Kathleen L. Strong, Ruth Bonita

Surveillance for Information and Policy Team, Chronic Disease and Health Promotion Department, Noncommunicable diseases and Mental Health Cluster, World Health Organization, Geneva

Investing in surveillance: a fundamental tool of public health

Summary

Objectives: The WHO Global InfoBase assembles country-level chronic disease risk factor prevalence data from WHO's member states.

Methods: The focus of this report is recent, nationally representative data. The risk factors of choice are those that make the greatest contribution to mortality and morbidity from chronic disease, can be changed through primary intervention and are easily measured in populations.

Results: Eight risk factors fit these criteria. They are: tobacco and alcohol use, patterns of physical inactivity, low fruit/vegetable intake, obesity, blood pressure, cholesterol and diabetes. Important to the data collection is the need to display prevalence data for these eight risk factors by age group(s) and sex and with some measure of the uncertainty of the estimate.

Conclusions: This tool can be used by countries to evaluate the quality of the data that they have for chronic disease surveillance. The aim is to improve risk factor data quality and to standardize data, either through common survey instruments or by using existing country data to model risk factor estimates. These "harmonized" estimates will allow for comparisons over time and between countries.

Keywords: Surveillance – Chronic disease – Risk factors – InfoBase.

Valid health statistics are essential for planning and implementing health policy in all settings. Information on chronic disease (noncommunicable disease [NCD]) risk factors is important for predicting the future burden of chronic disease in populations and also for identifying potential interventions to reduce the future burden. The World Health Report 2002: reducing risks, promoting healthy life highlights the importance of risk factors as indicators of future health status (WHO 2002). Even in the poorest countries, chronic disease risks such as raised blood pressure, cholesterol and tobacco use are accounting for increasing levels of chronic diseases and untimely deaths. In fact, five of the top 10 risks worldwide are chronic disease risk factors (WHO 2002). These are raised blood pressure, tobacco use, alcohol consumption, cholesterol and obesity/overweight and they exist in developing countries along with other important risk factors, including underweight, unsafe sex, unsafe water, iron and zinc deficiency (Tab. 1).

Unfortunately, country-level data on common, measurable chronic disease risk factors are sparse. Few countries have the resources, infrastructure, or on-going political commitment to collect this type of information in a sustainable manner. While some developed countries have regular national health surveys that include selected chronic disease risk factors, for many, the information comes from small, costly one-off surveys. This paper introduces a new tool for assessing chronic disease risk factor information, the WHO Global InfoBase. The InfoBase assembles current country-level data on chronic disease risk factors. The first use of this tool has been for the production of the SuRF report 1: *Surveillance of risk factors related to noncommunicable disease: current status of global data* (Strong & Bonita 2003). This report gives preference to data that is nationally representative, with valid sampling frames and explicit definitions for indicators of risk factors. It can be used as a basis for discussion on how to improve and/or augment this information so that it can provide the impetus for better health policy.

The WHO Global InfoBase

The WHO Global InfoBase is a timely tool for collecting and displaying country-level chronic disease risk factor data. The InfoBase brings together current country-level data on important chronic disease risk factors for WHO's 192 member states. Identifying country-level data and assessing its validity is the first step in developing better quality chronic disease data collections. Country-level risk

Soz.- Präventivmed. 49 (2004) 269–275
0303-8408/04/040269–7
DOI 10.1007/s00038-004-3076-0

Table 1 Leading 10 selected risk factors as selected causes of disease burden (WHO 2002). (bold: indicates major chronic disease risk factors)

	Developing countries		**Developed countries**
	High mortality	**Low mortality**	
1.	Underweight	**Alcohol**	**Tobacco use**
2.	Unsafe sex	Underweight	**Blood pressure**
3.	Unsafe water	**Blood pressure**	**Alcohol**
4.	Indoor smoke	**Tobacco use**	**Cholesterol**
5.	Zinc deficiency	**High body mass index**	**High body mass index**
6.	Iron deficiency	**Cholesterol**	**Low fruit and vegetable intake**
7.	Vitamin A deficiency	Iron deficiency	**Physical inactivity**
8.	**Blood pressure**	Low fruit and vegetable intake	Illicit drugs
9.	**Tobacco use**	Indoor smoke	Underweight
10.	**Cholesterol**	Unsafe water	Iron deficiency

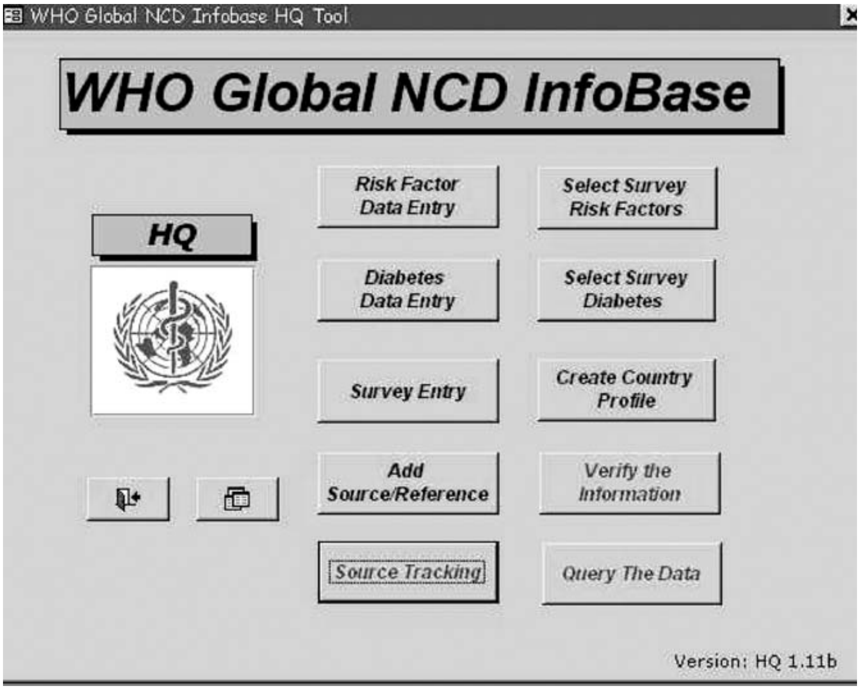
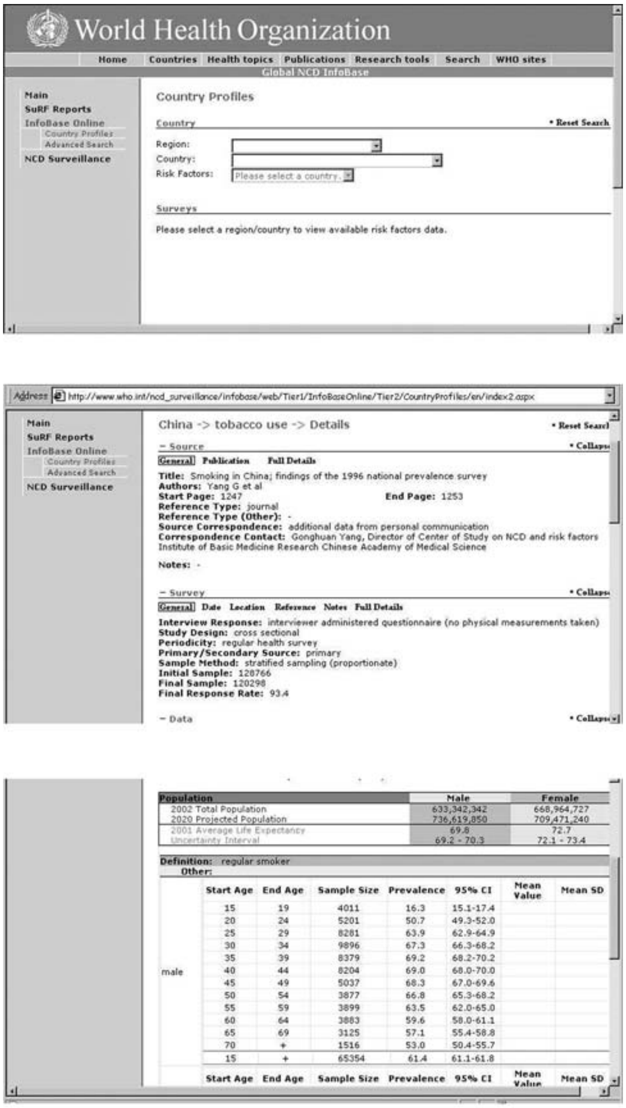


Figure 1 Source, survey and data entry tool for the InfoBase. The InfoBase has set templates for source, survey and data entry as well as a protocol for verifying, selecting and displaying the risk factor data

factor prevalence profiles help to identify a country's strengths in risk factor data collection and its gaps and deficiencies. The ultimate aim is to support and encourage the sustainable collection of good quality risk factor data and to promote the establishment of surveillance systems for chronic diseases and their risk factors as an alternative to costly "one time only" surveys.

The key feature of the InfoBase is that all metadata are available without the researcher having to refer to the original source publication. The InfoBase has templates for the entry of source, survey and risk factor data (Fig. 1) and ensures that data cannot be entered without full information about the source, contact person/Institution and survey information. This is important because risk factor data is most useful when presented with information about how the data were collected, including sampling frame, interview method and intended population. In addition, individuals interested in the chronic disease risk factor data often lack access to necessary journals and/or other data sources. Thus, the InfoBase is designed to be a "one stop" resource for these data needs.

The InfoBase collects sources and surveys of varying standards. It works on the premise that all data provide some level of information, provided that the limitations of the data are understood and taken into account. However, InfoBase tool allows the user to preferential select the sources/surveys that they choose from the data search form (Fig. 2) and display this data as a country profile (Fig. 3). The



Country Profiles

Country

Surveys

Please select a region/country to view available risk factors data.

http://www.who.int/ncd_surveillance/infobase/web/Tier1/InfoBaseOnline/Tier2/CountryProfiles/en/index2.aspx

China -> tobacco use -> Details

– Source

Title: Smoking in China; findings of the 1996 national prevalence survey
Authors: Yang G et al
Start Page: 1247 **End Page:** 1253
Reference Type: journal
Reference Type (Other): -
Source Correspondence: additional data from personal communication
Correspondence Contact: Gonghuan Yang, Director of Center of Study on NCD and risk factors Institute of Basic Medicine Research Chinese Academy of Medical Science

Notes: -

– Survey

Interview Response: interviewer administered questionnaire (no physical measurements taken)
Study Design: cross sectional
Periodicity: regular health survey
Primary/Secondary Source: primary
Sample Method: stratified sampling (proportionate)
Initial Sample: 128766
Final Sample: 120298
Final Response Rate: 93.4

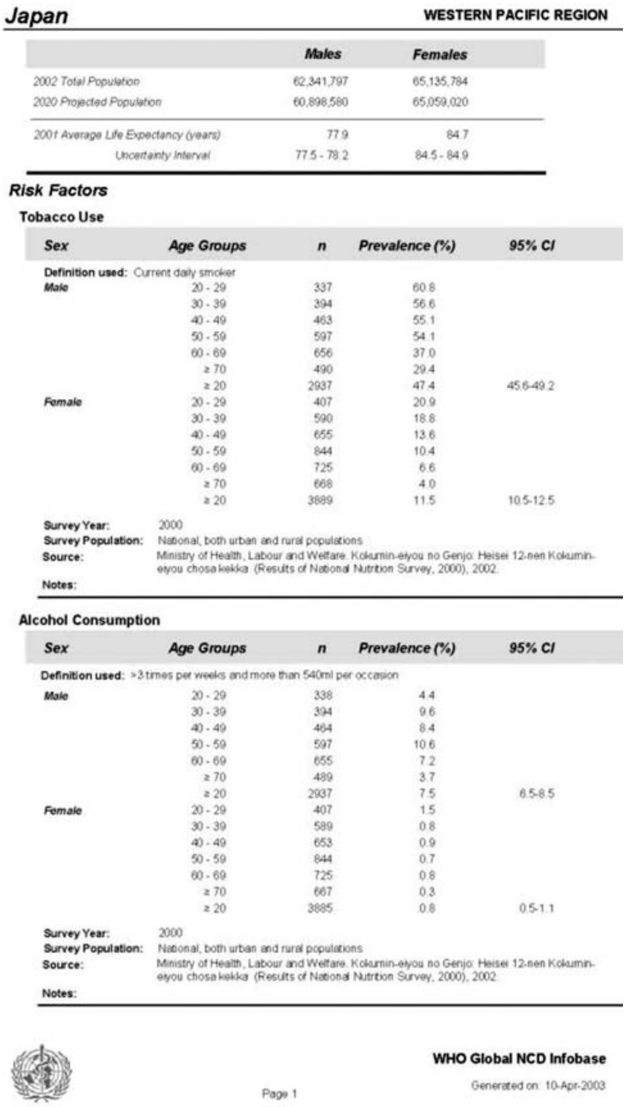
– Data

Population	Male	Female
2002 Total Population	633,342,342	668,964,727
2020 Projected Population	736,619,850	709,471,240
2001 Average Life Expectancy	69.8	72.7
Uncertainty Interval	69.2 - 70.3	72.1 - 73.4

Definition: regular smoker
Other:

	Start Age	End Age	Sample Size	Prevalence	95% CI	Mean Value	Mean SD
male	15	19	4011	16.3	15.1-17.4		
	20	24	5201	50.7	49.3-52.0		
	25	29	8281	63.9	62.9-64.9		
	30	34	9896	67.3	66.3-68.2		
	35	39	8379	69.2	68.2-70.2		
	40	44	8204	69.0	68.0-70.0		
	45	49	5037	68.3	67.0-69.6		
	50	54	3877	66.8	65.3-68.2		
	55	59	3899	63.5	62.0-65.0		
	60	64	3883	59.6	58.0-61.1		
	65	69	3125	57.1	55.4-58.8		
	70	+	1516	53.0	50.4-55.7		
	15	+	65354	61.4	61.1-61.8		
	Start Age	End Age	Sample Size	Prevalence	95% CI	Mean Value	Mean SD

Figure 2 Examples of the InfoBase on-line tool now available at www.who.int/ncd_surveillance/infobase/en. Note that the tool connects the data, with its source and survey information



Japan WESTERN PACIFIC REGION

	Males	*Females*
2002 Total Population	62,341,797	65,135,784
2020 Projected Population	60,898,580	65,059,020
2001 Average Life Expectancy (years)	77.9	84.7
Uncertainty Interval	77.5 - 78.2	84.5 - 84.9

Risk Factors

Tobacco Use

Definition used: Current daily smoker

Sex	*Age Groups*	*n*	*Prevalence (%)*	*95% CI*
Male	20 - 29	337	60.8	
	30 - 39	394	56.6	
	40 - 49	463	55.1	
	50 - 59	597	54.1	
	60 - 69	656	37.0	
	≥ 70	490	29.4	
	≥ 20	2937	47.4	45.6-49.2
Female	20 - 29	407	20.9	
	30 - 39	590	18.8	
	40 - 49	655	13.6	
	50 - 59	844	10.4	
	60 - 69	725	6.6	
	≥ 70	668	4.0	
	≥ 20	3889	11.5	10.5-12.5

Survey Year: 2000
Survey Population: National, both urban and rural populations
Source: Ministry of Health, Labour and Welfare. Kokumin-eiyou no Genjo: Heisei 12-nen Kokumin-eiyou chosa kekka (Results of National Nutrition Survey, 2000), 2002.
Notes:

Alcohol Consumption

Definition used: >3 times per weeks and more than 540ml per occasion

Sex	*Age Groups*	*n*	*Prevalence (%)*	*95% CI*
Male	20 - 29	338	4.4	
	30 - 39	394	9.6	
	40 - 49	464	8.4	
	50 - 59	597	10.6	
	60 - 69	655	7.2	
	≥ 70	489	3.7	
	≥ 20	2937	7.5	6.5-8.5
Female	20 - 29	407	1.5	
	30 - 39	589	0.8	
	40 - 49	653	0.9	
	50 - 59	844	0.7	
	60 - 69	725	0.8	
	≥ 70	667	0.3	
	≥ 20	3885	0.8	0.5-1.1

Survey Year: 2000
Survey Population: National, both urban and rural populations
Source: Ministry of Health, Labour and Welfare. Kokumin-eiyou no Genjo: Heisei 12-nen Kokumin-eiyou chosa kekka (Results of National Nutrition Survey, 2000), 2002.
Notes:

WHO Global NCD Infobase
Page 1
Generated on: 10-Apr-2003

Figure 3 Example of country chronic disease risk factor profile produced by the WHO Global InfoBase. For more complete information for WHO member states contact ncdsurf@who.int for a free copy of the SuRF report 1

SuRF report 1 published country profiles of the most recent, most nationally representative data from surveys with valid sampling strategies and specified measurement criteria (Strong & Bonita 2003). The InfoBase search and data display tools are now available to all users on the Internet as the InfoBase on-line tool (www.who.int/ncd_surveillance/infobase/en). The on-line tool includes a basic search tool for searching by WHO region, country and risk factor. The more advanced search function allows the user to search by survey geographic area, publication year and risk factor definition.

Much effort has gone into deciding which type of information is most useful for surveillance of chronic disease risk factors. The InfoBase collects data that is strictly relevant to chronic disease outcomes; i.e., mortality and morbidity from NCDs. The risk factors of choice are those that:

- make the greatest contribution to mortality and morbidity from chronic disease;
- can be changed through primary intervention; and
- are easily measured in populations.

Eight risk factors that relate to cardiovascular disease and other NCDs fit these criteria. These are tobacco and alcohol use, patterns of physical inactivity, low fruit/vegetable intake, obesity (measured by BMI [body mass index]), blood pressure, cholesterol; and diabetes (measured by blood glucose). Table 2 shows the chronic diseases causally associated with these risks to health.

Table 2 Chronic disease risk factors contribute to a range of related chronic conditions as shown below

	Cardiovascular disease*	Diabetes	Cancer	Chronic-obstructive pulmonary disease
Risk factor				
Tobacco use	✓		✓	✓
Alcohol use	✓	✓	✓	
Physical inactivity	✓	✓	✓	
Nutrition	✓	✓	✓	
Obesity	✓	✓	✓	✓
Raised blood pressure	✓	✓		
Dietary fat/blood lipids	✓	✓	✓	
Blood glucose	✓	✓	✓	

* Including heart diseases, stroke, hypertension

Of principal importance to the data collection is the need to display prevalence data for these eight risk factors by age group(s) and sex and with some measure of the uncertainty of the estimate for each member state.

Existing data sources

The availability of risk factor data varies from country to country. There is the additional issue of risk factor data collection sustainability where existing collections are held. For the InfoBase, data have been obtained primarily from published sources. The InfoBase now contains data from 2 600 sources for 176 countries. For some countries, available data are restricted to small, one-off surveys published in academic journals. WHO regions with many developed countries have a wider variety of survey instruments available to them.

Gaps and deficiencies in data

There are several limitations to current existing chronic disease risk factor data collections. It is difficult to compare survey data across countries. Even within a country, when data from different years are available, this data may not be comparable because of different survey instruments, resulting in different measurement methods and different criteria for clinical events. These problems can be solved to some extent by using standardized survey instruments.

Valid, reliable and comparable data are essential for research and population health surveillance. Valid data are produced by using the best possible sampling frame, the most representative sample, measurement techniques free of systematic bias and common training of field staff. Reliability of data refers to the level of confidence that can be placed in the estimate obtained from the survey. For surveillance of risk factors at the level of populations, more emphasis is placed on measurement accuracy (validity) than reliability. If the measurement and sampling is without systematic bias and of adequate size, then the resulting average estimate should approximate the true population estimate with a known level of uncertainty.

Even valid and reasonably reliable data may not be comparable within and between countries. Comparability of data is based on the use of standard measurement methods and also a common understanding in all survey settings of what is meant by survey questions. The need for comparable data is being addressed by WHO with its recommended survey instruments. A common set of indicators as well as standardized measurement methods for those indicators have been developed and the survey instruments holding these indicators are presented below. A global move towards data harmonization would require collection of data on a small set of core indicators that have measurement validity and are related directly to health outcomes, i.e., mortality or morbidity from NCD.

WHO's surveillance tools

Surveillance systems require standardized and sustainable collection of data over time as well as the flexibility to meet country needs. For surveillance to be useful to policy makers and health promotion specialists, it needs to be sustainable over time to provide information on trends. Changes in population risk factor prevalence, as discovered through trend data, may be predictive of future chronic disease epidemics. Of course, there will be a time lag between changes in chronic disease risk factors in a population and disease occurrence. For example, an increase in tobacco use for young women in developing countries will eventually lead to an increase in chronic health problems, including lung cancer, as the cohort ages. The World Health Organization has developed three survey instruments that promote valid and comparable chronic disease surveillance trend data, the STEPwise approach to Surveillance of chronic disease risk factors (STEPS), the Global Youth Tobacco Survey (GYTS) and the World Health Survey (WHS).

STEPwise approach to surveillance of chronic disease risk factors (STEPS)

STEPS is a survey package that includes not only the survey instrument but also complete information on how to train interviewers, appropriate sampling frames, sample sizes, data analysis and reporting. This approach to chronic disease risk factor surveillance focuses on building country capacity in low and middle income countries to collect on a periodic basis, small amounts of high quality risk factor data. The survey instrument is divided into three steps that relate directly to the complexity of risk factor measurement. Step 1 collects questionnaire-based information about tobacco use, alcohol consumption and nutrition and physical activity. Step 2 moves to objective standardized physical measurements to collect data on blood pressure, height and weight. Step 3 carries physical measurements even further with the collection of blood samples for measurement of lipids and glucose status (for diabetes). Countries can start at Step 1 and add steps sequentially as their capacity improves (Bonita et al. 2001).

The Global Youth Tobacco Survey (GYTS)

The GYTS, in collaboration with the CDC ([Centers for Disease Control and Prevention] available at www.cdc.gov/tobacco/global/GYTS.htm), is now the largest global surveillance system for any major public health risk. This questionnaire-based survey on tobacco use and attitudes towards tobacco focuses on adolescents aged 13 to 15 years. It is operational in 150 countries and questionnaires have been completed on over 1 million children in the target age range in randomly selected schools (Warren et al. 2000). The GYTS sampling frame and questionnaire is now being expanded to surveillance of other major risks that can be effectively measured in the school setting. The resulting survey instrument, the Global School Health Survey (GSHS), will follow the Stepwise approach to Surveillance of chronic disease risk factors. It will begin with risk factors that are easy to measure (i.e., by questionnaire) and move to physical and biochemical measurements as country capacity improves (http://www5.who. int/tobacco).

The World Health Survey (WHS)

The WHS is a measurement instrument that is collecting health information, including chronic disease risk factor information, from WHO member states. The survey is designed to provide up-to-date, country-level data to measure population health levels, health expenditure, coverage of health interventions and health system responsiveness. The WHS risk factor module and the STEPS survey instrument share a common set of indicators at Step 1 (health behaviours) as well as standardized measurement methods for those indicators. (www.who.int/WHS).

Data from the above mentioned surveys as well as from available published information and government reports are held in the WHO Global InfoBase.

Estimates of risk factor prevalence

Standardizing survey instruments and collecting data on a small set of core indicators may go a long way to providing better data in the future, but valid, existing data should not be ignored. The InfoBase collects all surveys and sources that have data on chronic disease risk factors without excluding any because of concerns about data quality. The reasoning behind this collection strategy is that all data can be used at some level, even if systematic biases occur in the data collection or the sample population is not fully representative of a larger population. None of the data that currently exists are directly comparable across countries. However, bringing the available data together in one database is the first step towards building better quality chronic disease risk factor data by displaying the country-level data that currently exists. The second step is to use this data to develop comparable estimates of national prevalence for each risk factor and member state. The development of comparable country-level estimates is well under way by the InfoBase team, using the InfoBase as the primary resource. These estimates along with an update of available country-specific data will form the basis of the next surveillance of risk factors report.

The estimation process involves first standardizing the age groups used to report the risk factor prevalence data. The eight risk factors of interest are (with the exception of fruit/vegetable intake) strongly associated with age and thus the prevalence values increase or decrease depending on the starting and ending age used for the survey. For example, tobacco use in males decreases with age, so starting with age 35 years and ending with age 64 years will give a lower prevalence of tobacco use than starting at age 15 years and ending with age 64 years. The age risk factor relationships can be used in regression models to extrapolate values for ages not surveyed and then to produce these estimates in standard age groupings. For some countries with national health statistics reporting systems and measured risk factor data, producing standard age groups solves the major data comparability issue. In these cases, the data in standard age groupings become the base line prevalence values and together with trend information, can be adjusted to predict future risk factor prevalence. For other countries with no risk factor data or data not representative of their national population, a more complex model (with more covariates) must be used to produce comparable risk factor estimates.

For the data harmonization process, the age-specific rates obtained from country-level data sources have been used to develop the models. The use of country data sources to derive the estimates is a transparent process because the data are freely accessible in the InfoBase. The estimated prevalence values will also be held in the InfoBase, thereby maintaining the structure of the InfoBase as a "one-stop" source for all data users. This is a major improvement over previous WHO estimates, which, in the absence of such a relational database, relied on studies selected by experts which may have excluded much of the available data sources and which lacked transparency.

In turn, these estimates are the inputs needed to produce estimates of the burden attributable to these risk factors and the burden that could be avoided if these risk factor levels were lower. This approach to analysis takes risk factor data to a new level, providing direct information that can be linked to cost-effectiveness of health promotion programmes to produce better, country-specific health policy. The methodology for this type of analysis has been already used by WHO for the Comparative Risk Assessment exercise completed for the World Health Report 2002.

Country health priorities should, of course, drive the collection of larger, perhaps non-harmonized data collections, where feasible. These would focus on monitoring and evaluating specific health promotion programmes.

Conclusions

In summary, there is an urgent need for a review of the status and sustainability of risk factor data collections for countries. This will be particularly revealing where there are no data collections for a country or where only sub-national surveys exist.

A further concern is that the uses of risk factor data for health policy be well-defined and related not only to how governments currently use the data but also to how they could use data in the future; i.e., including the attributable and avoidable burden due to risk factors, as used in the World Health Report 2002. Understanding this type of analysis is important to promote effective public health interventions.

Given that there are enormous financial constraints on data collection, a distinction needs to be made between general, population-level data that is routinely collected for surveillance and more specific information that may be needed to evaluate and monitor health promotion activities or interventions. Both sources of information are needed but the uses of the data will drive the collection time-frames and survey instruments. Regardless of data uses, standard data definitions, age group designations and sampling frames are all useful tools for making already valid data comparable within and between countries. There is a strong need to expand the WHO surveillance tools to ensure global coverage. This, in turn, will help countries to use the analysis tools that have been developed to promote better health policy decisions.

The collection of chronic disease risk factor data is continuing in many countries and the InfoBase will need to reflect this new work as it becomes available. In fact, for those countries that have embarked on a STEPS survey, a method for direct transfer of aggregate, core risk factor data to the InfoBase has been developed. The development of the InfoBase is meant to support the process of data collection and to hold the collection in a central place. At the same time, the current collections need to be used to transform the data into a useful, comparative tool for advocacy, policy and research.

Acknowledgements

The authors would like to acknowledge the InfoBase team at WHO Geneva, including Jaclynn Lippe, Carina Marquez, Tomoko Ono, and Maria Filimonenko. The authors would also like to acknowledge Dr Andy Wielgosz from the WHO Collaborating Centre at the University of Ottawa and Dr Hongbo Liang from Health Canada. Their initial work on the CVD InfoBase was a precursor for the WHO Global InfoBase. The SuRF Report 1 can be obtained by contacting ncdsurf@who.int.

Zusammenfassung

Investition in Surveillance: ein wichtiges Instrument für Public Health

Fragestellung: Die WHO Global InfoBase sammelt Daten zur Prävalenz von Risikofaktoren chronischer Krankheiten der WHO-Mitgliederstaaten.

Methode: Es wird auf aktuelle, nationale und repräsentative Daten fokussiert. Die gewählten Risikofaktoren sind diejenigen, die am meisten zur Mortalität und Morbidität chronischer Krankheiten beitragen, die durch Primärintervention geändert werden können und in der Bevölkerung einfach zu messen sind.

Ergebnisse: Acht Risikofaktoren erfüllen diese Kriterien: Tabak- und Alkoholgenuss, körperliche Inaktivität, geringer Früchte- und Gemüsekonsum, Übergewicht, Blutdruck, Cholesterin und Diabetes. Die Prävalenzdaten für diese Risikofaktoren werden nach Altersgruppen und Geschlecht und mit Schätzfehlern präsentiert.

Schlussfolgerungen: Dieses Instrument kann dazu benutzt werden, um die Qualität der Daten für die Beobachtung chronischer Krankheiten nach Ländern zu überprüfen. Ziel ist es, die Datenqualität von Risikofaktoren zu verbessern und zu standardisieren, entweder indem gemeinsame Surveyinstrumente gebraucht werden oder indem existierende Länderdaten für die Datenschätzung modelliert werden. Derart „harmonisierte" Schätzungen erlauben Vergleiche über Zeit und zwischen Ländern.

Résumé

Investir dans la surveillance: un outil important pour la santé publique

Objectifs: L'InfoBase mondiale de maladies non transmissibles (MNT) de l'OMS rassemble des données nationales de prévalences des facteurs de risque des MNT des états membres.

Méthodes: Le données sont récentes et portent sur des échantillons nationaux représentatifs. Les facteurs de risque choisis sont ceux qui contribuent le plus à la mortalité et à la morbidité par maladies chroniques, qui peuvent être modifiées par la prévention primaire et facilement mesurées dans les populations.

Résultats: Huit facteurs de risque répondent à ces critères. Ce sont: l'utilisation de tabac et d'alcool, l'inactivité physique, la consommation insuffisante de fruits et de légumes, l'obésité, la tension artérielle, le cholestérol et le diabète. Les prévalences pour ces huit facteurs de risque sont présentées par groupes d'âge et de sexe avec une certaine incertitude de la mesure.

Conclusions: Cet outil peut être employé par les pays pour évaluer la qualité des données qu'ils ont pour la surveillance de MNT. Le but est d'améliorer la qualité des données de facteurs de risque et d'harmoniser les données, par les instruments communs d'enquête ou en employant des données existantes du pays pour modéliser les évaluations de facteur à risque. Ces évaluations „harmonisées"permettront des comparaisons dans le temps et entre les pays.

References

Bonita R, de Courten M, Dwyer T, Jamrozik K, Winkelmann R (2001). Surveillance of risk factors for noncommunicable diseases: the WHO STEPwise approach: summary. Geneva: WHO.

Strong K, Bonita R (2003). The SuRF Report I: surveillance of risk factors related to noncommunicable diseases: current status of global data. Geneva: WHO.

Warren C W, Riley L, Asma S, et al. (2000). Tobacco use by youth: a surveillance report from the Global Youth Tobacco Survey project. Bull World Health Organ *78*: 868–75.

WHO (2002). The World Health Report 2002. Reducing risks, promoting healthy life. Geneva: WHO.

Address for correspondence

Dr. Kathleen L. Strong
World Health Organization
Avenue Appia, 20
CH-1211 Geneva 27

Soz.- Präventivmed. 49 (2004) 269–275