

# Claims data from health insurance programmes in sub-Saharan Africa: an untapped resource to promote Universal Health Coverage

Angela E Apeageyi , Maitreyi Sahu

**To cite:** Apeageyi AE, Sahu M. Claims data from health insurance programmes in sub-Saharan Africa: an untapped resource to promote Universal Health Coverage. *BMJ Glob Health* 2024;**9**:e015310. doi:10.1136/bmjgh-2024-015310

**Handling editor** Helen J Surana

Received 8 February 2024  
Accepted 21 June 2024

## WHAT IS THE PROBLEM?

In the effort to advance towards Universal Health Coverage (UHC), several African countries are implementing public health insurance programmes to increase financial risk protection, reduce catastrophic health expenditure and broaden access to services for treatment and prevention of disease.<sup>1,2</sup> In sub-Saharan Africa (SSA), 8 out of 49 countries have implemented some form of national-level contributory public health insurance system, and at least 7 others have passed legislation or are in the process of planning a national health insurance programme (Figure 1).<sup>1</sup> However, programmes face challenges with low enrolment, limited uptake of services and uncertain financial sustainability. These programmes can greatly benefit from analyses to identify opportunities for improvement and to support national policymaking.

Health insurance claims data, which are primarily collected for administrative and financial purposes, capture a wealth of information on beneficiary demographics (eg, age, sex, geographic region), beneficiary utilisation patterns (including health services accessed) and total and out-of-pocket expenditure across beneficiaries, providers and health facilities. Globally, many countries use these data for both internal programme monitoring and in publishable reports and journal articles to support national-level programme design. Academic researchers may also use these data, when available, to study ways to make healthcare systems more efficient and equitable.

Currently, among 15 countries in SSA with existing or planned national health insurance programmes, we identified only three—Ghana, Tanzania and Ethiopia—which have used claims data in published research (Figure 1). Of these studies, only the

## SUMMARY BOX

- ⇒ Many sub-Saharan African countries are implementing national public health insurance programmes to achieve Universal Health Coverage. However, insurance programmes face challenges including low enrolment 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 inadequate technological infrastructure, privacy concerns and challenges with data quality and representativeness. Additionally, the potential of these data to provide valuable insights for programmes may not fully be acknowledged.
- ⇒ We highlight examples of analyses of national health insurance programmes in Ghana, Tanzania, Indonesia and the USA that used claims data that provided insight into programme sustainability, quality of care and distributional equity of health services.
- ⇒ We urge national health insurance programmes in sub-Saharan Africa to invest in developing their infrastructure for analysing their claims data, to partner with external organisations where beneficial, and to consider making samples of their claims data available for research to provide insights towards sustainably achieving Universal Health Coverage.

research from Ghana<sup>3–5</sup> and Tanzania<sup>6–8</sup> cover national programmes; the Ethiopian study covers a pilot community-based programme in two districts.<sup>9</sup> Moreover, none of these programmes have released a sample of their data for public access, missing an opportunity to solicit external data processing and analytic capacity to understand healthcare utilisation and spending patterns and their drivers.

## WHAT ARE THE CHALLENGES?

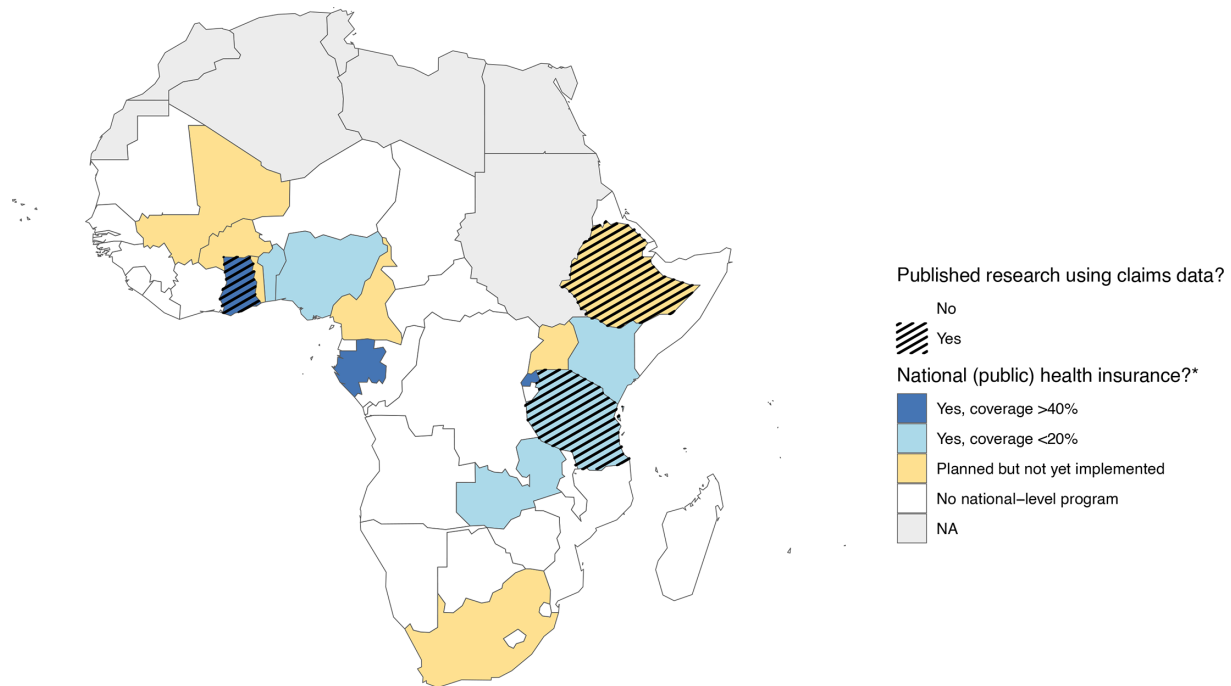
There are several logistical and political reasons why analysis of health insurance



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

Institute for Health Metrics and Evaluation, Seattle, Washington, USA

**Correspondence to**  
Dr Angela E Apeageyi;  
amicah@uw.edu



**Figure 1** Does a national public insurance program exist, and has there been published research using claims data? \*Status of insurance programs adapted from Cashin and Dossou.<sup>1</sup>

claims data from SSA for both national programme monitoring and research purposes may not have been pursued.

### Technical infrastructure

Historically, claims were documented in paper records and could not easily be described. However, many countries in SSA have switched to and benefitted from a system of entering claims into an electronic database, even if initially recorded on paper—for example, a Ghanaian study showed that electronic claims review saved programme costs compared with paper-based review.<sup>3</sup> Data size is also a challenge: not all programmes may have access to the computational capacity to store and process databases, which may include millions if not billions of claims per year. 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. These beneficiary privacy concerns may be alleviated by de-identifying and suppressing data so that it cannot be made re-identifiable.

### Data completeness and quality

Since claims data are collected primarily for processing reimbursements and not for programme monitoring or research, poor data completeness and accuracy for beneficiary and claim details can be a challenge. In addition, claims data often cover the provider charge to the insurer but not the adjudicated reimbursement, which represents

total spending. Claims records from SSA can sometimes also be challenging to aggregate by category because health conditions, drugs and procedures may be entered as free-text instead of using standardised coding such as International Classification of Diseases (ICD) codes.<sup>6 7</sup> Extensive manual cleaning may be required to leverage these data for programme monitoring, and further statistical adjustments may be required for research purposes.

### Applicability for research

Studies using claims data to draw national-level inference or forecasts should be aware of potential bias. As national health insurance programmes are being scaled up in SSA, they often include a small sample of primarily formal sector employees whose spending patterns may not be nationally representative.<sup>6</sup> Furthermore, claims data do not capture health spending by uninsured patients and may not cover spending by enrolled beneficiaries who pay out-of-pocket for expediency or when they access out-of-network providers or pharmacies. Studies describing health spending may need to incorporate adjustments when scaling claims data to the national level, or to supplement or compare claims data with nationally representative surveys, demographic data and/or pharmaceutical sales data.

### Political will

Finally, a critical factor is that key players in national health insurance programmes may not fully acknowledge the valuable insight that analysis of claims data can yield

to inform national programme design and peer country programmes. Furthermore, there may be well-founded scepticism of partnering with external researchers from institutions in high-income countries who may not have traditionally given the appropriate credit to country programmes and researchers.

### WHAT ARE THE OPPORTUNITIES?

Globally, a number of countries invest in their technical infrastructure for national health insurance data and analyse these data internally and in partnership with external researchers. Here, we highlight a few national programmes using claims data to understand programme coverage, cost drivers and sustainability.

#### Ghana

Ghana's National Health Insurance Authority (NHIA), which manages its National Health Insurance Scheme covering 40% of Ghanaians, maintains a robust electronic claims database with standardised disease and procedural coding.<sup>4</sup> Researchers within the NHIA use claims data for programme monitoring, NHIA annual reports include totals for annual inpatient and outpatient utilisation and spending, and published research includes a financial assessment of medical claims review to ensure sustainability.<sup>3</sup> Additionally, other studies, collaborations between NHIA researchers and external scientists, have leveraged claims to describe utilisation and quality of care for certain prescribed medicines as well the assessment of efficiency in claims processing.<sup>4,5</sup>

#### Tanzania

Tanzania's National Health Insurance Fund (NHIF) covers civil servants and their dependents comprising approximately 7% of the population. Claims are stored electronically; however, diagnoses, drugs and procedures have been recorded using free text. Accordingly, NHIF reports contain little analysis of claims beyond enrolment and spending summaries. Despite data limitations, equity researchers from Tanzanian and Swiss institutions used the 2015–2016 claims data to describe service take-up across geography, health conditions and beneficiary demographics.<sup>6,7</sup> These analyses also informed the financial sustainability of the programme when scaled to the full population and projected over time.

#### Indonesia

The Indonesian Social Health Security Agency *BPJS Kesehatan* manages its national health insurance programme covering 75% of the population and maintains claims in an electronic database, which includes ICD diagnosis codes. However, few analyses of these data existed until recently, in part due to privacy concerns.<sup>10</sup> In 2019, BPJS Kesehatan publicly released (at no cost) the 2015–2016 social health insurance sample data (SHISD), a 1% representative sample of their claims data covering nearly 1.7 million individuals. Academic researchers have since used the SHISD to describe hospital utilisation and costs

at the diagnosis level, for example, for dengue hemorrhagic fever.<sup>11</sup> Researchers also used the SHISD to evaluate Indonesia's performance-based financing scheme for primary care providers, which they found modestly increased primary care utilisation.<sup>12</sup> In 2023, the SHISD was extended longitudinally to cover 2015–2022.

#### USA

The Centers for Medicare and Medicaid Services (CMS), the federal agency which supports two major US government-funded health insurance programmes, internally analyses and reports trends and forecasts in national-level and state-level healthcare costs and utilisation. In addition, to leverage outside research capacity to inform policy, CMS makes from 1% up to 100% of its de-identified Medicare and Medicaid insurance claims available for purchase.<sup>13</sup> These fees offset data storage and processing requirements to handle the large data volumes and ensure data security and quality and allow for free trainings and support from the CMS Research Data Assistance Center. A large body of literature has been published using claims data to help CMS better understand health spending prioritisation; cost drivers and cost containment; and quality, equity and efficiency in service delivery. Moreover, since claims include ICD codes, they have even been aggregated by disease and linked to population and prevalence estimates to understand how factors including population growth and disease burden drive costs.<sup>14,15</sup>

### WHAT IS THE WAY FORWARD?

In recognition that resources are limited, we recommend that countries with national health insurance programmes develop priorities tailored to their operational needs. For instance, we call on programmes with gaps in health information infrastructure to prioritise capturing claims electronically using standardised disease-coding procedures, which can facilitate analyses and also save programme costs in medical claims review.<sup>3</sup> These countries can also partner with external organisations with technological expertise in large database management and data security to ensure patient privacy. Similarly, we encourage countries with existing health information and computational infrastructure to use their claims for internal monitoring and to consider making a sample of that data publicly available for research to facilitate leveraging of relevant external statistical and other research expertise to draw inference on their programmes. With appropriate political commitment, countries can gain valuable insights into how to expand coverage and ensure financially sustainable and equitable programmes on the road to achieving UHC.

**Acknowledgements** We thank Justice Nonvignon, PhD and Joseph Dieleman, PhD for their valuable feedback on a draft of this article, and Noemi Kreif, PhD for her helpful insight on the Southeast Asian context.

**Contributors** AEA and MS conceptualised the commentary. AEA wrote the first draft and supervised all aspects of the study. MS contributed to writing the

commentary, reviewed literature and produced the figure. All authors approved the study.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Map disclaimer** The inclusion of any map (including the depiction of any boundaries therein), or of any geographic or locational reference, does not imply the expression of any opinion whatsoever on the part of BMJ concerning the legal status of any country, territory, jurisdiction or area or of its authorities. Any such expression remains solely that of the relevant source and is not endorsed by BMJ. Maps are provided without any warranty of any kind, either express or implied.

**Competing interests** AEA declares no competing interests. Outside the submitted work, MS declares payments to the Institution by the National Pharmaceutical Council and consulting services to the Pharmaceutical Care Management Association.

**Patient consent for publication** Not applicable.

**Ethics approval** Not applicable.

**Provenance and peer review** Not commissioned; externally peer-reviewed.

**Data availability statement** Data are available upon request.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iD

Angela E Apeagyei <http://orcid.org/0000-0002-0190-1017>

## REFERENCES

- Cashin C, Dossou JP. Can national health insurance pave the way to universal health coverage in sub-Saharan Africa. *Health Syst Reform* 2021;7:e2006122.
- Barasa E, Kazungu J, Nguhiu P, *et al*. Examining the level and inequality in health insurance coverage in 36 sub-Saharan African countries. *BMJ Glob Health* 2021;6:e004712.
- Nsiah-Boateng E, Asenso-Boadi F, Dsane-Selby L, *et al*. Reducing medical claims cost to Ghana's national health insurance scheme: a cross-sectional comparative assessment of the paper- and electronic-based claims reviews. *BMC Health Serv Res* 2017;17:115.
- Ankrah D, Hallas J, Odei J, *et al*. A review of the Ghana national health insurance scheme claims database: possibilities and limits for drug utilization research. *Basic Clin Pharma Tox* 2019;124:18–27.
- Donneyong MM, Chang T, Pottegård A, *et al*. Prevalence and quality of antihypertensive therapy among hypertension patients enrolled in the Ghana national health insurance scheme. *Pharmacoepidemiol Drug* 2021;30:1566–75.
- Durizzo K, Harttgen K, Tediosi F, *et al*. Toward mandatory health insurance in low-income countries? An analysis of claims data in Tanzania. *Health Econ* 2022;31:2187–207.
- Osetinsky B, Fink G, Kuwawenaruwa A, *et al*. Investigating sustainability challenges for the national health insurance fund in Tanzania: a modelling approach. *BMJ Open* 2023;13:e070451.
- Kapungu I, Evarest E, Shaban N, *et al*. Modelling and forecasting claim payments of Tanzania national health insurance fund. *Tanz J Sci* 2023;49:911–20.
- 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;22:1072.
- Ng JYS, Ramadani RV, Hendrawan D, *et al*. National health insurance databases in Indonesia, Vietnam and the Philippines. *Pharmacoecoon Open* 2019;3:517–26.
- Fuad A. Indonesian social health insurance sample data are now available and accessible for research. *Pharmacoecoon Open* 2019;3:643–5.
- Sambodo NP, Bonfrer I, Sparrow R, *et al*. Effects of performance-based capitation payment on the use of public primary health care services in Indonesia. *Soc Sci Med* 2023;327:115921.
- data.CMS.gov. Centers for medicare & medicaid services. Available: <https://data.cms.gov/> [Accessed 15 Mar 2024].
- Dieleman JL, Cao J, Chapin A, *et al*. US health care spending by payer and health condition, 1996–2016. *JAMA* 2020;323:863–84.
- Dieleman JL, Squires E, Bui AL, *et al*. Factors associated with increases in US health care spending, 1996–2013. *JAMA* 2017;318:1668–78.