A systematic review of Demographic and Health Surveys: data availability and utilization for research

Madeleine Short Fabic,a YoonJoung Choi*a & Sandra Birda

Objective To systematically review the public health literature to assess trends in the use of Demographic and Health Survey (DHS) data for research from 1984 to 2010 and to describe the relationship between data availability and data utilization.

Methods The MEASURE DHS web site was searched for information on all population-based surveys completed under the DHS project between 1984 and 2010. The information collected included the country, type of survey, survey design, fieldwork period and certain special features, such as inclusion of biomarkers. A search of PubMed was also conducted to identify peer-reviewed articles published during 2010 that analysed DHS data and included an English-language abstract. Trends in data availability and in the use of DHS data for research were assessed through descriptive, graphical and bivariate linear regression analyses.

Findings In total, 236 household surveys under the DHS project were completed across 84 countries during 2010. The number of surveys per year has remained constant, although the scope of the survey questions has expanded. The inclusion criteria were met by 1117 peer-reviewed publications. The number of publications has increased progressively over the last quarter century, with an average annual increment of 4.3 (95% confidence interval, CI: 3.2–5.3) publications. Trends in the number of peer-reviewed publications based on the use of DHS data were highly correlated with trends in funding for health by the Government of the United States of America and globally.

Conclusion Published peer-reviewed articles analysing DHS data, which have increased progressively in number over the last quarter century, have made a substantial contribution to the public health evidence base in developing countries.

Introduction

The Demographic and Health Surveys (DHS) project, funded primarily by the United States Agency for International Development (USAID) with support from other donors and host countries, has conducted over 230 nationally representative and internationally comparable household surveys in more than 80 countries since its inception in 1984. USAID has invested approximately 380 million United States dollars (US$) in the project, and each dollar has leveraged approximately US$ 0.33 in donor and host-country contributions. The first three of the DHS’s six phases were implemented between 1984 and 1997. Thereafter the project was folded into a family of USAID monitoring and evaluation projects and was renamed MEASURE DHS.

DHS evolved from World Fertility Surveys and Contraceptive Prevalence Surveys implemented in the 1970s and 1980s. Like their predecessors, the DHS originally collected comparable population-based data on fertility, contraception, maternal and child health and nutrition in developing countries. Today, DHS core questionnaires cover a wider range of population and health topics. Moreover, through optional modules – sets of questions on particular topics, including human immunodeficiency virus (HIV) and acquired immunodeficiency syndrome (AIDS), malaria, and domestic violence – countries choose to collect data on special health issues. Through collection of biomarker data, the prevalence of conditions like anaemia, HIV infection and parasitemia can also be measured (Fig. 1).

Each DHS is conducted by an in-country institution, typically the national statistics office. The institution receives technical assistance from the DHS project to ensure data quality and international comparability. Technical assistance varies in accordance with the data needs and existing capacities of the national institution and is typically provided across all phases: survey planning and design, data collection, data processing, and report writing and dissemination. Thus, DHS provide opportunities to strengthen the survey and statistical capacity of national institutions.

USAID financial support ensures that, with few exceptions, all DHS data files and final reports are freely available to the public through the MEASURE DHS web site (www.measuredhs.com). Data files follow a standardized format to maximize comparability across surveys and improve ease of use for analysis. Final reports provide data summaries, describe the survey design and contain a copy of the questionnaires. The web site also provides in-depth information on data collection and processing to facilitate analysis.

The transparency of DHS, including the free availability of data files and of descriptions of the survey methods and metrics, allows for external review and feedback beyond the internal quality assurance and control procedures in place.

A key aim of DHS is to provide quality data for policy development and programme planning, monitoring and evaluation. To build the policy and programmatic evidence base, DHS data must first be transformed into information, which must then be made accessible to decision-makers. Robust analyses of the data, beyond those published by the DHS project in its survey reports, are essential for transforming data into information. Research alone, however, will not influence policies and programmes. The results of such analyses must be accessed, understood and assimilated by policy-makers and programme managers.

While DHS are widely acknowledged as data sources among researchers, no study has methodically reviewed how
We obtained data from www.measuredhs.com on all completed population-based surveys conducted under the DHS project between 1984 and 2010. The information collected included country, survey type, a brief description of survey design, fieldwork period and special features (i.e. the inclusion of biomarker(s) and/or optional modules). We searched PubMed to identify peer-reviewed publications that analysed DHS data (i.e. data from any population-based household survey conducted under the DHS programme, including standard DHS, interim or special DHS, AIDS Indicator Surveys [AIS] and Malaria Indicator Surveys [MIS]). Keyword search terms included demographic and health survey, demographic AND health survey, DHS, AIDS Indicator Survey, AIS, malaria indicator survey, and MIS. We conducted additional searches on non-standard official survey names used in specific countries such as National Family Health Survey in India. In all, the search strategy included 63 unique search terms (Box 1). Based on this initial search, we found 1326 publications as of 31 December 2010.

We then identified publications that met the following eligibility criteria: (i) involved an analysis of DHS data (i.e. data derived from surveys conducted with technical assistance under the DHS project supported by USAID); and (ii) contained an English-language abstract. We excluded publications that only referenced DHS in the background or discussion.

Of the 1326 publications, 1160 were initially considered for further review based on the eligibility criteria. All 1160 publication citations were downloaded to EndNote version X2.1, a bibliographic reference database (Thomson Reuters, Carlsbad, United States of America). During the subsequent review process, described in the following sections, we determined that an additional 43 publications did not meet the eligibility criteria. This resulted in a final total of 1117 articles published across 232 journals (Fig. 2).
Measurement and analysis

Before reviewing the publications, three reviewers, who are also this study’s authors, drafted a list of research topics reflecting major survey themes. To refine the list, the three reviewers independently examined a subset of 25 abstracts and categorized them on the basis of the main outcome and independent variables. They then compared among themselves the topic areas that each had assigned and resolved any discrepancies. Ultimately, the reviewers identified 13 research topics for classification of the publications they identified (Table 1). Topic categories were intentionally made broad to minimize misclassification and interreviewer variation.

All eligible publications were systematically divided into three subsets that were randomly assigned among the three reviewers. The reviewers assigned up to three research topics to each publication and noted when a given publication did not fall clearly within any of the 13 pre-defined topics. These papers (n = 134) were discussed by two reviewers and categorized based on reviewer consensus. The reviewers then created two broad categories constituting the two primary themes of DHS data: population and reproductive health (PRH) and maternal and child health and nutrition (MCHN) (Fig. 3). The two categories were not mutually exclusive, since publications frequently included both PRH and MCHN outcomes and/or independent variables. The PRH category included papers examining one or more of the following three topics: fertility and family planning; sexual behaviour and knowledge; and sexually transmitted infection, including HIV/AIDS. The MCHN category included papers examining one or more of the following three topics: maternal health and/or mortality; child health and/or mortality; and nutrition, except for studies on adult obesity. Both categories were assigned to summary papers of DHS results.

Reviewers identified the year of publication and the country and region of residence of each study’s population. The five regional classifications were: Asia; Europe and Eurasia; Latin America and the Caribbean; the Middle East; and sub-Saharan Africa. Study populations that included more than one country were identified as “multi-country” and study populations that included more than one region were identified as “multi-region”. Publications that used data from multiple DHS for a given country to examine time trends were identified as “trend”. Reviewers also took note of special study populations, such as neonates, youth (aged 15–24 years), couples and men.

As a final step we conducted descriptive analyses of the number of publications per year, country, region and research topic, and of the number of publications using data from multiple surveys for cross-country comparisons.
and/or trend analyses. We also calculated the total number of completed surveys per year and per country and further assessed the correlation between the number of surveys and the number of publications. We conducted graphical and bivariate linear regression analyses using STATA 11.0 (Stata Corporation, College Station, USA).

**Results**

**Survey data**

A total of 236 household surveys were completed across 84 countries between 1985 and 2010, including 197 standard DHS. The annual number of standard DHS remained constant, while the annual number of all DHS increased slightly due to the introduction of AIS and MIS in the mid-2000s (results not shown).

By region, 49% of the surveys were conducted in sub-Saharan Africa, followed by Asia (20%) and Latin America and the Caribbean (18%) (Table 2). On average, 2.8 surveys were conducted per country (standard deviation, SD: 2.1; median: 2). One survey was conducted by 30 countries; two or more surveys were conducted by 54 countries, and Egypt, with 10 surveys, conducted the highest number.

Although the annual number of standard DHS remained constant, the amount and scope of the data collected in a standard DHS has increased over time, both through revisions to the household and women’s core questionnaires, which occur every five years, and through the addition of optional modules, typically at country request (Fig. 1). For example, the household core questionnaire began with 25 questions in the first phase of the DHS project, but in the current phase, the sixth, it contains 131 questions, down from a peak of 226 in phase five. Much of the increase is attributable to a handful of topics, including HIV/AIDS, malaria and maternal and child health (MCH). Accordingly, 155 standard DHS collected data on knowledge and/or behaviours related to HIV/AIDS and 36 also collected HIV biomarker data; 62 collected data on behaviours related to malaria and 6 also collected malaria biomarker data; 97 collected data on gender-related issues and 76 collected data on adult health.

**Publications using DHS data**

Analyses using DHS data have been published in over 200 journals. About 40% of the articles were published in six journals, including Studies in Family Planning (n = 179), Journal of Biosocial Science (n = 103), Social Science and Medicine (n = 56), Demography (n = 28), International Family Planning Perspectives (n = 24) and Social Biology (n = 24). As Fig. 3 shows, the annual number of publications has increased substantially since data from the first DHS became available in 1985, with an average annual increment of 4.3 publications. Of the 1117 publications identified in total, we classified 566 (51%) as PRH and 605 (54%) as MCHN. About 13% of the publications (n = 145) dealt with neither PRH nor MCHN. Although the number of publications on PRH and MCHN in-
increased from year to year, publications dealing with MCHN increased more (2.09 additional papers per year on average) than those dealing with PRH (1.37 papers per year on average) (Table 3). The recent increase in publications not dealing with PRH or MCHN reflects primarily an increase in publications on adult health (n = 21) and in studies on survey methods, including data quality (n = 70) (results not shown).

Within the PRH category, publications related to HIV/AIDS have increased substantially during the last decade – by 1.9 publications annually on average between 2001 and 2010 (P-value: 0.007, bivariate linear regression). The annual number of publications dealing with fertility and family planning peaked in 1996 and has barely changed since 1997 (Fig. 4).

Within the MCHN category, publications on child health showed the largest increment (Fig. 4, Table 3), particularly over the last decade (3.6 additional publications a year, on average, between 2001 and 2010) (P-value: 0.007, bivariate linear regression).

The number of country-specific publications is associated with the number of DHS surveys conducted in that country, even including the two outliers, Bangladesh and India (Fig. 5). Bivariate linear regression analysis suggests that each additional survey is associated with 3.5 more country-specific publications (n = 84). Publications that analysed data from multiple DHS surveys either within or across countries, which represented 34% of all publications, also increased dramatically over time (Fig. 6).

Discussion
As DHS data proliferate, so does peer-reviewed research analysing DHS data. This expansion brings increased availability of information, which provides the basis for evidence-based policy and programme development. Our analyses reveal several important trends in the use of DHS data for research publication, as well as several gaps in data utilization.

First, although the number of publications has increased progressively in all health-related subject areas, publications dealing with MCHN have increased more than those dealing with PRH. The reason may be the inclusion of additional MCHN questions to the core questionnaire during the phase five questionnaire revision of 2003, and of additional MCHN-related modules, or it may lie in funding trends, as discussed further on.

Second, the data show that the number of studies on certain topics increased after questions and modules on those topics were added to the surveys. This suggests that the modifications made to the surveys are meeting emerging data needs in global health. For example, HIV/AIDS publications based on DHS data proliferated following the 2001 introduction of HIV testing during surveys and the 2006 introduction of AIS. Similarly, gender-related publications using DHS data increased following the introduction of the women’s status module in 1995 and of the domestic violence module in 1998. Third, recent increases in adult health publications using DHS data following the introduction of questions related to adult health (e.g. on tobacco use or diabetes testing) also suggest that changes to the DHS are meeting the emerging need for data in certain areas of health. Moreover, the fact that such questions have been added

| Table 3. Association between total number of publications and number of publications on specific research topics, 1985–2010: bivariate linear regression analyses (n = 24) |
|---------------------------------|----------------|-----------------|------------------|
| Research topic                  | Coefficient per year | 95% CI          | Adjusted R²     |
| All                             | 4.27            | 3.24–5.30       | 0.76            |
| Population and reproductive health | 1.37            | 1.02–1.72       | 0.74            |
| Fertility and family planning   | 0.66            | 0.37–0.95       | 0.48            |
| Sexual behaviour and knowledge  | 0.41            | 0.31–0.51       | 0.75            |
| HIV/AIDS                        | 0.64            | 0.41–0.86       | 0.59            |
| Maternal and child health and nutrition | 2.09            | 1.48–2.70       | 0.68            |
| Maternal                        | 0.62            | 0.37–0.87       | 0.53            |
| Child                           | 1.40            | 0.95–1.85       | 0.64            |
| Nutritional deficiencies        | 0.76            | 0.49–1.03       | 0.58            |
| Other                           | 0.85            | 0.59–1.11       | 0.66            |

AIDS, acquired immunodeficiency syndrome; HIV, human immunodeficiency virus.
* Nutritional deficiencies: Publications focused on nutrition, excluding studies on adult obesity.
+ Other publications focused neither on population and reproductive health nor on maternal and child health and nutrition.

Note: Each row represents an individual regression analysis.

Fig. 4. Annual number of publications on population and reproductive health (PRH) and maternal and child health and nutrition (MCHN), by topic, in 1985–2010

AIDS, acquired immunodeficiency syndrome; HIV, human immunodeficiency virus.
* PRH encompasses three non-mutually-exclusive topics: fertility and family planning, sexual behaviour and knowledge and HIV/AIDS.
+ MCHN encompasses three non-mutually-exclusive topics: maternal health and mortality, child health and mortality and nutritional deficiencies.
indicates that the DHS project is responsive to the data needs of host countries, not just those of USAID and other international donors, for whom adult health issues are typically of lesser priority. In short, the observed trends demonstrate that DHS are adapted to meet data needs, and that researchers are using DHS data to build the knowledge base. However, even though data are available, are being turned into information by the research community, and are being disseminated through peer-reviewed publications, it is difficult to measure just how much of this information is taken up by decision-makers.

Trends in funding for a particular issue provide a clue as to how much importance decision-makers attach to the issue. We collected data on annual United States Government funding for specific international health domains – family planning, HIV/AIDS and MCHN – and compared funding trends to trends in DHS publications on the same issues. Our analysis revealed high correlation between the two. For example, HIV/AIDS publications using DHS data correlated highly with United States Government funding for global HIV/AIDS programmes (Pearson’s correlation coefficient: 0.93; n = 16; P-value < 0.01). To confirm our findings, we also collected data on annual worldwide development assistance funding for health and compared these data with the total annual number of DHS publications on all topics. Again, the trends in funding and publications were highly correlated (Pearson’s correlation coefficient: 0.90; n = 18; P-value < 0.01). Our findings do not, however, imply causality; the correlation may reflect an increase in funding stemming from a rise in the number of publications using DHS data, or it may reflect an increase in publications as a result of increased funding. It may also reflect a bidirectional relationship in which an increase in the number of publications triggers increased funding, which generates, in turn, an increase in the number of publications. The correlation may even reflect a larger, sector-wide consensus on health priorities between researchers and the donor community. Finally, the correlation may simply be spurious. Regardless of where the truth lies, the strength of the correlation does point to the likelihood that research based on DHS data provides information that is relevant for international health and development programme officials and policy-makers.

Despite the large number of published analyses using DHS data, research gaps remain. For example, we found few publications on water and sanitation (n = 6). While DHS provide data on water source, they do not collect data on storage container or other water quality factors – a limitation of DHS. Similarly, although DHS are among the few sources that provide data on unmet need for family planning, we found only 13 publications that examined this topic or that included it as a covariate in analysis. This finding may reflect a true gap or
a weakness in our search term—unmet need and demand are often conflated. Published research based on data from special surveys is also limited. Although Peru has implemented the DHS Continuous Survey annually since 2004, our search discovered only one publication using those data. Several factors may have contributed to this gap, including, for example, the unique challenges involved in the statistical analysis of continuous survey data. These examples highlight not only the need for more published research in specific areas, such as water and sanitation or unmet need for family planning, but also, the importance of increasing awareness of non-traditional data sources, such as special surveys, among researchers.

Our study has limitations. First, we did not conduct a comprehensive search of studies analysing DHS data in both the peer-reviewed and grey literature. Further, our reliance on PubMed may have limited our results, since PubMed does not index certain non-biomedical but health-oriented journals, such as Population and Development Review, that are known to publish research using DHS data. Second, we may have underestimated researchers’ use of DHS data for trend and cross-country analyses, in part because we excluded trend analyses using one DHS data set in comparison with one or more non-DHS nationally-representative data sets (e.g. Multiple Indicator Cluster Survey), and perhaps also because our search terms were too restrictive and failed to identify several well-known global systematic analyses that employed DHS data.6–15 Third, we did not conduct any formal statistical test to assess inter-rater variation among reviewers. Nevertheless, we believe misclassification was minimal since our classifications were broad and not mutually exclusive. Finally, there is no benchmark against which to judge the amount of research generated from DHS data. Without a benchmark, assigning value to the figures presented herein is difficult.

Despite these limitations, our findings show that researchers are using DHS data more and more and that, as a result, health programme officials and policy-makers have increasing access to vital health information. In light of the world’s changing epidemiologic profile, the DHS project must continue to be responsive to the need for new types of data while maintaining data quality and comparability. Finally, countries and donors must find ways to capitalize on the opportunities provided by DHS to strengthen statistical capacity and improve health information systems.

**Funding:** No support from any organization for this submitted work. The views expressed herein do not necessarily reflect the views of the United States Agency for International Development.

**Competing interests:** None declared.
Revue systématique des enquêtes démographiques et de santé: disponibilité et utilisation des données pour la recherche

Résumé

Objectif Établir une revue systématique des publications dans le domaine de la santé publique pour évaluer les tendances dans l’utilisation des données de l’Enquête démographique et de santé (Demographic and Health Survey: DHS) pour la recherche de 1984 à 2010, et décrire la relation entre la disponibilité des données et leur utilisation.

Méthodes On a consulté le site internet MEASURE DHS pour obtenir des informations sur toutes les enquêtes sur la population, réalisées dans le cadre du projet DHS entre 1984 et 2010. Les données recueillies comprenaient le pays, le type d’enquête, la structure de l’enquête, la période de travail sur le terrain et certains paramètres particuliers, tels que l’inclusion de biomarqueurs. Une recherche sur PubMed a également été menée pour identifier les articles soumis à comité de lecture publiés en 2010, qui analysaient les données DHS et comprenaient un résumé en langue anglaise. Les tendances de la disponibilité des données DHS et de leur utilisation pour la recherche ont été évaluées par le biais d’analyses descriptives, graphiques et de régression linéaire à deux variables.

Résultats Au total, 236 enquêtes ont été réalisées auprès des ménages dans le cadre du projet DHS dans 84 pays en 2010. Le nombre d’enquêtes annuel est resté constant, bien que la portée des questions des enquêtes ait été élargie. Les critères d’inclusion ont été repris par 1117 publications à comité de lecture. Le nombre de publications a progressivement augmenté au cours du dernier quart de siècle, avec une augmentation annuelle moyenne de 4,3 publications (intervalles de confiance de 95%, CI : 3,2 à 5,3). Les tendances du nombre de publications évaluées par comité de lecture basées sur l’utilisation des données du DHS étaient fortement corrélées à l’évolution du financement de la santé par le gouvernement des États-Unis d’Amérique et plus généralement dans le monde.

Conclusion Les articles publiés après révision par comité de lecture analysant les données DHS, qui ont progressivement vu leur nombre augmenter au cours du dernier quart de siècle, ont apporté une contribution conséquente à la base de données de santé publique dans les pays en voie de développement.

Systematic review of surveys on demographic and health: availability and use for research

Abstract

Objective Establish a systematic review of literature on the use of data from the Demographic and Health Survey (DHS) for research purposes from 1984 to 2010, and describe the relationship between data availability and their use.

Methods Search of the website MEASURE DHS for all surveys on population conducted in the context of the DHS from 1984 to 2010. Data collected included country, type of survey, survey structure, fieldwork period, and certain characteristics, such as inclusion of biomarkers. A search on PubMed was also conducted to identify articles published in 2010 that analyzed the DHS data and included a summary in English. Trends in data availability and their use for research were evaluated using descriptive, graphical, and linear regression analysis.

Results In total, 236 surveys were conducted on 84 countries in 2010. The annual number of surveys remained constant, although the scope of the questions of the surveys expanded. The criteria for inclusion were carried over to 1117 publications at peer review. The number of publications progressively increased over the last quarter of the century, with an annual average increase of 4.3 publications (95% confidence interval 3.2 to 5.3). The trends in the number of publications evaluated by peer review based on survey data were highly correlated with the evolution of government spending on health in the United States and more generally in the world.

Conclusion The publication of articles reviewed after peer review analyzing the DHS data, which progressively increased over the last quarter of a century, contributed significantly to the research data base in developing countries.

Revisión sistemática de las encuestas demográficas y de salud: disponibilidad de los datos y su uso en investigaciones

Resumen

Objetivo Revisar de forma sistemática la bibliografía sobre salud pública para evaluar las tendencias en el uso de los datos de la Encuesta demográfica y de salud (DHS, por sus siglas en inglés) en investigaciones desde 1984 hasta 2010 y describir la relación entre la disponibilidad y el uso de los datos.

Métodos Se examinó la página web MEASURE DHS en busca de información acerca de todas las encuestas demográficas completadas dentro del proyecto DHS entre 1984 y 2010. La información recogida incluyó el país, el tipo de encuesta, el diseño de la encuesta, el periodo de trabajo de campo así como ciertas características especiales tales como la inclusión de biomarcadores. También se examinó PubMed con objeto de identificar los artículos revisados por especialistas externos y publicados a lo largo de 2010 que analizaran datos de DHS e incluyeron un resumen en inglés. Se evaluaron las tendencias en la disponibilidad y en el uso de datos de DHS en investigaciones a través de análisis de regresión lineal descriptivos, gráficos y bivariados.
Resultados En total, a lo largo del año 2010 se habían completado 236 encuestas de hogares en 84 países dentro del proyecto DHS. El número de encuestas anuales se ha mantenido constante, si bien se ha ampliado el alcance de las preguntas de las encuestas. De las publicaciones revisadas por especialistas externos, 1117 cumplieron con los criterios de inclusión. El número de publicaciones ha aumentado de manera progresiva durante el último cuarto de siglo, con un aumento medio de 4,3 publicaciones al año (intervalo de confianza, IC, del 95%: 3,2 – 5,3). Las tendencias en el número de publicaciones revisadas por especialistas externos y basadas en el uso de datos de DHS guardaron una considerable correlación con las tendencias en la financiación para la salud del Gobierno de los Estados Unidos de América y a nivel mundial.

Conclusión Los artículos revisados por especialistas externos y publicados que analizan los datos de DHS, cuyo número ha aumentado de manera progresiva durante el último cuarto de siglo, han contribuido considerablemente a la base de datos sobre la salud pública en los países en desarrollo.

References