainable Development Goal has indicators for monitoring its progress to achievement. With the help of 230 indicators distributed over 169 targets this monitoring is performed. Out of the 17 goals one goal is explicitly dedicated to health. This is the Sustainable Development Goal 3 (SDG3): "Good Health and Well-Being" by 2030. Fifty of the 169 targets are relevant for the attainment of SDG3 [3]. In this paper I explain how statistics can help in the attainment of SDG3 in Africa.

Statistics is the science of learning from data and it is also a modern methodology that is part of the standards of this information age we live in [1]. Statistics requires data as input. Collection of health data is an important activity in determining the health status of a population and finding ways to improve it. Health data are collected routinely in health facilities en masse. District Health Information Software 2 (DHIS2) is software that was designed for collecting data from district health

information systems. It is used in at least 56 African countries, to capture data from health facilities for use in decision making.

Statistics is often used to facilitate via simplification, the understanding of the complex processes that underpin a health problem in a population. As such, statistics is a tool that can be used to improve the health of people in society. In fact, at the beginning of the twenty-first century, the application of statistics in medicine was touted as one of the 11 important medical developments [4]. Improving health in a population should be a process that should include obtaining information about health problems in a population which will inspire leaders to formulate policies to tackle those problems. This is possible if correct, high quality and up to date statistical information about the health status of a population is made available to the right people; those that can make timely decisions using that information in order to improve people's health.

2. Where is Africa now? The current state of health in view of SDG3

Four years after the SDGs were adopted the status of health in Africa is not that rosy. A report by the WHO Office of the African Region summarized the state of health in Africa according to life expectancy, the burden of morbidity and mortality and the burden of risk factors of morbidity and mortality [3]. This report shows that as of 2018, Africa had experienced successes and failures in attaining SDG3. The report (WHO Regional Office for Africa (2018), page 14) [3] says:

– "The healthy life expectancy (a measure of life expectancy adjusted for years spent with disability) has been increasing in the Region, from 50.9 years to 53.8 years between 2012 and 2015, which represents the highest increase in any WHO region. Additionally, the gap in healthy life expectancy between the best and worst performing countries in the Region has reduced from 27.5 to 22 years. However, it still shows inequities, with healthy life highest in countries with better economies. The improvement is fastest in large population countries and in those with high population densities. Additionally, the levels of healthy life in the Region are still very low compared to other regions.
– The burden of disease is now driven by communicable conditions, non-communicable conditions and violence/injuries. However, lower respiratory conditions, HIV/AIDS and diarrhoeal diseases still

represent the top causes of both morbidity and mortality. Levels of morbidity and mortality are significantly reduced. DALYs [3] (WHO Regional Office for Africa (2018), page 34) due to the top 10 causes of morbidity have more than halved between 2000 and 2015, driven by reductions in malaria, HIV/AIDS and diarrhoeal diseases. The crude death rate due to the top 10 causes of mortality has also fallen, from 87.7 to 51.3 per 100 000 population in the same period. No significant reduction is seen for non-communicable diseases (NCDs).
– However, the burden of risk factors to morbidity and mortality is not seeing commensurate reductions. A person in the Region aged between 30 and 70 years has a 20.7% chance of dying from one of the major NCDs. All the four major risk factors identified in the Global Action Plan for the prevention and control of NCDs (2013–2020) [5] are high in the Region. These include alcohol abuse, insufficient physical activities, unhealthy diets and substance abuse."

Point number 2 above mentions NCDs. The key Non-Communicable Diseases which countries are expected to focus on in their preventive activities are chronic respiratory disease, cardiovascular disease, cancer and diabetes [5]. Three major risk factors of the key NCDs which countries are supposed to focus on in their preventive activities are alcohol abuse, insufficient physical activity, unhealthy diets and tobacco use [5].

The assessment by WHO revealed successes and failures. Successes imply that countries are trying their best to improve the health of their people. Failures mean that a lot more work needs to be done. The report shows that a lot more work needs to be done to increase life expectancy, reduce the burden of morbidity, mortality and the burden of risk factors of morbidity and mortality. It also shows that some emphasis must be placed on NCDs.

In their assessment for state of health in the African Region WHO reported that the burden of risk factors to morbidity and mortality is not seeing commensurate reductions [3]. Firstly, this may imply that interventions against morbidity and mortality which African countries in the African region implemented in the past are ineffective. Secondly this may imply that more preventive work is being done but such work is not based on data from the ground thereby making all efforts irrelevant. Lastly, this may imply that more information is needed in order to devise very effective interventions.

Interventions may be irrelevant or ineffective due to the fact that they are not designed using timely and

relevant statistical information from data collected from an area. It might also be due to the fact that some data that would have assisted in designing very efficient interventions is not used or is missing.

Extra work needed to achieve SDG3 by 2030 has to include computerization of routine health data collection, collection of extra health data, establishing extra sources of routine health data, increased use of health information from routine health data in decision making and in the design of interventions and the increased use of statistics in determining risk factors.

3. How can Africa attain health and well-being for all by 2030?

The status of health in Africa as revealed by the WHO report shows that African countries should maintain the current gains and improve areas that need to be redressed in order to reduce or eliminate problems that were discovered. This can be achieved by using 10 approaches which are summarized in three themes. The first theme is computerizing data collection and establishing sources of vital statistics and NCD morbidity data. The second theme is the identification of risk factors of health-seeking behavior, morbidity and mortality, and the determination of factors of risk factors of four NCDs. The third theme concerns strategies for speeding production of health information and its use.

Computerizing data collection and establishing sources of vital statistics and NCD morbidity data will help in timely data analysis, provision of information on vital rates for reducing mortality, computing precise estimates of healthy life expectancy and reduction of NCD morbidity and mortality. Determination of risk factors of health-seeking behavior, morbidity and mortality, major factors of risk factors of four NCDs will also help in provision of information for designing interventions for increasing access to health care and reducing the burden of NCD morbidity and mortality. Strategies for speeding production of health information and its effective use will be helpful in the timely provision of health information and increased use of health information in decision making and timely detection of health problems in a population. All these components are essential for achieving SDG3. The 10 approaches for achieving SDG3 are described in the sections that follow below.

3.1. Computerized data collection

Health facilities collect routine patient level data from people who come for medical attention. These data – if analyzed – can help in identifying health problems in a catchment population of a health facility. Therefore the starting point for improving health and well being is the proper management of health data from health facilities. One way to properly manage health data it to structure it via a computer application. An example of such a software application is DHIS2.

Health data should be entered in a computer database like the DHIS2 database. DHIS2 has an inbuilt set of standard indicators for monitoring various issues in a health district. Health facility utilization rate, clinic attendance, disease trends and outbreaks can also be detected using DHIS2. DHIS2 also has functions for performing simple analyses and aggregation of the patient-level data it collects.

Via DHIS2, massive data can be collected which can be used for various purposes. The data help in identifying disease trends, utilization of services, episodes of illness and epidemics. Disease trends provide insight into the progress of disease occurrence with time. Consequently, information from DHIS2 data should be used in health planning and identifying health problems in a catchment population of a health facility. Once the problem has been pinpointed, a follow up survey should be conducted to discover the risk factors of the problem. Remedial measures based on this information from the survey should be taken.

3.2. Establishing vital registration systems

In their assessment of the state of health in Africa in 2018, WHO used the healthy life expectancy indicator (HALE) as a measure of how healthy people are living [6,7]. HALE is computed from data from the vital registration systems and supplemented with information from alternative sources like surveys [7]. However overall HALE estimates from vital registration data are considered better than HALE estimates based on data from other sources because there is more uncertainty when estimating HALE from data from alternative sources [7].

Vital registration systems record births and deaths in the community and so provide data for computing rates like age-specific birth rates and death rates. Age-specific mortality rates [8] from vital registration systems are also used for compiling life tables. Detailed instructions for constructing life tables are in Elandt-Johnson and Johnson [8]. These rates and the life expectancy from life tables provide an overview of life expectancy in the community. A typical table constructed from vital registration data looks like the table below.

Table 1
Dummy table of vital registration data

Age group (i)	Number of deaths (d_i)	Number of people (p_i)	Crude incidence rate (d_i/p_i)
0–4	d1	p1	r1
5–9	d2	p2	r2
10–14	d3	p3	r2
15–19	d4	p4	r3
…	…	…	…
Total			

As of 2018, the levels of healthy life in Africa were low compared to other regions [3]. To use these rates most effectively for efforts to improve the life expectancy, data from vital registration systems must be monitored at country level. To effectively monitor life expectancy in view of the SDG3, each country should establish a vital registration system. Within each country, vital events should be monitored at district level.

Information from a vital registration system can for example help to reveal top causes of mortality, gender differentials of mortality and age groups with lowest life expectancy. A graph of crude death rate against age will reveal the age groups with highest mortality rate. This is the approach often used for finding the age group with the highest incidence rate of cancer [9].

3.3. Establishing supplementary data sources in order to improve care for patients with non-communicable diseases

Non-communicable diseases (NCD) are one of the hindrances to achieving SDG3 In 2017, the main driver for the burden of disability globally were non-communicable diseases which caused 80% of disability cases [10]. Globally, seven of the ten top causes of early death were non-communicable diseases [10]. In 2018, a WHO assessment revealed that between 2000 and 2015, there was no reduction in NCD mortality in Africa [3].

Four major NCDs are cardiovascular diseases, cancers, diabetes and chronic respiratory diseases [5]. These four NCDs share four behavioral risk factors namely tobacco use, unhealthy diet, physical inactivity and harmful use of alcohol [5]. Hypertension is also a non-communicable disease.

To adequately focus on NCDs, each NCD should have a register at district level in order to make the people that are suffering from NCDs known. This will enable better planning for their treatment and palliation and – as the magnitude of the problem will be now known – the development of strategies for their prevention, palliation or treatment.

Some NCDs like diabetes and cancer can be pre-

vented though better screening of the risk population. Screening will also help in the establishment of a disease register. Disease registers will provide estimates of NCD incidence and incidence rates. These will be important for monitoring effectiveness of preventive interventions.

A graph of crude incidence rate against age will help to pinpoint the age groups with high incidence of NCDs [9]. A graph of age-standardized incidence rates against time will show whether the incidence of NCDs is changing [9]. Estimates of incidence from disease registers should be monitored to see whether the trend of occurrence of the NCDs is increasing or decreasing.

3.4. Attempts to reduce the burden of four major risk factors of NCDs

The burden of four major risk factors of NCDs namely alcohol abuse, insufficient physical activity, unhealthy diets and substance abuse can be reduced by implementing interventions suggested by the Gobal Action Plan [5] and by formulating extra interventions using information from surveys conducted to determine the characteristics of the population vulnerable to these risks.

Alcohol abuse, insufficient physical activity, unhealthy diets and substance abuse are categorical variables. Insufficient physical activity can be categorical or continuous depending on how it is measured. If a categorical risk factor has two or more categories logistic or multinomial regression model [11] will help in determining the variables which influence it [11,12]. If a risk factor for NCDs is continuous and the covariates are all categorical, analysis of variance models will help in determining the variables which influence it [13]. If the outcome variable is continuous but the covariates are a mixture of categorical and continuous variables, analysis of covariance can be used to determine the factors which influence it [13].

Compartmental and multistate models are ideal for investigating the progression of acute and chronic diseases. Multistate models like the proportional intensity model [14,15] can be used to determine the risk factors which influence the transition from one disease state to another. The knowledge of these risk factors will help in formulating strategies for preventing transitions from one stage to another. The proportional intensity model can be fitted by using functions from the **msm** library in R-4.0.0 software [14]. Compartmental models [16] can be used to investigate the effects of various levels of treatment uptake on mortality. The models can be fitted in R-4.0.0 using functions from the **deSolve** library.

Table 2
Data required for computing stratum sample sizes

Township (h)	Population (N_h)	Weight (W_h)	Prevalence (P_h)	Cost (C_h)	S_h
Nancholi	20,000	0.364	0.1	MWK5	0.30001
Nyambadwe	15,000	0.273	0.12	MWK10	0.32497
Ndirande	30,000	0.545	0.15	MWK2	0.35708

*1 USDollar = 744.92 Malawi Kwacha (MWK).

Additional risk factors of NCD can be determined by using a Cox proportional hazards model [17] where the response is time to onset of a disease. Since age is usually measured as a discrete quantity, a discrete equivalent [17] of the Cox model can be fitted. The model can be fitted using the **glm** function in R-4.0.0 software.

3.5. Optimal use of available funding when conducting surveys

African countries have different levels of income. An analysis of the state of health in Africa in 2018 showed that there are differences in the state of health in some areas due to differences in countries' levels of income. Although this is so, there is a statistical approach for estimating a sample size for surveys which uses available resources optimally [18]. The sample size is estimated either to minimize variance, V of some principle survey outcome variable whilst constraining on survey cost or to minimize survey cost whilst constraining on variance, V of a principle survey outcome variable. Both approaches use one of two cost functions. An example below uses sample size formulae from Barnett [18].

Example 1: Estimating the sample sizes for a survey given a fixed amount of money.

A research agency has been hired to conduct a survey to determine the risk factors of a certain disease in three townships in Blantyre namely Nancholi, Nyambadwe and Ndirande. The agency is given MWK200,000* for the exercise. The overhead cost, c_0, of conducting the survey is MWK90,000. The populations of Nancholi, Nyambadwe and Ndirande are approximately 20000, 15000 and 30000 respectively. The costs of collecting data from a household in Nancholi, Nyambadwe and Ndirande are MWK5, MWK10 and MWK2 respectively. The proportions of people with disease in Nancholi, Nyambadwe and Ndirande are 0.10, 0.12 and 0.15 respectively. Calculate the sample size for each township and the total sample size for the survey. The data are presented in the table below. The cost function [15] is $C = c_0 + \sum_{h=1}^{k} n_h c_h$ where C is the total cost of the survey, c_0 is fixed overhead cost, n_h is

the sample size for the stratum h and c_h is the cost of collecting data in the stratum h.

Using the appropriate formula from pages 118–119 of Barnett (1991), the sample sizes are $n_1 = 6,711$ households for Nancholi, $n_2 = 3,855$ households for Nyambadwe and $n_3 = 18,945$ households for Ndirande. The optimal total sample size will be 29,511 households.

3.6. Universal access to medical help

Universal access to health care can be achieved by using interventions for increasing health facility utilization rate. These interventions should be based on a design developed by using information on health-seeking behavior gathered from population surveys conducted in an area. The reasons why some sick people go to a health facility or not can be discovered by conducting a survey to investigate the extent of this habit and associated risk factors. The outcome variable in such surveys should be health-seeking behavior. One way this can be phrased on a questionnaire is: "What did you in the last three months when you or a member of your household was ill?" Possible answers may be "Went to hospital", "Went to a herbalist", "Bought medicine from a grocery store", "Did nothing". Apparently, in this case health-seeking behavior is a polytomous response variable. Potential explanatory variables may include distance to health facility, health personnel's attitude to patients, quality of care provided at health facilities and patients' social, economic and demographic characteristics. Multinomial regression [11] of the health-seeking behavior on explanatory variables will result in regression coefficients for the explanatory variables and corresponding p-values. Regular mobile clinics and hospital trains can be seen as good examples of an attemp to provide medical care to remote populations.

3.7. Use of statistics and machine learning methods in diagnosis of illness

Clinicians diagnose illness with some assistance from medical laboratory technicians. Linear discriminant analysis, logistic regression, classification and regres-

sion trees and random forests [19] can help doctors diagnose illness. Artificial intelligence system (AI) based on modern analysis tools and found correlations is seen as a very accurate (and sometimes more accurate than doctors) tool in diagnosing breast cancer [20]. Obviously, for better performance, the training and validation data sets for such an AI system should be very big. The following is an example of the use of logistic regression [19] and regression tree for classifying breast tumor patients into two disease states.

Example 2: Logistic regression function used to classify breast cancer patients.

The Wisconsin Breast Cancer dataset with 11 variables is available for free on the web (archive.ics.uci.edu /ml/index.php). Six hundred and ninety-nine fine-needle aspirate samples were collected from women. Four hundred and fifty-eight are benign and 241 are malignant. Using R-4.0.0 software we divided the data into a training dataset which has 349 (222 benign and 127 malignant) observations and a validation dataset which has 350 (236 benign and 114 malignant) observations. A logistic regression model was fitted and validated in R-4.0.0. The regression model correctly classified 225 benign out of 227 benign cases (97.4% accuracy, 2.6% error rate and 107 malignant out of 112 malignant cases (95.6% accuracy, 4.4% error rate).

Example 3: Classification of the breast cancer patients using a regression tree.

Using the same breast cancer data, a regression tree was fitted and validation was performed. Two hundred and twenty-five out of 231 benign cases were correctly classified (97.4% accuracy), and 107 out of 119 malignant cases were correctly classified (89.9% accuracy, 10.1% error rate). The algorithm for diagnosis is as follows: If sizeUnif < 3, the tumor is benign. If the sizeUnif $\geqslant 4$, the tumour is malignant. If sizeUnif < 4 and bareNucl < 3, the tumour is benign. If sizeUnif < 4 and bareNucl is at least 4, the tumour is malignant.

3.8. Improvement of service delivery using evidence-based medicine methods

In the above paragraphs approaches for using statistics in improving the health and well-being of citizens of African countries have been described. Good quality medical care is an essential base requirement in improving health and wellbeing of people. There have been many advances in medicine in this and past centuries.

Quality of care given to patients will be high and treatment outcomes will improve when clinicians are able to use the best practice for treating patients, The best treatment to give to a patient is determined by using evidence from systematic reviews and meta-analysis [21]. These lead to evidence-based approaches for selecting the best treatment for treating a disease. Clinicians should be encouraged to offer good medical care which is supported by evidence from meta-analysis.

3.9. Timely production of statistical information for decision making

One of the reasons for the insufficient use of statistical information in decision making is that processing of health data can take too long, leading to decisions being made before a statistical report is available. This is especially relevant in case of unexpected outbreaks of a disease. If an outbreak has been detected, data collected if processed quickly will yield information which can be applied to control the epidemic. Data processing includes data entry, cleaning and analysis. Data entry and cleaning take a long time because of transportation of questionnaires and data entry forms and subsequent cleaning after data entry. Data entry and cleaning can be shortened by using Open Data Kit (ODK) software. Such software often includes automated consistency checks on the data entry.

3.10. Enhanced use of statistical information in decision making

Most leaders make decisions which are not backed by empirical data because either the necessary statistical information is not available, they do not understand statistical information provided to them or they opt for political correctness whilst defying empirical evidence from data. To make sure leaders use statistical information to make decisions, information should be condensed and made easy to understand. One way of compressing statistical information is to use indicators in reporting. Another way of making statistical information readable is to organize it into fact sheets [1].

DHS2 software has many indicators for monitoring. Beyond this, it is sometimes necessary to develop customized indicators which best serve the situation at hand. The identification and definition of such indicators is preferably done in a collaboration between the developers (statisticians) and the users (for example decision makers).

4. Discussion

In this paper various approaches have been discussed in using statistics to improving health and wellbeing in Africa in the context of SDG3.[1] The proposed approaches apply to all African countries regardless of the WHO region they belong to. The statistical approaches for achieving health and well being in Africa in the context of SDG3 are proposed with the understanding that countries will continue implementing the policies and measures which produced the gains which were described by the WHO team which assessed the state of health in the Africa Region in 2018 [3]. This paper builds on the successes and suggests methods for addressing the shortfalls that were pointed out by WHO in their assessment report.

Statistics is just a tool which needs good quality data to produce good quality information. Via applying recent IT applications (computers and software) data are more easily collected and made accessible for cleaning and analysis. In countries where DHIS2 is deployed, patient data are entered in a database on the same day or a day later. Health data collected in this way is almost real time data. Consequently, such data can be used anytime to detect outbreaks and trends. This implies the possibility to have timely and relevant information available for action from health data. Such data can also reveal areas that need in-depth investigation using surveys. Since computerized health data are bound to be timely, surveys conducted using information from such data will also be timely. Apparently, computerization of the collection of health data is necessary for achieving good health and well-being for all.

Vital statistics data are one of the types of health data collected by health systems in many developed countries. Such data are extracted from vital registration systems which many African countries do not have. Vital statistics data provides information on births, deaths, and life expectancy. Since populations are dynamic, they have to be monitored regularly in order to detect changes in morbidity, births, deaths and life expectancy with time. Attempts can then be made to improve life expectancy and reduce deaths. It is essential for African countries to establish vital registration systems as these will provide relevant data for improving the health and wellbeing of their people.

Disease registers are another source of health data suggested in this paper. These registers record data on people who have disease and socio-demographic characteristics. As such with time, newly recorded cases are incident cases in the population. The totality of new and old cases is the diseased subpopulation. When these two are known, the trend of incidence with time is known. The volume of medication and the disease-specific mortality rates are also known together with the socio-demographic characteristics. In the case of NCDs, These data can provide a clear picture of NCD incidence and burden. Consequently, it can be easy to plan for treatment, prevention and palliation because the extent of the NCD burden will be known from disease register data.

Surveys are another source of health data. Routine health data will primarily be essential in detecting outbreaks, trends and other health problems in a catchment population. When further investigation is necessary, surveys have to be conducted. The use of mobile tablets in surveys will shorten the time taken to enter data and so is recommended. In fact, mobile tablets helped greatly in saving lives in Guinea Bissau as data collection and entry tools through real time monitoring [22].

To achieve SGD3 using the approaches in this paper information on risk factors of health-seeking behavior, mortality, morbidity and factors that influence the four major risk factors of NCDs is essential. This information cannot be obtained by analyzing routine health data alone. Extra data from specific and dedicated surveys is required. When survey information has been used to devise interventions, it will be necessary to evaluate the interventions after some time in order to check if they have been effective. Therefore to achieve good health and well-being for all, health and survey data must provide the information for designing interventions and these should be followed by monitoring on an on-going basis.

Diagnosis is one of the exercises that generates health data. This is often performed by doctors. Statistical and machine learning methods can help reduce doctor's workload and false positives and negatives.

Evidence-based methods (EBM) allow for the use of best practices in treating patients to improve treatment outcome. In many cases, without EBM physicians stick

---

[1] In this paper, African countries are defined as all countries in Africa and its associated territories. Note that this is different from the way WHO groups countries. WHO groups countries of the world into regions. The WHO African region does not include Egypt, Morocco, Djibouti, Sudan, Somalia, Tunisia and Libya. These seven countries are in the WHO Eastern Mediterranean region. The assessment of the health status of the WHO African region did not include these seven countries. Nevertheless, since the data that were analysed to produce the WHO Africa region health status assessment report came from the majority of African countries we can safely generalize the results to the seven excluded countries.

to routine treatments which are not always the best for patients. About this, there is a well known example of the treatment of myocardial infarction using beta-blockers [21]. Because of the current ubiquity of internet in the world and the availability of smart phones and mobile tablets, systematic reviews and meta-analyses that inform EBM can be conducted almost anywhere in Africa. The implementation of EBM in patient handling is a conscious decision which has to be taken by those who treat patients and their bosses.

Multistate models and survival analysis techniques are helpful in the identification of factors that significantly influence time to the occurrence of an event or a health state in general. These methods can be applied on routine health state data as well as well as on data from vital registration systems, disease registers and surveys. Multistate models and survival analysis techniques can help in the reduction of the burden of NCDs and their principle risk factors.

Increased use of up-to-date health state information will help in addressing health problems at the right time. It allows decisions to be made using latest information gathered from the health district. Consequently, evaluation of interventions will give correct feedback of the effect of interventions.[2]

5. Conclusion

Statistics (data, indicators) and statistical methods (measures, tools) can be used to help to achieve the target for SDG3 by 2030. Computerization of the collection of health data is necessary for this. There are several computer software applications which can be used for capturing health data in health facilities. Data captured by such software applications can be used to detect significant health problems or issues that need redress. Deep insight into such health problems can be

obtained from data from surveys conducted in catchment areas of health facilities. Information obtained from such surveys can be used to devise interventions for reducing or eliminating health problems. After some time, evaluations of interventions can shed more light on their effectiveness. It can be helpful to coordinate monitoring of population health and well-being and implementation of activities for achieving change from the district health office.

Societies are dynamic, therefore, attempting to improve the status of health in a population needs to be an on-going process, in which the data-driven monitoring of health status and implementation of remedial activities aimed at addressing gaps needs to take a central position in supporting decision making. Apparently a culture of evidence-based decision making is needed to ground such an approach.

Acknowledgments

References

[1] Radermacher WJ. Official Statistics 4.0: Verified Facts for People in the 21st Century. Cham: Springer; 2019.

[2] The Sustainable Development Agenda. New York: United Nations (UN); 2017 [cited 2020 April 2]; Available from: https://www.un.org/sustainabledevelopment/development-agenda/.

[3] WHO Regional Regional Office for Africa. The state of health in the WHO African Region: An analysis of the status of health, health services and health systems in the context of the Sustainable Development Goals. Brazzaville: World Health Organisation; 2018.

[4] Angell MKJ, Relman AS. Looking back on the millenium in medicine[editorial]. New England Journal of Medicine. 2000; 342(6): 42–9.

[5] World Health Organisation (WHO). Global action plan for the prevention and control of noncommunicable diseases 2013–2020. Geneva, Switzerland: World Health Organisation; 2013.

[6] Stieffel MC, Perla RJ, Zell BL. A healthy bottom line: healthy life expectancy as an outcome measure for health improvement efforts. The Milbank Quarterly. 2010; 88(1): 30–53.

[7] World Health Organisation. WHO methods for life expectancy and healthy life expectancy. Geneva: World Health Organisation, 2014.

[8] Elandt-Johnson RC, Johnson NL. Survival Models and Data Analysis. New York: Wiley InterScience; 1980.

[9] Misiri H. Estimation of incidence from cross-sectional data, risk factors for early sexual debut and cancer trends in a population where HIV is prevalent: the case of Malawi, PhD Dissertation, University of Oslo, 2016.

[10] Institute for Health Metrics and Evaluation (IHME). Findings from the Global Burden of Disease Study 2017. Seattle, 2018.

---

[2]All the approaches suggested in this paper can help African countries to achieve good health and well-being by 2030 if everybody in the catchment population of a health district has access to health care. This is important. If some of the members of the population have no access to health care, SDG3 will not be achievable. It will just be a pipedream. Sick people do not present at a hospital because it is there. They go there because they are willing to be attended to in a health facility. Some people prefer to consult traditional doctors. Others prefer to use medicine bought from grocery stores. In fact, there are many factors that affect health-seeking behavior. Statistical methods will be handy in helping identify factors of health-seeking behavior. With interventions devised from relevant information, it is still possible for everybody in the population to have access to health care.

[11] Agresti A. Categorical Data Analysis. New York: John Wiley; 1990.

[12] Hosmer DW, Lemeshow S. Applied Logistic Regression. New York: Wiley; 1989.

[13] Neter J, Kutner M, Nachtsheim CJ, Wasserman W. Applied Linear Statistical Models. Boston: IRWIN; 1996.

[14] Cook RJ, Lawless J. Multistate Models for the Analysis of Life History Data. Boca Raton: Taylor and Francis; 2018.

[15] Keeling MJ, Rohani P. Modeling Infectious Diseases In Humans and Animals. Princeton: Princeton University Press; 2008.

[16] Brauer F, Castillo-Chavez C, Feng Z. Mathematical Models in Epidemiology. New York: Springer; 2019.

[17] Fahrmeir L, Tutz G. Multivariate statistical modelling based on generalized linear models. New York: Springer 1994.

[18] Barnett V. Sample Survey: Principles and Methods. Kent: Edward Arnold; 1991.

[19] Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning Data Mining, Inference, and Prediction. New York: Springer; 2008.

[20] McKinney SM, Shetty S. International evaluation of an AI system for breast cancer screening. Nature. 2020; 577: 89–94. doi: 10.1038/s41586-019-1799-6.

[21] Egger M, Smith GD, Altman DG. Systematic Reviews in Health Care. London: BMJ Publishing Group; 2001.

[22] Can Data Save Lives? Digitizing The Malaria Response In Guinea Bissau. Guinea Bissau, UNDP; 2019 [cited 2020 April 2]; Available from: https://stories.undp.org/is-the-digital-divide-hampering-the-malaria-response-in-africa.