**Claims data from health insurance programs in sub-Saharan Africa – an untapped repository for achieving Universal Health Coverage**

Authors: Angela E. Apeagyei, Maitreyi Sahu

Address: Institute for Health Metrics and Evaluation, Hans Rosling Center for Population Health, 3980 15th NE, Seattle, WA 98195

Word count: *1459*

***What is the problem?***

Despite the looming deadline for the Sustainable Development Goals in 2030, many low- and middle-income countries are falling short of the health-related Universal Health Coverage (UHC) goal which aims to ensure access to health care services without risk of financial hardship.1 The pandemic has exacerbated these challenges, pushing countries even further behind in their progress. In the effort to advance towards UHC, many African countries have implemented public health insurance programs to increase financial risk protection, reduce catastrophic health expenditure, and broaden access to affordable health care services for treatment and disease prevention.­2,3 In sub-Saharan Africa, eight countries out of 49 have implemented some sort of national-level contributory public health insurance system, and at least seven others have passed legislation or are in the process of planning a national health insurance program (Figure 1).4 However, programs face challenges with low enrollment, limited uptake of services, and uncertain financial sustainability. Such programs can greatly benefit from greater insight and analysis of their enrollment, utilization, and payment measures to better identify gaps and opportunities for improvement. Despite several countries having a wealth of health insurance claims data on beneficiaries and their health care utilization and spending patterns, little analysis of these data exist in the published literature. Currently, among the fifteen countries with existing or planned national health insurance programs, we identified only two (Tanzania and Ethiopia) which have used claims data in publicly accessible published research (Figure 1); of these studies, only the research from Tanzania covers the national health insurance program – the Ethiopian study covers only a community-based health insurance program in two districts.5–7 Moreover, neither of these programs have made their data publicly available. This represents a missed opportunity for the majority of health insurance programs in sub-Saharan Africa, as these rich datasets could be leveraged to support national-level policymaking. Publishing these analyses could provide valuable lessons learned for peer countries facing similar challenges in sustainably achieving UHC.

***What is the opportunity?***

Health insurance claims databases generally include beneficiary demographic information (e.g. age, sex, geographic region, salary), beneficiary utilization patterns (including health services accessed and potentially also the health condition being treated), and total and out-of-pocket expenditures for health services across providers. These data, while primarily collected for administrative and financial payment purposes rather than for research, can be analyzed for a number of purposes. First, by comparing enrollment patterns with population demographics from census data which are available for every country as well as survey datasets such as the Demographic and Health Surveys (DHS) and Multiple Indicator Cluster Surveys (MICS), it is possible to identify gaps in beneficiary enrollment by age, sex, and socioeconomic indicators at the regional level. Such analyses have been undertaken using survey data alone in countries such as Ghana which have comprehensive national surveys identifying health insurance enrollment,8 but countries could also leverage enrollment data from claims databases for this purpose if it were available. Second, using claims data for equity-focused analyses can identify utilization and spending patterns and gaps across age, sex, income, and geography. For example, a recent study of Tanzania’s National Health Insurance Fund found that just five hospitals in Dar es Salaam out of a total of 7390 health centers accounted for 30% of total program spending.6 Third, understanding the take-up of services by health condition can help to understand the financial burden of each disease. In high-income countries, such analyses of national insurance programs have been used by policymakers to better understand not only gaps in coverage but also health spending prioritization and how to control health care costs.9–11 Furthermore, they have even been linked to disease registries and prevalence estimates to understand the relationship between disease burden and spending, and factors driving spending.9,12 Fourth, these analyses can inform policymakers on the financial sustainability of programs including by highlighting the spending breakdown for primary and preventive care compared with tertiary and specialist care, and by allowing for projections over time and scaled up to the full country population.5–7 Finally, analyses of health insurance claims could be used for understanding health care quality and identifying opportunities for quality improvement.13 However, despite the enormous potential for these analyses, no African health insurance program has published any comprehensive and comparable assessments of spending by disease, demographic group and/or region.

***Why have claims data not been analyzed historically?***

There are a number of political and logistical reasons why these analyses may not have been pursued. Historically, claims were documented in paper records; however, currently the standard is to enter the claims into an electronic database, even if they are initially recorded on paper. There may also be valid concerns related to disclosing potentially sensitive beneficiary health and payment information which could invite both privacy concerns and external scrutiny. Ensuring beneficiary privacy by decrypting and suppressing data so that it cannot be made re-identifiable is key. In addition, not all programs might have access to the computational capacity to store and analyze these large databases which may include millions if not billions of claims per year as programs are scaled up. Further investment in the appropriate health information technologies may be warranted. Importantly, the enormous opportunity these analyses can wield for internal improvement of their programs as well as the implications for policy makers and other stakeholders may not be fully acknowledged by key players in the national health insurance programs. Finally, even if the political will and technical capacity to analyze these data exist, there are several steps required related to cleaning and standardizing the data so it can appropriately be used for national-level analyses which can sometimes make this exercise a non-trivial time commitment.

***What are some challenges with analyses of claims data?***

A review of the limited examples of published research using claims data in sub-Saharan Africa highlights some further challenges in using claims data to understand national health care spending and utilization. In terms of data coverage, the public health insurance systems we identified despite being national are in various phases of scale-up (Figure 1) and tend to cover a small segment of formally employed public sector or other employees and their dependents who likely are not representative of the full population. Furthermore, there are documented cases that even beneficiaries who are enrolled in health insurance programs may prefer to pay out-of-pocket for services as providers may prefer direct payment rather than waiting for reimbursement through the insurance system14 – therefore, insurance claims may not capture all utilization even among enrolled beneficiaries, and the claims may be skewed towards higher cost treatments. Pharmaceutical claims data in particular may be limited in their coverage because of the large proportion of medicines purchased out-of-pocket in private pharmacies, which may be better captured using retail pharmaceutical sales data.15 Furthermore, since data are collected for claims processing rather than for research, there is also less quality control of the data which may lead to missing or incorrect beneficiary or claim information, or values that do not appropriately correspond to the research question.6 For example, a common issue with insurance claims from any country is that they often capture just the charge to the insurer but not the total spending related to the health service provided. Finally, claims records for existing insurance programs in Africa may be handwritten and then entered as free-text and therefore highly challenging to analyze by health condition due to non-standardized reporting,6,7 in contrast to the International Classification of Disease codes which are used in many countries including the US. However, despite the many challenges with these data, the successful examples from the US, Europe, Tanzania and Ethiopia suggest that many of these data issues can be effectively resolved especially in light of recent technological developments.

***What is the way forward?***

We strongly advocate for countries to recognize the immense potential in analyzing health insurance claims data and emphasize that the benefit that countries will gain from making de-identified health insurance data accessible to researchers far outweigh the technical and logistical obstacles which currently exist. Ensuring that the data are stored electronically is crucial. Technological advances in cloud computing, data security, and artificial intelligence can further help to mitigate challenges around data storage and processing, patient privacy, and the standardization of disease coding, respectively. Programs can also leverage a number of existing statistical methods and adjustments exist to fill in data gaps, better represent national populations, and estimate total spending using charge data.10 Alternatively, programs can partner with external organizations with greater capacity to take on these technical tasks and facilitate knowledge and skill sharing, including in overcoming common challenges in analyzing claims data. With political will to surmount these hurdles, countries can gain valuable insights into how to expand coverage and ensure financially sustainable and equitable programs on the road to achieving UHC.

**Summary Box**

* Many sub-Saharan African countries are implementing national public health insurance programs in order to achieve Universal Health Coverage. However, insurance programs face challenges including low enrollment and questions around financial sustainability. Despite the availability of claims data, few analyses of these data exist in the published literature.
* Analyses of claims data are hindered by privacy concerns, computational capacity, and challenges with data quality and representativeness. Additionally, the potential of these data to provide valuable insights for programs may not fully be acknowledged.
* We highlight examples of analyses of claims data from Ethiopia and Tanzania which have provided insight into financial sustainability and distributional equity of health spending and utilization, as well as studies from the US and Europe which used claims to understand health care utilization and spending by disease and demographic group.
* We urge national health insurance programs in sub-Saharan Africa to invest in developing their infrastructure for analyzing their claims data, to partner with external organizations where beneficial, and to use the wealth of claims data they are collecting to provide insights into how to sustainably achieve Universal Health Coverage.

**Acknowledgements**

We thank Justice Nonvignon and Joseph Dieleman for their valuable feedback on a draft of this article.

**References**

1. United Nations Department of Economic and Social Affairs. *2030 Sustainable Development Goals. Goal 3: Ensure Healthy Lives and Promote Well-Being for All at All Ages*. Accessed November 26, 2023. https://sdgs.un.org/goals/goal3#targets\_and\_indicators

2. Ly MS, Bassoum O, Faye A. Universal health insurance in Africa: a narrative review of the literature on institutional models. *BMJ Glob Health*. 2022;7(4):e008219. doi:10.1136/bmjgh-2021-008219

3. Barasa E, Kazungu J, Nguhiu P, Ravishankar N. Examining the level and inequality in health insurance coverage in 36 sub-Saharan African countries. *BMJ Glob Health*. 2021;6(4):e004712. doi:10.1136/bmjgh-2020-004712

4. Cashin C, Dossou JP. Can National Health Insurance Pave the Way to Universal Health Coverage in Sub-Saharan Africa? *Health Systems & Reform*. 2021;7(1):e2006122. doi:10.1080/23288604.2021.2006122

5. Hussien M, Azage M, Bayou NB. Financial viability of a community-based health insurance scheme in two districts of northeast Ethiopia: a mixed methods study. *BMC Health Serv Res*. 2022;22(1):1072. doi:10.1186/s12913-022-08439-8

6. Durizzo K, Harttgen K, Tediosi F, et al. Toward mandatory health insurance in low‐income countries? An analysis of claims data in Tanzania. *Health Economics*. 2022;31(10):2187-2207. doi:10.1002/hec.4568

7. Osetinsky B, Fink G, Kuwawenaruwa A, Tediosi F. Investigating sustainability challenges for the National Health Insurance Fund in Tanzania: a modelling approach. *BMJ Open*. 2023;13(8):e070451. doi:10.1136/bmjopen-2022-070451

8. Salari P, Akweongo P, Aikins M, Tediosi F. Determinants of health insurance enrolment in Ghana: evidence from three national household surveys. *Health Policy and Planning*. 2019;34(8):582-594. doi:10.1093/heapol/czz079

9. Riley GF. Administrative and Claims Records as Sources of Health Care Cost Data. *Medical Care*. 2009;47(7\_Supplement\_1):S51-S55. doi:10.1097/MLR.0b013e31819c95aa

10. Dieleman JL, Cao J, Chapin A, et al. US Health Care Spending by Payer and Health Condition, 1996-2016. *JAMA*. 2020;323(9):863. doi:10.1001/jama.2020.0734

11. Kinge JM, Dieleman JL, Karlstad Ø, et al. Disease-specific health spending by age, sex, and type of care in Norway: a national health registry study. *BMC Med*. 2023;21(1):201. doi:10.1186/s12916-023-02896-6

12. Dieleman JL, Squires E, Bui AL, et al. Factors Associated With Increases in US Health Care Spending, 1996-2013. *JAMA*. 2017;318(17):1668. doi:10.1001/jama.2017.15927

13. Parente ST, Weiner JP, Garnick DW, et al. Developing a Quality Improvement Database Using Health Insurance Data: A Guided Tour with Application to Medicare’s National Claims History File. *Am J Med Qual*. 1995;10(4):162-176. doi:10.1177/0885713X9501000402

14. Dalinjong PA, Laar AS. The national health insurance scheme: perceptions and experiences of health care providers and clients in two districts of Ghana. *Health Econ Rev*. 2012;2(1):13. doi:10.1186/2191-1991-2-13

15. Guadamuz JS, Shooshtari A, Qato DM. Global, regional and national trends in statin utilisation in high-income and low/middle-income countries, 2015–2020. *BMJ Open*. 2022;12(9):e061350. doi:10.1136/bmjopen-2022-061350