Working paper SLE-25304 July 2026

Quality testing of commonly purchased drugs in Freetown

Fatu Conteh
Michael Dinerstein
Sofia Gallo
Anne Karing
Aaron Leonard

DIRECTED BY LSE UNIVERSITY OF OXFORD

FUNDED BY UKaid from the British people

Quality Testing of Commonly Purchased Drugs in Freetown*

Fatu Conteh† Michael Dinerstein‡ Sofia Gallo§ Anne Karing¶ Aaron Leonard‖

July 2026

Abstract

Whether a medicine helps or harms a patient depends on many factors, including the match between drug and condition, access and affordability, and the quality of the medicine itself. In this study, we focus on the latter by evaluating the quality of commonly purchased drugs in Freetown, Sierra Leone. Through mystery shopping, we collect 80 samples of four such drugs and submit them for pharmacopoeial testing. We find no evidence of falsified drugs and no dangerous contaminants. One in four samples is nonetheless substandard, failing at least one quality test, primarily by releasing too little of the active ingredient. These failures concentrate in painkillers rather than in the dewormer or antibiotic. In principle, both storage and manufacturer differences could explain this pattern. We find suggestive evidence for the latter: failures cluster by specific brands and variation in how often vendors restock is not predictive of drug failure. Targeted manufacturer-level enforcement may be more effective as a policy response than vendor-level restrictions, since those may limit an important channel of health access for consumers in low-income countries.

JEL codes: I11, I12, I18, O12

Keywords: drug quality, medicine access, pharmacopoeial testing, health, Freetown

*We thank the Ghana Food and Drugs Authority Centre for Laboratory Services and Research for conducting the pharmacopoeial testing. This research was funded by the International Growth Centre (IGC).

†RAMLabs, Freetown, Sierra Leone. Email: fconteh@ramlabs.org.

‡Duke University. Email: michael.dinerstein@duke.edu.

§University of Chicago. Email: skgallo@uchicago.edu.

¶University of Chicago. Email: akaring@uchicago.edu.

‖University of Chicago. Email: aaronleonard@uchicago.edu.

1 Introduction

Whether a given medicine helps a patient depends on many factors: the medicine has to be appropriate for the underlying health condition and affordable, and the patient has to comply with taking it. Another aspect, which this paper focuses on, is the quality of the medicine itself.

Drug quality is particularly important to policymakers in low- and middle-income countries, as existing evidence shows that medicines on the market can be falsified (i.e., have a misrepresented identity) or can simply contain substandard active ingredients and fail quality tests [World Health Organization, 2017, Ozawa et al., 2018]. In sub-Saharan Africa, studies of antimalarial quality have reported failure rates as high as 35–40% [Nayyar et al., 2012, Bate et al., 2008a, Tabernero et al., 2014], and poor-quality antimalarials have been linked to an estimated 122,000 under-five deaths each year [Renschler et al., 2015].

In this paper, we focus on studying the quality of commonly purchased drugs in Freetown, Sierra Leone. To answer this, we collected 80 drug samples through mystery shopping across four commonly purchased active pharmaceutical ingredients (APIs): Diclofenac Potassium and Paracetamol (painkillers), Mebendazole (deworming), and Metronidazole (antibiotic). Mystery shopping lets us measure the quality of medicine a consumer would actually receive when purchasing. All samples were submitted, with blinded labels, to the Ghana Food and Drugs Authority (FDA) Centre for Laboratory Services and Research for pharmacopoeial testing, including identification, assay (amount of active ingredient), dissolution (whether the tablet releases its contents quickly enough), uniformity of weight, and impurities.

Three findings emerge. First, we find no falsified drugs, as no sample failed identification or contained impurities. However, substandard drugs are somewhat common. 20 of 80 samples (25%) failed at least one pharmacopoeial test, and every failure involved dissolution. Failing dissolution testing means that the tablet did not release enough of the active ingredient in a specified time frame, so the effective dosage is smaller than what is listed on the label.

Second, substandard samples are heterogeneous across the APIs we test. Paracetamol fails most often (45%), followed by Diclofenac (40%) and Mebendazole (15%). Metronidazole has no failures, though it follows a different testing protocol without dissolution and so is not directly comparable. The shortfalls therefore concentrate in painkillers, which manage symptoms, while the deworming pill and the antibiotic, APIs that target underlying conditions, perform better.

Third, failures cluster in a small number of manufacturer-product combinations. Further, there is not a significant difference in substandard drugs between the sources of drugs

(informal vs. formal providers). Relatedly, informal providers source drugs from wholesale pharmacies frequently, and variation in how often drugs are sourced is not predictive of failure rate. Overall, we interpret this as suggestive evidence that manufacturing issues rather than storage issues explain the difference between providers. We cannot rule out that upstream storage issues common to both formal and informal vendors contribute to the drug testing failures.

The paper contributes to three literatures. First, we contribute to work on drug quality in developing countries [Johnston and Holt, 2014, Ozawa et al., 2018, World Health Organization, 2017, Almuzaini et al., 2013], where we add evidence by assessing the medicine consumers take in Freetown, Sierra Leone through mystery shopping. Next, our work relates to informal healthcare markets [Dagrou and Chimhutu, 2022, Das et al., 2016, Hughes et al., 2013], where prior work has focused on provider quality and information rather than product quality directly. Finally, we relate to work on drug quality as a market outcome [Björkman-Nyqvist et al., 2018, Bennett and Yin, 2018, Bate et al., 2011], where our finding that quality gaps concentrate in specific brands rather than channels suggests targeted manufacturer-level regulatory approaches may be more successful.

The paper proceeds as follows. Section 2 provides background context and describes how we collected and tested the drug samples. Section 3 describes the results. Section 4 discusses our interpretation and potential policy implications.

2 Background and Methods

2.1 Context

In Freetown, Sierra Leone, there are several channels for accessing medicine. These channels include formal providers, such as pharmacies, clinics, and hospitals, as well as informal providers known as *marketers*, who sell medicines on public buses, and *peddlers*, who sell medicines along streets.[1] Due to their larger presence and convenience of access, we focus on samples from pharmacies and marketers. Marketers largely source medicines from the same wholesale pharmacies that supply retail pharmacies. Both informal and formal channels play an important role in providing access to medicine as 14% of all respondents reported buying a medicine from a marketer in the past month, compared with 58% who reported buying from a pharmacy. It is common to purchase from both sources: 57% of those who purchased from a marketer in the last month also bought from a pharmacy.

For drug testing, we focus on three categories: painkillers (Diclofenac Potassium and

[1]Marketers have a larger share of the informal market: 6.8% of respondents reported buying from a peddler in the past month, compared with 14% from a marketer.

Paracetamol), deworming pills (Mebendazole), and antibiotics (Metronidazole). These are among the medicines most commonly sold by both informal and formal providers, with painkillers being the most common. We do not test the other high-volume categories—such as cough/cold remedies and antifungals—because of testing restrictions: cough/cold products are typically multi-ingredient combinations, and antifungals are usually sold as creams rather than tablets, neither of which was feasible for our available testing options.

2.2 Methods

Sample collection We collected 80 drug samples, 20 per API, through a convenience sample of vendors that Freetown residents commonly use in October–November 2025 (Table 1).

Pharmacy samples came from 29 distinct locations, in areas that consumers report commonly purchasing from in our survey evidence. We took several steps to ensure our sample collection reflected the actual quality that consumers face. First, to ensure we collected medicines that consumers receive in reality, enumerators did not specify brands, so the brand mix reflects what each channel actually offered. For each sample, we purchased approximately 80 tablets to ensure there was enough for formal testing (for deworming this was only 40 tablets, since there is only one tablet per package). Second, we did not specify the large purchase amount when initially procuring from vendors to avoid inference about our intentions, thus getting the most comparable product to what consumers receive.[2]

In sampling drugs, we first collected drugs from marketers. We started with marketers because we believe their supply is more unpredictable and our testing requires a common pharmacopoeia, per API. For example, if one sample of the same API was certified to the British Pharmacopoeia (BP) standard but another was certified to the United States Pharmacopoeia (USP), then laboratory testing would not be able to compare them. Thus, we started with the more unpredictable set of providers, and then when we collected drugs from pharmacies we could target identical brands, or if not available, those with the same pharmacopoeia.

Laboratory testing All 80 samples were submitted to the Ghana Food and Drugs Authority (FDA) Centre for Laboratory Services and Research, coordinated through RAMLabs, a Freetown-based research firm. Testing was conducted between January and March 2026 with blinded labels. Each sample underwent the full panel of tests required by the relevant pharmacopoeia, including identification, assay, dissolution, uniformity of weight, and impurities. Identification refers to whether the correct API is present, assay refers to whether the sample

[2]After securing a normal consumer size of sample, the mystery shoppers then ask for more tablets to reach the required number.

Table 1: Sample Overview

API	Pharmacopoeia	Total
Diclofenac Potassium (50mg)	USP	20
Mebendazole (500mg)	BP	20
Metronidazole (200mg)	BP	20
Paracetamol (500mg)	BP	20
Total		**80**

Notes: Samples were collected via mystery shopping in Freetown, Sierra Leone, between October and November 2025. All samples were submitted to the Ghana FDA Centre for Laboratory Services and Research for testing against the indicated pharmacopoeial standards. Dissolution was not tested for Metronidazole (BCS Class I under BP 2025).

contains the stated amount of active ingredient (i.e. a 50mg tablet contains 50mg within a permitted range), dissolution measures whether the active ingredient is released in sufficient time and quantity, uniformity of weight refers to whether the tablets in a sample have a consistent weight, and impurities refer to any unwanted ingredients in the tablets.[3] Between purchase and shipment, samples were stored at RAMLabs' Freetown offices in individually labelled bags in appropriate temperature and humidity conditions.

3 Results

3.1 Overall Quality

Our main result is that no sample failed identification, uniformity of weight or impurity testing, indicating there were no outright falsified drugs in any of the samples. At the same time, 20 of the 80 samples (25%) are considered substandard since they failed at least one pharmacopoeial quality test. Dissolution was the dominant, and essentially exclusive, failure among these samples: all 20 failures involved dissolution.[4] Two of those (both Mebendazole) also failed the assay test.

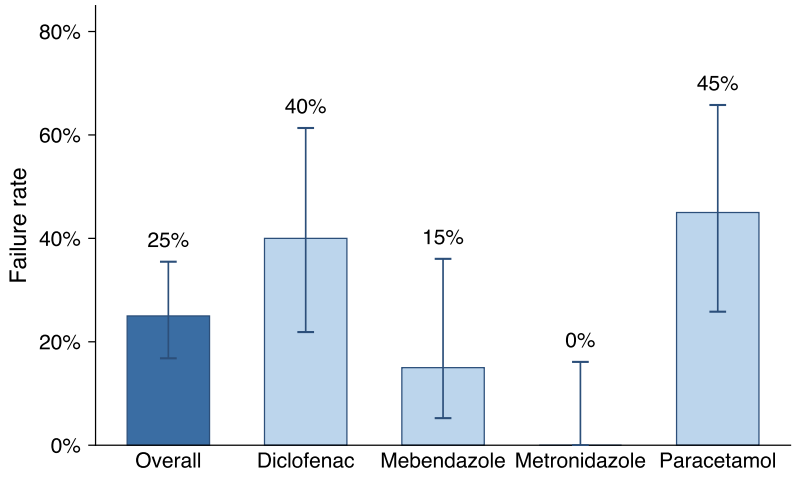
Failure rates vary substantially across APIs (Figure 1). Paracetamol has the highest failure rate at 45% (9 of 20), followed by Diclofenac at 40% (8 of 20), Mebendazole at 15% (3 of 20), and Metronidazole at 0% (0 of 20). The metronidazole result is not directly comparable

[3]Dissolution was not tested for Metronidazole, classified as BCS Class I under BP 2025, since metronidazole is considered highly soluble.

[4]Dissolution testing is sequential. At Stage S1, six tablets are tested individually and each must release at least 75% of labelled content. If any tablet fails, the lab proceeds to S2, adding six more tablets (12 total): the group average must be at least 75% and no individual unit may release less than 60%. If S2 fails, S3 adds 12 more tablets (24 total) under slightly looser criteria.

to the others since, as noted in Section 2.2, dissolution is not tested for Metronidazole as BP 2025 treats it as BCS Class I and exempts it from dissolution testing since it is already highly soluble. To understand these failures, Table 2 compares failed and passed samples on several continuous measures. We find that failed and passed samples carry similar (but still statistically significantly different) amounts of active ingredient: mean assay is 95.5% of label among failures versus 98.2% among passes. Where they differ more sharply is in dissolution: failed samples release about 54% of their labelled content on average, against about 89% for passes. Failed samples are also older (about 18 vs. 14 months at the time of testing) and closer to expiry (about 18 vs. 23 months at the time of testing), suggesting that the quality of medicine may slowly degrade over time.

Figure 1: Failure Rate by API



Notes: Share of samples failing at least one pharmacopoeial test, overall and by API. Error bars are 95% Wilson confidence intervals. $N = 80$ overall, $N = 20$ per API. Metronidazole was not tested for dissolution.

As mentioned, dissolution failures imply that the drug does not release enough of the API quickly enough, practically rendering the effective dosage to be lower. While this is less harmful than falsified drugs, dissolution failure risks the consumer not taking enough of the API for effective treatment. However, as shown, these failures are overwhelmingly concentrated among pain medications, which tend to manage and improve symptoms rather than treat underlying conditions. Substandard medicine would likely be more costly for deworming or antibiotics which treat conditions directly, where we find low failure rates. The full distribution of dissolution results is shown in Appendix Figure A1.

Table 2: Failed vs. Passed Samples

Measure	Failed	Passed	p-value
Mean assay (% of label)	95.5 ($n = 20$)	98.2 ($n = 60$)	0.02
Mean dissolution avg (%)	54.5 ($n = 20$)	89.2 ($n = 40$)	0.000
Mean age (months)	17.7 ($n = 19$)	14.1 ($n = 55$)	0.02
Mean time to expiry (months)	17.9 ($n = 20$)	23.0 ($n = 60$)	0.01

Notes: p-values from two-sided two-sample t-tests on continuous measures.

3.2 Quality by Manufacturer and Brand

We find that failures are highly concentrated among manufacturers. In particular, there are three main brands (i.e., manufacturer-product combinations) that account for 75% of failures. Interestingly, we find that quality problems are product-specific: one manufacturer's paracetamol fails most of the time, while its deworming product almost never does. Full manufacturer-level failure rates are reported in Appendix Table A1.

4 Discussion

Manufacturing versus storage Both manufacturing and storage conditions could cause the dissolution and assay failures we observe in the collected samples. Storage conditions only produce substantial dissolution failure under prolonged exposure (weeks to months) to elevated temperature and humidity. Manufacturer issues, such as incorrect formulation choices, may cause similar dissolution degradation. While we cannot distinguish the reason from the testing outcomes themselves, several patterns in the data suggest that manufacturing issues are the most likely explanation.[5]

First, failures cluster by manufacturer and product rather than by the vendor. In fact, when including manufacturer-API fixed effects, we find no statistical difference in failure rates between informal and formal vendors, and if anything, higher rates in the formal sector (Appendix Table A4). Second, if storage were the primary driver we would expect divergent outcomes for the same batch sold under different storage conditions. Instead, we find that when the same production batch is sampled multiple times across vendors, the

[5]Storage degradation of dissolution performance typically requires several weeks to months at conditions exceeding labelled limits (e.g., > 30°C / 70% relative humidity); short-term exposure of hours to a day or two rarely produces major dissolution failure unless conditions are extreme. Manufacturing-side failures, by contrast, can arise from in-process manufacturing, excipient interactions, formulation choices, raw-material storage, or chemical degradation of the active ingredient.

lab outcome agrees 14 times out of 16 (Appendix Table A3). We also consider how often vendors source medicines from wholesale pharmacies. We see no difference in how often samples that subsequently failed the lab tests were restocked versus those that passed. The short time that samples sit with vendors between purchase and sale is unlikely to generate the magnitude of failure we observe given the time required for dissolution degradation.

Does price signal quality? If consumers can sort good drugs from bad at the point of purchase, prices should track quality and regulation may matter less. We find that they do not. Pooling across channels, failed samples are not cheaper than passed ones within each API (Appendix Table A2): for Diclofenac the mean price per tablet is 1.19 vs. 1.01 NLe for failed and passed, for Paracetamol 0.84 vs. 0.72 NLe, and for Mebendazole 13.33 vs. 7.72 NLe—in each case the failed samples are, if anything, slightly more expensive, so a low price does not flag a bad drug. This is consistent with manufacturing defects: the tablets look, weigh, and present identically to compliant ones, so neither sellers nor buyers have a basis on which to discount the bad drugs.

Benchmarking and limitations Our 25% failure rate falls within the range reported by prior drug quality audits in sub-Saharan Africa (Table A5). It is close to the 25.4% reported for Nigeria by Gabel et al. [2024], the 22.6% from the multi-study African review of Mekonnen et al. [2024], and similar to the 18.6% of Schäfermann et al. [2020] in Cameroon and DR Congo, and the 35% Bate et al. [2008b] report for antimalarials across six African countries. Overall, our findings are, if anything, relatively favourable on drug quality in Freetown: we find no falsified drugs, whereas the comparable studies report small but non-zero falsification rates, and the failures we do find are concentrated in painkillers rather than the more critical antibiotic and deworming categories.

One limitation is that our sample size cannot rule out modest effects. Our sample size is primarily limited by the high cost of testing: each sample costs approximately $400 USD to test and shipping costs are substantial to ensure appropriate handling conditions. Future research should be conducted at scale when possible. We also restrict our testing to solid tablet form pharmacopoeial drugs in three high-volume categories (painkillers, dewormers, and antibiotics). Thus, we do not speak to the quality of products that fall outside this scope, such as topical antifungal creams, herbal medications and vitamins and supplements. These categories represent a non-trivial share of medicine sales and warrant separate testing methodologies that we leave to future work. Within pharmacopoeial drugs, we further restricted testing to single-API formulations: pharmacopoeial assay and dissolution methods require the laboratory to know the manufacturer's formulation in order to separate and

quantify each active ingredient, and this information is not available for combination products sold through these channels. We therefore are not able to test the commonly sold cold and cough tablets in our setting. Our final sample was also constrained by what vendors had on hand at the time of mystery shopping, rather than a fully prespecified product list.

Policy implications We close with a few observations that may be useful in future policy discussions. The patterns in our data point towards manufacturing rather than storage as the source of the drug failures we find. If that is correct, then the quality of a tablet is largely set before it reaches the point of sale and does not change much as it moves through the supply chain inside the country. This suggests some flexibility in where quality monitoring might happen: testing could be carried out at whatever point in the in-country chain is cheapest and most practical, rather than only at the retail end. Relatedly, because the shortfalls we observe concentrate in a small number of brands that are available across all types of seller, monitoring at the level of the brand or product may be more informative than distinctions based on the type of vendor or even manufacturer. However, drug quality is only one of several things that determine whether people take appropriate medicines and other important factors may differ across vendor channels in ways our testing does not measure [Das et al., 2016] but that future work should address.

References

Tariq Almuzaini, Imti Choonara, and Helen Sammons. Substandard and counterfeit medicines: A systematic review of the literature. *BMJ Open*, 3(8):e002923, 2013. doi: 10.1136/bmjopen-2013-002923.

Roger Bate, Peter Coticelli, Richard Tren, and Amir Attaran. Antimalarial drug quality in the most severely malarious parts of Africa – a six country study. *PLoS ONE*, 3(5):e2132, 2008a. doi: 10.1371/journal.pone.0002132.

Roger Bate, Philip Coticelli, Richard Tren, and Amir Attaran. Antimalarial drug quality in the most severely malarious parts of Africa – a six country study. *PLoS ONE*, 3(5):e2132, 2008b.

Roger Bate, Ginger Zhe Jin, and Aparna Mathur. Does price reveal poor-quality drugs? evidence from 17 countries. *Journal of Health Economics*, 30(6):1150–1163, 2011. doi: 10.1016/j.jhealeco.2011.08.006.

Daniel Bennett and Wesley Yin. The market for high-quality medicine: Retail chain entry and drug quality in India. *The Review of Economics and Statistics*, 100(4):691–703, 2018.

Martina Björkman-Nyqvist, Jakob Svensson, and David Yanagizawa-Drott. Can good products drive out bad? evidence from local markets for (fake) antimalarial medicine in Uganda. *Journal of the European Economic Association*, 16(3):659–693, 2018.

A. Dagrou and V. Chimhutu. I buy medicines from the streets because i am poor: A qualitative account on why the informal market for medicines thrive in Ivory Coast. *Inquiry: A Journal of Medical Care Organization, Provision and Financing*, 59:469580221086585, 2022. doi: 10.1177/00469580221086585.

Jishnu Das, Abhijit Chowdhury, Reshmaan Hussam, and Abhijit V. Banerjee. The impact of training informal health care providers in India: A randomized controlled trial. *Science*, 354(6308):aaf7384, 2016. doi: 10.1126/science.aaf7384.

F. Gabel, L. Heide, M. Lalani, P. N. Newton, and B. Schiavetti. Quality of essential medicines from different sources in Enugu and Anambra, Nigeria. *Tropical Medicine & International Health*, 2024. Forthcoming.

R. Hughes, C. R. Chandler, L. J. Mangham-Jefferies, and W. Mbacham. Medicine sellers' perspectives on their role in providing health care in north-west Cameroon: A qualitative study. *Health Policy and Planning*, 28(6):636–646, 2013. doi: 10.1093/heapol/czs103.

A. Johnston and D. W. Holt. Substandard drugs: A potential crisis for public health. *British Journal of Clinical Pharmacology*, 78(2):218–243, 2014. doi: 10.1111/bcp.12298.

Biset Asrade Mekonnen, Muluabay Getie Yizengaw, and Minichil Chanie Worku. Prevalence of substandard, falsified, unlicensed and unregistered medicine and its associated factors in Africa: A systematic review. *Journal of Pharmaceutical Policy and Practice*, 17(1): 2375267, 2024. doi: 10.1080/20523211.2024.2375267.

Gaurvika M. L. Nayyar, Joel G. Breman, Paul N. Newton, and James Herrington. Poor-quality antimalarial drugs in southeast Asia and sub-Saharan Africa. *The Lancet Infectious Diseases*, 12(6):488–496, 2012. doi: 10.1016/S1473-3099(12)70064-6.

Sachiko Ozawa, David R. Evans, Sophia Bessias, Deson G. Haynie, Tatenda T. Yemeke, Sandra K. Laing, and James E. Herrington. Prevalence and estimated economic burden of substandard and falsified medicines in low- and middle-income countries: A systematic review and meta-analysis. *JAMA Network Open*, 1(4):e181662, 2018. doi: 10.1001/jamanetworkopen.2018.1662.

Joseph P. Renschler, Kelsey M. Walters, Paul N. Newton, and Ramanan Laxminarayan. Estimated under-five deaths associated with poor-quality antimalarials in sub-Saharan Africa. *The American Journal of Tropical Medicine and Hygiene*, 92(6 Suppl):119–126, 2015. doi: 10.4269/ajtmh.14-0725.

Simon Schäfermann, Eckart Hörmann, Eduard Wagner, Anke Werntz, Mwiya Mukosha, and Lutz Heide. Substandard and falsified antibiotics and medicines against noncommunicable diseases in western Cameroon and northeastern DR Congo. *The American Journal of Tropical Medicine and Hygiene*, 103(2):894–908, 2020. doi: 10.4269/ajtmh.20-0184.

Patricia Tabernero, Facundo M. Fernández, Michael Green, Philippe J. Guerin, and Paul N. Newton. Mind the gaps – the epidemiology of poor-quality anti-malarials in the malarious world – analysis of the Worldwide Antimalarial Resistance Network database. *Malaria Journal*, 13:139, 2014. doi: 10.1186/1475-2875-13-139.

World Health Organization. A study on the public health and socioeconomic impact of substandard and falsified medical products. Technical report, World Health Organization, Geneva, 2017.

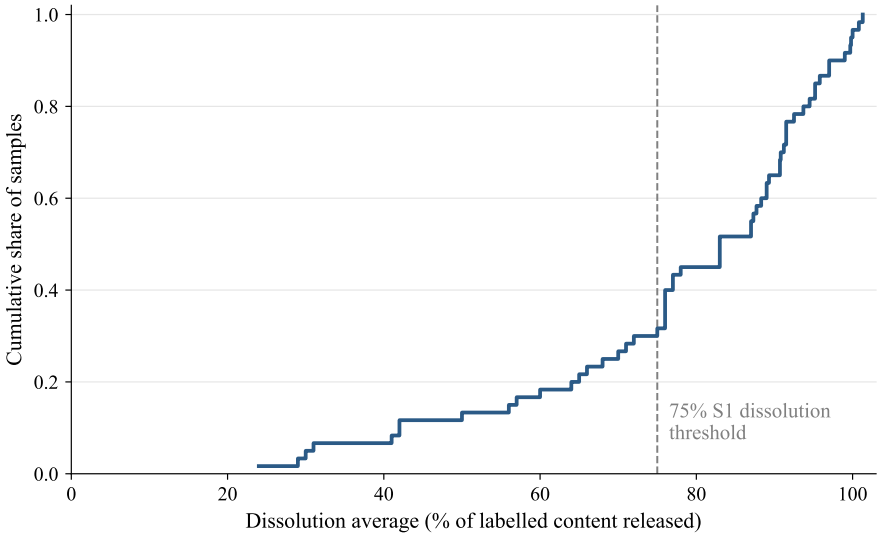
A Appendix Tables and Figures

Table A1: Failure Rates by Manufacturer-Product

API	Samples	Failed	Fail %
Mebendazole	2	2	100%
Paracetamol	10	8	80%
Diclofenac	2	1	50%
Paracetamol	2	1	50%
Diclofenac	2	1	50%
Diclofenac	11	5	45%
Mebendazole	16	1	6%
Diclofenac	2	0	0%
Metronidazole	2	0	0%
Metronidazole	2	0	0%
Metronidazole	3	0	0%
Metronidazole	4	0	0%
Metronidazole	2	0	0%
Metronidazole	2	0	0%
Paracetamol	2	0	0%
Paracetamol	3	0	0%
Metronidazole	2	0	0%
Diclofenac	2	0	0%

Notes: Each row is one manufacturer-product with two or more samples, ordered by failure rate.

Figure A1: Distribution of Dissolution Results



Notes: This graph plots the empirical cumulative distribution of the dissolution average (% of labelled content released) across all dissolution-tested samples; Metronidazole is excluded as it is exempt from dissolution testing. Dissolution is assessed under a sequential, multi-stage pharmacopoeial protocol. At Stage S1, six tablets are tested individually and each must release at least 75% of labelled content; if any tablet fails, testing proceeds to Stage S2, adding six more tablets (12 in total), where the group average must be at least 75% and no unit may release less than 60%; if S2 fails, Stage S3 adds twelve more tablets (24 in total) under slightly looser criteria. The value plotted is each sample's average release at the final stage reached.

Table A2: Mean Price per Tablet, Failed vs. Passed Samples, by API

	Failed		Passed	
API	N	Mean price	N	Mean price
Diclofenac	8	1.19	12	1.01
Mebendazole	3	13.33	17	7.72
Metronidazole	–	–	20	0.77
Paracetamol	9	0.84	11	0.72

Notes: Mean price per tablet (NLe) for failed and passed samples within each API, pooled across both channels.

Table A3: Within-Batch Quality Consistency

Batch outcome	Batches	Samples in batches
All samples passed	11	31
All samples failed	3	9
Mixed (some pass, some fail)	2	5
Total	**16**	**45**

Notes: The 16 batches that appear in two or more of the 80 samples (45 samples in total; remaining samples come from singleton batches).

Table A4: Failure on Market Channel: Linear Probability Models

	(1) Full sample	(2) Full, Mfr×API FE
Marketer	0.100	-0.051
	(0.097)	(0.114)
	$[p = 0.31]$	$[p = 0.66]$
Mfr×API fixed effects	No	Yes
Observations	80	80

Notes: LPM of failure on a marketer indicator across the 80-sample dataset. Standard errors in parentheses; p-values in brackets.

Table A5: Benchmarking: Drug Quality Studies in Sub-Saharan Africa

Study	Setting	N	Falsified ("fake") rate	Substandard rate
This study	Sierra Leone (Freetown); 4 APIs	80	0%	25%
Schäfermann et al. (2020)	Cameroon, DR Congo; antibiotics and other NCD medicines	506	0.6%	18.6%
Gabel et al. (2024)	Nigeria; 13 essential medicines	260	1.5%	25.4%
Bate et al. (2008)	6 African countries; antimalarials	210	Not reported	35%
Mekonnen et al. (2024)	Review across 7 studies in Africa; variety of medicines	7,592	Not reported	22.6%

Notes: Studies differ in drug classes, sample sources, and testing methods, so the rates are not directly comparable across rows. "Falsified rate" is the share of samples with the wrong or no active ingredient; "substandard rate" is the share failing at least one pharmacopoeial test. Set against these comparators, our results are relatively favourable: we find no falsified drugs, the most clinically critical drugs we test (the antibiotic and dewormer) fail less often than the painkillers, and our overall substandard rate is broadly in line with the comparable studies.

theigc.org